



Topic Builder

NGIC Site Visit

December 2016

UNCLASSIFIED // FOUO

DISTRIBUTION STATEMENT C - Distribution authorized to U.S. Government Agencies and their contractors to protect information and technical data that advance current technology or describe new technology in an area of significant or potentially significant military application, 15 July 2014. Other requests for this document shall be referred to Air Force Research Laboratory, Rome Research Site, AFRL/RIEB, 525 Brooks Road, Rome NY 13441-4505.

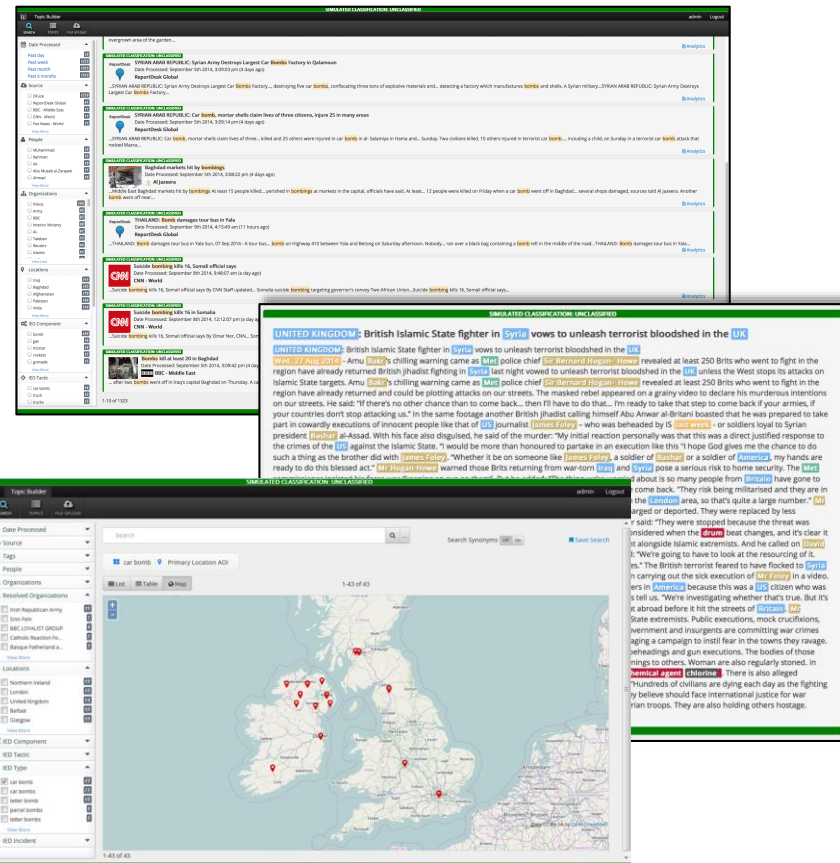
Contents

- What is Topic Builder?
 - Leveraging Unstructured Data
 - Topic Builder Interface
 - Multi-Lingual Support
 - Semantic Enrichment
- Why Topic Builder?
 - Topic Builder Interface
 - How Taxonomies Help
 - Entity and Concept Extraction
 - Geospatial Extraction
- Benefits of Topic Builder

What is Topic Builder

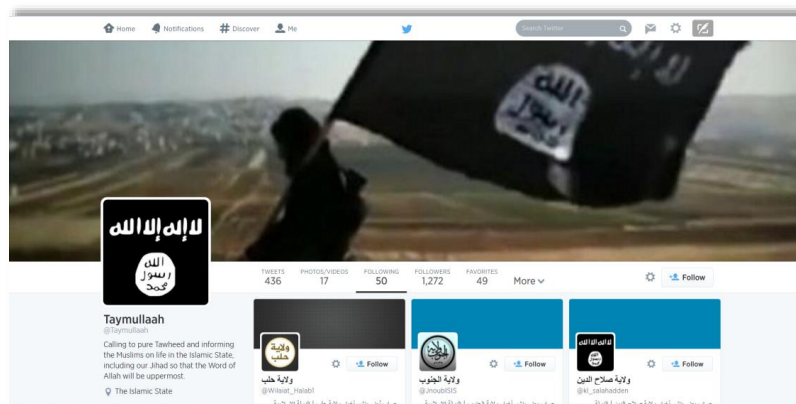
- **Data Integration:** Multi-source ingest – text files, email, RSS feeds, textual database fields, Microsoft Office and SharePoint, PDF, email, etc.; includes semantic enrichment/entity extraction
- **Search and Discovery:** Keyword, taxonomic, geospatial and tag search with faceted navigation
- **Monitoring Data Sources:** Topic management and multimodal alert notification
- **Knowledge Management:** Source metadata and textual content which works within classified and unclassified environments
- **Collaboration:** Shared topics and tags

Automated Content Acquisition with Continuous Monitoring and Alerting.



Leveraging Unstructured Data

- Many organizations are overwhelmed by *large volumes of unstructured data* without the tools to manage them
- Keeping up with *open source reporting* is difficult and tedious
- Multiple search engines and web sites need to be *individually accessed and searched* to achieve comprehensive coverage
- Tracking known high-value entities (groups, individuals, incidents, devices, etc) and tactics is hard, but *recognizing unknown entities is even harder*
- Seeing *larger patterns* found across many reports and spread over time is a challenge

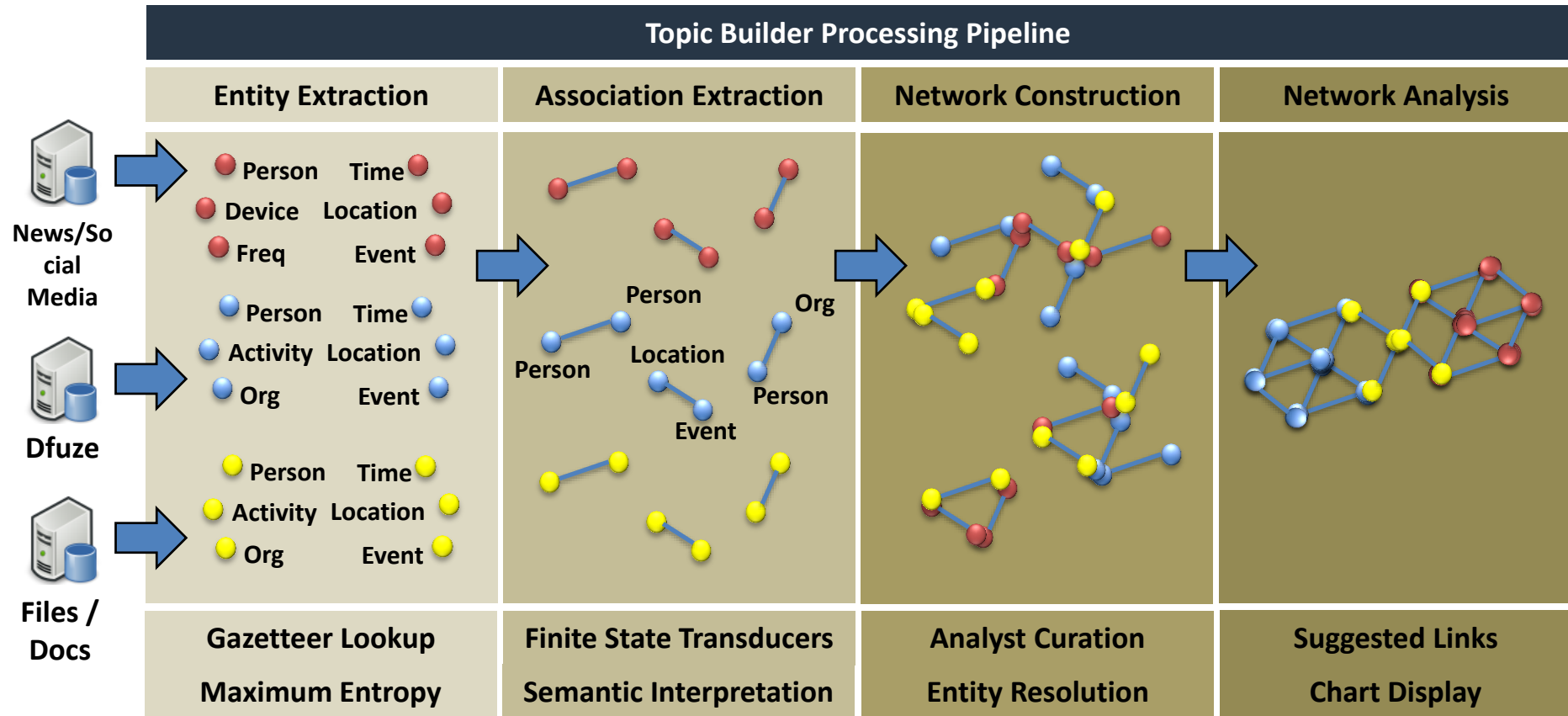


Multi-Lingual Support

- Topic Builder has the ability to handle multiple languages in several ways:
 - Synonyms in multiple languages can be added to taxonomies
 - Gazetteers support multiple languages
 - Natural language processing engine supports English*, Spanish*, Dutch*, German, Danish, Swedish and Portuguese
 - Additional languages can be handled through a process of training new machine learning models
 - Google Translate or other translation engines can be used during content acquisition
 - Some source, e.g., Opensource.gov, provide translation services

* These languages have named entity recognition models for Persons, Organizations and Locations

Semantic Enrichment



Topic Builder's Natural Language Pipeline uses text analysis techniques with reference data (Dfuze, Gazetteer, Taxonomies) to enable the user to "Shrink the Haystack"

Why Topic Builder?

- ***Automate Content Acquisition.*** Topic Builder brings multiple unstructured and structured data sources into a platform that **organizes and integrates your data** based on your preferences and priorities
- ***Shrink the Haystack.*** It is used to **search for information relevant to** an organization's Information Requirements, Social Network Development, Strategic Communication, Information Operations, etc
- ***Monitor and Notify.*** It is used to **continuously monitor multiple data sources** for information relevant to your mission
- ***Analyze with Context.*** It provides a methodology and tool for recognizing relevant information that is **much more valuable than traditional search**

Spend more time analyzing data, less time searching for it

Topic Builder Interface

The screenshot displays the Topic Builder Interface, a web-based tool for managing and analyzing data. The interface is divided into a sidebar on the left and a main content area on the right.

Sidebar (Faceted Navigation):

- Date Processed:** Filters for Past day, Past week, Past month, and Past 6 months.
- Source:** Filters for various news sources like dFuse, ReportDesk Global, BBC - Middle East, CNN - World, and Fox News - World.
- People:** Filters for individuals such as Muhammad, Rahman, Ali, Abu Musab al-Zarqawi, and Ahmad.
- Organizations:** Filters for entities like Police, Army, BBC, Interior Ministry, Al-, Taleban, Reuters, Islamic, and others.
- Locations:** Filters for geographical areas like Iraq, Baghdad, Afghanistan, Pakistan, and India.
- IED Component:** Filters for components of Improvised Explosive Devices (IEDs) such as bomb, gas, mortar, rockets, and grenade.
- IED Tactic:** Filters for tactics like car bomb, truck, and trucks.

Main Content Area:

- Simulated Classification: UNCLASSIFIED:** A header indicating the current filter status.
- Search Results:** A list of news articles with simulated classifications and unclassified status. Each article includes a title, date processed, and a brief description.
- Analytics:** A button next to each article, likely linking to a detailed analysis or report.

The interface is designed for faceted navigation and web-based access, allowing users to refine their search results based on various criteria.

Faceted Navigation

Web-based access

Intuitive Interface

How Taxonomies Help

The screenshot shows the HME (Heterogeneous Metadata Explorer) interface. A search bar at the top contains the text 'HME'. To the right of the search bar is a 'Search Synonyms' toggle set to 'On'. Below the search bar, the results are displayed in a list view. The first result is titled 'SIMULATED CLASSIFICATION: UNCLASSIFIED' and contains the text '...00 12 dead, 43 injured Ammonium nitrate, which it was to have used... The police suspect that ammonium nitrate - a chemical used in explosives... the blast was caused by an 18-tonne truckload of ammonium nitrate. The... chemical is a common fertiliser, but when mixed with diesel or kerosene can...'. The second result is also titled 'SIMULATED CLASSIFICATION: UNCLASSIFIED' and contains the text '...Improvised Explosive Device (IED), from a shopping complex in Anantnag. The... explosives included aluminum powder, sulphur, potassium nitrate, gunpowder...'. Callouts from the text point to 'Aluminum Powder' and 'Ammonium Nitrate'. The interface also includes a sidebar with filters for Date Processed, Source, People, and Organizations.

Mission-specific taxonomies are used as taxonomic synonym trees:

- Makes searching for entire categories easy, e.g., IED components, law enforcement terms, air defense systems, cyber attack vectors
- Handles spelling variations, acronyms, abbreviations and misspellings
- Handles multi-lingual synonyms
- Taxonomies defined for counter-IED, WMD proliferation, law enforcement, targeting, and humanitarian assistance/disaster relief

Entity and Concept Extraction

SIMULATED CLASSIFICATION: UNCLASSIFIED

UNITED KINGDOM: British Islamic State fighter in **Syria** vows to unleash terrorist bloodshed in the **UK**

UNITED KINGDOM: British Islamic State fighter in **Syria** vows to unleash terrorist bloodshed in the **UK**
Wed, 27 Aug 2014 - Amu **Bakr**'s chilling warning came as **Met** police chief **Sir Bernard Hogan- Howe** revealed at least 250 Brits who went to fight in the region have already returned British jihadist fighting in **Syria** last night vowed to unleash terrorist bloodshed in the **UK** unless the West stops its attacks on Islamic State targets. Amu **Bakr**'s chilling warning came as **Met** police chief **Sir Bernard Hogan- Howe** revealed at least 250 Brits who went to fight in the region have already returned and could be plotting attacks on our streets. The masked rebel appeared on a grainy video to declare his murderous intentions on our streets. He said: "If there's no other chance than to come back... then I'll have to do that... I'm ready to take that step to come back if your armies, if your countries don't stop attacking us." In the same footage another British jihadist calling himself Abu Anwar al-Britani boasted that he was prepared to take part in cowardly executions of innocent people like that of **US** journalist **James Foley** - who was beheaded by IS **last week** - or soldiers loyal to Syrian president **Bashar** al-Assad. With his face also disguised, he said of the murder: "My initial reaction personally was that this was a direct justified response to the crimes of the **US** against the Islamic State. "I would be more than honoured to partake in an execution like this "I hope God gives me the chance to do such a thing as the brother did with **James Foley**. "Whether it be on someone like **James Foley**, a soldier of **Bashar** or a soldier of **America**, my hands are ready to do this blessed act." **Mr Hogan-Howe** warned those Brits returning from war-torn **Iraq** and **Syria** pose a serious risk to home security. The **Met** commissioner insisted his force was "keeping an eye on them". But he added: "The thing we're worried about is so many people from **Britain** have gone to **Syria** and potentially **Iraq** to get involved. We think at least 500 to 600 went. Some have come back. "They risk being militarised and they are in a network of terrorists. "It's a worry that up to three-quarters of them come from the **London** area, so that's quite a large number." **Mr Hogan-Howe** also called for the Home Office to issue arrest orders for suspects who cannot be charged or deported. They were replaced by less restrictive Terrorism Prevention and Investigation Measures by the Coalition **Government**. The officer said: "They were stopped because the threat was reduced and quite properly they were not allowed to return. "But I think these things have got to be considered when the **drum** beat changes, and it's clear it has." **Mr Hogan-Howe** backed growing calls to strip citizenship from Brits who go to war zones to fight alongside Islamic extremists. And he called on **David Cameron** to plough more cash into the battle to prevent terrorism on our streets. The policeman said: "We're going to have to look at the resourcing of it. "The Australians have just invested an awful lot of money in the same problem in their security services." The British terrorist feared to have flocked to **Syria** and **Iraq** include "John the Beagle", the masked figure at the centre of an international manhunt seen carrying out the sick execution of **Mr Foley** in a video. **Mr Hogan-Howe** said: "We are making progress in that investigation. "We are working with our partners in **America** because this was a **US** citizen who was murdered by a British citizen, it appears. "It sounds like somebody from **London**, that's all the experts tell us. "We're investigating whether that's true. But it's not straightforward." **Last week**, **SAS** hero **Chris Ryan** warned the **UK** needed to tackle the IS threat abroad before it hit the streets of **Britain**. **Mr Hogan-Howe** spoke as **UN** investigators **yesterday** revealed the ultra-brutal inside world of Islamic State extremists. Public executions, mock crucifixions, amputations and whippings are all being used in their campaign of terror in **Syria**. Both the Syrian government and insurgents are committing war crimes and crimes against humanity in their bloody battle with each other. IS forces in **northern Syria** are waging a campaign to instil fear in the towns they ravage. Children as young as 10 are joining the terror training camps and have been forced to watch **Friday** beheadings and gun executions. The bodies slaughtered for allegedly violating religious laws are then left on crucifixes for up to three days as warnings to others. Women are also regularly retaliation, Assad forces have dropped barrel bombs on civilian areas. Some are said to contain the **chemical agent** **chlorine**. There is also all "systematic torture and deaths of detainees" in detention centres. **Probe** chief **Paulo Pinheiro** said: "Hundreds of civilians are dying each day a goes on with no regard to law or to conscience." Investigators have drawn up four lists of suspects they believe should face international justice crimes. Pictures posted by IS fighters on social media **yesterday** showed the group have executed Syrian troops. They are also holding others h Reported: **Sat, 30 Aug 2014 07:35:31 GMT**

Navigate
via Entities
and
Concepts

Turn Entity
Highlighting
On or Off

Entity Reference Highlighting ☐ Off ☒ On

<input checked="" type="checkbox"/> Date	<input checked="" type="checkbox"/> Location
<input checked="" type="checkbox"/> Person	<input checked="" type="checkbox"/> Organization
<input checked="" type="checkbox"/> IED	<input checked="" type="checkbox"/> WMD
<input checked="" type="checkbox"/> Target	<input checked="" type="checkbox"/> Cyber

SIMULATED CLASSIFICATION: UNCLASSIFIED

Geospatial Extraction

The screenshot displays the 'Topic Builder' interface, which is used for geospatial data analysis. The interface includes a sidebar on the left with various filters and a main map area on the right. The sidebar contains sections for 'Date Processed', 'Source', 'Tags', 'People', 'Organizations', 'Resolved Organizations', 'Locations', 'IED Component', 'IED Tactic', and 'IED Type'. The 'Locations' section is currently selected, showing a list of locations including Northern Ireland, London, United Kingdom, Belfast, and Glasgow. The main map area shows a map of the United Kingdom with several red location pins. A search bar at the top right contains the text 'car bomb' and 'Primary Location AOI'. A callout bubble labeled 'Faceted Navigation' points to the sidebar. A callout bubble labeled 'Change Report Locations' points to the 'Locations' section in the sidebar. A callout bubble labeled 'AOI Query' points to the search bar. A callout bubble labeled 'Extract Geospatial Areas from Text' points to the map area. The interface also includes a 'SIMULATED CLASSIFICATION: UNCLASSIFIED' banner at the top and bottom, and a 'Logout' button in the top right corner.

Topic Builder

SEARCH TOPICS FILE UPLOAD

admin Logout

Date Processed

Source

Tags

People

Organizations

Resolved Organizations

Irish Republican Army

Sinn Fein

BBC LOYALIST GROUP

Catholic Reaction Fo...

Basque Fatherland a...

View More

Locations

Northern Ireland

London

United Kingdom

Belfast

Glasgow

View More

IED Component

IED Tactic

IED Type

car bomb

Search

car bomb Primary Location AOI

Search Synonyms Off On

Save Search

1-43 of 43

List Table Map

1-43 of 43

1-43 of 43

SIMULATED CLASSIFICATION: UNCLASSIFIED

Faceted Navigation

Change Report Locations

AOI Query

Extract Geospatial Areas from Text

Benefits of Topic Builder

- Topic Builder users connected to Opensource.gov (IC), RSS/web services, News Outlets, and other OSINT sources will have ***access to the most relevant open data***
- ***User-defined topics*** monitor all the connected sources for new information relevant to the users' missions
- ***Powerful search capability*** goes well beyond key word search to identify related information
- Users can easily ***tag relevant reports*** with their own key terms and phrases for ***quick, future lookup in a collaborative environment***
- Users can also ***search for relevant local content*** uploaded by them and others in their organizations
- Faceted navigation ***exposes the frequency of important entities mentioned across reports*** providing the ability to drill into the data

Program Contacts

Wyatt Bora

ASPERE COTR/CIDNE PM

AFRL/RIEBB

Comm 315-330-4944

DSN 312-587-4944

Mobile 315-617-3619

wyatt.bora.1@us.af.mil

Peter M. LaMonica

WebTAS/CIDNE Program

AFRL/RIEBB

Email: Peter.LaMonica@us.af.mil

COM: (315) 330-4088

DSN: 587-4088

Mobile: (315) 240-7258

SIPR: peter.m.lamonica.civ@mail.smil.mil

JWICS: peter.lamonica@rome.ic.gov

Topic Builder

BACK-UP SLIDES