# Making Sense of Evolving Threats: A Novel AI Approach to Monitor Violent Organisations and their Defining Characteristics Over Time from Open Sources

**Thomas Powell[1] * Anne Merel Sternheim[1] Ioannis Tolios[1] Koen van der Zwet[1,2,3] Freek Bomhof[1]**

[1]Netherlands Organisation for Applied Scientific Research (TNO)
TNO Defence, Safety & Security
[2]Science Faculty, University of Amsterdam, Amsterdam
[3] Institute for Advanced Study, Amsterdam
THE NETHERLANDS

*tom.powell@tno.nl

## *ABSTRACT*

*Today's increasingly unstable geo-political environment presents opportunities for violent organisations to thrive – whether terrorist, criminal, insurgent or militant. It is therefore important for NATO and its nations to monitor and understand potentially threatening organisations as they evolve. At the same time, digitalization has led to an explosion in the amount of open source information, bringing with it analytical opportunities but also the challenge of information overload – whilst developments in Artificial Intelligence (AI) offer new technologies to deal with this challenge. In this paper we build on knowledge of organisational theory and AI to present a first step towards an approach for automatically extracting information about threatening organisations from large amounts of open sources over time. Specifically, we use an existing database of violent organisations as a starting point, and combine three different AI methods for extracting several specific organisational characteristics: (1) Supervised Machine Learning (ML) for extracting activity types and target types; (2) Unsupervised ML for extracting ideology; and (3) Natural Language Processing (NLP) for extracting organisation size, number of leaders, and number of deaths and injured in attacks. We evaluate the performance of these methods and reflect on their generalizability. By doing so we take a step towards tools for automatically monitoring threatening organisations and their defining characteristics as they evolve. Future steps, such as implementing models to 'fill in' missing or incomplete information and automatically detecting new threatening organisations as they emerge, are discussed.*

## 1.0 INTRODUCTION

Today's threat environment is populated by a complex and constantly evolving array of violent organisations – whether terrorist, criminal, insurgent or militant. These organisations need to be monitored and understood for NATO and its nations to ensure their societies are safe and secure. Several factors complicate these efforts. First, the increasingly hybrid nature of conflict means that organisations, often acting as proxies for state actors, continuously emerge, evolve and disappear again, sometimes at a high rate (Treverton, 2014). These organisations operate under the threshold of open military confrontation, often covertly and across geographical borders or in cyberspace, which makes attributing activities to these organisations a challenge (Cullen, 2018; NATO, 2019). Resurgent great-power competition has contributed to an unstable geopolitical environment, which is only compounded by global crises such as the COVID-19 pandemic and the increasingly visible pressures from climate change (Bekkers, Meessen & Lassche, 2018). At the same time, digitalization has produced an overload of potentially relevant information for understanding violent organisations. Challenges enough, then, for efforts by NATO and its nations to make sense of evolving threats.

Innovative methods are required to overcome these challenges and improve our ability to understand violent organisations as they evolve. Any proposed method needs to be able to deal with insights that can be gleaned from open sources, and deal with the associated information overload. Traditional intelligence methods are manpower intensive, and are poorly suited to the complex and constantly evolving operational environment and adversary within it (Eisler, 2012). In order to support the analyst in improving commanders' intelligence picture, we explore automated methods to extract, update and structure incoming open source information.

Any novel approach should recognise that an abundance of (historical) knowledge regarding violent organisations already exists, including in openly-available knowledge bases (e.g., NCTV, 2021; Stanford CISAC, 2021). These sources are valuable not only for their information value, but also because the information is structured in a meaningful way, based on the qualitative expertise of the analysts who contributed to them. For instance, the structure of these knowledge bases correspond to important characteristics of violent organisations, such as their modus operandi, size, ideology, structure, etc. (van der Vecht & Keijser, 2018). This existing structured knowledge can be extremely useful in guiding the analysis and monitoring of large amounts of incoming unstructured and often incomplete information. The methodological challenge lies in combining this existing structured knowledge with incoming information. In other words: fusing analyst-driven with data-driven insights to improve situational understanding of violent organisations.

In this paper we aim to address this challenge whilst drawing on existing work into combining data-driven (data science and AI) and analyst-driven (qualitative, theory-driven) insights (Pherson & Pherson, 2020; Westerveld, Powell & Eles, 2020). The result is the illustrated approach, which uses a combination of Artificial Intelligence (AI) techniques to extract information about the defining characteristics of violent organisations over time from open sources. By doing so we hope to support analysts in dealing with information overload and make a step towards decision support tools for defence practitioners for making sense of evolving threats. Specifically, we build on and extend literature on organisational theory, intelligence analysis and AI to address the following research questions (RQs):

- RQ1: Which AI techniques are well suited to extracting important characteristics of violent organisations? This question is addressed by the literature reviewed in Sections 1.1 to 1.4.
- RQ2: Are the automatically extracted characteristics an accurate representation of the knowledgebase from which they came? This question is addressed by the approach outlined in the Method in Section 2 and Results in Section 3.
- RQ3: To what extent can these AI techniques generalise to extract and monitor characteristics of violent organisations from other open sources? This is addressed in the Discussion in Section 4.

## 1.1 Characteristics of violent organisations

In order to monitor violent organisations, one needs to identify and understand their defining characteristics. Organisations are collections of human systems of cooperation and coordination assembled within identifiable boundaries to pursue shared performance goals or objectives (Hodge, Anthony & Gales, 2003). This general definition contains elements (e.g., coordination, boundaries, shared goals) possessed by all organisations (Mintzberg, 1985) – including violent ones.

In the context of violent extremist organisations, the literature about organisational characteristics is illustrative of at least four main points (Asal & Rethemeyer, 2008; Ligon et al., 2013). First, larger organizations are more dangerous. Second, coordination between members varies in organizations. Third, organizational mission or purpose matters. Fourth, connectedness or alliances are related to performance. These findings highlight the importance of studying organizational characteristics such as size, activity types, coordination, mission and connectedness in relation to ideological performance outcomes. Multiple levels of analysis are therefore needed to understand the violent organisation as a whole.

These multiple levels of analysing violent organisations were assimilated by van der Vecht & Keijser (2018). In their work on quantitatively modelling different types of violent organisations – including terrorist, criminal, insurgent or militant – they developed a generic typology. Drawing on Ligon et al. (2013), Ganor (2008) and Schultz (1978), amongst others, four major categories were identified to describe the characteristics of violent organisations:

(1) *Activities* – how the organisation fund, recruit and carry out violent activities, the latter being equivalent to Modus Operandi, an organisation's distinct pattern or method of operation (Combs, 2017);
(2) *Structure* – a set of relatively static organisation parameters including hierarchy, chain of command and labour division;
(3) *Design* – the design of organisational decision-making processes, including the degree of formalisation, centralization, and nature of communication (e.g., Ligon et al., 2013); and
(4) *Other attributes* – reflecting a catch-all category of important additional characteristics such as size, ideology, strategy and financial means.

Van der Vecht & Keyser (2018) applied this typology to two different types of violent organisations – the Colombian FARC and a fictitious Dutch jihadist network. In doing so, they illustrated how characteristics of violent organisations can be summarised through the assimilation of available (open source) data and information. Once summarised, they suggest that this information can be used in quantitative models that will help intelligence analysts in examining potential effects of interventions against a violent organisation, and address questions such as: "what if we prevent a certain type of violent activity or fundraising source?", or "what if we undermine the group's ideology?" Structured information about violent organisations can also be of substantial value in answering other types of intelligence questions (Pherson & Pherson, 2020): For instance of an explanatory nature ("why is this organisation in resurgence?"), or a descriptive nature ("how is this organisation evolving?").

We focus on this latter descriptive-type question by understanding violent organisations through the lens of organisation theory (e.g., Ligon et al., 2013). In order to test our approach to using different AI techniques to automatically extract organisational information, we make a sub-selection of important organisational characteristics. To ensure a good representation of the typology of van der Vecht & Keijser (2018), we chose the following characteristics: *Activities* (including *type*, *target*, and *number of deaths and injuries*); *Ideology*; *Size* and *Number of leaders*. As well as reflecting different types of characteristics, these are features that can evolve over time and at different rates (e.g., activity type evolves faster than ideology). These characteristics also lend themselves to extraction using different types of AI techniques – helping us to test our approach.

## 1.2 Existing information sources about violent organisations

As noted in the previous section, summarising data and information on violent organisations is a well-known task for intelligence analysts. Take, for instance, the Intelligence Preparation of the Operational Environment (IPOE; NATO JP 2-01.3, 2014). IPOE is a process used to "analyse all relevant aspects of the environment, including the adversary and other actors […] and political, military, economic, social, information, and infrastructure (PMESII) systems and subsystems". This holistic understanding enables the commander to leverage aspects of the environment to achieve operational objectives.

A major part of the IPOE is in evaluating the adversary and determining potential adversary courses of action. Understanding the adversary organisation at multiple levels is therefore highly relevant. Given the operational environment and the adversary within it is subject to constant change and evolution, the speed at which information is turned into intelligence is often crucial (Eisler, 2012). By constantly extracting, updating and structuring incoming open source data and information, the commanders' intelligence picture can be improved (see, for instance, Conklin, Bechtel & Goebel, 2020).

We contribute to this endeavour by presenting an approach that extracts information from open sources that, once further developed, can contribute a continuously updating understanding of violent adversary organisations. Two key challenges are associated with this goal: how to efficiently extract and analyse an abundance of potentially relevant information? (See Section 1.3.) And, how should this information be structured to provide meaningful insights about violent organisations?

Regarding this latter question, several open source knowledge bases exist that not only assimilate large amounts of information on violent organisations, but also structure this in an informative way. Take, for instance, the Stanford Mapping Militants knowledge base[1], or the Dutch NCTV knowledge base of terrorist organisations[2]. These knowledge bases contain rich qualitative information – descriptions, histories, summaries – and quantitative information – statistics, estimates, timelines – regarding many of the key organisational characteristics summarised by van der Vecht and Keijser (2018). For instance, the Mapping Militants knowledge base contains several categories: Organisation (leadership, name changes, size estimates, resources, locations); Strategy (ideology, aims, politics, targets and tactics); Major Attacks (first, largest, notable attacks); and Interactions (foreign designations, community relations, relationships with other militant groups, state sponsors, external influences). Clearly this type of knowledge base provides a comprehensive and authoritative summary of many violent organisations (albeit Stanford and NCTV have a focus on Islamic militant groups) according to their key characteristics.

The main drawback of these knowledgebases, however, is that they are highly manpower intensive to maintain. To the best of our knowledge these structured knowledgebases of violent organisations are not automatically and continuously updated from open sources and cost many hours of analytical work. Other continuously updated databases do exist, such as the GDELT[3] or Uppsala[4] databases, but these are not tailored to understanding violent organisations. The aim of this research is to fill this gap by presenting a method that is: (1) specifically tailored to understanding important characteristics of violent organisations, based on organisational theory; and (2) able to automatically and continuously extract information, thereby reducing the informational overload of the analyst.

## 1.3 Linking automatic extraction techniques to characteristics of violent organisations

In our approach to extract relevant organisational characteristics from open sources, we draw on the large body of existing work on three types of AI techniques: (1) supervised Machine Learning (ML); (2) unsupervised ML; and (3) Natural Language Processing (NLP). Existing literature on these methods and their link to the violent organisational characteristics of interests are described in this section.

### 1.3.1 Supervised Machine Learning to extract violent activities

In order to extract information on violent organisation activities (specifically *activity type* and *target type*) we apply supervised ML. Supervised learning uses a training dataset to teach models to yield the desired output. This training dataset includes inputs and correct outputs, which allow the model to learn over time. The model is then able to categorise new input data (e.g., other open source data) that have a relatively similar structure to the training dataset (Torres Torres, Hart & Emery, 2019). Some examples of supervised ML applications to understanding violent groups include identifying the potential risk of specific attacks by perpetrators based on historic events. For example, the work of Mo et al. (2017) attempts to predict the possibility of terrorist attacks using various machine learning methods such as Naïve Bayes and Logistic Regression. More recently, van Hensbergen (2020) trained different ML models on the Global Terrorism Database and was able to categorise new input data in terms of the Modus Operandi deployed in a violent

[1] https://cisac.fsi.stanford.edu/mappingmilitants
[2] https://www.nctv.nl/onderwerpen/kennisbank-terroristische-organisaties
[3] https://www.gdeltproject.org/
[4] https://ucdp.uu.se/

attack. These results show that machine learning algorithms are a feasible approach to the extraction and structuring of the activities of violent organisations.

### 1.3.2 Unsupervised Machine Learning to extract ideology

To extract information about *ideology* of a violent organisation, we used topic modelling, which is a related method to cluster analysis and part of the larger family of unsupervised ML techniques. Topic modelling is a statistical method that "analyses the words of the original texts to discover the themes that run through them, how those themes are connected to each other, and how they change over time" (Blei, 2012). In statistical terms, a topic is a probability distribution over words, in which the most relevant words having the highest probability of membership to a topic (Aggarwal & Zhai, 2012). To estimate them, the co-occurrence of words over a set of documents is used (Grimmer & Stewart, 2013). When applied to documents describing the ideology of violent organisation, topic modelling can distinguish organisations with semantically similar or different collections of words. This is useful since ideologies of violent organisations often reflect a mix or combination of different ideologies, or a sub-set of an overarching ideologies (e.g., the multitude of slightly differing ideologies possessed by Sunni Islamic militant groups in Pakistan's Khyber Pakhtunkhwa province). Moreover, shifts in an organisation's ideology may be measured by comparing the distance to the 'centre' of a topic over time.

### 1.3.3 Basic NLP to extract organisation size, number of leaders, number of deaths and injured

We used basic NLP for extracting information about the *size* of an organisation, the *number of leaders*, and the *number of deaths and injured in attacks*. Within the field of NLP (Cambria & White, 2014) there are several elementary, often rule-based, methods. These techniques can typically be used without the need of a (large) set of training data like in ML. Such methods are often used to pre-process a text before other more complex techniques (such as topic modelling) are applied, but can also be useful as a standalone method to extract information. We use one of these methods, regular expressions, in which special characters are used to match patterns in text (Friedl, 2006). For example, '`a`' matches the letter 'a', '`\s`' matches any whitespace character, '`[0-9]`' matches any number. Regular expressions are especially useful when information is extracted from a text which is fairly structured, or when the information is always in the same kind of textual context. The Mapping Militants knowledgebase, and the characteristics of size and number of leaders in particular, are well suited to extraction using regular expressions.

The preceding text provides an answer to our first research question (RQ1): different AI techniques are suited to extracting qualitatively different characteristics of violent organisations. Table 1 shows an overview of the different characteristics and their linked techniques. In the next sections 2 and 3, we address our second research question: Are the characteristics extracted using AI methods an accurate representation of the knowledgebase from which they come? To answer this question we apply the identified AI techniques to extract violent organisation characteristics from our chosen dataset, as illustrated in Figure 1.

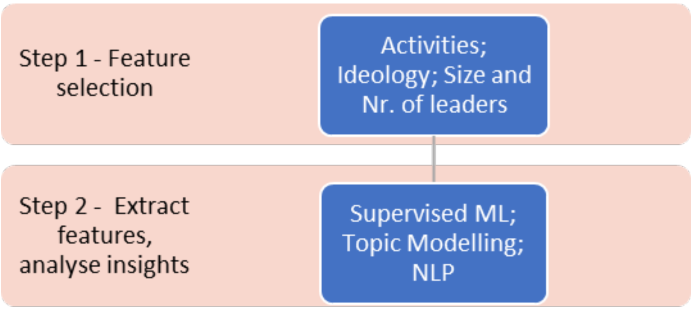

Step 1 - Feature selection
Activities; Ideology; Size and Nr. of leaders
Step 2 - Extract features, analyse insights
Supervised ML; Topic Modelling; NLP

**Figure 1 – Overview of the organisational features and AI methods in our analytical approach**

# 2.0 METHOD

The approach developed involves the steps outlined in Figure 1. Before explaining these in more detail in the sections below, we first explain the data we relied on in developing the AI methods.

## 2.1 Data: Stanford Mapping Militants database

As of July 25th 2021, the Stanford Mapping Militants knowledge base contained descriptions of 86 different organisations, and a mapping of the relationships between them. For the current research we are only interested in the description of the organisations.

All information on a single organisation is summarised in a portable document file (PDF). Amongst others, this file contains sections on the organisation's organisational structure (leadership, size, resources, geographic locations) and strategy (ideology and goals, major attacks), including how this has developed over time. For example, the organisation's size might have been estimated in 2014, and again in 2020. Both estimations will be listed under the 'Size' section.

In order to extract any information from the 86 PDF files automatically, the files were first parsed using a Python port of Apache Tika[5] and then the text of interest per organisation was extracted. At the end of the process, the extracted data was converted to JavaScript Object Notation (JSON)[6] format. In our case, a section header (e.g., 'size') is mapped to a value and section content (e.g., 'There are no publicly available size estimates for this group.'). Such files are easy to parse, are language independent and the mapping between the headers and the content makes extracting the different sections from the file very convenient.

## 2.2 Identification of organisational characteristics

Several relevant organisational characteristics (from here on also called features) were identified, being: The number of leaders, size, ideology, and for an organisation's activities: the type of event, the type of target, and the number of deaths and injuries. Different methods were used to extract these features from different sections of the data. Table 1 provides an overview.

**Table 1 – Overview of the methods used to extract organisation features from the relevant section of our dataset**

| Extraction method | Extracted feature | Data section |
|---|---|---|
| Supervised Machine Learning: Logistic regression model | Activity: event type | Major attacks |
| | Activity: target type | Major attacks |
| Unsupervised Machine Learning: Topic modelling | Ideology | Ideology and goals |
| Natural Language Processing: Regular expressions | Number of leaders | Leadership |
| | Organisation size | Size |
| | Activity: Number of deaths | Major attacks |
| | Activity: Number of injuries | Major attacks |

## 2.3 Extraction of organisational characteristics

Since this paper aims to illustrate the value of combining AI techniques to extract information about violent organisations, the methods employed and described in this section are based on pragmatic choice.

[5] https://tika.apache.org/index.html
[6] https://www.json.org/json-en.html

### 2.3.1 Supervised ML to extract Activity type and Target type

We use the approach and models from van Hensbergen (2020). In this work, a model was trained on input data in the form of summaries of incidents from the Global Terrorism Database (GTD). The model was trained to categorise these incidents into different characteristics relevant to the concept of Modus Operandi, including *activity type* and *target type*. The GTD contains 9 different categories of activity types, and 22 different categories of target types.[7] After pre-processing, van Hensbergen (2020) compared different methods and Logistic Regression and Stochastic Gradient Descent were chosen for training models. Internal validation of the main model resulted in accuracy scores of 90% for activity type and 93% for target type (van Hensbergen, 2020).

We applied the models for activity type and target type from van Hensbergen (2020) to the 'major attacks' sections of the Mapping Militants data. The major attacks section lists a selection of attacks executed by the organization, each in a new paragraph. By splitting the major attacks on paragraphs, each activity could be categorised independently. A qualitative evaluation of the categorisation performance is given in the Results Section 3.1.

### 2.3.2 Unsupervised ML to extract Ideology

We applied a specific implementation (Yao et al., 2009) of the popular Latent Dirichlet Allocation (LDA) topic modelling approach (Blei, Ng & Jordan, 2003). This method dictates that the user has to pick a number of expected topics, even when not sure about the actual number of topics. At the end of the process each document is represented by a distribution of topics and each topic by a distribution of words.

Before applying the LDA technique to the extracted 'ideology and goals' section of the documents, the documents were pre-processed. Initially, each document was converted into a list of lowercase tokens. Furthermore, the set of tokens was stripped in order to construct a useful and relevant set. First, we ignored tokens that are shorter than three characters and longer than fifteen characters. Second, stop words, punctuation and accent marks were removed from the list of tokens to reduce the dataset size, while keeping the most important information. Subsequently, sequences of 2 words and 3 words were extracted called bi-grams and tri-grams. Lemmatization and removal of the unnecessary parts of speech was then conducted. Only nouns, adjectives, verbs and adverbs were kept for analysis.

We then created a corpus and a dictionary, which are required by LDA.[8] The dictionary contains a list of terms in our collection of documents, in which each unique term is mapped to an index. The corpus is the document-term matrix representation of the dictionary. It indicates the frequency of a term per document. Since, we did not have specific expectations about the number of existing ideology topics in our corpus, we ran LDA model for a varying number of topics ranging from 1 topic to 12 topics. We used the topic coherence pipeline and the 'Cv' coherence measure introduced by Röder et al. (2015) to assess performance. The typical values of the Cv score are in the range [0,1]. The optimal number of topics was 4 with the highest coherence score of 0.39. This coherence score is relatively low which means interpretation of the resulting topics should be made with caution (Röder et al. (2015). Potential ways to improve this score could be a better pre-processing approach or a further finetuning of LDA's parameters. Based on this outcome, paragraphs describing ideology and goals were grouped into 4 different topics. We hypothesised that these topics would represent the different types of ideologies belonging to the violent organisations in the dataset.

[7] For the full list of attack types and target types in the GTD, see https://www.start.umd.edu/gtd/search/BrowseBy.aspx?category=attack

[8] Note that the use of a dictionary can limit the ability to detect large changes over time, which requires more sophisticated methods than the ones illustrated here (e.g., Tahmasebi et al., 2012).

### 2.3.3 Basic NLP to extract Size, Number of leaders and Number of deaths and injured

The *group size* feature was extracted from the size section, using a regular expression. The regular expression returned a list of size estimates throughout the years. The most recent estimate was returned as the currently known *group size*.

The *number of leaders* feature was also determined with the help of a regular expression. The leaders of the organisation were listed in different paragraphs, and for each leader, the reigning period was given in a predictable format. A regular expression was used to extract these periods, and the length of the extracted list was returned as the *number of leaders*.

Regular expressions were also used to extract the *number of deaths* and *number of injured* features. Both were matched against the appropriate paragraphs in the 'major attacks' section. The regular expression for the number of deaths and injured returned a number, and non-numerical signs were removed from the result (e.g.: '6,000+ injured' would return the number 6000).

# 3.0 RESULTS

In order to determine the performance of the feature extraction methods, the results were evaluated in a qualitative manner. For some of the feature extraction methods, evaluation was supported by examining the results of a single organisation in detail. Al Qaeda was chosen for this since it is a particularly well-known violent organisation.

## 3.1 Supervised ML for the extraction of Activity type and Target type

Because the Stanford Mapping Militants data was not labelled, we manually checked the performance on a sample of 13 paragraphs from the major attacks section. Table 2 below provides three examples of attacks conducted by Al Qaeda and the application of our manual coding schema to assess the validity of the result.

**Table 2 – Examples of manual coding of activity type and target type based on three texts about major attacks conducted by Al Qaeda**

| **Example text** | **Result** | | **Valid (*explanation*)** | |
|---|---|---|---|---|
| | **Activity type** | **Target Type** | **Activity type** | **Target Type** |
| November 15, 2003: Carried out over two days (November 15 and November 20, 2003), four truck bombs ran into 2 Jewish synagogues, a bank, and the British Consulate in Istanbul, Turkey. The bombing at the British Consulate may have been coordinated with U.S. President Bush's meeting with Tony Blair, which occurred the day of the second bombing (11/20/2003). (67 killed, 700+ wounded) | Bombing/ Explosion | Government (Diplomatic) | Yes (*the paragraph mentions truck bombs*) | No (*the paragraph mentions multiple targets. The given label does not cover all of them*) |

| | | | | |
|---|---|---|---|---|
| July 7, 2005: Four British men detonated 3 bombs on the London Underground and one on a double-decker bus during morning rush hour in London. Al Qaeda claimed the bombings, but there is no direct evidence that shows that AQ directed the attack. (56 killed, 770+ injured) | Bombing/ Explosion | Transportation | Yes (*the paragraph mentions bombs*) | Yes (*the paragraph mentions the underground and busses, which are transportation*) |
| October 2010: AQAP sent bombs through cargo mail, attempting to down planes over the U.S. The bombs were discovered before the planes left for the U.S. but had successfully passed through several cargo screening facilities in different countries. (No casualties) | Bombing/ Explosion | Unknown | Yes (*the paragraph mentions bombs*) | No (*the paragraph mentions planes as the intended target*) |

Across the 13 paragraphs that were manually coded, the correct activity type was extracted 100% of the time (Table 3). The correct target type was extracted 69% of the time (Table 4). Although we only coded a handful of paragraphs, supervised ML models used to extract activity type performed better than the external validation score of 89% in van Hensbergen (2020). For activity type, performance in this study was highly similar to van Hensbergen's external validation score of 67%.

**Table 3 – Performance (validity) of the ML model used to extract *activity type* based on 13 manually evaluated paragraphs from the major attacks section of the dataset**

| **Activity type** | **Count** | **Valid (%)** | **Invalid (%)** |
|---|---:|---:|---:|
| Bombing/ explosion | 9 | 100 | 0 |
| Hijacking | 1 | 100 | 0 |
| Hostage taking (Kidnapping) | 1 | 100 | 0 |
| Assassination | 1 | 100 | 0 |
| Armed Assault | 1 | 100 | 0 |
| **Total** | **13** | **100** | **0** |

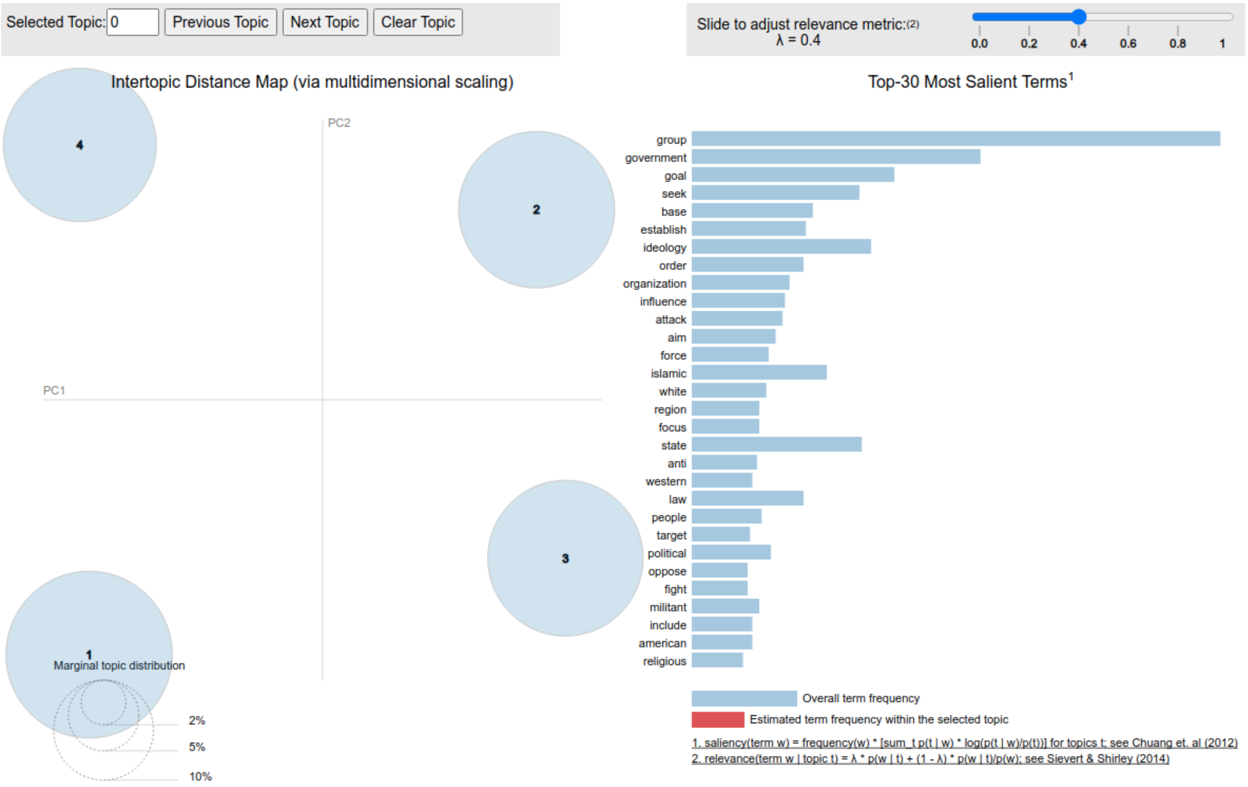
**Table 4 – Performance (validity) of the ML model used to extract *target type* based on 13 manually evaluated paragraphs from the major attacks section of the dataset**

| **Target type** | **Count** | **Valid (%)** | **Invalid (%)** |
|---|---:|---:|---:|
| Military | 1 | 0 | 100 |
| Maritime | 1 | 100 | 0 |
| Business | 2 | 100 | 0 |
| Government | 2 | 0 | 100 |
| Private Citizens & Property | 2 | 100 | 0 |
| Government (General) | 1 | 100 | 0 |
| Transportation | 1 | 100 | 0 |
| Airports & Aircraft | 1 | 100 | 0 |
| Unknown | 1 | 0 | 100 |
| Journalists & Media | 1 | 100 | 0 |
| **Total** | **13** | **69** | **31** |

## 3.2 Unsupervised ML to extract Ideology

Our LDA topic modelling method identified four different topics. We hypothesised that these topics would represent different ideologies of violent organisations in the dataset. Since topic modelling does not apply labels to the extracted topics, the results require considerable subjective interpretation.

We used pyLDAvis[9] a Python library for interactive topic model visualization to obtain a better understanding of the generated topics, as shown in Figure 2 below. The inter-topic distance map on the left shows that the identified topics are well spaced apart – suggesting that the topics should differ considerably from one another. Each circle represents a topic. The size of the circle represents how prevalent the topic was in the data, the larger the circle the more prevalent the topic. The bar chart on the right shows the most salient terms across all four topics.



Selected Topic: 0

Slide to adjust relevance metric:(2) λ = 0.4

1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)

**Figure 2 – Summary of the four identified topics based on the ideology and goals paragraphs of all violent organisations in the dataset**

To interpret whether the four identified topics identified meaningfully different types of violent group ideology, we inspect the most central organisation and the most salient words for each topic. These are shown in Table 5.

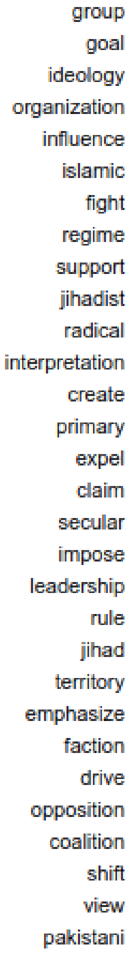
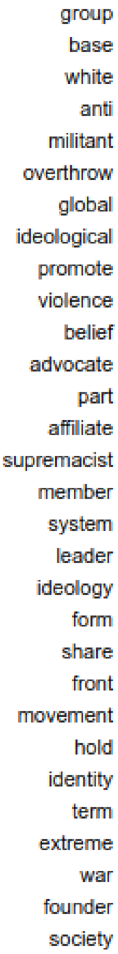
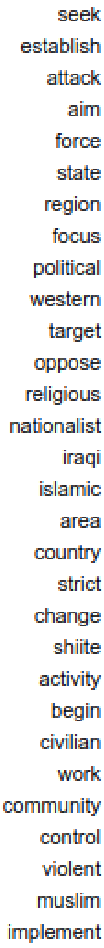
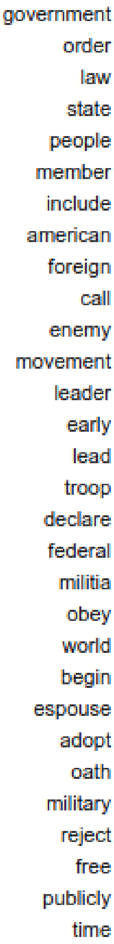
Two Islamic extremist organisations (Topic 1: Islamic Movement of Kurdistan; and Topic 3: Asa'ib ahl al Haq) and two far-right extremist organisations (Topic 2: The Base; and Topic 4: Oath Keepers) are identified as most central. Inspection of these organisations suggests meaningful differences between the topics. The Islamic Movement of Kurdistan's (topic 1) members share a predominantly Sunni Islamist ideology, whilst

[9] https://pyldavis.readthedocs.io/en/latest/readme.html

Asa'ib ahl al Haq (topic 3) is a Shiite militant and political organisation. These ideologies are somewhat represented by the top 30 salient words for these topics, for instance with salient words like “Islamic” and “radical” for topic 1, and “nationalist” and “state” for topic 3. However, the differences are not especially noticeable from these words alone.

Focusing on the two right-wing topics identified: The Base (topic 2) is a white supremacist organization that adheres to anti-Semitic conspiracy theories, whilst the Oath Keepers (topic 4) ideology aligns with the anti-government patriot/militia movement. These ideologies match fairly well with the salient words for these topics. For instance with “white” and “overthrow” for topic 2, and “government” and “American” for topic 4.

**Table 5 – The most central organisation and the top 30 salient words for the four topics based on the ideology and goals paragraphs of the dataset**

| | **Topic 1** | **Topic 2** | **Topic 3** | **Topic 4** |
|---|---|---|---|---|
| Most central Organisation | Islamic Movement of Kurdistan | The Base | Asa'ib ahl al Haq | Oath Keepers |
| 30 most salient words | <br>group<br>goal<br>ideology<br>organization<br>influence<br>islamic<br>fight<br>regime<br>support<br>jihadist<br>radical<br>interpretation<br>create<br>primary<br>expel<br>claim<br>secular<br>impose<br>leadership<br>rule<br>jihad<br>territory<br>emphasize<br>faction<br>drive<br>opposition<br>coalition<br>shift<br>view<br>pakistani | <br>group<br>base<br>white<br>anti<br>militant<br>overthrow<br>global<br>ideological<br>promote<br>violence<br>belief<br>advocate<br>part<br>affiliate<br>supremacist<br>member<br>system<br>leader<br>ideology<br>form<br>share<br>front<br>movement<br>hold<br>identity<br>term<br>extreme<br>war<br>founder<br>society | <br>seek<br>establish<br>attack<br>aim<br>force<br>state<br>region<br>focus<br>political<br>western<br>target<br>oppose<br>religious<br>nationalist<br>iraqi<br>islamic<br>area<br>country<br>strict<br>change<br>shiite<br>activity<br>begin<br>civilian<br>work<br>community<br>control<br>violent<br>muslim<br>implement | <br>government<br>order<br>law<br>state<br>people<br>member<br>include<br>american<br>foreign<br>call<br>enemy<br>movement<br>leader<br>early<br>lead<br>troop<br>declare<br>federal<br>militia<br>obey<br>world<br>begin<br>espouse<br>adopt<br>oath<br>military<br>reject<br>free<br>publicly<br>time |

To more closely evaluate the performance of the topic modelling method, the ideology text for Al Qaeda is shown in Table 6. This suggests a mis-categorisation of Al Qaeda to topic 3. Al Qaeda expounds a predominantly Sunni ideology, whereas topic 3 appears to be predominantly Shiite ideology. This may have been due to Shiite being included in the Al Qaeda ideology text twice, and the inability of the topic model to accurately recognise negative references to this ideology. Moreover, inspection of the most salient words suggest that the topic may have been defined based on words describing goals (such as seek, establish, attack, aim, force, state) rather than ideology (such as Sunni or Shia Islam). Taken together, these results suggests that the topic modelling method has some success in identifying violent organisation ideologies, but the subjective nature of labelling the topics make it less suitable as an automated method.

**Table 6 – The dominant topic and most salient words for the Al Qaeda ideology text**

| Ideology text | Dominant topic | Most salient words |
|---|---|---|
| Al Qaeda aims to rid the Muslim world of Western influence, to destroy Israel, and to create an Islamic caliphate stretching from Spain to Indonesia that imposes strict Sunni interpretation of Shariah law. However, not all AQ members and affiliates agree on the same laws. Some consider Shiite Muslims to be apostates, while others do not. This disagreement has caused rifts between AQ and its affiliates — for example, when AQI targeted Shiites in Iraq against the instructions of bin Laden. | 3 | seek,<br>establish,<br>attack,<br>aim,<br>force,<br>state,<br>focus,<br>region,<br>political,<br>western. |

## 3.3 Basic NLP to extract Size, Number of leaders and Number of deaths and injured

The use of regular expressions to extract the number of deaths, the number of injured, group size and number of leaders produced 100% accurate performance, based on the qualitative evaluations. When a number is sought the appropriate number is returned, when zero is present the number zero is returned, and when no number is present *Not-a-Number* (NaN) is a returned (the latter is not shown in the examples in Table 7).

Table 7 below shows the extracted number of deaths and number of injured for three of the 13 paragraphs in the ‘major events’ section. The group size and number of leaders features were tested for Al Qaeda, and the correct size (32,000-44,000 as of 2018) and number of leaders (7) were returned.

**Table 7 – Extracted number of deaths and number of injured in a sample of three paragraphs in the major events section of the dataset**

| | Number of death | | Number of injured | |
|---|---|---|---|---|
| | Actual | Extracted | Actual | Extracted |
| **(67 killed, 700+ wounded)** | 67 | 67 | 700+ | 700 |
| **(56 killed, 770+ injured)** | 56 | 56 | 770+ | 770 |
| **(No casualties)** | 0 | NaN | 0 | NaN |

## 4.0 DISCUSSION AND CONCLUSION

In this paper we presented a first step towards novel approach to extract the defining characteristics of violent organisations. We applied different AI methods – supervised Machine Learning (ML), unsupervised ML and Natural Language Processing (NLP) – to the extraction of organisational characteristics from the Stanford Mapping Militants database – namely activities, ideology, size and number of leaders. By doing so we illustrate the utility of our approach to analysts whose task is to monitor violent organisations as they evolve from open sources.

At the end of the Introduction (Section 1), we addressed our first research question (RQ1) and argued that that specific AI methods are suited to extracting particular violent organisational characteristics (see Table 1). Here, we take time to address RQ2: Are the automatically extracted characteristics an accurate representation of the information in the knowledgebase from which they came? Our answer to this is dependent on the AI method and organisation characteristic: The use of supervised ML to extract activity type and target type was shown to be reasonably accurate. The same can be said for NLP-based extraction of number of deaths, number of injured, organisation size and number of leaders. By contrast, the success of unsupervised ML (topic modelling) in extracting the ideology of violent organisations was not immediately evident. The identified topics did seem to represent distinct ideologies or goals but required substantial subjective interpretation. Future research could seek to improve this method, and/or evaluate its appropriateness for identifying other organisational characteristics such as goals/objectives or strategy.

Our third research question (RQ3) focused on the extent to which these AI methods generalise to other open sources than the knowledgebase used in this study. The performance of the supervised ML model indicates that it should generalise fairly well to extracting activities from other open source texts. The generalisability of our unsupervised ML approach to extracting ideology is likely dependent on the corpus of text used as input. The input text in this research contained both ideology and goals and so was perhaps not specific enough. If the input text is strictly focused on ideology then it is likely to perform well. However, given that open sources (such as news articles) can contain multiple different subjects in a single article, it is possible that topics may be distinguished based on the other characteristics of the text. This is a general limitation of topic modelling, and this suggestion should be examined in future research. Lastly, our NLP methods are the least generalisable to other data since they are highly dependent on the structure of the Mapping Militants dataset. Future work should seek to make these approaches as generalisable as possible beyond the limited data in this paper, for instance by training models on a wider dataset or by using word embeddings supplemented with context information (van Luenen, 2020).

This research opens up several avenues of future study. First, the outputs of the approach could be summarised and/or visualised in a dashboard to offer up-to-date information about violent organisations to intelligence analysts and commanders. For instance, Conklin, Bechtel & Goebel (2020) demonstrate how various data streams could be combined to provide a live and updating Intelligence Preparation of the Operational Environment (IPOE). Such an approach would need to account for the dangers in automatic extraction, such as the questionable credibility of some open sources – something which is not addressed in this paper. Another interesting possibility builds on innovative AI models that allow the imputation of missing features in text – for instance by combining word embeddings with grounded information in a knowledge graph (Yang, Zhu, Sachidananda & Darve, 2019). Imagine that new incoming information includes some relevant features (includes activity types, number of deaths and ideology) but not others (missing target type and size). An imputation approach could ‘fill-in’ the missing information based on the large amount of existing complete data.[10] In practice this could help answer time-sensitive questions such as those regarding unfolding attacks. Such as: what was the intended target of the attack and which organisation is responsible? Finally, our approach does not keep track of collectives or loose networks of individuals who

[10] Note that distinguishing imputed information from actual information would be important in this approach.

pose a threat but cannot be labelled as an existing violent organisation (Ligon et al., 2013). To address this, data science and AI methods, potentially relying on graph (network) type data, could be used to identify new and emerging collectives as they coalesce into increasingly structured groups or organisations.

To conclude, the approach presented illustrates the value of combining different types of AI methods to extract and monitor key characteristics of violent organisations. We aimed to support analysts dealing with some of the challenges of today's operational environment – particularly instability, complexity, hybrid threats and information overload. By doing so we hope to contribute towards decision support for defence practitioners who are responsible for making sense of violent organisations as they evolve.

## 5.0 REFERENCES

[1] Aggarwal C.C., Zhai C. (2012). "A Survey of Text Clustering Algorithms" in: *Aggarwal C., Zhai C. (eds) Mining Text Data*. Springer, Boston, MA. https://doi.org/10.1007/978-1-4614-3223-4_4

[2] Asal, V., & Rethemeyer, R. K. (2008). "The nature of the beast: Organizational structures and the lethality of terrorist attacks". *The Journal of Politics,* 70(2), 437-449.

[3] Bekkers, F., Meessen, R. & Lassche, D. (2018). "Hybrid conflicts: The new normal?" in The Hague: TNO & The Hague Centre for Strategic Studies. Available via https://repository.tudelft.nl/view/tno/uuid: c280883b-13af-4654-aa8a-9a8e7a0d2a99. Accessed on 11th of August, 2021.

[4] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). "Latent dirichlet allocation" in: *Journal of Machine Learning Research*, 3, 993-1022.

[5] Blei, D. M. (2012). "Surveying a suite of algorithms that offer a solution to managing large document archives" in: *Communication of the ACM*, 55(4), 77-84.

[6] Cambria, E., & White, B. (2014). "Jumping NLP curves: A review of natural language processing research." in: *IEEE Computational intelligence magazine*, 9(2), 48-57.

[7] Cullen, P. (2018). "Hybrid threats as a new 'wicked problem' for early warning." in: *Strategic Analysis, Number 8. Helsinki, Finland: The European Centre of Excellence for Countering Hybrid Threats.* available via https://www.hybridcoe.fi/publications/hybrid-coe-strategic-analysis-8-hybrid-threats-as-a-new-wicked-problem-for-early-warning/. Accessed on 11th of August, 2021.

[8] Conklin, B., Bechtel, B. & Goebel, M. (2020). "Realizing a Live Intelligence Preparation of the Operational Environment." in: *Trajectory Magazine*, https://trajectorymagazine.com/realizing-a-live-intelligence-preparation-of-the-operational-environment/

[9] Eisler, C. D. F. (2012). "Counter-IED strategy in modern war." in*: The Australian Army Journal is published by authority of the Chief of Army*, 9(2), 51.

[10] Friedl, J. E. (2006). "Mastering regular expressions". 2006. O'Reilly Media, Inc.

[11] Ganor, B. (2008). "Terrorist Organization Typologies and the Probability of a Boomerang Effect" in*: Studies in Conflict & Terrorism*, 31: 4, p. 269-283.

[12] Grimmer, J., & Stewart, B. M. (2013). "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." in: *Political analysis*, 21(3), 267-297.

[13] Hensbergen (2020). "Modus Operandi van Subversieve Groeperingen." B.S. Thesis. University of Amsterdam. 2020.

[14] Ligon, G. S., Simi, P., Harms, M., & Harris, D. J. (2013). "Putting the "O" in VEOs: What makes an organization?" in: *Dynamics of Asymmetric Conflict*, *6*(1-3), 110-134.

[15] Mason, R., McInnis, B., & Dalal, S. (2012, June). "Machine learning for the automatic identification of terrorist incidents in worldwide news media." in: 2012 IEEE International Conference on Intelligence and Security Informatics (pp. 84-89). IEEE.

[16] Mintzberg, H., & Waters, J. A. (1985). "Of strategies, deliberate and emergent." in: *Strategic management journal*, 6(3), 257–272.

[17] Mo, H., Meng, X., Li, J., & Zhao, S. (2017, March). "Terrorist event prediction based on revealing data." in: *2017 IEEE 2nd International Conference on Big Data Analysis* (ICBDA) (pp. 239-244). IEEE.

[18] NATO JP 2-01.3 (2014). "Joint Intelligence Preparation of the Operational Environment." *US Joint Chiefs of Staff publication.*

[19] Pherson, K. H., & Pherson, R. H. (2020). "Critical thinking for strategic intelligence." CQ Press.

[20] Röder, M., Both, A., & Hinneburg, A. (2015, February). "Exploring the space of topic coherence measures" in: *Proceedings of the eighth ACM international conference on Web search and data mining* (pp. 399-408).

[21] Sun, A., Naing, M. M., Lim, E. P., & Lam, W. (2003, June). "Using support vector machines for terrorism information extraction." in: *International Conference on Intelligence and Security Informatics* (pp. 1-12). Springer, Berlin, Heidelberg.

[22] Shultz, R. (1978). "Conceptualizing Political Terrorism: A Typology", in: *Journal of International Affairs*, 32: 1, p. 7-15.

[23] Stanford University (2021). Mapping Militant Organizations. https://cisac.fsi.stanford.edu/mapping militants

[24] Tahmasebi, N., Gossen, G., Kanhabua, N., Holzmann, H., & Risse, T. (2012, December). Neer: An unsupervised method for named entity evolution recognition. In Proceedings of COLING 2012 (pp. 2553-2568).

[25] Torres Torres, M., Hart, G. & Emery, T. (2019). "The Dstl biscuit book: Artificial Intelligence, Data Science and (mostly) Machine Learning". *Dstl publication 115968.*

[26] Treverton, G.F. (2014). "The future of intelligence: Changing threats, evolving methods." in: *The Future of Intelligence: Challenges in the 21st Century*. Duyvesteyn, I., de Jong, B., van Reijn, J. (Eds.), (pp. 27-38). London: Routledge.

[27] Van Luenen, A. F. (2020). "Recognising moral foundations in online extremist discourse: A cross-domain classification study." M.S. Thesis, Uppsala University, Department of Linguistics and Philology.

[28] Westerveld, J., Powell, T. E. & Eles, P. T. (2020). "Making best use of survey data for operational

analysis: Pattern analysis using supervised machine learning." in: *Proceedings of the 14th NATO Operations Research & Analysis Conference.*

[29] Yang, Z., Zhu, C., Sachidananda, V., & Darve, E. (2019). "Out-of-Vocabulary Embedding Imputation with Grounded Language Information by Graph Convolutional Networks." arXiv preprint arXiv:1906.03753.

[30] Yao, L., Mimno, D., & McCallum, A. (2009, June). "Efficient methods for topic model inference on streaming document collections." in: *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 937-946).