Deploying Semantic Web Technologies for Information Fusion of Terrorism-related Content and Threat Detection on the Web

MITZIAS, Panagiotis, KOMPATSIARIS, Ioannis, KONTOPOULOS, Efstratios, STAITE, James, DAY, Tony <http://orcid.org/0000-0002-3214-6667>, KALPAKIS, George, TSIKRIKA, Theodora, GIBSON, Helen <http://orcid.org/0000-0002-5242-0950>, VROCHIDIS, Stefanos and AKHGAR, Babak <http://orcid.org/0000-0003-3684-6481>

Deploying Semantic Web Technologies for Information Fusion of Terrorism-related Content and Threat Detection on the Web

The Case of the TENSOR EU funded Project

Panagiotis Mitzias
Information Technologies Institute
CERTH
Thessaloniki, Greece
pmitzias@iti.gr

Efstratios Kontopoulos
Information Technologies Institute
CERTH
Thessaloniki, Greece
skontopo@iti.gr

James Staite
CENTRIC
Sheffield Hallam University
South Yorkshire, England, UK
J.Staite@shu.ac.uk

Tony Day
CENTRIC
Sheffield Hallam University
South Yorkshire, England, UK
T.Day@shu.ac.uk

George Kalpakis
Information Technologies Institute
CERTH
Thessaloniki, Greece
kalpakis@iti.gr

Theodora Tsikrika
Information Technologies Institute
CERTH
Thessaloniki, Greece
theodora.tsikrika@iti.gr

Helen Gibson
CENTRIC
Sheffield Hallam University
South Yorkshire, England, UK
H.Gibson@shu.ac.uk

Stefanos Vrochidis
Information Technologies Institute
CERTH
Thessaloniki, Greece
stefanos@iti.gr

Babak Akhgar
CENTRIC
Sheffield Hallam University
South Yorkshire, England, UK
B.Akhgar@shu.ac.uk

Ioannis Kompatsiaris
Information Technologies Institute
CERTH
Thessaloniki, Greece
ikom@iti.gr

ABSTRACT

The Web and social media nowadays play an increasingly significant role in spreading terrorism-related propaganda and content. In order to deploy counterterrorism measures, authorities rely on automated systems for analysing text, multimedia, and social media content on the Web. However, since each of these systems is an isolated solution, investigators often face the challenge of having to cope with a diverse array of heterogeneous sources and formats that generate vast volumes of data. Semantic Web technologies can alleviate this problem by delivering a toolset of mechanisms for knowledge representation, information fusion, semantic search, and sophisticated analyses of terrorist networks and spatiotemporal information. In the Semantic Web environment, ontologies play a key role by offering a shared, uniform model for semantically integrating information from multimodal heterogeneous sources. An additional benefit is that ontologies can be augmented with powerful tools for semantic enrichment and reasoning. This paper presents such a unified semantic infrastructure for information fusion of terrorism-related content and threat detection on the Web. The framework is deployed within the TENSOR EU-funded project, and consists of an ontology and an adaptable semantic reasoning mechanism. We strongly believe that, in the short- and long-term, these techniques can greatly assist Law Enforcement Agencies in their investigational operations.

WI '19 Companion, October 14–17, 2019, Thessaloniki, Greece

ACM ISBN 978-1-4503-6988-6/19/10…$15.00
https://doi.org/10.1145/3358695.3360896

CCS CONCEPTS

• **Computing methodologies** → **Knowledge representation and reasoning**; **Ontology engineering**; *Description logics*; • **Applied computing** → *Law*.

KEYWORDS

Ontology, Semantic Web, Information Fusion, Counterterrorism

ACM Reference Format:
Panagiotis Mitzias, Efstratios Kontopoulos, James Staite, Tony Day, George Kalpakis, Theodora Tsikrika, Helen Gibson, Stefanos Vrochidis, Babak Akhgar, and Ioannis Kompatsiaris. 2019. Deploying Semantic Web Technologies for Information Fusion of Terrorism-related Content and Threat Detection on the Web: The Case of the TENSOR EU funded Project. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI '19 Companion), October 14–17, 2019, Thessaloniki, Greece.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3358695.3360896

1 INTRODUCTION

Today's Web and social media play an increasingly crucial role in spreading terrorism-related content [2, 7, 26]. The most typical uses of the Internet for terrorism-related purposes include propaganda (including recruitment, radicalization, and incitement to terrorism), financing, training, planning (including through secret communication and open-source information), execution, and cyberattacks [23]. In order to mitigate the widespread usage of the Internet for such malevolent intentions, authorities are deploying sophisticated systems that automatically analyse content on the Web and on social media [18]. Thus, several counterterrorism systems already exist for analysing text [8], multimedia [17], as well as social media content [4, 13, 14] on the Web.

However, since each of these systems is an isolated solution, online intelligence and security investigators often face the key caveat of having to cope with a diverse array of heterogeneous sources and formats [1]. In addition, important information might be available, but only in unstructured sources, which are difficult to access and retrieve. Moreover, as data volumes increase, extracting intelligence and knowledge from it becomes even more challenging. Thus, the challenge of fusing all this available information under one uniform format would entail substantial benefits for facilitating authorities in their operations.

Semantic Web technologies [6] can alleviate this problem by assisting investigators in all pivotal aspects [15]: information fusion and processing of vast amounts of seemingly irrelevant data, smart/semantic search, terrorist network analysis [25], spatiotemporal analysis, sharing and aggregating information. The pivotal role of the Semantic Web technologies backbone is played by ontologies, which are controlled vocabularies of terms relevant to a domain of discourse, along with a set of relations on the terms of that vocabulary that enforce a logical structure. In diverse and complex systems, ontologies serve as the knowledge representation platform for semantically integrating information from multimodal, heterogeneous sources [32]. An additional benefit is that ontologies are typically coupled with powerful tools for semantic enrichment and reasoning, which can as well be applied in the context of counterterrorism, and analytics and threat detection.

This exactly is the focus of the paper, which presents a unified semantic infrastructure for information fusion of terrorism-related content and threat detection on the Web. This framework is deployed within the TENSOR EU-funded project (https://tensor-project.eu/), and consists of an ontology and an adaptable semantic reasoning mechanism. Our ultimate motivation is to assist in their investigational operations both individual Law Enforcement Agencies (LEAs), as well as large-scale initiatives, like, e.g., the EU Internet Forum that brings together governments, Europol and technology companies to counter online terrorist content.

The paper is organised as follows: Section 2 presents key related work paradigms, including ontologies for modelling terrorism-related concepts as well as systems in th domain that deploy semantic technologies. Section 3 presents the TENSOR ontology, which is the main contribution of this work, followed by the deployed mechanisms for semantic integration (Section 4), semantic enrichment (Section 5), and semantic reasoning (Section 6). The paper concludes with some final remarks and directions for future work.

2 RELATED WORK

In literature there are several attempts to model terrorism-related concepts as ontologies, with the work by Mannes and Golbeck [19, 20] being one of the first endeavours. The authors propose an ontology for representing terrorist activity and address the key issues they encountered during the development of the ontology, mostly focusing on the description of sequences of events and the representation of social networks underpinning terrorist organizations. They also stressed the need for lightweight ontologies, so that the non-expert user can easily understand and expand them. These ideas have greatly assisted us during the development of our ontology.

A more recent paradigm is the Adversary-Intent-Target (AIT) ontology [28]. AIT is a model for semantically representing adversary groups and their intentions, classification of their weapons and attack types, and the relationship between the outcomes of an attack and the various recognized intentions of the adversary group. Quite similarly to our intentions within TENSOR, the AIT model focuses on structuring knowledge in a way that will allow reasoning about which groups would be likely to choose what kinds of weapons to perform which kinds of attack. The commonalities with our proposed TENSOR ontology include the representation of adversaries (terrorist groups), their intents and capabilities, and their use of weapons against targets.

Other relevant approaches include the following:

- An ontology of terrorism for automating the characterization and the classification of terrorist threats at early stages, aiming at a more efficient threat mitigation [5];
- An ontology for monitoring terrorist threats, focusing on monitoring subjects and objects (targets) of interest [12];
- A fuzzy ontology (relationships have degrees of membership) for representing terrorism events [16];
- An ontology for uncovering terrorism-related hidden semantic associations, which was the result of knowledge fusion from several existing ontologies and open knowledge systems, like, e.g., the Global Terrorism Database [11].

Finally, Veerasamy et al. [29] presented an ontology specifically developed for cyberterrorism, which is aimed at identifying whether a cyber-event can be classified as a cyberterrorist attack or a support activity, providing a rich semantic representation of underlying relationships, interactions, and influencing factors.

In comparison, our TENSOR ontology has a wider focus and semantically represents more aspects than the above models. In fact, none of the other models represents online artefacts, online users and user communities. On the other hand, our ontology is currently missing aspects for representing weapons and outcomes of terrorist attacks, as they were not considered relevant within the scope of the TENSOR project, but could be easily added if deemed necessary.

On the other hand, regarding systems that are built on semantic technologies, SemanticSpy [21] attempts to harness the abundance of available data, focusing on the significance of semantically representing relationships as a vital element of criminal and terrorist organization analysis, i.e., on relationships among members, relationships among events etc. Following the approach of having an

ontology serve as a common model for semantically integrating heterogeneous information, the Early Warning System (EWS) [22] is a simulation-based diagnostic support tool that collects information relevant to terrorism threat estimation and intelligence data analysis and attempts to predict terrorism threats as well as the stability of the threat factors. EWS also deploys low-level ontological inference and graph structure analysis. In the same context, the Global Criminal and Terrorist Tracking Framework (GCTTF) [33] deploys Semantic Grid technologies, aiming to incorporate the advantages of the Grid, Semantic Web, and Web Services.

Though promising, the above approaches all share the drawback of not taking full advantage of semantic technologies, largely ignoring the inherent capability of Semantic Web infrastructure for deploying semantic enrichment and reasoning mechanisms. Our proposed framework attempts to deliver such a combination of features that can be fully parameterized by end users.

3 THE TENSOR ONTOLOGY

Ontologies constitute the key component of the deployed Semantic Web technologies and play the role of the uniform model for semantically integrating the heterogeneous information coming from other components. The design of the TENSOR ontology relied heavily on close collaboration with Law Enforcement Agencies (LEAs) in the project's consortium, all of whom have extensive experience in countering online terrorism.

The design of the TENSOR ontology is based to some extent on established existing models. We rely on *SIMMO*, a model for describing socially interconnected multimedia-enriched objects that integrates in a unified manner the representation of multimedia and social features in online environments [27]. Since SIMMO is not available as an ontology, we simply adapted its key constructs and properties into the TENSOR ontology.

Moreover, we applied ontology reuse, by importing two well-established ontologies: *SIOC* (Semantically-Interlinked Online Communities) that provides the vocabulary for describing information from online communities [9], and *FOAF* (Friend of a Friend), which is an ontology describing persons, their activities, and their relations to other people and objects [10]. In more detail, we rely on FOAF's `Agent` class to represent groups and persons, and extend it with more specialized representations, like, e.g., terrorist groups and persons of interest. Furthermore, we extend SIOC's `Item` definition representing arbitrary online content items with our definition of an `Artefact`, adopting related properties inherited from SIOC, like, e.g., `mentions` and `sharedBy`.

3.1 Ontology Formalization and Overview

The TENSOR ontology is formalized in OWL 2 [30], the Web Ontology Language, which is a W3C recommendation used by the Semantic Web community as the de facto standard for developing ontologies. Thus, we capitalize on its wide adoption as well as its formal structure and syntax, which is based on Description Logics (DLs), a family of knowledge representation formalisms characterised by logically grounded semantics and well-defined reasoning services [3].

Figure 1 depicts the core concepts and associations of the TENSOR ontology. As already described, its primary aims are to (a) integrate intelligence from online heterogeneous content and, (b) establish links between online content items and human entities. As illustrated in the diagram, a key entity in the TENSOR ontology is `Agent`, which may represent an individual (i.e., a `Person`) or a `TerroristGroup`. Each individual is associated with one or more `UserAccounts` that contain information regarding the person's online presence, e.g., avatars and nicknames. Additional information linked to agents may refer to their `Ideology`, `Cause`, and `ModusOperandi`. Online content, on the other hand, is represented via the `Artefact` entity, and may refer to web pages, social media posts, images, videos and audio. Intelligence extracted from the artefacts may link an `Artefact` entity to physical locations, events, actors etc. Furthermore, each `Artefact` is also linked to an `Intent`, which represents the purpose behind the specific online content item.

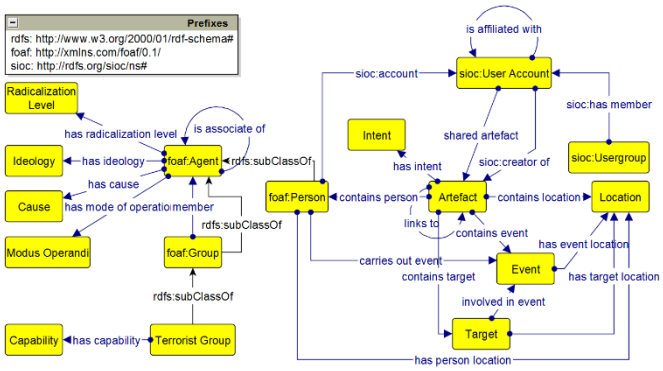

Figure 1: TENSOR ontology core concepts and associations.

3.2 Ontology Metrics

Table 1 includes some key metrics for the TENSOR ontology. Metrics # 1-4 and # 6-8 are self-explanatory and do not require further explanations. The DL expressivity metric refers to the underlying variety of the Description Logics used for representing the ontology: *AL* indicates the base attributive language that allows atomic negation, concept intersection, universal restrictions, and limited existential quantification. Moreover, *U* indicates the presence of concept union, *I* indicates the presence of inverse properties, while *(D)* denotes the use of datatype properties, data values or data types.

Attribute richness is defined as the average number of attributes (slots) per class, giving an indication of both the ontology design quality and the amount of information pertaining to instance data. The more slots defined, the more knowledge the ontology conveys. The value of 2.126984 demonstrates a relatively high attribute richness for the TENSOR ontology, especially when taking into account the fact that there are several classes that are merely enumerations of values (i.e., they correspond simply to sets of instances, like e.g., classes `Intent` and `Modus Operandi`), without having attributes associated to them, and thus do not contribute to the metric's value.

Inheritance richness is defined as the average number of subclasses per class and describes the distribution of information across different levels of the ontology's inheritance tree. It is a good indication of how well knowledge is grouped into different categories

and subcategories in the ontology. This metric distinguishes a horizontal from a vertical ontology; thus, the value of 0.555556 indicates an ontology that is rather horizontal, namely, it represents a wide range of diverse aspects, without delving too deep into their specialisations. On the other hand, any value higher than 1 would indicate an ontology that tends to be vertical.

Relationship richness is defined as the ratio of the number of (non-inheritance) relationships divided by the total number of relationships in the ontology and reflects the diversity of the types of relations. An ontology containing only inheritance relationships conveys less information than an ontology that contains a diverse set of relationships. A value of 0.825 indicates an ontology that conveys rich information to the user, since more than 4/5 of the relationships are non-inheritance (i.e. non-taxonomic).

Average population corresponds to the number of instances compared to the number of classes and is an indication of the ontology population quality. The only instances included in the TENSOR ontology are the types of `Intent`, `Modus Operandi` etc., while the ontology does not include any actual data, namely, instances of analyses, potential suspects etc. The latter pieces of information will be added after the TENSOR platform is being deployed in practice. Therefore, the value of 0.555556 for this metric is considered low.

Finally, *class richness* is related to how instances are distributed across classes. The number of ontology classes that have instances is compared with the total number of classes, giving an overview of how well the knowledge base utilises the knowledge modelled by the schema classes. For the same reason as for the average population, the low value for the TENSOR ontology (0.095238) is reasonable, since the ontology is mostly a schema and is not yet populated with instances, i.e. the ontology does not contain data that exemplifies the class knowledge existing in the schema. This process takes place during the semantic integration operation that is described next.

Table 1: Key ontology metrics.

#	Metric	Value
1	Class count	63
2	Object property count	133
3	Data property count	67
4	Individual count	35
5	DL expressivity	$ALUI^{(D)}$
6	SubClassOf axioms count	35
7	Equivalent classes axioms count	6
8	Disjoint classes axioms count	26
9	Attribute richness	2.126984
10	Inheritance richness	0.555556
11	Relationship richness	0.825
12	Average population	0.555556
13	Class richness	0.095238

4 SEMANTIC INTEGRATION

Semantic integration (also often referred to as "*semantic fusion*") refers to the integration of data gathered from different (and typically multimodal) sources into one common form, allowing thus for data management to be performed in a uniform manner [32]. The process of translating incoming data into class instances, relationships and properties is called "*ontology population*" [24], and results in having the diverse knowledge derived from various TENSOR components semantically represented and interconnected into complex semantic graph structures. This allows its further exploitation by elaborate semantic reasoning rules (see Section 6). For instance, in the case of an incoming Twitter post (i.e., a tweet), the received information would look like this:

```
{
  "id": " abc123",
  "domainId": " twitter:tweet:abc123",
  "created": "2019-01-11T15:14:55Z",
  "processed": "2019-01-11T15:14:55Z",
  "type": "twitter:tweet",
  "source": "Twitter",
  "attributes": {
    "retweetCount ": [10].
    "replyCount": [20].
    "quoteCount": [30].
    "favouriteCount": [100].
  }
}
```

In order to insert the above information into the ontology, a respective instance of class `SocialMediaPost` is created, and the following SPARQL query [31] performs the ontology population:

```
INSERT {
  ?artefact rdf:type owl:NamedIndividual .
  ?artefact rdf:type tensor:SocialMediaPost .
  ?artefact sioc:id "abc123" .
  ?artefact tensor:domainId "twitter:tweet:abc123" .
  ?artefact tensor:hasCreationDate "2019-01-
            11T15:14:55Z"^^xsd:dateTime .
  ?artefact tensor:retweetCount 10 .
  ?artefact tensor:replyCount 20 .
  ?artefact tensor:quoteCount 30 .
  ?artefact tensor:favouriteCount 100 .
} WHERE { BIND(tensor:tweet_abc123 AS ?artefact) }
```

As soon as the new artefact is added to the ontology, all associated entities are also retrieved, such as the Twitter profile and tweet location, along with all relationships (links) they are involved in. All this information is similarly translated into SPARQL queries and added to the ontology.

Additionally, the textual content in the tweet, as well as any attached images or videos, are analysed by respective components in the TENSOR architecture. These outputs are again forwarded to the ontology, in order to populate it (via respective SPARQL queries) with additional instances of concepts.

5 SEMANTIC ENRICHMENT

The process of *semantic enrichment* refers to the process of appending knowledge from external sources to the knowledge already existing in the ontology, thus, augmenting its semantics and leading to richer derivations. The rich variety of interconnected, publicly available resources operating under the principles of the Semantic Web allows the establishment of incoming knowledge streams for the discovery of relevant information. To this direction, we have deployed within TENSOR a semantics-enabled mechanism for normalizing textual location descriptors and for specifying the coordinates from locations discovered by TENSOR's text analysis components in, for example, tweets and Twitter profiles. More specifically, by querying established repositories, like *GeoNames*[1] and *OpenStreetMap*[2], the component retrieves, whenever possible, meaningful geographic metadata, such as a full location name, latitude and longitude, and more. For instance, given a detected location "*Delray Beach*", a look-up to OpenStreetMaps provides the following structured data:

```
{
  "place_id": 197514548,
  "licence": "Data © OpenStreetMap contributors,
              ODbL 1.0. https://osm.org/copyright",
  "osm_type": "relation",
  "osm_id": 117912,
  "boundingbox": [
    "26.420527",
    "26.4914774",
    "-80.1302588",
    "-80.0553486"
  ],
  "lat": "26.4614625",
  "lon": "-80.0728201",
  "display_name": "Delray Beach, Palm Beach County,
                   Florida, USA",
  "class": "place",
  "type": "city",
  "importance": 0.627235305540595,
  "icon":"https://nominatim.openstreetmap.org/
          images/mapicons/poi_place_city.p.20.png"
}
```

Consequently, the retrieved information is added to the ontology and is appropriately associated with the existing location instances. The presence of such diverse knowledge about the location of detected artefacts enables a direct filtering of entities based on location name (string matching), actual distance, etc. Moreover, it allows the application of more complex semantic reasoning rules, as presented in the next section.

6 SEMANTIC REASONING

With an underlying ontology in place, intelligent algorithms can perform automated reasoning analysis in order to reveal hidden correlations between content items. This operation is called *semantic reasoning* and is typically based on a set of rules that run on top of the knowledge stored in the ontology and infer the newly derived knowledge. This section presents some indicative semantic reasoning scenarios, which are a result from discussions with LEA partners in the TENSOR project. All of these scenarios have been implemented via SPARQL-based rules, but the actual rule excerpts are omitted from the paper for reasons of brevity.

A rather simple scenario involves the automated detection of common interests between individuals, which nevertheless entails high value for police investigators. Such associations can be derived from posting or re-posting common topics on social media. Since the topics of tweets are mostly expressed via hashtags, a respective semantic reasoning rule detects and relates user profiles that have posted content with more than, say, five common hashtags; this would be an indication of two or more users sharing similar interests. The execution of the rule results in establishing associations between user profiles that satisfy the rule, via property `sharesInterestsWith`.

Moreover, whenever similar topics are published by multiple profiles within a narrow timeframe, and these posts gain vast popularity on the social media platform (e.g., upvotes, retweets and favourites), there is a strong indication that the respective users have a shared source of information, e.g., a common influencer that disseminates material towards distinct popular communication channels. Another reason for this could be that the same person is behind the different social media accounts. A dedicated rule detects the shared sources and associates the respective user profiles via property `hasCommonSourceWith`. More advanced correlation mechanisms could also attach a certainty factor to the aforementioned association property.

Concerning more sophisticated scenarios, our proposed semantic reasoning framework provides an expandable set of services for more elaborate inferences, which permit the entities of interest and other parameters to be given as arguments. This feature allows deeper and more specialized investigations. For instance, the following excerpt accepts as input two social media profile IDs, and investigates the existence of direct connections (relationships) between those, in order to determine the way the two profiles are related (the so-called "*predicates*").

```
SELECT DISTINCT ?predicate WHERE {
  ?user_account_1 rdf:type sioc:UserAccount .
  ?user_account_1 tensor:domainId 'twitter:profile:A'.
  ?user_account_2 rdf:type sioc:UserAccount .
  ?user_account_2 tensor:domainId 'twitter:profile:B'.
  ?user_account_1 ?predicate ?user_account_2 .
}
```

The above rule may return a relationship of type `repliesTo` or `sharesInterestsWith` that connects Twitter profiles A and B.

Furthermore, as described in Section 5, entity locations are enriched with geographic information like latitude and longitude, which enables our implemented dedicated service to perform a search for entities located within a given radius from an entity of interest. Specifically, a call to the service requires the domain ID of an entity and a radius value in kilometres. A SPARQL rule

[1] https://www.geonames.org/
[2] https://www.openstreetmap.org

then performs the necessary calculations and the results contained within the response present a variety of information concerning both the location of the entity of interest, as well as the detected nearby entities. The excerpt below illustrates an indicative result.

```
{
  "entity_of_interest": {
    "domain_id": "twitter:profile:A",
    "location": {
      "name": "Florida, USA",
      "latitude": 27.7567667,
      "longitude": -81.4639835
    }
  },
  "nearby_entities": [{
    "domain_id": "twitter:profile:B",
    "location": {
      "name": "Delray Beach, Palm Beach County,
               Florida, USA",
      "latitude": 26.4614625,
      "longitude": -80.0728201
    }
  }]
}
```

The ability to retrieve the social network of a person of interest arguably constitutes a handy investigation tool. The following rule implements such a feature.

```
SELECT DISTINCT ?user_account_2_domain_id ?predicate WHERE {
  ?user_account_1 rdf:type sioc:UserAccount .
  ?user_account_1 tensor:domainId 'twitter:profile:310376509' .
  ?user_account_2 rdf:type sioc:UserAccount .
  ?user_account_2 tensor:domainId ?user_account_2_domain_id .
  { ?user_account_1 ?predicate ?user_account_2 .}
  UNION { ?user_account_2 ?predicate ?user_account_1 .}
  FILTER (?user_account_1 != ?user_account_2)
}
```

The result of this query, which investigates a given Twitter profile with ID `twitter:profile:310376509`, presents a detailed list of other profiles that have been associated with it, along with the corresponding type of association.

Additionally, we have also implemented a specialisation of the previous service, which discovers the key influencers within the social network of an individual, according to certain conditions. The prerequisites for a profile to be considered as "influencer" is satisfied by the existence of published content with high popularity (e.g., more than a thousand retweets and upvotes). As with all the rules mentioned in this section, these variables can be easily customised by the human operators to get the optimum results.

7 CONCLUSION AND FUTURE WORK

This paper presented the deployment of Semantic Web technologies in the battle for counterterrorism and threat detection on the Web. As confirmed by the LEA partners in the TENSOR consortium, the techniques presented in this paper will be of great assistance to investigators. As noted, specifically the fact that the semantic reasoning rules (see Section 6) are highly customisable by the users is of particular value to LEAs. We are now in the process of evaluating the presented framework and we will soon be in a position to publish our findings. Moreover, after the end of the project, we will work towards making the semantic framework more flexible, in order to be easily integrated to existing systems deployed at LEA premises, and to collaborate more tightly with other third-party systems generating outputs to be appended to the ontology.

ACKNOWLEDGMENTS

This work was supported by the project TENSOR (H2020-700024), funded by the European Commission.

REFERENCES

[1] Syed Ahsan and Abad Shah. 2008. Data mining, semantic web and advanced information technologies for fighting terrorism. In *2008 International Symposium on Biometrics and Security Technologies*. IEEE, 1–5.
[2] Imran Awan. 2017. Cyber-extremism: Isis and the power of social media. *Society* 54, 2 (2017), 138–149.
[3] Franz Baader, Ian Horrocks, and Ulrike Sattler. 2005. Description logics as ontology languages for the semantic web. In *Mechanizing mathematical reasoning*. Springer, 228–248.
[4] Leslie Ball. 2016. Automating social network analysis: A power tool for counterterrorism. *Security Journal* 29, 2 (2016), 147–168.
[5] Khelifa Benahmed and Nora Assouli. 2017. An Ontology Design for Terrorism Activities. *International Journal of Management and Applied Science* 3, 9 (2017), 90–95.
[6] Tim Berners-Lee, James Hendler, Ora Lassila, et al. 2001. The semantic web. *Scientific american* 284, 5 (2001), 28–37.
[7] Lisa Blaker. 2015. The Islamic State's use of online social media. *Military Cyber Affairs* 1, 1 (2015), 4.
[8] Roger B Bradford. 2006. Relationship discovery in large text collections using latent semantic indexing. In *Proceedings of the Fourth Workshop on Link Analysis, Counterterrorism, and Security*.
[9] John G Breslin, Stefan Decker, Andreas Harth, and Uldis Bojars. 2006. SIOC: an approach to connect web-based communities. *International Journal of Web Based Communities* 2, 2 (2006), 133–142.
[10] Dan Brickley and Libby Miller. 2015. FOAF Vocabulary Specification 0.99 (2014). *Namespace Document. Available online: http://xmlns.com/foaf/spec/ (accessed on 23 November 2018)* (2015).
[11] Mariusz Chmielewski, Andrzej Gałka, Piotr Jarema, Kamil Krasowski, and Artur Kosiński. 2009. Semantic Knowledge Representation in Terrorist Threat Analysis for Crisis Management Systems. In *International Conference on Computational Collective Intelligence*. Springer, 460–471.
[12] Phillip Galjano and Vasily Popovich. 2009. Theoretical Investigation of Terrorism. Ontology Development. In *Information Fusion and Geographic Information Systems*. Springer, 227–239.
[13] Ilias Gialampoukidis, George Kalpakis, Theodora Tsikrika, Symeon Papadopoulos, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2017. Detection of terrorism-related twitter communities using centrality scores. In *Proceedings of the 2nd International Workshop on Multimedia Forensics and Security*. ACM, 21–25.
[14] Ilias Gialampoukidis, George Kalpakis, Theodora Tsikrika, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2016. Key player identification in terrorism-related social media networks using centrality measures. In *2016 European Intelligence and Security Informatics Conference (EISIC)*. IEEE, 112–115.
[15] Jennifer Golbeck, Aaron Mannes, and James A Hendler. 2006. Semantic Web Technologies for Terrorist Network Analysis.
[16] Uraiwan Inyaem, Phayung Meesad, Choochart Haruechaiyasak, and Dat Tran. 2010. Construction of fuzzy ontology-based terrorism event extraction. In *2010 Third International Conference on Knowledge Discovery and Data Mining*. IEEE, 391–394.
[17] George Kalpakis, Theodora Tsikrika, Foteini Markatopoulou, Nikiforos Pittaras, Stefanos Vrochidis, Vasileios Mezaris, Ioannis Patras, and Ioannis Kompatsiaris. 2015. Concept Detection in Multimedia Web Resources About Home Made Explosives. In *2015 10th International Conference on Availability, Reliability and Security*. IEEE, 632–641.
[18] Thomas H. Kean, Lee H. Hamilton, and Nicholas Danforth. 2018. Digital counterterrorism: Fighting jihadists online. *Washington DC: Bipartisan Policy Centre* (2018).
[19] Aaron Mannes and Jennifer Golbeck. 2005. Building a terrorism ontology. In *ISWC Workshop on Ontology Patterns for the Semantic Web*, Vol. 36.

[20] Aaron Mannes and Jennifer Golbeck. 2007. Ontology Building: A Terrorism Specialist's Perspective. In *2007 IEEE Aerospace Conference*. IEEE, 1–5.
[21] Amit Mathew, Amit Sheth, and Leonidas Deligiannidis. 2006. SemanticSpy: suspect tracking using semantic data in a multimedia environment. In *International Conference on Intelligence and Security Informatics*. Springer, 492–497.
[22] Andrzej Najgebauer, Ryszard Antkiewicz, Mariusz Chmielewski, and R Kasprzak. 2008. The prediction of terrorist threat on the basis of semantic association acquisition and complex network evolution. *Journal of Telecommunications and Information Technology* (2008), 14–20.
[23] United Nations Office on Drugs and Crime. 2012. The Use of the Internet for Terrorist Purposes. *Publishing and Library Section, United Nations Office at Vienna. Available online: https://www.unodc.org/documents/frontpage/Use_of_Internet_for_Terrorist_Purposes.pdf (accessed:July 2019)* (2012).
[24] Georgios Petasis, Vangelis Karkaletsis, Georgios Paliouras, Anastasia Krithara, and Elias Zavitsanos. 2011. Ontology population and enrichment: State of the art. In *Knowledge-driven multimedia information extraction and ontology evolution*. Springer-Verlag, 134–166.
[25] Amit Sheth, I Budak Arpinar, and Vipul Kashyap. 2004. Relationships at the heart of semantic web: Modeling, discovering, and exploiting complex semantic relationships. In *Enhancing the Power of the Internet*. Springer, 63–94.
[26] Emily B Stacey. 2019. *Utilization of New Technologies in Global Terror: Emerging Research and Opportunities: Emerging Research and Opportunities*. IGI Global.
[27] Theodora Tsikrika, Katerina Andreadou, Anastasia Moumtzidou, Emmanouil Schinas, Symeon Papadopoulos, Stefanos Vrochidis, and Ioannis Kompatsiaris. 2015. A unified model for socially interconnected multimedia-enriched objects. In *International Conference on Multimedia Modeling*. Springer, 372–384.
[28] Matthew D Turner. 2011. A simple ontology for the analysis of terrorist attacks. *Electrical and Computer Engineering Technical Report. University of New Mexico. Available online: http://digitalrepository.unm.edu/ece_rpts/41 (accessed: July 2019)*. (2011).
[29] Namosha Veerasamy, Marthie Grobler, and Basie Von Solms. 2012. Building an ontology for cyberterrorism. In *Proc. of the 11th European Conference on Information Warfare and Security*. 286–295.
[30] W3C. 2012. OWL 2 Web Ontology Language Document Overview (Second Edition). *W3C Recommendation 11 December 2012. Available online: http://www.w3.org/TR/owl2-overview/ (accessed: July 2019)* (2012).
[31] W3C. 2013. SPARQL 1.1 Overview. *W3C Recommendation. Available online: https://www.w3.org/TR/sparql11-overview/ (accessed: July 2019)* (2013).
[32] Holger Wache, Thomas Voegele, Ubbo Visser, Heiner Stuckenschmidt, Gerhard Schuster, Holger Neumann, and Sebastian Hübner. 2001. Ontology-Based Integration of Information-A Survey of Existing Approaches.. In *Proceedings of the IJCAI-01 Workshop on Ontologies and Information*.
[33] Wai Ming Wong, Dickson K. W. Chiu, and Patrick C. K. Hung. 2008. A Global Criminal and Terrorist Tracking Framework Using Semantic Grid Technologies. In *2008 IEEE World Congress on Services, SERVICES II, Beijing, China, September 23-26, 2008*. 157–162. https://doi.org/10.1109/SERVICES-2.2008.12