2017

Automatically Detecting the Resonance of Terrorist Movement Frames on the Web

Ugochukwu O. Etudo
Virginia Commonwealth University

AUTOMATICALLY DETECTING THE RESONANCE OF TERRORIST MOVEMENT FRAMES ON THE WEB

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

By

Ugochukwu Etudo

Director: Dr. Victoria Yoon
Professor, Information Systems

Virginia Commonwealth University
Richmond, Virginia
May, 2017

Acknowledgement

For Mummy and Daddy

First and foremost, I would like to thank my parents who have always been my rock. I simply do not know what I would be without them. There is no sacrifice that they have not made for me, their patience is never ending. I love you both. Dr. Yoon, my dissertation chair, for inspiring and motivating me, for guiding me and for being the best mentor one could hope for. I thank you. My committee members, for taking the time to read this dense manuscript, for agreeing to be my mentors, for giving me invaluable advice, I thank you all. Mycal, Amy, Michelle, Mike, Bunga, Kim, Christiana, and Alex have always made me feel at home in America; they have believed in me when I did not believe in myself. I do not know that I would have had the mental fortitude to do this without them and so, I thank you.

Table of Contents

Abstract

AUTOMATICALLY DETECTING THE RESONANCE OF TERRORIST MOVEMENT FRAMES ON THE WEB

By

Ugochukwu Etudo

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2017

Director: Dr. Victoria Yoon
Professor, Information Systems

The ever-increasing use of the internet by terrorist groups as a platform for the dissemination of radical, violent ideologies is well documented. The internet has, in this way, become a breeding ground for potential lone-wolf terrorists; that is, individuals who commit acts of terror inspired by the ideological rhetoric emitted by terrorist organizations. These individuals are characterized by their lack of formal affiliation with terror organizations, making them difficult to intercept with traditional intelligence techniques. The radicalization of individuals on the internet poses a considerable threat to law enforcement and national security officials. This new medium of radicalization, however, also presents new opportunities for the interdiction of lone wolf terrorism. This dissertation is an account of the development and evaluation of an information technology (IT) framework for detecting potentially radicalized individuals on social media sites and Web fora. Unifying Collective Action Framing Theory (CAFT) and a radicalization model of lone wolf terrorism, this dissertation analyzes a corpus of propaganda documents produced by several, radically different, terror organizations. This analysis provides the building blocks to define a knowledge model of terrorist ideological framing that is implemented as a Semantic Web Ontology. Using several techniques for ontology guided information extraction, the resultant ontology can be accurately processed from textual data sources. This dissertation subsequently defines several techniques that leverage the populated ontological representation for automatically identifying individuals who are potentially radicalized to one or more terrorist ideologies based on their postings on social media and other Web fora. The dissertation also discusses how the ontology can be queried using intuitive structured query languages to infer triggering events in the news. The prototype system is evaluated in the context of classification and is shown to provide state of the art results. The main outputs of this research are (1) an ontological model of terrorist ideologies (2) an information extraction framework capable of identifying and extracting terrorist ideologies from text, (3) a classification methodology for classifying Web content as resonating the ideology of one or more terrorist groups and (4) a methodology for rapidly identifying news content of relevance to one or more terrorist groups.

Chapter 1 – Introduction

1.1 Introduction

The Web today is characterized by emergent socio-technical interactions (or even assemblages) that result in phenomena, both social and technical (Luczak-roesch 2015), intended and unintended, positive and negative. The unprecedented information diffusive capabilities inherent in Web environments have facilitated a shift in human activities towards digital spaces. In other words, an increasing proportion of our lives are represented in one way or another on the Web. This state of affairs has created a bounty of opportunities for the advancement of the human condition on a global scale. It has been noted that "the benevolent intention of the Internet is to enable global communication virtually distance-free and to make the life and business of global inhabitants more convenient and happier" (Lee 2015, p.iii). Such "distance-free" and often synchronous communication opportunities available on the Web, combined with the growing digitization of human activities have prompted an entrepreneurial boom as embodied in the Web 2.0 concept (O'Reilly 2007). However, this has also created opportunities for malfeasance. This introduction discusses the character of the socio-technical Web, the opportunities it provides for the advancement of humanity and the threats it poses.

1.2 A Socio-Technical Web

The sociotechnical concept was introduced to the information systems (IS) discipline in the inaugural 1977 edition of MIS Quarterly. In this inaugural edition, (Bostrom and Heinen 1977) write about a proposed conceptualization of the information systems development problem as one amenable to a sociotechnical approach. Their perspective can be gleaned from the following excerpt: "The typical goal of an intervention into the technical system is an improvement in task accomplishment, while interventions which focus on the social system tend to look for improvement in [quality of working life] QWL. The [sociotechnical systems] STS approach argues that any intervention must deal with these goals simultaneously" (Bostrom and Heinen 1977, p. 18). Incomprehensibly, Bostrom and Heinen (1977) disregard the studies which brought about the sociotechnical systems perspective. This idea dates back to (Emery and Trist 1969) who write that the core of this perspective involves a matching process characterized "as the *joint optimization of the social and technical systems*. The technical and social systems are *different from* each other in the sense that the former follows the laws of the natural sciences while the latter follows the laws of the human sciences and is a purposeful system. Yet they are *correlative* in that one requires the other for the *transformation* of an input into an output, which comprises the functional task of a work system" (Emery and Trist 1969).

Today's Web mirrors the characteristics of a sociotechnical system. Conceiving the Web as such is crucial to understanding the motivation of this thesis. To conceive of the Web as a sociotechnical system, one must accept the basic notion that the technological and social components of the web are engaged in a correlative relationship where emergent phenomena are characterized by both social world and technological features of the Web. I begin by discussing the Web 2.0 concept.

1.3 Web 2.0

The concept was first publicized by Tim O'Reilly (O'Reilly Media) and his colleagues as they sought to characterize a new wave of innovations and practices on the Web during a conference organized for that purpose (O'Reilly 2007). O'Reilly writes that "you can visualize Web 2.0 as a set of principles and practices that tie together a veritable solar system of sites that demonstrate some or all of those principles, at varying distance from that core" (O'Reilly 2009, p. 6). These core principles are summarized in Figure 1. A key defining feature of Web 2.0 is the recognition amongst internet companies that the Web is a platform on which service-oriented software ecosystems can be delivered for consumption by end users. The conceptualization of the end user by these companies is also crucial to understanding the concept. Web 2.0 is characterized by a view of users as publishers of content, where the more content these users publish, the more engaging and useful the eco-system becomes. Further, Web 2.0 applications transcend any notion of device dependence. This device transcendence is in lock-step with the view of the Web as a platform.

Strategic Positioning
- The Web as a Platform

User Positioning
- You control your own data

Core Competencies
- Services not packaged software
- Architecture of participation
- Cost-effective scalability
- Device-independent software
- Harnessing collective intelligence

Figure 1. Web 2.0 Core Principles (adapted from (O'Reilly 2009))

I note that the Web 2.0 concept is not an academic one. It is born of industry hype and touted by market analysts and the news media. Indeed, many of the technologies that are used to characterize this concept had existed and were in use long before the moniker was coined. Web 2.0 is, however a useful umbrella term for the set of applications and corresponding socio-technical phenomena that characterize them. For example, Ellison and Boyd (2013, p. 160) write of social media (social media is perhaps the most visible and pervasive of Web 2.0 applications): "what makes social media significant as a category is not the technology but rather the sociotechnical dynamics that unfolded as millions of people embraced the technology and used it to collaborate, share information, and socialize." This astute observation is true not only of social media but of all of those "categories" which can be said to fall under the Web 2.0 umbrella. In other words, the Web is increasingly characterized by socio-technical dynamics as user participation, near ubiquitous, and device independent access have become the foci of its technological offerings. The remainder of this section will explore a few, key examples of sociotechnical phenomena characteristic of this trend.

1.3.1 Online Communities

Early conceptualizations of online communities resulted in the identification of four types: communities of transaction (see Ou et al. 2014 for an example of the chinese concept of Guanxi unfolding in online communities of transaction), communities of interest, communities of fantasy and communities of relationships (Armstrong and Hagel 2000). A rich field of research has emerged in the domain of online communities and participation in such communities is widespread. For example it is estimated that 15% of male Internet users are active participants in Reddit, a diverse online community where individuals can find "subreddits" that focus on topics of personal interest (Duggan and Smith 2013). This development is indicative of the extent to which human activity has shifted to and is increasingly patterned by the Web. In marketing research, for example, online communities have been studied using novel methodologies, such as netnography and a Web-adapted variation of ethnographic method (Kozinets 2002), and are positioned as cheaper alternatives to traditional focus group approaches. Online communities, then, span virtually all aspects of human life. Health oriented online communities are particularly indicative of this trend (Conrad et al. 2016). In this domain such communities are used to promote not only positive interventions such as in (Myneni et al. 2016) where individuals are encouraged to adopt healthier behaviors, but also negative interventions such as “proanorexia” communities (Oksanen et al. 2016) where individuals may be inspired to engage in very unhealthy behaviors. An uncontentious conclusion here is that online communities are pervasive, emblematic of the sociotechnical nature of today’s Web, and have real impacts on human life.

1.3.2 Social Network Sites

Social media is arguably the most visible of Web 2.0 innovations; indeed many online communities are hosted on social media platforms. While the technology existed prior to the coining of the Web 2.0 concept (think Friendster) its widespread adoption is a fairly recent phenomenon. Social network sites (SNSs) have been defined as: “a networked communication platform in which participants 1) have uniquely identifiable profiles that consist of user-supplied content, content provided by other users, and/or system-level data; 2) can publicly articulate connections that can be viewed and traversed by others; and 3) can consume, produce, and or interact with streams of user-generated content provided by their connections on the site” (Ellison and Boyd 2013, p. 158). According to the Pew Internet Research Center, the use of social medial has jumped nearly tenfold between 2005 and 2015 such that 65% of adults currently use SNSs compared to 5% in 2005 (Perrin 2015). Such a meteoric rise in usage rates indicates both a sharp resonance of the features of social media with the public and a deeply penetrating shift in the fabric of society as socialization moves online. Of course, the widespread adoption of social media platforms entails the adoption of the three focal features of social media previously identified. In other words a rapidly growing number of individuals are creating uniquely identifiable profiles consisting of their own supplied content, are broadcasting their connections with other individuals, and are consuming, sharing and otherwise interacting with streams of content generated by other users. One particularly visible and often cited consequence of such widespread adoption is a dramatic increase in information propagation potentiality. This stems from the low publication barriers characteristic of social media and the (generically speaking) ability for individuals to interact with and reproduce streams of content generated by other users (Blanquart and Cook 2013; Pentland 2014). It has been shown that individual behavior is heavily influenced the characteristics of information propagation in their online social networks. Using data from the social trading (investment) platform eToro, Shmueli et al (2014) discovered regularities that characterize the manner in which social network topologies change over time, and Pentland (2014) demonstrates associations between these changes and the behaviors of members of the social network. More specifically these authors show that there exist optimal configurations of the social network topology of eToro that are facilitative of idea flows which lead to significantly increased returns for traders. Social network sites rapidly accelerate idea flow. In this environment, idea flow strongly influences individual behavior and decision making.

The example of eToro (i.e. Web technology influencing real world human behavior) is not idiosyncrasy of social investing. It pervades social digital activity. A non-controversial and important conclusion of this discussion is that *the dynamics of human interactions on social networking sites strongly influence real-world human behavior*.

1.4 Research Motivation – Lone Wolf Terrorism

Accepting the Web as an emergent sociotechnical assemblage enables the argument that robust systems capable of capturing information on a Web-wide scale are imperative in the study of radicalization as it concerns the lone wolf. That if the migration of social activities to digital, Web-enabled spaces is evident in legitimate uses of the Web, it is also evident in illegitimate uses of the Internet. This simple proposition serves as an imperative to study radicalization on the Web especially as it concerns the lone wolf. These ideas are developed further in the following sections.

On February 11, 2003, then director of the Federal Bureau of Investigation (FBI) Robert S. Mueller, III appeared before the Select Committee on Intelligence of the United States Senate to deliver remarks on the then two-year-old war on terror. Discussing the nature of the terrorist threat to the US, the director remarked: "I am particularly concerned about loosely affiliated terrorists and lone offenders, which are inherently difficult to interdict given the anonymity of individuals that maintain limited or no links to established terrorist groups but act out of sympathy with a larger cause" (Mueller III, 2003)[1]. Three important points are made here, where each point maps to a definitional characteristic of lone-wolves. The first point is that lone offenders are "**loosely affiliated**," the second is that this loose affiliation leads to **anonymity** and difficulty in detecting lone attackers, and the third is that formal affiliation to terrorist organizations is eschewed for **sympathy to a larger cause**. The director's remarks on lone wolf terrorists are in keeping with the consensus in the academic literature where available definitions focus on these individuals' lack of affiliation with terrorist organizations, individualized (syncretic) ideologies, and anonymity (Brynielsson et al. 2012; McCauley and Moskalenko 2008; Moskalenko and McCauley 2011; Spaaij 2010, 2012). For example the definition provided in (Spaaij 2010, p. 856) states that "lone wolf terrorism involves terrorist attacks carried out by persons who (a) operate individually, (b) do not belong to an organized terrorist group or

[1] Robert S. Mueller III, "War on Terrorism." Testimony before the Select Committee on Intelligence of the United States Senate, Washington, DC, 11 February 2003

network, and (c) whose modi operandi are conceived and directed by the individual without any direct outside command or hierarchy."

The threat of lone wolf terrorism is becoming the new global frontier in the fight against terrorism (Sageman 2011). Where traditional terrorists affiliated with terrorist organizations are being combated via traditional means (i.e. infiltration of terrorist networks) (Brynielsson et al. 2012; Tucker 2001), novel techniques need to be developed and deployed to meet the ever growing law enforcement challenge posed by lone wolves. To understand the need for and ideal nature of such novel techniques, a closer analysis of the nature of lone offenders is required. The following paragraphs discuss each of the three definitional characteristics of lone wolves that were identified by former FBI Director Robert S. Mueller, III and corroborated in the academic literature on lone wolf terrorism: (1) loose affiliation (2) anonymity (3) sympathy to a higher cause.

Lone wolves eschew traditional affiliations with organized terrorist groups and instead are ideologically sympathetic to some higher cause. These higher causes are, however, drawn from organized terrorist groups. The ideological model of lone wolf terror has been discussed by a number of authors and many common threads arise in their respective analyses. Spaaij writes of lone wolf ideology: "the protagonists of lone wolf terrorism often combine the broad structures of a more prevalent extreme ideology with their own personal grievances. Thus, lone wolves tend to create their own individualized ideologies from a mixture of broader political, religious or social aims and personal frustrations and aversion" (Spaaij 2012, p. 38). Smelser[2] writes of the

[2] It is important to note that Smelser is not writing about lone wolves but about terrorist ideology more broadly and in the group context particularly. His ideas in this regard are quite relevant to this thesis even though they refer to a group context and to a broader conception of terrorist ideology. This is the case because it is a core assertion of this work that the collective action framing that has been documented to occur in extremist groups is highly relevant to the lone wolf context in the digital age. For lone wolves, this process occurs online.

commonalities across a hand-picked number of instances of lone wolf terrorist behavior that they "all arose in the context of outside political domination, all envisioned the disappearance of the oppressors, and all had a vision of the dramatic creation of an ideal life" (Smelser 2007, p. 57). Elaborating, Smelser remarks that the impulses behind these episodes are constant while their content is diverse: "though differing radically in context, they all have the same impulses: explaining the suffering of a people, assigning responsibility for it to inimical agents or forces, anger at and punishment for the agent, and the vision of an ideal and often blissful future condition free from pain and suffering" (2007:57). In essence there appears to be a basic structure that is characteristic of the ideologies behind terrorist activity. When compared, the contextually and substantively diverse events share remarkable parallels surrounding the structure of their ideology. Smelser notes that there exists a "widespread human reaction to experienced dispossession and deprivation" and that it is important to "appreciate what these systems of belief contribute to the political situation of believers, to their motivation, and to their agendas for action" (2007:58). The above discussion highlights the importance of the concept of ideological **syncretism** in understanding the commonalities found in the ideological structure of lone wolves. Droogers (1989, p. 7) defines syncretism as that which "refers neutrally and descriptively to the mixing of religions". The term, as employed here refers, instead, to the mixing of ideologies. Syncretism is highly characteristic of the ideologies of individuals who have engaged in terrorist activities as there is a tendency amongst these individuals to develop ideology that is constructed from a mix of ideas sourced from multiple sources (Smelser 2007; Spaaij 2012). The result is a manufactured hodgepodge of beliefs drawn from disparate cultural, historical and theological sources.

The anonymity of lone wolf perpetrators of terror has been of particular concern to law enforcement. Anonymity is closely tied to the lack of formalized affiliations that is characteristic of such actors, or what Spaaij (2010) described as modi operandi [that] are conceived and directed by the individual without any direct outside command or hierarchy. The presence of a direct outside command hierarchy permits the use of traditional counter terror mechanisms such as group infiltration. Where this hierarchy is absent, new approaches must be developed and deployed. While formal command and control systems are irrelevant in the lone wolf condition, there does exist an important commonality with traditional terrorist networks; that is, in both traditional and lone wolf conditions, the processes of collective action framing[3] unfold in a similar fashion. Lone wolves display a strong tendency to identify with extremist groups or more informal collections or pockets of **online** sympathizers. In recent times, that is post 9/11, there is strong indication that "lone wolves may be seeking direction through venues other than organizations: namely, via networks of like-minded activists found online or on cable television." (Hamm and Spaaj 2015, p. 9). In the traditional condition, collective action framing occurs via well-established channels, within the context of a formally defined social movement with a command and control hierarchy. It is in the context of organized movements that the framing perspective on collective action was developed. However, this thesis argues that the framing perspective is also relevant to lone wolves where such framing occurs in “more informal collections or pockets of online sympathizers.”

[3] More on collective action framing in chapter 2. For now, I quote Snow & Byrd who write that “the framing perspective on collective action and social movements views movements not merely as carriers of existing ideas and meanings, but as signifying agents actively engaged in producing and maintaining meaning for constituents, antagonists, and bystanders” (Snow and Byrd 2007, p. 123).

To conclude, the Internet has provided an unprecedented avenue for troubled individuals to discover groups with ideologies capable of explaining away their troubled existence. This trend is relevant to this work when those ideologies present beliefs that may be considered as 'extreme'[4]. Indeed it has been reported in (Hamm and Spaaj 2015, p. 8) that "lone wolves may be seeking direction through venues other than organizations: namely, via networks of like-minded activists found online or on cable television." Hamm and Spaaij found, in their study of lone wolf attacks, that while 63% of pre-9/11 lone wolf attackers were affiliated (either through ideological affinity of direct affiliation) with organized extremists only 42% of post 9/11 lone wolf attackers had such affinities. Some (numerous examples exist) specific examples include the Madrid bombers that gained ideological inspiration from a document posted on the Global Islamic Media Front website in December 2003 and the Hofstad Group that originated in the Netherlands employed online media for recruitment and radicalization efforts (Sageman 2011).

1.5 Research Issue

Lone-wolf radicalization is increasingly taking place on the Web. In response to this trend, much work has been done towards semi-automated methods that rely on text analytics to identify individuals espousing terrorist ideologies on the Web. These works are based on the implicit (hardly if ever stated) assumption that the writings of an online persona can be used to link that persona with terroristic intent and/or terroristic affinity. The framework presented in this dissertation joins this stream of work in attempting to identify terroristic content on the Web. To

[4] "Extremism can be used to refer to *political ideologies* that oppose a society's core values and principles. In the context of liberal democracies this could be applied to any ideology that advocates racial or religious supremacy and/or opposes the core principles of democracy and universal human rights. The term can also be used to describe the *methods* through which political actors attempt to realise their aims, that is, by using means that 'show disregard for the life, liberty, and human rights of others" (Stevens and Neumann 2009)

be clear, any classifications of text in this dissertation *do not indicate that the individual responsible for the text is a lone wolf terrorist*. The proposed artifact simply identifies when text resonates the ideological framing of a group-based terroristic actors. Only human agents should make the final determination with respect to whether or not an individual poses a high risk of terroristic activity. This is because any machine (and indeed human) classifications are subject to Type I and Type II errors. In this domain, the cost of such errors can be grave and life altering for monitored individuals (in the case of false positives) and be disastrous for innocent people (in case of false negatives). The designs presented here, when used to monitor social networks, should be seen as winnowing down a massive amount data. They should never be interpreted as capable of identifying terrorists.

While there exist numerous theoretical and empirically backed frameworks in the literature on social science, on terrorism in general, and on lone wolf perpetrators of terror in particular (Hamm and Spaaj 2015; McCauley and Moskalenko 2008, 2014; Moskalenko and McCauley 2011), these *theoretical* frameworks have not been incorporated into any *IT* frameworks for detecting radical, terroristic content on the Web towards the interdiction of lone wolf terrorism. This implies (and we show) that extant IT frameworks for detecting such content on the web are a-theoretical[5]. Consider the following examples. In a study presenting IT frameworks for extracting insight from free-text Web data for security informatics, Agarwal et al. (2015) describe a tripartite framework of (1) intelligence and security informatics, (2) online social media platforms and (3) text mining and analytics, effectively constraining attempts at intelligence gathering on the Web to these three topic areas. However, social scientific,

[5] I intentionally skirt around discussions of kernel theories and design theories (see Gregor and Jones 2007). By a-theoretical, I mean lacking any theories of human behavior.

behavioral theory has never been considered as a useful component in developing IT frameworks for the detection of radical terrorist content online. In another study, Agarwal and Sureka (2015) present an IT framework for detecting radicalization on Twitter. Again, this research presents neither robust conceptualization of the notion of radicalization, nor guiding theoretical framework (yet the paper is entitled "Using KNN and SVM Based One-Class Classifier for Detecting Online Radicalization on Twitter"). Indeed, all that is known to the classifiers described in the study regarding radicalization is a set of Tweets seeded with hashtags #Terrorism, #Islamophobia and #Extremist which promote the undefined concepts of "hate" and "extremism." In (Alizadeh and Cioffi-Revilla 2014) computational models are presented that examine the distributional properties of radical opinions on social media. This analysis is, again, uninformed by any insight from behavioral theories on radicalization, extremism and social movements. Brynielsson et al. (2012) describe a semi-automatic system for detecting potentially radicalized personas on the Internet. Their approach is multi-faceted with tasks such as website discovery, website classification, and online persona merging amongst others. However, the a-theoretical approach persists. As an example, the authors propose a set of decision nodes that lead to a determination about a persona's potential radicalization. These decision nodes are foundational to their IT-framework. However, the development of the decision nodes appears to have been done rather arbitrarily, using "common-sense." In another, perhaps more striking example, Chen et al. (2008) presents a collection of research in the domain of terrorism informatics. The second half of the text, titled "terrorism informatics to support prevention, detection, and response" 14 research papers are presented as chapters. None of these 14 works include any theoretical foundation from the behavioral sciences on terrorism. In the comprehensive review of this literature presented in Chapter 2, each of these works will be examined. It shall be the finding that there is no extant study that develops an IT framework for detecting radical, terroristic content, (and eventually radicalized individuals) on the internet that takes seriously the rich body of theories on terrorism found in the social sciences.

1.6 Research Questions

The overarching question that motivates this research is as follows: *How can an IT artifact be developed to detect online personas that pose a risk of violent behavior in the name of one or more radical ideologies?* An individual poses a risk of violent behavior when they espouse violent terroristic ideologies. Such individuals are potentially lone-wolf terrorists. However, the final, actionable determination in that regard *should never* be made by automated means. Determining that an individual espouses violent terroristic ideologies is done using a theoretical framework grounded in the framing perspective of collective action. The framework is also capable of identifying radical terroristic content in general. Our claim that individuals who espouse terroristic ideologies are *potentially* lone wolves is undergirded by the proposition that lone wolves, while not part of a formal organization, engage in and are influenced by collective action framing that occurs in online communities. Given the overarching question motivating this research, the following six specific research questions arise:

RQ1: *What are the concepts and inter-concept relationships that are characteristic of radical ideologies and how can they be used in articulating the structure of a knowledge representation of ideologies?*

RQ2: *How may websites that promote radical ideology be detected?*

RQ3: How can the knowledge representation produced as a result of RQ1 be populated based on free text derived from Web-based sources?

RQ4: How may the knowledge representation populated in RQ3 above be used to automatically detect the ideology of an online persona?

RQ5: How can events relevant to individual, radical ideologies stored as part of the knowledge representation produced by RQ1 be detected from news sources available on the internet?

RQ6: What coordinating framework is necessary to integrate all of the above in order to effectively identify online personas for additional scrutiny by security operatives?

The first research question, RQ1, addresses the need for a sound, theoretically grounded ontological representation for extremist ideologies. That is, identifying those primitive (general) concepts that are relevant to all ideologies and identifying the relationships between them to the extent that the resultant concepts and inter-concept relationships are sufficient to model all facets of ideology. The process of developing this representation shall be guided by ontology engineering formalisms, with a focus providing a concise, minimally committed ontological representation of the domain, a thorough review of the literature on collective action framing, and a detailed content analysis of the writings of a selected sample extremist thought leaders.

The second research question, RQ2, addresses an important aspect of the proposed system, that is, its ability to identify radical content on websites. While this research question is satisfied with the same technique used to satisfy our fourth research question, we highlight it here as an important initial step towards the eventual development of an autonomous web-crawling system. In (Chen 2011) a labor intensive (manual) framework for identifying radical websites on the

"Dark Web." The automatic terrorist frame resonance detection framework proposed here is designed with this framework in mind. That is, it is capable of automating the manual components of the (Chen 2011) framework. While the development and evaluation of an automatic crawler for identifying radical content on the dark web is beyond the scope of this dissertation, the artifact developed herein is discussed in the context of a future crawler.

The third research question, RQ3, prompts an investigation into the domain of knowledge acquisition from text, towards the end of populating the ontological representation resulting from RQ1. To satisfy the requirements of RQ3 we are required to develop an automated system for knowledge discovery. This system must be able to identify and extract textual components that are semantically equivalent to the concepts and inter-concept relationships defined in the ontology from RQ1. To ensure generalizability of the approach we rely primarily on statistical, machine learning techniques for processing textual data sources.

The fourth research question requires us to develop strategies for classifying content using populated ontologies designed in RQ1 and populated in RQ3. We design an ensemble classification technique towards these ends. The ensemble classification technique applies to numerical representations of the resonance of terrorist movement frames asserted in baseline ontologies and discovered from websites and online personas of concern. We develop a resonance scoring method that produces a vector of resonance scores (aforementioned numerical representations) for any combination of ideology and content of concern. These vectors form the observations for which classification is performed.

Meeting the requirements of the fifth research question, RQ5, entails the development of tools that can be used to monitor news sources for events relevant to the radical ideologies which populate the ontology resulting from RQ1. To satisfy this research question, the ontology developed for RQ1 and populated via RQ3 is examined for its usefulness with respect to detecting news that may trigger individuals radicalized along one or more of its stored frames. News articles can be represented in terms of the ontology developed for RQ1. Articles represented in this way can be queried using any of the structured languages available for ontologies. We meet RQ5 by illustrating how a corpus of news articles can be easily analyzed for triggers relevant to one or more radical ideologies stored in our ontology, using our framework.

Finally, to satisfy the requirements of RQ6, an integrating framework will be developed to govern the interactions of the components necessary to realize the vision laid out in RQs 1-5.

1.7 Research Contributions and Significance

It has long been acknowledged that the Web is a key battleground in the fight against terrorism. Indeed, using the Web as a source of counter terrorism intelligence is amongst the primary foci of the interdisciplinary Information Security Informatics (ISI) field (Chen 2011). Numerous research publications have been produced that in some way leverage Web data to gain some counter terrorism insight (e.g. Agarwal and Sureka 2015; Alizadeh and Cioffi-Revilla 2014; Brynielsson et al. 2012; Johansson et al. 2013, 2015; Qin et al. 2008). However, much of the work done in this vein is without solid (if any) theoretical foundation. In particular, there are no studies that have used the theoretical insights available in the literature on collective action framing to inform the analysis of extremist propaganda. The notion of collective action framing seeks to illuminate the processes through which social movements generate and diffuse mobilizing and counter-mobilizing ideas and meanings in addition to elucidating the nature of these ideas and meanings (Benford and Snow 2000). Embracing this rich body of theory enables the forthcoming analysis to be cognizant of the widely heterogeneous nature of the ideologies that motivate social movements. Using this theoretical lens the IT artifact under development can be sensitive to nuanced but important differences in espoused ideologies even in instances where, on the surface, these ideologies appear to be similar, such as in the case of Islamic movements (Snow and Byrd 2007).

Further, the proposed design artifact is also a pioneering attempt in conceiving an ontology of extremist ideology. The ontology, as explained earlier in this introduction, will define a set of primitive concepts relevant to the description of any extreme ideology that forms the basis of a social movement. The ontology shall reflect the invariants that are characteristic of the theoretical perspectives employed in its development. As an example, the framing perspective asserts that *problem identification and locus of attribution* is invariably present and indispensable part of any movements framing task (Benford and Snow 2000) while (Hamm and Spaaj 2015) argue that radicalization begins with a *personal and political grievance*. Thus the manner in which a movement's ideology attributes a problematic situation to responsible parties is a modeling invariant that will be populated with individual instances of movement problem attributions. There are several benefits to such an ontology. One such benefit is that the ontology informs automated data collection efforts. Data collection from the text of selected websites can be focused on the entities and relations defined in the ontology. Another benefit is the rich reasoning capabilities of modern reasoners (see, for example, Sirin et al. 2007) which enable comprehensive rule-sets to be defined over such an ontology with fully decidable ontology languages available and well supported. Reasoning over the instances (individuals) of the proposed ontology allows the inference of previously unknown facts. In other words, the initial ontological representation ostensibly enables the discovery of new knowledge in the domain. An additional benefit is that the ontological representation enables information fusion capabilities that are indispensable to the framework proposed here. Briefly, because data on extreme ideologies sourced from propaganda sites is represented in the same dialect as the postings of monitored Internet personas the artifact has a direct means of associating the postings of a persona with espoused ideologies. The final benefit of this proposed representation is the fact that it is an application-independent representation and can thus be extended to other systems and be used for rich analyses of extreme ideologies.

The Internet pervades daily life. This pervasiveness has led to a digital shift in several social processes unto the web. While society has accrued many benefits from this shift, unintended consequences abound. One such unintended consequence is the provision of a mass communication platform with global reach for those who hold extreme, dangerous beliefs. This relatively new and almost unfettered ability to broadcast radical propaganda has drawn many disturbed, frustrated individuals into radical social movements they otherwise would not have had ready access to. Motivated by this disturbing and widely recognized trend, this dissertation puts forth an IT enabled framework for detecting radicalized individuals with violent potential. It proposes to do so by developing a novel knowledge base of extreme ideologies which can be used to identify individuals who may be radicalized. This research is the first of its kind to explicitly take into account the rich body of theory available in the framing perspective on collective action, as well as empirically validated process models of lone-wolf radicalization. This robust framework leverages its knowledge base to detect radical ideology in the writing of online personas, to quarantine such personas for further analysis while simultaneously monitoring news sources for potential triggering events.

1.8 Organization of Dissertation

The remainder of this dissertation is organized as follows. The second chapter is a comprehensive, thematic review of the literature relevant to the key themes in this work. It includes discussion on the definitional challenges of terrorism in general and lone wolf terrorism in particular. A model of lone wolf radicalization is presented. The chapter also introduces the framing theory of collective action and details its role within this thesis. Gaps in the literature on terrorism informatics are identified. The third chapter presents a detailed view of the proposed IT framework. In this chapter, a general discussion of the techniques and methods that we employ to develop a framework for automatically detecting the resonance of terrorist movement frames is provided. Given the method descriptions, we discuss the details of our implementation of the proposed system. The fourth chapter evaluates the proposed system with respect to our 6 research questions. We use well-established information extraction and retrieval metrics to evaluate the framework's capabilities with respect to research questions 2 and 4. We respond to the remaining research questions by illustrating exemplar use cases. Finally, in chapter 5, we provide our concluding remarks, focusing on our theoretical and practical contributions and our future research agenda.

Chapter 2 - Literature Review

2.1 Structure of the Review

Below is a thematic literature review spanning three interrelated themes. Each theme represents a major area of inquiry that is constitutive of the research presented herein. The first theme, *lone-wolf terrorism* reviews the academic literature on lone-wolf terrorism. The objective here is to elucidate the meaning of the concept as it has appeared in academia. Conclusions from the review of the first theme will, on the one hand, include a working definition and a deeper understanding of the genealogy of the concept and on the other, a working model of lone wolf terrorism that will undergird the design of an IT artifact for the detection and monitoring of lone wolf radicalization on the Internet. The second theme, *the framing theory for terror movements* is explored in a manner similar to the first. That is, the object of reviewing the literature on the framing perspective on terror movements is to elucidate the meaning of the concept as it appears in the relevant literature, paying particular attention to its evolution over time and its standing in the state of the art on the one hand. On the other hand, the exploration of the framing perspective on terror movements shall result in the formulation of a coherent theory of ideological framing in terrorist movements. The third and final theme, the *detection of radical content and radical individuals on the Internet* will review the relevant academic work towards automatically or semi-automatically detecting either or both of radical content or radical individuals on the Internet. This portion of the thematic review will focus on the kernel theories employed in the various extant designs. The detection of radical content and radical individuals on the Internet is, indeed, the core of the review as it is the overarching concern of this research. This review shall result in the articulation of a theoretical model of lone wolf radicalization on one hand and terrorist ideological framing on the other. In addition, a clear motivating gap in the literature on terrorism informatics for the identification of radicalization on the Internet will be presented.

2.1.1. A Note on Terrorism

A discussion of lone-wolf terrorism necessarily commences with a discussion of the broader concept of terrorism. While an in-depth discussion on terrorism in general is out of scope, there is a need to stake a position on the working definition of the term that will be implicit in its use throughout this dissertation. The academic discipline which claims the study of terrorism as its principal domain of inquiry is generally referred to as terrorism studies. Much of terrorism studies has been characterized by a definitional crisis, where criticism for existing definitions of the term have been that these definitions are directed at policy making to handle terrorism without and excluding "a social scientific understanding of political violence" (della Porta and Haupt 2012). Della Porta and Haupt cite (Goodwin 2004, p. 260)who writes that "many who have written about terrorism have directly or indirectly involved in the business of counterterrorism, and their vision has been narrowed and distorted by the search for effective responses to terrorism" and (George 1991, p. 92)who writes that "terrorology is intellectually sterile, if not bankrupt, because the construct of terror employed by terroristologists was not developed in response to honest puzzlement about the real world, but rather in response to ideological pressure." Cooper (2001) identifies another critical issue with defining terrorism, that is, the "one person's terrorist is another's freedom fighter" issue. The formulation of the issue is quite self-explanatory. Terrorism appears to be in the eye of the beholder; what might be considered a violent act of terrorist aggression by one individual may be considered warfare, or liberation struggle by another individual. The problem is worsened by the political motives which are present in the choices of what is called terrorism and what is not (Ganor 2002). (Stohl

2008) identifies ten “myths” that are taken as given in both popular and academic contrivances of terrorism, where these myths contribute significantly to the definitional confusion in the field. These are (1) political terrorism is exclusively the activity of non-governmental actors, (2) all terrorists are madmen (3) all terrorists are criminals (4) one person’s terrorist is another’s freedom fighter (5) all insurgent violence is political terrorism (6) the purpose of terrorism is the production of chaos (7) governments always oppose non-governmental terrorism (8) political terrorism is exclusively a problem relating to internal political conditions (9) the source of contemporary political terrorism may be found in the evil of one or two major actors and (10) political terrorism is a strategy of futility.

These definitional difficulties are at least in part attributable to the concept’s social construction. Each new incidence of terror appears to alter the meanings attributed to the concept, as old definitions are “stretched” to fit new contexts. (Collier and Mahon 1993) note of such conceptual stretching that “when scholars take a category developed from one set of cases and extend it to additional cases, the new cases may be sufficiently different that the category is no longer appropriate in its original form. If this problem arises, they may adapt the category by climbing the ladder of generality… As they increase their extension, they reduce the intension to the degree necessary to fit the contexts” The social construction of terrorism has been recognized here (Harre 2004) where a clear analysis of this construction is presented. Some consistencies do appear across the available definitions of terrorism. One key area of consensus is that terrorism manifests as violence intended to induce fear or anxiety in an audience that is different from the targets of violent acts.

The definition developed in (Schmid and Jongman 1988) is based on reactions elicited from terrorism scholars on an initial, all-inclusive definition provided on a questionnaire. The resulting definition takes into account the input provided by the field of scholars:

> "Terrorism is an anxiety-inspiring method of repeated violent action, employed by (semi)clandestine individual, group, or state actors, for idiosyncratic, criminal, or political reasons, whereby—in contrast to assassination—the direct targets of violence are not the main targets. The immediate human victims of violence are generally chosen randomly (targets of opportunity) or selectively (representative or symbolic targets) from a target population, and serve as message generators. Threat- and violence-based communication processes between terrorist (organizations), (imperiled) victims, and main target (audience(s)), turn [an audience] into a target of terror, a target of demands, or a target of attention, depending on whether intimidation, coercion or propaganda is primarily sought."

This definition points to that which delineates terrorism from other types of violent action. Those who are the immediate victims of terrorist violence are not its intended targets. Such acts of violence are therefore symbolic in nature. Violence as symbolism indicates a communicative intention behind acts of terror where the perpetrators seek to deliver a message to a wider audience than the immediate victims by inspiring anxiety in that audience (Stohl 2008). Sluka (1999) recognizes the importance of acknowledging that the social effects of the terrorist act are substantially more important the act itself. The communicative intention behind acts of terror is that which delineates it from other acts of violence (Young and Findley 2011). In addition, this definition includes states as potential perpetrators of terrorist acts of violence. (Stohl 2008, p. 6) observes that "while a primary purpose of terrorism as practiced by challengers to governmental authority, is the production of chaos to accelerate social disintegration to demonstrate the inability of the regime to govern or impose order, it remains the case that the most persistent and successful use of terror both in the past and in the modern era has been demonstrated by governments for the purpose of creating, maintaining and imposing order."

The Schmid and Jongman definition quoted above is the operative meaning of "terrorism" as used in this dissertation. Many other definitions of terrorism exist, and the field of terrorism studies is not in agreement with respect to any full definition. However, there appears to be little debate about the core component of terrorism, that which dilineates it from other forms of violence – communicative intent. By selecting a definition that focuses on this communicative intent, much of the definitional debate can be skirted while having a working conceptualization that is both useful and minimally controversial.

2.2. Lone Wolf Terrorism

Terrorism is typically viewed through the lens of collective and organized action. However, this dissertation is concerned primarily with individuals who conduct terrorist acts independently of organized groups, so-called lone wolves. The degree of attention paid to terrorism as collective and organized action is not the result of some oversight on the part of academicians. Indeed, the lone breed of terror, according to one account, makes up a mere 1.28% of all (global) incidents of terror between 1968 and 2007 (Spaaij 2010). Given that this breed of terror is (with global scaling) a marginal phenomenon at worst, the question arises – why should one pay special attention to lone wolf perpetrators of terrorism? One reason is that lone-wolf terror is a marginal phenomenon only when taken as a proportion of global terrorist incidents. Lone-wolf terror incidence displays striking variations (with respect to frequency) across countries. In the United States, for example, lone wolf incidents account for about 42% of all cases of terrorist violence, the highest incidence of any nation (ibid.). Accordingly, the problem of lone-wolf terror presents a problem of varying magnitude where this variation occurs on a country-by-country basis. Another reason why it is important to study lone wolf terrorism is that its incidence, while not yet on the rise in terms of lethality (Hamm and Spaaj 2015), is poised to rise per the consensus of a number of terrorism scholars and anti-terrorism practitioners (Simon 2013). It is important to note that in terms of frequency, lone wolf terrorism has been on the rise, peaking in the US in 90s (Bakker and de Graaf 2010). (Hamm and Spaaj 2015), for example, note that there is reason to believe that individual radicalization pathways are becoming increasingly more accessible. In chapter 1, it was noted that concern over lone wolves pervaded intelligence circles as evidenced by the remarks in 2003 by then FBI director Robert S. Mueller, III who expressed his view that the greatest terror threat to the US homeland was from lone wolves. This sentiment of a growing apex threat is also echoed by US President (POTUS) Barack Obama who remarked that “the risk that we’re especially concerned over right now is the lone wolf terrorist, somebody with a single weapon being able to carry out wide-scale massacres…” (McCauley and Moskalenko 2014, p. 70).

The operational definition of lone wolf terrorism used in this dissertation is taken from Spaaij’s conceptualization of the phenomenon – lone wolf terrorism represents acts of terror perpetrated by a person who (1) operates individually (2) is not a member of any terrorist group or organization and (3) conceived of his/her act of terror without an outside command hierarchy (Spaaij 2012). This definition is favored here over other popular conceptualizations because it is somewhat more restrictive in its formulation. Another popular definition of lone wolf terrorism is provided in (Stewart and Burton 2008) who write that a lone wolf terrorist is “a person who acts on his or her own without orders from – or even connections to – an organization.” This definition is essentially identical in its content to the one proposed in Spaaij’s work. It is important to note that the definition of lone wolf used here must be taken in tandem with the definition of terrorism presented earlier in this section, Spaaij’s three definitional elements will be considered in depth in the paragraphs below.

Lone wolves plan and execute attacks alone. This aspect of the definition, in addition to highlighting the solitary nature of lone wolves, misleadingly implies that these individuals are neither ideologically nor materially enabled by others. Lone wolves operate alone in as far as not receiving orders from an organizational hierarchy. Lone wolves are enabled by others either wittingly (by ideological inspiration or by leading through example) or unwittingly (Hamm and Spaaj 2015). This seeming paradox is easily dispelled by articulating a clear stance with respect to the implied meaning of "lone." The lone wolf metaphor seeks to bring to the fore the category of terrorist whose actions come about without orders from an organized terrorist group. In fact, it is a core principle used and implied throughout this dissertation that lone wolves *do* demonstrate affinity with groups of sympathizers. The case of Anders Behring Breivik illustrates this distinction. On July 22nd 2011 Anders Behring Breivik detonated a car bomb in the center of the Norwegian capital of Oslo killing 8 people. He subsequently drove to the island of Utøya where a Norwegian Labour Party Youth Camp was being held and opened fire. On that day, Breivik killed 77 people (Bangstad and Books 2014; Berntzen and Sandberg 2014). It has since been reported, based on verified claims from his manifesto and findings from court proceedings, that Breivik had at different times been a member of a number of right wing organizations and was active in Norwegian mainstream politics. Breivik espoused his affinity for certain Right-wing anti-Islamic bloggers as well as the opinions expressed on anti-Muslim and anti-immigrant websites in his manifesto (Pantucci 2011). By all indications, however, Breivik acted alone, his peripheral ties to far-right groups representing European anti-Muslim ideology notwithstanding; he did not share his mission with any of these groups, feeling that they could not be trusted with such knowledge (Pantucci 2011).

The concept of leaderless resistance has been invoked in describing the second component of Spaaij's definition: absence of formal affiliation to an organized extremist group that is characteristic of lone wolf actors. Leaderless resistance also provides insight into the third component of the definition: the absence of a command and control hierarchy in the contrivance of attacks. Both components will be discussed here. Leaderless resistance as a concept, or rather, tactic, emerged from the radical right, white supremacist movement in the United States in the later half of the 20th century (Kaplan 1997). Kaplan writes that as a result of the recognition amongst certain white supremacist groups that drumming up sufficient public support for the cause was unlikely, the strategy of leaderless resistance was born. The unwinding of the 'enemy state' would not be forthcoming with mass social action. However, the state could be damaged by the actions of a lone and resolute revolutionary. Further, "in the milieu of the radical right, no one is to be trusted, anyone could be (and probably is) an informer either for the government or for one of the many watchdog organizations monitoring radical right wing activity, and short of divine intervention, public support would not be forthcoming…yet in this state of weakness…[and] with nothing left to lose, a man is totally free to act as he will. For while the state had proven over and over again that it could effortlessly penetrate any right-wing organization, it had yet to develop the ability to thwart the will of one man acting alone" (Kaplan 1997, p. 82). Leaderless resistance is a more inclusive concept than what is intended here by "lone wolf terrorism" as it also permits into its conceptual umbrella "very small, highly cohesive" groups which carry out anti-establishment violence independent of any movement or support network (Kaplan 1997). In fact, the concept of leaderless resistance has been replaced in more contemporary academic literature by the concept of lone wolves (Berntzen and Sandberg 2014) even though the distinctions between the two concepts are not entirely clear (the commonalities abound). Examining the genesis of leaderless resistance provides insight to modern day lone wolf terrorism by highlighting the general conditions in which such a strategy "makes sense," so to speak. The tactic is associated with "weaker actors who are engaged in asymmetrical struggle, and thus is viewed [as] a sign of desperation and failure" (Joosse 2015, p. 2), indicative of a tactic of last resort, a tactic that comes at the heels of a perceived failure of organized, hierarchical movements (ibid.). Leaderless resistance is often conceptualized as the result of counterterrorism efforts to decimate the leadership structure of organized terrorist groups, thus, creating a "leaderless nexus" (Dishman 2005). While command and control structures have degraded, ideological adherents have become more loosely governed, resulting in small and autonomous tightly knit cells and lone individual actors becoming characteristic of larger ideological movements (ibid.). The breakdown of hierarchical terrorist organizations directly contributes to the proliferation of lone wolf attacks as seen in recent times. Leaders of these organizations change tactics as law enforcement decimates their command and control structures. Once an intimate part of the meticulous planning involved in spectacular terrorist attacks, these individuals have become ideological figureheads, neither providing any direct support nor issuing orders, but disseminating an ideology of violence to overcome perceived oppression.

A recurring theme in the literature on lone wolf terrorism is the centrality of the Internet to the threat (Bates 2012; Berntzen and Sandberg 2014; Hamm and Spaaj 2015; Michael 2012; Simon 2013). Above, the consequences of breaking down the command structure of hierarchical terrorist organizations were discussed. A shift in the role of the leaders of terrorist organizations from planning and command to ideological figureheads follows from this breakdown. The Internet is the most often cited enabler of this strategy, making leaderless resistance in the

Internet age more lethal than in its genesis amongst radical right wing organizations. Beginning in 1940 and ending in 2013, (Hamm and Spaaj 2015) identify 98 cases of lone wolf terrorism in the united states. For each of these cases, the authors coded a locus of radicalization. Loci of radicalization are the means through which an individual is radicalized prior to perpetrating a terrorist act. In their analysis Hamm and Spaaij show that prior to the September 11th 2001 (9/11) attacks in the United States, the primary locus of radicalization for lone wolf terrorists was the extremist group, accounting for the radicalization of 26% of these individuals (see figures 3 and 4). Post 9/11, however, the Internet is tied for the for the most frequently occurring locus of lone wolf radicalization, being the locus for 20% lone wolf attackers compared to 3% pre 9/11. It becomes an imperative to examine how and why this shift is occurring.

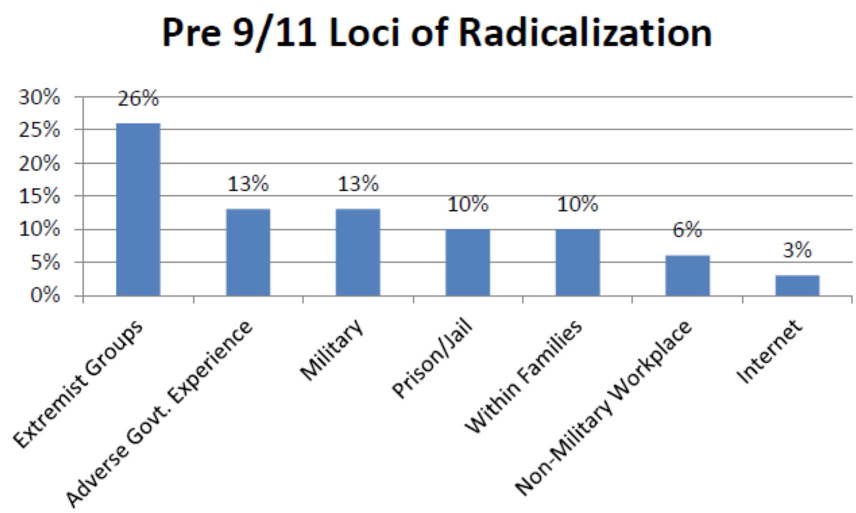

Figure 3. Loci of Lone Wolf Radicalization Pre-9/11 (Hamm and Spaaj 2015)

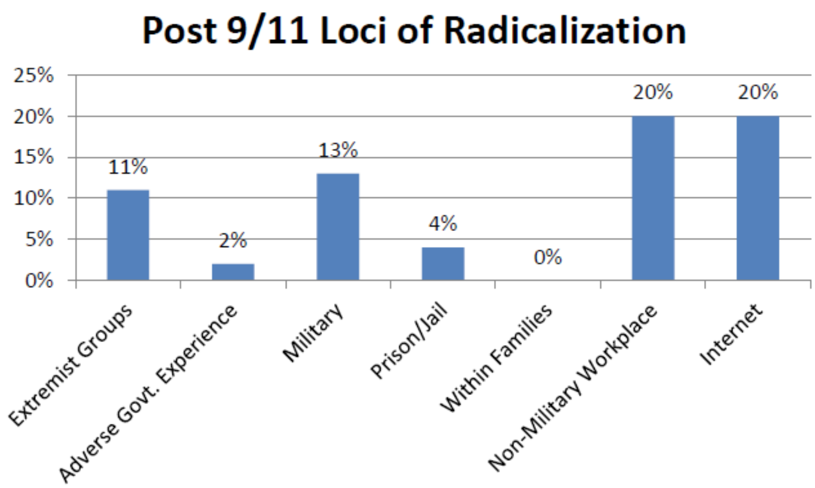

Figure 4. Loci of Lone Wolf Radicalization Post-9/11 (Hamm and Spaaj 2015)

"Within the last 10 years, the Internet has become the principal platform for the dissemination and mediation of the culture and ideology of jihadism" writes (Awan 2007, p. 389). Tucker (2001) alerts to the threat of a new terrorism (there is actually nothing new about the tactics he writes about) that is facilitated by ICTs (however, the facilitative role of the Internet is new). He writes: "one reason amateurs and ad hoc groups can operate as they do is an often unremarked-upon aspect of the communication revolution (see chapter 1 on Web 2.0). In addition to facilitating networks, the communication revolution also facilitates… the mobilizing of resources – political and individual support, and knowledge as well as money – that all terrorist organizations must do. The declining cost and increasing ease of communicating over great distances means that terrorist groups have greatly increased the potential pool of resources [- this especially includes people - that] they can draw on" (p.2). Writing for RAND Europe, (von Behr et al. 2013) report that the Internet is relevant to virtually every national security investigation conducted by security operatives in the United Kingdom (UK). Terrorism scholars are virtually unanimous in their view that lone wolf radicalization is significantly more likely to occur on the IInternet than through face to face interactions with other like-minded individuals. The role of the Internet as the primary locus of radicalization of would-be lone attackers has been variously characterized as a device for the dissemination of radical ideologies (Pantucci et al. 2011) and as a device for the dissemination of terrorist know-how (Weimann 2012).

This thesis argues that using an explicit, theoretical understanding of lone wolf terrorism and radicalization leads to better design choices and outcomes where the development of an IT framework to detect online radicalization is concerned. The processes of radicalization are nuanced and complex, but there exists a large body of research into these processes. It behooves IT researchers in the security informatics field to incorporate the findings from this body of work into their designs. A radicalization model of lone wolf terrorism has been proposed in (Hamm and Spaaj 2015).

Figure 5. Radicalization Model of Lone Wolf Terrorism (adapted from Hamm and Spaaij 2015)

The model is a six-part cyclical framework that describes a generic radicalization pathway of lone wolf terrorists. Its development is based on a detailed analysis of a number of terrorism databases where the authors of the model sought to extract common strands in post-9/11 lone wolf activity in the USA. *Personal and political* grievance refers to the motives behind an individual's radicalization. Both personal *and* political grievances have to be present in this model as it is shown from empirical data that in 80% of cases of lone-wolf radicalizations the radicalized individual harbored both personal and political grievances. *Affinity with online sympathizers or extremist group* lone wolves display a strong tendency to identify with extremist groups or more informal collections or pockets of online sympathizers. In recent times, that is, post 9/11 there is strong indication that "lone wolves may be seeking direction through venues other than organizations: namely, via networks of like-minded activists found online or on cable television." (p.9). This finding highlights the importance of being able to detect both radical social movements and radicalized individuals. Just as importantly it generates the implication that being able to generate accurate links between individuals and radical social movements is key. *Enablers* are people who help lone wolves in either direct or indirect ways. Indirectly by providing inspiration for terrorism, and directly by unwittingly assisting in planning activities. Data supports that a majority of lone wolves were enabled by individuals who either unwittingly provided direct assistance or who served as examples to be followed. *Broadcasting Intent* lone wolves almost always broadcast their ideology and threats via numerous channels, increasingly, via online channels. The recorded broadcasting intent of lone wolf actors is a key motivation for the design presented in this thesis. *Triggering event*s serve as catalysts for acts of terror by lone wolves. Personal, political events are in play, here. Events are sometimes immediate happenings that lead to violent acts of terror or they may "accumulate over time (through a series of **escalation thresholds**) until the lone snaps under pressure…" (p.11).

A model such as **this** can be employed by researchers who seek to build IT frameworks to detect radicalization on the Internet. Insight from this model can lead to a clearer understanding of the problem domain, and add assurance that the design resultant IT framework is based on empirical findings. Specific examples of the consequences of the above radicalization model of lone wolves terrorists for artifact design are immediately apparent. For example, the escalation thresholds that are characteristic of triggering events point to the need to monitor events that may be relevant to lone actors and map these events to some hierarchy of severity; the ubiquitous existence of personal and political grievances can be incorporated into a predictive model of radicalized individuals; the existence of enabling individuals or organizations may prompt an IT researcher to develop tools that track where amateur lone wolves may seek out aid in an effort to identify the such actors. This thesis seeks to fill, in a broad sense, the "theory-void" in the academic literature on security informatics.

2.3. Framing Theory

In this section, a framing theory of collective action is introduced. The framework serves to provide a well-researched, common dialect for analyzing critical aspects of the dissemination of terrorist ideologies[6] and providing theoretical support for the various positions staked in this work. Framing theories are associated with a critical sociology of social movements referred to as social movement theory (SMT).

[6] The word "ideology" is used here in lieu of frame, as the concept of framing has not yet been introduced. "Frame" and "ideology" should not be taken to mean the same thing.

The discussion on collective action framing cannot commence without some attention to the social movement context with which it has an existential dependence. The literature on social movements is vast and a review of this literature is beyond scope. However, the meaning of the concept must be given, at least as it is to be understood in this dissertation. What constitutes a social movement is still very much in debate (Karagiannis 2009). However contentious the literature may be, a useful starting point for conceptualizing social movements is given by Klandermans who writes that social movements are "collective challenges by people with common purposes and solidarity in a sustained interaction with elites and authorities" (Klandermans 1997, p. 2). Social movements are first and foremost a type of collective or group action. That which accounts for such action is the domain of social movement scholarship. Three recurring themes appear in the literature on analyzing the emergence and development of social movements, and taken together, begin to account for group action:[7] (1) Political opportunities (2) organizational form (mobilizing structures), and (3) framing processes. Inquiry into the first of these three themes is undergirded by the conviction amongst social movement scholars that "social movements and revolutions are shaped by the broader set of political constraints and opportunities unique to the national context in which they are embedded" (McAdam et al. 1996, p. 3). Mobilizing structures characterize the organizational configurations that groups may adopt towards engaging in collective action (ibid.). McAdam and colleagues write of two theoretical frames that are widely employed in inquiry into social movement mobilizing structures. These are resource mobilization theory, which seeks to study the aggregation of resources into a coherent organization for the purposes of collective action (McCarthy and Zald 1977)) and the political process model which instead embraces the concept of revolution and grassroots

---

[7] A distinction is being made here between group and individual action. (Van Stekelenburg and Klandermans 2013)offers an excellent introduction to the social psychology of the individual in the context of group action.

organization, rejecting the mechanistic propositions of resource mobilization (McAdam et al. 1996; Tilly 1978). The third and final category, framing processes, is the subject of the remainder of this section.

Framing refers to the signifying work in which so-called signifying agents are engaged (Benford and Snow 2000). Signification is the creation and maintenance of meaning for movement stakeholders (constituents, antagonists, bystanders). Signifying work leads to the production of frames. Signifying agents are movement actors (i.e individuals with some non-trivial degree of influence within a movement) who are engaged in meaning making for movement stakeholders. Framing "denotes an active, processual phenomenon that implies agency and contention at the level of reality construction" (p. 614). Framing is thus an activity that involves movement activists asserting their agency to generate new frames which differ from and challenge existing frames. Framing activities result in the production of *collective action frames*.

For Goffman, the concept of frame entails labelling schemata; schemata of interpretation that enabled people perceive and label occurrences in their life-spaces and (Goffman 1974). Moving forward, "frame" and "framing" refer explicitly to Benford and Snow’s collective action frame theory and not Goffman's frames. These schemata govern the meanings that people attach to events and are therefore quite important in guiding action. Collective action frames are a specialized and refined version of this broader framing conception with the purpose of "mobilizing potential adherents and constituents, to garner bystander support, and to demobilize antagonists" (Benford and Snow 2000 citing Snow and Benford 1988). Collective action frames, therefore, are beliefs and schemata of attribution with an orientation towards action. Such frames lend legitimacy to the actions of social movement organizations (Benford and Snow 2000). This study takes heed of B&S's warning to not take collective action frames to mean static schema. While collective action frames *are* schemas of interpretation, they are constructionist in nature, existing not in the heads of movement actors but dialogically in the intersubjective space between these actors (Snow and Benford 2000). This conception centers the contested nature of the frame production process (collective action framing). Quoting Gamson (1992, p. 111), Benford and Snow (2000) write that "collective action frames are not merely aggregations of individual attitudes and perceptions but also the outcome of negotiating shared meaning."

2.3.1 Framing Features

This subsection concerns the features of framing theory, focusing on *core framing tasks, discursive processes, and framing contests*. While distinct, these three core features complement each other and lead to a comprehensive and useful framework for understanding frames *and* their active construction.

Core framing tasks are a tripartite framework of diagnostic framing, prognosis framing and action mobilization. *Diagnostic framing* (injustice frames) is commonplace across social movements, supported by empirical work. Not all collective action frames contain an injustice component (p5). In fact, there appears to be little need for injustice frames in "religious, self-help, and identity movements" (p5). Injustice frames are ubiquitous when the movement pertains to political/economic change. There is an attributional component to diagnosis framing. It seeks to attribute blame or responsibility, assigning some causal linkage from the problematic situation to culpable agents. In addition there exists a concept of boundary or adversarial framing where movement adherents seek to delineate boundaries between good and evil, protagonist and antagonist. As the introductory remarks on framing theory suggests, diagnostic framing is often a contentious activity. *Prognostic framing* is a proposed solution to the diagnosed problem. Prognostic framing highlights, therefore, that which the movement believes should be done or accomplished (Snow and Benford 1988). There are suggestions of a linkage between diagnostic and prognostic such that there appears to be a tendency for diagnostic activities to constrain possible prognosis. Within a given movement (however the analyst chooses to delineate movement) multiple diagnostic framings may exist. There is no requirement that these framings be complementary, coherent or non-contradictory. Each diagnostic frame, however, does require an appealing, complementary prognostic frame, indicating that there is analytic merit to associating prognostic frames with corresponding diagnoses. Prognostic framing also involves counter framing (Benford 1987: 75 for more information) where the prognostic framing task is constrained or influenced by competing ideologies/movements. This finding is intuitive as a social movement is not typically alone in seeking to change a particular state of affairs. Other SMOs are typically involved and their framings influence and are influenced by the framings of other relevant SMOs. For example, (Wiktorowicz 2004) makes an account of such "framing contests" between SMOs engaged in *Jihad*. (Benford and Snow 2000) (and this particularly important for the purposes of this dissertation) identify prognostic framing as the core framing task that serves as the primary differentiator between SMO's. *Motivational framing* deals with agency in collective action frames. It provides the basic rationale for engaging in corrective action and includes rhetoric for constructing and rationalizing motive. The issue of agency has been addressed in (Benford 1993)by "identifying four generic vocabularies of motive that emerged in the course of interaction among movement activists, rand and file supporters, recruits and significant others: vocabularies of **severity urgency efficacy** and **propriety**" p.7. Important questions remain about the conditions that lead to the formulation and adoption of certain vocabularies of motive. What is clear, however, is that these vocabularies provide movement participants with elaborate rationale for collective action and movement adherence.

Discursive processes refer to spoken and written communications of the members of SMOs in the context of everyday movement activity. Discursive processes, as part of a coherent framing theory, are the means through which frames are *articulated*, and *amplified*. Frame articulation refers to the process of stringing together events and experiences in a coherent manner. Each of diagnostic, prognostic and motivational frames are strung together through frame articulation. Novel frames are not necessarily constructed from novel ideas. It is the manner in which these are "spliced together and articulated, such that a new angle of vision, vantage point, and/or interpretation is provided" (p.623) that lend novelty to emerging diagnostic, prognostic and motivational action frames. Frame amplification (or punctuation) processes concern the centering of movement discourse on a number of themes, events, or ideas that are particularly representative of the movement. Doing so entails "the idealization, embellishment, clarification, or invigoration of existing values or beliefs" (p.624) in a deliberative strategic and goal oriented way.

Another dimension of framing theory is that of contest framing processes. Given any collective action space, there will exist contention amongst movement actors regarding the ideal frames that should be generated. This is often referred to as the politics of signification and comprises a non-trivial portion of the signification work in which movement actors are engaged. Benford and Snow (2000) cite the three forms in which framing contests manifest: counterframing by movement opponents, bystanders, and the media; intra movement framing disputes; the "dialectic between frames and events" (p.625). Indeed framing contests are implicit to the very existence of social movements, as such movements are generally representative of the existence of differences in signification around a particular issue within a particular society. Counterframing generally refers to attempts by bystanders, the media and movement opponents to undermine the extant frames of a particular movement. Of course, movements do not sit idly by in the presence of active counterframing, leading to the re-articulation of existing frames by movement actors (Benford and Snow 2000). Intramovement framing contests are also common and have been examined in a number of movement contexts (Wiktorowicz 2004). Intramovemet framing disputes are constitutive of the emergent movement frames as these disputes involve debates over diagnostic and prognostic frames as well as over motivational framing, that is "how reality should be presented so as to maximize mobilization" (p. 626). Finally, framing contests can manifest in the dialectic between existing collective action frames and collective action events. The outcomes of movement actions serve to constrain future discursive and frame generative processes. While initial framings make certain actions possible, the resultant actions place constrains on future framings, thus a dialectic B&S 2000.

The last major consideration in the framing perspective is that of frame resonance (Benford and Snow 2000). The potency of manufactured frames attends to the matter of variation in effectiveness of different frames. Effectiveness, here, refers to ability of a frame to mobilize individuals either to action or to belief. Three crucial factors lead to potent framings by movements (strongly resonating frames): (1) empirical credibility, (2) experiential commensurability, and (3) narrative fidelity (Snow and Benford 1988). If a movement's framings are verifiable against real events that can be considered evidence of their validity, those framings are empirically credible and are more likely to resonate with the target audience. Of course, "empirical credibility" to a given individual is a subjective matter. Evidence is filtered through an interpretive screen characterized by the personal experiences of individuals. As such, experiential commensurability is a function of the question "does [the framing] suggest answers and solutions to troublesome situations which harmonize with the ways in which these conditions have been or are currently experienced" (Snow and Benford 1988)? Finally, the extent to which a frame resonates with an individual is governed by the degree to which framing narratives are faithful to or evocative of cultural narrations, "the stories, myths and folk tales that are part and parcel of one's cultural heritage and that thus function to inform events and experiences in the immediate present" (Snow and Benford 1988). It follows that narrative fidelity is an important predictor of frame resonance. For example, (Zuo and Benford 1995) study how framings used by the student Democracy Movement in China "were interpreted and articulated such that activists' claims resonated with the experiences and cultural narratives of Chinese citizens" (p. 133) and contrast the potent frames produced by the student led Chinese Democracy Movement with the less potent frames produced by the Chinese Communist Party. The student-led Chinese Democracy Movement was a response to increasing inequality in China following a series of economic reforms that simultaneously privatized much of the Chinese economy while maintaining strong State control over the resultant market economy. This economic configuration led to massive corruption at the State level, fewer opportunities for the common people, and economic inequality. Such a state of affairs sewed discontent amongst the masses.

The students behind the Chinese Democracy Movement used cogent framings in the form of critiques of the injustices of the new Chinese economy. The students' framings resonated with the Chinese masses because the frames were not only "grounded … in three Chinese cultural traditions or narrations: Confucianism, communism and nationalism" but were "consistent with citizens' observations and experiences" (Zuo and Benford 1995). Further enhancing this resonance were the ineffectual counter-framing employed by the Chinese Communist Party.

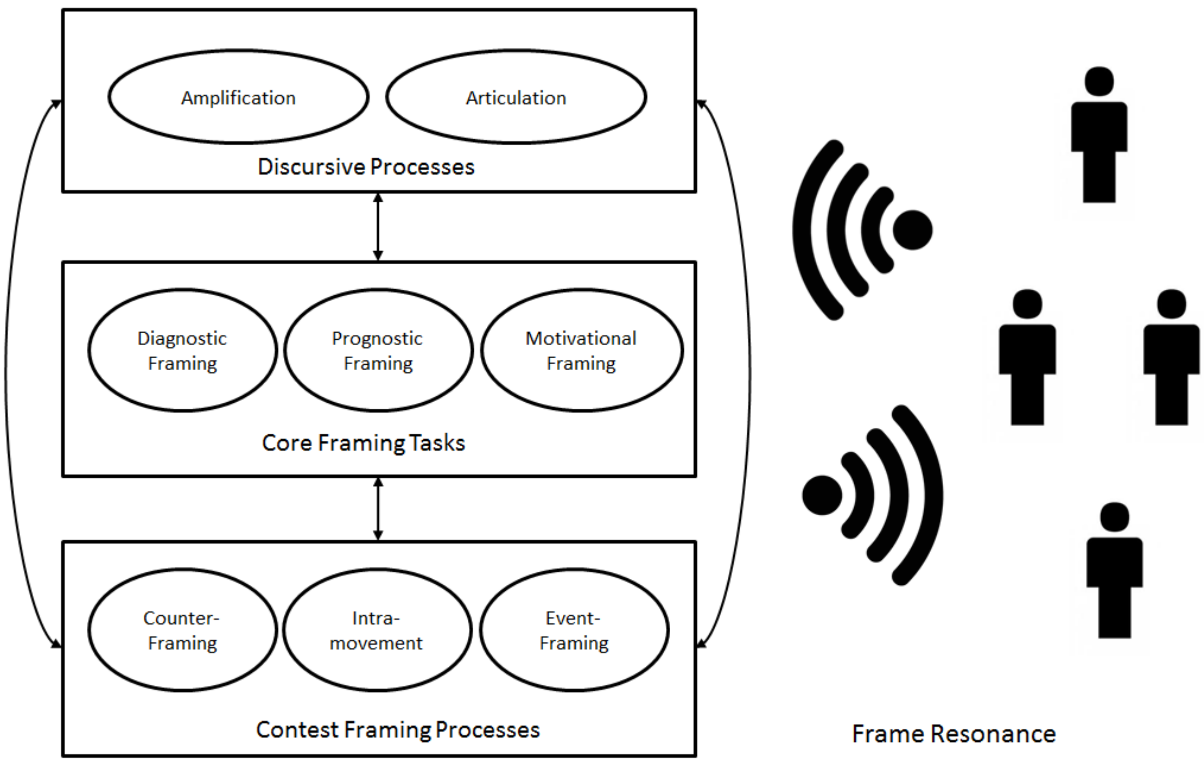

Figure 6. Collective Action Framing Theory

Collective action framing has been deployed extensively in social movement studies as well as media studies as a lens to theorize about various aspects of social movements especially where action is concerned. The utility of this lens in the study of group based terrorism (the general form of terrorism from which lone wolf terror is distilled) was notably argued in Jackson et al eds. (2009) edited volume calling for a critical (in the sense of critical social theory) perspective to terrorism research. Gunning (2009), writing in this volume, explains that much of terrorism studies has been focused on interest-based, rational choice explanations of terrorism on one hand and ideological based explanations that take ideology as static on the other. Gunning puts forward framing theory as a means to simultaneously problematize ideology (that is regard it as dynamic and contested as opposed to static and given) and take into account resource- (interest-) based perspectives. (Berntzen and Sandberg 2014)concur, pointing to additional benefits of framing theory in terrorism studies: "framing theory in social movement studies offers a detailed and specific tool for comparing the rhetoric of different political actors. The theoretical framework emphasizes actors' way of thinking, rationale, and motivation, and has previously been used to study Al Qaeda and racist nationalism in North America and Europe" (p.761). Mirroring Gunning, they go on to note that framing theory is inclusive and amenable to complementary theoretical perspectives and explanations frequently used in the study of terrorism. Such perspectives include "individual psychology, social factors, political opportunities, or the existential attractions, excitement, and feelings of terrorists" (p. 761). Indeed, framing theory enables these aspects to be integrated into a comprehensive understanding of terrorism (ibid.). A number of terrorism researchers have employed framing theory beyond the use of a framing metaphor in understanding various dimensions of terrorism. Collective action framing, however, has been scarcely employed in the literature on lone wolf terrorism due principally to the surface-level construction of lone wolf terrorism as a solitary activity. Berntzen and Sandberg (2014) are amongst the first (and potentially only) to argue that lone wolf terror has a collective nature. Their argument is central to the thesis proposed herein. **Regardless of the sparse use of collective action framing in the literature on lone wolf**

terrorism, there are some instances of collective action framing theory being used as a theoretical lens to study a group (or movement) based terrorism. These instances are reviewed in Table 1 below.

The reviewed papers are selected using a rather strict set of criteria. First, to be included in this review, a study must empirically investigate terrorism (group based or lone). Second, in its exposition of terror, the study must employ the theoretical framework illustrated in Figure 6, derived from the scholarship of (Benford and Snow 2000). Where the study uses a conception of framing not directly sourced from Benford and Snow's work, it should be possible to map the analysis, at a minimum, to the core framing tasks shown in Figure 6. Further, the study must be focused on the manner in which social movements frame *raison d'etre* as opposed to how governments and the media frame the movement. Finally, this review focuses only on articles published in the top terrorism journals identified in (Ranstorp 2006): (1) Terrorism and Political Violence (2) Studies in Conflict and Terrorism. Articles are retrieved from these journals by searching for "collective action" and/or "framing." Citations are also analyzed in an attempt to increase the breadth of the review.

Paper Citation	**Discussion**
(Berntzen and Sandberg 2014)The collective nature of lone wolf terrorism	The main thrust of this work is a contention that there is merit in seeing lone wolf terrorist as enacting a rhetoric that is couched in larger social movements. All framing tasks in this paper are considered from two angles. The first is the collective action framing of the Norwegian anti-Islamic movement and the second is the framing of Anders Breivik an enactor of the movement's rhetoric (i.e. lone wolf). (1) Diagnostic Framing: the Norwegian

	anti-Islamic movement diagnoses Islamism as a posing an existential threat to western values and the Norwegian welfare state. In addition the Norwegian state and ideological elites are oppressive forces against the movement. State and Muslims are the enemy. (2) Prognostic Framing: the various anti-Islamic social movement organizations propose a halt to non-western immigration, assimilation of already present ethnic groups over multiculturalism. These groups explicitly reject violence. (3) Motivational Framing: positions passivity as one's acceptance of the process through which islamization occurs. Personal freedoms of choice are at stake.
(Wiktorowicz 2004) Framing Jihad: Intramovement Framing Contests and al-Qaeda's Struggle for sacred authority	In a study of the jihadi movement known as al-Qaeda, Wiktorowicz employs framing theory as a means to examine intramovement conflict. This study focuses on how the credibility of movement ideological figure heads comes into play in such framing conflicts. More specifically, this research examines how *crediting* and *discrediting* strategies are used towards these ends. These strategies are refined into 4 basic strategic themes relevant to the credibility of a movement's popular intellectuals" (1) credentialing (2) decredentialing (3) vilification and (4) exaltation (p.159).
(Page et al. 2011) Al Qaeda in the Arabian Peninsula: Framing Narratives and Prescriptions	Page uses selected narratives of principal figureheads of Al Qaeda in the Arabian Peninsula (AQAP) to delineate the diagnostic, prognostic and motivational framings used in their media operations. (1) Diagnostic framing: the Yemeni administration led by Salih has failed and this failure is evidenced by the suffering of the people of Yemen due to corruption, socioeconomic collapse, inflation and joblessness; In addition, Western aggression against Muslims, regardless of location, is cited as a problem. The West is cited as the responsible party for all socioeconomic problems in Yemen. (2) Prognostic Framing: Violent jihad must be waged against the West and any apostate regimes, this is the only way to achieve the sought after political change. Jihad is a means to remove Western elements from the peninsula, overthrow Salih and establish an Islamic caliphate. (3) Motivational Frames: Humiliation is heavily employed. Those Muslim men who remain apathetic in the face of the global and local humiliation of the *umma* are being outdone by women and children

	engaging in violent jihad serves. This motivational frame serves to spur men into action in an effort to protect their "manhood."
(Karagiannis 2009) Hizballah as a Social Movement Organization - A Framing Approach	The framing approach is used to characterize the means through which the terrorist organization Hizballah (Party of God) creates and recreates its collective identity towards garnering support. (1) Diagnostic Framing: Hizballah draws heavily on injustice frames. These frames have taken the form of attributing poor conditions in Lebanon to Israeli occupation of Lebanese land as well as the dominion of an unjust, secular government. (2) Prognostic Framing: initially this took the form of an Iranian style Islamic republic established through revolution and the dissolution of borders in the middle east, portraying political change as a religious imperative. However in more recent times the prognosis has abandoned the pan-islamic rhetoric, focusing more on the local, Lebanese, condition. (3) Motivational Framing: The group has employed strategies such as casting it's slain leaders as descended from Shia heroes, overlaying the lives of these leaders with the mystique of Shia heroes. Violent Jihad against Israel is framed as *wajib* (religious duty), that Muslims have a sacred duty to liberate Palestinian lands from infidels.
(Rogan 2010) Jihad Against Infidels and Democracy - A Frame Analysis of Jihadist Ideology and Jurisprudence for Martyrdom and Violent Jihad	This work uses a framing notion of obscure origin, that is, it neither derives from Goffman nor from Benford & Snow. However, the "dominant frame structures" presented in the paper are congruent with framing theory conception in use in this dissertation. Rogan Identifies 4 principal frames (note that he is referring to frames and not framing processes) (1) Whole Story (2) Characterization of Other (3) Self-Presentation (4) Moral Judgment. (1) roughly maps to frame punctuation or amplification, (2) roughly maps to diagnostic frames as the characterization of the other is essentially singling out those who are responsible for the problematic situation, (3) roughly maps to diagnostic framing as well given that "Self-Presentation denotes how an individual defines the self and the image of the self as it is being presented to others relative to the other party" (p. 399) In other words, taken together, (2) & (3) present a comprehensive framework for diagnosis frames. Finally, (4) roughly correlates to motivational and prognostic frames but

	presents a narrower category focusing only on moral remedies (prognosis) and moral impetuses (motivational). The frames used in two documents authored by Ayman al-Zawahiri (ostensive leader of al-Qaeda) were analyzed and so-called "micro-frames" were extracted for each of the four categories see pages 402 & 404.
(Wright 2009)Strategic Framing of Racial Nationalism in North America and Europe An Analysis of a Burgeoning Transnational Network[8]	Here, the framing strategies of radical far-right, racial nationalist groups (white supremacists) in Europe and the US are analyzed using a comparative approach. (1) Diagnostic Framing: both U.S. and European groups blame Islamic terrorism on their respective governments' support of Israel and their entanglements in international conflict, thus blaming insecurity at home Jews. Further, liberal immigration policies and multiculturalism have been linked to insecurity in the homelands of both US and European contexts. (2) Prognostic Framing: some groups promote the "Third Position" perspective which calls for a homogeneous (mono-cultural) nation characterized by a supremacist religion and race. Additional prognoses include an end to multiculturalism, pro-Israeli policies and immigration. (3) Motivational Framing: attributions of threat from immigration pervade the motivational framing of these groups. This is attained by evoking images of Western countries as becoming "Third World" slums in the face of "Third World" immigration. Immigration is also cast as invasion.

Table 1. Collective Action Framing and the Study of Terrorism

A few striking observations emerge from this review. The first is that collective action framing theory (at least as conceptualized here) has principally been used to examine Islamic[9] extremism in the literature. Only one case of the perspective's application in non-Islamic movement contexts could be included (with caveat, at that) in this review. Ranstorp proffers a tenable

[8] This paper is not quite focused on terrorism. However it is included here because it concerns the framing strategies of white supremacist movements the ideologies of which are commonly associated with terrorist attacks both in the US and in Europe.
[9] I make no claim that Islamic extremism is a monolith, quite the contrary.

explanation for this state of affairs – 9/11 (Ranstorp 2006). Prior to 9/11 Islamic terrorism accounted for a small minority of terrorism articles published in the two major terrorism journals (less than 10%). However, Ranstorp notes that somewhere around half of all terrorism publications after 9/11 were related to Islamic terrorism. The collective action framing perspective described in this chapter has only been applied post 9/11. Terrorism research has attained notoriety for being descriptive in nature, lacking a coherent theoretical framework for understanding terrorism as a phenomenon. (Ranstorp 2009) writes about his review of terrorism research that the "principal cause for this critique is the surprisingly few research inventories conducted over the years designed to fundamentally question theories, assumptions, and knowledge production" (p.13). However, there is evidence that this theoretical vacuum may begin to be filled by social movement theory (Gunning 2009). This dissertation focuses on framing theory within the context of a broader social movement theory. The framing perspective is well suited to the purposes here for many reasons. First, framing provides a common vocabulary for characterizing the syncretic process of ideology formulation and delivery across disparate movements. Second, the framing perspective provides constructs that are sufficient to discriminate between movements. In other words, framing empowers the analyst to view movement ideology and strategy as problematic and nuanced, varying between movements that ostensibly have the same goals. Third, the framing perspective proffers theoretical insight on the propagation of frames to potential movement adherents. The concept of frame resonance lends rigor to this dissertation's core contentions. In particular, the contention, mirrored in (Berntzen and Sandberg 2014), that lone wolf terrorists tend to demonstrate affinity with social movement frames, is supported by the notion of frame resonance

2.4. Detecting Lone Individual Radicalization on the Internet

This work aims to develop and evaluate an IT artifact the purpose of which is to automatically detect the resonance of terror movement frames in textual content, a first step towards identify radicalized, potential lone terror operatives on the Web. Post 9/11, academicians have given some attention to the topic of detecting various forms of radicalization and extremist activities on the Web. A review of these works is presented below (see Table 2).

Paper	**Non-IT Theory**	**IT Techniques**	**Data Source**	**Type of Extremism**	**Objective**
(Wadhwa and Bhatia 2013)	Not Present	A pipeline is presented that includes: • Standard text pre-processing • Word frequency for topic identification	Twitter	Islamic	To capture the dynamics of online extremist communities
(Berger and Strathearn 2013)	Not Present	Manual inspection of tweets to determine affinity towards an arbitrarily defined white nationalism ideology	Twitter	White Nationalism	Studies the nature of white extremist social interactions on Twitter.
(Kwok and Wang 2013)	Not Present	Bag-of-Words model with a Naïve Bayes classifier	Twitter	Anti-Black	Reports on an effort to learn a binary classifier for the identification of hate speech against blacks
(O'Callaghan et al. 2013)	Not Present	Social network analysis	Twitter (as starting point)	White Nationalism	The goal is to use Twitter as a starting point to

		Temporal network analysis Social network clustering			discover other extremist communities
(Reid et al. 2004)	Not Present	Social network analysis Natural language processing	Existing terrorism DBs; Terrorist Websites; Social Media	Varied	Multi-dimensional research project to develop terrorism knowledge portals
(Sun et al. 2008)	Not Present	Named entity recognition Relation extraction Directed graph (ontology) Multi-slot information extraction Entity disambiguation (core-reference resolution)	News websites; varied	Islamic	Given a corpus of documents, what is a suitable method for ranking them by information gain given an objective of extracting terrorism event related entities and relations? These entities and relations are user defined.
(Cohen et al. 2014)	Warning behaviors (Meloy and O'Toole 2011)	Sentiment analysis Author recognition	No data is analyzed. Paper is conceptual	Varied	Using linguistic markers to sense so-called "weak signals" left by lone wolves on Internet fora. No actual artifact is produced, simply suggestions about a technology stack
(Dahlin et al. 2012)	Not Present	Field matching Network analysis Spatio-temporal matching	No data is analyzed. Paper is conceptual	Varied	Fusing together aliases on discussion boards with a broader aim (not actualized) of providing analysts with coherent watch-lists

(Glass and Colbaugh n.d.)	Not Present	Tensor decomposition (PARAFAC decomposition) Web crawling Sentiment analysis	Websites; Twitter	Islamic	Detecting large scale Internet disruptions potentially associated with major events;
(Brynielsson et al. 2012, 2013)	Warning behaviors (Meloy and O'Toole 2011)	Web crawling/mining Author recognition with alias matching Naïve Bayes classifier	Websites; Twitter	Varied	Searching for digital traces on the Web to identify potential lone wolf terrorists
(Wiil et al. 2009)	Not Present	Data mining Social network analysis Graph theory Data visualization	Terrorism databases; Websites	Varied	Development of a toolbox of various models to harvest and analyze terrorist information; Of particular interest is the use of triples to store terrorism event data;
(Scanlon and Gerber 2014)	Not Present	Supervised learning algorithms (naïve Bayes, logistic regression, support vector machines, classification trees) Natural language processing for feature extraction	Web forum	Islamic	To learn classifiers for recruiting activities on violent extremist websites
(Agarwal and Sureka 2015)	Not Present	NLP for feature extraction purposes K-nearest neighbor and support vector machine classifiers	Twitter	Islamic	Learning k-nearest neighbor and support vector machine based classifiers for detecting hate and extremism promoting tweets
Abbasi & Chen (2008)	Not Present	Affect analysis (sentiment analysis)	Website fora	Islamic; White Nationalism	Using a probabilistic disambiguation approach, this

					study creates an affect lexicon that is used to gauge affect intensities on white supremacist and jihadist fora.

Table 2. Review of Existing Approaches to the Detection and Analysis of Internet Extremism and Radicalization

Table 2 summarizes some key takeaways from the security informatics literature on detecting radicalization and extremism on the Internet. Limitations in the literature can be inferred from these summaries and are the subject to the following discussion. The first major limitation is that the extant work in this domain is atheoretical. While Brynielsson and colleagues (Brynielsson et al. 2012, 2013; Cohen et al. 2014) use the notion of warning behaviors derived from the work of (Meloy and O'Toole 2011) as a theoretical backbone for their work on detecting so-called "weak signals" associated with individuals undergoing radicalization on the Web, their work is an exception and not the rule. Brynnielsson et al leverage the warning behaviors conception to justify the existence of weak signals, leakages, or traces inadvertently left by radicalized individuals on the Web. One such weak signal is termed "identification" and refers to "the desire to be like an influential role-model, warrior identification, or identification with a group or larger cause" (Brynielsson et al. 2013, p. 3). Identification is an important signal for detecting potential terrorists on the Web as such individuals are likely to espouse the views of major terror figureheads, concepts and themes. Author recognition techniques are suggested as a means to assess the extent to which an individual identifies with such figureheads, concepts and themes. However, this formulation is quite limited in a number of ways. First, automatic authorship identification aims to discover an author's writing style independent of the content of a text

(Houvardas and Stamatatos 2006). To use such algorithms effectively to the ends described in (Brynielsson et al. 2013) would entail the use of a massively comprehensive list of training documents corresponding to major extremist figures who may be the object of a lone-wolf's fixation. In addition, this approach is based on the assumption that potential terrorists writing on the Web will write using the same style as those whom they seek to emulate. This assumption is unfounded. Further, there is no evaluation of the efficacy of the author identification approach to this end. The solution approach proposed in this dissertation is based on assumptions consistent with the "identification" concept presented in (Brynielsson et al. 2013). However, it uses a more general conception of frame resonance to detect the affinity of an individual towards radical ideologies. The use of a knowledge base of radical framings to such an end enables an automated agent to detect radical writing that echoes the *ideas* and not the authorship style of particular individuals. The contention here is that this is a less restrictive approach. Further, the proposed knowledge base, in and of itself, offers advantages beyond the tracing individuals who identify to radical ideology. One such advantage is that it permits inference with respect to events that may be relevant to such individuals.

Another limitation with the existing literature is that there exist no attempts to associate social media postings with specific beliefs, ideologies and movements. There are a number of benefits to doing so, chief of which is the ability to associate short social media postings with detailed ideological conceptions. Consider a tweet which is 140 characters long. Such a short message might point to, say, a particular diagnostic frame used by a particular extremist group. It could be inferred that the writer of the post is likely to be sympathetic to this group. Having a detailed knowledge base of the frames, people, organizations, countries etc. that are related to the group's diagnostic framings extends what is known about an individual from a starting point of just 140 characters. Wiil et al. (2009) propose a knowledge base of “terrorist-related data and information” that is populated using Web data sources. The knowledge base is, however, restricted to the representation and storage of people, places, and events relevant to terrorist organizations. It has little to no use with respect to identifying unknown future cases of terrorism.

Another shortcoming of the existing literature is that it proffers few instantiated and evaluated design artifacts. The majority of approaches presented are either un-instantiated conceptual models of yet-to-be realized systems (thus they lack an evaluation) or simply report the results of data mining approaches to security informatics. A final observation of the literature is the lack of any approach to cataloging in a machine accessible way the many extant ideologies espoused by the numerous extant terrorist organizations. Such a catalog would provide numerous opportunities to discover critical relationships between terrorist espoused ideologies, potentially predict alliances, determine previously unknown affinities, detect changes in espoused ideologies over time, and so on. A machine accessible representation of terrorist espoused ideologies holds great promise both for the academic study of terrorism but also for counter terrorism efforts. Such a representation also allows for linkages to be asserted and represented (in the same machine accessible representation) between ideologies, events, people, groups, movements and, indeed, other ideologies.

This study seeks to fill these gaps by providing an instantiated and evaluated design artifact towards (1) the cataloging of radical ideologies (more precisely frames) with the aim of (2) automatically detecting radicalized individuals in online social spaces, as well as automatically detecting radical content in general. The proposed artifact will also enable (3) the monitoring of events relevant to the ideologies within the catalog with the goal of creating an early warning system (of sorts) for potential lone wolf terrorist attacks. This is the first attempt of this nature (i.e. including all three components) in the academic literature. A key advantage to this approach to detecting potential lone wolves is the creation and maintenance of a knowledge-base of terrorist framings that may be useful to both academics and security practitioners.

Chapter 3 – Research Method

3.1. Structure of Chapter

The purpose of this chapter is to elaborate a set of design objectives towards an IT artifact for detecting lone wolf radicalization on the Web. In addition, and subject to the aforementioned objectives, this chapter proposes and defends a number of evaluation strategies. The structure of the chapter is as follows. First, an introduction to the Information Systems' conception of design science, the methodological paradigm into which this work fits, is proffered. Next, the alignment between the objectives of this research, the selected theoretical perspective, and the resultant design choices are expounded. The study objectives motivate the selection of informing theory. Informing theory, by definition, affords a number of design choices. Given a set of design choices, evaluation strategies are proposed such that evidence of the efficacy of the design with respect to the objectives can be developed.

3.2. Research Objectives

The overarching objective of this research is the automatic detection of Web content that potentially indicates the radicalization of an online persona towards one or more radical terroristic ideologies. This objective is motivated by the well documented belief within intelligence circles that the threat of lone wolf attacks is the most serious posed by the phenomenon of terrorism due to the inherent difficulties in detecting such behaviors prior to perpetration. Stemming from this objective, this research aims to answer 6 specific research questions:

RQ1: What are the concepts and inter-concept relationships that are characteristic of radical ideologies and how can they be used in articulating the structure of a knowledge representation of terrorist framings?

RQ2: How may websites that frame and promote radical ideology be detected?

RQ3: How can the knowledge representation produced as a result of RQ1 be populated based on free text derived from the websites discovered in response to RQ2?

RQ4: How may the knowledge representation populated in RQ3 above be used to automatically detect the ideological affinities of an online persona?

RQ5: How can events relevant to individual, radical ideologies stored as part of the knowledge representation produced by RQ1 be automatically detected from news sources available on the internet?

RQ6: What coordinating framework is necessary to integrate all of the above in order to effectively identify online personas for additional scrutiny by security operatives?

The first research question, RQ1, seeks to develop a conceptualization of the domain of radical terrorist framings. Given such a conceptualization, an ontological specification is designed that approximates the conceptualization as closely as possible. The resultant ontology of radical framings thus approximates the empirical domain of radical, terrorist framings. Such an ontology enables the representation (in the vocabulary of the ontology) of particular instances of terrorist framings allowing for detailed analyses of terrorist organizations erstwhile considered monoliths. This is the first such effort towards conceptualizing and casting in a computer readable language the framings (espoused ideologies) of terrorist social movement organizations. There are many motivations for doing this, (discussed in detail in the next section) however, one stands out. In 2013, Lisa Stampnitzky's book *Disciplining Terror: How Experts Invented 'Terrorism'* proffered a detailed analysis of terrorism scholars with special emphasis on how due to the politicization of the terrorism concept "by the late 1970s, the role of concrete political grievances and motivations in understanding political violence had become highly contentious; so much so that the focus on 'understanding' terrorism could expose experts to charges of 'sympathy' with terrorists" (Stampnitzky 2013, p. 54). This trend, far from subsiding, has continued well into the 21st century and is markedly worse post-9/11. The proposed ontology, in keeping with Stampnitzky's call, is at least in part an attempt to scrutinize the grievances and strategies of terrorist organizations with genuine curiosity vis-à-vis the phenomenon of terrorism.

This question requires that our approach to identifying radical ideologies be suitable for use in an automated web crawler. This question imposes several important requirements on the design of our framework. The most important of these requirements is that the framework should be unsupervised. That is, we do not need to retrain classifiers to adapt the framework to new types of radical content.

The third research question points to the need for a knowledge base construction (KBC) system that takes as its inputs a set of unstructured or semi-structured text documents and outputs a structured data containing asserted facts extracted from the inputs. There are two key classification tasks that characterize this problem. The first is entity resolution (named entity recognition), where the instances of domain entities defined in an ontology are identified within textual data sources. The second is typically referred to as relation extraction wherein the presence of a particular relationship between two entity instances is examined. Given suitable entity resolution and relation classification procedures, the knowledge representation produced as a result of RQ1 can be populated using the collection of data retrieved from RQ2. The goal here will be to systematically examine the efficacy of various classification algorithms and feature extraction techniques given a training corpus of text.

The fourth research question entails, first, the techniques developed as a result of the third research question. The data source here will be Twitter postings (tweets) retrieved using twitter's firehose service. An ontology reasoner will be applied to these new triples to retrieve a set of extremist social movement organizations (SMO) that share the framings espoused by the online persona. How such reasoning takes place, and the rules that support such reasoning, is a topic for the description of the proposed ontology.

3.3. Design Science Research in Information Systems

This discussion begins with a brief presentation of design science. Simon (1996) refers to the science of the artificial, the very notion we (the academic discipline of information systems) term design science, and describes it as being "a body of knowledge about artificial… objects and phenomena designed to meet certain desired goals." Takeda et al. (1990) provide a representation of the reasoning process associated with the general design cycle (GDC) Vaishnavi & Keuchler (2007). Their model structures, in parallel, the concepts of knowledge flows, process steps and their logical formalism to provide a well-rounded picture of the reasoning process in the GDC. The process presented by Takeda et al. (1990) starts with an *awareness of a problem* to which a *suggested solution* is applied. This solution to the problem is then *developed* and *evaluated* to reach a *conclusion*. Their model highlights the iterative nature of design reasoning where the development, evaluation and conclusion pertaining to a problem-solution combination may lead to *circumspection* that iterates the reasoning process back to the awareness stage. Moving from being aware of a problem to *suggesting* a solution corresponds to the logical formalism of abduction. Moving from developing the proposed solution to evaluating the outcomes corresponds to the logical formalism of deduction.

Next, the outputs of design science research, borrowing heavily from the synthesis provided by Vaishnavi & Keuchler (2007), are discussed. March & Smith (1995) argue that there are four types of products (outputs) of design science research, that is, constructs, models, methods and implementations. The notion of *artifact* is central to design science research where an artifact is the result of the design scientist's production and application of knowledge concerned with tasks or situations (March & Smith ,1995). *Constructs* are a conceptual vocabulary applied to a problem/solution domain*; models* represent propositions or statements articulating the relationships amongst constructs*; methods* are sets of steps that are used to carry out a task such that the method is the problem and solution statement expressed in terms of the construct vocabulary; *instantiation* "operationalizes constructs, models and methods (March & Smith ,1995)," thus realizing "the artifact in an environment" (Vaishnavi & Keuchler, 2007). This dissertation instantiates a design artifact.

3.4. A Framework for Automatically Detecting the Resonance of Terror Movement Frames

We propose a framework composed of three major systems: a Terror Beliefs Ontology (TBO), a Frame Discovery System (FDS), and a Frame Resonance Detection System (FRDS). These three systems work together towards our objectives, and are controlled by a coordinating component. First, TBO is purposefully designed to expressively represent the framings of terrorist organizations while remaining amenable to frame discovery by FDS. Second, FDS is a collection of supervised, multi-class classifiers and rule or regular expressions based classifiers designed to extract information from text subject to the constraints put in place by TBO. Finally, we develop FRDS for determining when an inputted text from some online persona, website, or news article expresses semantics in line with the frames of one or more terror movements represented in TBO. We demonstrate that these three components are sufficient to perform the core tasks associated with our research objectives: (1) detecting online personas with potentially radical and terroristic views, (2) detecting websites spreading radical, terroristic propaganda[10] (3) identifying potentially triggering news articles.

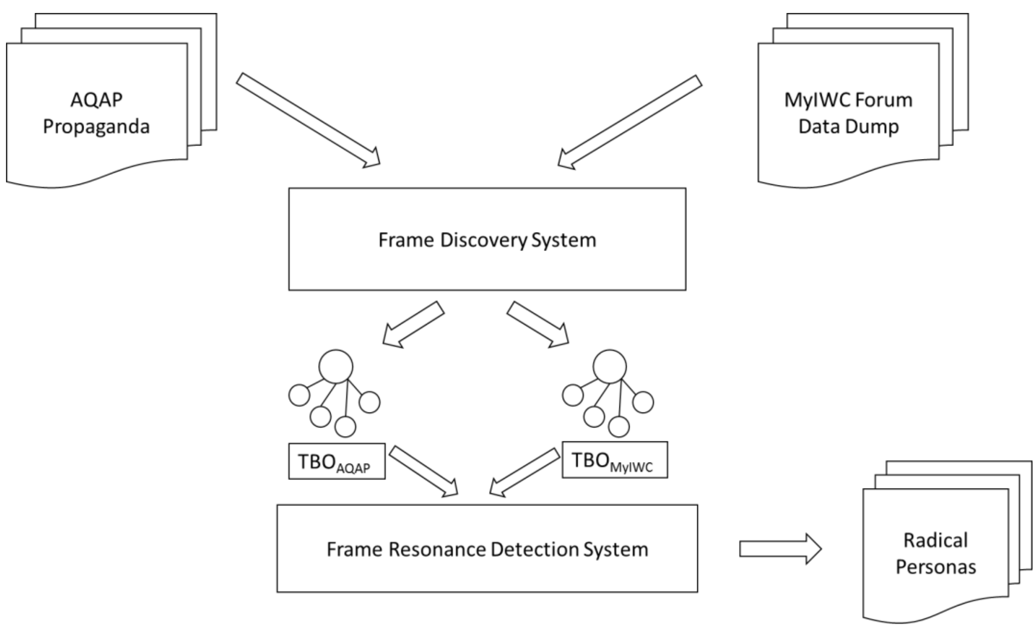

Figure 7. Knowledge Based Framework for Discovering Radical Content on the Web

[10] The first two objectives are compressed into a single design goal, that is, the ability to automatically detect radical content.

3.5 Ontology Development Methodology

A note on what is meant by "ontology:" The term "ontology" is used here to refer to an engineering artefact. This stands in contrast to the philosophical term typically associated with philosophical assertions vis-à-vis "the categorical structure of reality." It is, instead, what Gruber(1993) describes as an "explicit specification of a conceptualization" where a conceptualization is an "abstract, simplified view of the world that we wish to represent for some purpose" (p.1). This "view of the world" is typically referred to as the domain of the ontology (not to be confused with domain and range constraints on relations). Guarino (1998) describes an ontology as an "engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words." Ontologies are computable as they are implemented in any one of many formal ontology computer languages.

Clearly, the notion of conceptualization is central to the objectives of an ontology. In order to clearly describe ontologies we must first give a formal definition of a conceptualization. Following Guarino (1998) we define a conceptualization as a set of conceptual relations defined on a domain space. A domain space is defined as a structure <D, **W**> such that D is the domain relevant to our intended conceptualization and **W** is a set of possible worlds within that domain. Say we intend to conceptualize the domain of highway traffic patterns. The set **W** would contain all possible spatial and inertial configurations of cars traversing a highway. Now, if we want to conceptualize this domain, D, of highway traffic patterns, we apply a set of conceptual relations on the domain space creating a tripartite structure C = <D, **W**, $\mathfrak{R}$> where $\mathfrak{R}$ is the set of conceptual relations. A conceptual relation is a transformation over possible worlds **W** and the domain D into a set of valid ordinary relations. Given our example domain of highway traffic patterns, a conceptual relation over the domain and descriptive of a particular spatial and inertial configuration of cars traversing a highway (a possible world $w \in \mathbf{W}$) would be descriptive of that state, defining a set of permissible speeds and spaces between vehicles that satisfy the conceptual relation. In this way a conceptualization is set of conceptual relations defined for a domain space (Guarino 1998).

Given the above definition of a conceptualization we require a language suitable for representing domain conceptualizations. We would specify a logical language **L** constrained by a vocabulary V. For a given conceptualization our domain of interest we would specify a model using the vocabulary of **L** defined in terms of the conceptualization $C = \langle D, \mathbf{W}, \mathfrak{R} \rangle$ and an interpretation function that assigns elements of the domain to constant symbols of the vocabulary and elements of $\mathfrak{R}$ to the predicate symbols of V. A representation language thus specified has an *ontological commitment* to the conceptualization of the domain. It is in these terms that (Guarino 1998) defines an ontology:

> An ontology is a logical theory accounting for the intended meaning of a formal vocabulary, i.e. its *ontological commitment* to a particular *conceptualization* of the world. The intended models of logical language using such a vocabulary are constrained by its ontological commitment. An ontology indirectly reflects this commitment (and the underlying conceptualization) by approximating these intended models.

As implemented in knowledge bases such as ours, a more comprehensible definition is provided in (Gruber 1993) who writes:

> An ontology is an explicit specification of a conceptualization... [making] a systematic account of Existence. For knowledge-based systems, what “exists” is exactly what can be

> represented. When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the *universe of discourse*. The set of objects and the describable relationships among them, are reflected in the representational *vocabulary* with which a knowledge-based program represents knowledge.

This vocabulary comes in the form of classes, functions, relations and other objects whose interpretations and use are constrained by logical axioms. The vocabulary of an ontology is organized into a taxonomic class hierarchy. The hierarchy must have a single root concept usually called “Thing” from which all other concepts in the vocabulary follow. Each concept (class) is related to other concepts with “is-a” relationships indicating subsumption. While the ontology’s vocabulary is arranged in a taxonomic, subsumption hierarchy, the concepts that make up this vocabulary can be instantiated. Instances of these concepts are generally referred to as individuals. While individuals of a particular concept retain that concept’s taxonomic context, they can participate in a much broader set of relationship types (that is, relationships other than “is-a”). An ontology engineer or designer can specify these relationships, often referred to as object properties, that may exist between class instances. The designer can also constrain the ranges and domains of object properties such that these properties are *directed* and *restricted*. The domains and ranges of an object property are each a set of classes the individuals of which may participate as domain instances or range instances. The domain of an object property specifies the individual types that may participate on the left side of the object property. The range of an object property specifies the individual types that may participate on the right side of the object property. Relationships between individuals resemble network structures whereas relationships between classes are hierarchical. Several enriching axioms can be added to object properties in the most expressive of ontology languages.

A critical component of this research involves conceptualizing the domain of terrorist movement framings. In the previous chapter, the framework of collective action framing theory and a profile of its usage in the study of terrorism were provided. The application of this theory in this dissertation is primarily with respect to guiding the analytic process of conceptualizing the domain. Because an ontology draws from a conceptualization, it is vital that the conceptualization be analytically useful, and accurate. A sound, empirically grounded overarching theoretical framework can guide the emerging conceptualization towards being both analytically useful and accurate. Framing theory identifies three core framing tasks that all social movements must engage in. Because terrorist organizations are considered to be social movements, the framing perspective is easily ported to terrorism studies (Gunning 2009). Each of the framing tasks results in corresponding frames. The three core framing tasks are (1) diagnostic framing, (2) prognostic framing, and (3) motivational framing. Diagnostic framing processes result in diagnostic frames, which are schemata for attributing blame for a problematic situations. Prognostic frames, produced as a result of prognostic framing processes, posit solutions, in the form of action, to the problematic situation, as diagnosed. Motivational frames provide mobilization impetus for potential movement followers. Given a theory through which the domain can be conceptualized in an analytically useful and accurate way, a methodology is required for developing an ontology that commits to the conceptualization. METHONTOLOGY (Fernández-López et al. 1997a) is amongst the earliest and potentially best known of the available ontological engineering methodologies and is still much in use today (see for example Jaskolka et al. 2015). METHONTOLOGY will be used to inform the systematic design of the ontology of terrorist movement framings. While METHONTOLOGY provides support for the development strategy employed in this dissertation, it is not adopted wholesale. In particular, the approach presented here draws from the “specification” and “knowledge acquisition” stages of METHONTOLOGY.

The specification stage is intended as a means to define the requirements, purpose, and scope of the ontology, including the parameters for acquiring the knowledge required to specify the ontology. METHONTOLOGY provides a format for conducting requirements analysis and that format is adopted here in the table below.

Ontology Requirements Specification	
Domain	Terrorism and the framings used by terrorist movements to inspire action and mobilize supporters.
Purpose	An ontology about the framing of terrorist movements' extremist ideologies. The ontology has many potential user groups: First, counter terrorism practitioners can use the ontology to estimate which terrorist movements are likely to react to unfolding world events. In addition, counter terrorism practitioners can use the ontology to analyze the manifestos and written documents of individual perpetrators of terrorism to determine which terrorist groups may have contributed to the individual's radicalization. Further, these experts could also use the ontology to detect potential alliances and conflicts between different terrorist organizations. Second, academics may use the ontology to conduct empirical, comparative studies on terror movement framings. The ontology provides numerous actionable variables that may be of interest to this user group. Finally, the ontology is proffered in a serializable and machine agnostic formal language such that it can be integrated into varied AI systems for counterterrorism.
Level of Formality*	Rigorously-Formal: the ontology will be expressed in Web Ontology Language (OWL), a computer language for specifying formal semantics, making it at least semi-formal. The use of intensional semantics (see Guarino 1998 p. 5) ensures that the meaning of an ontological term in the proposed ontology is implicit in its representation. Rigorous evaluation criteria are integrated into a requirements oriented framework such that the objectives of the ontology. The criteria include concepts such as correctness, completeness and minimal ontological commitment.
Scope	The high-level scope of the proposed ontology is given by collective action framing theory. More specifically, the framing construct of *core framing tasks* and its sub-constructs of *diagnostic*, *prognostic*, and *motivational* frames must be adequately represented in the classes, relations, and properties of the ontology. The scope is defined by

	empirically investigating the following: what are the concepts, relations, and properties sufficient to model the conceptualization of terrorist movement core framing tasks.
Sources of Knowledge	Primary sources are used to determine the concepts, relations, and properties of the ontology. Theoretical sampling is used to select these sources, where the theory is collective action framing, and such that the analysis of initial sources determines the selection of additional sources. Particular care is taken to ensure that there is variety with respect to the movement types included in the sample. Generally, knowledge sources take the form of articles published on websites verifiably belonging to terrorist organizations. Preliminary analysis indicates that these sites usually contain pages dedicated to articulating the foundational philosophy of the movement. Such pages contain detailed information regarding the diagnostic and prognostic framings of the movement. In addition, terrorist organization websites invariably include a newsfeed through which the organizations are able to apply their schemata of interpretation on real-world events as they unfold. Preliminary analysis indicates that such pages are rich information sources regarding the motivational framings of terrorist movements. In the grounded theory tradition, a theoretical sample of such Web publications is collected and coded. The coding process is patterned by the informing theoretical perspective such that the main analytical categories are predetermined. Sub categories within the main analytical categories are discovered through the coding process.

*Uschold and Gruninger (1996)

Highly informal - expressed loosely in natural language

Semi-informal - expressed in a restricted and structured form of natural language

Semi-formal - expressed in an artificial formally defined language

Rigorously formal - meticulously defined terms with formal semantics, theorems and proofs of such properties as soundness and completeness

Table 3 – Requirements Specification for Ontology of Terror Movement Framings

Given the above requirements, it is now possible to more closely sketch out the knowledge acquisition phase. According to Fernández-López et al. (1997), knowledge acquisition for ontology development may entail “formal and informal analysis of text techniques in books and handbooks in conjunction with structured and non-structured interviews with experts” (p. 37) to construct a glossary of terms relevant to the domain. To conceptualize a domain entails garnering information regarding the form and content of the domain. This requires the collection and analysis of data, either or both primary and secondary sources. The ontology development process documented here relies upon primary data sources to determine the concepts, relations and properties of the ontology. The collection of ontological concepts, relations, and properties may be referred to as the glossary[11] of the domain conceptualization.

To ensure rigor in the development of this glossary, the analysis of knowledge sources is informed by well-established qualitative coding methodology. When qualitative coding is discussed, grounded theory typically comes to mind. The grounded theory (GT) approach to qualitative data analysis was initially proposed in (Glaser and Strauss 1967). Since that initial book, the concept has evolved greatly and with much controversy. GT has been described as "the discovery of theory from data systematically collected and analyzed without structured forcing through predetermined theoretical frameworks" (Holton and Walsh 2016, p. 12) such that emergent theory is grounded in empirical data and free of preconception. Preconception is essential to the development of the proposed ontology, and accordingly, GT, taken "as-is" is not appropriate for the purposes here. Within the classic GT framework, however, a number of useful coding techniques have been developed (Saldaña 2015). (Urquhart 2012) describes coding as the tagging of conceptual labels onto chunks of data.

Thematic coding is one such approach to qualitative coding (Braun and Clarke 2006; Urquhart 2012) and is adopted here for the analysis of knowledge sources. Thematic coding can be approached in one of two ways – top-down or bottom-up (also called inductive thematic

[11] Formal ontologies are constrained by ontology modelling formalisms and the capabilities of the selected ontology modelling language. Accordingly, there is bound to be 'slippage' between the final ontological representation and the conceptualization embodied in the glossary. The inverse of this slippage may be thought of as the ontological commitment of the final ontology.

analysis). The thematic coding strategy used here is the top-down approach (also referred to as theoretical thematic analysis) because it uses predetermined theoretical categories extracted from the literature (Urquhart 2012). It also ensures that the resultant codes are applicable beyond the data, thus avoiding overfitting the data to the analysis. This is important here because framing theory has been positioned as a sense-making framework for conceptualizing the ideologically informed frames espoused by terror movements. According to (Braun and Clarke 2006), a "theme captures something important about the data in relation to the research question and represents some level of *patterned* response or meaning within the data set" (p. 11). Three themes are identified from the framing literature and applied to the data: diagnostic, prognostic and motivational frames. These concepts are discussed in detail in Chapter 2. Each of these three themes is further distilled into codes within the data. In addition to the decision regarding inductive vs. theoretical thematic analysis, a decision needs to be made about the conceptual level at which themes may be identified (Braun and Clarke 2006). Given a suitable conceptual level of coding, all codes must adhere to the choice. One such conceptual level is referred to as a semantic approach. Here, the codes are focused on surface meanings, focusing on what study participants and/or selected data sources say. This stands in contrast to an interpretive approach where the goal is to identify underlying concepts represented within the data (Braun and Clarke 2006). The semantic approach is selected for this analysis because of the purpose of the coding exercise. Extracted codes will eventually be converted into a glossary and later into formal semantics. The resultant formal semantics will be used to guide information retrieval from textual data sources. A coding approach that is strictly interpretive may lead to complications when attempting to use the resultant formal semantics for information retrieval. This is because the conceptual codes are less likely than semantic codes to be directly observable in text. An indispensable requirement for ontology guided information retrieval is that the concepts asserted within the ontology be common features of the candidate texts. Given a set of thematic codes (created as described above), a concise but complete glossary of classes, relations, and properties relevant to the domain can be defined. Doing so involves an analytical process where extracted codes are organized by their respective theme, areas of overlap are identified, relationships between codes are elaborated and a distilled set of constructs are extracted.

Of course, no coding can be done without data. This paragraph explores the theoretical sampling approach used to collect data sources for the above coding exercise. Theoretical sampling is a process whereby data gathering co-occurs with theory building. This sampling approach is native to the GT paradigm (Glaser and Strauss 1967) but is not exclusive to the domain. As codes are generated, the researcher acquires insight into how data sources should be selected. This dissertation relies on published writings of terror movement organizations as data sources for thematic coding. The Profiles of Perpetrators of Terrorism in the United States (PPT-US) dataset[12] is used as a starting point for the identification of terrorist groups. The dataset includes a variable that captures the date of the last US attack conducted by any given group. The 10 groups that have most recently perpetrated attacks against the US homeland are selected as a starting point for the proposed analysis. The reason behind this is the intuitive contention that active groups are more likely to maintain detailed websites from which data can be gathered. Given a terrorist group amongst the 10 selected, its official website is identified and scoured for data that might glean insight into the three theoretical constructs of import here. Preliminary coding has indicated that (1) all the websites visited contain a section of pages which describe in great detail, the terror movement's *raison d'etre*, (2) that these descriptions are valuable sources

[12] https://www.start.umd.edu/data-tools/profiles-perpetrators-terrorism-united-statesppt-us

for insight into the diagnostic and prognostic frames espoused by the movement but suffer from a dearth of insight into motivational framings. Given this preliminary observation, it is observed (1) that all the websites visited contain a section of pages which may be considered "newsfeeds" through which the movement is able to apply its schemata of interpretation on current events, and (2) that these newsfeeds are valuable sources to understand the movement's motivational frames. As thematic coding progresses, it is expected that additional sampling insights will emerge.

3.6 Frame Discovery

The ontology described above is an ontology of terrorist movement frames, which should be populated with instances (individuals) of such frames. This dissertation proposes an automated approach to frame discovery wherein an automated agent analyzes an imputed text and returns a collection of frames discovered within the text, if any. Frame discovery therefore is an information extraction problem. Information extraction can be understood as a process which "selectively structures and combines data which is found in one or more texts" (Cowie and Lehnert 1996). The frame discovery system proposed herein is tasked with doing just that: selectively structuring data which is found in one or more texts where this structure is defined in a knowledge base (ontology). An approach for discovering frames from English texts is presented below in more detail.

Information extraction systems typically combine two interdependent information extraction tasks: entity extraction and relation extraction in a pipeline architecture. The entity and relation extraction view of information extraction has its roots in Conceptual Dependency Theory (CDT), originally proposed by Roger Schank (Schank 1972). De Busser writes of the basic assumptions undergirding CDT where "there exists a conceptual base that is interlingual, onto which linguistic structures of a given language map during the understanding process and out of which such structures are created during generation" (Moens 2006). CDT's main thrust, as stated above, has major implications for modern conceptualization of information extraction implicitly bearing upon architectures and tasks in the domain. Schank's "conceptual base," or conceptualizations are structures based on a set of primary concepts and interconnections that may exist between these primary concepts where these interconnections are governed by a "closed set of universal *conceptual syntax rules* and a larger set of concept specific *conceptual semantic rules*" (Moens 2006). This model of concepts and concept interrelationships is the conceptual base capable of mapping linguistic structures and is also the conceptual basis from which linguistic structures are created. CDT has provided a comprehensive model (schema) that makes it possible to develop artificial intelligence systems capable of semantically analyzing entire texts.

The widespread use of CDT to undergird information extraction research inspired the rise of methods to represent the knowledge extracted from texts. CDT essentially specifies a universal schema meaning conveyance. Incidentally this schema lends itself well to machine representation. Beginning with frames[13] (these frames are distinct from the rhetorical frames referred to throughout this dissertation and as such will be referred to as representation frames), AI researchers have made several advances in creating versatile machine readable knowledge representation structures. Knowledge representation and information extraction are thus inexorably linked: CDT, which makes information extraction possible, also underlies how such information is represented.

[13] According to Marvin Minsky, "[a] frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room or going to a child's birthday party" (Minsky 1974). The idea of a stereotyped situation is analogous to the idea of a type and its instances.

It follows, then, that what is referred to as frame discovery in this dissertation is a method of information extraction and subsequent representation. The next section presents the technical details regarding the design of terror frame discovery system implemented in this dissertation.

3.6.1. Named Entity Recognition (NER)

Named Entity Recognition (NER) is a critical subtask in many information extraction systems. For the purposes here (and quite generally), the NER problem is a classification problem wherein for a collection of tokens we attempt to estimate the probability that a sequence of these tokens $(t \geq 1)$ signify a predefined category. In Figure 8, the ellipses correspond to tokens. Assuming that a hypothetical analyst would like to identify those sequences of tokens which signify a named mention of a *person*, an entity extraction system should label the tokens "Osama" at t_0 and "Bin-Laden" at t_1 respectively as *person*.

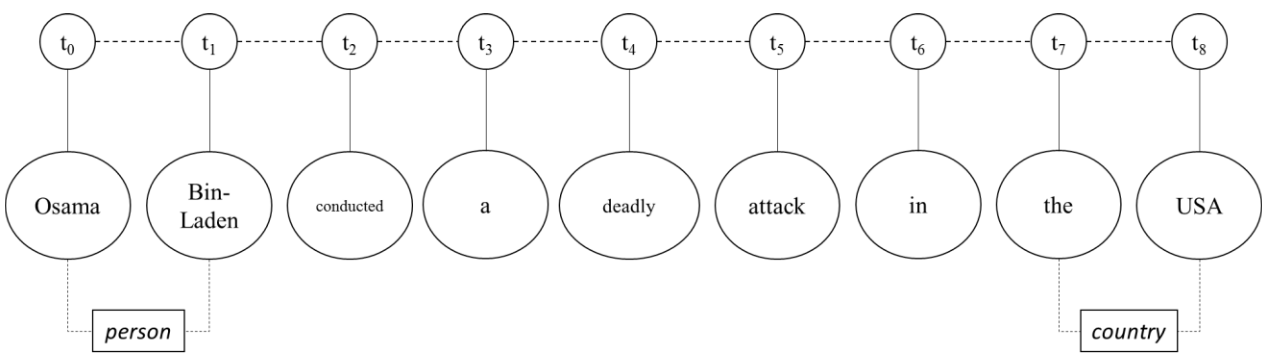

Figure 8 – Tokens Occurring in a Sequence

As a classification problem, NER may be performed using any of a wide array of classification techniques. The choice of classification technique depends on the domain of inquiry amongst other things. The literature on computational approaches to NER is vast and thus cannot be fully reviewed here. However, it is critical to provide an historical context to NER to better expound upon the design choices made in this dissertation. The tenability of NER as a robust and tolerably accurate technique for extracting information from documents is rather recent. In 1996, MUC-6[14] in response to concerns that the systems presented in prior MUCs "were tending towards relatively shallow understanding techniques (based primarily on pattern matching), and that not enough work was being done to build up the mechanisms needed for *deeper understanding.*" (Grishman and Sundheim 1996, p. 468) shifted focus to techniques that could enable such understanding by automated agents. Since MUC-6, the pace of work in NER accelerated with particular interest being paid to biomedical and defense applications (Sekine and Ranchhod 2009a).

An important consideration with NER systems is the scope of the categories of labels for which available techniques are useful. Not every category of thing can be successfully extracted by an NER system. Many advances in NER classification emerge from NER conferences such as the MUCs. These conferences furnish standard corpora and standard evaluation designs for a (other machine understanding problems apply) set of NER classification problems and attendees submit systems that compete for the best performance. This structure of scientific acceptance so to speak has led to narrowly scoped NER systems developed with rather constrained textual genres or domains. As these gatherings provide both corpora and named entity category, a consensus has developed in NER community regarding the named entity categories that lend themselves to discovery by NER systems. The most prominent of these categories outside of the biomedical field are *person*, *location*, and *organization*. Indeed, the word "Named" in NER "aims to restrict the task to only those entities for which one or many rigid designations… stands for the referent" (Sekine and Ranchhod 2009b). To understand what this means we turn to Saul Kripke who writes "Let's call something a *rigid designator* if in every possible world it designates the same

[14] MUC stands for Message Understanding Conference where MUC-6 was the sixth conference in the series.

object” (Kripke 1972, p. 200). The *referent* of a rigid designator is some unique object that satisfies the properties of the designator. Kripke goes further by stating that only proper names may be rigid designators. His reasons for this are compelling but further description is out of scope. No matter, Kripke’s work underlies the tendency for NER systems to focus on entity labels that correspond to names. This study joins others (Alfonseca and Manandhar 2002) in taking the position that such restrictions unduly constrain the applications of NER classification systems and their encompassing information extraction systems. Restricting the breadth of an information extraction systems’ NER subsystems’ label set to *person*, *location*, and *organization* also restricts knowledge that may be discovered using such a system.

The set of entities required to adequately support the knowledge acquisition requirements of the frame extraction process is significantly broader than (but inclusive of) *person, location* and *organization*. It includes entity categories such as *ideology*, *peoples*, *problempractice*, *religion* and more. Indeed, one of the critical arguments of this dissertation is that the set of designators amenable to discovery by an NER classification system extends beyond proper names or proper nouns when that system is couched within a well constrained knowledge acquisition framework equipped with a well-designed ontology.

There are several available methods for conducting the classification tasks at the heart of a named entity recognition system. As with many classification tasks, the classification objective of NER systems is the recognition of entities that are not known in advance. This is accomplished by defining a set of rules for classification that fire upon observation of certain discriminative features of the domain (Sekine and Ranchhod 2009b). State-of-the-art systems are imbued with machine learning algorithms which assign weights to such discriminative features. The literature is rich with supervised, semi-supervised, and (less so) unsupervised approaches that either “automatically induce rule-based systems or sequence labeling algorithms starting from a collection of training examples” (Sekine and Ranchhod 2009b). Considering the submissions from the last CoNLL shared task concerned with NER provides a snapshot of the machine learning approaches typical in NER systems. CoNLL is the Conference on Computational Natural Language Learning and is generally considered to be amongst the most influential conferences in the genre. CoNLL 2003[15] is the last of the series that dealt with NER. Its shared task concerned language independent NER and focused on four classical named entity categories: *person*, *location*, *organization*, and *miscellaneous*. Sixteen systems participated in the shared task using a diverse set of machine learning techniques. Florian et al. (2003) used an ensemble of machine learning techniques to produce the best results on the shared task. Their system combined maximum entropy, hidden Markov, a robust risk minimization classifier based on a regularized winnow method and a transformation-based learning classifier (Florian et al. 2003). Klein et al. (2003) focus on using character level tokens as opposed to words and classify to the four named entity categories using both hidden Markov and conditional Markov models. Another participant, (McCallum and Li 2003) show early results for a system using conditional random fields (CRFs) for NER.

For the NER subsystem within our frame discovery information extraction system we choose to use is a CRF based NER classifier capable of incorporating non-local (the ability to incorporate non-local information is key benefit of CRFs and Finkel et al. makes this computationally tenable) information described here (Finkel et al. 2005). This classifier is part of the CoreNLP suite maintained by NLP group at Stanford. This is chosen for several compelling reasons. First, CoreNLP suite is a comprehensive product with software comprising many essential NLP

[15] http://www.clips.uantwerpen.be/conll2003/ner/

subtasks. Its adoption ensures that integration of these tasks into a suitable architecture can be achieved in the same runtime environment. Choosing CoreNLP for the NER subsystem thus facilitates other portions of this work that rely upon NER. A second reason is that CoreNLP's NER component exposes an excellent CRF classifier to the java API. It is open source and thus customizable to domain specifics. Per our choice of a CRF classifier, it is not feasible to justify that choice based on past performance of CRF classifiers observed in the literature. While (Finkel et al. 2005) report that CRFs generally outperform other classifiers on the NER task, it is quite evident that several other, interrelated factors affect classifier performance on the NER task. Such factors include the choice of features (more on features later) used for the task, the unit of analysis (i.e. what is considered a token, be it words or character n-grams, or even single characters), and the domain of interest. As such, the choice of Stanford's CRF NER classifier is based on a combination of considerations regarding the superior performance of CRF classifiers on the NER task, and the characteristics of Stanford's CoreNLP suite that make it an appropriate choice for the frame discovery system more broadly.

Building a CRF, as with any classifier, first requires that a set of feature functions be defined where each feature function f_k takes a series of inputs and produces a real number (e.g. 0 or 1) that expresses whether the conditions of the feature have been met. The Stanford implementation of CRF is a linear chain CRF such that there is a restricted window of labels, $l \in \boldsymbol{L}$, on which any given feature may depend. Labels correspond to the named entity categories in which we are interested (e.g. *person*). Let $\boldsymbol{o} = \langle o_1, o_2, \dots\ o_T \rangle$ be the input sequence. In our case this sequence is of words appearing in a document. There also exists a sequence of states that we are interested in $\boldsymbol{s} = \langle s_1, s_2, \dots s_T \rangle$ such that a state is assigned to each outputted observation and can take some value in $\boldsymbol{L}$. Note that we output as many states as there are observations where each state takes some value in $\boldsymbol{L}$. A CRF will then estimate the conditional probability of a state sequence given an input sequence as:

$$\text{Equation 1: } P(\boldsymbol{s}|\boldsymbol{o}) = \frac{1}{Z_{\boldsymbol{o}}} \exp(\sum_{t=1}^{T} \sum_{k} \lambda_k f_k(s_{t-1}, s_t, \boldsymbol{o}, t))$$

A normalizing factor Z_o is used in conjunction with exponentiation and defines the above probability as some value between 0 and 1:

$$\text{Equation 2: } \sum_{\boldsymbol{s} \in S^T} \exp(\sum_{t=1}^{T} \sum_{k} \lambda_k f_k(s_{t-1}, s_t, \boldsymbol{o}, t))$$

The normalizing factor is computed over all possible state transitions (state sequences) such that the number of such transitions exponentiated on T. For each state transition, the sum of all weighted feature scores for every possible state is taken. The resultant sums are summed over all possible states and subsequently summed over all possible state transitions. Note that, because this is a special case of CRFs where the window for state transitions is constrained, the Equation 1 (in light of Equation 2) is computationally tractable using dynamic programming (typically the Viterbi algorithm). The above yields:

$$\text{Equation 3: } P(\boldsymbol{s}|\boldsymbol{o}) = \frac{\exp\left(\sum_{t=1}^{T} \sum_{k} \lambda_k f_k(s_{t-1}, s_t, \boldsymbol{o}, t)\right)}{\sum_{\boldsymbol{s} \in S^T} \exp\left(\sum_{t=1}^{T} \sum_{k} \lambda_k f_k(s_{t-1}, s_t, \boldsymbol{o}, t)\right)}$$

It was noted earlier that each feature is weighted where this weight is denoted by λ_k for the k^{th} feature f_k. The weights of a CRF are developed by training the CRF (recalling that CRFs are supervised) on a corpus of examples. These weights are developed such that they "maximize the conditional log-likelihood of labeled sequences" (McCallum and Li 2003) within the training set. Let the weights of a CRF be $\Lambda = \{\lambda_j \ldots\}$ and let the training set be $D = \{\langle \boldsymbol{o}, \boldsymbol{l} \rangle^1, \ldots \langle \boldsymbol{o}, \boldsymbol{l} \rangle^j, \ldots \langle \boldsymbol{o}, \boldsymbol{l} \rangle^N\}$. Several approaches may be taken to maximize the conditional log-likelihood of labeled sequences within a labelled training document. The idea is to recursively, and for each feature, calculate the gradient of the log probability of the training label given the observed sequence beginning from some arbitrary weight. The arbitrary weight is then adjusted towards the gradient at each pass with a prespecified rate.

Feature selection is a critical determinant of the performance of an NER system. In this study, we opt to define features for groupings of named entity categories (or label), where the expectation is that each grouping contains categories amenable for the selected set of features. The ideal scenario would be to define a unique feature set for each label but this significantly slows down the system when the frame discovery process is initialized. CoreNLP exposes a diverse set of features and allows for the customization of features. In subsequent chapters, we discuss the features chosen for each grouping labels. The choice of training documents is also critical. The objective is to select a training corpus that is as representative as possible of the domains relevant to our study. The training set is described in ample detail later in the dissertation.

3.6.2. Relation Extraction

We define the relation extraction problem narrowly. First a relation extraction or relation detection system processes an input sequence of tokens annotated by an NER system. That is, each token in an input sequence of tokens has one or more NER labels in $L = \{l_1, \dots l_j, null\}$. The relation extraction problem, then, is to estimate some conditional probability that a pair of annotations within a given sentence expresses a specified relationship.

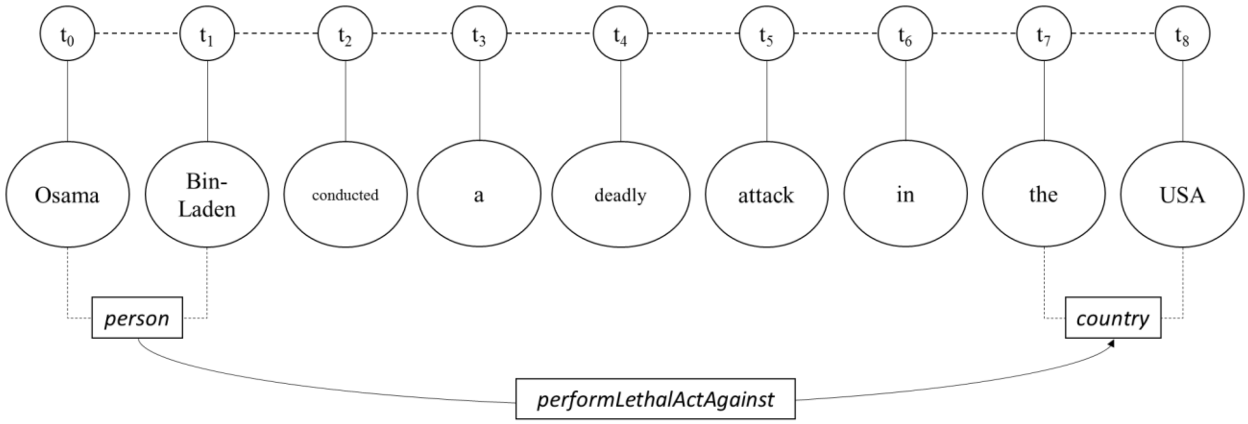

Figure 9 – Relation Extraction Illustrative Example

Figure 9 provides a basic illustration of the problem, the tokens t_0 and t_1 as well as t_7 and t_8 have been sequentially labelled *person* and *country* respectively by the NER subsystem. It is the task of the relation extraction subsystem to generate conditional probabilities (across its trained relation space given the observed tokens in the sequence) that estimate the likelihood that the tokens labelled *person* and *country* participate in a particular relation. The problem is direction sensitive; the classification performLethalActAgainst(*person, country*) is different from the classification performLethalActAgainst (*country, person*). In other words the classifier tasked with detecting relations between entities in a labelled sequence of tokens must be imbued with features sufficient to discriminate between, say, the occurrence of performLethalActAgainst (*person, country*) and performLethalActAgainst (*country, person*).

In the literature there exist two broad categories of relation extraction approaches: supervised and semi-supervised (Bach and Badaskar 2007). Supervised approaches mainly employ feature or kernel based classifiers and require considerable annotated text input to learn these classifiers. Feature based classifiers use feature functions like those discussed in the previous section. Kernel based approaches to relation extraction rely on parse trees as a basic structure on which similarity computations are made (Zelenko et al. 2003). Semi-supervised approaches are generally seeded with a minimal set of examples. These seeds are used to induce an initial set of patterns which are subsequently applied to unlabeled data thus enlarging the seed set and enabling the induction of additional patterns.

While there are various classification approaches used for relation extraction, there is much common ground in the text preprocessing required for relation extraction, such as “chunking” and “parsing” of text. Text chunking attempts to divide sentences into “nonoverlapping segments on the basis of fairly superficial analysis” (Ramshaw and Marcus 1999, p. 1). These target segments are phrases which encompass syntactically related words (Tjong Kim Sang and Buchholz 2000). Typically, systems that chunk texts will conform to the chunk types specified by one of the shared tasks on chunking, for example CoNLL-2000. Chunk segments are based on the syntactic category portion of a tree-bank where the chunk will contain everything left of the syntactic head delineated by the syntactic category. A full discussion of tree banks is beyond the scope of this dissertation; however, a brief discussion of the Penn Tree Bank’s syntactic tag set is critical to understanding how these tags play a pivotal role in chunking, parsing and subsequently, relation extraction. The syntactic tag set in the Penn Tree Bank (Marcus et al. 1993) defines the types of phrases that may exist in the English language. For example, NP denotes a noun phrase and VP denotes a verb phrase and so on. The decision to classify a phrase into any of these categories is based on the part-of-speech of that phrase. If the head word of a phrase is, for example, a singular common noun, then the phrase will be tagged as a noun phrase. A given phrase can be part of a larger phrase and so on, up to the level of a sentence. This characteristic is often represented as a parse tree, such as in Figure 10 below. Parsing thus provides a representation of the grammatical structure of a sentence, providing a constituent dependency pathway (sequence of words) between two named entities in a sentence. This pathway is an essential feature for many relation extraction systems.

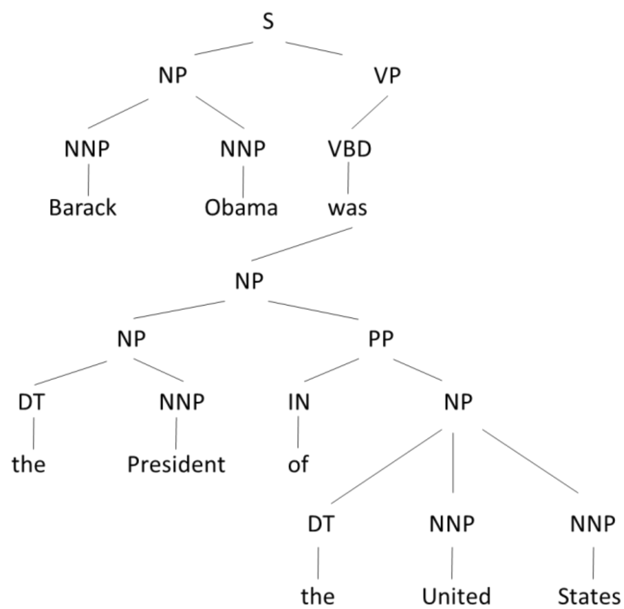

Figure 10 – Syntactic Parse Tree Illustration

For our purposes, we choose a supervised, feature based approach to relation extraction for several reasons. First, a suitable, trainable, supervised classifier for relation extraction is exposed via Stanford's CoreNLP Java API. Recalling that we also use Stanford's CoreNLP Java API for named entity recognition and other NLP subtasks, our choice here is partly based on a preference for continuity. More critical to our choice, however, is the requirement for this project to develop a large set of domain specific relations. As such, there was a need to hand-craft the examples to be used in developing a classifier. Stanford's relation extraction implementation is based on the work presented here (Surdeanu et al. 2011). A multi-class logistic regression classifier is employed for classification.

3.6.3. A Pipeline Architecture for Frame Discovery

The previous two sections describe the two major subsystems that comprise the Frame Discovery system: named entity recognition and relation extraction. However, these two subsystems are not merely dependent on each other but also on several text preprocessing subsystems. Stanford's CoreNLP Java API provides for the design of a pipeline architecture for information extraction. We customize the API slightly to accommodate our needs. Figure 10 illustrates the pipeline we propose herein.

#	**Task**	**Description** see (Manning et al. 2014)
1	Tokenization	This task breaks up text into a series of tokens (which roughly correspond to words). The Stanford Tokenizer is based on the Penn Tree Bank corpus but is also tuned to handle noisier data.
2	Sentence Boundary Detection	Given a sequence of tokens (boundary detection can be done before or after tokenization (Wilcock 2009)) sentence boundary detection splits those tokens into sentences.
3	Part-of-Speech Tagging	We cannot do justice to the concept of Part-of-Speech (POS) tagging here, however, this task involves assigning a token a grammatical function given the context of its use. For example, some words are nouns while others are verbs and so on.
4	Lemmatization	Lemmatization is the process of transforming an observed token (word) into its normalized form by removing or adding suffixes to the word. For example, the words *travelling*, *travels*, and *traveled* would be normalized to *travel*.
5	Parse	See section 3.2.2 for our discussion on parsing.
6	Dependency Parsing	The object of dependency parsing is the production of a dependency graph given some input sequence of tokens in a sentence. The dependency graph is a labeled and directed graph with nodes that correspond to tokens in the input sentence and arcs that are formed between nodes and labeled with some dependency category (Nivre et al. 2007). These labels are based on some dependency grammar that describes the set of grammatical relationships that can exist between tokens. Examples include *amod* - adjective modifier, *det* – determiner, *infmod* – infinitival modifier (De Marneffe et al.

		2006). We employ the Stanford parser here which uses grammars defined in (Carroll et al. 1999).
7	NER	See section 3.2.1 for our discussion on NER.
8	Relation Extraction	See section 3.2.2 for our discussion on relation extraction.

Table 4 – Pipeline Architecture for Relation Extraction

3.7 Implementation

The section within these subsections attend to the implementation details of the Terror Beliefs Ontology, the Frame Discovery System and the Frame Resonance Detection System. Chapter 3, until now, has focused on the methods that we have used to design a methodology towards automatically identifying radical content and populating a knowledge repository of radical ideologies. Below, we discuss the implementation details of the various system components individually.

3.7.1. Implementation of Terror Beliefs Ontology

A note on OWL 2, SPARQL DL and Pellet Reasoner: In previous sections, we described the theoretical form of ontologies. As this section is dedicated to implementation details, a brief discussion of the technology stack used to implement those theoretical ideas is warranted. At the bottom of the stack is Web Ontology Language 2 (OWL 2), a World Wide Web Consortium (W3C) Standard. OWL 2 provides classes properties individuals and data values such that they can be stored in Semantic Web documents (Motik et al. 2009). OWL 2 allows for a formal description of a domain of interest in terms of three broad syntactic categories: (1) *entities* including individuals (instances) properties and classes which form the primitives of the language, (2) *expressions* that provide for complex constraints to be placed upon OWL 2 entities, and (3) *axioms* which express assertions of truth within the domain (e.g. an axiom would be used to assert that a given class has a given instance or that a given class has a particular subclass).

Used in concert, these three elements can provide logical semantics in a model for a domain conceptualization to the extent that inferences can be drawn from the model (inferences are inferred axioms, that is, axioms that were not asserted by the model-maker). The Pellet Reasoner (Sirin et al. 2007) is the most capable, pure-java, open source licensed reasoner available. Its key strengths lie in its coverage of all OWL 2 DL semantics, its support for entailment[16] and datatype reasoning. Ontology reasoning provides several additional features to an ontology. A reasoner validates the logical consistency of an ontologies axioms and assertions. Reasoners such as Pellet can also provide decidability on user defined entailment rules. In fact "reasoning… [is] an important goal of DL language design (Krötzsch et al. 2012). Reasoners are crucial to our work as we specify a large set of entailment rules. Finally, SPARQL DL is a versatile query language that provides querying functionality over the full semantics of OWL 2 DL. It is the only well supported language of its kind to support both the graph structure of ontologies as well as the descriptive logics of OWL 2 DL. Further, SPARQL DL (and this is the most compelling feature for our use cases) provides querying over TBox, ABox and RBox components of a knowledge base (or ontology). An ABox axiom captures knowledge about instances asserted in a knowledge base (assertional axioms) and how these instances are related to each other. For example `Mother(julia)` is an assertional axiom stating that the named individual `julia` is an instance of the concept/class/type `Mother`. Another example is `worksFor(steve_jobs, apple_inc)` which is an assertional axiom stating that the named individual `steve_jobs` is related to the named individual `apple_inc` by the `worksFor` relation. A TBox axiom (a

[16] Entailment reasoning is the ability of reasoner to follow a collection of entailment rules on an ontology and infer a reasoned ontology. Entailement rules are rules which follow logical consequence, that is, when a set of conditions are meant, something follows from that. Semantic Web Rule Language (SWRL) is an entailment rule framework supported by Pellet.

terminological axiom) describes the relationships between concepts/types/classes. For example, a TBox axiom would be used to model the fact that all for-profit companies are organizations:

$$ForProfitCompany \sqsubseteq Organization$$

Such that the concept/type/class `Organization` subsumes `ForProfitCompany`. Finally, RBox axioms are the properties given to roles (relations/relationships) in a knowledge base. Roles, for instance, also support subsumption such that for the role `worksFor` we could assert that it subsumes the role isChiefExecutiveOf:

$$worksFor \sqsubseteq isChiefExecutiveOf$$

These axioms can become quite complex and enable a high degree of expressivity. SPARQL DL allows us design highly expressive ontologies with the knowledge that we are able to query over our complex axioms.

TBO, as implemented, balances two critical objectives: (1) providing high expressivity with respect to the representation of terror movement frames, and (2) guiding a frame discovery system capable of populating TBO. We sought to ensure that the design of TBO was intuitive, amenable to easy querying (we chose SQPARQL DL as our query engine as it supports ABox, TBox and RBox queries on OWL 2 DL), easily modifiable and easily integrated with other ontologies. First we defined a collection of classes as immediate children of `OWL:Thing` (that is the root node shared by all OWL ontologies). These top-level concepts are `Claimant`, `Frame`, `FrameRelation`, and `FrameEntity`. An instance of `Claimant` is some entity, a person or an organization, perhaps, that is the source of an ontological fact[17]. An instance of `Frame` is a

[17] An ontological fact is the assertion of an axiom within the ontology. The word fact, in this sense, has no mandatory relation to real-world truth.

representation of the framing activity that is responsible for a particular frame. Recalling that we are interested in three framing activities, diagnostic, prognostic and motivational framing, an instance of `Frame` can take one of these three forms. An instance of `FrameEntity` is some real world entity, one which may be invoked in a terror movement's frame emittance (more on this later). A `FrameRelation` instance is an instance of one of a collection of defined relationships that mobilize `FrameEntity` instances into meaningful terror belief frames. Figure 10 illustrates the top level hierarchy.

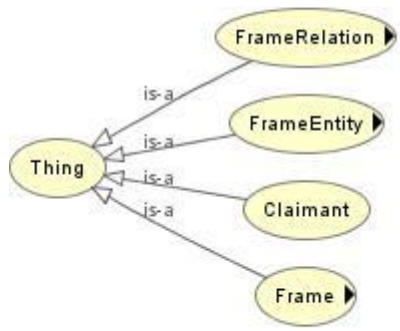

Figure 11 - Top Level Concepts in TBO

Figure 11 shows these top-level classes and their subsumption relationships with respect to `OWL:Thing`. `OWL:Thing` ensures that all ontologies have a common root node from which all other concepts flow. All ontologies implemented in OWL will have `OWL:Thing` as their root node. The above design provides for an expressive formalism for representing the substance of terrorist movement frames. We determined that the substance of a frame is expressed in the form of some action, forecast, label or practice. An action involves a prescription that some instance of `FrameEntity` takes some action with respect to another instance of `FrameEntity`. Similarly, a practice is some repeated and persistent behavior attributed to an instance of `FrameEntity` to the detriment of another instance of `FrameEntity`. A label is a characterization employed by some `FrameEntity` in describing another `FrameEntity`, and a forecast is a prediction that attributes blame to an instance of `FrameEntity` regarding the future state of some other `FrameEntity`. We define four subclasses of `FrameRelation` designed to capture these substantive elements – `Practice`, `Forecast`, `Label` and `Action` and a fifth class, `ConsistentFrameRelation` which is the superclass to all `FrameRelation` instances satisfying their constraints (more on this later). Each of `Practice`, `Forecast`, `Label` and `Action` has several subclasses, specific categories of each substantive frame element. `Forecast` `FrameRelations`, for example, fall into two types: `demiseOfProtectedGroup` and `victoryOverEnemy`. Figures 12 -15 illustrate the leaf nodes of the `FrameRelation` branch in TBO.

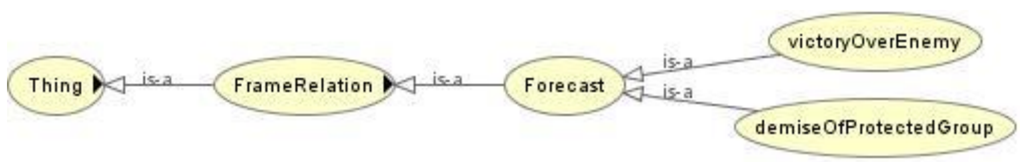

Figure 12 - `Forecast` `FrameRelation` in TBO

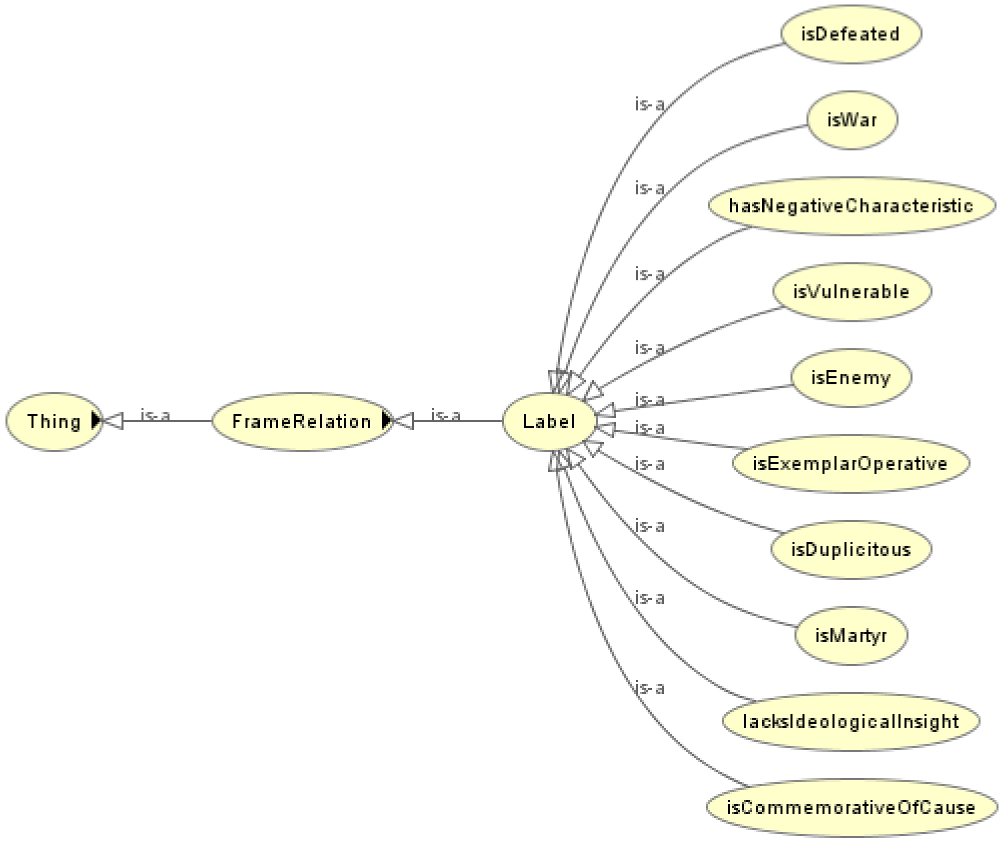

Figure 13 - `Label` `FrameRelation` in TBO

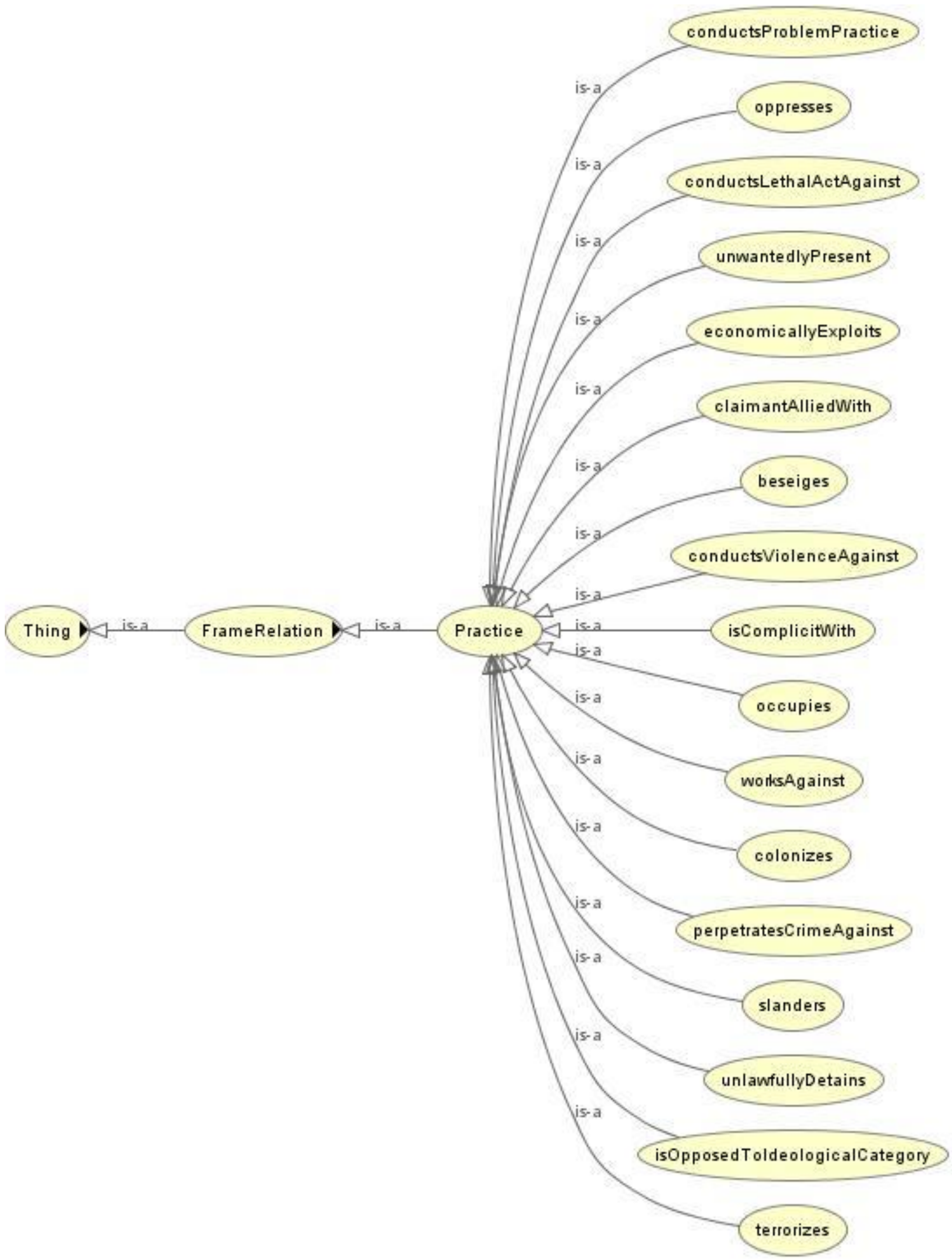

Figure 14 - `Practice` `FrameRelation` in TBO

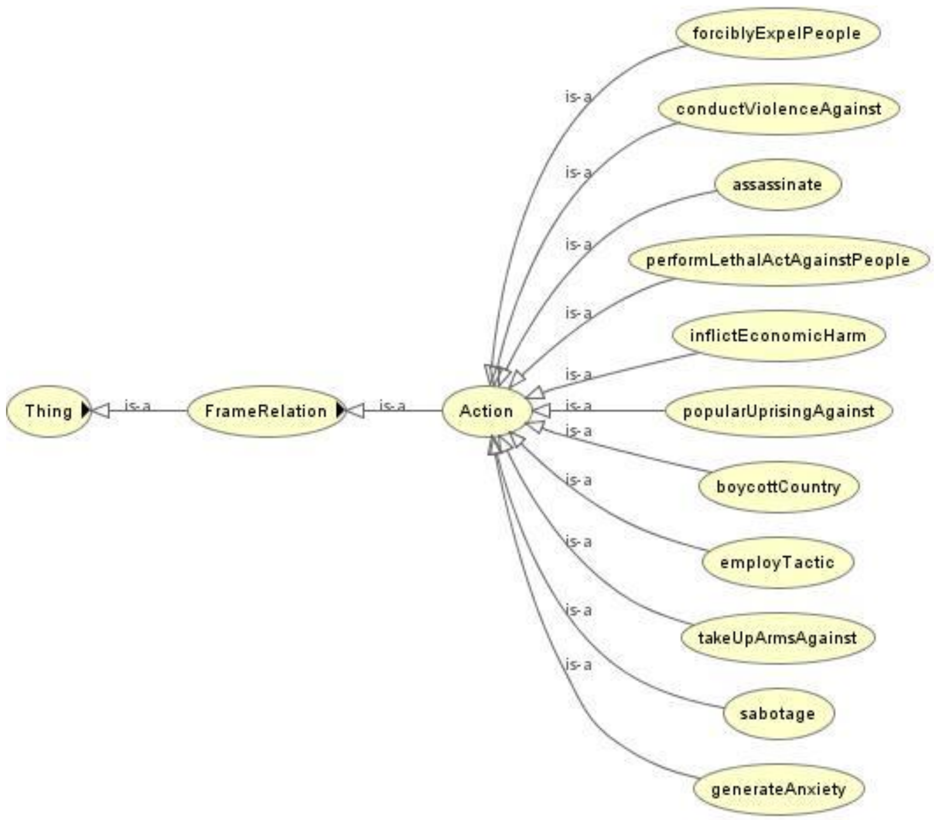

Figure 15 - `Action` `FrameRelation` in TBO

We make a notable design choice with respect to the representations in Figures 12 – 15. More specifically, the leaf nodes of the `FrameRelation` branch of TBO (that is, the subclasses of `Action`, `Practice`, `Label`, and `Forecast`) are defined as classes as opposed to relationships. OWL 2 DL provides constructs for specifying RBox axioms. In Figure 15, for example we define a class `takeUpArmsAgainst`. The class could have been represented as a role in the ontology, such that we would have, for example `takeUpArmsAgainst(Peoples, Peoples)`. Roles provide several modelling benefits. For example, with a role, it is easy to restrict the classes that are permissible as the first argument and the entities that are permissible in the second argument. Further, we could specify some helpful RBox axioms such as the transitivity axiom[18] (roles can have complex axioms such as symmetry, asymmetry, reflexivity, and irreflexivity (Krötzsch et al. 2012)). We have decided to forgo these RBox axiom benefits and instead opt for a TBox representation for a single critical reason. TBox axioms may have ABox axioms. In other words, it is essential for our representation to instantiate relationships such that instances of relationships may themselves become arguments in other relationships/roles. Next, we shall discuss the FrameEntity class in detail. After this discussion we shall return to this design choice.

The FrameEntity branch of TBO specifies the real world entities that may be mobilized by FrameRelation instances into meaningful frames. For example, take the FrameRelation takeUpArmsAgainst. The FrameRelation instances along their arguments define their meaning; `takeUpArmsAgainst(liberationists, vivisectors)` is something entirely different from `takeUpArmsAgainst(gaza_strip, Zionists)`. The former FrameRelation indicates a call to animal liberationists to engage in armed conflict against vivisectors, whereas the latter is a call to arms for the people of the Gaza Strip against Zionists. `liberationists`, `vivsectors`, and `zionists` are named individuals (instances) corresponding to the `Peoples FrameEntity`. Figure 16 below shows the hierarchical relationships of top-level elements, as denoted by ‘is-a’. Note that TBO has many other relationships among elements, such as hasDomain, has Range, hasFrame, etc. Tables 5 - 11 contain a full dictionary of all elements of the ontology and their definitions.

---

[18] $takeUpArmsAgainst \circ takeUpArmsAgainst \sqsubseteq takeUpArmsAgainst$ would be the representation for a transitive axiom

Figure 16 - `FrameEntity` Classes in TBO

Earlier, we discussed our design choice regarding the representation of `FrameRelation` as a collection of classes as opposed to a collection of roles. In Figure 17, we illustrate how the various components of this ontology come together to represent the semantics of terror organizations' collective action framing. It is the need for this representation that motivated our choice.

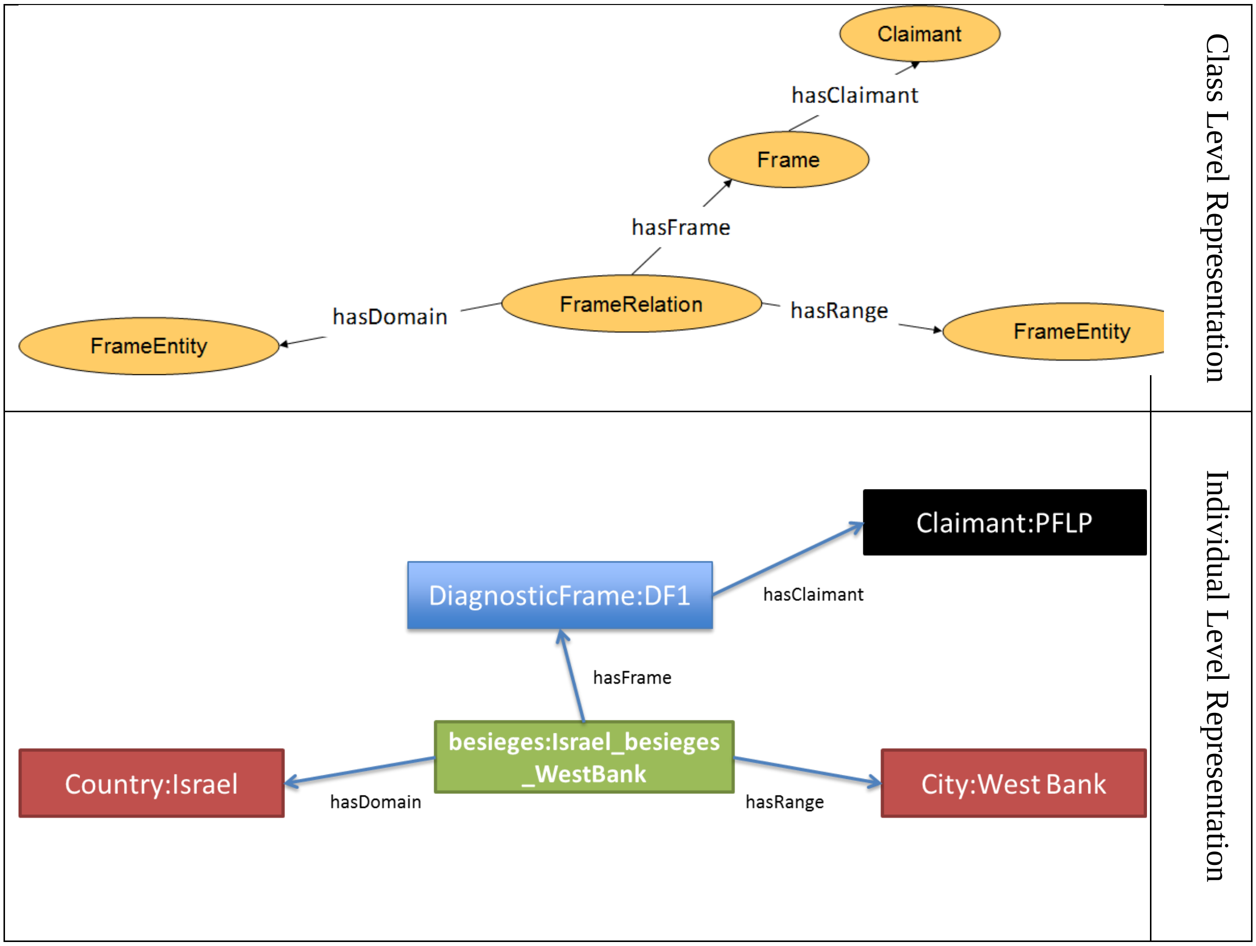

Figure 17 - Representing Terrorist Movement Frames

Our ontology is designed to work in the context of a framework for detecting the resonance of terrorist movement frames. Accordingly it was imperative that we index collected frames along their claimants. An instance of the `FrameRelation` class anchors the `FrameEntity` instances that mobilize a frame. This technique is roughly analogous to an association relation or associative entity in the relational model of data. It has been documented in a working group note of W3C and details can be found here https://www.w3.org/TR/swbp-n-aryRelations/. In Figure 17, the `FrameRelation` instance `Israel_besieges_WestBank` has three object property assertion axioms: `hasDomain, hasRange, and hasFrame`. Such axioms define associations between instances in the ontology. They are governed by RBox axioms (discussed earlier). The `hasDomain` object property points to (has a range) an instance of a class that makes up the first argument of the `FrameRelation`. The `hasRange` object property points to an instance of a class that makes up the second argument of the `FrameRelation`. Accordingly, we get `besieges(?domain, ?range)` such that we semantically render the representation in Figure 16 as: The Popular Front for the Liberation of Palestine (PFLP), the `Claimant`, issued a `DiagnosticFrame` stating that the `Country` Israel `besieges` the `City` West Bank.

The tables below give a full listing of all `FrameEntity`, and `FrameRelation` classes defined in TBO, including carefully chosen definitions for each class. The following section details our implementation of a Frame Discovery System, the system tasked with marshalling a group of trained classifiers towards the task of populating the ontology. Our Frame Discovery System does not cover all FrameRelations within the ontology. In particular, FDS does not cover Label FrameRelation classes, with the exception of the isEnemy class. We delve into the details of this and other frame detection design choices in the next section.

Subclasses of FrameEntity
Place - *A physical location*
NonHumanOrganism - *Animals, Plants and other such non-human entities that may be of concern to certain terrorist organizations*
ConceptualFrameEntity - *A category for entities not explicitly specified in this taxonomy*
Religion - *A system of faith and worship*

Organization - *A named entity that formally structures individuals towards a particular goal*
Bloc - *A bloc is some loosely-defined collection of countries, people or groups that share a common purpose but cannot be classified as a race of peoples, as an organization, nor a well defined place. E.g. "The West" is a bloc.*
Person - *A named human individual*
PoliticalOrientation - *A political orientation is a political philosophy*
IdeologicalCategory - *An instance of this class is an ideology not covered by a more specific FrameEntity. These are, typically distinctive doctrine's causes or theories. Words representing these ideological categories typically end in the suffix "-ism."*
SocialClass - *A social class is a designation given to a person or group of people regarding their socio-economic standing. E.g. "elite," "peasant," "middle income."*
HistoricalEvent - *A named event that has occurred in the past*
Civillians - *A specific categorization given to people who are non-combatants*
Combatants - *An instance of Combatants is a group of individuals who are not part of an official, State military but engage in combat*
People - *An instance of People is a people, delineated by ethnicity, religion, nationality, geographical location.*

Table 5. Subclasses of `FrameEntity`

Subclasses of Organization
LawEnforcement - *A civillian law enforcement organization*
Governmental - *A governmental organization or entity that does not fall into the category of InternationalGoverningBody, or LawEnforcement*
MediaOutlet - *A specific, named organization that disseminates news in any form of media on a mass scale e.g. CNN. The element can also be used to label a general reference to agglomerated media entities, e.g. "Western Media"*
Business - *A named organization constituted with the express purpose of making profit*
InternationalGoverningBody - A named international organization that can make pronouncements to which States must adhere.

SocialMovement - *A named, non-terroristic movement that seeks to bring about social change on some issue*
TerroristGroup - *A named organization that seeks to perpetrate violence with terroristic motive*

Table 6. Subclasses of `Organization`

Subclasses of Place
City - *A named geopolitical region that exists within a Country. Cities, states, towns, villiages are all examples.*
Continent - *One of a standard list of continents*
Country - *One of a standard list of countries*
Region - *A named designation of a geographic area that corresponds to a collection of countries, but is not a continent. E.g. West Africa, Arabian Peninsula, South East Asia.*

Table 7. Subclasses of `Place`

Subclasses of Person
PoliticalFigure - *A named person who occupies or has occupied political office*
TerrorOperative - *A named person who is a member of a terrorist organization or an adhereant to its ideology*

Table 8. Subclasses of `Person`

Subclasses of Practice
beseiges - *the notion of seige and barricade is recurrent especially in radical islamic rhetoric as a greivance perpetrated by an enemy FrameEntity upon a peoples, country, etc that the terror organization claims to be protecting*
colonizes - *a FrameEntity may be said to colonize another FrameEntity. The concept of colonization is evoked as grievance perpetrated by a malicious FrameEntity upon a protected frame enity*
conductsLethalActAgainst - *a FrameEntity may be accused of conducting a lethal act against another FrameEntity*
conductsViolenceAgainst - *a FrameEntity may be accused of perpetrating violence aginst another FrameEntity*
economicallyExploits - *a FrameEntity may be accused of plundering the wealth of another FrameEntity*

isComplicitWith - *this frame relation is manifest from the finding that many terrorist organizations accuse frame entities of being complicit with other frame entities*
occupies - *refers to the physical occupation of a place by some FrameEntity*
oppresses - *a FrameEntity may be said to oppress another FrameEntity*
slanders - *a FrameEntity can be said to slander another FrameEntity, in particular religious and political figures.*
terrorizes - *a FrameEntity can be accused of perpetrating acts of terrorism against another FrameEntity*
unlawfullyDetains - *a FrameEntity can be accused of unlawful detention of another FrameEntity, usually a people and in some cases non-human organisms.*
unwantedlyPresent - *a FrameEntity may be accused of being unwantedly present in a place.*

Table 9. Subclasses of `Practice`

Subclasses of Action
assassinate - *a call to assasinate a person FrameEntity*
boycottCountry - *a call to boycott a country FrameEntity*
forciblyExpelPeople - *a call for the forcible explulsion of a people*
generateAnxiety - *a call to adherents to generate anxiety in some FrameEntity. For example the animal liberation front calls for activists to generate anxiety for businesses that conduct tests on animals.*
inflictEconomicHarm - *a call to adherents to inflict economic harm on some FrameEntity*
performLethalActAgainstPeople - *a call to conduct an action with lethal intent targeted towars a people*
popularUprisingAgainst - *a general call for mass uprising against some FrameEntity. Protests can also fall into this category*
sabotage - *a call to carry out sabotage against some FrameEntity*
takeUpArmsAgainst - *a call for violent revolution*

Table 10. Subclasses of `Action`

Subclasses of Forecast
demiseOfProtectedGroup - *All terrorist organizations have a frame entity, or several different frame entities they profess to be protecting or advancing. Predicting the demise of these groups is typically used to spur action. For instance, the KKK frequently refers to a coming white genocide.*
victoryOverEnemy - *The target frame entity is any frame entity that has been identified as a terrorist organization's enemy and this FrameRelation indicates (usually exaggerated) optimism about the eventual defeat of the enemy*

Table 11. Subclasses of `Forecast`

3.7.2. Implementation of a Frame Discovery System

Given the above architecture of TBO, we define a series of classifiers operating within a pipeline architecture capable of populating TBO given a series of plain text inputs. This is the ultimate goal of FDS – populating TBO. Accordingly the classifiers we define for this system are tailored to the `FrameEntity` classes (NER classifiers) and the corresponding `FrameRelation` classes (Relation Extraction Classifiers). As we previously discussed we implement this information extraction framework with Stanford's CoreNLP API. We use its capabilities extensively, in particular, its API wrappers for featurization and classification. We were successful in defining classifiers for a subset of the concepts within the ontology. We gain good coverage over these concepts but faced some challenges in doing so.

In earlier sections, we described the relation extraction problem as follows: to estimate some conditional probability (conditional upon the features of the observed sequence) that a pair of NER annotations within a given sentence expresses a specified relationship. The key here is that relation extraction traditionally deals with pairs of NER annotations. However, subclasses of the `Label FrameRelation` are unary relations, such that, ideally, they have a single NER annotation as their argument. We experimented with work-arounds for this problem and implemented a novel approach that did not require changes to CoreNLP API. We experimented on the isEnemy relation and devised an anchoring technique to relation extraction on a unary relation. First we define a list of words denoting "enemy" that we noticed were typical in our corpus. We build a Regular Expression based NER classifier (citation to Stanford regexner) for these words. This classifier does not rely on machine learning and instead attempts to match tokens with the specified list of regular expressions.

```
851 O   0   O   CC  But O   O   O
851 O   1   O   VBZ is  O   O   O
851 Loc 2   O   NNP Israel  O   O   O
851 O   3   O   DT  the O   O   O
851 O   4   O   JJ  only    O   O   O
851 ENEMY/NEGATIVEWORD  5   O   NN  enemy   O   O   O
851 O   6   O   IN  that    O   O   O
851 O   7   O   PRP we  O   O   O
851 O   8   O   VBP are O   O   O
851 O   9   O   VBG facing  O   O   O
851 O   10  O   IN  in  O   O   O
851 O   11  O   DT  the O   O   O
851 O   12  O   NN  battle  O   O   O
851 O   13  O   .   ?   O   O   O

5   2   hasEnemy
```

Figure 18. Unary Relation Extraction Workaround

In Figure 17, we present a screen grab of our formatted training corpus for relation extraction. The above format is tab-delimited into columns. The first column is a sequence number associated with the sentence, indicating that the above is the 851st sentence in the corpus. The second column indicates the NER annotation associated with a token. If no annotation is available, the entry is "O." The third column indexes the sequence of tokens, such that the token at index 2 is the 3rd token in the sentence and has the NER annotation "Loc" which indicates a `Place`. Below the sentence block we have another tab-delimited structure. This block of text specifies that the hasEnemy relation (hasEnemy maps to `isEnemy`) has as its first argument the token at index 5 and as its second argument the token at index 2. We can use this binary relation to generate a unary relation in TBO: `isEnemy(Israel)`. The presence of the word "enemy" allows us to anchor the relation such that we do not violate the binary requirement. The ENEMY regular expression NER classifier comprises the regular expressions in Table 12. As it stands, it has a limited lexicon. However, we plan on expanding it in future work.

```
the occupation      ENEMY
occupier     ENEMY
enemy  ENEMY
enemies      ENEMY
traitor      ENEMY
adversary    ENEMY
adversaries  ENEMY
colonizer    ENEMY
kuff[^A-Za-z0-9]r   ENEMY
kufir  ENEMY
kuffir ENEMY
kuffar ENEMY
kaffir ENEMY
kafir  ENEMY
k[^A-Za-z0-9]fir    ENEMY
k[^A-Za-z0-9]ffir   ENEMY
mushrikin    ENEMY
mushrikeen   ENEMY
murtaddin    ENEMY
mushrik[^A-Za-z0-9] ENEMY
murtadd[^A-Za-z0-9] ENEMY
```

Table 12. ENEMY Regular Expressions

The ENEMY regular expression classifier is one of four regular expression NER classifiers defined within the FDS. The remaining three are: COMBAT, NEGATIVEWORD, and SALAFI. Each regular expression classifier, as with ENEMY, is a special case. Our preference is for probabilistic classifiers that take advantage of machine learning algorithms thus enabling the labelling of previously unobserved sequences during operation. We found that our linear chain CRF classifier could not reliably classify previously unseen instances of combatant. The basket of features (indicators) that we are able specify for a NER classify contains features of two broad kinds: (1) context-based features and (2) features based on candidate tokens. The first type of features captures aspects of the words surrounding a sequence being considered for classification. The second type is focused on the sequence itself.

We found that context-based features alone are not effective for NER classifications beyond the standard “organization” “person” and “location” labels. Adequate performance for labels outside of these categories requires the joint usage of both feature types. Even so, certain label categories will still be problematic. The “combatant” label is one such case. Even when we augment the classifier with gazetters, we still suffer unreliable performance. This is because features based on the candidate sequence are almost never fired for the “combatant” label. The tokens within universe of tokens that refer to “combatant” share little in common in and of themselves. Consider a training corpus where tokens “soldier,” “Warrior” “Air Force Captain” and “mujahideen” are all labelled as “combatant.” Word shape features on these tokens may take the form of “xxxxxxx” for “soldier,” “Xxxxxxx” for “Warrior,” “Xxx Xxxxx Xxxxxxx” for “Air Force Captain” and so on. Clearly no pattern is emerging. Character level n-grams on the candidate sequences won’t be very useful either. For example, if we specified n=3 for n-grams, we may get the following for the above: ior for Warrior, ier for soldier, and so on. Still, no pattern emerges. As the features on the candidate token offer little discriminative value, we augment the “combatant” statistical classifier with regular expression NER.

We also discovered that, in many cases, we needed to accommodate anglicized renderings of Arabic words prominent in Salafist Jihadi propaganda. As was alluded to in the previous paragraph, the performance of a statistical approach to NER is contingent on the ability of a feature space to capture patterns that are suggestive of NER labels. For anglicized renderings of Arabic words, the features specific to a candidate token or token sequence will be radically different from those we would expect in the English language.[19] This leads to increased misclassification. To circumvent this problem, we created the SALAFI regular expression NER classifier and were careful to not include anglicized renderings of Arabic words as positive training examples in any of our classifiers. The SALAFI regular expressions include any mention of a named entity for which a classifier is to be defined.

We defined a regular expression NER classifier, NEGATIVEWORD, as a collection of negative words defined in (Hu and Liu 2004). This regular expression classifier, similar to the others described above, also serves special purposes. First, for those negative words that are adjectives, we found the list in (Hu and Liu 2004) to be a useful source of negative characteristics ascribable to other named entities as asserted by radical propaganda. Figure 18 illustrates this.

[19] Indeed, this challenge is well known in the NER community and has spurred research streams on language independent NER. See the coNLL shared task on this topic for details (Tjong Kim Sang and De Meulder 2003)

```
856 Other/REL/peops 14  O   NNPS    Jews    O   O   O
856 O   15  O   IN  in  O   O   O
856 O   16  O   DT  all O   O   O
856 O   17  O   NNS parts   O   O   O
856 O   18  O   IN  of  O   O   O
856 O   19  O   DT  the O   O   O
856 O   20  O   NN  world   O   O   O
856 O   21  O   TO  to  O   O   O
856 O   22  O   VB  support O   O   O
856 Loc 23  O   NNP Israel  O   O   O
856 O   24  O   ,   ,   O   O   O
856 O   25  O   VBP protect O   O   O
856 O   26  O   PRP$    its O   O   O
856 NEGATIVEWORD    27  O   JJ  aggressive  O   O   O
856 O   28  O   NN  existence   O   O   O
856 O   29  O   CC  and O   O   O
856 O   30  O   VB  consolidate O   O   O
856 O   31  O   CC  and O   O   O
856 O   32  O   VB  expand  O   O   O
856 O   33  O   DT  this    O   O   O
856 O   34  O   NN  existence   O   O   O
856 O   35  O   .   .   O   O   O

0   14  isComplicit
0   23  isComplicit
14  23  isComplicit
23  27  hasNegativeCharacteristic
```

Figure 19. Illustration of the NEGATIVEWORD Annotation

The NEGATIVEWORD regular expression classifier labels the token (aggressive) at index 27 of the 856th sentence in the above training corpus as NEGATIVEWORD. The POS tag JJ is the Penn Tree Bank's tag for an adjective. The highlighted relation label "hasNegativeCharacteristic" is applied such that we assert that the Claimant accuses Israel of having the negative characteristic of aggression. The second application of the NEGATIVEWORD classifier is a feature for relation extraction. A critical feature for our relation extraction classifier is occurrence of an NER annotation along the dependency path between two candidate labels. For certain relations, we intuited that adding an explicit label for negative words would improve classification on the relation extraction task.

For the remainder of this subsection, we describe our various statistical classifiers paying special attention to the features selected as predictive of the classes. In Table 13, we describe the training and configuration of the classifier CIVILIANS+COMBATANTS. We developed a training corpus for this classifier based on the book *A World at Total War: Global Conflict and the Politics of Destruction 1937-1945*. The intuition supporting this choice is the expectation that the book would contain many examples corresponding both to "civilian" and "combatant" NER labels. This is the ultimate justification behind all our choices for training corpora.

Classifier Name:	CIVILIANS+COMBATANT
Supported Classifications:	civilian; combatant
Classifier Type:	Linear Chain Conditional Random Field NER
Training Document:	A World At Total War: Global Conflict and the Politics of Destruction (Chickering et al. 2005)
Feature List	*Description*
CRF Order 3	Means that features apply to at most the previous three tokens and the current token or at most the current token and the next three tokens
Class	Uses the supplied label class as a feature
Next Word	Uses the next word plus the class of the current word
Previous Word	Uses the previous word plus the class of the current word
n-Grams	We use character n-Grams of order 6, that is, n-Grams sequentially computed on characters present in tokens, 6 characters at a time. Combined with Next Word and Previous Word features, we also collect n-grams for the next and previous words as well as the current word.

Tags	Because previous word and next word are being used as features, this feature is the POS tag for the previous and next word
Word Tag	Use the POS tag of the current word as well as its class
Disjunctive Words	Disjunctives are words such as *but, or, either,* and *though* which express discord or differences between the words they are anchored on. This feature is fired within a window of 4 tokens around the observation
Class Sequences	This feature interacts with Next Word and Previous Word features to define features on the combination of the class labels of the previous word, the current word as the next word
Word Shape	Encodes a representation of the word's shape as given by the distribution of upper and lower case characters within the word
Gazette Features	We developed a gazette that includes specific examples of civilians and combatants. Gazette features can be drowned out if the CRF weights other features higher

Table 13. CIVILIANS+COMBATANTS Classifier Details

Classifier Name:	REL
Supported Classifications:	religion
Classifier Type:	Linear Chain Conditional Random Field NER
Training Document:	"The Global Religious Landscape: A report on the size and distribution of the World's major religious groups as of 2010" (Hackett and Grim 2012)
Feature List	*Description*
CRF Order 3	Means that features apply to at most the previous three tokens and the current token or at most the current token and the next three tokens

Class	Uses the supplied label class as a feature
Next Word	Uses the next word plus the class of the current word
Previous Word	Uses the previous word plus the class of the current word
n-Grams	We use character n-Grams of order 6, that is, n-Grams sequentially computed on characters present in tokens, 6 characters at a time. Combined with Next Word and Previous Word features, we also collect n-grams for the next and previous words as well as the current word.
Tags	Because previous word and next word are being used as features, this feature is the POS tag for the previous and next word
Word Tag	Use the POS tag of the current word as well as its class
Disjunctive Words	Disjunctives are words such as *but, or, either,* and *though* which express discord or differences between the words they are anchored on. This feature is fired within a window of 4 tokens around the observation
Class Sequences	This feature interacts with Next Word and Previous Word features to define features on the combination of the class labels of the previous word, the current word as the next word
Word Shape	Encodes a representation of the word's shape as given by the distribution of upper and lower case characters

Table 14. REL Classifier Details

Classifier Name:	IDE
Supported Classifications:	ideology
Classifier Type:	Linear Chain Conditional Random Field NER
Training Document:	A World at Total War: Global Conflict and the Politics of Destruction
Feature List	*Description*

CRF Order 2	Means that features apply to at most the previous two tokens and the current token or at most the current token and the next two tokens
Class	Uses the supplied label class as a feature
n-Grams	We use character n-Grams of order 6, that is, n-Grams sequentially computed on characters present in tokens, 6 characters at a time. Combined with Next Word and Previous Word features, we also collect n-grams for the next and previous words as well as the current word.
Tags	Because previous word and next word are being used as features, this feature is the POS tag for the previous and next word
Word Tag	Use the POS tag of the current word as well as its class
Word Shape	Encodes a representation of the word's shape as given by the distribution of upper and lower case characters

Table 15. IDE Classifier Details

Classifier Name:	PEOPS+PP+TACTIC
Supported Classifications:	peoples; problempractice(PP); tactic
Classifier Type:	Linear Chain Conditional Random Field NER
Training Document:	Manifesto for Radical Abolitionism The Coming White Genocide (KKK) Strategy for the Liberation of Palestine
Feature List	*Description*
CRF Order 3	Means that features apply to at most the previous three tokens and the current token or at most the current token and the next three tokens
Class	Uses the supplied label class as a feature
Next Word	Uses the next word plus the class of the current word

Previous Word	Uses the previous word plus the class of the current word
n-Grams	We use character n-Grams of order 6, that is, n-Grams sequentially computed on characters present in tokens, 6 characters at a time. Combined with Next Word and Previous Word features, we also collect n-grams for the next and previous words as well as the current word.
Tags	Because previous word and next word are being used as features, this feature is the POS tag for the previous and next word
Word Tag	Use the POS tag of the current word as well as its class
Disjunctive Words	Disjunctives are words such as *but, or, either,* and *though* which express discord or differences between the words they are anchored on. This feature is fired within a window of 4 tokens around the observation
Class Sequences	This feature interacts with Next Word and Previous Word features to define features on the combination of the class labels of the previous word, the current word as the next word
Word Shape	Encodes a representation of the word's shape as given by the distribution of upper and lower case characters

Table16. PEOPS+PP+TACTIC Classifier Details

The feature sets across the different NER classifiers depicted above are largely the same with the exception of the IDE classifier. This classifier is distinct because of the unique importance of word shape and n-Grams on the observed word for classification. CRFs work by assigning weights to features. Accordingly, it is essential to maintain a parsimonious set of discriminative features. We do not want a situation where non-discriminative features are inadvertently weighted such that they crowd out the influence of more discriminative features. This is especially crucial given that we make use of rather small training corpora. Ideologies typically have similar word endings (e.g. “ism,” “ite,” “ist” and so on) accordingly we wanted to give the CRF the best possible chance of weighting these features highly (these would be captured by n-gram features).

With respect to relation extraction, we train a single multi-class logistic classifier using the same features shown in (Surdeanu et al. 2011). Table 17 summarizes the classes of relations for which the model is trained.

IsComplicit
isOpposedToIdeologicalCategory
Besieges
Colonizes
conductsLethalActAgainst
conductsViolenceAgainst
economicallyExploits
occupies
Oppresses
perpetratesCrimeAgainst
Slanders
Terrorizes
unlawfullyDetains
unwantedlyPresent
worksAgainst
hasEnemy
worksFor
victoryOverEnemy
popularUprisingAgainst
takeUpArmsAgainst

employTactic
performLethalActAgainstPeople
hasNegativeCharacteristic

Table 17. Relation Extraction Classes

3.7.3. Implementation of Frame Resonance Detection System

FRDS is a user facing integrating system which is principally tasked with determining when some input text produces resonating frames. We develop an algorithm for computing frame resonance scores and based on our findings from analyzing several online data sources, we propose methods to interpret these scores as they pertain to the three design outputs.

There are two document collections required to initiate FRDS. These are its inputs. The first document collection is the baseline that FRDS uses to generate a baseline expectation regarding the framings issued by a radical organization. This document collection can be made up of speeches issued by prominent spokespeople of the organization, manifestos and other textual publications of the organization as well as websites managed by the organization. Ideally, any textual document produced by the terrorist movement under investigation works as an input to this collection. The second document collection is comprised of text obtained from sources of concern. Sources of concern are online personas of interest to the analyst. Users of a forum, for example, or a manifesto posted online by a troubled individual. Given these two document collections, FRDS can automatically determine whether the frames espoused by the organization of concern are resonating with the online personas of concern. FRDS also enables the analyst to take deep dive, so to speak, into the frames produced by either the organization of concern or by any of the individuals of concern. The analyst can query these frames using a structured language and gain insight into the actors and parties that are relevant to the organization of concern.

For each document collection, FRDS invokes FDS which extracts frames and writes them to TBO. TBO indexes frames by claimant so the system is aware to which persona or movement a frame belongs. Once FDS has traversed the document space, FRDS detects resonance. We developed a resonance scoring technique that provides strikingly accurate and discriminative results. The remainder of this section examines this scoring technique.

Looking back at our representation of terrorist movement frames depicted in Figure 16, each instance of a frame is characterized by a domain `FrameEntity`, a `FrameRelation`, and a range `FrameEntity`. The distinction between domains and ranges is critical. The representation in illustrated in Figure 16 makes this clear. A domain `FrameEntity` can be thought of as belonging on the left side of a `FrameRelation` and a range `FrameEntity` can be thought of as belonging on the right side of a `FrameRelation`. In Figure 16, we represented the PFLP framing of Israel besieging the West Bank as follows: `besieges(Israel, West_Bank)`, where the first argument to the `FrameRelation` besieges is the domain and the second argument the range. The inverse would be tantamount to an entirely different frame, one more likely to be emitted by the Israeli Defense Force than the PFLP.

We require some technique that can rapidly compare collections of frames, providing some numerical score that captures resonance. Say, as an example, we are examining the resonance of Al-Qaeda in the Arabian Peninsula (AQAP) frames with respect to a forum user. We have created representations in TBO for the frames discovered by FDS on AQAP magazines, speeches and other propaganda texts. We have done the same for the forum user. Say we look at any random `FrameRelation – performLethalActAgainst`, for example. For AQAP, there will exist some collection of `FrameEntity` instances on the domain of this relation and some collection of `FrameEntity` instances on the range of the relation. The same may be true for the forum user. Our goal is to assess the similarity between AQAP framing and the individual's narrative along the `performLethalActAgainst` relation. We construct a vector on the ranges of this relation for AQAP. We similarly construct a vector on the ranges of the relation for the forum user's narratives. We do the same on the ranges. We project these representations onto vector space and take the cosine of the resultant vectors to assess their similarity. We do this for the domain vectors and range vectors separately. We subsequently combine these scores using a weighted average. We use cosine similarity because it is apathetic to vector size. We expect that (and we indeed do) we will compare vectors of radically different sizes and cosine similarity offers a handy solution.

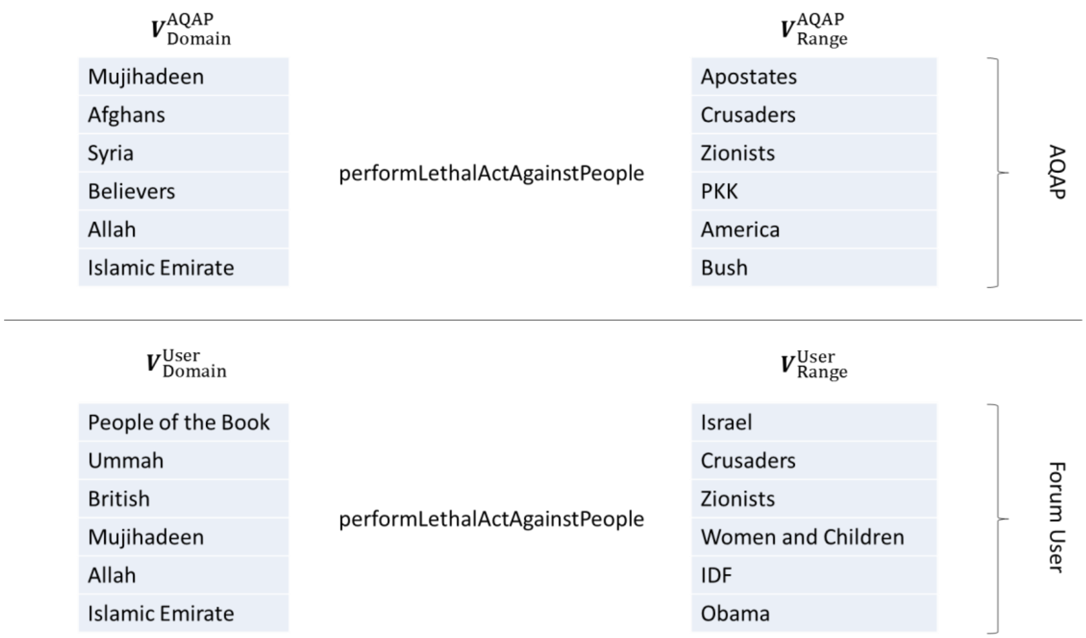

Figure 20. Calculating Resonance Scores on a `FrameRelation`

Using the visualization in Figure 18, say we have the above vectors domain and range vectors for AQAP and for some forum user. We define the range and domain vectors on some relation as $\boldsymbol{V}_j^i$

where i indexes the `Claimant` and $j \in \{domain, range\}$. For two such vectors $\boldsymbol{V}_j^i$ and $\boldsymbol{V}_j^{i\prime}$ where $i \neq i'$, we first convert the vectors to vector space using term frequency vectors. These term frequency vectors are defined using character n-grams where n = 2 for quick processing. Once we successfully project the two vectors into vector space, we then compute cosine similarity:

$$\cos(\theta) = \frac{\boldsymbol{V}_j^i \cdot \boldsymbol{V}_j^{i\prime}}{\|\boldsymbol{V}_j^i\| \| \boldsymbol{V}_j^{i\prime}\|}$$

Given the above, $\text{similarity}(\boldsymbol{V}_j^i, \boldsymbol{V}_j^{i\prime})$ results in a number between 0 and 1, inclusive of both bounds. If $\text{similarity}\left(\boldsymbol{V}_j^i, \boldsymbol{V}_j^{i'}\right) = 1$ then we say that the vectors are perfectly similar. If $\text{similarity}\left(\boldsymbol{V}_j^i, \boldsymbol{V}_j^{i'}\right) = 0$ then we say that the vectors are perfectly dissimilar. Values between these indicate varying levels of similarity/dissimilarity. Of course, the above computation only accounts for one of range or domain vectors. For any pair of `FrameRelation` instances that we wish to compare, we have **both** range and domain vectors. Say we compute cosine similarity on the range vectors and we compute cosine similarity on the domain vectors, we are left with two similarity figures (it turns out that over a corpus, these metrics are highly correlated) we need a way to harmonize the two figures. To this end we define constants in FRDS that assign intuition driven weights[20] to ranges and domains of each `FrameRelation` enabling a weighted average to cosine similarity to estimate the resonance of a framing category. The intuition behind the weights is based on which vectors we believe contain higher information content. For example, the `hasEnemy` relation has high informational content (variance) on the range vector

[20] Our ongoing work will use a formula for this. The formula will be based on the correlation between similarity scores on range and domain vectors for a `FrameRelation`. The formula will penalize lower correlations in determining weights such that the lower the correlation between two vectors, the more asymmetric the weights.

and low informational content on the domain vector (recalling that all domains of a hasEnemy relation are instances of the narrowly defined Enemy FrameEntity). Note that in cases where our intuition is inaccurate, we do not suffer much performance loss. Indeed, our experiments indicate that we can obtain highly discriminative performance when we only consider either range or domain vector similarity across all relation categories.

Accordingly, we define the resonance of a frame category (where frame categories are delineated by `FrameRelation` instances) as the weighted average of the range and domain cosine similarities on that category for a baseline and an online persona of interest. This computation results in a vector resonance scores along FrameRelation categories for a given persona of interest and a given terror movement baseline. The vector, alone, does not reveal much. We require some data driven approach to discriminate between vectors that indicate strong terror movement frame resonance on a persona and weak or no resonance. We use a combination of k-means clustering and discriminant analysis on a large corpus to estimate coefficients for each frame category. The coefficients work in a manner similar to beta coefficients in linear regression. Extending the metaphor, frame categories serve as predictors (independent variables) of a binary response variable (0, 1). When the sum of resonance/coefficient products for an observation exceed 0.5, we say that the observation (persona of concern) demonstrates resonance with baseline terror movement. We found that these coefficients are **not** universal such that a set of coefficients developed from an AQAP baseline will fail to predict resonance, on the same data set, with an ISIS baseline. Accordingly, a unique set of coefficients must be developed for each terror movement baseline ontology. Given a set of positive and negative examples of online personas, FRDS can estimate these coefficients.

Coefficient estimation begins with K-Means clustering, an approach that allows us to pre-specify the number of clusters. We found that $k = 3$ generally produced good results. The cluster with the highest centroids is the cluster which contains radicalized observations, that is, personas for whom the baseline movement's frames resonate. Given this 3-cluster representation, we label each observation with its corresponding cluster. We then recode the clusters into a binary response variable such that the bottom 2 clusters are labeled in the negative (0) and the top cluster is labelled in the affirmative (1). We then fit a discriminant function to the result and thus obtain coefficients. Because the coefficients are based on clusters computed on observed variance, the discriminant function is over-fitted to the baseline. This is the reason why our coefficients are tied to the baseline.

Chapter 4 – Evaluation

4.1 Introduction

In the previous chapter, we described a system that automatically generates knowledge representations of terrorist movement framings. This system is also designed to use those representations to detect the resonance of terrorist movement frames with respect to online personas, websites, and news articles. We argued in the previous chapters that the automatic detection of frame resonance on the Web is a critical counter-terrorism requirement towards mitigating the threat of lone wolf terrorism. We argued further that detecting frame resonance with respect to individuals makes an accurate search for radicalizing influences on large corpora of social media data possible; we noted that the frame resonance detection techniques used to that end is also capable of detecting websites with radical content as well as identifying news reports that may represent potential triggering events. In this chapter, we demonstrate the efficacy of the system towards these goals. We evaluate this design artifact along two dimensions. First, we provide measures of precision and recall with respect to the many classification subtasks undergirding FDS. Second we shed light on the accuracy with which the implemented system can detect frame resonance from textual data sources such as forum discussions and news articles and crawled websites. We also present examples of how TBO may be used to analyze news articles and terror movement frames to identify potentially triggering events.

This chapter is structured as follows. We begin by presenting descriptive statistics on the corpora used to train the classifiers in FDS. We follow this discussion by describing the performance of each classifier used in FDS. Next, we describe the textual data sources with which we evaluate our ability to detect the resonance of terrorist movement frames. This analysis will demonstrate that we can define accurate classifiers for frame resonance detection and that the results of the analysis are dependent on the terror movement baseline. We subsequently demonstrate the usefulness of TBO by running a series of queries against an exemplar instance of the knowledge base that has been populated by FDS using our AQAP and ISIS corpus. This demonstration also serves to illustrate how TBO can be used to find potentially triggering events in news corpora. Finally, we demonstrate the coordinating framework used to martial our three core components.

4.2 Performance of FDS Classifiers

Relation extraction systems, and this is true for NER, are typically trained on large standardized corpora desinged for use on shared tasks at events such as coNLL and the Machine Understanding Conference (MUC). However, our domain of interest is highly specific, requiring us to hand select training documents that maximize our opportunity for tagging positive examples of our relation categories and NER labels. There is some overlap in the documents we use for NER and for relation extraction, but very little. With respect to NER, our entity categories were not quite as specialized as the relation extraction categories, meaning that we were afforded the opportunity to train NER classifiers on varied documents. Relation extraction training and evaluation are done on the same collection of documents using k-fold = 5 cross validation at training time.

As was noted in Chapter 3, we trained a single Linear Chain Conditional Random Field model to extract mentions of religions in text. Our training corpus (see Chapter 3) included 335 mentions of religion out of 8,881 words used for training. We evaluated the model on a corpus of 11,152 words based on text extracted from Rumiyah Issue 6 and text scraped from the PFLP's website. Note that all NER classifiers are evaluated on the same corpus. The evaluation results are shown in table 18. Precision and recall values are used to summarize the results. Precision is a measure of the proportion of classifications made by a classifier that are correct. Recall is a measure of the proportion of positive observations labelled by a classifier to the total number of positive observations in the gold standard. The harmonic mean of both measures is known as F1, a widely accepted, balanced measure of the overall performance of a classifier.

Label (tag)	**Precision**	**Recall**	**F1**
Religion (REL)	81.16%	96.55%	88.19%

Table 18. Religion Classifier Performance

To develop a classifier of ideology mentions in text, we needed a document source rife with such mentions. During our qualitative analysis of terrorist propaganda, we noted that ALF documents were a particularly rich source of propaganda mentions. Accordingly, we included the Manifesto for Radical Abolitionism in our corpus. We augmented the manifesto with snippets from the book, *The Politics of Destruction* (Chickering et al. 2005). This book is a historical account of the political and ideological forces that contributed to the two World Wars. Accordingly, we intuited that it would be an excellent source of examples of ideology mentions in natural language. The resultant corpus contained 11,444 words, of which we tagged 340 as ideology mentions. We evaluated the resultant classifier on our NER evaluation corpus, with results shown in Table 19.

Label (tag)	**Precision**	**Recall**	**F1**
Ideology (IDE)	74.07%	51.28%	60.61%

Table 19. Ideology NER Classifier

We developed a single classifier to extract mentions of: Peoples, tactics and problem practices. All three categories are related. We anticipated that a text with ample mentions of resistance tactics and enemy problem practices would also have ample mentions of peoples (groups of people linked under a single category, e.g. fascists, vivisectors, crusaders and so on). We used text extracted from the ALFs Manifesto for Radical Abolition, the PFLPs Strategy for the Liberation of Palestine and the KKKs Coming White Genocide. We used all three of these documents as part of our document collection during the qualitative analysis. Accordingly, at classifier training time, we were aware that these three documents represent a rich source of examples for this classifier. The KKK document, for example, talks about how peoples of non-white racial backgrounds are orchestrating the demise of the white race. Certainly, the authors identified peoples, tactics and problematic practices in making their case. The resulting corpus contains 21,140 words. We identified 701 examples of peoples, 152 examples of tactics, and 189 examples of problematic practices. This is a challenging classifier to model. The main reason is that NER is designed to extract mentions of *rigid designators* (Kripke 1972), a concept we discussed extensively in Chapter 3. *Some* categories of peoples are identified by rigid designators, but not all. For example, “Muslims” is a rigid designator for a people of a certain faith, however, “apostates” is not. A good indicator that a word is not a rigid designator is whether or not it is a proper noun. Many of the entity mentions we trained this classifier on were not rigid designators, and accordingly, performance suffers. However, the evaluation results shown in Table 20 are not fatal, far from it. We have used FDS to successfully extract hundreds of mentions of these categories from text. However, the document space was massive, spanning millions of words. We are limited to rather small evaluation corpora that result in overly favorable or overly unfavorable performance figures. In Table 20, N/A means that we were unable to extract any examples of problem practices from our evaluation corpus. Indeed, manually labelling the corpus we only found 5 mentions of problem practices.

Label (tag)	Precision	Recall	F1
Peoples (peops)	74.19%	32.39%	45.10%
Tactics (tactic)	50.00%	14.24%	22.22%
Problem Practices (PP)	N/A	N/A	N/A

Table 20. Peoples, Tactics and Problem Practices Classifier

The results in table 21 with respect to the civilians label must be taken with caution. Only 9 mentions were classified with this label in the evaluation corpus. All 9 were rather easy targets for the classifier. We trained this classifier using Chickering's book on the politics of destruction. The classifier is also aided by gazetter features.

Label (tag)	Precision	Recall	F1
Civilians (civilian)	100%	50%	66.67%
Combatant (combatant)	87.50%	38.89%	53.85%

Table 21. Civilians and Combatant Classifier

Below are the results from Stanford's 4-class CRF classifier for Location, Miscellaneous, Organization and Person. Stanford's NLP team trained this classifier on a MUC corpus designed for that purpose. The feature set underlying the classifier is, however, proprietary. Again, our evaluation of this multi-class classifier is based on an issue of the Islamic State's Rumiyah magazine, as well as examples from PFLP propaganda. Locations, organizations and persons mentioned in the text have very distinct word shape features attributable to anglicized spellings of Arabic names. For example, an Arabic name such as al-Bara Ibn Azib has a markedly different word shape than, say John Smith. This issue is notable in names of places, locations and organizations as well. In addition, our evaluation data is extracted from PDF files for a glossy magazine. Article and section titles in the magazine generate features that severely hamper extraction of the organization label as capitalization features are heavily weighted. The miscellaneous category performs acceptably well. We also find the Location category to be acceptable.

Label (tag)	**Precision**	**Recall**	**F1**
Location (Loc)	54%	70.59%	61.28%
Miscellaneous (Other)	58.41%	91.67%	71.35%
Organization (Org)	30.16%	28.79%	27.97%
Person (Peop)	18.86%	54.10%	27.97%

Table 22. Stanford's 4 Class Location, Miscellaneous, Person and Organization Classifier

Label	Correct	Predict	Actual	Precn	Recall	F			
_NR				22820.0	23076.0	22834.0	98.9	99.9	99.4
besieges				2.0	2.0	2.0	100.0	100.0	100.0
colonizes				4.0	4.0	8.0	100.0	50.0	66.7
conductsLethalActAgainst				0.0	0.0	2.0	0.0	0.0	0.0
conductsLethalActiAgainst				0.0	0.0	1.0	0.0	0.0	0.0
conductsProblemPractice				1.0	1.0	1.0	100.0	100.0	100.0
conductsViolenceAgainst				5.0	6.0	9.0	83.3	55.6	66.7
economicallyExploits				17.0	30.0	31.0	56.7	54.8	55.7
employTactic				38.0	42.0	49.0	90.5	77.6	83.5
generateAnxiety				0.0	0.0	2.0	0.0	0.0	0.0
hasEnemy				29.0	33.0	34.0	87.9	85.3	86.6
hasNegativeCharacteristic				12.0	12.0	22.0	100.0	54.5	70.6
hasNegativeIdeology				0.0	0.0	2.0	0.0	0.0	0.0
hasProtectedGroup				1.0	1.0	1.0	100.0	100.0	100.0
isComplicit				142.0	150.0	241.0	94.7	58.9	72.6
isComplict				0.0	0.0	2.0	0.0	0.0	0.0
isComplidit				0.0	0.0	1.0	0.0	0.0	0.0
isOpposedToIdeologicalCategory				3.0	3.0	8.0	100.0	37.5	54.5
lacksIdeologicalInsight				3.0	3.0	4.0	100.0	75.0	85.7
occupies				19.0	22.0	24.0	86.4	79.2	82.6
oppresses				3.0	3.0	15.0	100.0	20.0	33.3
performLethalActAgainstPeople				33.0	40.0	46.0	82.5	71.7	76.7
perpetratesCrimeAgainst				0.0	0.0	1.0	0.0	0.0	0.0
popularUprisingAgainst				42.0	47.0	56.0	89.4	75.0	81.6
sabotage				1.0	1.0	3.0	100.0	33.3	50.0
slanders				2.0	3.0	12.0	66.7	16.7	26.7
socialist				0.0	0.0	1.0	0.0	0.0	0.0
takeUpArmsAgainst				85.0	101.0	121.0	84.2	70.2	76.6
takeUpArmsAgaint				0.0	0.0	1.0	0.0	0.0	0.0
terrorizes				2.0	2.0	3.0	100.0	66.7	80.0
unlawfullyDetain				0.0	0.0	1.0	0.0	0.0	0.0
unlawfullyDetains				0.0	0.0	2.0	0.0	0.0	0.0
unwantedlyPresent				6.0	6.0	14.0	100.0	42.9	60.0
victoryOverEnemy				27.0	30.0	34.0	90.0	79.4	84.4
worksAgainst				60.0	74.0	97.0	81.1	61.9	70.2
worksFor				22.0	22.0	29.0	100.0	75.9	86.3
Total	559.0	638.0	880.0	87.6	63.5	73.6			

Figure 21. Evaluation Results for Relation Extraction Classifier

Testing the performance of our relation extraction classifier proved to be challenging due to the specialized relation categories we trained. To train the relation extraction algorithm, we manually labelled a corpus of about 75,000 words, spanning 2,315 sentences. The corpus is made up of documents collected from the ALF website (including the manifesto), Strategy for the Liberation of Palestine, documents collected from the KKK website, and two issues of AQAP's Inspire magazine. We needed this larger corpus to increase our odds of finding examples of our specialized relations. We wanted to ensure that the multi-class L2 logistic regression we used for classification would be robust against NER classification errors. Accordingly, we labelled the training corpus using our NER classifiers prior to manually annotating relation examples. In some cases, we used manually added NER labels where our NER classifiers failed to do so. We did this especially for relation categories that did not occur frequently within the text. As we are limited in key resources, we do not develop a separate testing corpus for relation extraction. Instead, we use k-fold cross validation at training-time. Setting the number of folds to 5 meant that our training corpus is divided into 5 subsamples. One subsample is reserved for validation and the other four are used to train the model. Once the model is trained it is validated and results of that validation are stored. Another subsample is selected, and the other four are used to train the classifier. Validation is performed on the second selected subsample. This is done for all 5 folds. Validation results from each fold are averaged to produce the final result illustrated in Figure 21. This is the raw output of our validation script. Note that while labelling the corpus we occasionally misspelled relation names which is why we have two rows conductsLethalActiAgainst and conductsLethalActAgainst. These errors are trivial. Do note however that these performance metrics are likely to be slightly inflated. K-fold cross validation does not test the performance of a classifier beyond the domain of the data used to train it. As such, we suspect that classifier performance will suffer in radically different domains. Still, these performance metrics show excellent performance. We note that the classifiers appear to favor precision over recall.

4.3 Frame Resonance Detection Evaluation

We posed the following research questions: (2) *How may websites that promote radical ideology be automatically detected?* (4) *How may the knowledge representation populated in RQ3 above be used to automatically detect the ideology of an online persona?* This portion of the evaluation examines the extent to which our design can answer those questions. Both research questions 2 and 4 can be resolved using FRDS in concert with FDS and TBO. TBO, as always, represents the information extracted by FDS, lending it structure. FRDS subsequently makes inferences based on the populated TBO. We conduct several investigations to determine how well our system can identify radical content, the essence of research questions 2 and 4. To do this, we relied extensively on data provided by the Artificial Intelligence Laboratory at the University of Arizona. More specifically, we make use of their "dark-web forum dumps," a repository of scraped discussion threads from special interest Web forums. We also relied on other public data sources for source propaganda material produced by the Animal Liberation Front (ALF), the Popular Front for the Liberation of Palestine (PFLP), Al-Qaeda in the Arabian Peninsula (AQAP), and the Islamic State (ISIS).

Organization	Data Sources for TBO Baseline
Animal Liberation Front[21]	• Manifesto for Radical Abolitionism: By Any Means Necessary • Memories of Freedom • Our Next Moral Challenge • The ALF Mission Statement *Source:* http://www.animalliberationfront.com/
Al-Qaeda in the Arabian Peninsula	• Inspire Magazine Issue 1 (Jan 2010) • Inspire Magazine Issue 2 (Oct 2010) • Inspire Magazine Issue 3 (Nov 2010) • Inspire Magazine Issue 8 (May 2012) • Inspire Magazine Issue 9 (May 2012) • Inspire Magazine Issue 12 (March 2014)

[21] Material is not properly dated.

	• Osama Bin-Laden Tape Transcripts (~2004) *Sources*: http://www.aljazeera.com/archive/2006/09/200841012647537920.html http://jihadology.net/category/inspire-magazine/
The Islamic State (ISIS)	• Dabiq Issue 14 (Apr 2016) • Dabiq Issue 1 (July 2014) • Dabiq Issue 10 (Jul 2015) • Dabiq Issue 8 (Mar 2015) • Rumiyah Issue 1 (Sep 2016) *Source*: https://clarionproject.org/islamic-state-isis-isil-propaganda-magazine-dabiq-50/
Popular Front for the Liberation of Palestine	• Strategy for the Liberation of Palestine (1969) • Collection of PFLP Press Releases (~2016) *Source*: http://pflp.ps

Table 23. Terror Organizations and their Propaganda

The source materials listed in Table 23 are preprocessed and classified using FDS. As described in Chapter 3, classifier results are subsequently read into TBO where the resultant frames are indexed by claimant. This process is repeated for a collection of textual data comprised of forums and an entire dump of the PFLP's Website. Table 24 provides descriptive statistics on the forum data provided by University of Arizona's Artificial Intelligence Laboratory.

Forum Name (Period of Collection):	ANSAR (12/8/2008 – 1/20/2010)	MYIWC (11/5/2000 – 2/19/2010)
# of Posts:	29,492	25,016
# of Threads:	11,244	6,310
# of Members (Personas):	382	756
Description:	This forum is used to distribute news regarding the progress of Islamic State terror operations around the world. By all indications it is used *exclusively* for this purpose.	A well-intentioned forum for Muslims around the world to learn about their faith from one and other. However, conversations often become political with some individual overtly supporting Salafist ideologies, violent jihad, and martyrdom operations.

Table 24. University of Arizona AI Lab Forum Dump Descriptions (Artificial Intelligence Lab Management Information Systems Department University of Arizona 2010)

With over a thousand personas and about 54,000 individual posts between them, the two Web forums enable us rigorously examine the predictive and discriminative ability of our approach. We design an investigation that systemically examines how TBO representations of content can be used to (1) identify content that resonates terror organization frames, and (2) identify the terrorist ideology *most likely* resonated by a collection of frames. We run FRDS using TBO frames generated from the baseline against TBO frames generated from the forums. We are interested in identifying radical personas. Accordingly, we group forum posts by forum member. Within TBO, a `Claimant` instance is generated for each forum member so that extracted frames can be associated with their owning `Claimant` instance. For each Claimant instance of concern, (here, `Claimant` instances are forum members) we generate a Frame Resonance Profile (FRP), a vector of weighted cosine similarity scores calculated using the method presented in Chapter 3. For a given forum member, an FRP is created for each baseline frame collection in TBO such that, in this investigation, each forum member has 4 FRP vectors, one for each of ALF, PFLP, AQAP, and ISIS. Each frame resonance profile is uniquely identified by (1) a terror organization the frames of which are the baseline for its vector of scores and (2) the document of concern (say, for example, a collection of social media posts written by an online persona), such that these two documents functionally determine the values of the FRP.

For our method to reliably identify terror organization frames and terrorist ideologies most likely resonated by a collection of frames, the method should, first, be sensitive to the nature of a terror movement's ideology. That is, given a document of concern, and given its FRPs on two terrorist ideologies, say *Ideology A* and *Ideology B*, we should expect that:

1. If *Ideology A* and *Ideology B* are similar, then the FRPs generated on these two ideologies with respect to the document of concern should also be similar. This is a stringent test. We could determine that two ideologies represented in TBO are similar by running FRDS with one ideology as a baseline and the other as a representation of concern. However, this approach would only demonstrate that the content of our *representations* of the two ideologies are similar. We want to go further than this by simultaneously examining the sensitivity of represented content to the substance of an ideology and the sensitivity of resonance scores on a document of concern to the affinity a text displays to the baseline. Two similar ideologies must produce two similar FRPs on the same document of concern.
2. If *Ideology A* and *Ideology B* are dissimilar, then the FRPs generated on these two ideologies with respect to the document of concern should be dissimilar. This requirement is the other side of the coin. We could certainly directly compare the two ideologies, however we would have no insight regarding the ability of differing representations to produce differing FRPs on the same document of concern.

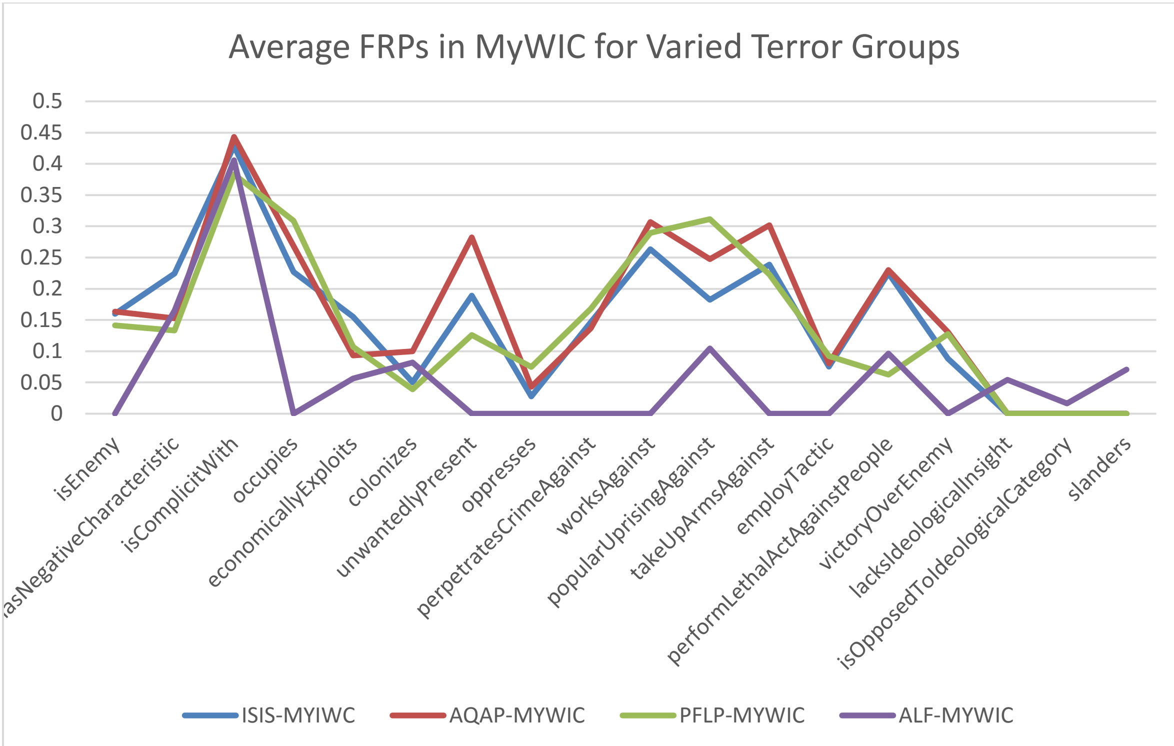

Figure 22. Resonance Profile Comparison – ISIS, AQAP, ALF, and PFLP[22]

We use the above chart to address these two concerns. The average FRP for MyWIC with ISIS, AQAP, PFLP and ALF as baseline ideologies is illustrated in Figure 21. Each contiguous line on the chart represents the average FRP produced with respect to MyWIC on a given terror organization. Each point on the chart is an average for the corresponding FrameRelation category across members of the forum *the posts of whom produced a reasonable number of frames*. There exist several users on each forum the posts of whom produced no or an unreliably small number of frames. We typically observe residual noise attributable to classification error in an FRP produced for a forum member for whom a reasonable number of frames is extracted. However, in these cases there exist enough frames to adequately bolster the probability that FRP scores

[22] Note that the FRPS represented in this chart do not span all available FrameRelation types. The baseline representations in TBO for a given radical organization do not necessarily span all FrameRelation types for which a classifier is available. For the head to head comparison the chart is limited to FrameRelation types along which frames were extracted for both organizations.

have a discernable signal amidst the noise. Where the number of extracted frames is small, we cannot be reasonably certain that a signal is present. Accordingly, for a given baseline ideology, and for a given forum member we sum over all the fields of the corresponding FRP. Only members who score over 0.5 on this test are included for Figure 21.

In developing this prototype system, we have been extensively exposed to the content of these four terrorist movements' ideologies. Al-Qaeda and ISIS both share a Salafist interpretation of Islam; they tend to decry the activities of Western powers in predominantly Muslim societies primarily on religious grounds. They both justify their violence with question interpretations of religious texts. They also propose the same collection of violent strategies to be conducted on the same collection of adversarial entities. They emit frames that evoke a similar view of an ideal world to their followers. The Popular Front (i.e. PFLP) shares much overlap with AQAP and ISIS, in terms of its adversarial stance towards Western powers. The PFLP is based on a secular vision and thus does not invoke the notion of sin, religious obligation, religious punishment, nor does it invoke imagery of an ideal future built on religious orthodoxy. Instead, the PFLP mobilizes Marxist revolutionary theory in diagnosing and prognosing the Palestinian question. A significant amount of terror movement framing is concerned with identifying and justifying an adversarial stance towards certain entities. Therein lies the major overlap between the Salafist, jihadi axis and the Popular Front. As such, we expect substantial similarity between the FRPs produced for *any* document of concern (its frames must resonate with the movements') on these three movement's frames. This trait, confirmed in Figure 22, indicates that FRPs are heavily influenced by the content of the baseline movement frames on which they are generated.

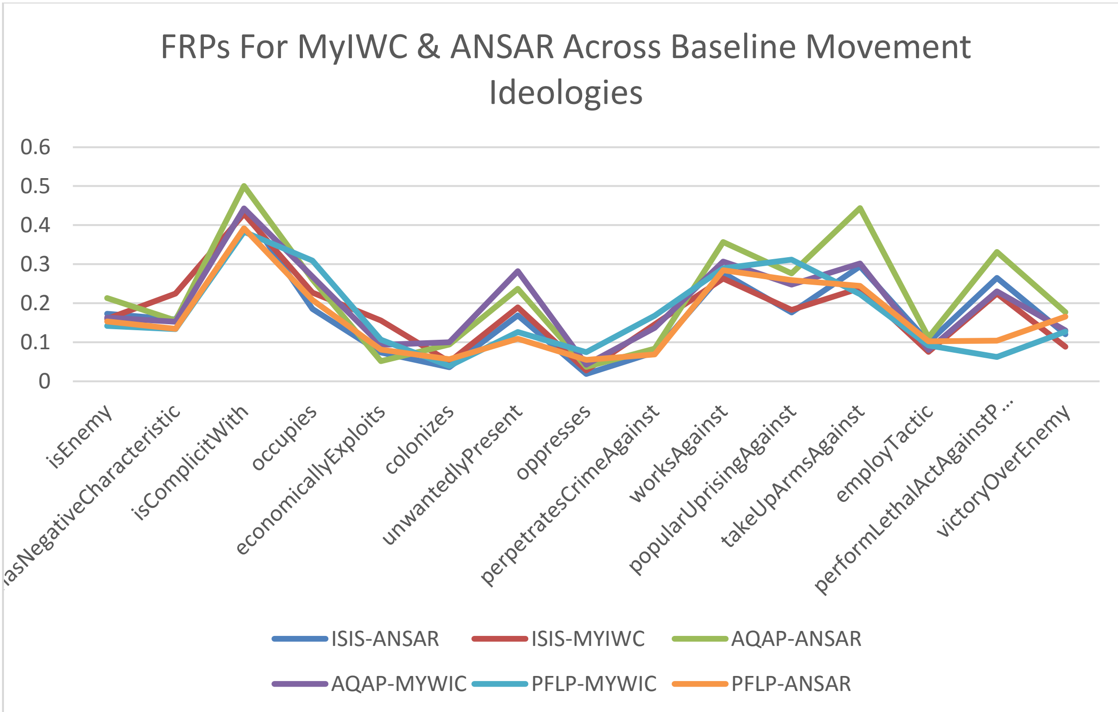

Figure 23. Resonance Profile Comparison – MyIWC and Ansar Forums on Four Movements

The implications of the results shown in Figure 21 bode well for the sensitivity of the FRPs produced by FRDS. These results indicate that FRPs reflect the substance of the representations of terrorist movement ideologies on which they are based. Figure 21 includes the average FRP from the MyIWC corpus with respect to all four terrorist movement ideologies. The Animal Liberation Front-based FRP outlies the others substantially. The ALF is an organization that encourages the use of illegal, terroristic force to discourage animal abuse especially as it relates to industrial scale animal abuse. Its adversaries include vivisectors, fur farmers, vegan pacifists, and "pseudo-abolitionist" actors who claim the banner of animal welfare but are complicit with the corporate forces driving animal cruelty. This ideological framework is almost entirely orthogonal to those of the other three movements under investigation. To be sure there is some common ground especially with respect to frames anchored on the `isComplicitWith FrameRelation`. Figure 21 indicates that the average significant FRP in MyWIC shares about resonance level across organizations on the `isComplicitWith` frames similarity A simple DL Query against TBO reveals this common ground:

```
isComplicitWith and hasFrame some( DiagnosticFrame and
hasClaimant some (Claimant and hasName value "ALF")) and hasRange
some  (IdeologicalCategory and hasName value "capitalism")or
hasDomain some (IdeologicalCategory and hasName value
"capitalism")
```

Toggling the literal "ALF," between "AQAP," "ISIS," and "PFLP" in the above query produces the following results:

Animal Liberation Front	Al-Qaeda in the Arabian Peninsula
capitalism__Eurocentric20172612032656 capitalism__consumerist20172612032656 capitalism__elitist20172612032656 capitalism__individual_consumers201726120 capitalism__white20172612032656	Instances (9 of 9) capitalism__%60_Iffah20173009023039 capitalism__Americans20173009023054 capitalism__Chicago20173009023054 capitalism__Jewish20173009023043 capitalism__Los_Angeles20173009023054 capitalism__Obama20173009023043 capitalism__Taqwah20173009023039 capitalism__Tawheed20173009023039 capitalism__Western20173009023039
Popular Front for the Liberation of Palestine	**Islamic State**
Instances (11 of 11) capitalism__Arab20175112025152 capitalism__Arab20175112025153 capitalism__Israel20175112025153 capitalism__bankers20175112025149 capitalism__capitalist20175112025149 capitalism__emirs20175112025149 capitalism__feudal_lords20175112025149 capitalism__feudalism20175112025152 capitalism__kings20175112025149 capitalism__nationalism20175112025156 capitalism__sheikhs20175112025149	No Result

Figure 24. DL Query Results – Exploring Ideological Overlap

In literal terms, the above query asks TBO to return all instances of the `isComplicitWith FrameRelation` asserted by each of the 4 movements for which either the range or domain `FrameEntity` instance is an `IdeologicalCategory` with the literal name "capitalism." Figure 23 assigns the various result sets to their respective claimant movement. All movements but ISIS assert isComplicitWith frames associating some FrameEntity with capitalism. The results can be interpreted as domain_frame_entity___range_frame_entity such that each instance expresses the claim that domain_frame_entity is complicit with range_frame_entity. Each of the movements with non-null result sets on the query is framing the same notion; capitalism is associated with concepts, people and places that their target audience is likely to hold in low esteem. The PFLP frames indicate that the movement is making a vigorous case against inequity and is attributing blame, at least in part, to capitalism. The ALF frames indicate that the group is associating capitalism with structures of domination and inequity, with Western European values, with greed, all of which are concepts frequently under attack in radical Leftist circles. AQAP is making a similar case and joins PFLP in associating capitalism with Israel and, more broadly, with Jews.

We have established in this section that our measure of frame resonance is sensitive to the substance of its baseline representation of a terrorist movement's frames. We explored this sensitivity by examining FRPs which estimate the similarity between a movement's frames and the frames extracted from a document of concern. Our examination covered four terrorist movements three of which share significant ideological overlap and one of which is tangential to the others. The examination revealed that orthogonality between the one tangential movement and three interrelated movements is captured by FRPs. This requirement is critical step towards providing satisfactory answers to our 2nd and 4th research questions. Given that our FRP measure meets this essential requirement, we now evaluate our approach to automatically identify radical content given the frames of four different terror organizations.

4.4. Automatic Identification of Radical Content

RQ2 and RQ4 both require an automatic method for identifying radical content. In Section 3.7.3, we described FRDS, the system component responsible for interpreting FRPs. FRDS classifies FRPs as espousing none, one, or many of the terrorist movement frames which are represented in TBO. For a document of concern, a frame resonance profile is generated on each collection of terror movement frames stored in TBO. FRDS labels each FRP with the collection(s) of frames with which it resonates, if any. To automate this process, we sought an unsupervised approach to FRP classification. We briefly described the approach in Section 3.7.3 and will delve into the details of the steps and models we employ. We evaluate our approach using standard classification metrics such as precision and recall.

To be useful in ad hoc situations or situations where data is scarce, our framework must be able to incorporate a new collection of terror movement frames into TBO in such a way that FRDS can use it as an FRP label immediately after the collection is added. As an FRP label is a vector of resonance scores, it is imbued with information potentially useful for classification. We can construct a standard classification problem based on FRPs by treating each one as an observation, its values as predictors. Each value in an FRP label represents the resonance of a document with a collection of terror movement frames along a `FrameRelation` . We show below that resonance levels along FrameRelation categories are good predictors of their enclosing FRP's label. Any supervised machine learning approach for classifying FRPs on a collection of terrorist movement frames would require that a considerable number of labelled FRPs be available to fit a classification function that can perform well on new FRPs. This is a time-consuming process involving a great deal of manual effort. Unsupervised classification approaches attempt to partition unlabeled observations in such a way that the differences between characteristics of the observations in one partition is maximized with respect to all other partitions. A partitioning function can be derived from this process. However, interpreting the contents of the resultant clusters is a manual task. For example, in our scenario, were we to generate clusters for a collection of FRPs such that two clusters emerge, we would ostensibly have no way of determining if either of the clusters formed around FRPs which should be labelled as resonating with their baseline movement frames. Fortunately, FRPs have properties which afford a solution.

It is intuitive to assume that all FRP predictors (FrameRelation-level resonance scores) are positively associated with the probability that a movement's frames resonate with an FRP. It is important to note that a different classifier must be learned for each collection of a terrorist movement's frames. This is because different collections of terror movement frames emphasize or de-emphasize different FrameRelation-level resonance scores. However, regardless of the classifier, the higher the resonance scores the higher the likelihood of a positive classification. This prior knowledge enables us to define a universal heuristics-based classifier to perform classifications until the system has encountered enough FRPs for each collection of terrorist movement frames. The classifier is universal because its standard can be applied to all collections of terrorist movement frames in TBO. The classifier labels as a positive example any FRP that meets two requirements: (1) it has an average FrameRelation-level resonance score that is greater than 0.3 where this average leaves out resonance scores that are equal to 0 and (2) it has non-zero scores on at least half the FrameRelation categories for which its referent radical ideology has instances. This rule-based classifier allows the system to instantly make use of a terror organization's frames towards finding radical content, and it performs well on our evaluation metrics. As more content is processed by the system against a newly added collection of terror organization frames, the rule-based classifier labels the FRPs produced from the content, growing a collection that can eventually be used to fit a classification function in with an unsupervised algorithm.

We now describe our unsupervised approach to estimating classification functions to identify radical content sympathetic to a given cause. As the rule-based classifier labels content with respect to a given radical ideology, a collection of labelled FRPs is developed. Once an adequate number of FRPs have been collected, we attempt to fit an unsupervised classifier to the data. All FRPs that are based on the same radical organization's frames share the same number FrameRelation resonance score entries. Recall that a document's resonance with a given radical ideology (collection of terrorist movement frames) is predicted by resonance scores on the FrameRelation categories used to describe the ideology. This means that FRPs based on different radical organizations will have a different collection of FrameRelation categories. Notice, for example, in Figure 21 how some terror organizations have non-zero resonance scores on certain FrameRelations while others do not. To ensure that each classification function is based on observations in the same vector-space, we ensure a 1:1 relationship between FRP classifiers and radical ideologies.

Once a critical mass of FRPs on a given radical strategy has been reached, our probability-based unsupervised classification algorithm is triggered. We define critical mass as a sample of FRPs in which the number of positive cases exceeds the number of FrameRelation categories active for the base ideology and the number of negative cases is at least equal to the number of positive cases. This critical mass ensures that all classification models meet the assumptions of the classification algorithms we employ. We perform *k*-means clustering with *k*=2 so that we define two clusters. *k*-means clustering is an unsupervised clustering algorithm that attempts to find exactly *k* clusters in a collection of observations. By setting *k*=2 we are assuming that the data is clustered into two distinct partitions. This assumption is intuitive given that we set criteria for our sample that significantly increases the probability of at least two distinct clusters existing in the data. *k*-means clustering starts from *k* arbitrarily defined cluster centers and minimizes a distortion metric such that cluster centers are iteratively updated. Clusters' centers are set when updates no longer change their values. In our experiments, we typically cycled through 4 updates before reaching this stopping point. We do not include the rule-based classifier's labels as variables in the cluster analysis so as not to fit the clusters to those labels. We do, however, use these labels to label our clusters. That is, the cluster labelled as positive (resonating) is that which has the largest number of positive labels assigned by the rule-based classifier.

Given cluster generated labels on the collection of FRPs, we can use 1-neares-neighbor (1NN) to define a decision boundary along the edges of a Voronoi tessellation. A Voronoi tessellation is a collection of non-overlapping, tessellated cells each of which defines a region around a cluster-labelled FRP. Each cell region is the collection of points around an FRP such that those points are closer (i.e. Euclidean distance) to the FRP than to any other FRP. This implies that the edges of Voronoi cells mark the half-way point between an FRP and its nearest neighbor. Those edges that mark the half-way point between FRPs in the positive cluster and FRPs in the negative cluster make up the decision boundary that 1NN learns. In this way, we can learn an unsupervised classifier.[23] However 1NN allows for extremely complex decision boundaries,

[23] This is somewhat misleading in that the classification function is derived from labelled data, technically making the approach a supervised one. However, labelling is not done manually and the initial clustering approach is unsupervised.

indicating that it tends towards overfitting. To overcome this, 1NN can be scaled to *K*NN where *K* is the number of neighbors a given point can look to for classification. This approach can smooth the decision surface and make the algorithm more robust with respect to overfitting. However, estimating an optimal *K* can be rather challenging (AUC and ROC curves can be used to those ends). Accordingly, we evaluate various classification functions for different levels of *K*. We also evaluate our rule-based classifier's usefulness along the KNN classifications for our clustered data. We conduct this analysis my examining ROC curves and AUC as well as recall and precision.

There are 756 individuals in the MyIWC forum. Each individual forum user is an observation for any one baseline ideology. Of the 756 individuals that were in the forum, 568 were usable examples. Each forum user typically made several posts to the forum. For any one individual, all posts are combined into a single collection of text that is eventually processed by my Frame Discovery System. In certain cases, no frames are discovered. These cases account for the disparity between the total number of forum users and the total number of usable examples. Of these 568, 73 are selected as what will become the gold standard testing collection. 568 - 73 = 495, and as such, the training sample consists of 495 observations. The 495 training observations are used to generate the classifiers (excluding the rule based classifier) that are evaluated in this dissertation. The resultant classification functions are evaluated on the testing set of 73 observations.

For every forum member in the gold standard collection, we read each post, labelling users as sympathetic to either AQAP, ISIS or neither. The manual labelling is informed by our extensive experience with these ideologies. We made frequent references to source propaganda materials to justify our labels. Of the 73 members, we confidently labelled 40 as resonating the frames of either ISIS or Al-Qaeda. We did not find any members who resonated one organization’s frames but not the other. Accordingly, *all* positive examples on AQAP are also positive examples on ISIS. This scenario implies that a good classifier of FRPs would need to mimic our gold standard on both ISIS and Al-Qaeda based FRPs.

We use ROC analysis as well as precision and recall on all 5 classification strategies examined. Receiver Operator Characteristic (ROC) curves plot *sensitivity* along *1-specificity* for a classification algorithm. ROC curves are derived from contingency tables generated at varying cut points on the classification algorithm’s probabilities. Below is a contingency table and the metrics that it defines.

	Predicted Positive	**Predicted Negative**	
Positive Condition	*a (true positive)*	*b (false positive)*	$a + b$
Negative Condition	*c (false negative)*	*d (true negative)*	$c + d$
	$a + c$	$b + d$	$a + b + c + d = N$

Table 25. Contingency Table

In the contingency table, positive and negative conditions are the gold standard. The a true positive, is a predicted positive observation that the gold standard also labels as positive. The count of such predictions is stored in the cell labelled *a (true positive)*. The other cells are interpreted in a similar way. The true positive rate (TPR), or the sensitivity of a classifier is defined as the proportion of truly positive predictions to all actual positive examples in the sample. It is computed as a/a+b (see contingency table). The false positive rate (FPR) or 1-specificity is the proportion of the number of falsely positive predictions to the number of actually negative examples in the sample, computed as c/c+d. It defines the probability that a classifier's signal is a false alarm. To plot a ROC curve with y-axis = TPR and x-axis = FPR, a contingency table is generated for each of several potential cut points on a classifier's output (usually a probability or score that separates positive predictions from negative ones).

We examine 5 algorithms on our data using the ROC curve analysis. First, for the rule based classifier, RBC, we develop a scoring function that yields some RBC_{score} for each FRP. The scoring function mimics the heuristics we defined earlier for the classifier, that is we label as positive any FRP such that: (1) it has an average FrameRelation-level resonance score that is greater than 0.3 where this average leaves out resonance scores that are equal to 0 and (2) it has non-zero scores on at least half the FrameRelation categories for which its referent radical ideology has instances. We define the RBC_{score} on the *j*-th FRP as follows:

$$RBC_{score}^{j} = \frac{\ln(\sum_i^N rs_i - \alpha(\sum_i^N rs_i) + 1)}{1 + (N - \theta)}$$

Where N is the number of FrameRelations active for a given ideology, rs_i is the i-th FrameRelation score in the FRP, α is a constant set to 0.3, and θ the count of non-zero FrameRelation-level resonance scores. Adding 1 to the denominator ensures that we do not divide by 0. Adding 1 to the numerator ensures that we don't take the natural log of a value less than or equal to 1.

The other classifiers are KNN (we discussed KNN and our use of it above) and Linear Discriminant Analysis, trained in the same way as our KNN process – using labelled clusters to train the classifier. We examine KNN for *K*=1, 2, and 3. The figure below summarizes our findings with the various classifiers.

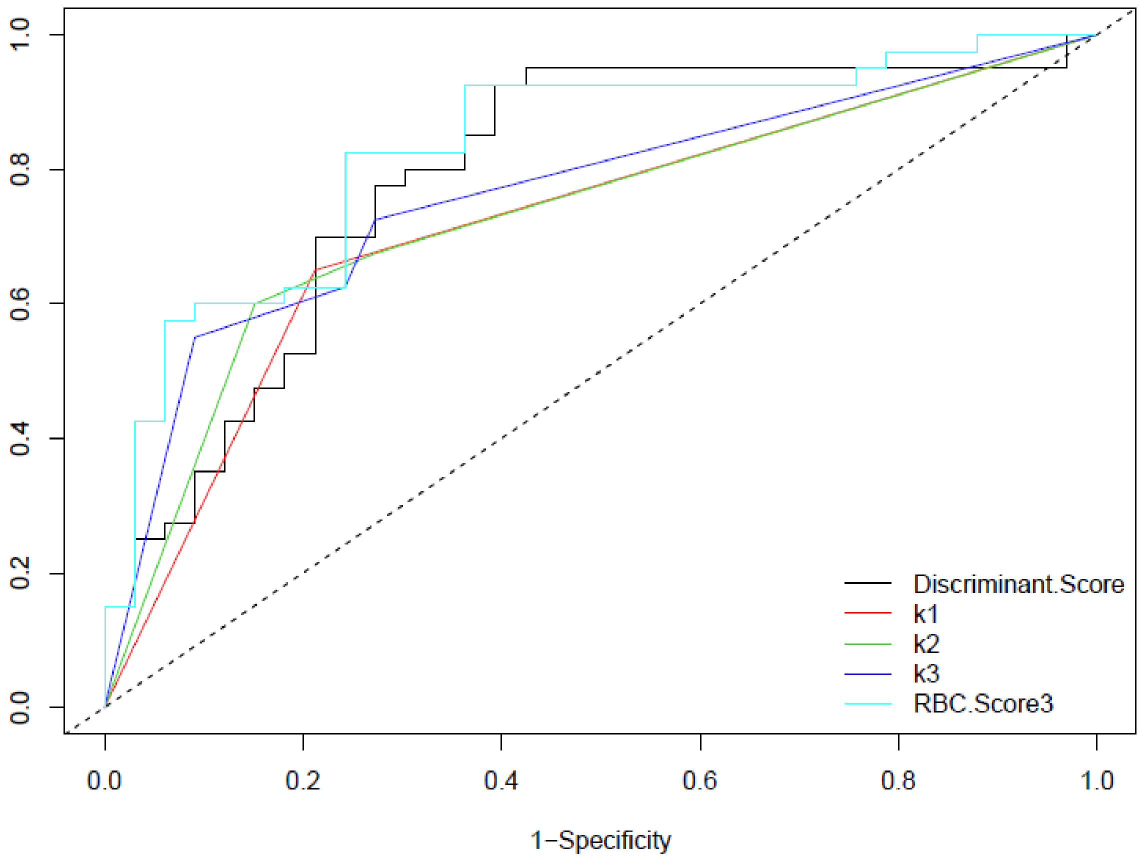

Figure 25. ROC Curves for 5 Classifiers MyWIC-AQAP

The above ROC curve is with respect to FRPs based on AQAP frames where each FRP represents similarity between those frames and the posts of a single member. The curve suggests that RBC outperforms the other classifiers. The closer a curve is to the top-leftmost section of the chart the better the classifier. This analysis can be confirmed by computing the area under the curve (AUC) as shown in the table below.

Marker	**AUC**

RBC.Score3	**0.8303**
Discriminant.Score	0.79091
k1	0.71894
k2	0.73182
k3	0.76174

Table 27. AUCs for 5 Classifiers AQAP-MyIWC FRPs

The contingency tables for discriminant analysis and RBC are shown below. The data suggests that reliable classification of FRPs is feasible. Results on the ISIS baseline enable us to reach a similar conclusion.

	Precision	**Recall**	**F1**	**TP(a)**	**FP(b)**	**FN(c)**	**TN(n)**	**Total**
Discrim.	0.72	0.95	0.833491	38	15	2	18	73
RBC	0.78	0.9	0.841304	36	10	4	23	73
	a/a+b	a/a+c						

Table 28. Contingency Table for Discriminant Analysis MyWIC-AQAP

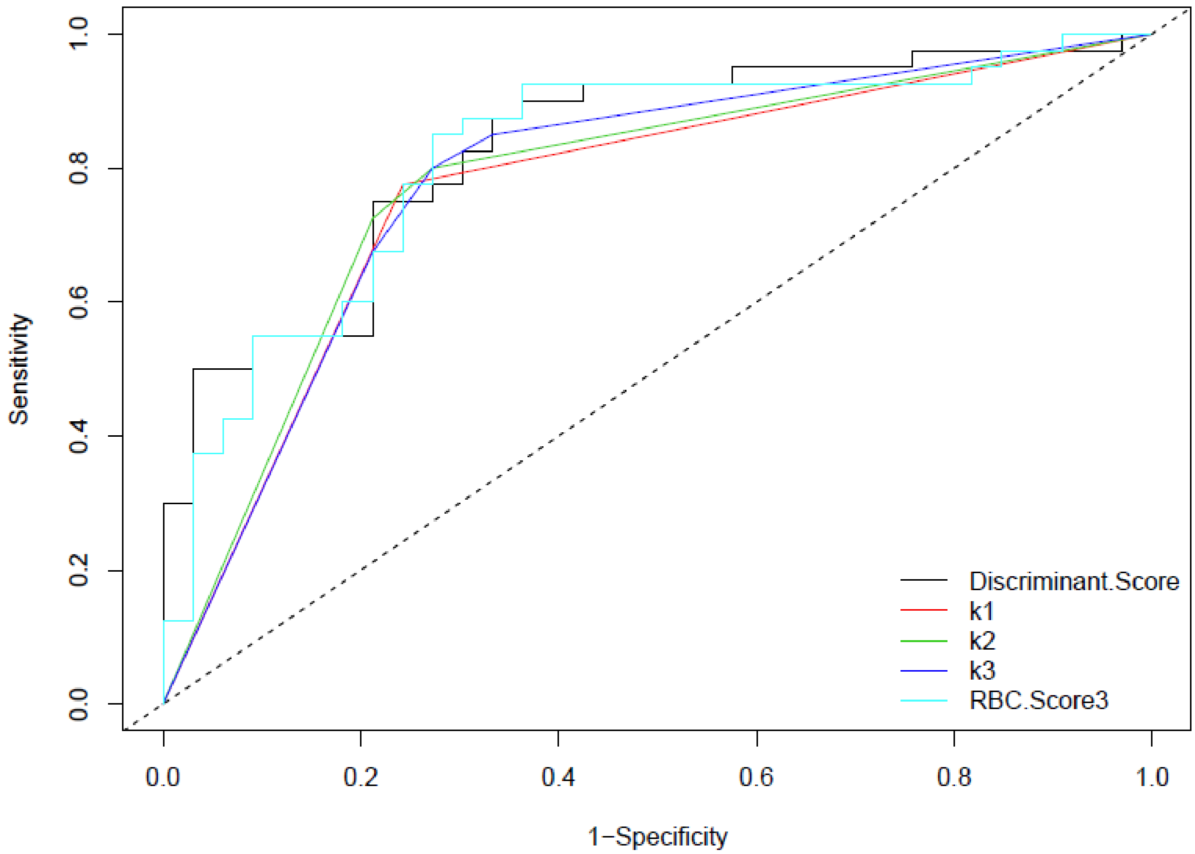

Figure 26. ROC Curves for 5 Classifiers MyWIC-ISIS

We perform the same analysis before, except that we use FRPs from MyIWC users based on ISIS frames. The discriminant analysis classifier performs slightly better than our rule based classifier on AUC analysis, is more recall sensitive in our contingency tables than RBC, but loses out to RBC with respect to precision.

Marker	**AUC**
RBC.Score3	0.81818
Discriminant.Score	**0.83106**
k1	0.76629
k2	0.77765

k3	0.78295

Table 29. AUC Analysis 5 Classifiers MyIWC ISIS FRPs

	Precision	Recall	F1	TP(a)	FP(b)	FN(c)	TN(n)	Total
Discrim.	0.711538	0.925	0.818269	37	15	3	18	73
RBC	0.810811	0.75	0.780405	30	7	10	26	73

Table 30. Contingency Table for Discriminant Analysis and RBC ISIS-MyIWC

When we designed RBC we did not expect that it would so much as compete with well-established parametric and non-parametric classifiers. As such we attempt to discover why RBC appears to perform so well in order to inform our future research on developing an even more accurate rule-based classifier and scoring system for FRP classification.

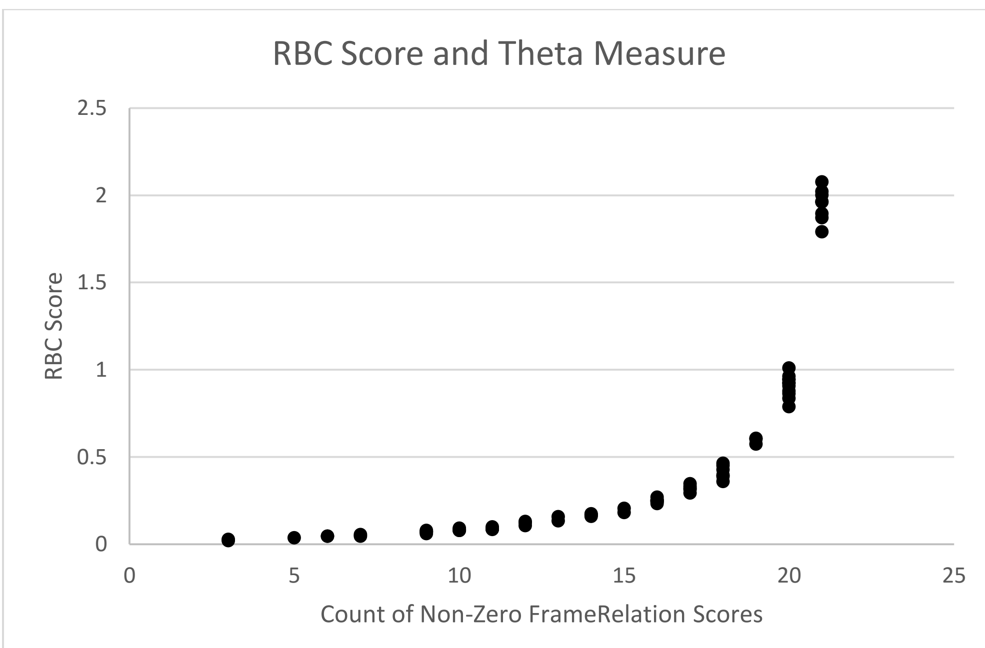

Figure 27. RBC Score and the Count of Non-Zero FrameRelation Scores, θ

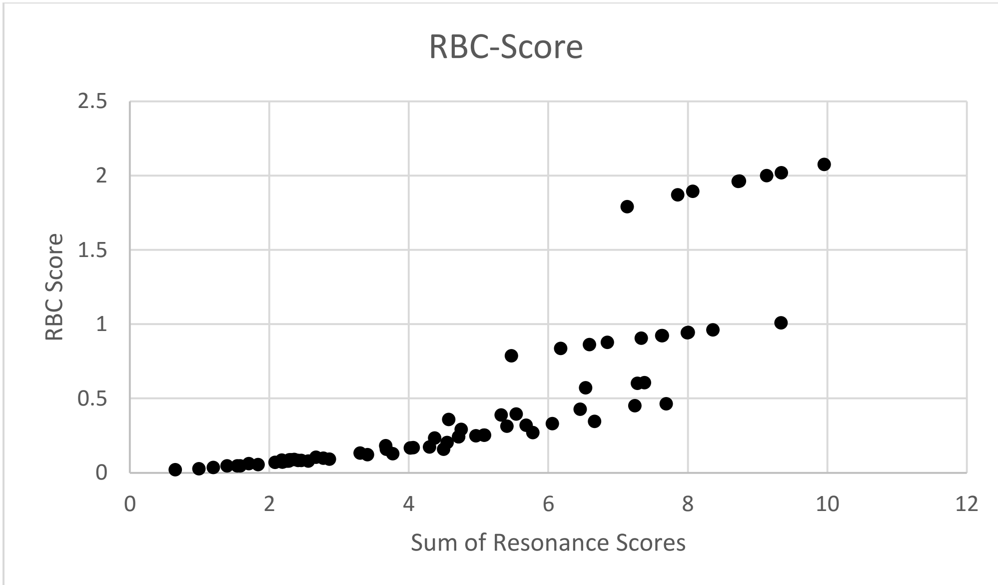

Figure 28. RBC Scores and FrameRelation Score Totals

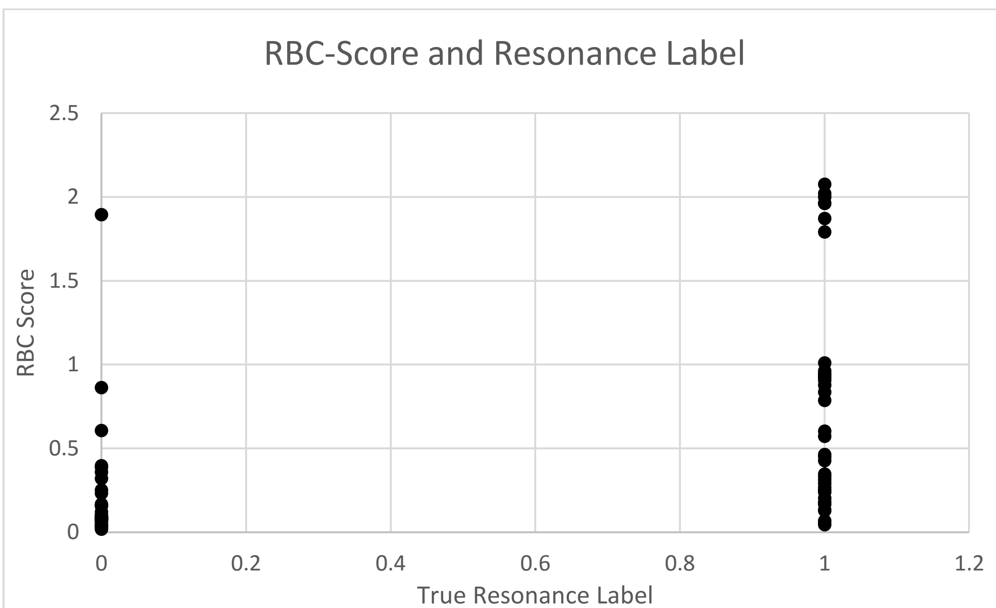

Figure 29. RBC Score and the True Resonance Labels

Figures 27 – 29 plot the relationship between RBC scores and θ, the relationship between RBC scores and the sum of resonance scores and the relationship between RBC scores and gold standard labels respectively. RBC relies extensively on the sum of resonance scores and the θ

measure as indicated by how closely the scores track these variables in Figures 27 and 28. Examining Figure 29, we see that the patterns in Figures 27 and 28 are repeated with respect to the plot of RBC scores and gold standard labels. This helps to justify our intuition that θ and the sum of resonance scores good predictors of FRP resonance. What is more, this positive, exponentiated relationship is much stronger with respect to θ than with the sum of resonance scores. The more non-zero FrameRelation scores in an FRP, the more likely a positive label. We can conclude this given that the distribution of θ along RBC scores is highly suggestive of the distribution of gold standard labels in Figure 29. However, with respect to the sum of resonance scores, we observe that low RBC scores are possible even with very high resonance scores, indicating lower predictive value.

4.5. Detecting Potentially Triggering News

As with every application of our system we begin with a collection of terrorist movement frames in TBO and read into TBO some "documents of concern." In this case, those documents are news articles. In Chapter 2, we presented a radicalization model for lone wolf terrorism. This model describes a process of radicalizing influences that culminates in an act (or several acts) of terrorism. The authors of the model (Hamm and Spaaj 2015) argue that both personal incidents and extraneous incidents may trigger radicalized individuals into action. Our framework encodes the ideological content of radicalizing messages to draw structured insights from them. Certainly, a counter-terror analyst may read news articles daily and make these determinations based on her extant knowledge. However, there is only so much news a single person can consume! We propose to encode massive streams of news in TBO, and illustrate here how themes extracted from news reports can be compared against terror movement frames using DL-Query on TBO.

To illustrate this, we queried Lexis Nexis' Global news service for all news within a three-month window related to the Keyword "Afghanistan." As this is an illustrative example, and given that we have good representations in TBO for ISIS and AQAP, we chose this keyword to capture news reports regarding activities likely relevant to either organization's frames. We executed FDS over the resultant corpus, storing the extracted frames in TBO. One way to begin the analysis once TBO has been populated is to search TBO for `Peoples` instances participating in violent `FrameRelation` instances. Instances linked to news articles may report acts of violence against certain peoples held sacred by these organizations. The DL-Query for this is simply `People`[24]. The news corpus produced 76 instances of `People`. We execute the same query against ISIS entries (576 instances) and AQAP (356 instances). Cross referencing the query results allows us to detect peoples mentioned in the news corpus who are of interest to AQAP and ISIS. The news corpus shares several People instances with ISIS and AQAP. We noticed, for instance, that Houthis, a separatist group in Yemen appeared several times for the news corpus and for ISIS.

We want to investigate the nature of ISIS' relationship with the Houthis. Accordingly, we run the following DL Query against the ISIS collection to learn more:

```
FrameRelation and hasRange some (People and hasName value
"Houthi") or hasDomain some (People and hasName value
"Houthi")
```

[24] Note that we partitioned TBO into separate ontologies for AQAP, ISIS, and the news collection. This greatly simplifies our DL queries.

This query asks for FrameRelation instances such that instances of People with the name "Houthi" participate either as ranges or as domains. We return the following result:

- Houthi__Sa20175311085318
- Houthi__Wil%C4%81yat_Sanaa20175
- Islamic__Houthi20175311085318
- Sa__Houthi20175311085318
- Wil%C4%81yat_Sanaa__Houthi20175

The first and fourth instances are erroneous (likely caused by formatting issues in the ISIS document corpus), the second instance is particularly interesting. The instance "Islamic___Houthi" is an instance of the FrameRelation takeUpArmsAgainst, a call to Muslims to take up arms against Houthis. The second instance in the list Houthi___Wilayat_Sanaa (the strange characters are URL encodings added by FDS to ensure that legible names for individuals can be stored as ontology URLs regardless of special characters), is an instance of takeUpArmsAgainst, claiming that Houthis are engaged in violent acts in Wilayat Sanaa (a region in Yemen that ISIS claims as its territory). The last frame is an inverted form of the second indicating that the classifier could not clearly state the direction of the relationship. In cases like this, we assume that there is fighting between both entities. Regardless, we can easily infer that ISIS has a negative attitude with respect to Houthis (future work involves doing this automatically). We ran the same query against the news collection and received the following individuals in response:

- Houthi__Saudi20171019021020
- Houthi__Yemen20171019021019
- Houthi__Yemenis20171019021019

The first two are individuals of takeUpArmsAgainst, and the last is an instance of worksAgainst. This suggest continued Houthi separatist violent activity in Yemen. Recall that Wilayat Sanaa is located in Yemen and claimed as an ISIS territory. Accordingly, continued Houthi fighting in Yemen may be a triggering event for ISIS operatives in Yemen.

We hope to, in future research, fully explore the potential for TBO to produce warnings in response to news articles. Currently the ontology provides a query-able structure to large collections of text. The quick insights that can be garnered in this way may be used to ease the monitoring burden of agencies tasked with interdicting terrorism. Accordingly, with respect to our fifth research question, we can produce a partial response. That is, while the functionality is not automatic, TBO is suitable for detecting news that is potentially triggering to radicalized individuals - by cross-referencing frame entities and exploring the relationships in which cross referenced frames participate.

4.6. Coordinating TBO, FDS and FRDS

TBO, FDS and FRDS are not very useful in isolation. They are designed to operate within a coordinated framework that determines which features of which subsystems need to be executed, and in what order in response to a user’s requests. We define a coordinating workflow our framework’s core task – classifying content. See the table below for a summary of this workflow.

Task	System	Processes
Evaluate Content Input: comma separated list of baseline organizations	FDS	1. Partition content into user specified units of analysis 2. Clean and process content 3. Classify and log extracted information 4. Send log to TBO

<table>
<tr><td rowspan="4">Input: directory containing the files for analysis

Input: comma separated list of file types

Input: java command line option -claimantIsFile=true when true each file in the folder represents a single claimant (one FRP is produced for each baseline organization)

Input: java command line option -learnNewOrganization=false

When false the list of organizations must already be known to TBO. When true, specify a directory where source propaganda can be found.</td><td>TBO</td><td>1. Discard information that does not comply with axioms
2. Write content to OWL file(s)
3. Index content by <code>Claimant</code></td></tr>
<tr><td>FRDS</td><td>1. Request frames from TBO, organized by Claimant
2. Wait if TBO is not ready</td></tr>
<tr><td>TBO</td><td>1. Executes query for list of user-supplied baseline organizations
2. Executes query for newly added frames
3. Send to FRDS</td></tr>
<tr><td>FRDS</td><td>1. Determine which classifier to use
2. Compute FRPs
3. Classify FRPs
4. Write FRPs, classifications, and query responses to a text file, and to a CSV file for user analysis</td></tr>
</table>

Table 31. Example Interaction between System Components

These interactions can become quite complex. For example, when a user sets -learnNewOrganization=true *and* specifies content for analysis. We defined a Java class – StartTBOPipe.java, the main method of which is the entry point to the entire solution. The main bottleneck here is FDS. For an input corpus of plain text roughly 300MB large (this is a lot of text!) FDS took 10 days to classify running on a Mac Pro with two Dual-Core Intel Xeon processors each clocked at 2.66GHz with the JVM’s heap size set to 6GB. This is not ideal. Our future work aims to move TBO to a cloud platform such that the coordinator class StartTBOPipe can make “sharding” decisions when faced with large requests. Future versions of StartTBOPipe will be able to shard input text files into smaller chunks and dynamically spool new servers when needed. The coordinator will also manage a log these sharding activities to learn, over time, optimal sharding strategies. We hope to develop algorithms for this soon.

4.7. Summary

In this chapter, we evaluated our design artifact with respect to its primary objectives of populating TBO in as accurate a manner as possible, and predicting frame resonance in radical content. We demonstrated that unsupervised classification of FRPs is possible with excellent results, that while our FDS classifiers are trained on relatively small corpora, they perform well, and that the representation of terror movement frames in TBO is sensitive to underlying ideology expressed by those frames. We also illustrate the how TBO may be used to identify potentially radical individuals, as well as how the interdependent components of TBO are coordinated towards our final goals. With these results in mind, we present our conclusions in the next chapter. More specifically, we address our current position with respect to all 6 research questions and discuss our future research plan for this promising line of inquiry.

Chapter 5 - Conclusions

5.1. Introduction

This dissertation is an account of the development and evaluation of a theoretically informed framework to detect the resonance of terror movement frames. The framework has been instantiated as prototype software application encompassing three main subsystems. At the heart of the framework is a hand-crafted knowledge base for representing the frames of any radical ideology. The knowledge base is implanted as an ontology, TBO, the terror beliefs ontology. We developed the ontology based on the results of thematic coding analysis of a sample corpus of radical, terroristic propaganda. The terror beliefs ontology guides and is populated by a frame discovery system (FDS). FDS is a collection of information extraction techniques spanning hand-crafted rules, gazetters, regular expressions and learned classifier functions. FDS reads input texts and sends the information it extracts from them to TBO for population. TBO is implemented in Web Ontology Language (OWL), a logical representational framework for domain conceptualizations. The structure we specified for TBO is axiomatic such that it imposes semantic constraints on the data that populates it. Accordingly, information extracted by FDS must meet those constraints to be asserted in TBO. Once TBO is populated with frames belonging to one or more organizations, the framework can be used to automatically identify content resonating those frames. TBO's structure ensures that each frame is linked to a `Claimant`, that is, some entity responsible for the frame. For each Claimant in TBO, there exists a collection of frames defined along `FrameRelation` types. A frame is a triple defined as `<FrameEntity`, `FrameRelation`, `FrameEntity>` where the order of that triple specifies the meaning of the frame. The triple `<Cosmetic_Companies`, `economicallyExploits`, `Animals>` is different from the triple `<Animals`,

`economicallyExploits, Cosmetic_Companies>`. A collection of such triples defines what TBO "knows" of the ideological framing of a terrorist organization. The third and final component of the framework, the frame resonance detection system (FRDS), is used to determine when inputted content resonates a Claimant instances collection of frames. To do so, FRDS must receive the inputted content (we refer to this content as document(s) of concern) as triples defined in TBO terms. Accordingly, documents of concern are first processed by FDS and inputted into TBO before FRDS begins processing. FRDS examines the similarity between documents of concern and the frames of a specified Claimant, or all Claimant instances stored within TBO. The basic representation of a single document with respect to a Claimant is a frame resonance profile (FRP). An FRP is a vector of similarity scores defined for a document of concern with respect to a TBO Claimant. The cells of an FRP are indexed by FrameRelation. Each cell value is the similarity between the Claimant and the document of concern for its corresponding FrameRelation. FRDS determines the best classification of an FRP in the set {resonates (positive), doesn't resonate (negative).

Our framework goes beyond this classification task. TBO is a repository of knowledge about radical organizations that is both structured and query-able. We discussed how this structure and resultant querying capabilities can be used to search news articles for potentially triggering events. Given the complex interactions required between TBO, FDS, and FRDS we defined a coordinating class, StartTBOPipe, that manages subsystem interactions. We provide a demonstration of StartTBOPipe for a prototypical process in TBO.

We evaluate the ability of the framework to satisfy our overarching research question: *How can an IT artifact be developed to detect online personas that pose a risk of violent behavior in the name of one or more radical ideologies?* We demonstrate that we can identify online personas who pose a risk of violent behavior in the name of one or more radical ideologies using a method that is, effectively unsupervised. We demonstrate that our representations of radical ideologies are sensitive to the underlying nature of those ideologies: similar ideologies produce similar resonance scores on the same content, and dissimilar ideologies produce dissimilar scores on the same content. We granularized our overarching question into 6 research questions the answers to which are discussed in the proceeding section.

In this chapter, we present the contributions of our research to theory and practice. We discuss limitations and prescribe an agenda for future work. Section 5.2 explores each research question, arguing the extent to which the work presented herein satisfies the question. Section 5.3 discusses the theoretical implications of work, and section 5.4 discusses the practical implications. We conclude in section 5.5 where we present a discussion on the future directions of our research.

5.2. Contributions

By addressing the following research questions, this study makes significant contributions to both theory and practice:

RQ1: What are the concepts and inter-concept relationships that are characteristic of radical ideologies and how can they be used in articulating the structure of a knowledge representation of terrorist framings?

RQ2: How may websites that frame and promote radical ideology be detected?

RQ3: How can the knowledge representation produced as a result of RQ1 be populated based on free text derived from the websites discovered in response to RQ2?

RQ4: How may the knowledge representation populated in RQ3 above be used to automatically detect the ideological affinities of an online persona?

RQ5: How can events relevant to individual, radical ideologies stored as part of the knowledge representation produced by RQ1 be detected from news sources available on the internet?

RQ6: What coordinating framework is necessary to integrate all of the above in order to effectively identify online personas for additional scrutiny by security operatives?

For each research question, we have attempted a solution. For RQ1 we developed TBO, the terror beliefs ontology. The development of TBO began with a qualitative, thematic coding exercise. We began coding a corpus of radical propaganda from organizations with various goals. We progressively and opportunistically increased our sample as we learned more about the domain. We defined three categories or themes to use in thematic coding based on our referent theory – Collective Action Framing Theory: diagnostic frames, prognostic frames, and motivational frames. Each code we extracted from the thematic exercise fit into one of these categories. Once we reached saturation, we began to analyze the extracted codes, synthesizing the list into additional themes. This exercise informed virtually all other aspects of this research. We found, for instance that diagnostic, prognostic and motivational frames could be described, generally, by a predicate interacting with two entities. We learned that these predicates only vary slightly from one terroristic movement to the next. Indeed, a single collection of predicates can be used to describe any radical movement's ideology for our purposes, varying only the subjects and objects of the predicate. Consider these two frames for example: (1) corporate interests exploit innocent animals; (2) Crusaders exploit Arabs. The notion of exploitation is universal to the framings, such that all that separates the two frames are the subjects and objects. The subject and object tell us that the first frame is likely derived from the Animal Liberation Front while the second frame is from AQAP or some other Salafist organization. We cycled through several versions of TBO attempting to balance the twin requirements of properly representing our conceptualization of terrorist movement framings and being useful for representing information automatically extracted from text.

We tackle research questions 2 and 4 simultaneously as we developed and evaluated a methodology for classifying radical content along its source. So long as the content can be converted into text, our framework is likely to identify it. By focusing on the resonance of frames we can use propaganda materials to classify previously unseen content. One of our unique contributions is the ability to do this in an unsupervised manner, a first, to our knowledge, in the related literature. While we manually trained several NER classifiers and multi-class relation extraction classifier, the training process does not have to be repeated. The method is also independent of ideological cause, such that neither retraining of the classifiers nor reconstruction of the terror beliefs ontology are necessary. Further, the system can update its knowledge about radical ideologies by reading new materials produced by the responsible terror movement. We go beyond simple classification. TBO is a query-able knowledge base enabling deep insights from large collections of text read by the approach. We ensured that the method could be used in ad-hoc situations. That is, the moment TBO has been populated with information regarding a movement it has not previously encountered, it can be immediately used to discover content resonating the movement’s frames. We achieved this using a combination of a rule based classifier and an unsupervised approach to learning class labels for frame resonance profiles.

Research question 3 demands that the knowledge representation (TBO) be populated automatically. We have accomplished this via FDS. FDS is able to label 10 different categories of entities which may serve as the subject and object arguments to `FrameRelation` types. While the resources available to us in the course of this research limited the coverage of our NER classifiers (that is, we cannot classify every `FrameEntity` type in the ontology), we find that the entities we are able to extract are sufficient for rich and varied frame representations across multiple ideologies. We used Stanford's linear chain CRF classifier and its CoreNLP API to these ends. We are also able to detect relationships between entity categories in text using a multi-class logistic regression classifier also exposed through the CoreNLP API. We defined relationship categories for this classifier to match the `FrameRelation` types defined in TBO. Controlled by our coordinating class, FDS uses these classifiers in pipeline architecture that entails several text preprocessing steps. Once FDS completes classifying an input, it writes the results of its classifications to temporary log files which the coordinating class uses to generate instances for TBO. TBO also includes several built-in axioms which reject inconsistent frames, that is, frames violating its rules. We examined the utility of representations in TBO by demonstrating that those representations are sensitive to the text on which they are based. We showed orthogonality between representations of radically different organizations as partial validation of our automatic population process. We also computed contingency tables to illustrate the results of NER and relation extraction classifiers.

While the resources available to us did not permit the development of an automated triggering event detection system, we demonstrated that we have partially answered research question 5. Using the querying capabilities of FDS, an analyst interacting with the system can rapidly identify news stories that directly touch on events which are pertinent to the representations in TBO. This technique can be extended in numerous ways, and we discuss them in the next section.

Finally, with respect to research question 6, we developed a coordinating class, StartTBOPipe which interfaces with the user and calls upon various system components to produce the requested results. The interactions between the three functional components of our framework are complex, requiring that we coordinate their interactions. While we did not develop a visual interface for the application, we expose its features via the command line. Using Java commands and StartTBOPipe options, a user can specify source materials to be read in, specify folders from which documents of concern can be collected and receive the results of the analysis. This class will also serve to coordinate distributed architecture for future iterations of our resonance detection and frame representation framework.

5.3. Theoretical Implications

We developed a theoretical framework that was suggestive of the design that we implemented. In implementing and evaluating the design herein we engage with several theoretical concepts. This work has implications for Collective Action Framing Theory – CAFT, in particular. First, we show that ontologies can represent the content and substance of collective action frames. Frame analysis (Goffman 1974) is nothing new. However, it has always been a field for qualitative analysis. While such analyses have provided numerous insights into several social movement over the decades, they preclude analysis of massive quantities of data. Our approach paves a pathway for large-scale analytics on collective action frames, terror related or not. (Vicari 2010) is the first (and only, to our knowledge) work to propose a structured linguistic analysis for representing collective action frames. Her approach, however, is geared towards manual coding. We demonstrated in this dissertation that there is legitimate potential to represent frames in a machine-readable format for very large-scale analytics. We showed that these analytics are specific enough to produce significantly different representations for significantly different social movements (in our case, terrorist movements) and highly similar representations for similar social movements. We have thus introduced the first large scale quantitative analysis of collective action frames. There are potentially myriad explanations for why data analytics have not previously been melded with CAFT. One explanation may be the philosophical provenance of the theory as a Critical Sociology of social movements. Another may be a dearth of data analytic know-how in the disciplines where CAFT is popular. Regardless of these reasons, there is much promise in such inevitably interdisciplinary studies as we show in this dissertation.

Our review of the literature in Chapter 2 revealed that extant literature on security analytics is oversupplied with binary classification approaches, and undersupplied with theoretically informed and robust artifacts. We also noted a tendency towards uninstantiated and unevaluated designs. The designs described in this dissertation are implemented in a prototype system written in pure Java code. We have evaluated the prototype with respect to the objectives of our research, filling a gap in literature. Our design is built, from conceptualization to instantiation, atop CAFT. Much of this research would not have been possible without the CAFT framework hinting at the empirics of social movements and collective action framing. We also rely extensively on the radicalization model of lone wolves in (Hamm and Spaaj 2015). Integrating this model with CAFT we postulated that when collective action frames resonate with their target audience, that audience will leave traces of this resonance by way of the "Broadcasting Intent" component of the radicalization model. Our results provide evidence that is consistent with our theoretical framework's predictions to those ends. Social movements (terror organizations are social movements) carry our framing activities which produce frames. Those frames can be distributed in myriad ways but are primarily communicated via text (Vicari 2010). The concept was introduce in (Snow and Benford 1988) who argue that "potent" frames resonate. Resonance is a conceptual continuous scale of acceptance by the target audience of the frames. When resonance is high, the audience internalizes the frames and interprets their world along the lines specified by the frames. When the frames do not resonate, the target audience does not interpret the world around them using those frames. Given that a resonating frame affects individual world views, we found it appropriate to claim that resonance would be detectable in an individual's writing given a robust understanding of the resonating frames. If this postulation is erroneous, we would expect that our representation of CAFT frames in TBO would not be able to identify potentially radicalized individuals based on what they write. Our results indicate that we can, indeed, marshal the representation of frames to identify radical content and that for our evaluation data, we can do this with reasonable accuracy.

In Chapter 1, we noted the work of Smelser who writes of the commonalities across a hand-picked number of instances of lone wolf terrorist behavior. Smelser notes that they "all arose in the context of outside political domination, all envisioned the disappearance of the oppressors, and all had a vision of the dramatic creation of an ideal life" (Smelser 2007, p. 57). Elaborating, Smelser remarks that the impulses behind these episodes are constant while their content is diverse: "though differing radically in context, they all have the same impulses: explaining the suffering of a people, assigning responsibility for it to inimical agents or forces, anger at and punishment for the agent, and the vision of an ideal and often blissful future condition free from pain and suffering" (2007:57). We believe TBO captures the essence of Smelser's findings that while group-based terroristic ideologies differ radically in context, they share a common form (what Smelser calls impulses). For instance, in the above quote, Smelser notes that in all the cases he studied, the overall group ideologies all had a tendency to assign responsibility to inimical forces for the suffering a protected group of people. TBO captures these impulses with a unifying collection of `FrameRelation` types. If radical ideologies express the same FrameRelation types (Smelser's impulses), the intuitive follow up question is: what distinguishes them? Smelser's answer is context and TBO mirrors this. While two radical ideologies may express the same FrameRelation types, they will express a diverse set of FrameEntity instances around the FrameRelation types. Our findings from populating TBO show this very clearly. For instance, in Table 32, we show filtered collection of isComplicitWith frames for the PFLP and for the ALF. These are two radically different terrorist organizations. However, they both extensively rely on the impulse to identify entities that are complicit in some problematic practice. However, their contexts are very different. We are encouraged by this. However, we can strengthen our work by assembling a panel of terrorism experts to evaluation the content and structure of TBO towards its ends. We have developed a modified Delphi methodology to do this and we are assembling the panel for our future work.

The design presented herein can be considered, in and of itself, to be a kernel theory that may be generalized to other domains. Collective Action Framing Theory has only recently been applied to terrorism studies. Our approach is applicable (1) in any domain where CAFT is applicable and (2) in any domain where there exists sufficient evidence to support the claim that frame resonance is discernable from text produced by entities resonating those frames. Put another way, we argue that our approach is a useful kernel theory whenever there exists a need to detect textual content on the Web that resonates the frames of some entity that seeks collective action. One possible generalization of this approach is an application to business marketing. While CAFT is understudied in business disciplines in general, there has been some interest in the marketing discipline (Peters 2004). The argument can be made that certain profit-oriented organizations attempt to spark social movements by emitting collective action frames. The extent to which those frames reach and influence their intended audience can cast as a frame resonance detection problem. Adapting the framework described in this dissertation to a new domain would require that new concepts at the leaf-node level of the ontological taxonomy be developed based on the frames of the entity seeking collective action. New NER and relation extraction classifiers would also need to be developed.

Query: `isComplicitWith` **`and`** `hasDomain` **`some`** `IdeologicalCategory`	
Interpretation: Which "isComplicitWith" frames have a domain that is some type of ideology	
Popular Front for the Liberation of Palestine	Animal Liberation Front
Rightist__bourgeoisie201751120251	abolitionism__Eurocentric20172612032656
capitalism__feudalism20175112025	abolitionism__capitalism20172612032656
capitalist__capitalism201751120251	abolitionism__consumerist20172612032656
capitalist__imperialist20175112025	abolitionism__elitist20172612032656
colonialism__Britain2017511202515	abolitionism__white20172612032656
colonialism__France2017511202515	consumerist__capitalism20172612032656
colonialism__imperialism20175112	elitist__capitalism20172612032656
colonialism__neo-colonialism201751	pacifists__terrorist20172612032657
community__bourgeoisie20175112	terrorist__pacifists20172612032657
feudalism__capitalism20175112025	
feudalists__colonialist_capitalism201	
imperialism__capitalism2017511202	
imperialist__Israel20175112025155	
imperialist__colonialist20175112025	
imperialist__forces2017511202515	
nationalism__capitalism2017511202	
neo-colonialism__United_States2017	
neo-colonialism__social_forces2017	
socialist__Israel20175112025155	
socialist__Soviet_Union2017511202	

Table 32 – Identical Impulses with Different Contexts

5.4. Practical Implications

We motivated our work in Chapters 1 and 2 as a response to the increased use of the Web as a tool for radicalizing individuals post 9/11. We noted in those chapters that while the characteristics of the Web make it an ideal tool for terrorist organizations to spread radical propaganda, it also presents opportunities to both learn about these ideologies and interdict these activities. The theme of lone-wolf radicalization is especially important when discussing the role of the internet for terrorist organizations. The use of the internet by these organizations is primarily as a recruiting tool. Accordingly, identifying radical content from websites, internet forums and social media is an important task in countering these organizations. The internet is made up of billions of web pages, and scouring them for radical content is an impossible task for human agents. It follows that automated techniques are needed to both identify this content and generate knowledge about the content. We demonstrated that our approach performs well as a classifier of radical content with F-Measures exceeding 80% for most tests. Our method can easily incorporate new radical ideologies with minimal human effort. For instance, an analyst charged with identifying radical content related to al-Qaeda need only point the system to, say, an instance of Inspire magazine. Once the system processes this magazine, it is immediately ready to identify resonating content without the costly need to identify numerous examples of Al-Qaeda materials. This has major implications for the application of our approach to big data contexts. Large and diverse corpora can be examined without little or no human intervention providing deep insights, structured representations and predictive capabilities. This combination of features is a first in the literature on identifying text espousing terroristic ideas. We delve into big data applications later in this chapter when we discuss our future work.

In addition, counter-terror practitioners can quickly glean valuable insights from TBO. TBO gives well defined labels and asserts relationships between these concepts such that a user can easily query it for insight. Since TBO is implemented as an OWL ontology, the rich collection of query languages available to such ontologies is also available to users of TBO OWL. For instance, users can quickly come up with lists of persons, peoples, countries, religions, and ideologies towards which a terrorist organization takes an adversarial stance. One could determine the persons, peoples, countries, religions, and ideologies that an organization encourages its followers to attack, the groups that the ideology appeals to and several other sources of insight. The framework presented here is the first of its kind in the domain to propose this unique combination of capabilities.

5.4. Limitations and Future Work

We consider the research reported within this dissertation to be the very beginning of a fruitful stream of research. That notwithstanding, there are several limitations to the current study. First, we need an approach to better correct the errors introduced to TBO by the classifiers. One such classification introduced error occurs when relation extraction is unable to definitively determine the direction of relationship. When this happens (as we demonstrated in our news analysis) two frames are recorded, each the inverse of the other. In many instances, hard coded axioms are sufficient to rid the ontology of these errors. In other cases, the axioms allow these frames through. In this preliminary research, TBO was used primarily to inform classification, and accordingly we did not want to unduly drop frames with informational content. However, we are considering using machine learning approaches in our future work to detect which of the inverted frames is the correct frame. When TBO has been populated with frames from a collection, it can end up with tens of thousands of instances. Beginning with the premise that most of these instances are correct, we may be able to identify outlying frames and correct them without direct human intervention.

Another area for future research involves increasing the density of TBO. This means collapsing multiple mentions of the same FrameEntity into a single FrameEntity individual. It also involves creating and automatically asserting ancillary but descriptive relationships in TBO. Consider our news example. We noted that Wilayat Sanaa is a city in Yemen. However, that it is a city in Yemen is not explicitly asserted in the ontology. Introducing such relationships can revolutionize the way in which we consider frames to be equivalent. For example, the AQAP frame in TBO stating that Saudi Arabia is conducting violence against Wilayat Sanaa, would be considered very dissimilar from another frame stating that Riyadh is conducting violence against Yemen. Improving the density of the ontology would eliminate such shortcomings.

One interesting area of future work is the development of a Web Crawler that implements our framework but autonomously discovers content on the Web. Indeed, our ability to classify on the fly is in anticipation of such a system. FRDS is designed to support automated identification of radical content on the Web when coupled with a Web crawler. FRDS classifiers FRPs automatically, it does not require prior labelled examples to produce excellent performance using the combination of RBC, *k*-means, *K*NN, and discriminant classifiers. In addition, the entire framework is "cause-agnostic." That is, we have demonstrated here that we can create descriptive representations in TBO for radically different terrorist organizations. We showed that these representations produce differing FRPs on the same documents of concern, where these FRPs differ in the manner that we would intuitively expect. As such, we anticipate that the crawler under development can traverse a diverse collection of websites representing several, different, terrorist ideologies. While this Web crawler is expected to provide valuable insight to responsible parties regarding the Web locations of terroristic propaganda, it will also serve to bolster our existing representations of these ideologies in TBO. That is, when we identify radical content on the Web, we can only do so with respect to the extant frame representations in TBO. This is the core strength of the method. While extant approaches to identifying terroristic or radical content in text require highly specific supervision entailing expert level NLP knowledge, ours simply requires that source propaganda be provided to the system. In light of this, the system performs better when it is imbued with diverse propaganda pertaining to any given radical ideology. The web crawler under development will enable the framework to automatically learn new facts about the frames employed by a radical organization. When the frames extracted from a webpage are shown to resonate with a particular ideology in TBO, those frames can be persisted in TBO as additional examples of that organization's frames and immediately mobilized in resonance detection. Following (Mitchell et al. 2015), we hope to develop a never ending learning framework for TBO.

Another area for additional research is the evaluation of TBO. Our results strongly suggest that TBO is useful for our purposes. However, we believe that interacting with the larger security informatics and terrorism studies community can improve TBO.

Finally, scalability is critical for the practical deployment of this prototype system. Practical use cases for this system all involve a large volume of input data in the form of documents of concern. For example, a Web forum administrator may want to incorporate this system to automatically monitor the forum for radical activity. The system must scale up if the forum produces large quantities of data with high velocity. Similarly, another use case for the framework is discovering web sites that promote radical ideologies. Our framework would receive entire websites as documents of concern. For this web crawling task to be practicable, an automated decision will need to be made for each website to determine if it harbors terrorist frames. This decision, in certain applications, will need to be made before the automated crawler can continue to other sites in its list. For instance, were the crawler employs backlink chaining to discover previously unknown radical sites (for instance the proposal in (Chen 2011)) backlink chaining would only occur when the framework is able to make a confident decision that the current website is promoting terroristic ideology. Without reasonable scalability, this web crawling task would be extremely time consuming, defeating the purpose of automation. However, our framework is computationally intensive. We have previously noted that 10 days were required on a well-equipped machine to classify and subsequently populate TBO with a 300MB plain text file. The frame discovery process is our main bottleneck, wherein candidate frames are extracted from text using a collection of NER and relation extraction classifiers. We are developing a plan to architect the frame discovery process (FDS) on the Amazon cloud, such that we can spool up machines as needed and shard large collections into serval small chunks. This, however, will not be sufficient. Our efficiency losses stem from the fact that we rely heavily upon the Java-based APIs provided by the NLP research group at Stanford University. While these API's provide state of the art features for NLP, they are slow. We made a few modifications to the code to suit our purposes and we are considering a major rework of the API's code to facilitate faster processing towards our ends.

References

Agarwal, S., and Sureka, A. 2015. "Using KNN and SVM Based One-Class Classifier for Detecting Online Radicalization on Twitter," *Distributed Computing and Internet Technology*, pp. 431–442.

Agarwal, S., Sureka, A., and Goyal, V. 2015. *Open Source Social Media Analytics for Intelligence and Security Informatics Big Data Analytics*, Springer International Publishing (doi: 10.1007/978-3-319-03689-2).

Alfonseca, E., and Manandhar, S. 2002. "An unsupervised method for general named entity recognition and automated concept discovery," in *Proceedings of the 1st international conference on general WordNet, Mysore, India*, pp. 34–43.

Alizadeh, M., and Cioffi-Revilla, C. 2014. "Distributions of Opinion and Extremist Radicalization: Insights from Agent-Based Modeling," *Social Informatics, Socinfo 2014* (8851), pp. 348–358.

Armstrong, A., and Hagel, J. 2000. "The real value of online communities," *Knowledge and communities*, Butterworth-Heinemann, pp. 85–95.

Artificial Intelligence Lab Management Information Systems Department University of Arizona. 2010. "Myiwc Forum Dataset," University of Arizona Artificial Intelligence Lab, AZSecure-data, Director Hsinchun Chen (available at www.azsecure-data.org).

Awan, A. N. 2007. "Virtual jihadist media: Function, legitimacy and radicalizing efficacy," *European Journal of Cultural Studies* (10:3), pp. 389–408 (available at http://ecs.sagepub.com/content/10/3/389.full.pdf).

Bach, N., and Badaskar, S. 2007. "A survey on relation extraction," *Language Technologies Institute, Carnegie Mellon University*.

Bakker, E., and de Graaf, B. 2010. "Lone wolves: how to prevent this phenomenon," *International Centre for Counter-Terrorism, The Hague* (November), pp. 1–8 (doi: 10.19165/2010.1.02).

Bangstad, S., and Books, Z. 2014. *Anders Breivik and the rise of Islamophobia*, Zed Books London.

Bates, R. A. 2012. "Dancing with wolves: Today's lone wolf terrorists," *The Journal of Public and Professional Sociology* (4:1), p. 15 (available at http://digitalcommons.kennesaw.edu/jpps/vol4/iss1/1/?utm_source=digitalcommons.kennesaw.edu/jpps/vol4/iss1/1&utm_medium=PDF&utm_campaign=PDFCoverPages).

von Behr, I., Reding, A., Edwards, C., and Gribbon, L. 2013. "Radicalisation in the Digital Era: The use of the internet in 15 cases of terrorism and extremism," *Rand*, p. 76 (available at http://www.rand.org/pubs/research_reports/RR453.html).

Benford, R. D. 1993. "'YOU COULD BE THE HUNDREDTH MONKEY,'" *The Sociological Quarterly* (34:2), Wiley Online Library, pp. 195–216.

Benford, R. D., and Snow, D. A. 2000. "Framing Processes and Social Movements: An

Overview and Assessment," *Annual Review of Sociology* (26:2000), pp. 611–639 (doi: 10.1146/annurev.soc.26.1.611).

Berger, J. M., and Strathearn, B. 2013. "Who Matters Online: Measuring Influence, Evaluating Content and Countering Violent Extremism in Online Social Networks," *Developments in Radicalisation and Political Violence* (March), p. 56.

Berntzen, L. E., and Sandberg, S. 2014. "The Collective Nature of Lone Wolf Terrorism: Anders Behring Breivik and the Anti-Islamic Social Movement," *Terrorism and Political Violence* (26:5), pp. 759–779 (doi: 10.1080/09546553.2013.767245).

Blanquart, G., and Cook, D. M. 2013. "Twitter Influence and Cumulative Perceptions of Extremist Support: A Case Study of Geert Wilders," *Australian Counter Terrorism Conference*, pp. 1–11.

Bostrom, R., and Heinen, J. 1977. "MIS Problems and Failures: A Socio-Technical Perspective PART I: THE CAUSES.," *MIS quarterly* (September), pp. 17–33 (available at http://search.ebscohost.com/login.aspx?direct=true&profile=ehost&scope=site&authtype=crawler&jrnl=02767783&AN=4680567&h=oT9ADtAbwdt307cUZW5ltRwV7OxHIrzyCCP9mREERRKNz6GW7d7xjcdJFzPgTbbRDd6LJlyapTTXFMIUnrR0aw%3D%3D&crl=c).

Braun, V., and Clarke, V. 2006. "Using thematic analysis in psychology," *Qualitative research in psychology* (3:2), Taylor & Francis, pp. 77–101.

Brynielsson, J., Horndahl, A., Johansson, F., Kaati, L., Martenson, C., and Svenson, P. 2012. "Analysis of weak signals for detecting lone wolf terrorists," *Proceedings - 2012 European Intelligence and Security Informatics Conference, EISIC 2012*, pp. 197–204 (doi: 10.1109/EISIC.2012.20).

Brynielsson, J., Horndahl, A., Johansson, F., Kaati, L., Mårtenson, C., and Svenson, P. 2013. "Harvesting and analysis of weak signals for detecting lone wolf terrorists," pp. 1–15.

Carroll, J., Minnen, G., and Briscoe, T. 1999. "Corpus annotation for parser evaluation," *Proceedings of the EACL workshop on Linguistically Interpreted Corpora (LINC).*

Chen, H. 2011. *Dark web: Exploring and data mining the dark side of the web* (Vol. 30), Springer Science & Business Media.

Chen, H., Reid, E., Sinai, J., Silke, A., and Ganor, B. 2008. *Terrorism informatics: knowledge management and data mining for homeland security* (Vol. 18), Springer Science & Business Media.

Chickering, R., Förster, S., and Greiner, B. 2005. *A world at total war: global conflict and the politics of destruction, 1937-1945*, Cambridge University Press.

Cohen, K., Johansson, F., Kaati, L., and Mork, J. C. 2014. "Detecting Linguistic Markers for Radical Violence in Social Media," *Terrorism and Political Violence* (26:1), pp. 246–256 (doi: 10.1080/09546553.2014.849948).

Collier, D., and Mahon, J. E. 1993. "Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Analysis.," *American Political Science Review* (87:4), Cambridge Univ Press, pp. 845–855.

Conrad, P., Bandini, J., and Vasquez, A. 2016. “Illness and the Internet: From Private to Public Experience,” *Health:* (20:1), SAGE Publications, pp. 22–32.

Cooper, H. 2001. “Terrorsim the problem of definition revisited,” *American Behavioral Scientist* (44:6), pp. 881–893.

Cowie, J., and Lehnert, W. 1996. “Information extraction,” *Communications of the ACM* (39:1), ACM, pp. 80–91.

Dahlin, J., Johansson, F., Kaati, L., Martenson, C., and Svenson, P. 2012. “Combining entity matching techniques for detecting extremist behavior on discussion boards,” *Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2012*, pp. 850–857 (doi: 10.1109/ASONAM.2012.154).

Dishman, C. 2005. “The Leaderless Nexus: When Crime and Terror Converge,” *Studies in Conflict & Terrorism* (28:3), pp. 237–252 (doi: 10.1080/10576100590928124).

Droogers, A. 1989. “Syncretism: The Problem of Definition,” *Dialogue and syncretism: An interdisciplinary approach* (1), Rodopi, p. 7.

Duggan, M., and Smith, a. 2013. “6% of online adults are reddit users. Pew Internet.,”

Ellison, N. B., and Boyd, D. 2013. “Sociality through social network sites,” *The Oxford handbook of internet studies*, Oxford University Press Oxford, pp. 151–172.

Emery, F., and Trist, E. 1969. “Socio-technical systems,” in *Systems Thinking: Selected Readings*F. Emery (ed.), Harmondsworth, Middlesex: Penguin Books, pp. 281–296 (available at http://brianwhitworth.com/hci-sts.pdf).

Fernández-López, M., Gómez-Pérez, A., and Juristo, N. 1997a. “METHONTOLOGY: From Ontological Art Towards Ontological Engineering,” *AAAI-97 Spring Symposium Series* (SS-97-06), pp. 33–40 (doi: 10.1109/AXMEDIS.2007.19).

Fernández-López, M., Gómez-Pérez, A., and Juristo, N. 1997b. “METHONTOLOGY: From Ontological Art Towards Ontological Engineering,” *AAAI-97 Spring Symposium Series* (SS-97-06), pp. 33–40 (doi: 10.1109/AXMEDIS.2007.19).

Finkel, J. R., Grenager, T., and Manning, C. 2005. “Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling,” (June), pp. 363–370.

Florian, R., Ittycheriah, A., Jing, H., and Zhang, T. 2003. “Named entity recognition through classifier combination,” in *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pp. 168–171.

Gamson, W. A. 1992. *Talking politics*, Cambridge university press.

Ganor, B. 2002. “Defining Terrorism: Is One Man’s Terrorist Another Man’s Freedom Fighter?,” *Media Asia* (29:3), pp. 123–133 (doi: 10.1080/1561426022000032060).

George, A. 1991. “The discipline of terrorology,” *Western state terrorism*, pp. 76–101.

Glaser, B., and Strauss, A. 1967. “The discovery of grounded theory: strategies for qualitative inquiry,” *Aldin, Chicago*.

Glass, K., and Colbaugh, R. (n.d.). “Web Analytics for Security Informatics.pdf,” pp. 7–9.

Goffman, E. 1974. “Frame analysis: An essay on the organization of experience.,” Harvard University Press.

Goodwin, J. 2004. “Review essay: What must we explain to explain terrorism?,” *Social Movement Studies* (3:910161110), pp. 259–265 (doi: 10.1080/1474283042000266155).

Gregor, S., and Jones, D. 2007. “The anatomy of a design theory,” *Journal of the Association for Information Systems* (8:5), pp. 312–335.

Grishman, R., and Sundheim, B. 1996. “Message Understanding Conference-6: A Brief History.,” in *Coling* (Vol. 96), pp. 466–471.

Gruber, T. R. 1993. “A translation approach to portable ontology specifications,” *Knowledge acquisition* (5:2), Academic Press, pp. 199–220.

Guarino, N. 1998. “Formal Ontology and Information Systems,” in *Proceedings of the First International Conference on Formal Ontologies in Information Systems, FOIS’98*, Amsterdam: IOS Press, pp. 3–15.

Gunning, J. 2009. “Social movement theory and the study of terrorism,” *Critical terrorism studies: A new research agenda*, Routledge Abingdon, pp. 156–177.

Hackett, C. P., and Grim, B. J. 2012. “The global religious landscape: A report on the size and distribution of the world’s major religious groups as of 2010.,”

Hamm, M., and Spaaj, R. 2015. “Lone Wolf Terrorism in America: Using Knowledge of Radicalization Pathways to Forge Prevention Strategies.,”

Harre, R. 2004. “The Social Construction of Terrorism,” in *Understanding Terrorism: Psychological Roots, Consequences, and Interventions*F. Moghaddam and A. Marsella (eds.), American Psychological Association, pp. 91–102.

Holton, J. A., and Walsh, I. 2016. *Classic Grounded Theory: Applications With Qualitative and Quantitative Data*, SAGE Publications.

Houvardas, J., and Stamatatos, E. 2006. “N-gram feature selection for authorship identification,” in *Artificial Intelligence: Methodology, Systems, and Applications*, Springer, pp. 77–86.

Hu, M., and Liu, B. 2004. “Mining and summarizing customer reviews,” in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177.

Jaskolka, J., MacCaull, W., and Khedri, R. 2015. “Towards an Ontology Design Architecture,” in *2015 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 132–135.

Johansson, F., Kaati, L., and Shrestha, A. 2013. “Detecting multiple aliases in social media,” *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining - ASONAM ’13*, pp. 1004–1011 (doi: 10.1145/2492517.2500261).

Johansson, F., Kaati, L., and Shrestha, A. 2015. “Timeprints for identifying social media users with multiple aliases,” *Security Informatics* (4:1), Springer Berlin Heidelberg, p. 7 (doi: 10.1186/s13388-015-0022-z).

Joosse, P. 2015. “Leaderless Resistance and the Loneliness of Lone Wolves: Exploring the Rhetorical Dynamics of Lone Actor Violence,” *Terrorism and Political Violence* (6553:July), pp. 1–27 (doi: 10.1080/09546553.2014.987866).

Kaplan, J. 1997. “Leaderless resistance,” *Terrorism and Political Violence* (9:3), pp. 80–95 (doi: 10.1080/09546559708427417).

Karagiannis, E. 2009. “Hizballah as a Social Movement Organization: A Framing Approach,” *Mediterranean Politics* (14:3), pp. 365–383 (doi: 10.1080/13629390903346863).

Klandermans, B. 1997. *The social psychology of protest*, Blackwell Publishers.

Klein, D., Smarr, J., Nguyen, H., and Manning, C. D. 2003. “Named entity recognition with character-level models,” in *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pp. 180–183.

Kozinets, R. V. 2002. “The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities,” *Journal of Marketing Research* (39:1), pp. 61–72 (doi: 10.1509/jmkr.39.1.61.18935).

Kripke, S. A. 1972. “Naming and necessity,” in *Semantics of natural language*, Springer, pp. 253–355.

Krötzsch, M., Simancik, F., and Horrocks, I. 2012. “A description logic primer,” *arXiv preprint arXiv:1201.4089*.

Kwok, I., and Wang, Y. 2013. “Locate the Hate: Detecting Tweets against Blacks,” *Twenty-Seventh AAAI Conference on Artificial Intelligence*, pp. 1621–1622 (available at http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CC0QFjAA&url=http://www.aaai.org/ocs/index.php/AAAI/AAAI13/paper/download/6419/6821&ei=e7hJUq2EAtKq4AOB04HoDg&usg=AFQjCNEi9mX0w71lUCo8tdxTnQJkR74MLg&am).

Luczak-roesch, M. 2015. “Socio-technical Computation,” pp. 139–142.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., and McClosky, D. 2014. “The stanford corenlp natural language processing toolkit.,” in *ACL (System Demonstrations)*, pp. 55–60.

Marcus, M. P., Marcinkiewicz, M. A., and Santorini, B. 1993. “Building a large annotated corpus of English: The Penn Treebank,” *Computational linguistics* (19:2), MIT Press, pp. 313–330.

De Marneffe, M.-C., MacCartney, B., Manning, C. D., and others. 2006. “Generating typed dependency parses from phrase structure parses,” in *Proceedings of LREC* (Vol. 6), pp. 449–454.

McAdam, D., McCarthy, J. D., Zald, M. N., and others. 1996. “Introduction: Opportunities, mobilizing structures, and framing processes—toward a synthetic, comparative perspective on social movements,” *Comparative perspectives on social movements: Political opportunities, mobilizing structures, and cultural framings*, Cambridge University Press Nova Iorque, pp. 1–20.

McCallum, A., and Li, W. 2003. “Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons,” in *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pp. 188–191.

McCarthy, J. D., and Zald, M. N. 1977. “Resource mobilization and social movements: A partial theory,” *American journal of sociology*, JSTOR, pp. 1212–1241.

McCauley, C., and Moskalenko, S. 2008. “Mechanisms of political radicalization: Pathways toward terrorism,” *Terrorism and Political Violence* (20:3), pp. 415–433 (doi: 10.1080/09546550802073367).

McCauley, C., and Moskalenko, S. 2014. “Toward a profile of lone wolf terrorists: What moves an individual from radical opinion to radical action,” *Terrorism and Political Violence* (26:1), pp. 69–85 (doi: 10.1080/09546553.2014.849916).

Meloy, R. J., and O’Toole, M. E. 2011. “The Concept of Leakage in Threat Assessment,” *Behavioral sciences & the law* (29), pp. 513–527.

Michael, G. 2012. *Lone wolf terror and the rise of leaderless resistance*, Vanderbilt University Press.

Minsky, M. 1974. “A framework for representing knowledge.,”

Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Betteridge, J., Carlson, A., Dalvi, B., and Gardner, M. 2015. “Never-Ending Learning,” in *AAAI*.

Moens, M.-F. 2006. “Information Extraction from an Historical Perspective,” *Information Extraction: Algorithms and Prospects in a Retrieval Context*, Springer, pp. 23–45.

Moskalenko, S., and McCauley, C. 2011. “The psychology of lone-wolf terrorism,” *Counselling Psychology Quarterly* (24:2), pp. 115–126 (doi: 10.1080/09515070.2011.581835).

Motik, B., Patel-Schneider, P. F., Parsia, B., Bock, C., Fokoue, A., Haase, P., Hoekstra, R., Horrocks, I., Ruttenberg, A., Sattler, U., and others. 2009. “OWL 2 web ontology language: Structural specification and functional-style syntax,” *W3C recommendation* (27:65), p. 159.

Myneni, S., Cobb, N., and Cohen, T. 2016. “In Pursuit of Theoretical Ground in Behavior Change Support Systems: Analysis of Peer-to-Peer Communication in a Health-Related Online Community,” *Journal of medical Internet research* (18:2), JMIR Publications Inc., Toronto, Canada, p. e28.

Nivre, J., Hall, J., Nilsson, J., Chanev, A., Eryigit, G., Kübler, S., Marinov, S., and Marsi, E. 2007. “MaltParser: A language-independent system for data-driven dependency parsing,” *Natural Language Engineering* (13:2), Cambridge Univ Press, pp. 95–135.

O’Callaghan, D., Greene, D., Conway, M., Carthy, J., and Cunningham, P. 2013. “Uncovering the wider structure of extreme right communities spanning popular online networks,” *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci ’13*, pp. 276–285 (doi: 10.1145/2464464.2464495).

o’Reilly, T. 2009. *What is web 2.0*, “ O’Reilly Media, Inc.”

O’Reilly, T. 2007. “What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software,” *Communivation & Strategies* (65:1), pp. 17–37 (doi:

10.2139/ssrn.1008839).

Oksanen, A., Garcia, D., and Räsänen, P. 2016. "Proanorexia Communities on Social Media," *Pediatrics* (137:1), Am Acad Pediatrics, pp. 1–3.

Ou, C. X., Pavlou, P. a, and Davison, R. M. 2014. "Swift guanxi in online marketplaces: The role of computer-mediated communication technologies," *MIS Quarterly* (38:1), pp. 209–230.

Page, M., Challita, L., and Harris, A. 2011. "Al Qaeda in the Arabian Peninsula: Framing Narratives and Prescriptions," *Terrorism and Political Violence* (23:2), pp. 150–172 (doi: 10.1080/09546553.2010.526039).

Pantucci, R. 2011. "What Have We Learned about Lone Wolves from Anders Behring Breivik?," *Perspectives on Terrorism* (5:5–6), pp. 27–42.

Pantucci, R., Rothman, K., and Meleagrou-hitchens, A. 2011. "A Typology of Lone Wolves : Preliminary Analysis of Lone Islamist Terrorists Editorial Assistants," (March), pp. 1–23.

Pentland, A. 2014. *Social physics: how good ideas spread-the lessons from a new science*, Penguin.

Perrin, A. 2015. "Social Media Usage: 2005-2015 | Pew Research Center," (available at http://www.pewinternet.org/2015/10/08/social-networking-usage-2005-2015/).

Peters, C. L. O. 2004. "Using vocabularies of motives to facilitate relationship marketing: The context of the Winnebago Itasca Travelers Club," *Journal of Vacation Marketing* (10:3), Sage Publications, pp. 209–222.

della Porta, D., and Haupt, H.-G. 2012. "Patterns of Radicalization in Political Activism: An Introduction," *Social Science History* (36:3), pp. 311–320 (doi: 10.1215/01455532-1595372).

Qin, J., Zhou, Y., Reid, E., and Chen, H. 2008. "Studying global extremist organizations' Internet presence using the darkweb attribute system," *Terrorism Informatics, Springer, USA*, pp. 237–266.

Ramshaw, L. A., and Marcus, M. P. 1999. "Text chunking using transformation-based learning," in *Natural language processing using very large corpora*, Springer, pp. 157–176.

Ranstorp, M. 2006. "Introduction: mapping terrorism research," *Mapping terrorism research. State of the art, gaps and future direction*.

Ranstorp, M. 2009. "Mapping terrorism studies after 9/11," *Critical terrorism studies: a new research agenda*, Routledge, pp. 13–33.

Reid, E., Qin, J., Chung, W., Xu, J., Zhou, Y., Schumaker, R., Sageman, M., and Chen, H. 2004. "Terrorism knowledge discovery project: a knowledge discovery approach to addressing the threats of terrorism," *Intelligence and Security Informatics Second Symposium on Intelligence and Security Informatics, ISI 2004, Tucson, AZ, USA, June 10-11, 2004. Proceedings*, pp. 125–145 (doi: 10.1007/978-3-540-25952-7_10).

Rogan, R. G. 2010. "Jihad Against Infidels and Democracy: A Frame Analysis of Jihadist Ideology and Jurisprudence for Martyrdom and Violent Jihad," *Communication Monographs* (77:3), pp. 393–413 (doi: 10.1080/03637751.2010.495949).

Sageman, M. 2011. *Leaderless jihad: Terror networks in the twenty-first century*, University of Pennsylvania Press.

Saldaña, J. 2015. *The coding manual for qualitative researchers*, Sage.

Scanlon, J. R., and Gerber, M. S. 2014. “Automatic detection of cyber-recruitment by violent extremists,” *Security Informatics* (3:1), pp. 1–10 (doi: 10.1186/s13388-014-0005-5).

Schank, R. C. 1972. “Conceptual dependency: A theory of natural language understanding,” *Cognitive psychology* (3:4), Academic Press, pp. 552–631.

Schmid, A. P., and Jongman, A. J. 1988. *Political Terrorism: A New Guide To Actors, Authors, Concepts, Data Bases, Theories, And Literature*, Transaction Publishers (available at https://books.google.com/books?hl=en&lr=&id=NgDks1hUjhMC&pgis=1).

Sekine, S., and Ranchhod, E. 2009a. *Named Entities : Recognition, classification and use*, Amsterdam, NL: John Benjamins Publishing Company (available at http://site.ebrary.com/lib/vacommonwealth/docDetail.action?docID=10318224).

Sekine, S., and Ranchhod, E. 2009b. *Named Entities : Recognition, classification and use*, Amsterdam, NL: John Benjamins Publishing Company.

Shmueli, E., Altshuler, Y., and Pentland, A. S. 2014. “Temporal Dynamics of Scale-Free Networks,” pp. 359–366.

Simon, J. D. 2013. “Lone wolf terrorism: Understanding the growing threat.,”

Sirin, E., Parsia, B., Grau, B. C., Kalyanpur, A., and Katz, Y. 2007. “Pellet: A practical owl-dl reasoner,” *Web Semantics: science, services and agents on the World Wide Web* (5:2), Elsevier, pp. 51–53.

Smelser, N. J. 2007. *The faces of terrorism: social and psychological dimensions PhD Proposal* (Vol. 1), New Jersey: Princeton University Press (doi: 10.1017/CBO9781107415324.004).

Snow, D. A., and Benford, R. D. 1988. “Ideology, frame resonance, and participant mobilization,” *International Social Movement Research.* (1), pp. 197–217.

Snow, D. A., and Benford, R. D. 2000. “Clarifying the relationship between framing and ideology in the study of social movements: A comment on Oliver and Johnston,” *Mobilization* (5:2), In press 5, pp. 55–60.

Snow, D. a., and Byrd, S. C. 2007. “Ideology, Framing Processes, and Islamic Terrorist Movements,” *Mobilization* (12:1), pp. 119–136 (available at https://webfiles.uci.edu/sbyrd/home/terroristmovements.snow.byrd.2007.pdf).

Spaaij, R. 2010. “The Enigma of Lone Wolf Terrorism: An Assessment,” *Studies in Conflict & Terrorism* (33:9), pp. 854–870 (doi: 10.1080/1057610X.2010.501426).

Spaaij, R. 2012. *Understanding Lone Wolf Terrorism Global Patterns, Motivations and Prevention*, London New York: Springer Dordrecht Heidelberg.

Stampnitzky, L. 2013. *Disciplining Terror: How Experts Invented’Terrorism’*, Cambridge University Press.

Van Stekelenburg, J., and Klandermans, B. 2013. “The social psychology of protest,” *Current*

Sociology, SAGE Publications, p. 11392113479314.

Stevens, T., and Neumann, P. R. 2009. *Countering online radicalisation: A strategy for action*, International Centre for the Study of Radicalisation and Political Violence.

Stewart, S., and Burton, F. 2008. “The Lone Wolf Disconnect,” *Stratfor* (30).

Stohl, M. 2008. “Old myths, new fantasies and the enduring realities of terrorism,” *Critical Studies on Terrorism* (1:1), pp. 5–16 (doi: 10.1080/17539150701846443).

Sun, Z., Lim, E.-P., Chang, K., Suryanto, M. A., and Gunaratna, R. K. 2008. “Document Selection for Extracting Entity and Relationship Instances of Terrorist Events,” in *Terrorism Informatics*, Springer, pp. 309–346.

Surdeanu, M., McClosky, D., Smith, M. R., Gusev, A., and Manning, C. D. 2011. “Customizing an information extraction system to a new domain,” in *Proceedings of the ACL 2011 Workshop on Relational Models of Semantics*, pp. 2–10.

Tilly, C. 1978. *From mobilization to revolution*, McGraw-Hill New York.

Tjong Kim Sang, E. F., and Buchholz, S. 2000. “Introduction to the CoNLL-2000 shared task: Chunking,” in *Proceedings of the 2nd workshop on Learning language in logic and the 4th conference on Computational natural language learning-Volume 7*, pp. 127–132.

Tjong Kim Sang, E. F., and De Meulder, F. 2003. “Introduction to the CoNLL-2003 shared task: Language-independent named entity recognition,” in *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*, pp. 142–147.

Tucker, D. 2001. “What is New about the New Terrorism and How Dangerous is It?,” *Terrorism and Political Violence* (13:3), pp. 1–14 (doi: 10.1080/09546550109609688).

Urquhart, C. 2012. *Grounded theory for qualitative research: A practical guide*, Sage.

Uschold, M., and Gruninger, M. 1996. “Ontologies: Principles, methods and applications,” *Knowledge engineering review* (11:2), Cambridge Univ Press, pp. 93–136.

Vicari, S. 2010. “Measuring collective action frames: A linguistic approach to frame analysis,” *Poetics* (38:5), Elsevier B.V., pp. 504–525 (doi: 10.1016/j.poetic.2010.07.002).

Wadhwa, P., and Bhatia, M. P. S. 2013. “Tracking on-line radicalization using investigative data mining,” *2013 National Conference on Communications, NCC 2013* (doi: 10.1109/NCC.2013.6488046).

Weimann, G. 2012. “Lone Wolves in Cyberspace,” *Journal of Terrorism Research* (3:2), pp. 75–90.

Wiil, U. K., Memon, N., and Gniadek, J. 2009. “Knowledge Management Processes, Tools and Techniques for Counterterrorism.,” *Kmis*, pp. 29–36 (doi: 10.5220/0002291900290036).

Wiktorowicz, Q. 2004. “Framing Jihad: Intramovement Framing Contests and al-Qaeda’s Struggle for Sacred Authority,” *International Review of Social History* (49:S12), pp. 159–177 (doi: 10.1017/S0020859004001683).

Wilcock, G. 2009. “Introduction to linguistic annotation and text analytics,” *Synthesis Lectures on Human Language Technologies* (2:1), Morgan & Claypool Publishers, pp. 1–159.

Wright, S. a. 2009. “Strategic Framing of Racial-Nationalism in North America and Europe: An Analysis of a Burgeoning Transnational Network,” *Terrorism and Political Violence* (21:2), pp. 189–210 (doi: 10.1080/09546550802544565).

Young, J. K., and Findley, M. G. 2011. “Promise and pitfalls of terrorism research,” *International Studies Review* (13:3), pp. 411–431 (doi: 10.1111/j.1468-2486.2011.01015.x).

Zelenko, D., Aone, C., and Richardella, A. 2003. “Kernel methods for relation extraction,” *Journal of machine learning research* (3:Feb), pp. 1083–1106.

Zuo, J., and Benford, R. D. 1995. “Mobilization Processes and the 1989 Chinese Democracy Movement,” (36:1), pp. 131–156.