State of ~~the~~ what Field?: A Topic Modeling Approach to Assessing the Impact of Digital Issues on Terrorism Research

Christopher Whyte (L. Douglas Wilder School of Government and Public Affairs)

Abstract

There is a pressing need to flesh out understanding of the contemporary research program on terrorism in the digital age. This article takes steps to do so by assessing the ideational landscape of such work. I employ a series of topic models on thousands of scholarly article abstracts to outline the thematic nature of such work, to assess the topical usage of cyber terror language in scholarly publications, and to map out the different legs of the research program. The results of this analysis show empirically that cyberterrorism is often cited as a qualifier but lacks thematic nuance and is tenuously linked to other major thematic topic areas.

Introduction

Over the past two decades, a host of scholarship in the security studies domain has problematized, conceptualized, and, where possible, examined cyberterrorism as a unique threat facing global society.[1] As any reader of such literature will quickly observe, however, there exists enduring division about cyberterrorism on several fronts.[2] Foremost among these is fundamental definitional disagreement among scholars about what constitutes cyberterrorism itself. Does the term describe cyberwarfare (high-level disruptive operations enabled via the employment of malicious code) conducted by terrorists for coercive

[1] For an early overview of the field of study see M. Conway, "What is Cyberterrorism? The Story so Far," *Journal of Information Warfare* 2, no. 2 (March 2003): 33–42. For a more recent account, see I. Awan, "Debating the Term Cyber-Terrorism: Issues and Problems," *Internet Journal of Criminology* 2045, no. 6743 (2014): 1–14.
[2] Awan, "Debating the Term," 2.

purposes?[3] Or is cyberterrorism simply constituted of actions taken by terrorists involving information and communications technologies (ICT), such as the mining and use of cryptocurrencies for nefarious purposes, the use of encryption to publish illicit content, or the use of social media to radicalize?[4] Though strong arguments can, and have, been made on both sides, there exists no clear consensus.

As one might expect given such a stark conceptual disagreement, scholars are also consistently split on the degree to which cyberterrorism actually threatens civil society and national security over and above the conventional shape of the terrorist enterprise. In recent work, Jervis, Lee, and Whiting (2016) array a series of potential cyberterrorist outcomes in noting that discourse invariably seems split between the sensationalist and the pedestrian.[5] A "cyber 9/11" or, as is often referenced more broadly in work on interstate cyber conflict, "cyber Pearl Harbor," is often held up as a scenario realistic enough that scholars and practitioners should continue to prepare for an inevitable manifestation of cyberterrorism as massive disruption to infrastructure and society.[6] At the same time, a common assessment is that guns and bombs are cheaper and, at least situationally, more effective than their cyber equivalents. If cyberterrorism is problematic, it is so partly because of the way that ICTs enable traditional terrorist operations and partly because of the way society perceives such an aid to political violence as a uniquely virulent problem set.

---

[3] See, for instance, D. Verton, "Black Ice: The Invisible Threat of Cyber-Terrorism," New York: McGrawHill Osborne, 2003; and M. Pollitt, "Cyberterrorism-Fact or Fancy?" (2001). Accessed at http://www.csgeorgetown.edu/~denning/infosec/pollitt.html on June 21, 2017.

[4] E.g, M. Kenney, "Cyber-Terrorism in a Post-Stuxnet World," *Orbis*, 59, no. 1 (2001): 111–28; J. Bronskill, "CSIS on Alert for Cyber Saboteurs: Spy Agency Monitors Threat to Computer Networks," *Ottawa Citizen* (2001): 3; G. Weimann, "The Sum of all Fears?" *Studies in Conflict and Terrorism*, 129, no. 135 (2005); G. Weimann, *Terror on the Internet: The New Arena, the New Challenges*, US Institute of Peace Press, 2006; and G. Weimann, "Al-Qaida's Extensive Use of the Internet," *CTC Sentinel (published by the Combating Terrorism C, US Military Academy at West Point)* 1, no. 2 (2008).

[5] L. Jarvis, L. Nouri, and A. Whiting, "Understanding, Locating, and Constructing Cyberterrorism," in *Cyberterrorism*, ed. T. Chen et al., 25–41, NY: Springer, 2014.

[6] For work in this vein focusing on the development and impact of cyber threat alarmism, see S. Lawson, "Beyond Cyber-Doom: Assessing the Limits of Hypothetical Scenarios in the Framing of Cyber-Threats," *Journal of Information Technology & Politics* 10, no. 1 (2013): 86–103; and M. D. Cavelty, "From Cyber-Bombs to Political Fallout: Threat Representations with an Impact in the Cyber-Security Discourse," *International Studies Review* 15, no. 1 (2013): 105–22.

Furthermore, scholarship on violent extremism and the web is also characterized by disagreement on the degree to which ICTs themselves enable radicalization, subversion, and recruitment by terrorist outfits. Do new global information substrates and tools enable new modes of radicalism and new formats of ideational diffusion among extremist communities?[7] Or are digital tools simply evolutionary steps forward in the modes of communication and logistical support always found at the operational heart of the terrorist enterprise?[8]

Given the degree to which scholarship focused on cyberterrorism is centered on existential and definitional debate, identification of distinct substantive research trajectories can be a difficult task. After all, prospective threats are often the most difficult to problematize. Moreover, a multitude of political activities and developments sit at the intersection of the information revolution and the affairs of contentious nonstate actors on the world stage, a fact that presently fuels definitional ruminations more than it does empirical analyses. Here, I take steps to assess the ideational landscape of work focusing on the terrorist enterprise in the digital age. Where recent efforts on expanding the field both qualitatively and quantitatively have largely focused on how researchers should organize their efforts in a methodological sense, I focus on identifying discrete research trajectories apparent in existing scholarship. Though new veins of focus will undoubtedly characterize the landscape of cyberterrorism research over time, existing pathways suggest the clearest areas in which the diverse researchers who occupy this space might direct their energies.

In the sections that follow, I consider two primary questions. First, what themes constitute work that lexically centers on "cyberterrorism?" Second, what has the impact of

[7] E.g., A. Stenersen, "The Internet: A Virtual Training Camp?" *Terrorism and Political Violence* 20, no. 2 (2008): 215–33; and C. Edwards and L. Gribbon, "Pathways to Violent Extremism in the Digital Era," *The RUSI Journal* 158, no. 5 (2013): 40–47.

[8] See, e.g., D. C. Benson, "Why the Internet is Not Increasing Terrorism," *Security Studies* 23, no. 2 (2014): 293–328.; and W. McCants, "Testimony, US House of Representatives, Subcommittee on Counterterrorism and Intelligence, Jihadist Use of Social Media: How to Prevent Terrorism and Preserve Innovation, 6 December 2011," retrieved Sept. 21, 2014, homeland. house. gov/sites/homeland. house. gov/files/Testimony% 20McCants. pdf.

the information revolution been on scholarship on extremism and political violence? I approach these questions in two primary ways. First, I obtain and assess topic models for scholarship taken from across numerous fields of academic study focused on "cyberterrorism" as a distinct concept. I then engage in an examination of both work focused on terrorism that distinctly engages with digital topics and research in the interdisciplinary cybersecurity field that covers terrorism and political violence. Here, I employ both traditional and dynamic topic modeling approaches to consider the impact of digital issues on the terrorism field, broadly construed. This effort speaks to a foundational disagreement still apparent in scholarly debates in the terrorism studies field on the nature of digital developments as either evolutionary or revolutionary.

The results of this analysis make several contributions to the literature on nonstate cyber conflict and extremists' use of ICT in world affairs. First, I demonstrate that cyberterrorism remains an indistinct concept that is nevertheless regularly employed in high-level descriptions of scholarship. Moreover, what thematic distinction exists in groupings that describe computer network security and risks to industry (and infrastructure) appears limited to specific academic fields, particularly computer science and business. These findings suggest that "cyberterrorism" is often employed as a qualifier in argumentation focused on semi-related topic matter. Moreover, while digital issues *do* appear as semantically distinct from other topics in the terrorism studies field, they manifest only when sufficiently expansive latent thematic model parameters are specified. The same is true of topics pertaining to terrorism studies found in the literature on cyber security and Internet governance, though such topics are more distinct than is the case with ICT-oriented work. Measures of topical similarity confirm that distinct semantic themes in this vein are referenced in small part across a wide swathe of research work but tend to have few major thematic connections to other topics, indicating thematically distinct engagement but limited

impact on the high-level topical landscape of scholarship. Dynamic models that track topics over time support this assessment and demonstrate that the significance of digital keywords to the lexicon of core terrorism studies topic areas is a limited and recent phenomenon, likely triggered by increased focus on the Islamic State and other Islamist online organization over the past half decade.

The State of the Cyberterrorism Field

Research focused on the impact of the information revolution on the terrorist enterprise might be split into two categories: research seeking to problematize "cyberterrorism" as a discrete concept with related empirical manifestations and work that generally ignores such frameworks in assessing the use of ICTs by nonstate actors. Both are of interest to this study. The first of these categories covers a broad gamut of topics ranging from the highly conceptual and nonspecific to thick accounts of particular operations.[9] At the heart of the research program on cyberterrorism, however, is a simple fact: large-scale disruptive acts of terror enabled and executed *entirely* via the use of ICTs have not yet taken place.[10] To a degree, one might reasonably think that this fact limits the ammunition of those scholars who claim that "cyberterror" is entirely about such disruption.[11] If there is no evidence to support the notion that terrorist organizations are set up to attempt coercion via cyberattack, then proponents of the idea that *all* digital elements of terrorist campaigns factor into the concept arguably boast a superior case for what should presently be considered to be

[9] For an overview, see *inter alia* Weimann, *Terror on the Internet*; Chen et al., *Cyberterrorism*; L. Jarvis et al., "The Cyberterrorism Threat: Findings from a Survey of Researchers," *Studies in Conflict & Terrorism* 37, no. 1 (2014): 68–90; McCants, "Testimony" (2011); and C. Archetti, *Understanding Terrorism in the Age of Global Media: A Communication Approach*, NY: Springer, 2012.
[10] Weimann, "The Sum of All Fears?"
[11] A perspective first significantly outlined in a series of works by D. Denning circa 2001. See, for instance, Denning, "Activism, Hacktivism, and Cyberterrorism: The Internet as a Tool for Influencing Foreign Policy," *Networks and Netwars: The Future of Terror, Crime, and Militancy* 239 (2001): 288; and "Cyberwarriors: Activists and Terrorists Turn to Cyberspace," *Harvard International Review* 23, no. 2 (2001): 70.

"cyberterror."[12] And yet, it is not apparent that the absence of a defining "cyber 9/11"-style event is negatively indicative of the probability of such incidents taking place in the future. Indeed, if anything, work covering the increasing global expansion of a black market malware economy,[13] the development of organized criminal service operations[14] and the work of hackers linked with groups like Islamic State[15] suggests that coercive cyber assaults from terrorist sources will inevitably be a feature of the landscape of international affairs.[16] Whether such features will be exceptional or common is unclear.

Beyond debate over the real shape of cyberterrorism, much work in this vein—at least in the security studies fields, broadly defined—also focuses on antagonistic actions taken by terrorists online from a critical perspective.[17] The construction of the cyberterrorist threat is of significance for security researchers and practitioners for a number of reasons. Work in this area has principally noted the significance of a constructivist set of critiques of the phenomenon as related to better understanding of how impactful cyberterror might be as differentiated from more traditional elements of the terrorist enterprise. Since terrorism is inherently about coercion of social and political functions of a given national and transnational system, understanding of how cyberterrorism is framed inevitably lends itself to

[12] Pollitt, "Cyberterrorism" ; also see Bronskill, "CSIS on Alert "and Gabriel Weimann, *Cyberterrorism: How Real is the Threat?* Vol. 31. United States Institute of Peace, 2004.
[13] See, among others, A. D. Romeo, "Hidden Threat: The Dark Web Surrounding Cyber Security," *N. Ky. L. Rev.* 43 (2016): 73; and M. Castelluccio, "The Silk Road on the Dark Web, *Strategic Finance* 99, no. 1 (2017): 55.
[14] E.g., C. Everett, "Ransomware: To Pay or Not to Pay?." *Computer Fraud & Security* 2016, no. 4 (2016): 8–12.
[15] See A. B. Atwan, *Islamic State: The Digital caliphate*, University of California Press, 2015.
[16] For work assessing the requirements and likeliness of cyber coercive operations, see D. Flemming and N. Rowe, "Cyber Coercion: Cyber Operations Short of Cyberwar," in *Proceedings of the 10th International Conference on Cyberwarfare and Security ICCWS-2015, Skukuza, South Africa,* March 2015, 95–101; C. Whyte, "Ending Cyber Coercion: Computer Network Attack, Exploitation and the Case of North Korea," *Comparative Strategy* (2016); and J. R. Lindsay and E. Gartzke, "Coercion through Cyberspace: The Stability-Instability Paradox Revisited," in *The Power to Hurt: Coercion in Theory and in Practice,* K. M. Greenhill and P. J. P. Krause, eds., New York: Oxford University Press, Forthcoming.
[17] E.g., M. Conway, "Media, Fear and the Hyperreal: The Construction of Cyberterrorism as the Ultimate Threat to Critical Infrastructures" (2008); L. Jarvis et al., "Constructing Cyberterrorism as a Security Threat: A Study of International News Media Coverage," *Perspectives on Terrorism* 9, no. 1 (2015); and L. Jarvis et al,. "Unpacking Cyberterrorism Discourse: Specificity, Status, and Scale in News Media Constructions of Threat," *European Journal of International Security* 2, no. 1 (2017): 64–87.

better comprehension of the motility of digital antagonism. In other words, threat construction matters a great deal because it can signify the nature and scope of sociopolitical responses to terrorism and because perceptions of threat constructs inform terrorist understandings of the relative value of different approaches.

In spite of some clear division in the field along definitional lines, some key contributors agree that "cyberterrorism" suffers as a distinct concept from a range of shortcomings. Foremost among these, as Jarvis and MacDonald note, is the fact that "cyberterrorism" is itself an extremely imprecise term that connotes a broad range of potential activities that don't map well onto the traditional parameters of research on the terrorist enterprise. In addition to the fact that several common terms (i.e., "cyberwar," "cyberjihad," "hactivism," etc.) are often used interchangeably with "cyberterrorism" in discussing general threats, more specific ruminations on techniques and tactics often gloss over the tenuous nature of the link between cyber actions and terrorist acts.[18] Certainly, the scenario of cyberattacks perpetrated by terrorist actors in order to realize some sort of political outcome would constitute cyberterrorism in the strictest sense of the term. But the use of ICTs more broadly to antagonize, organize, and mobilize is not unique to the terrorist enterprise.[19] Even where terrorists may be more likely to employ ICTs for illicit disruptive and circumventive reasons than the average non-state actor (such as inciting adherents to violence with hate speech, soliciting funding from criminal organizations, or stealing demographic data from governments in order to plan attacks),[20] the same might be broadly true of a host of other contentious nonviolent participants in world affairs, from countercultural advocacy groups and cultist organizations to insurgent movements and

[18] Jarvis et al., "Constructing Cyberterrorism."
[19] A point conceded by many scholars. See, among others, R. Heickerö, "Cyber Terrorism: Electronic Jihad." *Strategic Analysis* 38, no. 4 (2014): 554–65.
[20] For perhaps the best-known description of such activities in the aggregate, see G. Weimann, "Virtual Disputes: The Use of the Internet for Terrorist Debates," *Studies in Conflict & Terrorism* 29, no. 7 (2006): 623–39.

loosely-defined hacker collectives.[21] Such antagonism is not necessarily linked with the defining feature of terrorism, i.e., acts of terror designed to coerce. The "cyberterrorism" field suffers, in short, from a lack of conceptual ownership of much of what constitutes the phenomenon the term allegedly describes.[22]

The analysis that follows is intended to help remedy this problem with the cyberterrorism field and contextualize scholarly efforts by mapping out distinct threads of the field's emerging research program. In doing so, it aims to augment broad advice for research practices in this vein with a more detailed read of how distinct thematic lines of approach manifest and where there exists empirical (i.e. topical) diversity within them. The empirical assessment of thematic trends in the field aims to answer three questions. First, is there clarity of conceptual and empirical nuance, even if divided, in work that focuses on "cyberterrorism?" Second, have information security topics had a semantically meaningful impact on the content of core elements of the terrorism and political violence field of study? And finally, what are the distinct legs of the research program on digital issues within the broader terrorism studies field? Answering these questions speaks directly to longstanding divisions in how scholars have approached the study of cyber terror, broadly writ, and is critical for the empirical expansion of the field called for in recent years by Conway, Jarvis and others.

Topic Modeling as an Approach

Traditionally, any attempt to systematically examine a field of scholarship runs into distinct measurement problems, from selection bias that flavors small-scale qualitative studies of the "most important" pieces of scholarship in a given vein to an inability to infer meaning from certain bibliometric data. Bibliometric attempts to uncover trends in research over time are particularly tempting for scholars, as both categorical and usage statistics

[21] For an introduction to such perspectives, see Denning, "Activism, Hacktivism," 288.
[22] M. Conway, "Determining the Role of the Internet in Violent Extremism and Terrorism: Six Suggestions for Progressing Research," *Studies in Conflict & Terrorism* (2016).

commonly available at the article level ostensibly allow researchers to study the behavior of elements of the academy with regards to a specific topic. Among other things, such analyses can be useful for determining if there exist gender biases in citation practices, if institutional structure and imperatives prompt different modes of approach to publication of research on a topic, or if common funding discrepancies across academic fields lead to variably useful scholarly products.

As is often true of rating systems that do not derive directly from value intrinsically obvious in data, however, bibliometric approaches suffer from a poverty of meaningfulness of metrics.[23] Does an increased volume of work focusing on one particular topic over time really represent greater interest in the field or an expansion of perspectives being brought to the table? It may be the case that a field of study is undergoing a programmatic transformation of scholarship production expectations wherein research is being published in a more piecemeal fashion than was previously the case. Likewise, surging reference to a particular topic may stem from increased real-world focus on a subject and not a diversification of actual scholarly thought. The same may be said for direct usage of citation-derived metrics, where favor may be given to the "stars" of a given field or those at higher levels within professional cohorts.

One major development in the computer science field—and specifically the machine learning field—to remedy such shortcomings has involved the construction of algorithmic tools for broad-scoped lexical analysis of sources that essentially produces meaningful data directly from text. One major set of tools is the topic model. Topic models are statistical models that consider collections of documents to be made up of common themes irrespective of the breakdown of documents themselves.[24] In essence, documents can be thought of as

[23] For a survey of such arguments, see C. W. Belter, "Bibliometric Indicators: Opportunities and Limits," *Journal of the Medical Library Association: JMLA* 103, no. 4 (2015): 219.
[24] D. M. Blei et al., "Latent Dirichlet Allocation." *Journal of Machine Learning Research* 3, no. Jan (2003): 993–1022.

containing different themes and topics determined by the contours of the larger group of documents. Topic models assume that the broader corpus reflects a fixed vocabulary from which topics emerge. Distributions of topics are obtained simply via algorithmic assessment of term collocations and frequencies, with the result that individual documents are constituted of multiple themes that are only apparent at the higher corpus level.

Put more simply, probabilistic topic models allow for the latent discovery of thematic trends across an unstructured corpus of documents.[25] Topic models do not require prior input of parameters by researchers (though semi-structured models are not uncommon in research employing topic models), a fact that remedies issues of semantic bias introduced at the design phase that are common in the social sciences. Finding themes is an inductive process. Latent discovery of topics thus provides a unique capability for researchers to view the "ground truth" of trends in document construction. Thematic trends, existent across entire input corpuses but annotated as proportions across individual documents, are useful as an artifact for qualitative analysis *and* for quantitative investigations wherein semantically meaningful latent trends across a particular dataset can be referenced in coding of discrete sociopolitical phenomenon as easily as might researcher-determined variables (and without the requisite concern about input bias).[26]

Topic models can be obtained in either an unsupervised or semi-supervised fashion.[27] In other words, researchers *can* limit their inputs to setting parameters (topic frequency, smoothing, etc.) for topic generation without imparting any prior coding bias on the text-to-

[25] For full details on the evolution of topic models, see T. L. Griffiths et al., "Hierarchical Topic Models and the Nested Chinese Restaurant Process," in *Advances in Neural Information Processing Systems*, 17–24, 2004; D. M. Blei and J. D. Lafferty, "A Correlated Topic Model of Science," *The Annals of Applied Statistics* (2007): 17–35; D. Ramage et al., "Partially Labeled Topic Models for Interpretable Text Mining," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, 2011, 457–465; and Blei, "Probabilistic Topic Models," *Communications of the ACM* 55, no. 4 (2012): 77–84.
[26] In this way, topic models are more than simply a unique tool for general observation of trends in text distributions. They are a way to control for the traditional input biases that stand to skew the conveyance and results of research wherein researchers systematically adopt a subjective and potentially not replicable coding scheme.
[27] Blei, "Probabilistic."

data process. But they might also identify existing divisions in a document corpus by annotating documents accordingly (i.e., using a control variable to force topic generation artificially around a pre-set descriptor variable, such as single vs. multiple authors, "high" citation count vs. "low," etc.).[28] In this way, scholars are given tools to assess the viability of other tools and data available for a given topic. The sections below make additional use of the ability to force topic generation with minimal researcher input by taking the results of unsupervised topic models, using latent themes detected across the corpus to assign topical values at the document level,[29] and then further obtaining topic models for what is a multi-layered study of topicality in the cyberterrorism field.

Data

In order to examine the shape of the dispersed field of scholarly works that cover a range of issues at the intersection of the information revolution and the terrorist enterprise, I employ the texts of nearly 5,000 peer-reviewed article abstracts drawn from across nearly three decades and downloaded alongside relevant bibliometric data from Web of Science. The sections below assess the field in several different slices, from research that focuses centrally on the term "cyberterrorism" itself to the broader landscape of research into both terrorism and cyber security. Latter sections address the combined corpus and employ the semi-supervised design outlined above to produce a multi-layered map of existing arms of research on issues of digital age terrorism.

Source data was collected from Web of Science using a series of keyword collocation search inputs aimed at capturing all conjugations and variations of terminology related to cyberterrorism. Initially, this included specific reference to the term "cyberterrorism" and variants. Beyond this initial focus on the core term, however, the aim was to collect all relevant scholarly publications across the two main fields of study within which discussion of

[28] Ibid.
[29] As in Ramage et al., "Partially Labeled Topic Models."

cyberterrorism is most likely to appear, namely, terrorist studies and cyber security and governance research. In both instances, keyword collocation formulae were employed in order to capture all relevant source data (i.e., collocation of terms like cyber* within *n* words of defense, military, etc. and the additional provision of unique keywords like radical* or governance within *n* words of cyber*/Internet/digital, etc.) and specific unique terms (e.g., cyberjihad) were included to ensure complete coverage.

The choice to use abstracts as the primary input material follows precedent in a range of works in political science, business, computer science, and machine learning that have employed topic models in research.[30] Abstracts solve a challenge for machine learning algorithms that attempt to understand latent patterns in large amounts of input information wherein massive documents are too noisy for effective treatment. Not only is publication length variable across academic journals and other publications, but authors exhibit distinct writing styles and will diverge radically from their peers in how they reference relevant topics, establish theoretical foundations, and outline arguments. The abstract not only allows for a concise application of algorithms to source input information; it is also the element of any publication where authors must summarize their premises, research questions, and arguments in a concise and accessible fashion.

Finally, researchers employing a topic modeling approach must select three sets of parameters for model output—the number of topics desired and a pair of smoothing parameters. In the sections below, various topic models are presented.[31] For the most part, the size of each (i.e., the number of topics) is the result of trial-and-error on the part of the

[30] Abstracts are a common choice for such lexical analyses of scholarly bodies of work, as in J. Chang and D. M. Blei, "Relational Topic Models for Document Networks," in *International Conference on Artificial Intelligence and Statistics*, 2009, 81–88; A. J.-B. Chaney and D. M. Blei. "Visualizing Topic Models," in *ICWSM*, 2012; and J. D. Mcauliffe and D. M. Blei, "Supervised Topic Models," in *Advances in Neural Information Processing Systems*, 2008, 121–28.

[31] For further details on the nature of smoothing and how choices might be made by the researcher, see A. Asuncion et al., "On Smoothing and Inference for Topic Models," in *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, AUAI Press, 2009, 27–34. For a description of selection in social science work, see S. Kaplan and Keyvan Vakili, "Studying Breakthrough Innovations Using Topic Modeling: A Test Using Nanotechnology Patents," 2013.

researcher and validators to identify semantically meaningful proportions among the model outputs. Several different topic model sizes were chosen for presentation from a large number of produced models. The exception to this approach lies with the dynamic topic presented below. There, I obtain a Word2Vec model as a means for obtaining window topics without researcher input.[32] In essence, the number of topics is determined via an initial algorithmic bag-of-words assessment, a result that both informs the resultant topic model and is interesting as a measure of field diversity in itself. Smoothing terms—both term and topic terms—affect the granularity of word assignment to different topics. After experimentation and in line with other studies, I set both to 0.01.

What is Cyberterrorism?

What constitutes "cyberterrorism?" This question lies at the heart of much scholarly debate and investigation, particularly in those predominant elements of the terrorism studies community that hail from the political science, security studies, and psychology fields. The first task I turn to is a direct assessment of work that employs the term (and variants thereof, like "cyber-terrorism," "cyberterror," "cyber terror," etc.) itself. This is done, as described above, through application of a keyword, key term, and keyword collocation formula to all peer-reviewed publications available through the Web of Science research interface. The result is a corpus of around 500 articles that prominently reference or employ the term and its variants. A series of parameters for topic modeling via latent dirichlet allocation (LDA) were then experimented with such that a semantically meaningful breakdown of latent themes in such literature is represented. Figure 1 below displays initial results for a 12-topic model.

[32] As in D. Greene and J. P. Cross, "Exploring the Political Agenda of the European Parliament Using a Dynamic Topic Modeling Approach," *Political Analysis* 25, no. 1 (2017): 77–94.

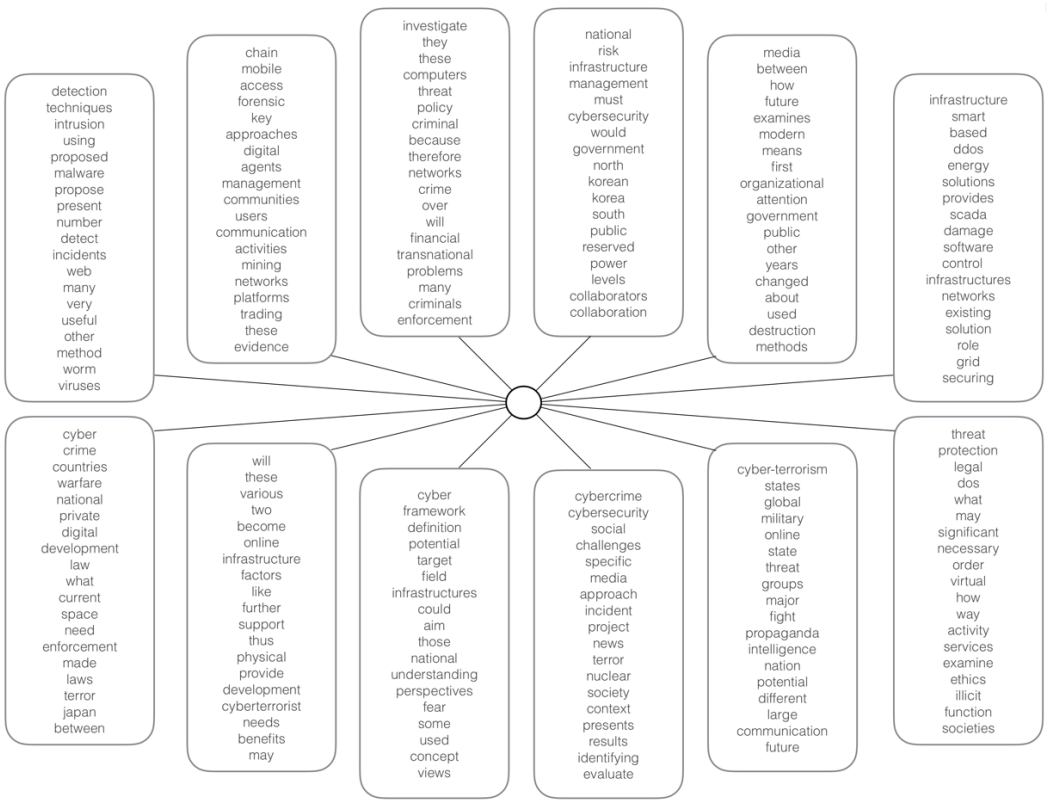

Figure 1. Results for a 12-topic LDA model obtained from nearly 500 article abstracts containing language" based around the term "cyberterrorism" and its variants.

The topics (the combinations of words grouped under each box) arrayed in Figure 1 contain several distinguishing features. In the top left, two topics clearly describe discrete thematic areas in technical cyber security. One, prominently grouping words like detection, malware, intrusion, worm and viruses together, clearly describes work focused on network security and the security of networked computers. The other, with strong topical relationship between words like chain, mobile, management, users, and platform, appears to describe the "lower" layers of computer security pertaining to the design of devices, the security of operating systems, and general issues in access control. Elsewhere, several topics make distinct reference to different threat modes or topic areas commonly differentiated in work on cyber security and cyberterrorism, including cyber threats to critical infrastructure (top right), cybercrime and legal responses (bottom left) and notable global hotspots of cyber tensions (top center).

However, the story of these model results is nevertheless one of vagueness about the "field" of cyberterrorism. Though there exist a few highly specific topics in the display in Figure 1, most are imprecise in their description of discrete themes. One possible explanation for this is that work containing reference to "cyberterrorism" is itself inherently multidisciplinary in nature. After all, the topics to be discussed under the cyberterror moniker (from actual computer/network security issues to infrastructure protection practices, law enforcement policies, and more) are likely variably of interest to scholars across different fields of academic study. Thus, in order to consider latent trends in work on cyberterrorism in greater context, Figure 2 breaks the corpus into different identifiable fields of academic research based on Web of Science categorizations and re-produces 6-topic models via LDA.

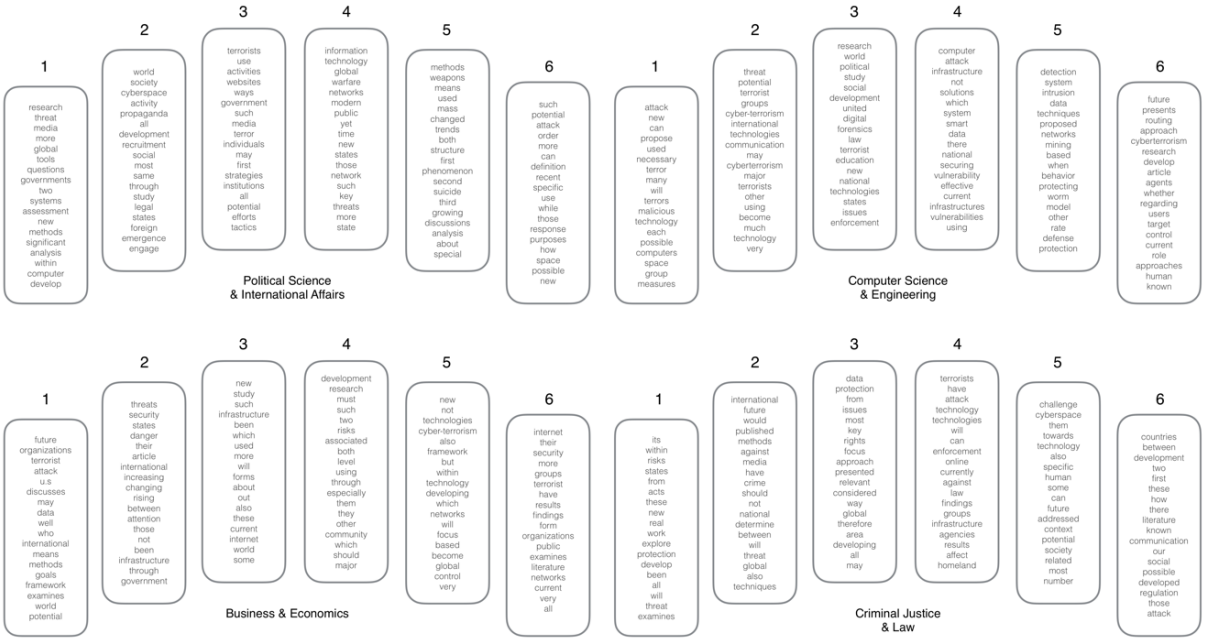

Figure 2. Results for 6-topic LDA models obtained from nearly 500 article abstracts containing language" based around the term "cyberterrorism" and its variants, split by (combined) Web of Science field categories.

Figure 2 arrays four 6-topic models by different academic fields of study. These are (1) political science and international affairs, (2) computer science and engineering, (3) business and economies, and (4) criminal justice and legal studies. Six topics were chosen following a series of experimental model production runs that suggested expanded models

exhibited diminished semantic clustering. This is not unexpected, given that any categorical narrowing of focus of a given corpus usually reduces the vocabulary parameters with which the algorithm will work.

What was unexpected is the degree to which the granularity of a limited number of topics in the Figure 1 model appears to diffuse across fields of academic research. Technical topics that address the content of security concerns regarding terrorist involvement in computer network attack *are* more readily apparent in work in the computer science and engineering fields of study. There, however, remains a single topic devoted to such relatively more granular discussion of the components of a computer network attack threat (Topic 5). Others variably and vaguely group together based on common terms linked with the threat of cyberterrorism, from *attack*, *malicious* and *target* to *vulnerability*, *infrastructure*, and *digital*.

Likewise, research referencing cyberterrorism in other fields of study presents vaguely across the topics produced. Latent thematic similarity in most cases is limited to apparent descriptions of a threat (generally phrased), underlying global and societal developments, and the need for research itself. Though some topics, such as those in the business and economics field that contain regular use of terms like risk and protection, do suggest a linkage between the content of research and the unique imperatives of a given field of academic study, almost no topics are granular enough to be considered as distinct from those around it.

This initial result is troubling, though perhaps not particularly unexpected. As authors like Jarvis and MacDonald have noted in recent work attempting to problematize cyberterrorism, there exists little in the way of consensus among scholars about what the term connotes. Moreover, it is common practice for academic work referencing the concept to treat the term as synonymous with others, like cyberjihad or even cyberwar. The models above supports this read of work focusing on “cyberterrorism” itself as being distinctly indistinct.

Indeed, insofar as the few more relatively granular themes in Figure 1 are not present across numerous identified topics, it seems reasonable to suggest that the term "cyberterrorism" and its variants are regularly employed to connote a broad-scoped threat potentiality and a distinct (though only at the highest possible level) cyber conflict concept referenced in yet-foundational discussions of the changing shape of security challenges in the digital age.

At the Intersection of Cyber and Terror

Naturally, the initial models discussed above do little to engage with the thematic granularity of work at the intersection of ICT and the terrorist enterprise. Rather, they focus on work that specifically contains the terminology of "cyberterrorism." Even insofar as the results above appear to qualitatively describe the vague condition of engagement with the concept of cyberterrorism, they cannot effectively describe the impact of the information revolution on work covering political violence, extremism, and terror in conflict. This section undertakes that task with a far more broad-scoped investigation of latent thematic trends in two sets of literatures: on (1) cyber security and Internet governance, and (2) terrorism, radicalization, and political violence. The primary aim here is to assess the degree to which ICT-specific content matter presents in the terrorism studies literature (and vice versa in the cyber/governance literature) in order that we might obtain a better understanding of how fully scholars have attempted to develop and diversify our understanding of digital age extremism. This is done in two ways: via (1) initial LDA assessments of the thematic scope of each field alongside similarness assessment of relevant topics and (2) use of non-negative factorized matric (NMF)[33] dynamic topic modeling to "view" the year-to-year evolution of scholarship focused on cyber issues in a combined corpus of both fields.

Cyberspace and Information Technologies within the Terrorism Literature

[33] See D. D. Lee and H. S. Seung, "Learning the Parts of Objects by Non-Negative Matrix Factorization," *Nature* 401, no. 6755 (1999): 788–91.

To what degree does content focused on ICT-related matters constitute a clear and distinct part of the terrorism studies field of study? To some degree, this might seem like a fool's errand insofar as ICT issues are almost inherently crosscutting. Discussion of a broad range of topics in the field might include the role of the Internet, the use of malware, or the part played by social media in enabling extremist activities. Nevertheless, topic modeling approaches have demonstrably been able to differentiate between the use of content in bodies of text as either pretext or thematically distinct and complex elements of literature. Table 1 and Figure 3 below demonstrate exactly this point.

Topic	Keys
1	*nuclear, weapons, will, threat, north, would, korea, mass, destruction, united, could, terrorists, u.s, against, such, use, proliferation, there, should, may*
2	*intelligence, been, crime, counter-terrorism, new, national, global, threats, counterterrorism, measures, such, community, policy, governance, threat, organized, through, cooperation, european, efforts*
3	*law, rights, human, against, legal, use, military, force, humanitarian, such, torture, under, freedom, criminal, armed, justice, right, protection, united, responsibility*
4	*suicide, social, organizations, violent, group, terrorists, radicalization, religious, movement, organization, how, study, organizational, analysis, members, ideology, individuals, case, extremism, who*
5	*september, after, bush, eleven, american, administration, president, 2001, nine, policy, years, iraq, since, new, over, afghanistan, george, were, had, taliban*
6	*data, study, countries, using, analysis, economic, domestic, results, civil, transnational, conflict, incidents, effect, find, than, effects, democracies, empirical, our, evidence*
7	*terrorists, they, can, government, when, may, counterterrorism, strategic, than, support, against, model, military, conflict, actors, terror, strategy, likely, will, use*
8	*islamic, qaeda, iraq, muslim, islamist, islam, al-qaeda, pakistan, western, against, syria, global, religious, jihadist, world, afghanistan, militant, radical, strategy, arab*
9	*policy, world, relations, regional, european, countries, economic, new, global, cooperation, asia, peace, east, foreign, development, region, europe, china, power, military*
10	*public, media, were, support, threat, study, news, about, citizens, how, survey, attitudes, policy, social, findings, perceptions, among, national, who, perceived*
11	*risk, information, internet, use, health, attack, such, biological, can, used, response, management, potential, cyber, technology, emergency, preparedness, communication, threats, training*
12	*how, research, analysis, studies, what, understanding, within, politics, discourse, about, can, critical, policy, been, argues, terror, paper, first, approach, through*

Table 1. Results for 12-topic LDA model obtained from 4,166 articles constituting the body of work in the Terrorism Studies field.

Table 1 displays the results for a 12-topic LDA model obtained from 4,166 article abstract taken from the Web of Science database. As with the previous section, articles were identified via reference to a unique formula of search keywords, key terms, and term collocations. Distinct from the models arrayed in Figures 1 and 2, there are clear high-level

themes in this model, related to distinct functional and substantive elements of the research program on terrorism and political violence. In particular, there are several functional sub-elements of the field that describe the methodological topology of the field (Topics 6 and 12), others that contain key concept groupings (Topics 4 and 5), and yet others that relate to significant substantive organizers (Topics 1, 2, 3, 8, and 9). With regards to those substantive organizers, there are distinct high-level topics centered on global jihad, humanitarianism and rule of law in conflict, operational counterterrorism, and macro threats to international security. This last set of topics (particularly Topics 1 and 9) demonstrate the consistent engagement of the broader security studies community with those focused on extremism and political violence, either via reference to terrorism in broader security assessments or discussion thereof in the context of evolving global security conditions.

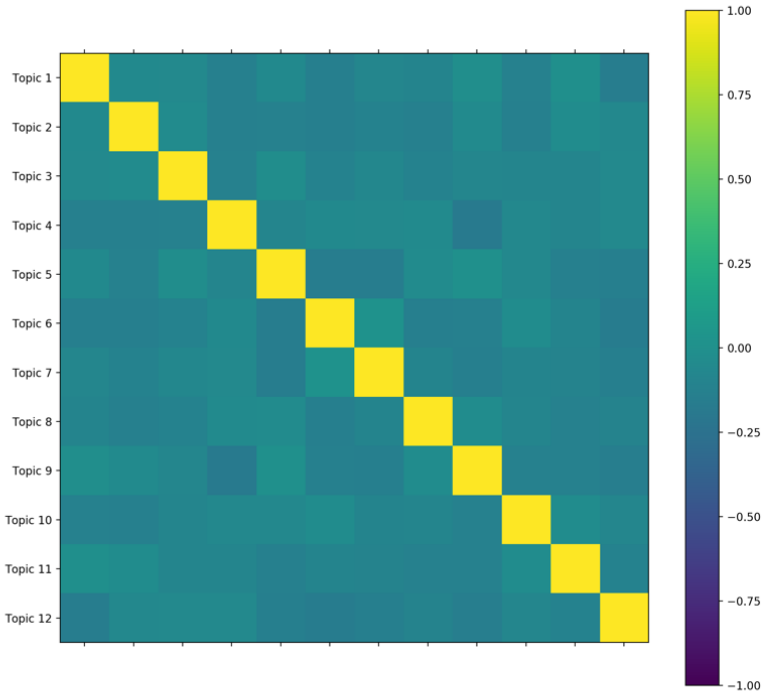

Figure 3. Correlation matrix visualization for Table 1's 12-topic LDA model.

Though keywords linked with ICT *do* appear prominently in Table 1's high-level modeling of the terrorism studies field (i.e., Topic 11's cyber, internet, technology, information, etc.), they appear alongside unrelated key terms like biological, health, and preparedness. This suggests that cyber or ICT topics remain indistinct in such high-level topical models of the field and are instead a significant element of a clear "new threats" theme. Interestingly, however, whereas one might be tempted to think that such "new threats" might be a regular feature in work studying extremism and political violence (i.e., a throwaway qualifier applied referentially in work more distinct at the field's highest levels), this appears not to be the case. Figure 3 visually displays Pearson's *r* correlation values for topic proportion similarities across the 12 topics produced in Table 1's LDA. With topic models, the general expectation is that low yield topic parameters (i.e., small requested numbers of topics to be returned) with appropriate smoothing parameter settings will return more thematically distinct topics than will more expansive ones. Figure 3 demonstrates this in showing the remarkably low, consistent *r* values associated with each topic pairing. Of specific interest here is the fact that Topic 11 is not notably "stickier" than are the other topics found. This would seem to provide evidence in favor of the null for any hypothesis assuming that focus on "new threats" means general engagement and qualification alone. At least at this high level, there appears to be reasonable evidence of the existence of distinct "new threat" topics in work in the field. Table 2 and Figure 4 below pick up further on this point.

Topic	Keys
1	*will, they, should, can, would, may, than, when, must, even, power, because, such, terrorists, all, many, those, them, likely, some*
2	*crime, economic, been, new, organized, trade, global, governance, challenges, maritime, transnational, development, attention, changes, institutions, issues, process, since, century, piracy*
3	*rights, human, law, legal, against, use, such, humanitarian, under, laws, torture, responsibility, force, right, intervention, norms, protect, freedom, detention, case*
4	*threats, law, system, actors, enforcement, order, new, global, development, regime, traditional, different, democratic, such, environmental, within, non-state, through, democracy, legal*
5	*pakistan, iraq, syria, afghanistan, iran, party, india, turkey, indian, taliban, islamic, against, parties, been, regional, country, pakistani, government, support, turkish*
6	*was, were, after, had, september, during, new, years, over, been, 2001, events, 11-Sep, since, first, time, period, attack, two, there*
7	*intelligence, counterterrorism, policy, counter-terrorism, been, government, community, police, measures, agencies, national, policies, prevent, strategy, agency, such, role, how, level, implementation*
8	*conflict, conflicts, armed, civilians, bin, insurgency, against, boko, african, killing, laden, haram, africa, insurgents, peace, non-state, nigeria, such, rebel, forces*
9	*conflict, conflicts, armed, civilians, bin, insurgency, against, boko, african, killing, laden, haram, africa, insurgents, peace, non-state, nigeria, such, rebel, forces*
10	*support, threat, public, survey, citizens, attitudes, study, perceptions, were, trust, opinion, toward, among, perceived, who, results, national, respondents, effect, americans*
11	*risk, health, public, response, emergency, preparedness, disaster, biological, risks, training, events, attack, resilience, management, event, bioterrorism, such, needs, natural, community*
12	*politics, discourse, through, how, global, critical, within, practices, identity, terror, relations, explores, narratives, power, argues, such, ways, argue, theory, narrative*
13	*economic, countries, data, transnational, study, effect, country, incidents, domestic, impact, results, using, analysis, electoral, show, growth, find, number, period, our*
14	*model, behavior, ireland, analysis, northern, models, study, present, republican, personal, activity, well, implications, specific, using, response, most, discussed, proposed, results*
15	*civil, than, democracies, likely, data, study, effects, using, democratic, countries, research, domestic, empirical, other, regime, argue, results, find, less, use*
16	*what, concept, first, about, question, second, one, other, how, warfare, terror, argues, new, contemporary, been, debate, his, both, essay, paper*
17	*military, operations, forces, strategic, force, against, use, strategy, capabilities, defense, efforts, effective, civilian, threats, nato, campaign, role, afghanistan, while, order*
18	*terrorists, terror, attack, government, target, targets, may, model, when, can, policies, counterterrorism, deterrence, actions, governments, strategic, they, against, potential, such*
19	*social, crisis, media, people, society, internet, become, world, crises, other, how, urban, public, new, one, can, action, all, communication, country*
20	*was, had, his, they, after, french, russian, movement, organization, eta, left, end, france, party, government, basque, revolutionary, terrorists, been, what*
21	*media, how, analysis, news, coverage, about, study, role, events, framing, public, official, were, discourse, examines, information, issues, rhetoric, paper, content*
22	*qaeda, network, islamic, jihadist, global, iraq, threat, networks, movement, group, al-qaeda, qaeda's, lone, ideology, isis, jihad, fighters, organization, structure, united*
23	*suicide, israel, women, palestinian, israeli, conflict, gender, female, hamas, bombing, women's, terrorists, hezbollah, lebanon, bombers, palestinians, peace, motivations, bombings, gaza*
24	*nuclear, weapons, destruction, mass, threat, proliferation, chemical, biological, wmd, could, use, terrorists, united, will, would, global, potential, such, materials, weapon*
25	*research, studies, analysis, literature, study, understanding, radicalization, theoretical, approach, case, empirical, how, framework, field, social, work, academic, our, some, into*
26	*religious, islamic, muslim, islam, islamist, religion, radical, world, muslims, western, arab, violent, identity, extremism, middle, movements, cultural, militant, among, east*
27	*organizations, group, organizational, violent, organization, they, factors, social, ideology, members, why, support, militant, goals, study, use, how, individuals, can, terrorists*
28	*policy, u.s, united, american, foreign, bush, administration, president, september, iraq, 2001, power, strategy, obama, his, policies, new, against, administration's, domestic*
29	*information, financial, technology, cyber, paper, system, infrastructure, can, money, will, use, activities, technologies, national, internet, new, used, systems, water, data*
30	*north, korea, china, law, korean, relations, chinese, sea, japan, rights, south, edited, human, diplomatic, united, u.s, diplomacy, korea's, china's, america*

Table 2. Results for 30-topic LDA model obtained from 4,166 articles constituting the body of work in the Terrorism Studies field.

Table 2 displays thirty topics produced via LDA from the corpus of terrorism studies scholarship. Thirty topics were chosen for several reasons. First, in experimentation the highest proportion of semantically meaningful themes was evident in models producing between twenty-five and thirty-seven topics. Second, and perhaps more importantly, this production of thirty topics represents the smallest number of topics wherein there is a clear "digital issues" subgrouping distinct from other terms. Topic models that include higher numbers of topics function differently than do those featuring small topics. That is not to say that they are produced via a different process or that the algorithm involved actively alters its treatment of variables. Rather, it is simply the case that topic models specifying higher numbers of topics to be produced feature greater stratification of significant topics wherein the first few specified enduringly describe themes that uniquely define documents in the corpus and later topics more truly describe distinct sub-themes that appear in (but do not singularly define) documents. Though the input challenge for the researcher is simply to produce topic models that maximize semantic meaningfulness among topics produced, different output formats allow for more or less granular assessments of the ideational construction of (in this case) an academic field.

Much as was the case with the 12-topic high level model of the terrorism studies literature, the topics displayed in Table 2 largely retain clear semantic meaning when considered alongside one another. Aside from a couple of functional topics (i.e., those dominated by common functional words not caught by a stop list in pre-processing, usually unique to a corpus), most topics describe the methodological topology of the field, key concept groupings, and significant substantive organizers. Naturally, these topics are more diverse in this expanded assessment of lexical tendencies in the terrorism studies field. Whereas Table 1 outlined a "new threats" topic that include cyber and ICT keywords among unrelated ones, Table 2 displays a more nuanced read of the field wherein issues like

biological threats to national security and cyber security are fleshed out. That said, the "digital issues" topic (Topic 28) is only distinct in this expanded model. The question, as the results above suggest, is whether ICT focus in research presents as a unique program of study or a regular reference point that retains semantic distinction in terminological terms only (i.e., the wording is unique, but there is no distinct scholarly engagement).

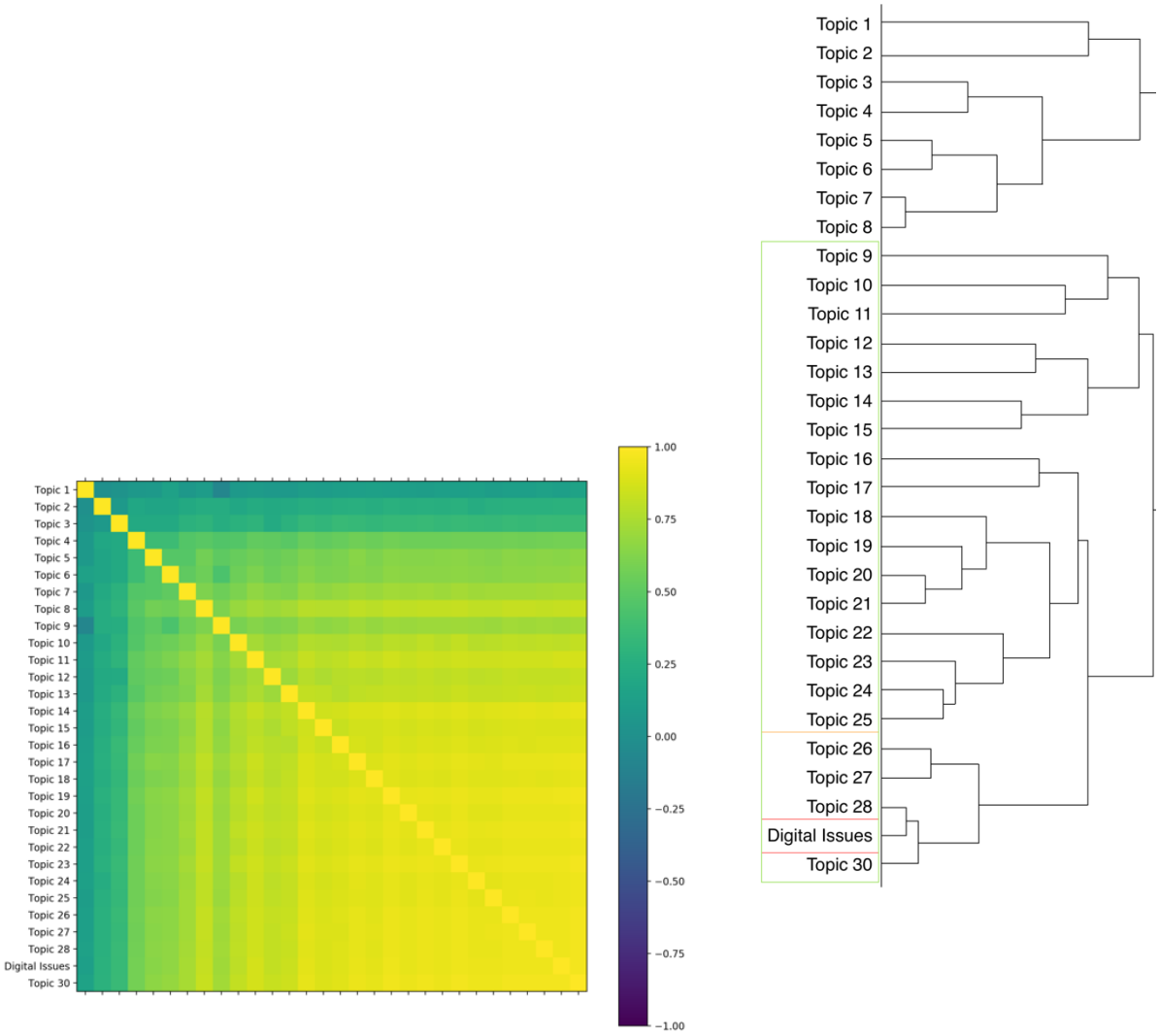

Figure 4. Correlation matrix and dendrogram visualization for Table 2's 30-topic LDA model.

The visualizations displayed in Figure 4 above address this point in two ways. On the left is displayed a correlation matrix visualization for Table 2's 30-topic LDA model. Where

the relatively low *r* scores illustrated in Figure 3 above are common for models wherein a low number of topics is specified for return, more expansive topic models should display a clear spectrum of topic correlation where initial topics reflect distinct semantic themes in a given corpus (such as dominating paradigms in an academic field of study) and lower ranked topics reflect distinct substance below the level of overarching themes. Figure 4 does just this, with functional and major substantive topics at a certain height exhibiting higher correlation with other topics at similar heights. In the dendrogram above, the sole "digital issues" topic is surprisingly divorced in similarity from other topics, suggesting that the topic is distinct but not thematically close to many others. At the same time, the correlation matrix (reflecting Pearson's *r* scores opposite the similarity matrix used to produce the clustering visualization) suggests that Topic 29 co-occurs regularly with all but the highest-numbered topics. The narrative that seems to emerge from these figures is, thus, that "digital issues" are a distinct sub-theme of the terrorism studies literature that are referenced in small part across a wide swathe of scholarly works but that have few major thematic connections to other topics.

Terror in the Literature on Cyber Conflict and Security

Flipping the prompt on its head, to what degree do topics in terrorism and political violence present in the body of scholarly work on cyber conflict and security? Here, it is perhaps most important to consider the scholarly communities that different corpuses of research documents (in this case, abstracts) proxy for. In attempting to look at terrorism topic matter in the literature on cyber conflict and security, we are essentially looking at the broad scope of an emerging field in which social scientists (primarily political scientists and related security policy-practitioner researchers) are actively self-organizing in line with emerging substantive frameworks. This stands in relatively stark contrast with literature in the terrorism studies field where the above analysis is essentially considering a more substantively focused

family of research communities. With both bodies of work, of course, the primary question is of impact of digital topics on the broader field such that we might adjudicate on the notion that ICTs have meaningfully altered the shape of different research programs.

Topic	Keys
1	*social, use, terrorist, online, groups, public, activities, they, communication, digital, media, world, who, people, through, organizations, terrorists, science, been, society*
2	*terrorism, war, state, strategic, global, world, digital, what, one, other, terrorist, china, organizations, foreign, current, united, military, become, how, also*
3	*threat, attacks, over, would, nuclear, world, most, while, power, threats, risk, military, against, technology, into, major, may, governments, russia, time*
4	*operations, military, defense, war, cyberspace, network, force, use, capabilities, than, offensive, model, but, attack, cyberwar, terms, been, work, order, computer*
5	*system, systems, data, attacks, attack, model, against, services, infrastructure, management, network, risk, service, our, based, detection, using, critical, threats, access*
6	*national, critical, government, challenges, space, all, infrastructure, countries, development, cybersecurity, threats, strategy, strategies, measures, global, nato, defence, leadership, must, needs*
7	*parties, media, between, websites, party, was, participation, campaign, two, both, study, were, social, than, democracy, using, news, data, however, public*
8	*governance, policy, power, conflict, cyberspace, development, how, actors, but, government, society, smart, between, theory, public, cyber-security, been, politics, city, complex*
9	*how, use, malware, they, there, attacks, used, software, but, also, two, research, analyze, level, most, between, tools, both, make, problem*
10	*law, human, rights, convention, protection, treaty, china, criminal, against, space, weapons, act, outer, sea, arms, jurisdiction, humanitarian, armed, territorial, laws*
11	*legal, but, law, state, electronic, north, there, privacy, including, technologies, data, private, cyberspace, should, korea, networks, traditional, some, most, south*
12	*cyberspace, research, theory, analysis, literature, cyberterrorism, policy, making, been, issues, decision, based, questions, there, then, threat, between, study, theoretical, national*

Table 3. Results for 12-topic LDA model obtained from nearly 600 articles constituting the body of work in the cyber security and conflict field.

Table 3 arrays a dozen topics obtained from a corpus of nearly 600 articles constituting the body of work in the cyber security and conflict field. As before, articles were identified via reference to a unique formula of search keywords, key terms, and term collocations. Interestingly, while the literature on cyber conflict presents as a distinct topology of research areas, several identifiable thematic threads appear to be constituted of general threat language more than of key terms one might expect. Though the table above reports only the top ~15 keywords associated with each topic, this trend holds through the top 40 keywords presented for each topic in the LDA production process.

Among the topics, there are several references to terrorism. However, while topics containing relevant keywords have distinguishing characteristics in relation to others, terrorism-inclusive topics are among those that contain the most generic terminology. While Topic 1 generally suggests the shape of research focused on terrorists' use of the web as an organizing and mobilizing tool, Topics 2 and 12 would be difficult to characterize beyond a label like "cyber threat terminology." By contrast, other areas of the field topically present as relatively distinct. Certainly there are other general terminology topics, but Topic 5 suggests the shape of work on threat management related to infrastructural threats, Topic 7 appears to describe scholarship on the democratic information environment, Topic 10 distinctly describes work on international law, and Topic 11 seems to emerge from analyses specifically focused on the Korean Peninsula.

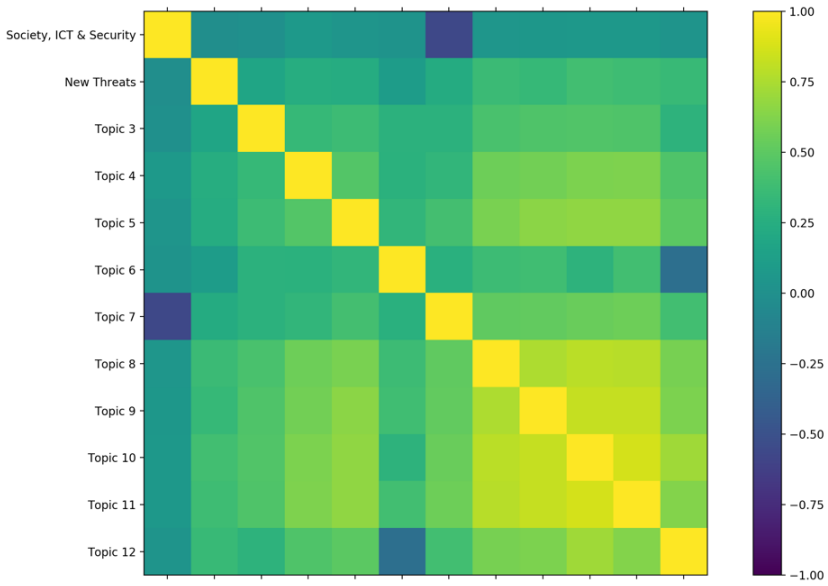

Figure 5. Correlation matrix visualization for Table 3's 12-topic LDA model.

Much as was the case with Table 1's initial high-level topic modeling of terrorism studies literature, Table 3's results suggest that terrorism is an indistinct element of the cyber

conflict field employed alongside other traditional threat nomenclature in scholarly works. Figure 5 above considers the similarity of the 12 initial topics. The results are quite interesting. Topic 1, perhaps the clearest and most distinct terrorism-oriented topic, presents as analytically distinct from other areas of the cyber conflict field. Again, as noted above, it is to be expected that lower-numbered topics will correlate only minimally with those that follow and that higher-numbered ones will likely co-occur increasingly strongly with other topics. This is largely to do with the way in which topic model outputs present themselves in the initial identification of thematically distinct major themes in a given corpus that then give way to, depending on the complexity of the text in the corpus, either less distinct major themes or minor themes that present as increasingly present across research in small, distinct amounts.

In Figure 5, Topic 1 is extremely dissimilar to other topics. This is potentially because Topic 1 describes the shape of one of the most relatively distinct program of study that exists in the cyber conflict literature: that of terrorists' use of the web to radicalize and organize. One exception to this is a high negative *r* score with respect to Topic 7. This is perhaps unsurprising if we assume that Topic 7, in dealing with democratic processes and political elements of democratic information environments, significantly contains work aimed at assessing e-governance and cyber-augmented information operations. Nevertheless, it is curious that there appears to be little crossover in two domains so concerned with the public sphere.

Topic	Keys
1	*communication, research, people, been, network, use, results, study, students, social, groups, they, being, become, studies, who, one, there, university, than*
2	*state, global, digital, one, but, all, european, sovereignty, within, threats, world, very, effective, system, systems, cyberspace, only, steps, hand, alliance*
3	*risk, threats, governments, framework, response, attacks, covert, tools, address, events, organizations, action, responses, disruptive, proposed, each, cost, industry, into, our*
4	*system, model, cyberspace, models, view, how, defense, deterrence, complex, work, modern, possible, concept, cyber-attack, assurance, better, make, deception, point, rather*
5	*government, critical, services, infrastructure, model, e-government, public, private, websites, sector, national, activities, implementation, well, ict, especially, protection, business, administration, role*
6	*space, under, systems, documents, performance, use, reports, principles, management, elements, all, strategy, control, european, activities, national, current, rights, measures, union*
7	*parties, party, campaign, small, media, major, election, minor, politics, chapter, than, both, news, canadian, within, web, competition, conservative, support, local*
8	*policy, strategies, cyber-security, research, discourse, how, social, practices, explores, both, argues, but, leadership, attempt, understanding, strategy, were, actors, meaning, safe*
9	*norms, efforts, actors, private, issues, should, national, attacks, approach, while, cyberspace, issue, they, deterrence, make, level, attention, question, clear, different*
10	*law, rights, human, convention, protection, treaty, criminal, space, weapons, against, sea, act, jurisdiction, outer, arms, armed, humanitarian, territorial, china, islands*
11	*electronic, north, south, korea, privacy, law, cyberspace, legal, korea's, regulations, health, surveillance, rules, concepts, including, traditional, present, ethics, difficult, companies*
12	*theory, cyberspace, decision, making, decisions, policy, process, been, using, was, explain, research, our, however, analysis, problem, concepts, key, model, issues*
13	*what, cyberterrorism, second, three, questions, first, two, analysis, traditional, between, threat, into, current, then, debate, concludes, strategies, literature, particularly, research*
14	*attacks, against, terrorism, critical, infrastructure, terrorist, groups, actions, different, various, networks, recent, other, attack, computers, components, also, part, infrastructures, conflicts*
15	*china, global, cooperation, united, power, economic, issues, policy, governance, china's, chinese, cyberspace, foreign, government, countries, among, order, context, russia, between*
16	*operations, capabilities, military, network, offensive, capability, support, approach, operational, use, defensive, terms, cyberspace, than, ability, concepts, defence, forces, attack, create*
17	*cybersecurity, development, strategic, also, city, countries, societies, smart, national, technology, need, through, resilience, yet, become, most, one, been, out, world*
18	*terrorist, use, world, legal, online, organizations, they, those, terrorists, who, activities, even, provides, through, violent, violence, all, governments, some, cyberspace*
19	*network, management, access, due, data, several, distributed, ddos, control, malicious, method, key, many, applications, users, based, system, services, number, service*
20	*data, systems, system, detection, intelligence, our, using, used, was, results, research, within, sources, intrusion, technologies, threat, knowledge, networked, process, time*
21	*public, media, social, between, democracy, governance, citizens, opinion, government, they, communication, study, how, show, was, content, were, regimes, about, authoritarian*
22	*politics, human, social, world, power, how, theory, virtual, environmental, what, movement, diplomacy, life, one, transnational, ways, real, about, relations, globalization*
23	*nuclear, most, many, two, important, but, between, weapons, when, what, become, state, into, yet, other, one, actors, against, iran, threat*
24	*attacks, attack, also, attacker, damage, traditional, future, power, particular, infrastructure, scenario, against, possible, cyberattacks, field, few, case, domain, platform, target*
25	*military, war, force, strategic, conflict, national, traditional, defense, e.g, than, cyberwar, cultural, concept, been, but, appropriate, resources, defend, kinetic, one*
26	*digital, crime, nato, science, prevention, forensic, region, including, strategic, legal, other, environment, through, part, national, challenges, area, energy, cybercrime, specific*
27	*malware, how, terrorism, used, organizations, analyze, use, software, problem, known, analysis, also, terrorist, set, better, files, about, there, structure, organizational*
28	*war, society, russian, world, what, 2007, russia, propaganda, conflict, use, western, ukraine, since, hybrid, against, military, order, conventional, itself, around*
29	*between, participation, study, public, activities, sphere, both, significant, online, culture, found, there, relationship, were, sample, procedural, findings, participants, effects, justice*
30	*technologies, technology, networks, military, over, but, may, computer, much, global, time, network, communications, than, less, only, term, control, loss, could*

Table 4. Results for 30-topic LDA model obtained from nearly 600 articles constituting the body of work in the cyber security and conflict field.

In order to assess the cyber conflict field and the place of terrorism and political violence language within it in a more granular fashion, Table 4 presents the results of a 30-topic LDA model obtained from nearly 600 articles constituting the body of work in the cyber security and conflict field. Again, though a large number of topic model sizes were explored, 12-topic and 30-topic models were chosen for presentation purposes as the correct standard for capturing different levels of semantic meaningfulness. Remarkably, the 30-topic model presents a relatively cohesive set of thematic research areas in the literature on cyber conflict. While there remain a broad number of topics constituted of relatively general terminology, there are distinguishing features between each. Several are clearly defined by empirical prompts in cyber conflict studies, such as the 2007 Russia-Estonia conflict,[34] the Olympic Games campaign against Iran's nuclear weapons development program,[35] and cyber tensions on the Korean peninsula.[36] In essence, these areas of the field are common case studies that dominate in research. Some clearly suggest the shape of work on global Internet governance (and governance disagreements), while yet others outline important organizational (i.e., the development of NATO cyber forensics and defensive capacity) and technical topics.

Where the relatively granular breakdown of the terrorism studies field in the above section included a very limited interface with ICT-defined themes, the breakdown of the cyber field includes four topics that are distinct from one another and focused on

[34] See P. Finn, "Cyber Assaults on Estonia Typify a New Battle Tactic," *Washington Post* May 19, 2007; R. Ottis, "Analysis of the 2007 Cyberattacks against Estonia from the Information Warfare Perspective," in *Proceedings of the 7th European Conference on Information Warfare*, 2008, 163; and A. L. Russell, *Cyber Blockades*, Washington, DC: Georgetown University Press, 2014.

[35] See A. Matrosov et al., "Stuxnet under the Microscope," *ESET LLC*, September 2010; J. R. Lindsay, "Stuxnet and the Limits of Cyber Warfare," *Security Studies* 22, no. 3 (2013): 365–404; D. Albright et al., *Did Stuxnet Take Out 1,000 Centrifuges at the Natanz Enrichment Plant?* Institute for Science and International Security, 2010; R. Langner, "To Kill a Centrifuge: A Technical Analysis of What Stuxnet's Creators Tried to Achieve," Hamburg: Langner Group, 2013; and K. Zetter, "How Digital Detectives Deciphered Stuxnet, the Most Menacing Malware in History," *Wired Threat Level Blog*, July 11, 2011, http://www.wired.com/ threatlevel/2011/07/how- digital- detectives-deciphered- stuxnet.

[36] See *inter alia* J. Jun et al., *North Korea's Cyber Operations: Strategy and Responses*, Rowman & Littlefield, 2016; and C. Whyte, "Ending Cyber Coercion: Computer Network Attack, Exploitation, and the Case of North Korea," *Comparative Strategy*, 2016.

terrorism/political violence. Topics 13 and 14 suggest that shape of research engages, respectively, with definitional questions on the nature of cyberterrorism and questions of the potential impact of cyberterrorist activities. Topic 18 appears to deal with terrorism from a law and governance perspective, with both the keyword list presented in Table 4 and the extended list (available in the appendix) relating the terrorist enterprise and terrorists' use of the web for organizational purposes to government oversight. Finally, Topic 27 appears to more directly deal with the integration of cyber tools and techniques into terrorist organizations.

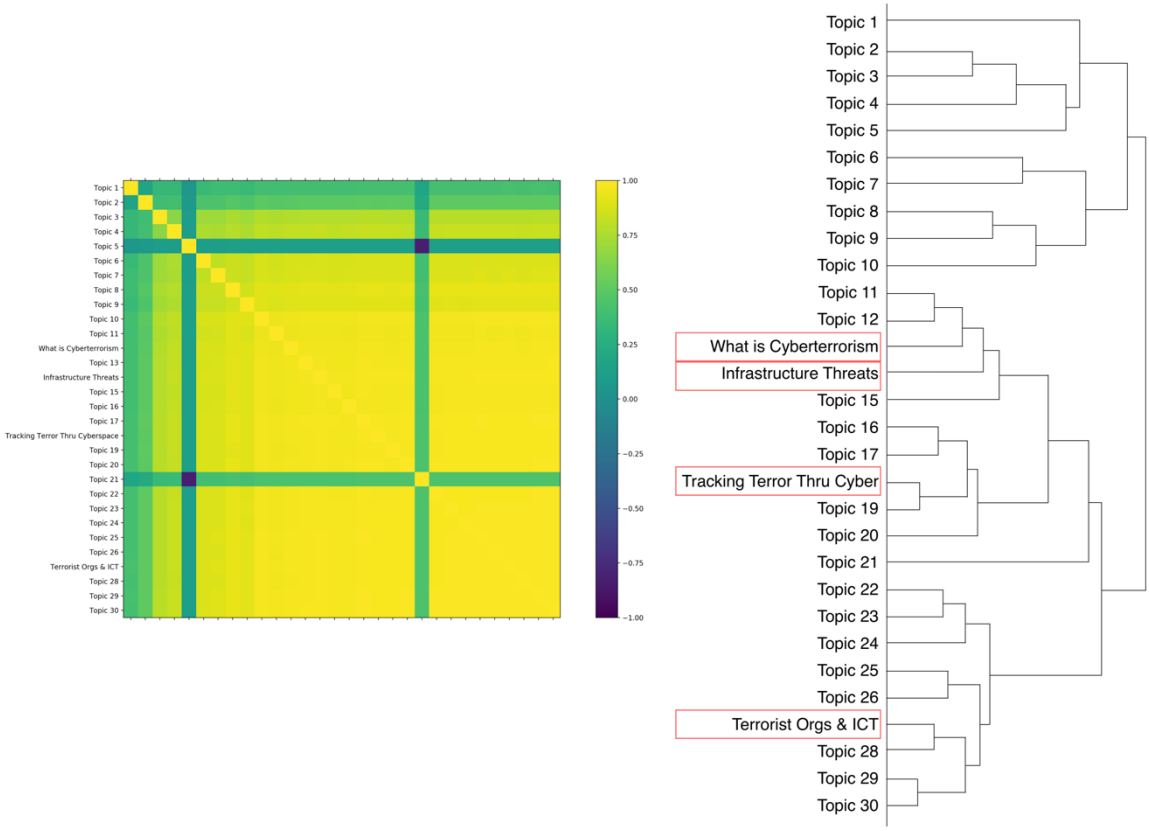

Figure 6. Correlation matrix and dendrogram visualization for Table 4's 30-topic LDA model.

Even without further visualization, the expanded model of cyber conflict research suggests that topics at the intersection of ICT and terrorism studies already constitute a distinct set of research agendas within the field. The question, however, remains as to the degree to which such themes are present in other parts of the cyber conflict field. Indeed,

given that two of the terrorist-oriented topics appear to describe issues broadly relevant to other cyber research efforts in some granularity (i.e., the incorporation of new techniques into legacy organizational structures and the legal-governance basis of a form of cyber conflict), one might expect these topics to co-occur broadly across the field in thematically significant ways.

As Figure 6 demonstrates, however, there is mixed evidence to support such a hypothesis. It is certainly the case that each digital topic co-occurs broadly in much the way the lone "digital issues" topic did in the terrorism studies literature. Much as was the case in the section above, however, there is no evidence of significant major similarity with other themes in the correlation analysis. The story is somewhat different as told by the agglomerative hierarchical clustering visualization on the right in Figure 6. Here, it appears that the definitional topics (i.e., what is cyberterrorism and how impactful might it end up being?) are reasonably similar to a number of other major themes. While the topics appear in small proportions frequently across the entire corpus, it also seems to be the case that they share vocabulary with a number of other topics. This is perhaps unsurprising given that Topics 13 and 14 deal with definitional issues being grappled with across the entire cyber conflict field of study, from the prospective risk nature of different modes of cyber threat to the applicability of different conceptual frameworks. Topics 18 and 27, on the other hand, exhibit distinctly fewer topical similarities with other areas of the cyber conflict field. Much as was the case with the "digital issues" element of the terrorism studies field, these topics are empirically distinct. And, much as was the case with the terrorism studies literature, these empirically distinct topics present as a distinct sub-theme of the literature that are referenced in small part across a wide swathe of research works but that have few major thematic connections to other topics.

The Scope and Impact of Digital Research in the Field

Clearly, the terrorism studies and cyber conflict bodies of work assessed above constitute, when taken together, the broad horizon of research that might analytically link the information revolution with distinct topics in terrorism, political violence, and non-state conflict. What the analysis above does not speak to, however, is the temporal nature of thematic developments in the broader combined fields of cyber conflict and terrorism studies. Though modern information technologies have their roots in innovations made in the 1960s and cyber conflict has occurred for at least three decades, the global expansion and continued adoption of ICTs at the level of fundamental societal functions has naturally had a cumulative motivational effect on scholars of world affairs. In other words, many scholars have only recently turned to consider the effects of the information revolution on traditional security topics, on the one hand, and the development of new and distinct socio-technical phenomena, on the other. This section thus turns to an alternative form of topic modeling, dynamic topic modeling, in order to consider topical development in scholarship over time. In doing so, we are qualitatively presented with the opportunity to adjudicate on apparent propellants of new topic formation in scholarship.

Dynamic topic modeling alters a simple assumption made in the processing of obtaining traditional topic models, essentially that relationships between documents don't matter. This is a reasonable assumption from various points of view, as many inter-document relationships are relatively meaningless for the task of discovering latent thematic trends in a corpus. If a researcher wants to better understand topic formation around particular characteristics of individual documents, they may undertake semi-supervised topic modeling (as is done in the section below this) to note a particular unifying attribute of *some* documents over others. However, the order of documents still does not matter.

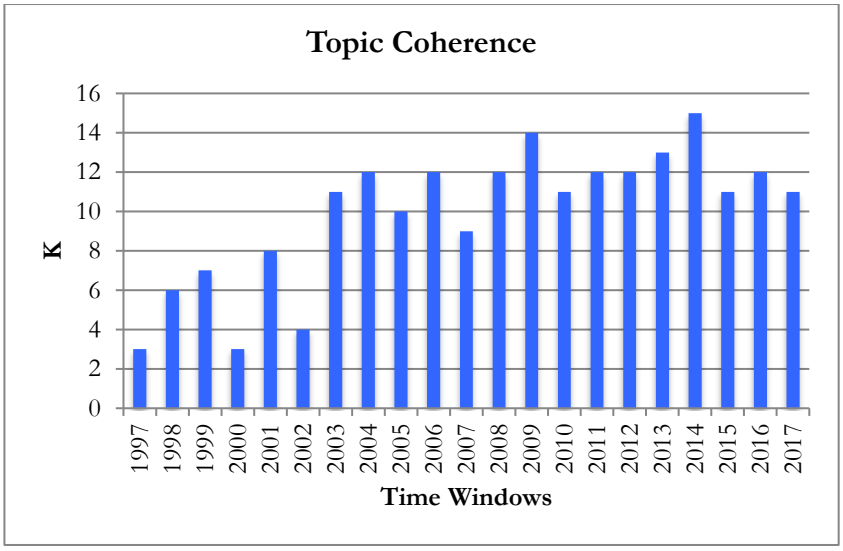

Figure 7. Topic coherence outputs of a Skipgram Word2Vec model obtained from a corpus of nearly 5,000 article abstracts drawn from Web of Science.

Naturally, any attempt to consider latent themes over time benefits from an ability to reference the order in which documents appear. Dynamic topic modeling allows researchers to do just that by obtaining topics in a two-step process. First, a series of topic models are produced based on the order of documents in a given corpus. These "window" topics, when the aim is to look at topics over time, reflect thematic trends within a given period of time, such as months, years, or decades. They also constitute the set vocabulary via which a broader set of dynamic topics can be generated that describes macro topics for the entire period of time under study. The result is an ability to study the formation of major themes in a corpus via dissection of contributing topics that present from time period to time period.

Figure 7 above presents the topic coherence outputs of a skipgram Word2Vec model[37] obtained from the combined corpus of nearly 5,000 article abstracts drawn from Web of Science related to terrorism studies and cyber conflict. Though not necessary, it is possible to

[37] For description of the approach, see Winnie Cheng, Chris Greaves, and Martin Warren. "From n-gram to skipgram to concgram." *International journal of corpus linguistics* 11, no. 4 (2006): 411-433; and David Guthrie, Ben Allison, Wei Liu, Louise Guthrie, and Yorick Wilks. "A closer look at skip-gram modelling." In *Proceedings of the 5th international Conference on Language Resources and Evaluation (LREC-2006)*, pp. 1-4, 2006.

relinquish further control over the parameter specifications needed to obtain topic models with dynamic topic modeling. In producing a skipgram model, it is possible to obtain a measure of topic coherence for each window specified by the researcher. In essence, by setting broad limits on the number of topics that the researcher might be interested in finding, it is possible to obtain a measure of cohesiveness of different term clusters. This can then be used in lieu of an otherwise arbitrary number of requested topics chosen by the researcher in the topic model generation process at the window topic stage. For our purposes, the output of the skipgram word2vec model is interesting because it signals the diversification of the combined literature under study. In particular, topic coherence (i.e., identifiable, semantically meaningful topic areas in the field) appears to have boomed in the early 2000s.

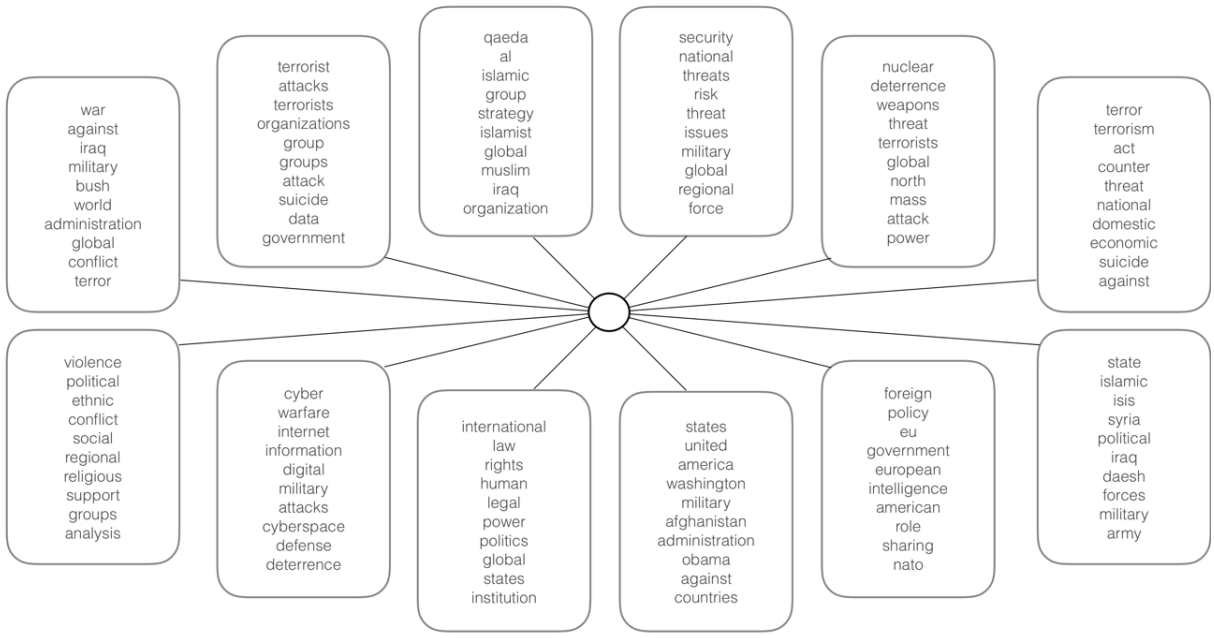

Figure 8. Results for a 12-topic non-negative matrix factorization dynamic topic model obtained from nearly 5,000 article abstracts drawn from Web of Science.

Figure 8 arrays the dynamic topics produced via initial assessment of window topics across a 21-year period from 1997 to 2017 (prior to 1997, literature did not present as topically coherent enough to model with the dynamic approach). There exist a clear set of

distinct issue areas. These include: (1) the War on Terror, (2) the shape of terrorism, (3) global jihadism and radicalization, (4) militaries and the use of force, (5) nuclear proliferation and terrorism, (6) counterterrorism, (7) ethnic violence and conflict, (8) cyber conflict and warfare, (9) international law and humanitarian issues, (10) American operations in Iraq and Afghanistan, (11) transatlantic security coordination, and (12) the Syrian/Islamic State conflicts. Naturally, for the purposes of this study, Topic 8, cyber conflict and warfare, is of prime interest.

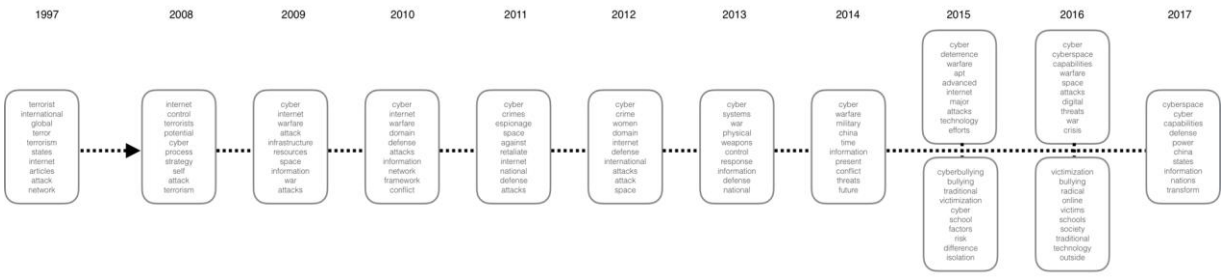

Figure 9. Window topic results for Topic 8 of Figure 8's 12-topic non-negative matrix factorization dynamic topic model.

Topic 8 quite naturally has its roots in the corpus in initial work on terrorism, broadly construed. This is unsurprising from two potential angles. First, the corpus of documents includes subject matter dominated by terrorism studies in the late 1990s and diversification of the field in the 2000s was largely driven by the events of the September 11th attacks on the United States. Second, early discussions of cyber conflict, or "netwar," in the late 1990s were often linked with the possibility of terrorist attack.[38] This was the case for a few reasons, primarily that Clinton-era efforts to address cyber security had their roots in non-state and semi-state security incidents. These included the Oklahoma City bombings that defined Clinton's early presidency and the later events of Moonlight Maze and SOLAR

[38] For narrative descriptions of the period, see *inter alia* R. A. Clarke and R. K. Knake. *Cyber War*, Tantor Media, Incorporated, 2014; and Jason Healey, ed. *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*, Cyber Conflict Studies Association, 2013.

SUNRISE[39] (wherein the United States government was the target of espionage operations, respectively, by state-linked and independent hackers) that drove policymakers to act on cyber issues.

Interestingly, diversification of the cyber conflict topic does not become distinct until 2008 when the language of cyber warfare studies appears to diverge (given the systematic absence of previously prominent keywords) from that of terrorism studies. In the period between 2008 and 2014, window topics most closely linked with the overall dynamic topic illustrate more distinct thinking on the nature of cyberspace as a unique domain akin and related to space. In 2015, the dynamic topic diverges insofar as it is most closely linked with distinct topics focusing on cyber conflict and cyber security as a social issue. Though the appearance of keywords like deterrence do suggest a diversification of research focus within the cyber security field in the social sciences, the differentiation made between such thematic work and focus on cyberbullying and victimization illustrates a lack of diversity among cohesive topics in what IR scholars would traditionally consider digital security issues. More to the point here, the cyber conflict field itself does not present as interfacing with terrorism and political violence studies terms following early ruminations.

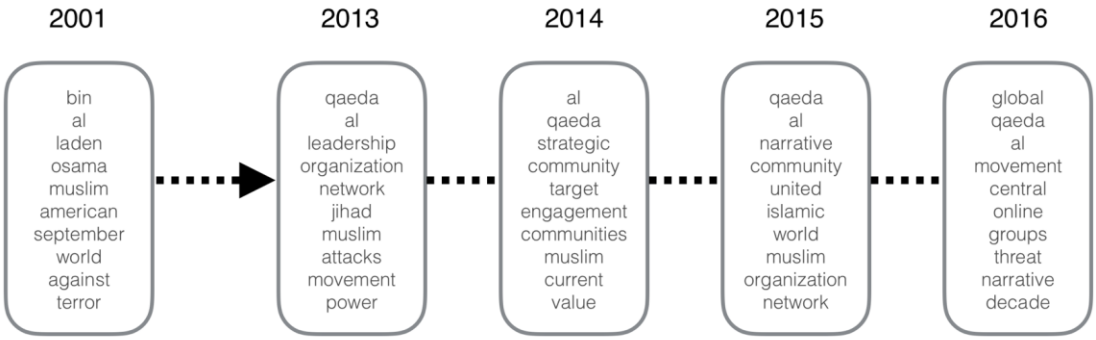

Figure 10. Window topic results for Topic 3 of Figure 8's 12-topic non-negative matrix factorization dynamic topic model.

[39] For perhaps the best in-depth descriptions of these events see F. Kaplan, *Dark Territory: The Secret History of Cyber War*. New York: Simon and Schuster, 2016.

Two other topics include keywords that smack of digital issues moving into the core of distinct thematic issue areas. Figure 10 demonstrates how window topics linked with the broader dynamic topic describing global jihad have shifted in focus over time in exhibiting community over organizational nomenclature. Where topics through 2013 rarely vary in key term usage beyond keywords like attacks, terrorize, organization, transnational, and network, core window topics from 2014 onwards reflect a shift in scholarly discourse focused on jihadi narrative and radicalization among global Muslim communities. In 2016, the word online was uniquely linked to the global jihad window topic, while the words digital, forum, and website appear in more expansive top twenty word lists in 2015 and 2016.

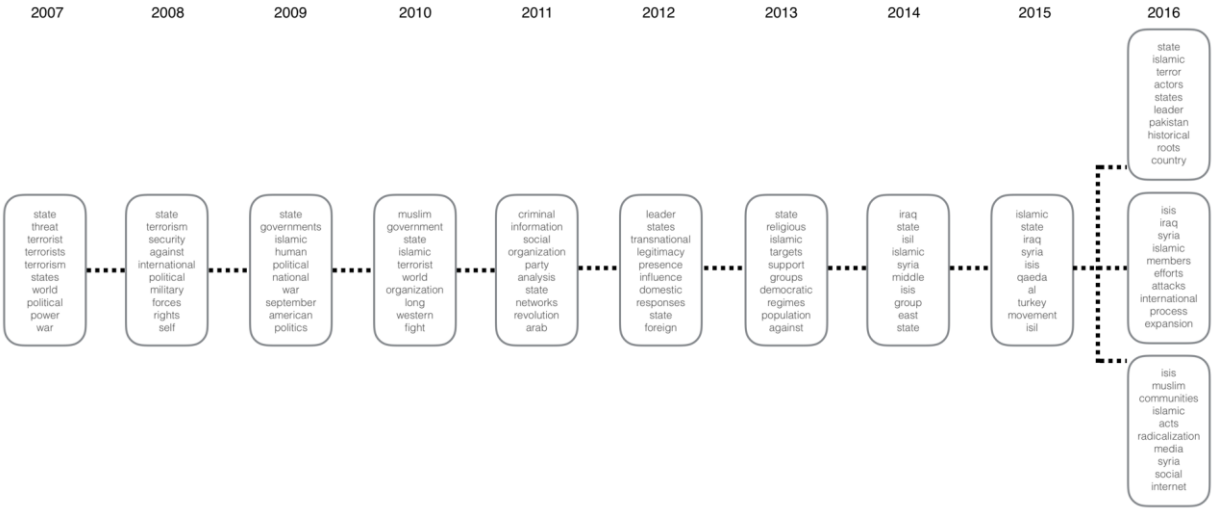

Figure 11. Window topic results for Topic 12 of Figure 8's 12-topic non-negative matrix factorization dynamic topic model.

Likewise, as Figure 11 shows, thematic rhetoric in the dynamic topic describing the Syrian and Islamic State conflicts has had a unique evolutionary trajectory from broader thematic topics that simply describe the shape of the terrorist enterprise. The dynamic topic appears to be defined by the intersection of discourse on specific conflicts, including those wars and the Arab Spring, and tensions between the Islamic and Western worlds more generally. In 2016, several window topics contributed to the construction of the dynamic

topic, seemingly describing the threefold focus on the Islamic State as a territorial adversary, a terrorist organization, and a community element in the Muslim world. With this last topic, the interlinkages between the most recent Syria-Iraq set of conflicts and Muslim communities via digital media are quite apparent.

Discussion

Topic modeling presents a unique opportunity for assessing the latent thematic content of large bodies of scholarly work. In presenting a range of evidence derived from diverse application of topic models to scholarship on terrorism and cyber conflict studies, this article demonstrates that the cyber terrorism subfield (spread across the fields of cyber conflict and political violence studies) is *both* nuanced and limited. This section discusses a range of implications of the evidence outlined above.

A lack of Conceptual Precision

Broadly, this study supports and illustrates past arguments that the field, beyond one or two specific issue areas, remains topically imprecise in several ways. First, the use of the term "cyberterrorism" (and variants) is linked with remarkably little in the way of thematic distinction across different fields of study. Perhaps most interestingly, there is a clear lack of analytic and technical language in the topics resulting from analysis of the "cyberterror" (by keyword) corpus. Whereas we might expect to see distinction in content focusing on the tools of cyberterrorism as either being ancillary or linked with major infrastructural threats to national security, little beyond general terminology defines most topics produced. This suggests that reference to cyberterrorism is often a qualifying crutch in research more broadly on cyber or terrorism issues and that foundational definitional debate remains the bread and butter of much fuller scholarly engagement. Second, ICT wordage appears only sparingly in high-level topic model results obtained from both the cyber conflict and terrorism bodies of work. Where it does appear, such wordage is found alongside other non-traditional

terminology, such as "biological." This strongly indicates that the state of the field is yet one of relative novelty.

Targeted Impact, Common Reference

More specifically, distinct semantic themes that describe digital issues areas of study in the literature are referenced in small part across a wide swathe of research work but tend to have few major thematic connections to other topics. This supports the narrative that cyberterrorism, cyber-jihad and various other digital issue terminologies are referenced as a qualifier often in research across the broader field(s) without meaningful conceptual engagement. This is not meant as a critique, as a central criticism of scholars focused on cyber terrorism subject matter is that such a lack of conceptual precision and consensus exists. Nevertheless, it seems clear that such a shortcoming has manifested in limited engagement by scholars focused on other parts of the terrorist enterprise or other conflict issues. "Cyber security" and Internet-based terror are of unique interest, but they are analytically difficult to break apart and problematize.

Event Driven

The field appears in analysis to be limited in another manner. Topic diversity often meets semantic meaningfulness in a topic's primary reference to specific events and conflicts. This is not surprising, to some degree, and even the broader terrorism studies takes reference at the highest level from defining events in recent history, such as the events of September 11th, 2001, and the resulting War on Terror. With cyber terrorism subject matter, however, granular assessment of the field reveals that disciplinarily unique themes (Islamic State military operations versus radicalization processes versus transnational terrorist network logistics, for instance) reference just a few conflicts. Specifically, much work in the field seems to focus on Israeli-Palestinian tensions, Islamic State and the Syrian conflict, and related terrorist activities around the world. While, again, it is not surprising that scholars

might focus on such prominent subject matter, the focus on just a few case issues belies the potential scope of cyber terrorist activities and further reinforces the notion that conceptual divisions in the field over definitions present a barrier to empirical investigation of cyber antagonism by nonstate actors in general.

Missing Techno-Analytic Engagement

In observing the results in sections above with an eye to what is missing, it seems strange that there is such a lack of analytic and technical language in topics. Specifically, it is unusual to see topic model results obtained from a corpus that represents a relatively cohesive subfield of study that contain no discrete analytic themes. Just as the topic model of the broader terrorism field above reflect several categories that describe the tools and terminology of scholars' approach to research, we might expect such terminology, whether technical or methodological, to appear among digital issues when bound as a subfield. This is not the case. Specifically, there appears to be almost no generalized inclusion of terminology linked with computer network attack or defense, such as "ddos," "malware," "rootkit," "virus," etc. While there is minimal evidence of such in the topic models produced from the narrow "cyberterrorism" keyword corpus, the same is not true for the broader field. Moreover, qualitative analysis of the "cyberterrorism" corpus suggests that such technical assessment occurs outside of the social sciences, where threat of digital terrorism is used to qualify the security implications thereof.

Missing Topical Focus

Beyond missing techno-analytic language, there are clear topical absences in the latent thematic areas defined in the figures above. In particular, there is surprising omission of content that might fall within the political violence field. Engagements on topics in digital activism and hacktivism, for instance, are curiously absent as themes in the broader corpus despite an arguable conceptual link with "cyberterrorism." After all, cyberattacks are not

inherently violent, even if aggressive,[40] and a broad set of works in the cyber conflict field have noted the conceptual similarities between digital terrorism and other forms of coercion online, like hacktivism. It seems reasonable to assert that the continued expansion and deepening of work in the research program on nonstate cyber conflict would thus benefit from new focus on topics not seen in the results above, such as (among others):

- Cyber threats to civil society organizations (CSOs);
- Nonstate CSO cyber conflict, including:
    - Cyber-enabled subversion and propaganda;
    - Cyber coercion by non-terrorist actors;
    - Cyber-augmented criminal advocacy (e.g. "naming and shaming" operations undertaken by social movements).
- Foreign-focused cyber mercenary operations;
- The use of ICT in state sponsorship of terrorism;
- Augmentation effects of ICT usage alongside traditional terrorist acts.

Moreover, scholars would do well to employ theoretical frameworks common to the political violence field in engaging on such topics. In many cases, it is likely that the above topics are not understudied or completely absent. Rather, they are found as objects of study in other academic fields where there is simply no thematic engagement with work in the terrorism studies field.

Conclusion

This article has provided empirical support for recent arguments that the focus on the intersection of digital and terrorism studies lacks analytic precision and diversity on several

[40] For those who argue that cyber weaponry is nonviolent in nature, see *inter alia* T. Rid, "Cyber War Will Not Take Place," *Journal of Strategic Studies* 35, no. 1 (2012): 5–32; and E. Gartzke, "The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth," *International Security* 38, no. 2 (2013): 41–73.

fronts. Given the evidence above, it seems reasonable to suggest that continued definitional squabbles and focus hamper the development of research on the Internet, political violence, radicalization, and cyber conflict. As such, I echo the call made by other scholars in this domain over the past two years to redouble efforts to add conceptual nuance to the field enabled by smarter development of empirical data collection and testing approaches.

Specifically, I argue that the best approach for achieving such conceptual and empirical diversity would be the adoption of a program of micro-foundational compartmentalization at the project level. By adopting less restrictive conceptual assumptions about the traditional bounds of extremism, political violence, and the terrorist enterprise, scholars can work to develop new understanding of how ICTs have impacted threat actors and processes. In particular, scholars should expand the range of activities that might be considered to constitute political violence or radicalization efforts. In doing so, they will inevitably expand focus on actors clearly not often considered in contemporary efforts to problematize conflict in the digital age, including subversive organizations, hacktivist outfits, organized criminal entities, and non-jihadi extremists. Likewise, researchers would do well to augment existing methodological strengths with new data production schemes. The information revolution has not only been immensely impactful on the subjects of study for the terrorism study field; it has also provided a host of new tools for assessing and operationalizing antagonistic, conflictual behavior. Finally, scholars in this domain should actively look to cross interdisciplinary boundaries in considering the socio-technical design side of issues (from the design of malware to the shape of relevant parts of the Darkweb ecosystem) and collaborating with those in the computational sciences. In doing so, future assessments of the shape of this subfield will hopefully find a more diverse, expansive, and thematically precise set of research programs.