(RUSSIAN) DETERRENCE, WE HARDLY KN*O*W YE…

STEPHAN DE SPIEGELREIRE, YAR BATOH, DORIAGORIACHEVA, GLIB VO-LOSKYI, KHRYSTYNA HOLYNSKA, ANASTASIIA VOZYVYCH, BOGDAN BALIUK, HRYHORII PAVLENKO, YEVHEN SAPOLOVYCH, ANASTASIIA SHCHEPINA – HCSS AND ANDREAS HEINER

VOL. II

OF THE PUBLICATION:

PENTTI FORSSTRÖM (ed.): **RUSSIAN CONCEPT OF DETERRENCE IN CONTEMPORARY AND CLASSIC PERSPECTIVE**

MAANPUOLUSTUSKORKEAKOULU
HELSINKI 2021

Pentti Forsström (ed.): *Russian concept of deterrence in contemporary and classic perspective*
Maanpuolustuskorkeakoulu
Sotataidon laitos
Julkaisusarja 2: Tutkimusselosteita nro 11

National Defence University
Department of Warfare
Series 2: Research Reports no. 11

Editor:

Pentti Forsström

Writers:

Stephan De Spiegeleire, Yar Batoh, Daria Goriacheva, Glib Voloskyi, Khrystyna Holynska, Anastasiia Vozovych, Bohdan Baliuk, Hryhorii Pavlenko, Yevhen Sapolovych, Anastasiia Shchepina – HCSS and Andreas Heiner – Trimes Oy, Finland

Recent publications in PDF format: http://www.doria.fi/handle/10024/73990

Cover image: FNDU

Pictures and illustrations: Authors

Maanpuolustuskorkeakoulu – Sotataidon laitos
National Defence University – Department of Warfare

(RUSSIAN) DETERRENCE, WE HARDLY KN*O*W YE…[1]

Stephan De Spiegeleire, Yar Batoh, Daria Goriacheva, Glib Voloskyi, Khrystyna Holynska, Anastasiia Vozovych, Bohdan Baliuk, Hryhorii Pavlenko, Yevhen Sapolovych, Anastasiia Shchepina – HCSS[2] and Andreas Heiner – Trimes Oy, Finland

Paper prepared for the Finnish National Defence University's Russia Seminar 2021.

This is the Volume II of the publication: Pentti Forsström (ed.): National Defence University, Department of Warfare Series 2: Research Report No. 11: Russian Concept of Deterrence in Contemporary and Classic Perspective, 2021, (https://urn.fi/URN:ISBN:978-951-25-3250-6).

1. INTRODUCTION

Deterrence is back on the international political agenda. The reappearance of geopolitical competition and great power brinkmanship has rekindled interest in the theory and practice of deterrence. Deterrence has also returned as a guiding concept in the strategic postures of major – and even not-so-major – military powers. It is central to NATO's efforts to meet Russia's resurgence; it permeates Russia's military (and other) efforts to hold off what it sees as a revisionist West; it remains a cornerstone of US grand strategy; and it is part and parcel of the doctrines and behaviors of emerging great powers like China and India. Last but not least, the need for (often massively) increased funding for deterrent capability options is now once again widely shared as a self-evident 'truth'.

But what do we actually know about deterrence? Do scholars and practitioners, amongst and across themselves, really share a common understanding of the term? Do we know whether deterrence 'works', and under which circumstances? When it comes to the use of deterrence by a country with, as many argue[3], a very different strategic culture like Russia – how confident are we that we have really done our homework on fully comprehending it? That we have come up,

---

[1] The title refers to the English saying that is sometimes used (in the past tense) to deplore insufficient knowledge (and/or interest) about somebody or someone that is no longer there: "we hardly knew ye", using the Middle English version of the personal pronoun you that is also still known from expressions like "gather ye rosebuds, while ye may". The origins of "we hardly knew ye" are unclear, but many refer to the Irish folk song "Johnny, we hardly knew ye" Wikipedia, "Johnny We Hardly Knew Ye," in *Wikipedia*, March 25, 2020, https://en.wikipedia.org/w/index.php?title=Johnny_We_Hardly_Knew_Ye&oldid=947351758. Since deterrence is – for better and/or (?) for worse – still very much among us, we changed the verb to the present tense.

[2] The authors gratefully acknowledge the financial support from the United States Defense Department's Minerva Research Initiative and the Carnegie Corporation of New York. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

[3] D. Adamsky, "From Moscow with Coercion: Russian Deterrence Theory and Strategic Culture," *Journal of Strategic Studies*, 2018, https://doi.org/10/gdxs6v; Stephen R. Covington, "The Culture of Strategic Thought Behind Russia's Modern Approaches to Warfare," Paper (Harvard Kennedy School, Belfer Center for Science and International Affairs, October 2016).

in genuinely open-minded ways, with rigorous *and* creative hypotheses to explain Russian deterrent thinking and behavior? That we have systematically amassed and parsed all publicly[4] available cues to find out which of these hypotheses appear better supported by the evidence?

This paper sets out to clinically examine the publicly available scholarship on these issues in a systematic way. After an introductory 'scene-setting' section follow two sections that attempt to provide an epistemic equivalent to an MRI scan for the field(s) of deterrence studies. The first one of those two consists of a more '*technical*' bibliometric analysis of the scholarly literature that embodies our (public) understanding of deterrence. It addresses questions like how much scholarly literature has been produced, how well is it being used and debated, how quickly and effectively knowledge is generated and propagated, how collaborative the field is, how thorough it is, etc. The second of the two main sections examines the actual *substance* of this literature using a variety of both older and more recent tools to map the epistemic landscape that is contained in the literature: how is deterrence defined in these Russian and Western scholarly publications, what are the key terms and topics, how have they changed over time, which ones are particularly popular today, etc. The paper concludes with some final conclusions and recommendations for both the demand and the supply side of knowledge on deterrence in an international security context.

2. SETTING THE SCENE

Before diving into our actual analysis of the field of deterrence, we feel that three scene-setting thoughts may prove useful for our readers to better understand some of the atypical choices that were made in this paper.

Different layers of deterrence research

In the field of international security the usage of the term 'deterrence' is automatically assumed to refer to the theory and practice of international (mostly state but in recent years also non-state) actors deterring each other. Most literature reviews, therefore, focus exclusively on the international security literature. In our own ongoing work on deterrence, of which this paper is one early result, we differentiate between 4 'Russian doll'-like layers of deterrence research. The first one of those is the broad inquiry into deterrence as a general phenomenon. Alongside the literature on deterrence in the field of international security, this layer includes other social science disciplines like criminology, health policy, economics, business management; but also numerous 'hard' sciences like ecology, agronomy, safety research, biochemistry, horticulture, and even primatology, many of which also contain deep (and often fascinating) pockets of knowledge that have barely been explored by scholars working in the field of international security. In this paper, we

[4] This paper is based exclusively on unclassified sources.

will denominate this most expansive layer of deterrence research as '***deterrence-broad***'[5]. The second layer is the subset of that broader literature that deals specifically with deterrence in the context of human agency (***deterrence-human***): human individuals or groupings trying to deter others from pursuing certain unwanted courses of action like engaging in crime, smoking cigarettes, bullying other students, entering into a market, invading other countries etc. This subset includes writings on deterrence in disciplines like political science, but also sociology, anthropology, law, health policy, policing, education, history, psychiatry, etc. The third layer is a further subset of the previous subset that readers of this paper will feel more familiar with because it hails from the field of international relations[6] and international security (***deterrence-IS***). The literature on the individual theory and practice of international[7] deterrence in one specific country – in this case, Russia – or a group of countries is once again a subset of the deterrence-IS layer of analysis (***deterrence-IS/Russia***).

[5] The only attempt we are aware of to draw some (still very shallow) policy lessons from this deterrence-broad literature for the assessment of the usefulness of international security deterrence is our own Stephan De Spiegeleire et al., *Reimagining Deterrence: Towards Strategic (Dis)Suasion Design* (The Hague, The Netherlands: The Hague Centre for Strategic Studies (HCSS), 2020). We remain convinced that a deeper understanding of the micro-, meso- and macro-dynamics of deterrence even in entirely different contexts like predator-prey relationships or ovipositional deterrents could still shed more light on this phenomenon.

[6] Also the international political economy subfield of international relations contains some writings on certain forms of dissuasion/deterrence like sanctions, etc. The best overview is probably David A. Baldwin, *Economic Statecraft*, New edition (Princeton: Princeton University Press, 2020), not in the least because he also put sanctions against the context of other forms of – in his case – *economic* statecraft, which is something that has still not been done with anything even approaching the same amount of rigor in the international *security* field, let alone across the entire broader '*external action*'-portfolio. For an early attempt, however, see Stephan De Spiegeleire et al., *Reimagining Deterrence: Towards Strategic (Dis)Suasion Design* (The Hague, The Netherlands: The Hague Centre for Strategic Studies (HCSS), 2020).

[7] We want to point out that the study of Russian 'domestic' deterrence is typically also excluded from examination, even though the purposive strategic manipulation of fear in families, universities, on work floors, in the criminal justice system, etc. is arguably still far more prevalent in 'modern' Russia than in most liberal (and increasingly 'post-modern' - on the distinction, see Robert Cooper, *The Post-Modern State and the World Order*, Paper / Demos, no. 19 (London: Demos, 1996) democracies. Examining these (in certain cases empirically better investigated) manifestations of Russia's domestic intimidation and deterrence might also provide interesting insights into Russia's approach to international deterrence (and maybe even in the extent to which domestic and international approaches to deterrence may be interconnected.

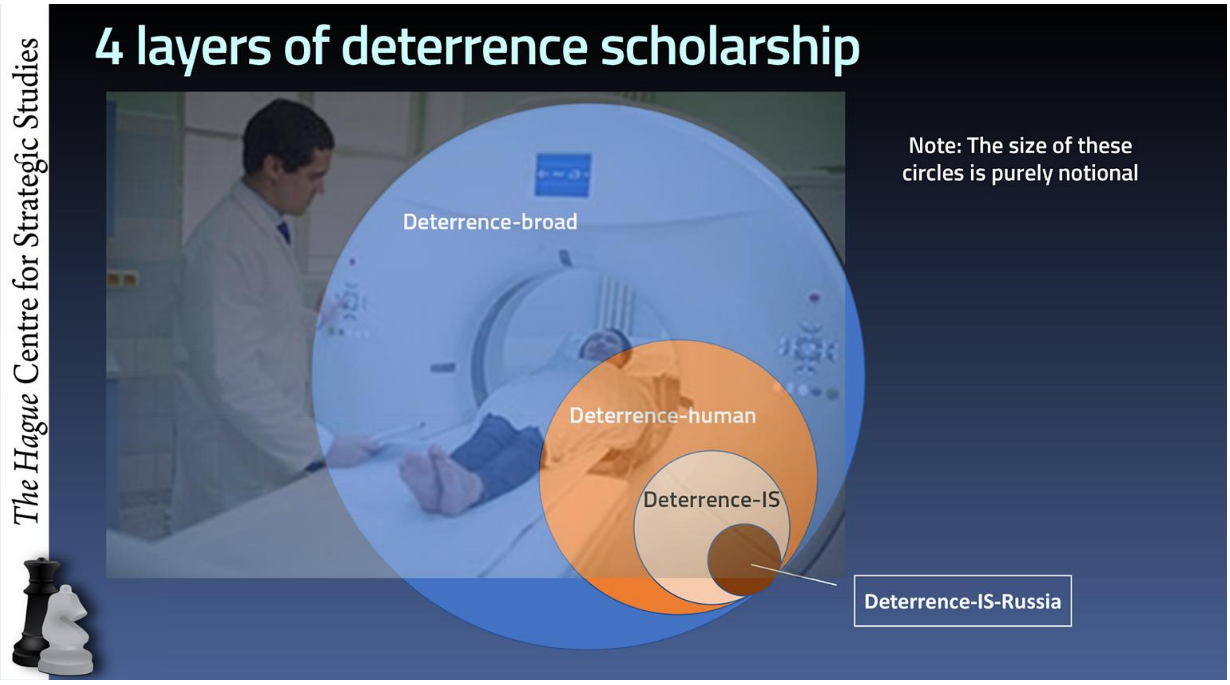


Figure 1: Four layers of deterrence scholarship

In most academic disciplines there is an inherent tension between these different layers/levels of aggregation. Most linguists working on the Russian language, for instance, probably spend most of their time focusing on the literature on that language. But they would be unlikely to shy away from also looking at new insights for the East-Slavic or the Slavic or the Indo-European language groups or indeed the broader field of linguistics, and especially computational linguistics – where specialists in Natural Language Processing are now making more progress in one year on almost all languages of the world than scholars working on any one language (group) have made in decades[8]. Area studies in general also have always struggled with the tension between the region-specific and the broader comparative or theoretical layers[9]. One of our starting points for this paper was that we at least wanted to explore insights from the other layers of deterrence research as well[10].

[8]A recent Google model, for instance, extended Google's (currently) most powerful T5 (Text-To-Text Transfer Transformer) language model, pre-trained on a new massive open-source dataset called the Colossal Clean Crawled Corpus (C4), into a multilingual setting whereby it obtained quality improvements across 101 (!) languages.
Linting Xue et al., "Mt5: A Massively Multilingual Pre-Trained Text-to-Text Transformer," *ArXiv:2010.11934 [Cs]*, March 11, 2021, http://arxiv.org/abs/2010.11934; William Fedus, Barret Zoph, and Noam Shazeer, "Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity," *ArXiv:2101.03961 [Cs]*, January 11, 2021, http://arxiv.org/abs/2101.03961.
[9]For an interesting recent treatment of some of these issues, see K. Mielke and A. K. Hornidge, "Area Studies at the Crossroads: Knowledge Production After the Mobility Turn" (books.google.com, 2017), https://books.google.com/books?hl=en&lr=&id=OLE-DgAAQBAJ&oi=fnd&pg=PR8&dq=%22area+studies%22&ots=P4el-y61uT&sig=vByHcbiSsw6Dh9NMs-7T6b6grWM.
[10] We should note that Figure 1 only shows one part of the scientific embedding of our extant knowledge on (Russian) deterrence theory and practice. Analysts working in this field should ideally also be well-versed in different areas of defense analysis (defense strategy, military doctrine, operational planning, targeting, military hardware, modeling, etc.), of political science (international relations, (mis)perceptions, civil-military relations, bureaucratic politics, domestic politics, etc.), and much more. Similar layers could therefore also be envisaged for each of those embeddings of deterrence. But this paper focuses on the knowledge base on the pure 'deterrence'-specific embedding of the (Russian) deterrence-IS literature and

Standing on the shoulders of giants

One of the principal catalysts of scientific progress has always been researchers' willingness and ability to 'stand on the shoulders of giants'[11]. Many will undoubtedly recognize this expression as the motto of Google Scholar[12], to date still (by far) the largest repository of metadata on the world's scholarly literature[13].

As humans, we currently still primarily encode most of what we know about the world around us in written 'natural'[14] language[15]. In this sense, the world's body of scholarly publications embodies the epistemic 'shoulders' on which successive generations of scholars can stand and can, in turn, start providing new shoulders for subsequent generations of knowledge-augmenters. The bad news in all of this is that we as humans are not particularly disciplined or systematic in the way in which we semantically encode our thoughts[16] and then decode them again into (in this paper's context mostly written and poorly structured) natural language. The good news is that homo sapiens has evolutionarily – by and large – learned to decode that transmitted natural language back into concepts its brain can cognitively grasp, absorb, regurgitate, build upon, and then transmit to others. With every paper we read, in every discussion we engage in, we all painfully realize that much gets lost in this iterative encoding and decoding process. But it is equally clear that much communication (and knowledge transfer) still gets through. It, therefore, stands to reason that epistemic progress still primarily requires human scholars to be able to tap into the accumulated knowledge ('Wissensanhäufung') that is embedded in scholarly writings.

therefore use these other onion-like deterrence layers as points of comparison: are there any differences between deterrence-IS and these other broader deterrence-layers, and if so, what do they reveal about the deterrence-IS knowledge layer?

[11]The metaphor of 'dwarfs standing on the shoulders of giants' refers to "[u]sing the understanding gained by major thinkers who have gone before in order to make intellectual progress." C W Trowbridge and Jan K Sykulski, "On the Shoulders of Giants" (2013), https://eprints.soton.ac.uk/355119/1/COMPUMAG%25202013%2520CTW%252BJKS.pdf.

[12]Google, "Google Scholar," 2021, https://scholar.google.com/intl/en/scholar/about.html.

[13] Alberto Martín-Martín et al., "Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A Multidisciplinary Comparison of Coverage via Citations," *Scientometrics*, 2020, 1–36; Holly Else, "How I Scraped Data from Google Scholar," *Nature*, April 11, 2018, https://doi.org/10.1038/d41586-018-04190-5.

[14]The term refers to language that is spoken and written by humans 'in the wild', as opposed to computer languages or other communication systems in nature (like whale songs).

[15]Wikipedia serves as an excellent illustration of this. To future historians, Wikipedia may very well prove to be the last (and best) example of a primarily 'natural language'-encoded source of cumulative information and knowledge; whereas the knowledge graphs that are trying to formalize some of this Wikipedia knowledge in a more computer-friendly language (like Wikibase or Wikidata) may be the harbingers of post-'natural language' knowledge encoding. Industry titans like Amazon, Baidu, Facebook, Google, IBM, Microsoft, Yahoo, Yandex now put knowledge graphs front and center in their respective knowledge building efforts. On knowledge graphs, see Jeff Z. Pan et al., eds., *Exploiting Linked Data and Knowledge Graphs in Large Organisations* (Springer International Publishing, 2017), //www.springer.com/us/book/9783319456522; Jianfeng Du et al., "Validation of Growing Knowledge Graphs by Abductive Text Evidences," *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (July 17, 2019): 2784–91, https://doi.org/10.1609/aaai.v33i01.33012784; Xin Luna Dong and Divesh Srivastava, "Knowledge Curation and Knowledge Fusion: Challenges, Models and Applications" (ACM Press, 2015), 2063–66, https://doi.org/10/gghvvh; Madalina Croitoru and GKR, eds., *Graph Structures for Knowledge Representation and Reasoning*, vol. 5th International Workshop, Gkr 2017, Melbourne, Vic, Australia, August 21, 2017: Revised Selected Papers (Berlin Heidelberg: Springer, 2018); Jun Zhao et al., eds., *Knowledge Graph and Semantic Computing. Knowledge Computing and Language Understanding* (New York, NY: Springer Berlin Heidelberg, 2019). For a more readable (generic) overview of Microsoft's (early) efforts in this area, see Stefan Weitz, *Search: How the Data Explosion Makes Us Smarter* (London, UNITED KINGDOM: Taylor & Francis Group, 2016).

[16]Douglas A. Bernstein, ed., *Psychology*, 9th Ed (Belmont, CA: Wadsworth, Cengage Learning, 2012), Chapter 7.

At that more abstract level the idea of cumulative knowledge building is widely accepted. The real problems start when scholars try to operationalize this lofty ideal of 'standing on the shoulders of giants' in their own research. How does one effectively and adequately identify the real 'giants' in a field? How many of these giants' 'shoulders' form a solid enough foundation and how does one know or find out? How many layers of deeply piled-up giants is one supposed to look down upon in our quest for validated or at least compelling, promising and/or inspiring knowledge? How does one deal with the fact that subfields or adjacent fields (see our matryoshka layers) 'worship' different giants? How does one reduce the risk of missing potentially unrecognized giants?

For most scholars, these questions currently come to a head in the 'literature reviews' that all scholarly contributions are expected to contain. The main idea behind these literature reviews is very much in line with the underlying intuition of Google Scholar's motto: any intended augmentation of humanity's epistemic record (i.e. to a field's knowledge base) is expected to demonstrate adequate proficiency and understanding of what other scholars have produced previously. Yet there are no widely accepted 'standards' for these literature reviews. A – fairly typical – Harvard University 'guide', for instance, suggests that a bibliography should consist of two parts: "works cited" (which we find back in virtually all scholarly publications) and "works consulted"[17] (which, for instance, not a single one of the items in our deterrence-IS datasets included). Furthermore, recommendations on literature reviews also differ across disciplines and are often influenced not only by the prevalent disciplinary norms in a field but also by journal guidelines that impose direct or indirect limits (either by explicitly capping the number of works that can be cited[18] or by including references into already tight word limits[19]). This forces authors to be extremely selective in whom and how to cite, leading to choices that may best suit authors' preferences of being recognized in their particular discipline (e.g. making sure they cite the smallest amount of influential/'expected' works that they think they can get 'away with') while also following prevailing trends (e.g. 'diversifying' references across schools of thought or even authors' attributes like nationality, gender[20], etc.).

For this paper, we pursued an unusually expansive approach to the two questions raised in this scene-setting section: 1) "what is our 'real' field?" (e.g., is it Russian approaches to deterrence, is it international security deterrence, or is it deterrence as a whole?) and 2) "which giants' shoulders are we supposed to stand on?" (i.e. how do we identify human giants or epistemic 'giant-ideas'?, how do we even survey the broader landscape in which these giants stand out? how many giants

[17]"The bibliography should include all works that were used in the development, research, and writing of your thesis, not just the works that are quoted".
[18]Springer Nature, "Formatting Guide | Nature," 2021, https://www.nature.com/nature/for-authors/formatting-guide.
[19]Cambridge Review of International Affairs, "Submit a Manuscript," January 25, 2016, https://www.cria.polis.cam.ac.uk/Submissions/submit-a-manuscript.
[20]Bart Penders, "Ten Simple Rules for Responsible Referencing," *PLOS Computational Biology* 14, no. 4 (2018): e1006036, https://doi.org/10.1371/journal.pcbi.1006036.

are enough?). On the first question, we decided to focus on both the deterrence-IS and the (significantly smaller) deterrence-IS-Russia layers of deterrence research, but to also use the deterrence-broad literature as a point of comparison. On the second question as well, our analytical aperture was much wider than usual in the sense that we did not try to 'filter' the available literature(s) based on 'quality'-criteria. Instead of mostly hunting for giants (whether personal or ideational), our goal was to map as much of the entire available literature as we could lay our hands on and to then explore what we might learn from that.

Fish vs fishing rods

Before we jump into our actual findings, we want to stress that our main ambition in this paper was not necessarily to add new knowledge to the body of literature on the theory and/or practice of (Russian) deterrence. Instead this research effort primarily endeavoured to present a 'state of the field' ("what do we know about the knowledge that has been and is being produced on (Russian) deterrence?") and to explore/showcase some new ways in which our community could start building new knowledge on top of the extant epistemic base ("are there better, smarter, faster, more integrative ways to generate more trustworthy knowledge on (Russian) deterrence"?).

Most of our bibliography is likely to look unfamiliar to most of our readers. That was very much done on purpose. We were eager to share some of the most useful readings we stumbled upon in our own tortuous journey of discovery into these often new or 'unknown' disciplines, tools, datasets, insights, etc. Our own resolute goal remains very much to catch more proverbial 'fish' – validated epistemic insights that, from our own vantage point, will hopefully improve our strategic policy-making ability at the national and international levels. We have just become increasingly convinced that to catch better fish, we will need better fishing rods.

3. AN 'MRI'-SCAN OF THE FIELD: THE NUMBERS

As part of the larger RuBase research program[21] we are working on, our team first of all conducted an extensive bibliometric[22] equivalent of a medical 'MRI-scan' on the field(s) of deterrence studies. This section of our paper reports on these (still ongoing) efforts. It describes what bibliometrics is, where it stands, which bibliometric datasets we collected and how. It then moves on to describe what this 'MRI' taught us about some key aspects of this field: what we know about its scholarly production, the uptake of this production by other scholars, the velocity and

[21]RuBase is a five-year research project (January 2018 – December 2019; January 2020-December 2022) that sets out to improve our understanding of Russia's multi-domain international behavior. The 'base' in RuBase refers both to the knowledge base the team is building and to its (aspirationally) foundational nature. Existing and new text- and numbers-based datasets and -tools which will be collated/built and explored through both human analysis and (supervised *and* unsupervised) machine-learning algorithms. This should generate a new visual and interactive knowledge base on Russia that will be positioned as a platform for new collaborative ways of cumulative knowledge building on this topic that is only gaining in policy relevance. The project is funded by the Carnegie Corporation of New York and by the Minerva Research initiative.

[22]Bibliometrics is defined as "a set of quantitative methods used to measure, track, and analyze print-based scholarly literature" Robin Chin Roemer and Rachel Borchardt, *Meaningful Metrics: A 21st Century Librarian's Guide to Bibliometrics, Altmetrics, and Research Impact* (Chicago: Association of College and Research Libraries, A division of the American Library Association, 2015), 28.

effectiveness with which this happens, the degree of collaboration this process exhibits, its entropy and last but not least its thoroughness. The best way to describe how we would characterize our findings is that we found them to be profoundly discomfiting and discomforting.

Bibliometric datasets

Bibliometric databases – the state of the field

This section sets out to map some key ('technical') aspects of the overall knowledge landscape on (Russian) deterrence in the field of international security based on what we can glean from the openly published 'knowledge production[23]' on this topic. As we noted, a significant part of extant manifest[24] knowledge on any topic rests encoded in written scholarly publications of various sorts. These include academic ones, but also non-academic, yet still high-quality publications from think tanks, governments, international organizations: the so-called 'grey literature'[25]. One might therefore reasonably expect any serious knowledge-mapping effort to start by analyzing at least the bibliographic metadata[26] of all available scholarly publications on a given topic. Unfortunately, it is currently still impossible to create a single bibliographical dataset containing even just the metadata for all available academic (let alone non-academic, but still high-quality) publications on a topic.

Until recently, access to this bibliographic meta-data was severely hampered by the typical pathologies of what remained essentially a duopolistic market. This market consisted of on the one hand *Scopus*[27], owned by the Netherlands-based *Elsevier* (also the owner of the content of many academic publications)[28] and on the other hand *Web of Science*[29], the more recent incarnation of

[23]J.Y. Tsao et al., "Galileo's Stream: A Framework for Understanding Knowledge Production," *Research Policy* 37, no. 2 (March 2008): 330–52, https://doi.org/10.1016/j.respol.2007.10.004.

[24]We suspect that only a fraction of all actual knowledge on this topic has been made manifest – e.g. through publication. We reckon (but have not tested whether) that is as much the case for the classified realm as it is for the non-classified one. Much (most?) of our real knowledge is probably still lurking in the minds of myriad scholars, military planners, (also intelligence) analysts, etc. but eliciting this tacit knowledge through different knowledge discovery tools has so far remained beyond our current abilities even in more 'open' contexts. For a recent review, see Jia Hao et al., "A Review of Tacit Knowledge: Current Situation and the Direction to Go," *International Journal of Software Engineering and Knowledge Engineering* 27, no. 05 (June 2017): 727–48, https://doi.org/10.1142/S0218194017500279.

[25]Harris M. Cooper, Larry V. Hedges, and Jeff C. Valentine, eds., *The Handbook of Research Synthesis and Meta-Analysis*, 2nd ed (New York: Russell Sage Foundation, 2009). For a more recent meta-analysis of the grey literature, see Avijit Mahala, Gayatri Dwivedi, and Manorama Tripathi, "A Bibliometric Study of Grey Literature (2007-2019)," *Collection and Curation* (January 1, 2020), https://doi.org/10.1108/CC-12-2019-0043.

[26]Metadata include the usual bibliographical fields such as author(s), publication title, source, year, abstract, etc. The 'open metadata' movement that tries to 'open' these metadata for general public use has made significant progress (especially recently – see below), but still remains far removed from its goal. Ginny Hendricks et al., "Crossref: The Sustainable Source of Community-Owned Scholarly Metadata," *Quantitative Science Studies* 1, no. 1 (February 1, 2020): 414–27, https://doi.org/10.1162/qss_a_00022; European Commission. Directorate General for Research and Innovation et al., *Open Metadata of Scholarly Publications: Open Science Monitor Case Study.* (LU: Publications Office, 2019), https://data.europa.eu/doi/10.2777/132318.

[27]Elsevier, "Scopus," 2019, https://www.scopus.com/standard/marketing.uri.

[28]Scopus contains nearly 78 million records from 34,346 peer-reviewed journal.

[29] Clarivate Analytics, "Web of Science," *Web of Science Group* (blog), 2019, https://clarivate.com/webofsciencegroup/solutions/web-of-science/.

the pre-digital 'Science Citation Index'[30] that is currently owned by US-based *Clarivate* (that does not have ownership over any academic publications, but specializes in scientific 'intelligence').[31] Both are subscription-based, expensive, relatively hard to access (and especially to retrieve datasets from in large quantities), of disappointing quality (with many inconsistencies and omissions even within databases), with significant overlap across them (see the left Venn-diagram on Figure 2), and with surprisingly few value-added services beyond bare-bones access to their database (i.e. no faceted search, no or very limited visualizations, etc.). In recent years the situation in this market has improved somewhat, not in the least by the appearance of some new entrants into this market like CrossRef, Google Scholar, Microsoft Academic, Semantic Scholar, but especially the far more intuitively accessible, 'open' and advanced Lens.org and Dimensions.ai websites.

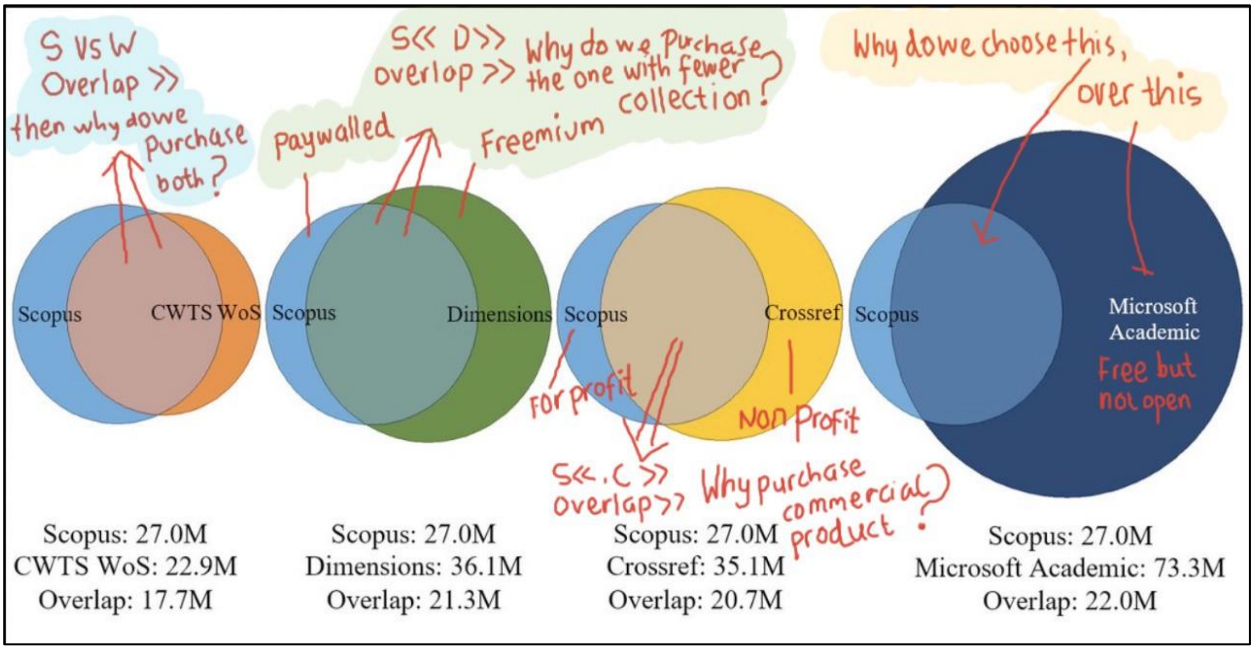


Figure 2: Overlap of selected databases[32], annotated by a sharp-witted Indonesian scientist[33]

While access to actual bibliographic metadata has therefore improved, the absence of any systematically available common field (like the unique DOI – Digital Object Identifier – field) that could be used to reliably identify identical articles[34], or – in the absence of such an

[30]The first citation index in science was created in 1960 by Eugene Garfield and the Institute for Scientific Information to "be a spur to many new scientific discoveries in the service of mankind". Eugene Garfield, "Can Citation Indexing Be Automated ?," in *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*, ed. Laurence B. Heilprin et al. (Washington D.C.: U.S. Government Printing Office, 1966), 189–93; Eugene Garfield, "Letter of May 21, 1959 from Dr. E. Garfield to Dr. J. Lederberg," May 21, 1959, http://www.garfield.library.upenn.edu/lederberg/052159.html.

[31]The core collection of the Web of Science currently contains citation data for 79 million records from 21,100 peer-reviewed journals.

[32]Ludo Waltman, "Q&A about Elsevier's Decision to Open Its Citations," *Leiden Madtrics* (blog), December 22, 2020, https://leidenmadtrics.nl/articles/q-a-about-elseviers-decision-to-open-its-citation.

[33]Dasapta Erwin Irawan, "Berlawan Dengan Prinsip Umum Para Pembeli Barang [Contrary to the Principle of Buying Goods in General]," Medium, January 19, 2021, https://medium.com/open-science-indonesia/berlawan-dengan-prinsip-umum-para-pembeli-barang-5c4893fa362a.

[34]The obvious solution to this predicament – using so-called (unique and consistent) DOIs or Digital Object Identifiers – is widely-accepted. Unfortunately, both Scopus and Web of Science have so far refused to provide a (shared) DOI field for all

unambiguous field in all databases – of a reliable matching algorithm that would still allow us to achieve the same goal[35], makes it essentially impossible to work with one consolidated dataset[36]. Our analysis in this section is therefore unfortunately based on separate analyses of datasets culled from a number of these databases[37].

Our bibliometric datasets

To examine the key features of the literature on deterrence, we collected three different bibliometric datasets (see also Figure 1) based on different queries. The first query was a purely 'international relations/international security' query that included more traditional terms like nuclear or conventional deterrence, as well also more recent ones like hybrid deterrence, cyberdeterrence, or cross-domain deterrence(***'Deterrence-IS'***).[38] In order to be able to put our findings from these Deterrence-IS datasets in perspective, we also used a second query that aimed to collect metadata from the entire body of scholarly literature on deterrence in general, including from other disciplines than political science/international relations such as sociology, criminology, biology, economics, etc.[39]: ***Deterrence-broad***. Given the fact that many of these fields are also dealing with various aspects of human agency, we thought this broader dataset would provide an interesting point of comparison[40].

---

of their publications. Crossref, the official Digital Object Identifier Registration Agency of the International DOI Foundation, is definitely making inroads here, but unfortunately at this moment, we are not yet in a position to use DOIs to match and deduplicate all identical records shared in both databases.

[35]Also here, progress has been made by one of the internationally leading research teams in this field – Leiden University's CWTS (the Dutch acronym for 'Center for Science and Technology Studies'). Martijn Visser, Nees Jan van Eck, and Ludo Waltman, "Large-Scale Comparison of Bibliographic Data Sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic," *ArXiv:2005.10732 [Cs]*, January 17, 2021, http://arxiv.org/abs/2005.10732. HCSS is currently working with Dr. Chaomei Chen from Drexler University to implement this algorithm (or another one) in CiteSpace. In future iterations of this work, we should therefore be able to present the overall findings based on one single, consolidated bibliometric dataset culled from these various databases.

[36]We should add, however, that the past few months have seen unprecedented progress on this: as of January 2021, the fraction of publications with open references now stands at 83% (in 2016, it used to be 1% – Dario Taraborelli, "A Good Progress Bar," Tweet, *@ReaderMeter* (blog), January 20, 2021, https://twitter.com/ReaderMeter/status/1351943177293369349) because Elsevier decided to make the reference lists of all of its publications – hundreds of millions of them – openly available in Crossref (Waltman, "Q&A about Elsevier's Decision to Open Its Citations.").

[37]Access to our team's (far more expansive) full analysis of all datasets can be arranged upon request.

[38]The precise query was ("nuclear deterrence" OR "military deterrence" OR "conventional deterrence" OR "hybrid deterrence" OR "cyber deterrence" OR cyberdeterrence).

[39]See also De Spiegeleire et al., *Reimagining Deterrence*.

[40]We were unable to identify an intelligent way to create a 'deterrence-human' subset, but still intend to pursue this line of inquiry at a later time.

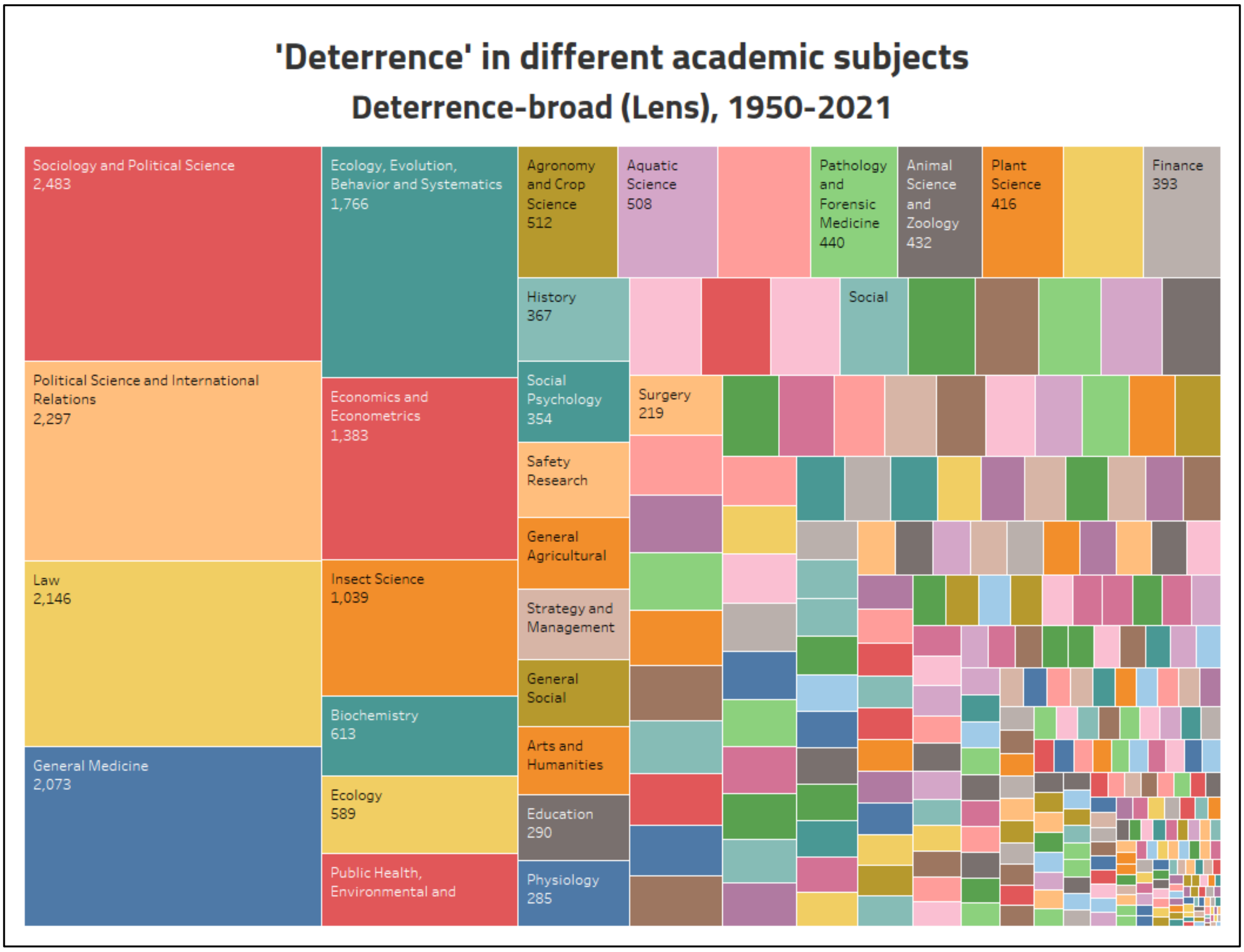


Figure 3: Deterrence in different academic subjects

Because our primary interest was in the (extremely limited) literature on Russian deterrence, we also created three additional Russia-specific bibliometric datasets. Two were a subset of our (mostly) English-language deterrence-IS datasets that was filtered to only include publications with the word 'Russia' in publications' titles and/or abstracts. In Web of Science, for instance, this yielded 266 Russia-related articles from our Deterrence-broad dataset, 65 of which were on international security. In Scopus, it gave us 91 Russia-related articles on international security. we will refer to those datasets as '***Deterrence-IS-Russia/English'.*** To compare both the state of research and the main themes in the Russian literature on international security-related deterrence, we also created a second Russia-specific bibliometric dataset from the Web of Science's Russian Science Citation Index. We refer to these datasets as '***Deterrence-IS-Russia/Russian***'[41].

The following table presents an overview of the different datasets we collected.

[41]Query TOPIC: ((устраш* OR сдерживани* OR запугивани* OR шантаж*) AND (ядерн* OR атомн* OR кримин* OR преступ* OR оруж* OR вооруж* OR воен*)) Timespan: All years. Indexes: RSCI.

Available data → Databases ↓	Focus	Years covered[42]	Number of documents
Lens	Deterrence-Broad	1883-2020	49,687
Scopus	Deterrence-Broad	1899-2020	22,259
WoS Core	Deterrence-Broad	1910-2020	19,524
Dimensions	Deterrence-Broad	1871-2018	20,024
Lens	Deterrence-IS	1958-2020	2,550
Scopus	Deterrence-IS	1958-2020	810
WoS Core	Deterrence-IS	1958-2020	695
Google Scholar	Deterrence-IS	1990-2020	44,888
Dimensions	Deterrence-IS	1958-2020	1,247
WoS Core	Deterrence-IS-Russia/English	1994-2020	65
Scopus	Deterrence-IS-Russia/English	1994-2020	91
WoS RSCI	Deterrence-IS-Russia/Russian	2005-2020	127

Table 1: Our bibliometric datasets

Although we will present visuals from a number of these different databases, we will mostly focus on Lens and on Google Scholar[43] in this section, because of a) their overall significantly greater coverage even on the purely 'academic' side; but also b) because they (and especially Google Scholar) also include not-purely-academic publications. Taking this into account, we used Google Scholar to create a think tanks dataset. We extracted all of the top-level domains out of the Google Scholar json-file, and then used the 2019 *Global Go To Think Tank Index Report*[44] to make sure we included at least the top-20 think tanks, which we then still augmented by some of the other think tanks we knew and that ranked highly in the extracted list. The final list can be found in Annex A – List of think tanks included in the analysis. We also include some analysis from Web of Science RSCI, because it is the only database that has coverage of Russian bibliometric information through the Russian Science Citation Index (2005-present).

Our analysis of these various datasets allows us to draw several robust conclusions across these different datasets, which we have listed, numbered, illustrated, and interpreted in the following

[42]The time coverage is based on the earliest (typically the oldest one in the database) and latest available (typically the most recent one at the time of downloading) publications in a given dataset.

[43]Google Scholar makes it exceptionally difficult to scrape large numbers of items from its cache, but the appearance of the python 'scholarly' package allowed us to retrieve all relevant results into a json-file, which we could use for bibliometric analysis. Luciano Bello, *Scholarly-Python-Package/Scholarly*, Python (2014; repr., scholarly-python-package, 2021), https://github.com/scholarly-python-package/scholarly.

[44] (James G. McGann, "2019 Global Go To Think Tank Index Report," 2020, https://repository.upenn.edu/cgi/viewcontent.cgi?article=1018&context=think_tanks.)

sections.

Research production

What do our datasets tell us about the sheer volume of scholarly publications in these fields?

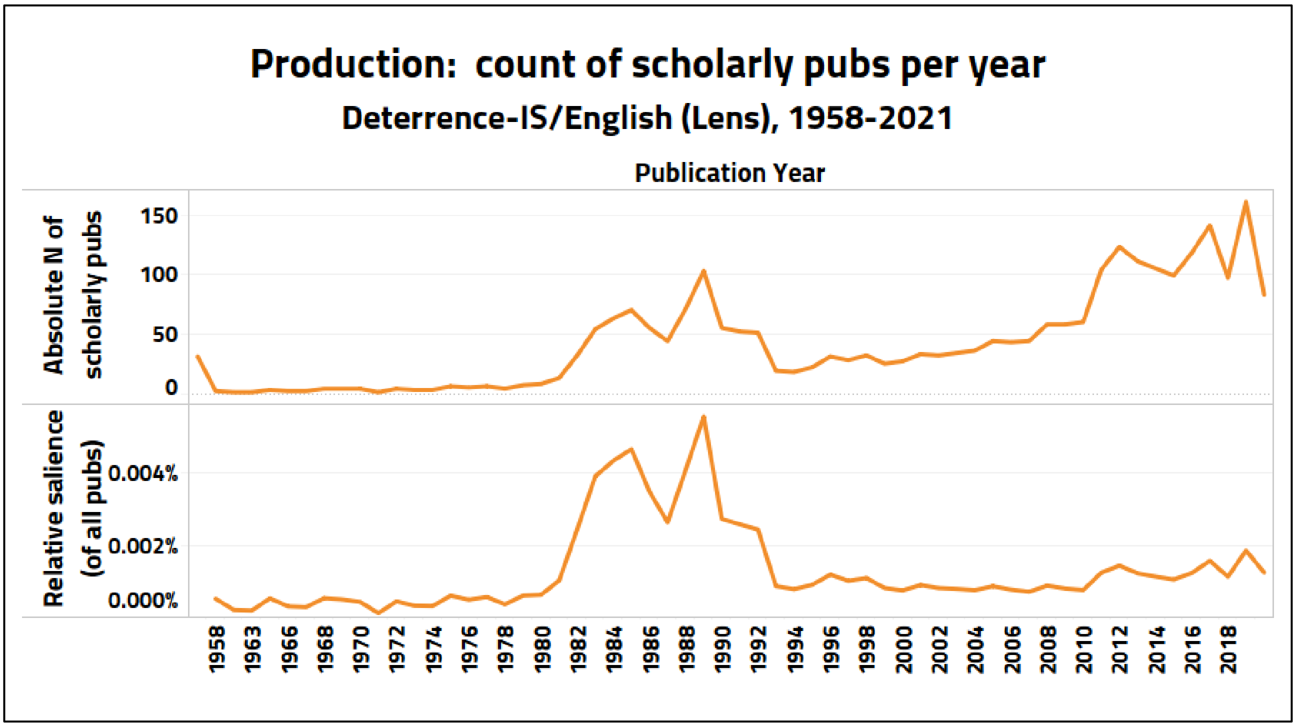


Figure 4: Number of scholarly documents on deterrence/international security per year: absolute vs relative

The heydays of research production in the field of deterrence in an international security context (deterrence-IS) occurred in the late- and immediate post- Cold War era (1982-1992) – both in absolute terms (the top of Figure 4: how many documents were published that year) and in relative terms (the bottom of Figure 4: what percentage does that number represent of all scholarly documents produced that year);

In ***absolute*** terms, the teens of this century saw the highest numbers of publications. This reflects primarily the massively increased number of publication venues that are currently available for authors to get their material published. But this also still means that since 2011, about 100 scholarly documents are being produced on this topic per year – more than ever before.

In ***relative*** terms, however, we have not even come close to the peaks of the 1982-1992 period. One way to think about this is that the relative focus of the global scientific community on this topic has declined since the topic's heydays a few decades ago.

While the number of publications remains low, we can observe that Russian language literature on deterrence in the Web of Science RSCI (Deterrence-IS-Russia/Russian) starts to increase around 2015. This increase does not appear to be directly related to real-world events like Ukraine or Syria, since the titles and abstracts talk about strategic stability, nuclear aspects of deterrence as before. Russia as a topic of interest in WoS Core (Deterrence-IS-Russia/English) peaked in 2014 and interest has been increasing ever since. If we look at the country of origin of the first authors of the articles (information present in 78 out of 105 articles of the two datasets (English and Russian language) combined), we can see that in both cases, Russian authors contributed to peaks the most.

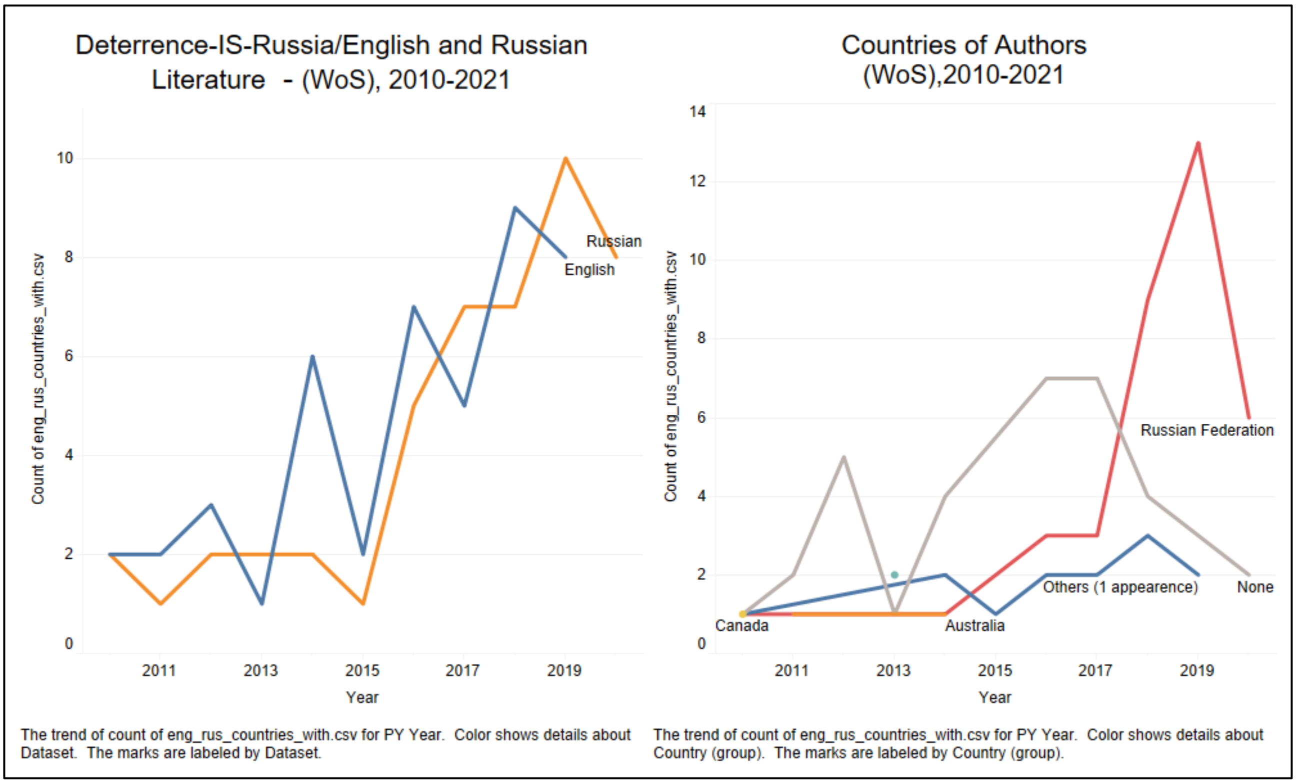


Figure 5: Production - comparison Russian vs. English publications; publications by country

Research uptake

Our team took a look at the extent to which this (quite modest) research production has been and is being leveraged by other scholars working in this or in other fields ("standing on the shoulders"). This part of our analysis is based on the numbers that are included in our bibliometric datasets for every publication: how often that publication has been cited until the moment of our analysis (in our case in early January 2021) – e.g. 'Times cited' (TC) in the Web of Science, 'Scholarly Citation Count' in the Lens, 'Cited by' in Google Scholar, etc.[45]

[45]These citations include 'global' citations across the entire database from which our datasets were culled (e.g. possible also citations in NON-deterrence-related publications) and not only 'local' citations in 'our' datasets,

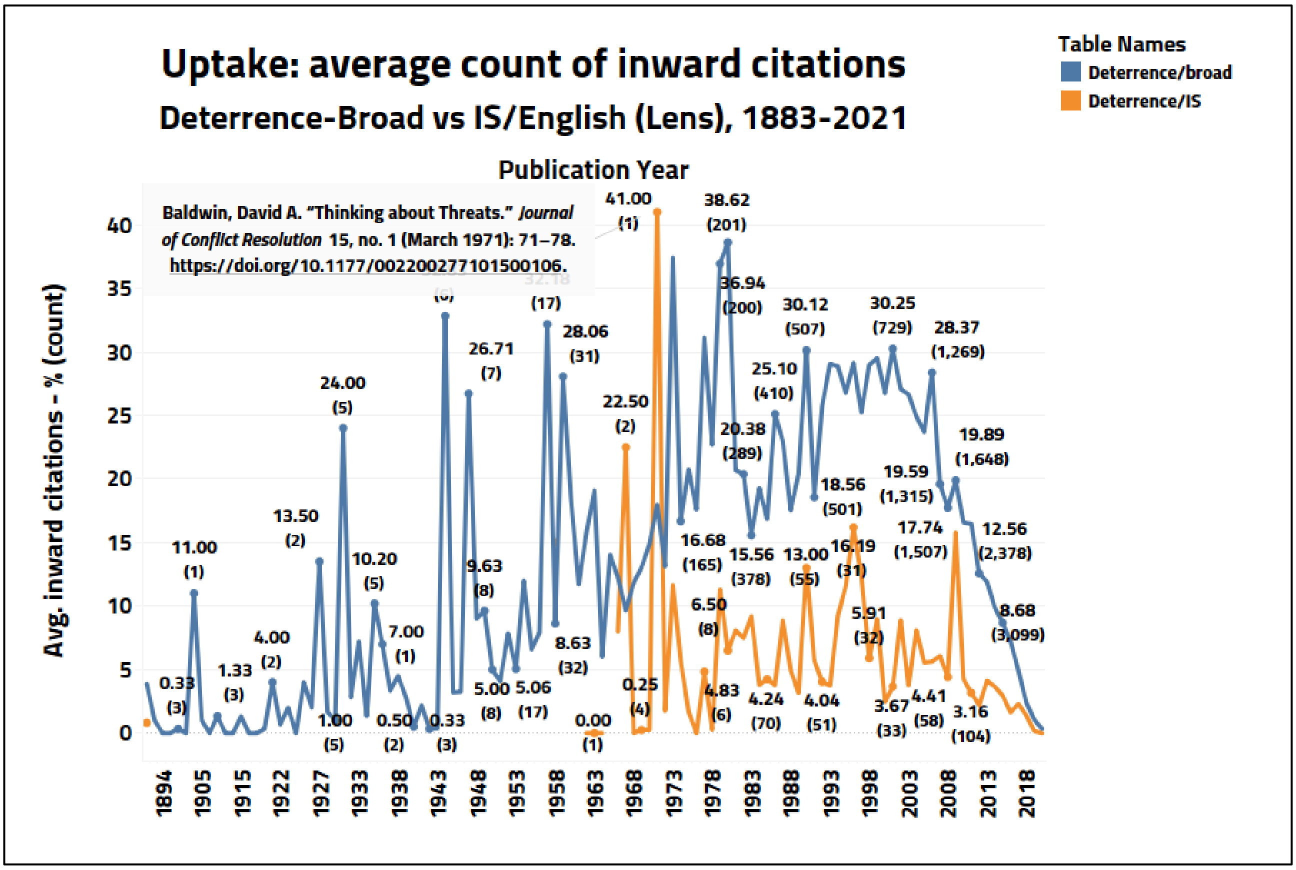


Figure 6: Uptake - average count of inward citations

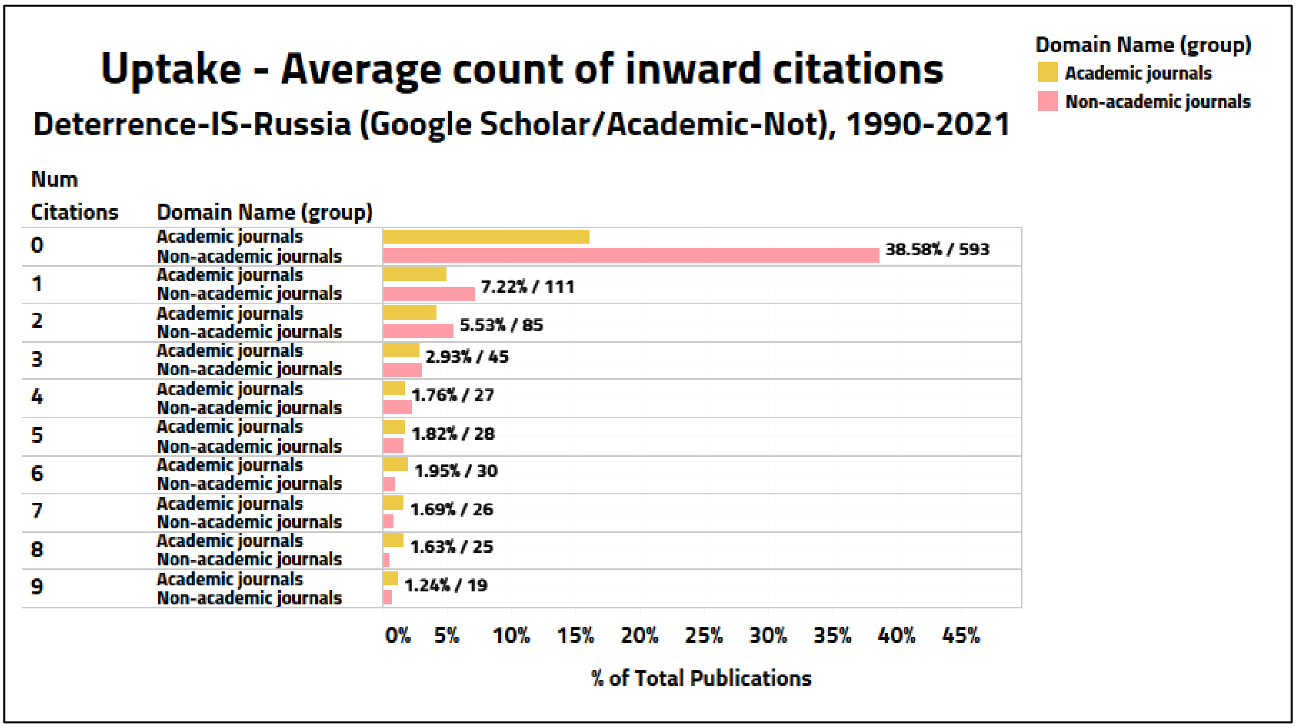


Figure 7: Uptake - average count of inward citations: academic vs non-academic

The largest academic scholarly dataset we have, from the Lens, reveals that on average, an article about deterrence in the field of international security (in orange) is only cited about 4.5 times. This compares to an average that is more than 3 times higher (15.5) for the broader deterrence literature in that same database. The only dataset we collated that also includes non-academic scholarly publications (Google Scholar) indicates that 60% (!) of all articles published in the non-academic subset of our deterrence-IS dataset never once get cited, compared to 50% – still only one in two gets cited at least once – in the academic subset.

The trend over time in citations shows a clear upwards trend in the broader deterrence literature, but not in the IS-specific one. [We want to point out that the decline in the average number of citations in recent years is expected due to the fact that it takes a (surprisingly) long time before publications get cited (in general, but especially in both of these fields, which always puts more recent publications at a disadvantage – see below)].

There are various ways of interpreting these disappointingly low numbers. One could be that authors working in this field are not particularly diligent in looking up, reading and/or using/citing their colleagues' work (and we remind our readers, that the Lens deterrence-IS dataset does contain over 2500 publications). Another could be that much of this literature may not be available to them. A third possible explanation is that many articles may be judged by authors to be of insufficient quality to even deserve being cited. Whatever the real cause(s) may be, we submit that this is certainly a point of attention for the field as such and for the scholars working in it. Funders (and all of this work is funded one way or another) and publishers may also wish to think more strategically about different incentive structures that would either disincentivize the publication of possibly sub-par work; incentivize more cooperation between (especially some of the most value-added) contributors; or some combination thereof.

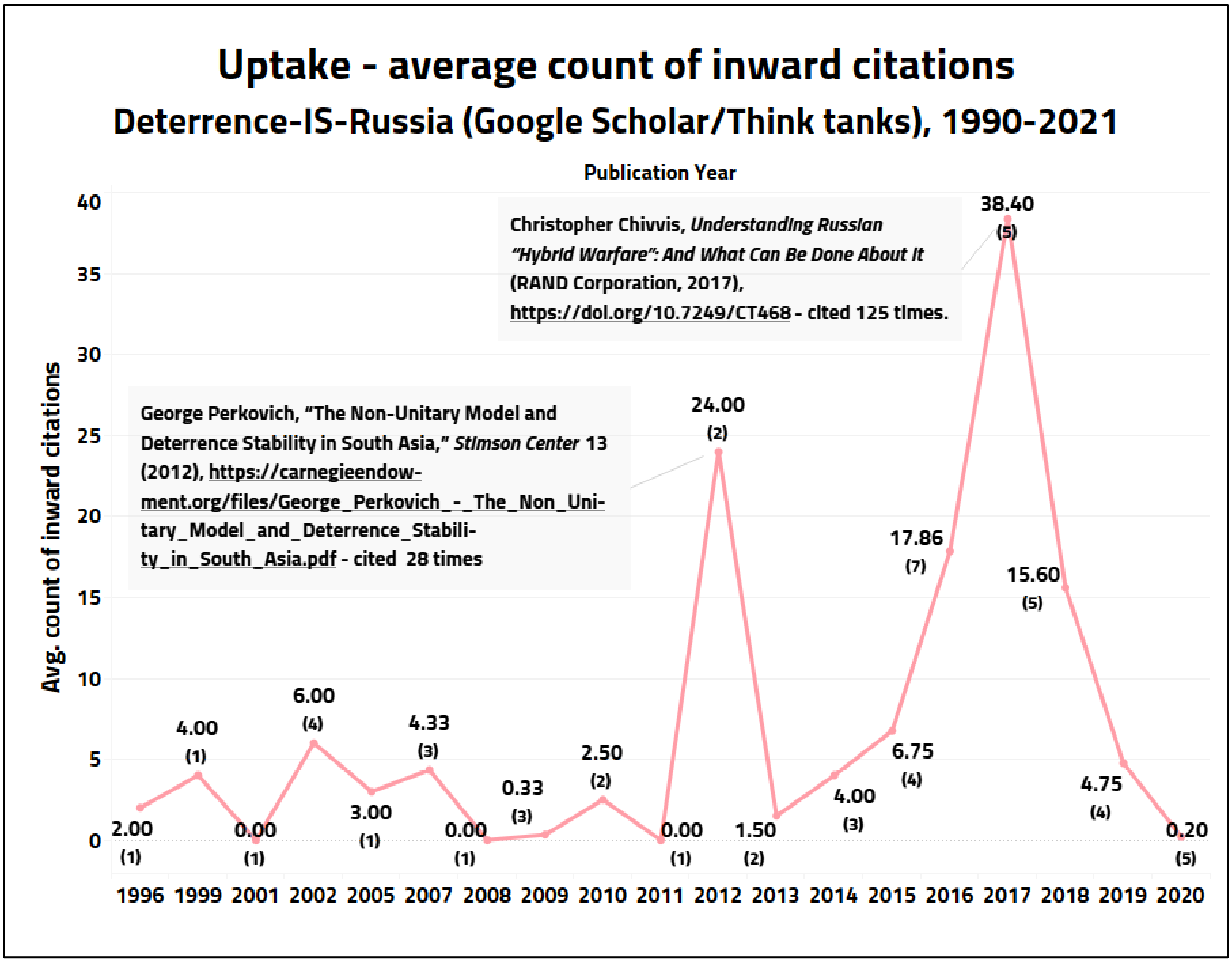


Figure 8: Uptake - average count of inward citations: think tanks

If we filter out of all these academic international security publications dealing with deterrence only the ones that contain the term 'Russia' in their abstracts (IS-deterrence-Russia/English) we are left with a paltry 65 publications in our largest dataset (Google Scholar). In this dataset 30% (almost one-third) of publications have never been cited, which may strike our readers as surprisingly high (read: 'bad'), but it is, as we have seen, actually still significantly better than the non-Russia-specific dataset. Looking at the entire dataset (and not just at the cited vs non-cited subsets) the average amount of times a publication is cited, however, drops to 3.1, which is even lower than the non-Russia-specific ones. So a larger part of these publications is cited than their non-Russia-related counterparts, but still only a very few times on average. We hasten to add that, as we will see, this does not include the (more numerous) publications on this topic in the non-academic scholarly literature by think tanks and others (see below), as well as the non-publicly available ones. So whereas the 'real' overall public[46] 'knowledge' situation is probably slightly less

[46]Although it is difficult to say much about the classified literature, it is probably fair *and* safe to point out that that part of 'human knowledge' on this topic is mostly focused on more technical/tactical/operational issues, and not on the broader conceptual/strategic issues that the 'open source' literature mostly focuses on. We know for a fact that there is significantly more collaboration on those issues in government labs and other research facilities – certainly nationally, and to a much smaller

dire than these data suggest, we would still submit that this situation should give all of us pause.

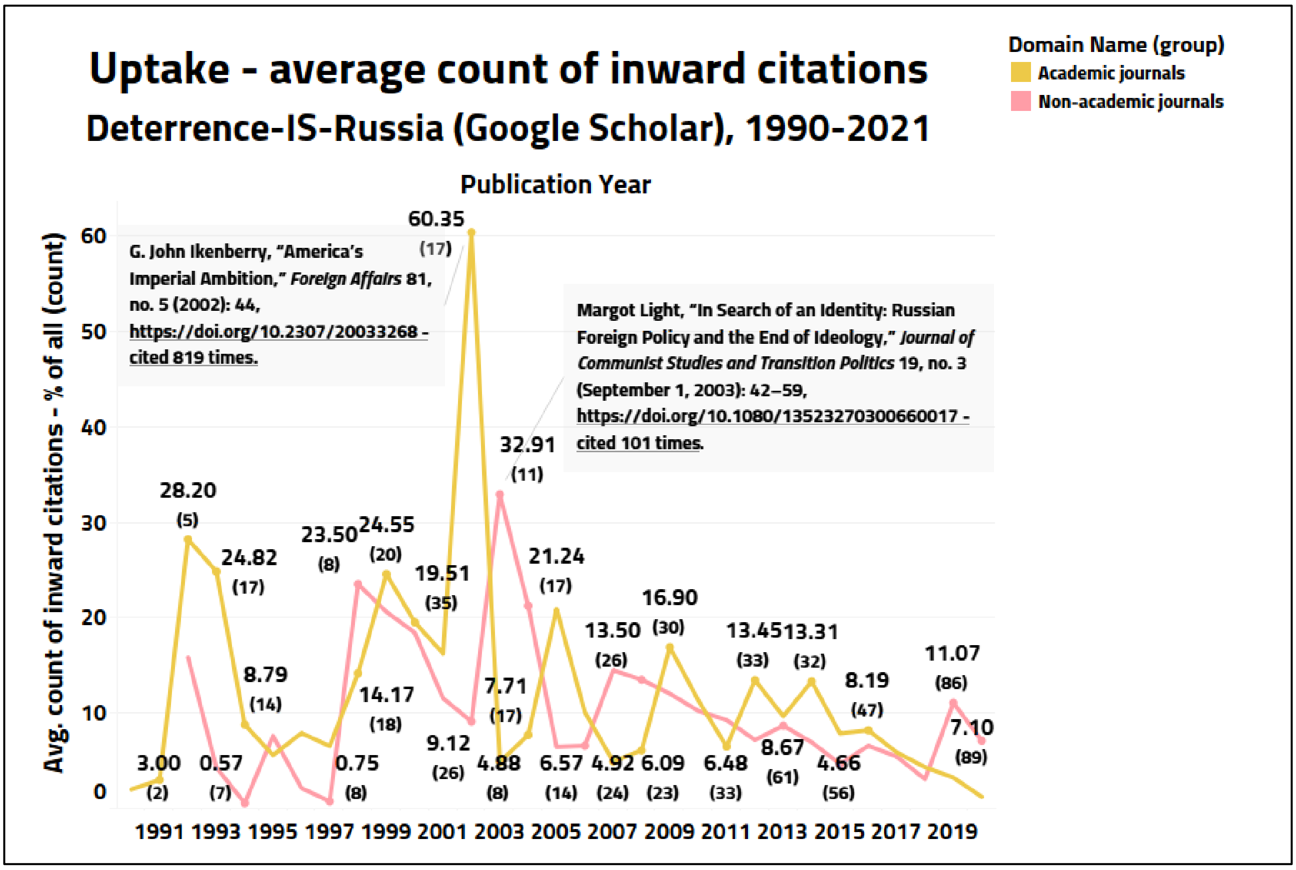


Figure 9: Uptake - average count of inward citations: Russia-specific publications

extent also internationally. This still begs the question, however, how confident the scholarly *and* the public decision-making communities can be about the more fundamental/conceptual underpinnings of the very concept of deterrence.

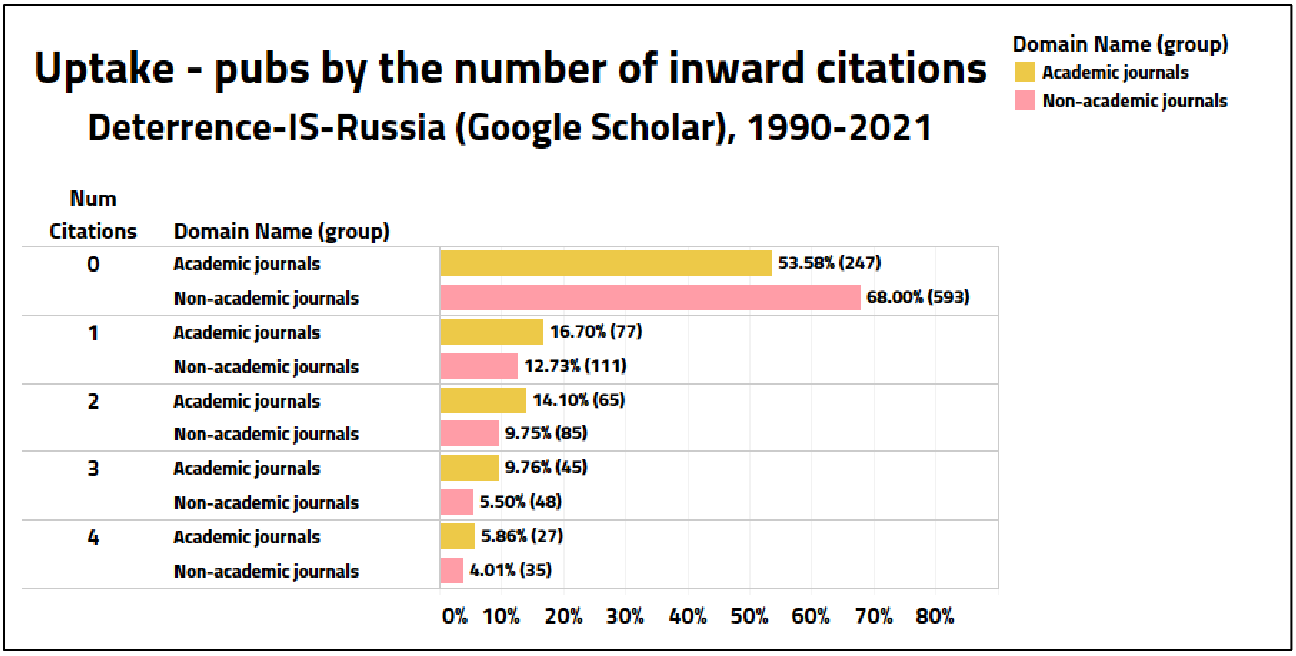


Figure 10: Uptake - average count of inward citations: academic vs non-academic, per author count

In the Web of Science RSCI, Russian-language literature is naturally cited significantly less than English-language publications on Russian deterrence. Overall, the number of citations remains low with some outliers such as the 2016 article by A. Lanoszka on Russian deterrence in Eastern Europe (within the English-language dataset) that was cited 44 times even though it is relatively recent. To remove such discrepancies, we looked at the average number of citations for this small dataset and noticed trends similar to those in the other corpora.

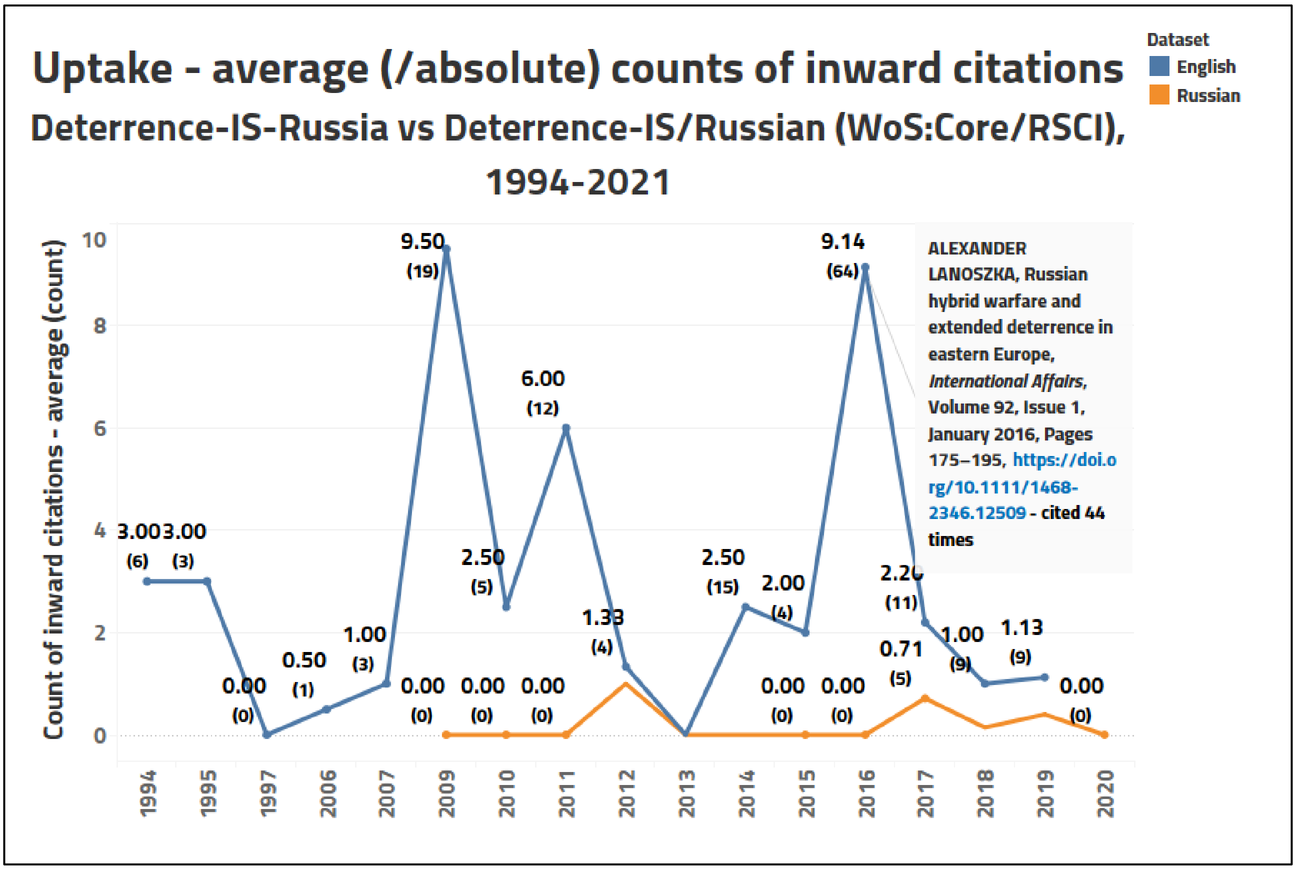


Figure 11: Uptake - average count of inward citations: English vs Russian publications

The only dataset that allows us, as we already saw, a glimpse into the differences between academic and non-academic publications is Google Scholar. Figure 7 shows us this breakdown, revealing that academic publications are cited 10.4 times on average, while non-academic ones come in at only 7.8. There does not appear to be a clear trend over time for either of these two categories. When we look at the differences between academic and non-academic sources dealing with both international security *and* with Russia in our Google Scholar dataset, however, we find out that 41% of academic publications on this topic never get cited vs. 62% of non-academic ones.

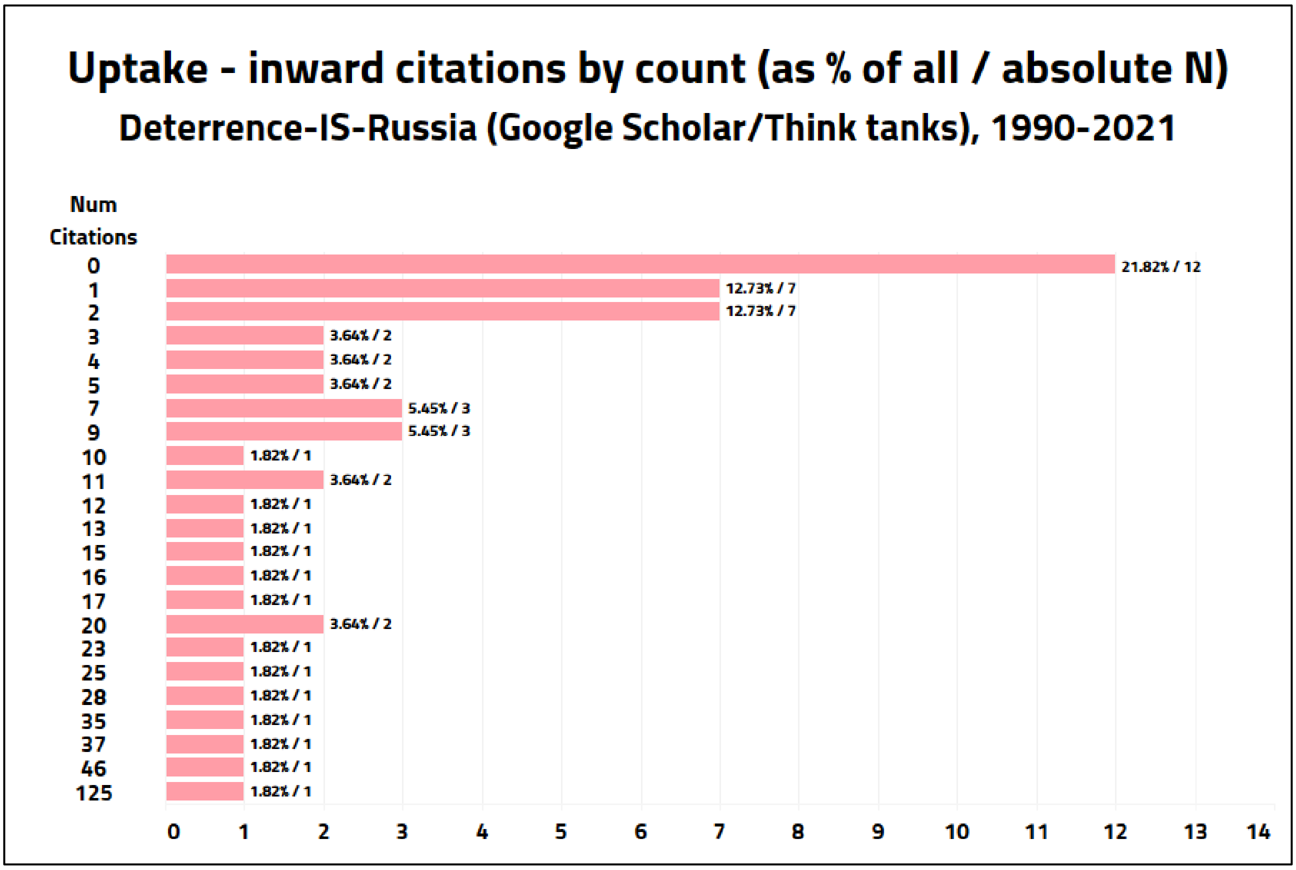


Figure 12: Uptake - average count of inward citations: Russia-specific publication from think tanks

If we just single out some of the major think tanks in the Google Scholar dataset (the only one that includes them, we find that the average think tank publication in this field is cited 10.3 times – almost exactly the same as the academic publications in that same Google Scholar dataset and therefore significantly higher than their non-academic counterparts. The one big difference that sets think tanks apart from all others seems to be that only 22% of these publications have never been cited – a significantly lower number than the other datasets. It may also be worth noting here that the intended value of those think tank reports are not primarily focused on research uptake, but on policy uptake. What little evidence we have on this, in general, is also not particularly uplifting[47], but honestly compels us to admit that our metrics on this score are significantly less reliable than the more 'technical' scientometric evidence we have been presenting so far.

[47]Generally speaking, see Andrew Rich, *Think Tanks, Public Policy, and the Politics of Expertise* (Cambridge: Cambridge University Press, 2004), https://doi.org/10.1017/CBO9780511509889. For more recent discussions (and some evidence) in our field, see Donald E. Abelson, *Do Think Tanks Matter? Assessing the Impact of Public Policy Institutes*, Third edition, revised and expanded (Montreal: McGill-Queen's University Press, 2018); James G. McGann, *Think Tanks and Policy Advice in the United States: Academics, Advisors and Advocates*, Routledge Research in American Politics 1 (Abingdon [England] ; New York: Routledge, 2007). There is a general recognition, however, that actual evidence on the direct impact of inputs from think tanks on official government policy in various areas (and especially the security and defense area) is relatively slim.

Research velocity

The extraordinary 'pandemic' year 2020 has shown all of us how the international scientific community is capable of generating cumulative (and actionable) knowledge at unprecedented speed on issues that are widely recognized as being of 'existential' importance. Over the past few decades, the overall pace of scientific progress has proved to be fairly glacial in many disciplines. Because of the currently surrealistically inefficient academic publication 'market', it still takes – in most disciplines – an extraordinarily long gestation period before a publication (and – more importantly – the new knowledge it contributes) sees the light of day. Once published, the propagation of the new knowledge that is encoded in that publication throughout all relevant epistemic communities also tends to proceed (as we have seen – IF it even proceeds at all) at a surprisingly leisurely pace. Information on the half-life of these publications is scarce, but what little we do know about this suggests that much of the presumed 'knowledge' tends to be quickly overtaken.

Our bibliometric Deterrence-Broad and Deterrence-IS-Russia/English (WoS) datasets reflect this trend to a certain extent. 'Older' articles have a significantly larger average half-life (up to 4.5 years) which shows their importance to the field even today. Within the IS-Deterrence/Russia dataset, two articles, published in 1984 and 2001 keep their half-life at this level, which may highlight their significance today. As the number of articles on the matter started to increase in the second half of 2010s, the average half-life, naturally, decreases, but some papers maintain high 'scores' for their half-life throughout the entire dataset.

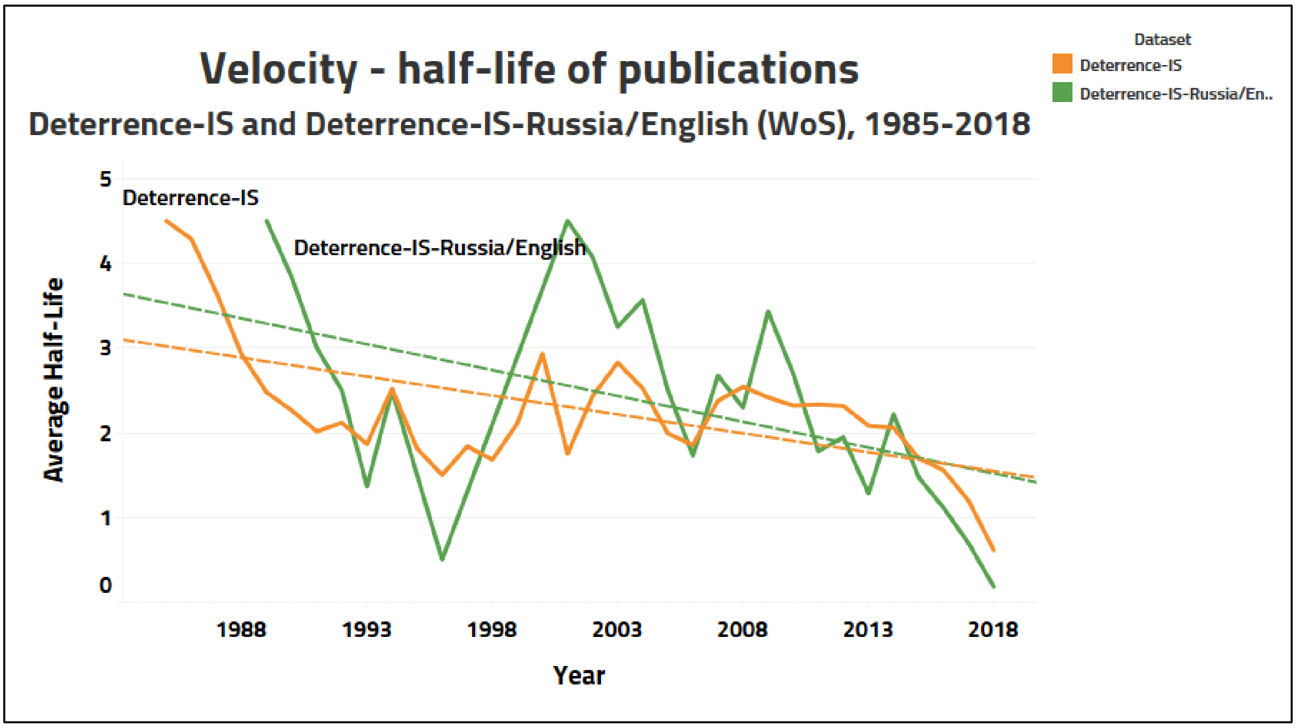


Figure 13: Velocity - half-life of publications

This comparison between half-life metrics in our Deterrence-Broad dataset and Deterrence-IS highlights another important finding. The precipitous drop in Deterrence-IS publications' 'half-life' with the end of the Cold War reflects how quickly output in this area was overtaken with the USSR disintegration. Moreover, Deterrence-Broad 'half-life' metrics are higher on average (2.7 versus 2.2 for 1985-2018) and overtake Deterrence-IS with the decline of the Cold War. The trendline especially highlights the divergence of 'half-life' metrics comparing two datasets. It reflects the higher temporal vitality of research on deterrence in criminology, health studies, and 'hard sciences' rather than in those disciplines, studying international security which restricted with higher temporal boundaries of relevance.

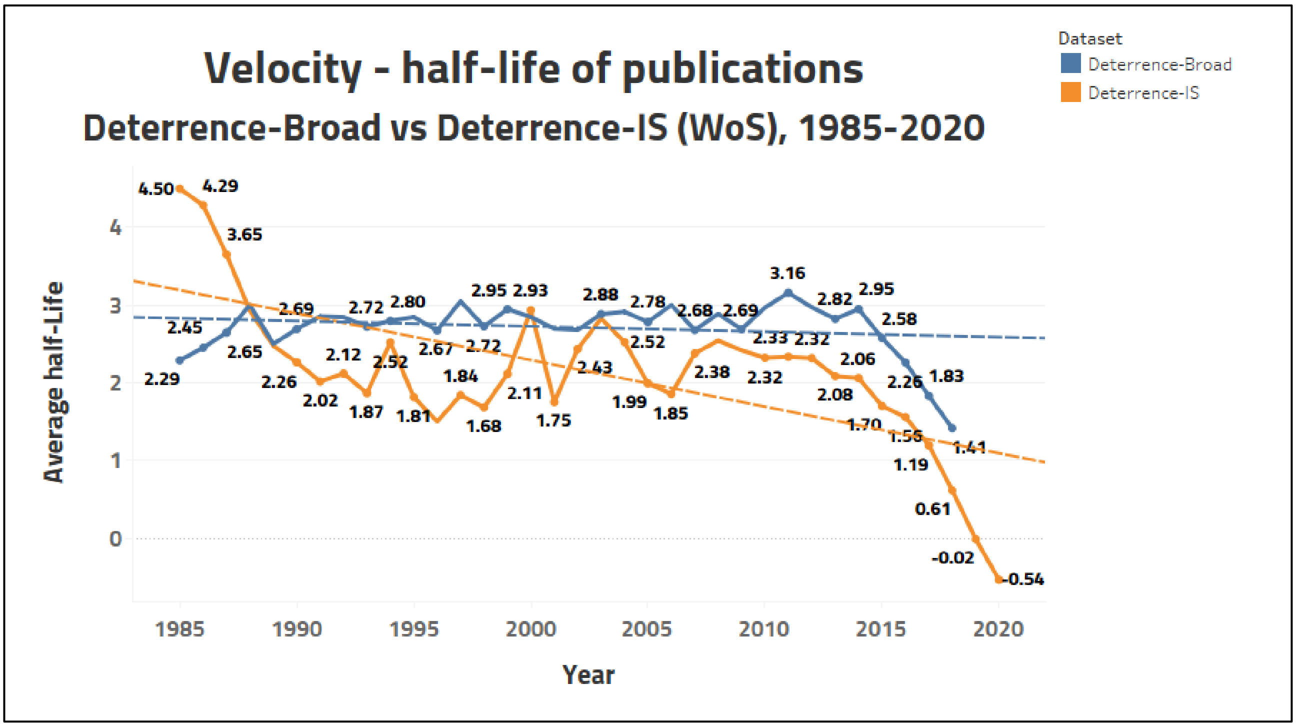


Figure 14 Velocity - half-life of publications: deterrence-broad vs deterrence-IS

We also computed captures how quickly a publication can get cited in our datasets – another metric capturing the speed of scientific turn-around. For that purpose, we found the year of publication of the most recent cited reference in a publication and calculated how many years had lapsed between that year and the year when the citing publication was published. The data (Figure 15) shows that the 'youngest' cited reference in our dataset was on average 1.3 years old – for both Deterrence-broad and for Deterrence-IS.

If one considers that the gestation period of a scientific publication (the time it takes to produce a research paper, to submit it and get it accepted for publication and then the actual final publication) in the social sciences is almost two years[48], this means that knowledge is being built at a surprisingly leisurely pace[49] in which it takes at least 3 years from the time an author submits a paper to be included in the scholarly record to the first publication that builds on it.

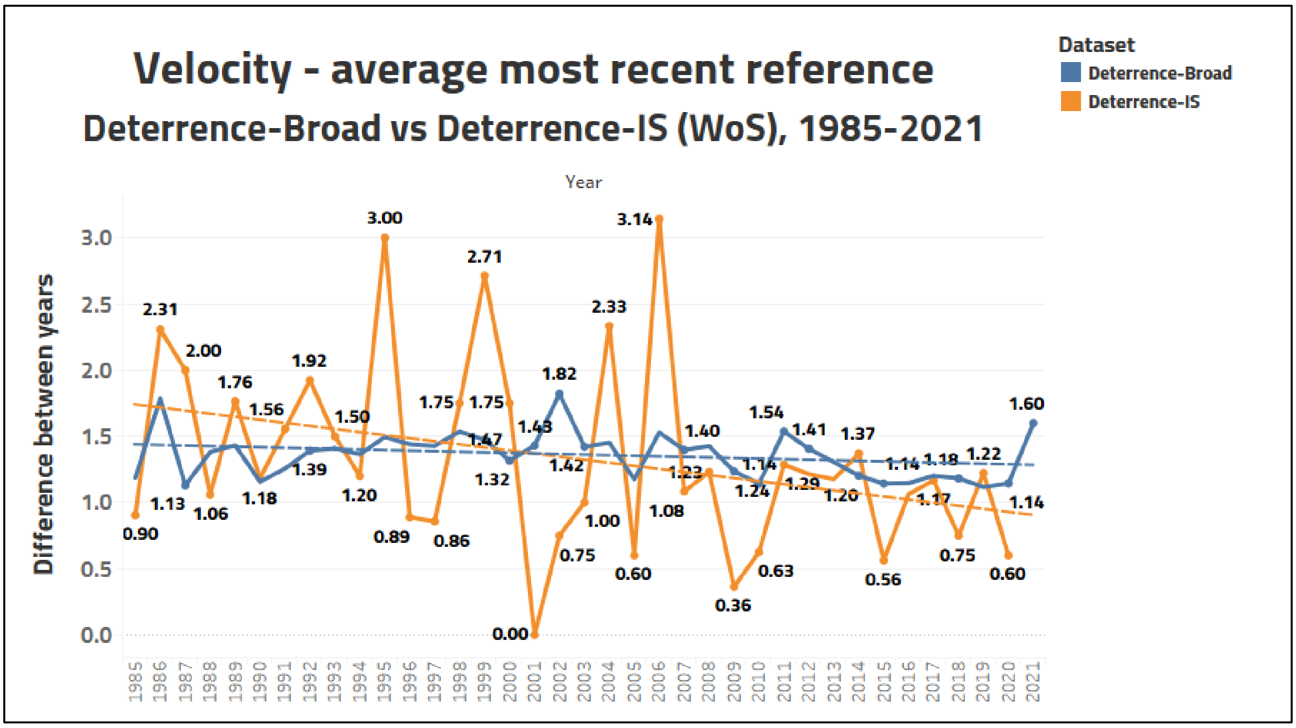


Figure 15: Velocity – average fastest citation: deterrence-broad vs deterrence-IS

Research collaboration

Our team also took a closer look at the important aspect of collaboration in the deterrence field(s)[50]. The research literature is sometimes described as a 'conversation among scholars'[51]. In this sense, scholars who communicate with one another (emailing, sharing early drafts, soliciting comments, etc.), cite, or even just consult each other's work could already be said to 'collaborate'

---

[48]Bo-Christer Björk and David Solomon, "The Publishing Delay in Scholarly Peer-Reviewed Journals," *Journal of Informetrics* 7, no. 4 (October 1, 2013): 914–23, https://doi.org/10.1016/j.joi.2013.09.001, shows that the publication process itself (from submission to publication) takes about 14 months in the social sciences. If we add a – fairly heroic – 6 months for the actual 'production' of the research paper, that leads to some 20 months.

[49]By comparison, in many STM disciplines the production process goes significantly faster: for chemistry, for instance, it takes about 5 months (Björk and Solomon); and for covid 19 just about 2 months (Serge P. J. M. Horbach, "Pandemic Publishing: Medical Journals Strongly Speed up Their Publication Process for COVID-19," *Quantitative Science Studies* 1, no. 3 (August 1, 2020): 1056–67, https://doi.org/10.1162/qss_a_00076).

[50]Ludo Waltman, Robert J.W. Tijssen, and Nees Jan van Eck, "Globalisation of Science in Kilometres," *Journal of Informetrics* 5, no. 4 (October 2011): 574–82, https://doi.org/10.1016/j.joi.2011.05.003; Marcin Kozak, Lutz Bornmann, and Loet Leydesdorff, "How Have the Eastern European Countries of the Former Warsaw Pact Developed Since 1990? A Bibliometric Study," *Scientometrics* 102, no. 2 (February 1, 2015): 1101–17, https://doi.org/10.1007/s11192-014-1439-8., L. E. Mindeli and V. A. Markusova, "Bibliometric Studies of Scientific Collaboration: International Trends," *Automatic Documentation and Mathematical Linguistics* 49, no. 2 (March 1, 2015): 59–64, https://doi.org/10.3103/S0005105515020065.

[51] Dave Harris, *Literature Review and Research Design: A Guide to Effective Research Practice*, 1st ed. (Routledge, 2019), https://doi.org/10.4324/9780429285660.

in some way. Unfortunately (and despite what we teach university students when they are being trained), journal editorial boards do not require authors to list the scholarly works they have consulted, but only the works they have cited.[52] As a result, the overwhelming (99.9%?) majority of authors of scholarly publications only list the latter. We are therefore unable to trace a large chunk of that broader ongoing 'conversation among scholars'. The field of bibliometrics can, therefore, essentially say nothing about (positive or negative) co-inspiration during the gestation process of any epistemic contribution to any field. It can only provide evidence-based on the output of that process by way of co-authorship of the final publication, which has therefore become the primary metric for scientific collaboration in bibliometrics. Let us take a look at what our datasets reveal on this.

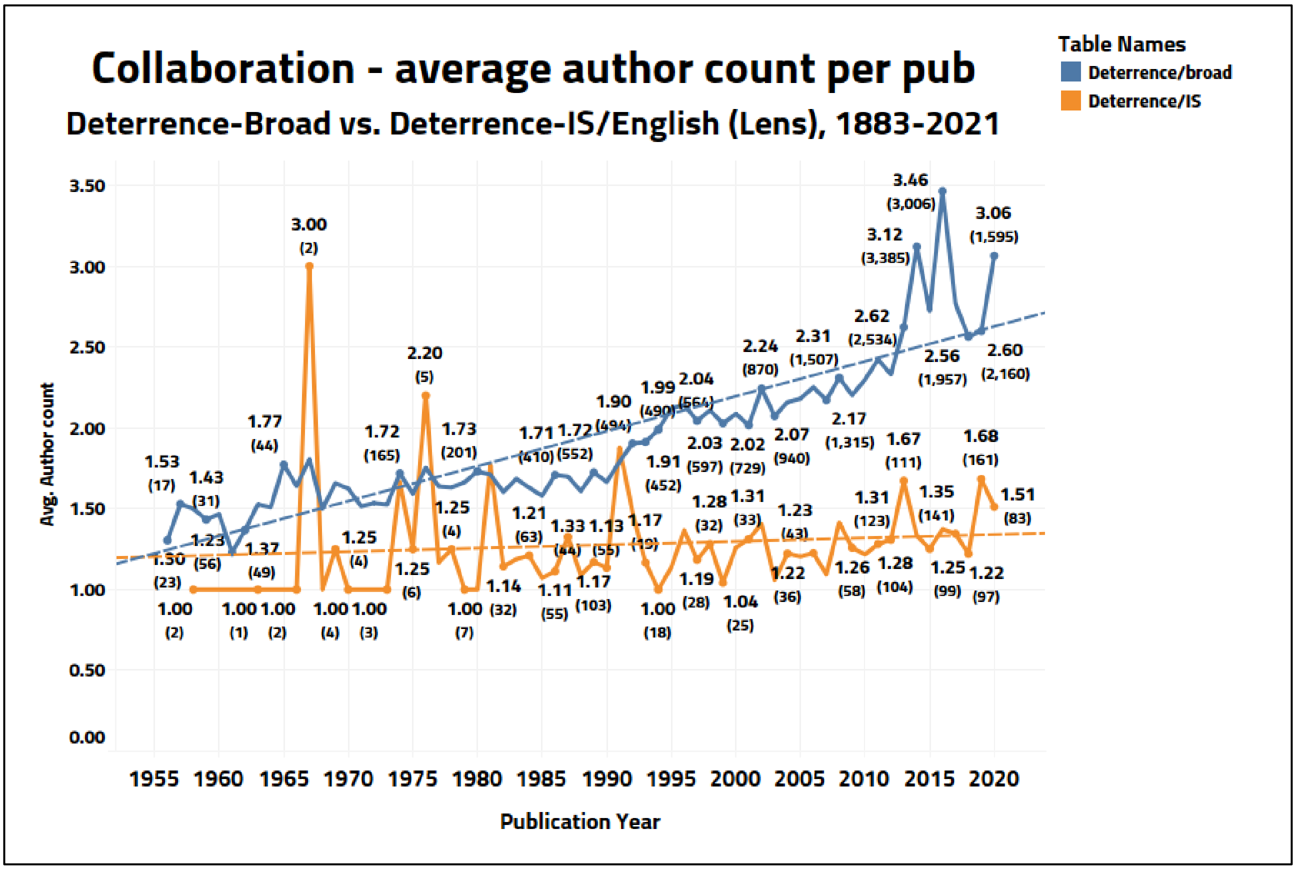


Figure 16: Collaboration - average author count per pub: deterrence-broad vs deterrence-IS

On this metric too, the Deterrence-IS (Lens) literature's track record is dispiriting. The average number of co-authors in the field over the entire time period is 1.3, whereas the equivalent value for the broader deterrence literature is almost two times higher (which is – compared to other disciplines – still quite low): 2.41[53].

[52]In our experience, most authors (and reviewers) tend to see footnotes more as proof of academic integrity (not stealing others' ideas) than as a way to traceably document epistemic linkages.

[53]The same holds true for our Deterrence-IS-Russia/English (WoS-Core) and Deterrence-IS-Russia/Russian (WoS-RSCI) datasets. The collaboration is very low with 3 being the maximum number of co-authors and 87% of articles having a single author.

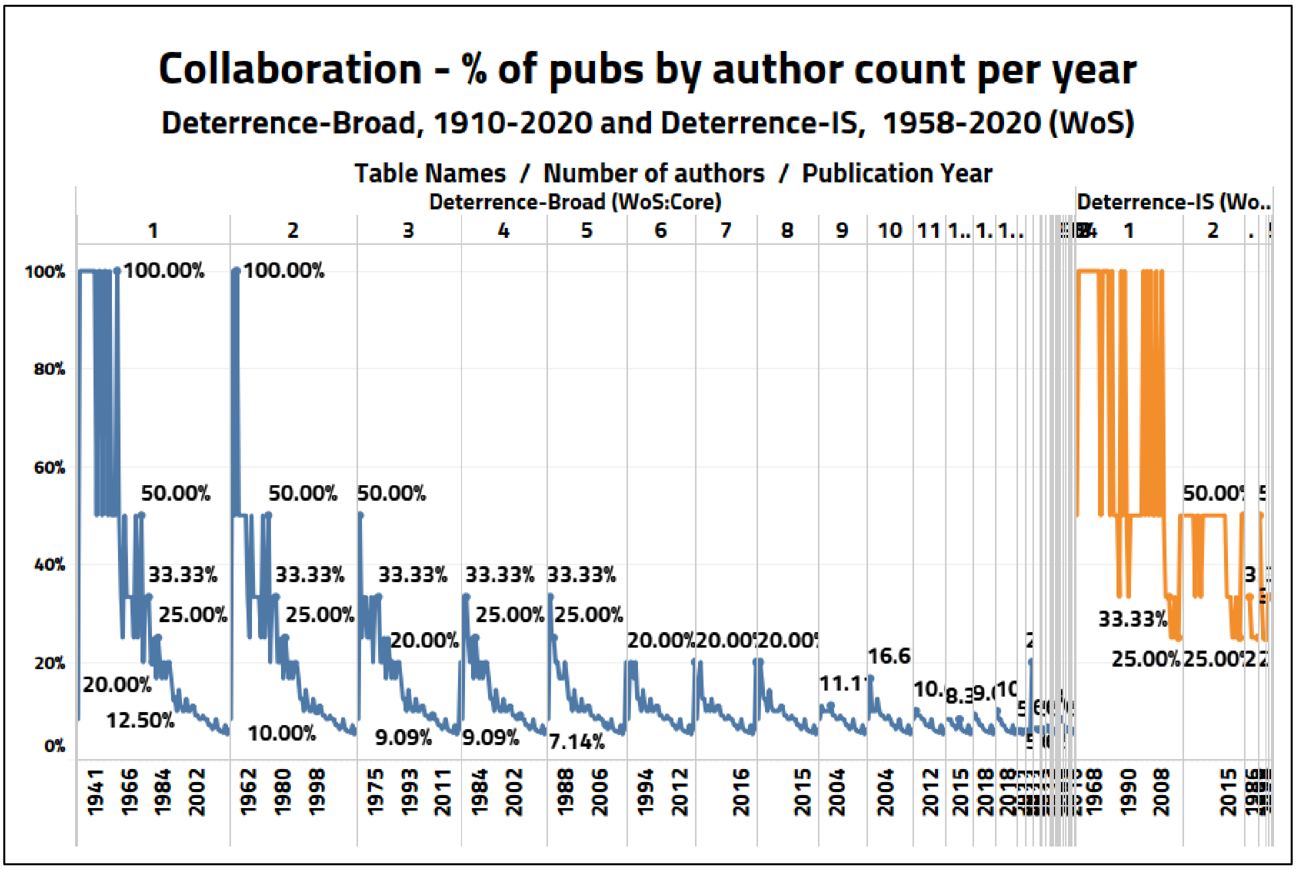


Figure 17: Collaboration - average author count per pub: deterrence-broad vs deterrence-IS, by author count

This disappointing showing also holds when we break down the previous visual over time. That shows us that the 'broad' (blue) deterrence literature goes much further to the right in terms of the number of authors; and that it also drops down much lower in the average numbers of single-, double-, triple- etc. authored publications.

As we saw was the case for the numbers of citations per document, also here we see the broader deterrence literature moving upwards and onwards, whereas the international security deterrence literature essentially stagnates at a very low baseline.

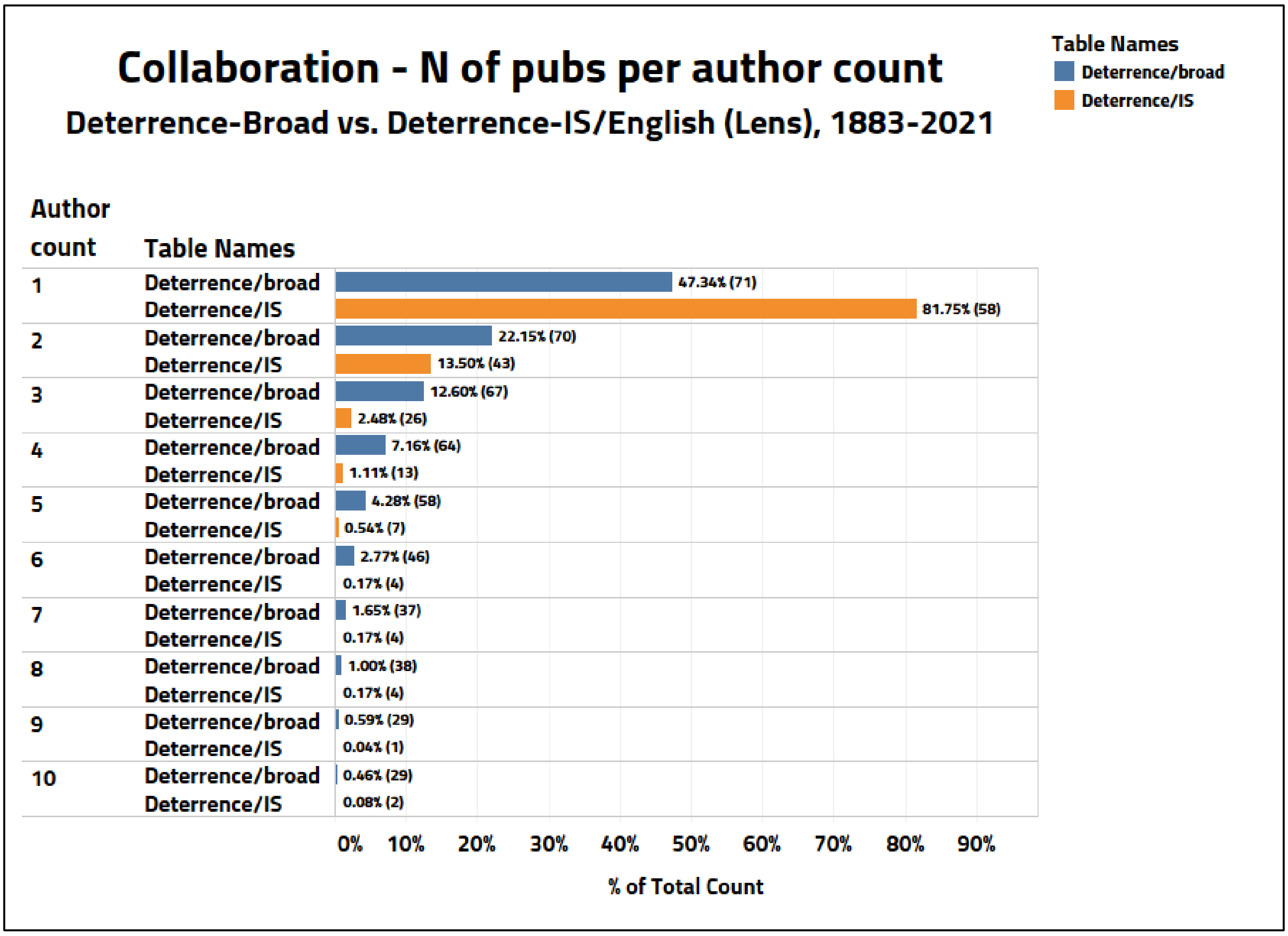


Figure 18: Collaboration - number of publications per author count: deterrence-broad vs deterrence-IS

The – in our view – most depressing finding is that more than 80% (!) of all Deterrence-IS (Lens) focused documents are single-authored, compared to about 50% (much better, even if still relatively low) of Deterrence-broad (Lens) documents. Adding dual-authored publications brings the tally up to 95% (compared to 70% for deterrence-broad).

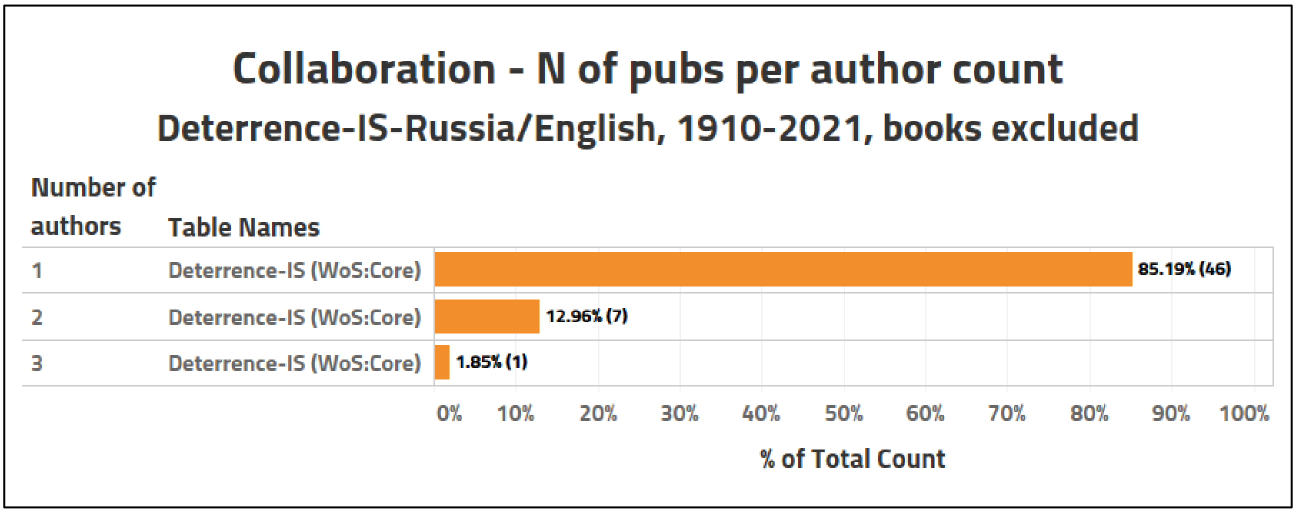


Figure 19: Collaboration - number of publications per author count: deterrence-IS-Russia

If we single out those security-related publications that contain the word Russia in their abstracts, the percentage of single-authored publications even goes up to 85%. None of them have more than 3 co-authors.

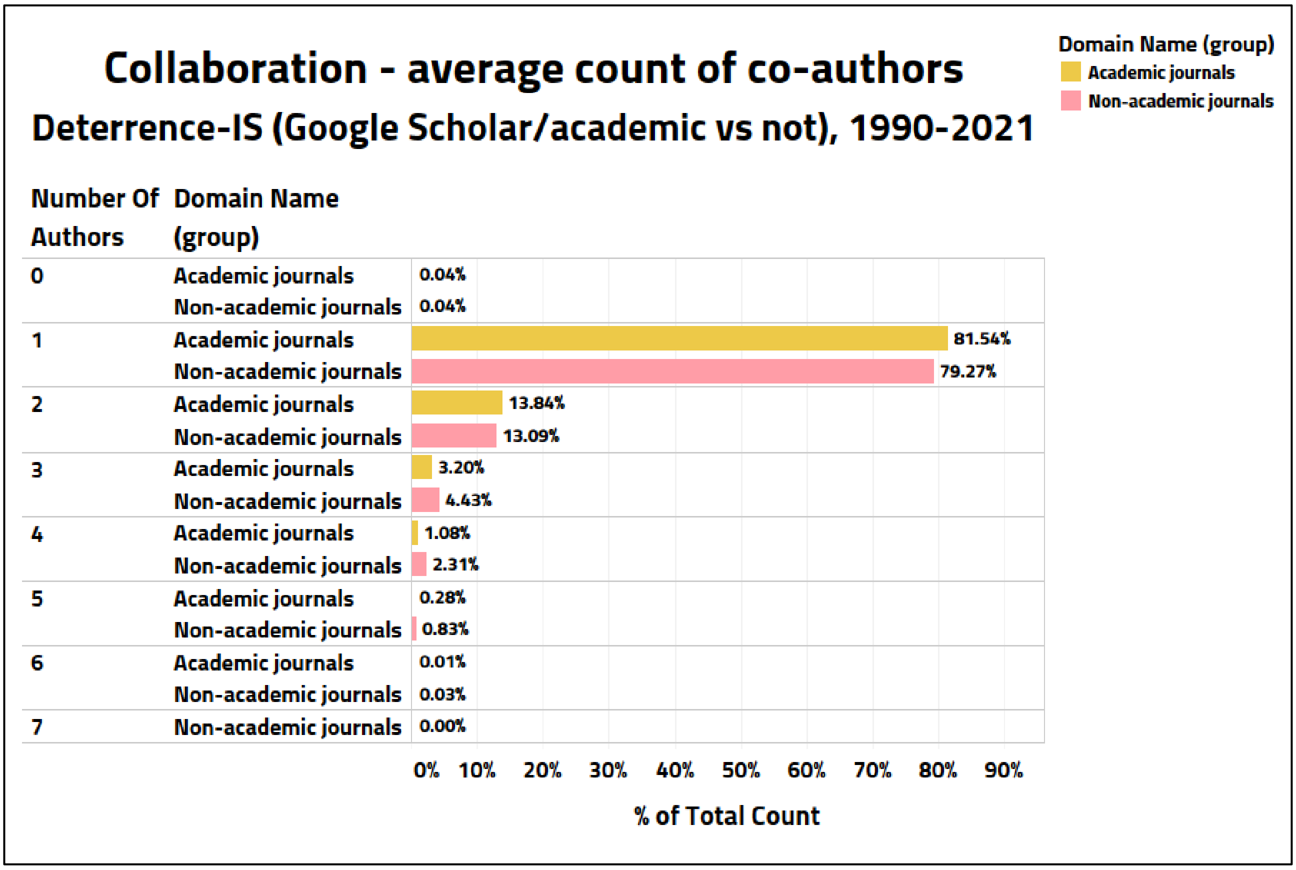


Figure 20: Collaboration – number of academic vs non-academic publications per author count

Our more 'expansive' Google Scholar deterrence-IS dataset shows that the difference between academic and non-academic collaboration is actually quite small.

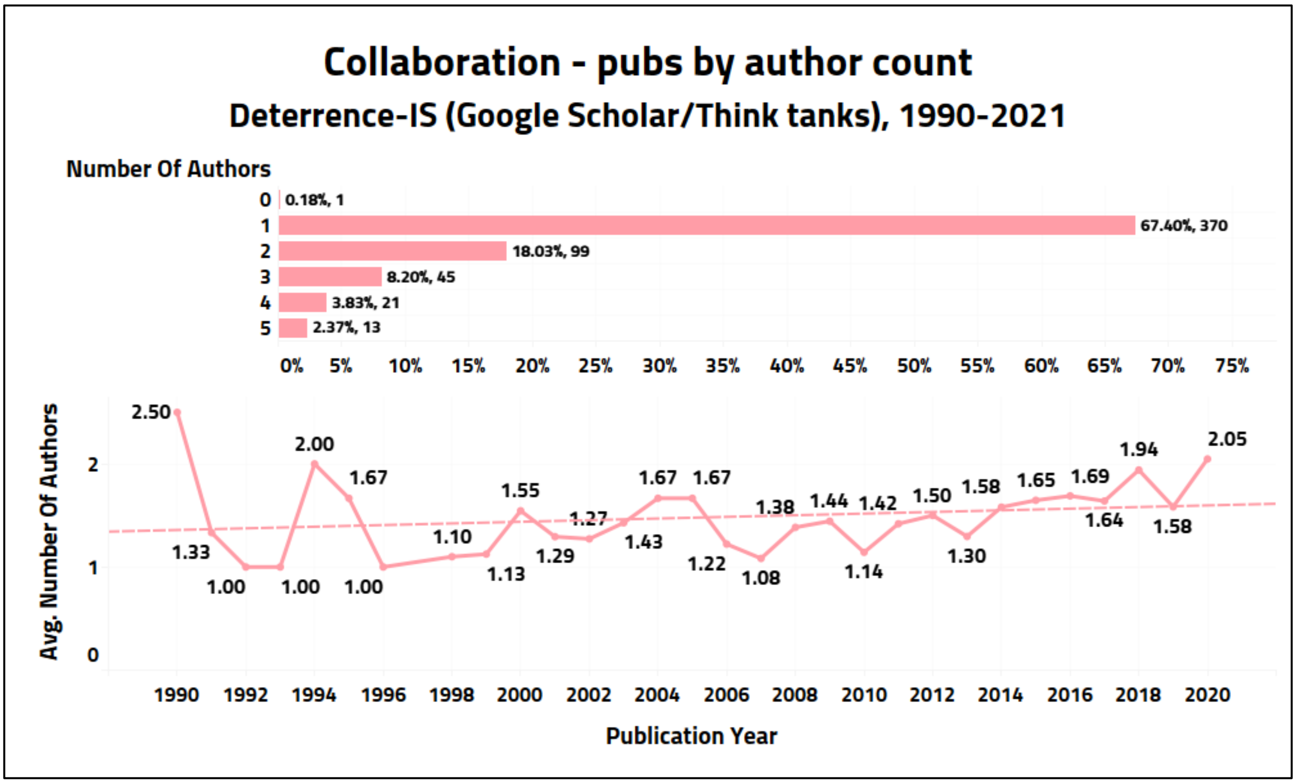


Figure 21: Collaboration – number of think-tank publications per author count+

The same Google Scholar data reveal that think-tanks are slightly more collaborative than others, but still less so than we ourselves had expected/hoped.

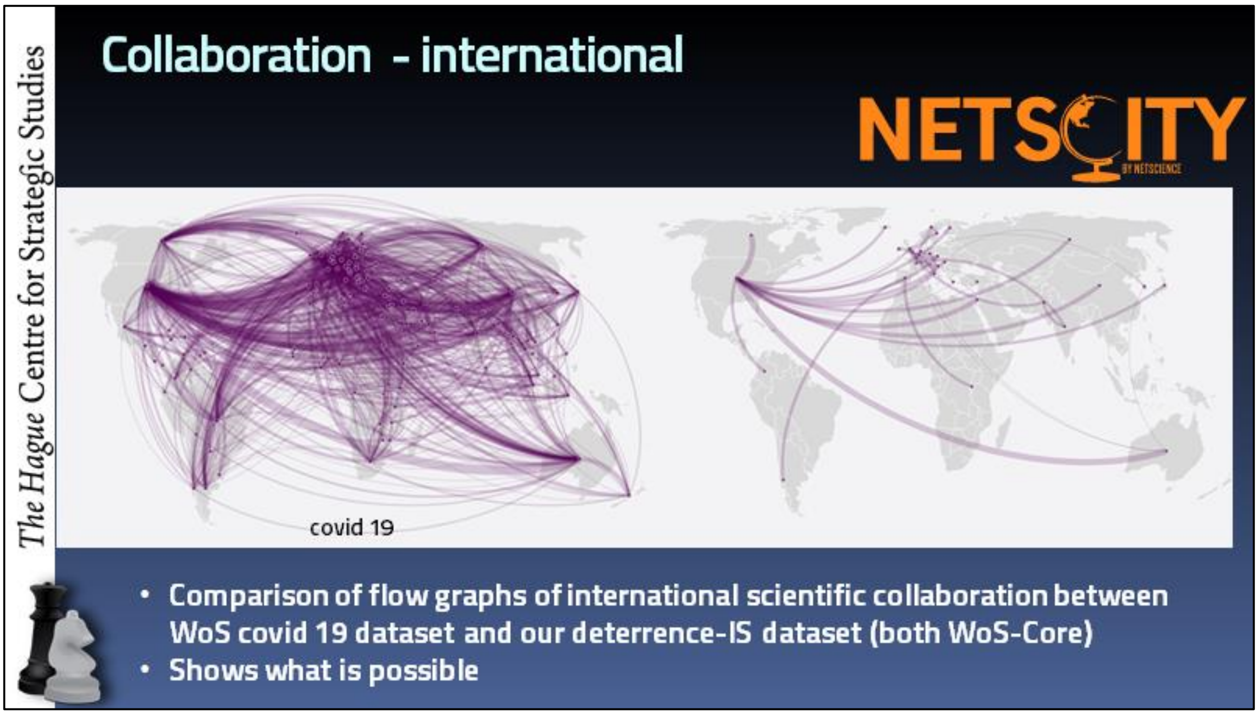


Figure 22: International collaboration - comparing deterrence-IS and covid-19

One of the most salient indicators of scientific collaboration is probably the extent to which global scientists manage to collaborate across geographical boundaries to get a better analytical grip on extreme (also policy) challenges. Figure 22 shows, on the right, a visualization that is based on HCSS' own 'Russian foreign, international and security policy' dataset[54] that was downloaded from Clarivate's 'Web of Science' and then put through Netscity's[55] author geolocation-extraction tool. To put this finding in perspective, we did the same for the 8715 peer-reviewed articles in the Web of Science database on covid 19 that were published in 2020.

The left visual shows how scientists from all over the world worked together that year to deal with the coronavirus challenge. One of the reasons global policymakers (as well as individual citizens) are interested in/concerned about deterrence is the significant loss of life it might trigger. Covid 19 has so far killed 2.5 million people[56]. The best estimate we have for the likely death-toll of a Russia-US nuclear exchange is 50 million people[57]. We would submit that this visual comparing the international 'scientific effort' to tackle these two challenges is one of the most eloquent testaments to the need for more international scientific collaboration in the FSDP/Deterrence-IS-field as well.

Research entropy[58]

One of the interesting and important aspects of any research field (like deterrence) – also for policy-makers – is the amount of epistemic 'confidence' it exhibits. Some fields may prove to be far more 'established' than others, with a majority of scholars 'buying into' the established paradigm[59]. Those fields will therefore display low entropy, and evidence-based decision-makers are likely to feel more confident basing decisions on that (mostly) 'agreed' epistemic record. Other fields, on the contrary, may be in the throes of intellectual turbulence – decay *or* effervescence – with high degrees of entropy. These may present bigger challenges to decision-makers, who then

[54]This is a dataset that we are working with for other ongoing research and that is different from the deterrence-specific datasets we used in the rest of this paper.

[55]Marion Maisonobe et al., "NETSCITY: A Geospatial Application to Analyse and Map World Scale Production and Collaboration Data between Cities" (ISSI'19: 17th International Conference on Scientometrics and Informetrics, Rome, Italy, 2019), 12.

[56] statista.com, "Coronavirus Deaths Worldwide by Country," Statista, accessed February 20, 2021, https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country/.

[57]Luisa Rodriguez, "How Many People Would Be Killed as a Direct Result of a Us-Russia Nuclear Exchange?," Rethink Priorities, June 29, 2019, https://www.rethinkpriorities.org/blog/2020/6/19/how-many-people-killed-nuclear-war; Seth Baum and Anthony Barrett, "A Model for the Impacts of Nuclear War," SSRN Scholarly Paper (Rochester, NY: Social Science Research Network, April 3, 2018), https://doi.org/10.2139/ssrn.3155983.

[58]"The information entropy of a term can be seen as a measure of its associated uncertainty. If we consider the appearance of a term as an event that transmits a message, then observing a rare event taking place is more information than observing a common event. Entropy is zero when we have nothing to learn from the occurrence of an event. The entropy reaches its maximum when the uncertainty is the highest, or, the occurrences of an event are completely random." Chaomei Chen and Min Song, *Representing Scientific Knowledge: The Role of Uncertainty* (Cham: Springer International Publishing, 2017), 168, https://doi.org/10.1007/978-3-319-62543-0. The seminal contribution on entropy in information retrieval is Claude Elwood Shannon, "A Mathematical Theory of Communication," *ACM SIGMOBILE Mobile Computing and Communications Review* 5, no. 1 (2001): 3–55.

[59]Thomas S. Kuhn, *The Structure of Scientific Revolutions* (University of Chicago Press, 2012).

have to decide how to treat this uncertainty when they ponder the impact of this new knowledge on the decisions they have to make. These tensions between different degrees of entropy do imply that any metric that can capture the amount of entropy that is inherent in any policy-relevant epistemic field should be of interest to both scholars working in this field *and* to decision-makers.

The CiteSpace bibliometric software package[60] allows us to calculate information entropy.[61] It is based on the extraction of noun phrases from the text elements of bibliometric datasets (titles, abstracts, and keywords) and measures changes in the 'vocabulary' used by the authors in the dataset. As new terms and concepts appear, they enrich the language of the discussion on this topic, thereby increasing entropy.

Russian language dataset (WoS RSCI: Deterrence-IS-Russia/Russian) has this introduction of the new vocabularies after 2014 with the largest jump between 2014 and 2016, which actually corresponds to the overall increase in the number of documents that explored different faces and forms of deterrence. English language dataset on Russia (IS-Deterrence/Russia) from Web of Science Core has a much smoother trend although some peaks are still evident, especially in the more recent year.

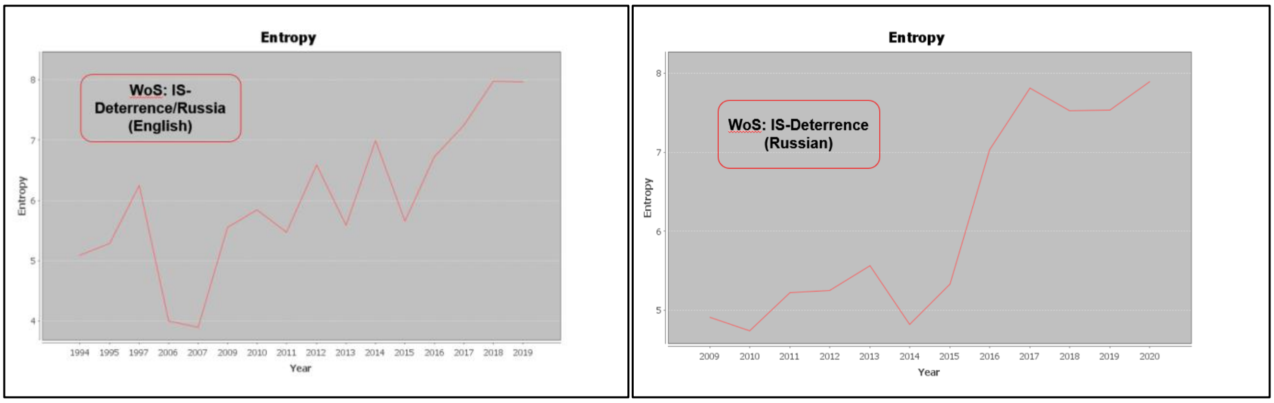


Figure 23: Entropy - changes over time: deterrence-IS and deterrence-IS-Russia

The expansion of the language in the Deterrence-IS (WoS-Core) dataset was caused by the entrance of the terms on Ukrainian conflict and economic warfare. The authors start to talk about

[60]"CiteSpace, developed by Chen Chaomei, a Chinese-American researcher at Drexel University, is one of the most distinctive and influential visualization software programs in information analysis in the United States. It comprehensively utilizes the theories and methods of the disciplines of information science, scientometrics, statistics as well as philosophy and sociology of science of science to achieves the purpose of using graphical representation of knowledge framework, structure, interaction, intersection, and derivation through the steps of data mining, processing, measurement, and drawing". (Xiaoyu Wang et al., "Visual Analysis on Information Theory and Science of Complexity Approaches in Healthcare Research," *Entropy* 22, no. 1 (January 2020): 109, https://doi.org/10.3390/e22010109.). On the program, see Chen and Song, *Representing Scientific Knowledge*, 2017; Chaomei Chen, "Science Mapping: A Systematic Review of the Literature," *Journal of Data and Information Science* 2, no. 2 (March 21, 2017): 1–40, https://doi.org/10/gfc4wk.
[61]Chaomei Chen and Min Song, *Representing Scientific Knowledge: The Role of Uncertainty* (Cham: Springer International Publishing, 2017), https://doi.org/10.1007/978-3-319-62543-0 168-169.

'economic-military power balance', 'economic sanctions' and discuss the situation in Ukraine which was not present before. Interestingly, the concept of 'conventional deterrence' is also mentioned relatively frequently (in comparison to other phrases extracted by the CiteSpace) which might signify the interest to traditional understanding of deterrence and its strengths and weaknesses. The expansion of language in the Deterrence-IS-Russia/Russian (WoS RSCI) dataset is mostly caused by more frequent usage of the concepts already present in the discussion but with higher frequency. The only exception is 'arms control' that was not used up to 2019 and then in 2020 has the frequency of 4. We discuss the evolution of terms in more detail in the ['Burst in the literature'] section.

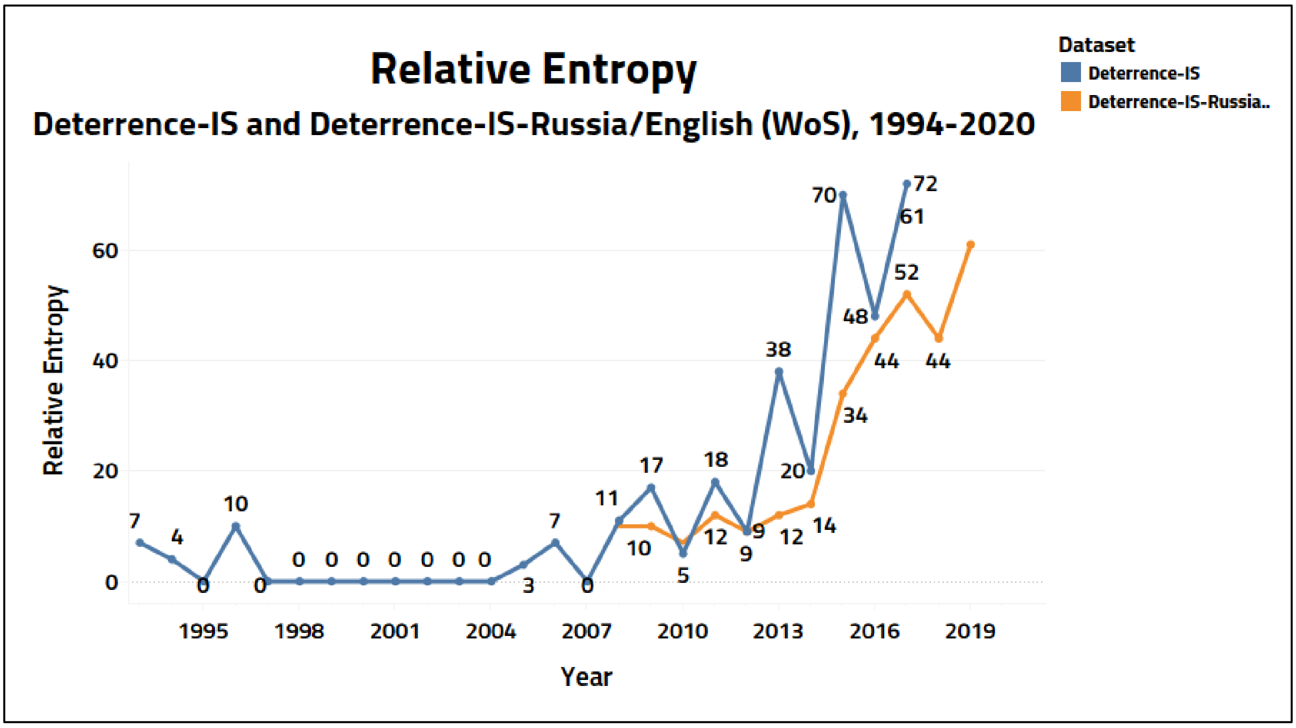


Figure 24: Entropy - deterrence-IS vs deterrence-IS-Russia

Research thoroughness

One would expect decision-makers and scholars alike to be interested in how 'reliable' findings from different disciplines are. There are, unfortunately, no available metrics (yet) for how 'thoroughly' a certain academic field has 'done its homework' or for how reliable its main findings (and/or policy prescriptions) are. Our team did nevertheless engage in an effort to uncover at least a few basic interesting indicators for those (few) aspects of the scientific 'thoroughness' of the field that can be partially collected, measured, and analyzed.

The recall challenge

Many of the metrics we have presented so far have painted a fairly bleak picture of the epistemic state of deterrence research. There is, however, at least one – slightly – more encouraging aspect of this field that our team examined in terms of 'technical' bibliometrics. It centers on the – in our opinion – greatly underestimated 'recall' problem: the extent to which the field really 'recalls' all important or useful knowledge that has been generated in it.

The field of information retrieval (IR[62]) uses the important concepts of 'precision' and 'recall' to evaluate the results of a search query[63]. Although the terms themselves may not be well-known outside of 'that' IR-field, the underlying concepts will be very familiar to anybody doing research. 'Precision' is essentially what all of us aim for when we enter search queries in a search engine. We want to be able to retrieve publications that deal as accurately as possible with the precise topic that we want to investigate. 'Recall', on the other hand, is in the first instance about what fraction of ALL available relevant documents a search query retrieves. But it is secondly also about what part of all of those available relevant publications – and therefore also of the knowledge encoded in them – is actually read/used in subsequent analyses.

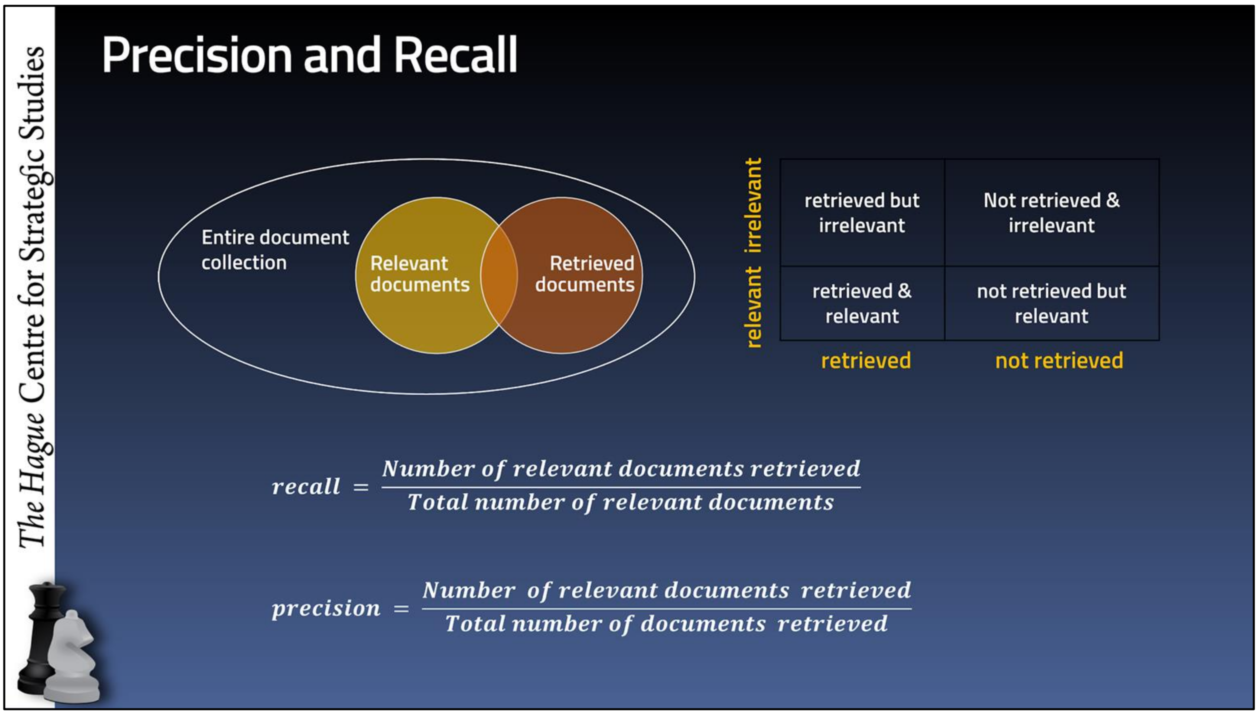


Figure 25: Thoroughness - precision vs recall

Our Deterrence-IS query, for instance, yielded 25,200 results on Google Scholar. That means 2,520 pages with 10 results each to scroll through – something Google Scholar does not allow researchers to (easily) do in the first place; and that would (until recently) have been a prohibitively labor-intensive proposition. But even now that we are (by and large) able to do this, how

[62]Not to be confused with the other (very much smaller) 'IR'- field of International Relations.
[63]Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze, *Introduction to Information Retrieval*, Reprinted (Cambridge: Cambridge Univ. Press, 2009).

can we identify all not only relevant but also truly useful (/interesting/original/solid/persuasive/validated, etc.) ones without first retrieving and reading and evaluating all of them?

In many disciplines bibliometrics can provide some analytical succor on that score. The aforementioned low citation counts in this field make this much more of a challenge. The search within our full-text corpora on Deterrence-IS finds only a few occasional mentions of the word 'bibliometrics' and none of them regarding the application of bibliometrics research to a study of deterrence. Scholars therefore typically resort – either consciously or not – to proxy 'authority' measures like an author's affiliation (e.g. reports by certain respected think tanks or publications by scholars from prestigious universities are more likely to be used/cited), some bibliometric author-level metrics like an author's g-, h- or m-index[64] (higher is deemed to be better); nationality (e.g. a scholar may want to make sure to include at least one French or Chinese point of view, but may decide to skip a Romanian or Chilean or Dutch one), etc. But such crude filters will score much better on 'quality' *precision* (i.e. they are likely to yield high-quality sources) than on quality *recall* (i.e. they will still not know what fraction of all high-quality sources the thus selected sources really represent).

The underlying question here is what we really know about the extent to which authors working in this field really do 'build' on as much of the available (especially high-quality) 'secondary' research or 'primary' sources as proves retrievable and relevant. This is an empirical question that is not easy to answer. We know that an author who starts working on some knowledge contribution to a field typically already possesses a significant amount of tacit knowledge[65] that she will draw upon in her contribution. Some of this tacit knowledge may end up in the cited references, but most of it probably will not. That author is furthermore also likely to have done some additional research for her new publication. We already mentioned how in some elite undergraduate programs, students are encouraged to list these consulted publications in their bibliography alongside the ones they actually end up citing. We also pointed out that this is not at all common practice in our field[66]. This leaves the recall issue as a potentially wide-open gaping wound in many 'international security' areas – and one that we have depressingly few diagnostic let alone therapeutic tools for.

One of the few ways to find out what part of the relevant text-encoded knowledge base au-

---

[64]Anne-Wil Harzing, "Reflections on the H-Index," Harzing.com, April 23, 2008, https://harzing.com/publications/white-papers/reflections-on-the-h-index.

[65]Tacit knowledge or implicit knowledge (as opposed to formal, codified or explicit knowledge) is the kind of knowledge that is difficult to transfer to another person by means of writing it down or verbalizing it. See Michael Polanyi, *Personal Knowledge: Towards a Post-Critical Philosophy*, 2005, https://ebookcentral.proquest.com/lib/uqac-ebooks/detail.action?docID=179903. The main author of this report never ceases to be amazed at how some authors of scholarly documents, when prodded in personal discussions, typically know so much more than what they managed to condense in their written publication.

[66]In our own experience with commissioned research, we have found our sponsors barely interested in the bibliography; and supremely uninterested in the 'consulted' lists, which, in our recent reports in which we use bibliometric research, ended up with not just a few pages of bibliographies but 100s of additional pages (in small font) of 'used' references.

thors/knowledge contributors have tapped into in their research is through their footnotes/bibliographies – the references they cite in their paper. Bibliometric databases have engaged in valiant efforts to extract these footnotes/bibliographies from scholarly publications[67]. A few (also open source) software solutions exist that allow users to do the same[68]. Unfortunately, these efforts remain decidedly suboptimal. Recent developments such as Crossref DOI, open citations[69]/metadata/abstracts, etc.[70] inspire hope, but for the time being any serious investigations into the recall problem remain extremely difficult.

There is also another reason why the recall issue may matter more than most researchers (not only in this field) acknowledge. Over the past few years, a number of fairly remarkable findings have started to raise profound questions about some of the main pillars of traditional 'quality control' mechanisms in the academic world. A more in-depth treatment of this issue is impossible within the confines of this paper, but those scholars who may be less familiar with these matters are warmly invited to take a closer look at the references we provide here on a) the peer review process[71]; and b) the so-called replicability crisis[72].

---

[67]It is important to recognize that the presence of a cited reference is in and of itself not a guarantee that this reference is accurate. The literature on this is rife with evidence of distorted citations (excluding contradictory studies – S. A Greenberg, "How Citation Distortions Create Unfounded Authority: Analysis of a Citation Network," *BMJ* 339, no. jul20 3 (July 23, 2009): b2680–b2680, https://doi.org/10.1136/bmj.b2680); reference citing publications that reached the opposite conclusions the original authors made – Kare Letrud and Sigbjorn Hernes, "Affirmative Citation Bias in Scientific Myth Debunking: A Three-in-One Case Study," *PloS One* 14, no. 9 (2019): e0222213–e0222213, https://doi.org/10.1371/journal.pone.0222213.

[68] Patrice Lopez, "Introduction – GROBID Documentation," GROBID Documentation, 2020, https://grobid.readthedocs.io/en/latest/Introduction/; Laurent Romary and Patrice Lopez, "Grobid-Information Extraction from Scientific Publications," *ERCIM News* 100 (2015); Animesh Prasad, Manpreet Kaur, and Min-Yen Kan, "Neural ParsCit: A Deep Learning-Based Reference String Parser," *International Journal on Digital Libraries* 19, no. 4 (2018): 323–37; Mark Grennan et al., "GIANT: The 1-Billion Annotated Synthetic Bibliographic-Reference-String Dataset for Deep Citation Parsing.," in *AICS*, 2019, 260–71; Dominika Tkaczyk et al., "CERMINE: Automatic Extraction of Structured Metadata from Scientific Literature," *International Journal on Document Analysis and Recognition (IJDAR)* 18, no. 4 (2015): 317–35.. For an overview, see Dominika Tkaczyk et al., "Machine Learning vs. Rules and Out-of-the-Box vs. Retrained: An Evaluation of Open-Source Bibliographic Reference and Citation Parsers," in *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*, JCDL '18 (New York, NY, USA: Association for Computing Machinery, 2018), 99–108, https://doi.org/10.1145/3197026.3197048.

[69]I4OC, "I4OC: Initiative for Open Citations," I4OC, 2021, https://i4oc.org/.

[70]Lutz Bornmann et al., "Which Aspects of the Open Science Agenda Are Most Relevant to Scientometric Research and Publishing? An Opinion Paper," *Quantitative Science Studies*, February 10, 2021, 1–31, https://doi.org/10.1162/qss_e_00121; David Shotton, "Elsevier Endorses DORA and Opens Its Journal Article Reference Lists," *OpenCitations Blog* (blog), December 20, 2020, https://opencitations.wordpress.com/2020/12/20/elsevier-opens-references/.

[71]Jeffrey Brainard, "What a Massive Database of Retracted Papers Reveals About Science Publishing's 'Death Penalty,'" *Science*, October 25, 2018, https://doi.org/10.1126/science.aav8384; Cat Ferguson, Adam Marcus, and Ivan Oransky, "Publishing: The Peer-Review Scam," *Nature News* 515, no. 7528 (November 27, 2014): 480, https://doi.org/10.1038/515480a.

[72] Initially the crisis focused primarily on psychological studies (e.g. Open Science Collaboration, "Estimating the Reproducibility of Psychological Science," *Science* 349, no. 6251 (August 28, 2015): aac4716–aac4716, https://doi.org/10.1126/science.aac4716; Stuart Ritchie, *Science Fictions: How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth*, First edition (New York: Metropolitan Books ; Henry Holt and Company, 2020).), but subsequently many similar studies also appeared in different disciplines (J. E. R. Staddon, *Scientific Method: How Science Works, Fails to Work, and Pretends to Work* (New York, NY: Routledge/Taylor & Francis Group, 2018); Committee on Reproducibility and Replicability in Science et al., *Reproducibility and Replicability in Science* (Washington, D.C.: National Academies Press, 2019), https://doi.org/10.17226/25303; D. A. Eisner, "Reproducibility of Science: Fraud, Impact Factors and Carelessness," *Journal of Molecular and Cellular Cardiology* 114 (January 1, 2018): 364–68, https://doi.org/10.1016/j.yjmcc.2017.10.009). One of the most heavy-hitting (and shocking) recent publications on this topic for the social and behavioral sciences include Duncan J. Watts, "Should Social Science Be More Solution-Oriented?," *Nature Human Behaviour* 1, no. 1 (January 10, 2017): 1–5, https://doi.org/10.1038/s41562-016-0015; Garret S. Christensen, Jeremy Freese, and Edward Miguel, *Transparent and*

The following visuals represent the findings of our analysis of the recall challenge in the field, as measured by the average number of cited references per scholarly publication.

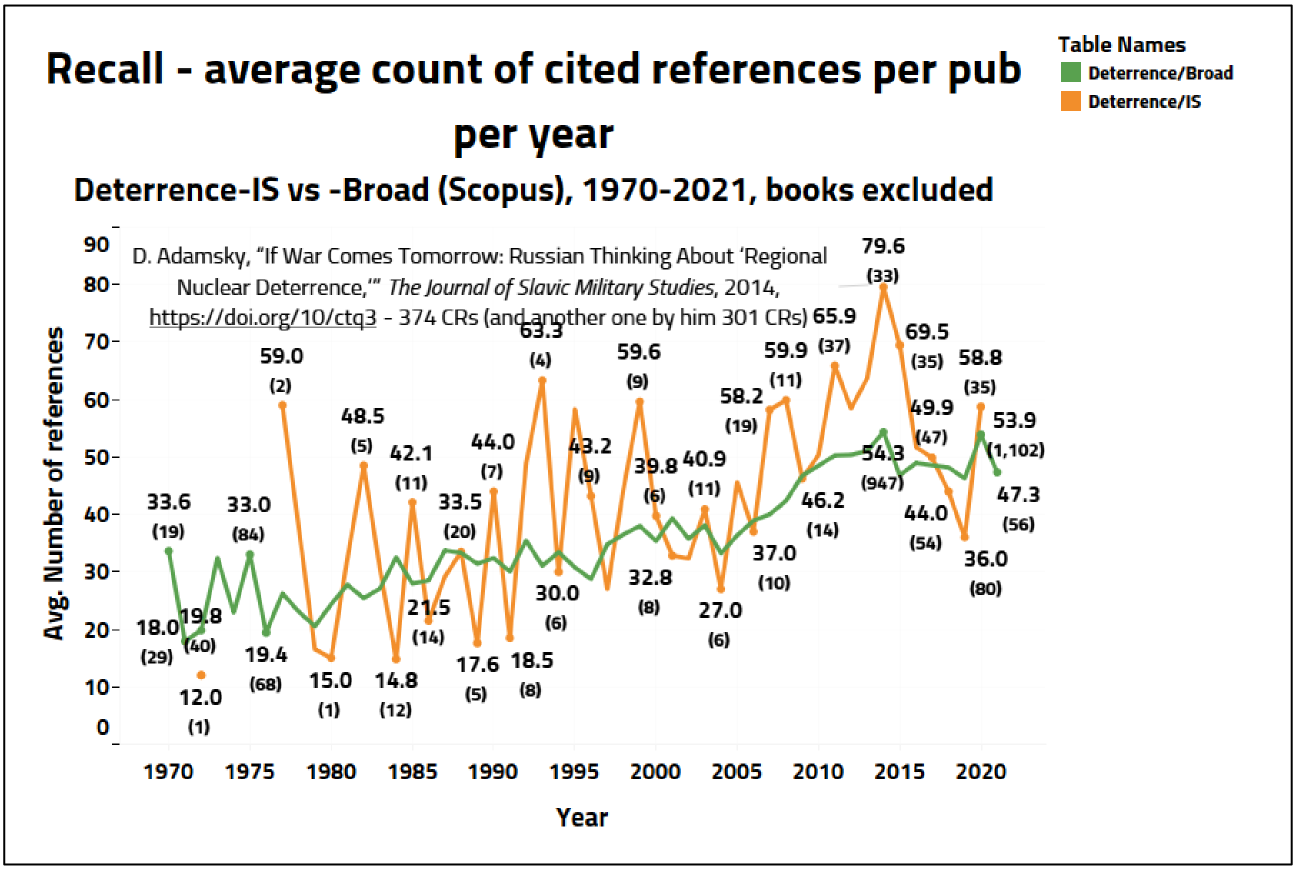


Figure 26: Thoroughness - recall: deterrence-broad vs deterrence-IS

We observe in Figure 26, that the average number of references per publication in our Deterrence-IS (Scopus) datasets [73] does not differ that much from the ones in our Deterrence-broad (Scopus) dataset (although with much more variability over time). Over all years the average of the Deterrence-broad (Scopus) dataset is 43.7, and of the Deterrence-IS (Scopus) one – 32.6. We also note a similar upward trend over time for both datasets. This suggests that, at least on this metric, the field of security deterrence aligns with the 'standard practice' in the broader deterrence field. Given the much larger overall number of publications in the 'broad deterrence' literature,

---

Reproducible Social Science Research: How to Do Open Science (Oakland, California: University of California Press, 2019). and especially the ongoing large-scale scientific research project financed by the US defense research agency DARPA on this problem Alyssa Foote, "Darpa Wants to Solve Science's Reproducibility Crisis With AI," *Wired*, February 15, 2019, https://www.wired.com/story/darpa-wants-to-solve-sciences-replication-crisis-with-robots/; Adam Rogers, "Darpa Wants to Build a BS Detector for Science," *Wired*, July 30, 2017, https://www.wired.com/story/darpa-bs-detector-science/. Initial unofficial leaks about the outcomes of this research suggest that the problem may very well prove to be as bad as initially feared (Alvaro de Menard, "What's Wrong with Social Science and How to Fix It: Reflections After Reading 2578 Papers," Fantastic Anachronism, September 11, 2020, https://fantasticanachronism.com/2020/09/11/whats-wrong-with-social-science-and-how-to-fix-it/index.html.)

[73] Also here we would like to point out that we have performed this analysis for all of our datasets, and that we will be happy to provide interested readers with access to this much more expansive analysis on our Rizzoma knowledge base.

this is definitely an encouraging data point.[74]

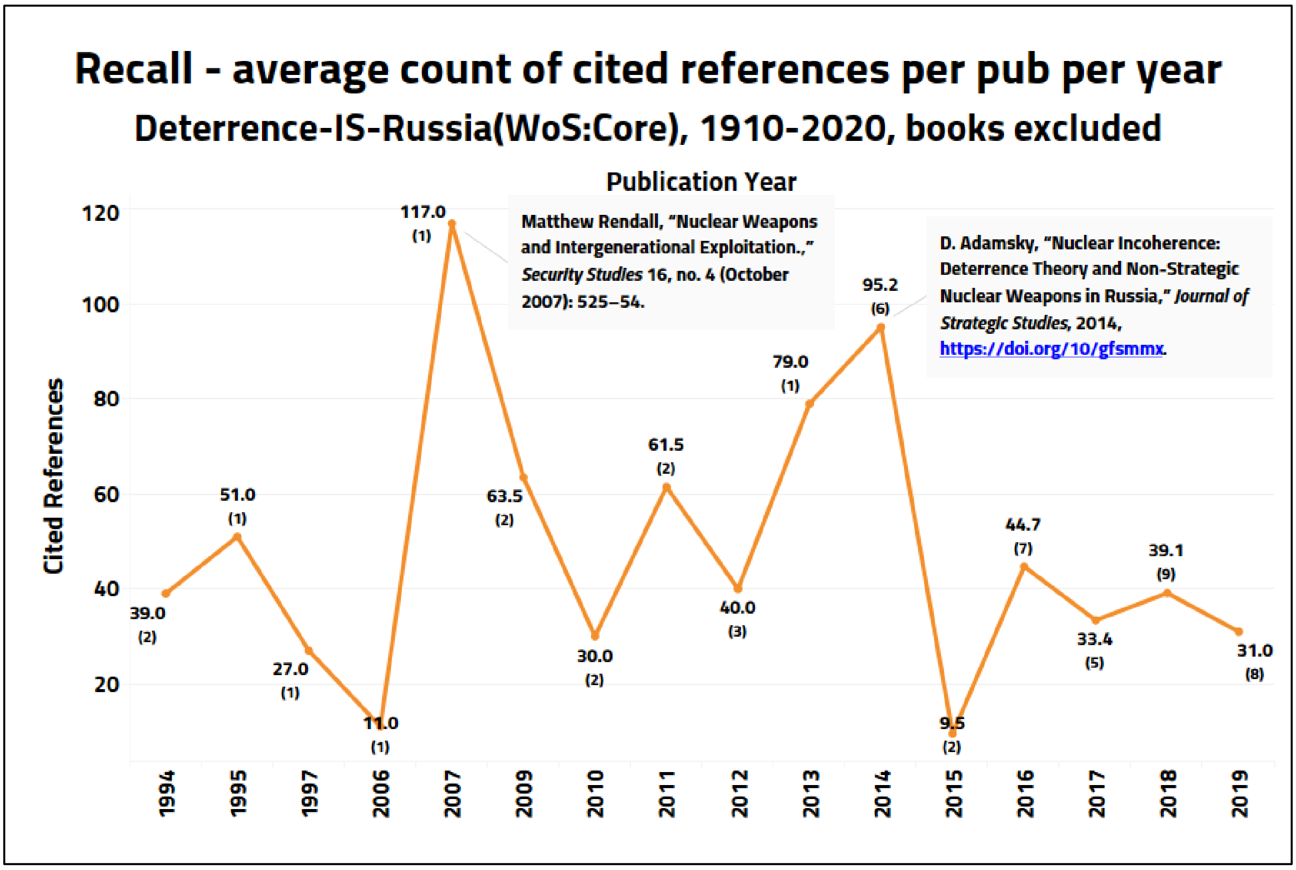


Figure 27: Thoroughness - recall: deterrence-IS-Russia

For the 65 Russia-related deterrence publications (Deterrence-IS-Russia/English) our results are even slightly more positive: we find that the average number of cited references over all these years is 46.5 – so higher than the averages for the Deterrence-Broad (WoS-Core) – 44.26 and Deterrence-IS (WoS-Core) – 32.77.

What remains unclear, however, is to what extent those numbers of references can be deemed adequate in terms of recall.

Table 1 (p. 16), which showed the numbers of documents that can be retrieved by similar search queries from different bibliometric databases, already revealed that Google Scholar beats all 'academic' databases hands down in terms of coverage. The much higher number of results reflect the fact that Google Scholar is multilingual (so searches for the term 'deterrence' will also yield Chinese, Japanese, Russian, etc. results) and also includes the so-called 'gray literature' – which includes, among others, publications from the world's think tanks – as well as sources from non-standard repositories (like Cyberleninka in the Russian case). It is therefore safe to assume that

[74]Unfortunately, Google Scholar does not provide information on cited references, which makes it impossible for us to present the evidence on academic vs non-academic publications or for think tanks.

the 25k+ sources that we were able to retrieve from Google Scholar probably represent a fair quantitative approximation of the overall amount of at least possibly 'serious'[75] publications on security deterrence. It therefore still remains an open question whether an average number of cited references of about 50 out of these 25k+ 'available' one is 'sufficient', especially when considering that many publications' cited references also include a significant number of 'primary sources' that are not included in Google Scholar – like official Russian statements on deterrence; all Russian military publications dealing with deterrence; etc. (see also our next section).

The trustworthiness challenge

Having established that the recall problem, while still representing a formidable challenge in the field(s) of deterrence-IS, at least does not appear to be worse than in the broader 'deterrence' layer of scholarship, we still have to raise the question about how reliable/validated the available knowledge in the entire 'scholarly record' in this field really is. It is at least conceivable that scholars working in some smaller research field like this one may score exceptionally well on leveraging all available knowledge in 'their' field. But if that knowledge is not trustworthy in the first place, that accomplishment would still be of limited value for further genuine knowledge building, let alone for basing expensive and/or life-or-death policy decisions on it.

One excellent example of this underlying veracity challenge was vividly illustrated by Google researchers working on the Google Knowledge Graph[76](/Vault[77]). They demonstrated that using 'popularity' in regular web search (based on just Google's page-rank algorithm, which is – mutatis mutandis – akin to the frequently used bibliometric influence metrics) misses a significant amount of high-trust web pages that a person formulating a query would be interested in, but would be unlikely to find in her results[78].

---

[75] Certainly compared to the more expansive (and problematic) Google Search. On Google Search's problem with trustworthiness, see Figure 28: Thoroughness - 'popularity' vs trustworthinessFigure 28 on Google's Knowledge Graph in the next section.
[76] Amit Singhal, "Introducing the Knowledge Graph: Things, Not Strings," Google, May 16, 2012, Knowledge graph construction techniques.
[77] Xin Dong et al., "Knowledge Vault: A Web-Scale Approach to Probabilistic Knowledge Fusion," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14 (New York, NY, USA: ACM, 2014), 601–10, https://doi.org/10/gfvn2s.
[78] Xin Luna Dong et al., "Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources," February 12, 2015, https://arxiv.org/abs/1502.03519v1; Hal Hodson, "Google Wants to Rank Websites Based on Facts Not Links," New Scientist, February 25, 2015, https://www.newscientist.com/article/mg22530102-600-google-wants-to-rank-websites-based-on-facts-not-links/.

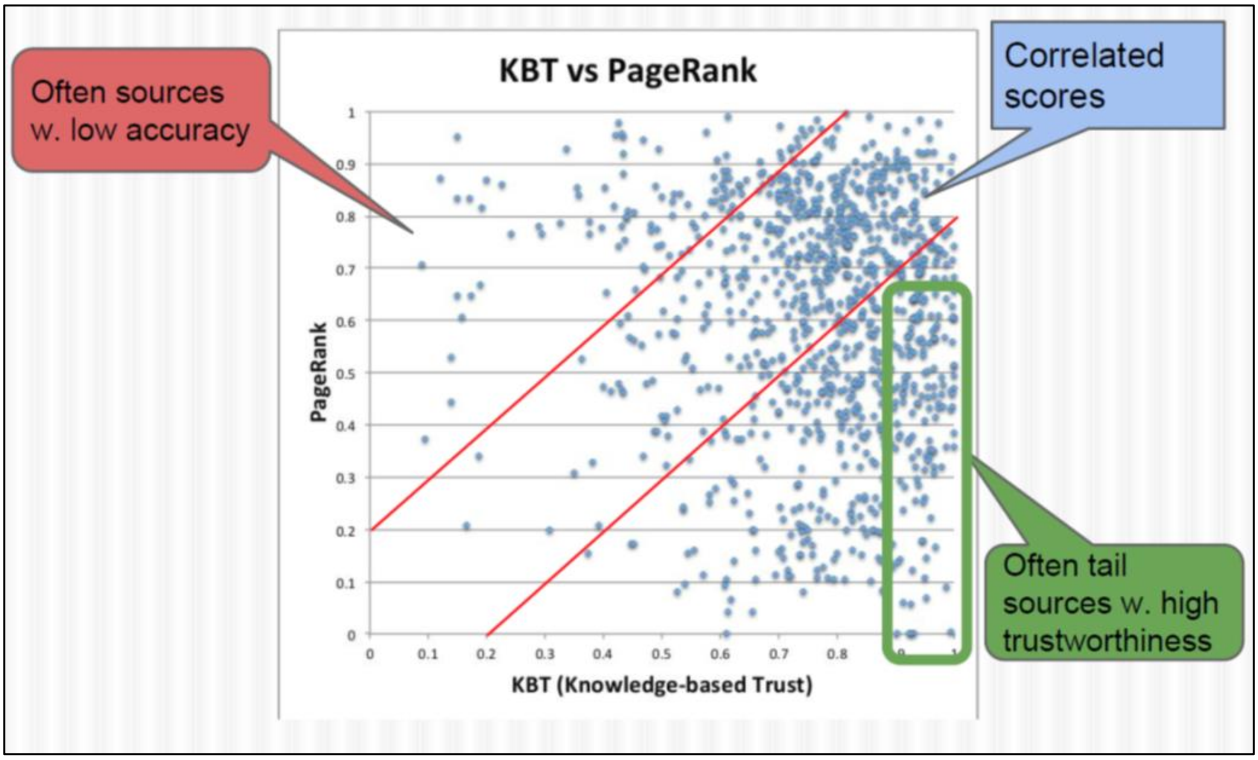


Figure 28: Thoroughness - 'popularity' vs trustworthiness[79]

Optimists might feel confident that professional scholars perform much better than 'naive' web searchers depicted on such a 'popularity vs trustworthiness' scatterplot, and that at least in the scholarly arena, the correlation between article-level bibliometric metrics like the number of citations and actual trustworthiness is significantly higher. Realists, however, are more likely to prefer to empirically test this optimistic – and widely shared – claim.

As we have indicated, we are not currently in a position to 'verify' a scholarly publication's epistemic (positive, value-added) contribution, deductive rigor or empirical trustworthiness. But one of the exceptionally promising developments in the explosion of creativity around the nexus of open-source access to the scholarly record and NLP has been the scite.ai team's recent more systematic (if still imperfect) efforts to use NLP and deep learning to ascertain whether specific empirical claims in scholarly publications are supported or disputed by other scholarly publications[80]. We performed a simple query on the term 'deterrence' in the currently 23M full-text articles in the scite_database, yielding 18,467 results. Since scite_ records do not contain any subject-related fields, we proceeded to download all of them and to use the Lens' API to obtain and join the (Microsoft Academic Services) 'Fields of study' (FoS) for all of the scite_ records in

[79]Xin Luna Dong, "Leaving No Valuable Data Behind: The Crazy Ideas and the Business" (MLConf, Seattle, Wash., May 2017), https://www.slideshare.net/SessionsEvents/luna-dong-principal-scientist-amazon-at-mlconf-seattle-2017.

[80]The tool "used machine learning and other techniques to develop a "smart citation index" called scite, which categorizes citations based on context." J.M. Nicholson et al., "Scite: A Smart Citation Index That Displays the Context of Citations and Classifies Their Intent Using Deep Learning," preprint (Scientific Communication and Education, March 16, 2021), https://doi.org/10.1101/2021.03.15.435418.

our dataset, based on their DOI-field. We then filtered out all publications with one or more verifiable claims, which left 2,924 – or about 15% – of them.

Our FoS analysis revealed that about 3% (84 out of 2,859) of all Political Science publications on deterrence contained verifiable claims. This compares to 3.4% in the subfields of international relations OR international security, with- astonishingly – absolutely no articles with verifiable claims in only the international security FoS. If we look at the ratio between supporting and disputing findings, we find about 85% of all claims supported by other publications (12 supporting/2 disputing) in International relations OR International security as well as in Political science. We can therefore find some comfort in the apparent (positive) finding that most claims that are made in this literature do seem to be confirmed in other publications.

The far weightier (and depressing) finding from Scite, however, clearly remains that only a very small number of publications in the Deterrence-IS layer of scholarly publications seems to contain any verifiable claims at all. There can be no doubt that these are very early days for these types of semantic 'meta-analyses' of epistemic claims in different fields of study. We know that the Scite_ database remains limited. We know that the algorithmic determination of publications' 'Fields of study' by MAS remains suboptimal. We know that the NLP algorithms used by Scite_ to identify these claims, and to then also determine whether they are confirmed and disputed will still be improved – potentially leading to different outcomes. We can only state that these early findings are extremely worrisome.

The empirical evidence-challenge

The final – and in our opinion most impactful – part of our examination of the scientific 'thoroughness' of the deterrence literature concerns its empirical validity: not based on discursive claims made in the literature, but on actual real-life historical, or at least synthetically generated data sources.

One of the first questions that is typically asked when a person or an organization has to make an important decision – especially in the current increasingly 'evidence-based' decision-making environment – is: what does the evidence tell us? One of the primary policy-relevant knowledge questions one might therefore legitimately expect to be addressed in the literature is whether deterrence actually 'works'. And so what do we actually 'know' about the world's experience with international deterrence? Which 'real-life' datasets do we have and what do they tell us? What have we learned from attempts to formally model (short-, medium- and long-term) dynamics and outcomes of deterrent interactions between human 'agents'? What do these findings from real-life and simulated findings tell us about whether deterrence works in the ways the deterrer intended and yields the intended results? What do we know about deterrence's Nth-order effects? Both in general, but of course especially in the field of international security?

Deterrence-broad

In a previous paper[81], some of us have surveyed the empirical findings from a number of different disciplines other than international security that have taken a close and deep look at what different disciplines have learned about real-life human deterrence in their respective contexts. We focused that overview on one of the most visible manifestations of this 'show me the beef'-line of inquiry in an epistemic field: 'meta-analyses'[82] that attempt to summarize the research findings (often in the form of ad-hoc experiments, but sometimes also based on different datasets) from various empirical research efforts/papers.

One of the most striking findings of our own (mini-)meta-analysis of different available meta-analyses of deterrence was that in many areas in which human deterrers have endeavored to strategically manipulate 'fear' in order to dissuade others from doing something unwanted (in family relations, education, the business workfloor, medicine, public health, criminal justice, policing, etc.), we found multiple, rich meta-analyses that summarized and built on ("stood on the shoulders of") rigorous large-N empirical investigations into real-life experiences with deterrence. We found none for the deterrence-IS – let alone the deterrence-IS-Russia – field(s).[83]

The substantive findings of the non-Deterrence-IS meta-analyses suggested, to our own surprise, that over time, (existential) fear may have started playing a much less dominant – or at least altogether different – role in homo sapiens' life. Concomitantly (and secondly), the *instrumentalization* of said fear in other humans in order to achieve one's own goals also appears to have been increasingly at least supplemented and, in some cases, even mostly supplanted by other – more subtle, more diversified, more 'focused', more mixed (sticks *and* carrots) – *non*-fear based (dis)suasive stratagems. Thirdly, we observed that humans have started widening the aperture of their investigations in all of these areas – both in a temporal sense and in a societal sense. The vestigial urge to immediately strike back and punish an offender whenever some injustice was deemed to have been perpetrated – also to teach that actor a lesson 'for the future' – has increasingly made way for a more considered recognition that these recurrent and reciprocal tit-for-tat retaliations impose significant long-term costs on the justice-seeker as well. At the same time, the repeated interactions with various others across different societal (also international) cleavages made humans realize that today's enemy (perpetrator, criminal, bully, etc.) may become tomorrow's (or

---

[81]De Spiegeleire et al., *Reimagining Deterrence.*

[82]For more background (and a number of caveats) about 'meta-analyses' in different disciplines, see Michael Borenstein, ed., *Introduction to Meta-Analysis* (Chichester, U.K: John Wiley & Sons, 2009); Aïda Bafeta et al., "Analysis of the Systematic Reviews Process in Reports of Network Meta-Analyses: Methodological Systematic Review," *BmJ* 347 (2013), https://doi.org/10/gb3shj; John Wallace, Bosah Nwosu, and Mike Clarke, "Barriers to the Uptake of Evidence from Systematic Reviews and Meta-Analyses: A Systematic Review of Decision Makers' Perceptions," *BMJ Open* 2, no. 5 (2012); Shannon Kugley et al., "Searching for Studies: A Guide to Information Retrieval for Campbell Systematic Reviews," *Campbell Systematic Reviews* 13, no. 1 (2017): 1–73, https://doi.org/10.4073/cmg.2016.1; Gillian Petrokofsky, "Guidelines and Standards for Evidence Synthesis in Environmental Management: Version 5.0," Monograph (Collaboration for Environmental Evidence, 2018), https://eprints.soton.ac.uk/420195/; Julian P. T. Higgins and Cochrane Collaboration, eds., *Cochrane Handbook for Systematic Reviews of Interventions*, Second edition, Cochrane Book Series (Hoboken, NJ: Wiley-Blackwell, 2020).

[83]The field of International Political Economy shows that it is certainly possible to do this – see footnote 6.

even already be today's ally) in other fields, therefore greatly mitigating the urge to 'play hardball'. Fourth, in most of these areas we detect a growing recognition of the many also negative first and Nth-order effects of a fundamental deterrent posture on actors' mentalities and behaviors. To give but one example: high incarceration rates are increasingly recognized as not only not being an effective deterrent that might help in lowering future crime rates, but as even radicalizing criminals – leading to more and worse future recidivism at an unusually high cost to those societies with high incarceration rates. All of these – very recognizable, also in the deterrence-IS-field – dynamics feature quite prominently in the 'does deterrence even work?' debates in other Deterrence-human knowledge layers.

Deterrence-IS

What is the actual evidence base for deterrence in international security? Either based on actual longitudinal 'historical' datasets or modelling efforts?

Deterrence-IS: Real-life datasets

There is a (very) small set of scholars that have attempted to construct datasets about international deterrence and to investigate these. The first systematic attempt to construct a dataset on Deterrence-IS we were able to track down was published back in 1988 by Paul K. Huth, who collected 58 cases of extended-immediate deterrence and evaluated whether these ended up in success or failure[84]. He was followed by Erich Weedy, who identified roughly the same number of cases (although for a much narrower period of time) and defined success of deterrence as the prevention of military conflict[85]. In the second decade of the 21st century Todd S. Sechser[86] and Johnson et al[87] compiled two bigger datasets on topics adjacent to deterrence: extended deterrence threats with defense pacts and compellent threats. The most recent (still small-N) effort to meticulously analyze the phenomenon of deterrence was made by Michael Mazarr and some of his colleagues from RAND[88], who analyzed 39 cases of US-led extended deterrence between

---

[84]Paul K. Huth, "Extended Deterrence and the Outbreak of War," *The American Political Science Review* 82, no. 2 (1988): 423–43, https://doi.org/10.2307/1957394.
[85]Erich Weede, "Extended Deterrence, Superpower Control, and Militarized Interstate Disputes, 1962-76," *Journal of Peace Research* 26, no. 1 (1989): 7–17.
[86]Todd S. Sechser, "Replication Data for: Militarized Compellent Threats, 1918-2001" (Harvard Dataverse, January 2, 2013), https://doi.org/10.7910/DVN/VDJQ1E; Todd S. Sechser, "Militarized Compellent Threats, 1918–2001," *Conflict Management and Peace Science* 28, no. 4 (September 1, 2011): 377–401, https://doi.org/10.1177/0738894211413066.
[87]Jesse Johnson, Brett Ashley Leeds, and Ahra Wu, "Capability, Credibility, and Extended General Deterrence" (Harvard Dataverse, May 22, 2015), https://doi.org/10.7910/DVN/LT9TPB; Jesse C. Johnson, Brett Ashley Leeds, and Ahra Wu, "Capability, Credibility, and Extended General Deterrence," *International Interactions* 41, no. 2 (March 15, 2015): 309–36, https://doi.org/10.1080/03050629.2015.982115.
[88]Michael Mazarr et al., What Deters and Why: Exploring Requirements for Effective Deterrence of Interstate Aggression (RAND Corporation, 2018), https://doi.org/10.7249/RR2451.

1945-2018. The table below shows the numbers of cases in the datasets compiled by the above-mentioned analysts and the reported success rates of deterrence (or of the application of other related strategies).

Source	Timespan	Number of cases	Number of successful cases	Number of failed cases	Success rate
Extended Deterrence and the Outbreak of War, Huth, 1988	1885-1984	58	34	24	59%
Extended Deterrence, Superpower Control, and Militarized Interstate Disputes, Erich Weede, 1989	1962-1976	57	53	4[89]	93%
Escalation of Great Power Disputes: Deterrence Versus Structural Realism, Huth et al.[90], 1993	1832-1984	97	54	43	56%
Militarized Compellent Threats, Todd S. Sechser, 2011	1918-2001	242	109[91]	133	45%
Capability, Credibility, and Extended General Deterrence, Johnson et al., 2015	1816-2000	1085	We were unable to extract data, but the conclusion of the authors is the following: "We find that defense pacts with more capability and more credibility reduce the probability that a member state will be a target of a militarized dispute. We also find that states can affect the capability and credibility of their extended deterrent threats through alliance design"		

[89]This extremely low number of failures compared to other works probably can be attributed to a very narrow definition of deterrence success. In this work only war is considered as a failure of deterrence, while other works define a failure of deterrence as a situation, when the target of deterrence changes the status quo (which does not necessarily mean war), despite the threats.

[90]Paul Huth, Christopher Gelpi, and D. Scott Bennett, "Escalation of Great Power Disputes: Deterrence Versus Structural Realism, 1816-1984: Version 1" (ICPSR – Interuniversity Consortium for Political and Social Research, 1995), https://doi.org/10.3886/ICPSR06355.V1; Paul Huth, Christopher Gelpi, and D. Scott Bennett, "The Escalation of Great Power Militarized Disputes: Testing Rational Deterrence Theory and Structural Realism," *The American Political Science Review* 87, no. 3 (1993): 609–23, https://doi.org/10.2307/2938739.

[91]Author differentiates between instances where the target fully complied (95 cases) and partially complied (14 cases).

What deters and why, Michael J. Mazarr et al., 2018	1945-2018	39[92]	24	11	62%

Table 2: Deterrence-relevant datasets

As we see from the table deterrence is far from being a 100% successful strategy. If we exclude one outlier, the success rate hovers around 60% – not terrible, but certainly also not great. [It is also noteworthy that the application of suasion for changing someone's behavior (compellence threats) in contrast to preventing someone from changing behavior (deterrence) has even lower success rates.]

Deterrence-IS: Synthetic datasets

Attempts at purposive international deterrence, like many other defense and security topics, clearly do not lend themselves easily/ethically to real-life 'randomized' experimentation[93]. In cases like this, one would expect an extra-intensive focus on building rich synthetic formal models or simulations that could still enlighten decision-makers (and scholars) about the various dynamics or trade-offs that are inherent to any deterrent interactions between 'agents'. Our team set out to look for such publications but came back – also here – quite disappointed.

The past decade did witness the appearance of a few research efforts that aimed at building predictive models for deterrence. We present these works briefly in the table below.[94]

Source	**Type of modeling**	**Main findings**
An Agent-Based Approach on Conditional Deterrence, Yang et al., 2018	Agent-based modeling	However, it is still difficult to find an integrated predictive model for deterrence that takes account of both global and regional interactions in an increasingly globalized and concurrently fractured security environment.

[92]Four cases were still ongoing at the time of publishing.

[93]We also want to point out that we were quite astounded to find that a topic as important as this one has not received, to the best of our knowledge, any attempt at quasi-experimentation. "Like randomized experiments, quasi-experiments are used to estimate the effects of one or more treatments on one or more outcome variables. The difference is that quasi-experiments do not have random assignment to treatment conditions. Instead, the treatment effect is estimated by making comparisons across cases that are exposed to different treatment conditions in some nonrandom fashion, and/or by comparisons across time (before and after treatment implementation), and/or by other kinds of comparisons discussed later." Melvin M. Mark and Charles S. Reichardt, "Quasi-Experimentation," in *Handbook of Applied Social Research Methods*, ed. Leonard Bickman and Debra J. Rog (SAGE, 1998), 183. As the next generation of automated event datasets comes online, we suspect we will be seeing much more of this.

[94]While these try to use a more rigorous mathematical approach to modeling, the method does not remain uncontested. RAND's Paul Davis, for instance, advocates a more limited approach and claims that "the most important insights gained from decision modeling could be obtained with simple models that could be reduced to figures, tables, and a story". For example,Paul K. Davis, *Simple Models to Explore Deterrence and More General Influence in the War with Al-Qaeda*, RAND Occasional Paper Series (Santa Monica, Calif.: RAND National Defense Research Inst, 2010).

Deterring the development and use of nuclear weapons: a multi-level modeling approach, Carley et at., 2018	Agent-based dynamic network modeling	The model supports development of generic models for a problem domain that can be rapidly re-used through their lack of specificity and their general application of the underlying theories. With additional effort, and appropriate use of collected and valid data, additional and more specific models can support policy analysis as well as virtual experimentation. The results presented are illustrative although congruent with other models using the same scenario and interventions. As such, the results should not be taken as providing definitive strategic guidance for the Pacific Rim in the current time period.
Sociocultural Models of Nuclear Deterrence, Morgan et al., 2017	Agent-based sociocultural modeling	Models generated through this process are intended to serve as points of discussion of possible unanticipated consequences or benefits of a suggested set of interventions, not to serve as definitive guide to future events. The reception of the Northeast Asia scenario models by independent SMEs suggests that the approach may be useful in this capacity.
Testing Deterrence. An Agent-Based Modeling Approach, Kang and Compton, 2008	Agent-based modeling	• Regions of deterrence often exist, but under some conditions there is no deterrence at all; • WMD do not stop highly risk acceptant actors from initiating conflict; • Salience for political vs. economic interests plays a big role; • With increase in severity of conflict comes increase in potential gain; • Model demonstrates that conditions for initiation exist where the attacker has a nuclear disadvantage. Real-world example: Israel.

Table 3: Deterrence-relevant modelling efforts

Deterrence-IS: Where's the beef?

So also here, with respect to deterrence's evidence base, the 'recall' problem seems to loom large. What exactly is the 'universe' of deterrent events in international relations? And do we just want to focus on the 'tip of the iceberg' – the analytical attempt to figure out whether strategic deterrence (deterring 'Hitler', 'Stalin', 'Xi', 'Putin', 'Al Qaeda' – but – mutatis mutandis – also 'Bush', 'Blair', 'Sarkozy', 'Merkel', 'Trump', etc.) worked when looked at from the proverbial 30'000 feet vantage point? Or should we also try to identify and analyze the probably tens of thousands of daily (!) attempts by probably all international actors to 'deter' third parties at tactical, operational *and* strategic levels? Are the examples of constant sable-rattling we currently witness from Russia over NATO-territory (and vice-versa), from China over Taiwan (and from 'the West' over China), between China and India, etc. single deterrence cases with binary outcomes (deterrence works or not)? Or are they better viewed as consisting of constant stream of a number of different types and gradations of deterrence that are constantly and adaptively being modulated by the participating actors based on tension dynamics? Is it even methodologically acceptable to 'pick on the dependent variable[95]' – i.e. investigate cases where deterrence was attempted, and then find out whether that succeeded or not? Or should we not instead strive to also include all cases where actors opted for other strategies – to be able to compare the 'relative' utility of deterrent options?

These questions are not purely theoretical. We are currently witnessing the emergence of much larger datasets of 'international events'[96], that are leveraging the availability of ever more media sources *and* of more powerful and sophisticated NLP-algorithms that can extract events from them. Event datasets are the result of monitoring the news of the world and distilling the different events that took place by source and target actor as well as the nature of the event itself (i.e. a triple: 'who (1) did what (2) to whom (3)?')[97]. This basic 'triple' is also augmented with other features like 'when' (4) and 'where' (5). All of these events are extracted from media sources that also have metadata like the publication source, date, article title, URL, etc. Historically, in order to create a dataset of events taking place, the sources (e.g. news articles) had to be coded manually, but thanks to technological advancements this process has become more automated, making it possible to create enormous, all-encompassing databases with truly global sources lists and a long timeframe.

The largest event dataset, called GDELT (Global Database of Events, Language, and Tone) has been developed by Kalev H. Leetaru. It covers the period from 1979 to the present and is updated every fifteen minutes.[98] GDELT has two versions – English and Translingual. The latter includes media sources in 65 other languages. Integrated Crisis Early Warning System (ICEWS) covers the period since 1995. It was developed by Lockheed Martin for the Defense Advanced Research Projects Agency (DARPA) and the Office of Naval Research (ONR) and later released for public

---

[95]On the dangers of selection bias (a long-standing issue in econometrics) in the social sciences, see Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, N.J: Princeton University Press, 1994), 128–38. On selecting on the dependent variable, see also Barbara Geddes, "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics," *Political Analysis*, 1990, 131–50; David Collier, James Mahoney, and Jason Seawright, "Claiming Too Much: Warnings About Selection Bias," *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2004, 85–102.

[96]Khrystyna Holynska et al., "Events Datasets and Strategic Monitoring" (The Hague, 2020).

[97]Philip A Schrodt, "Analyzing International Event Data: A Handbook of Computer-Based Techniques" (March 22, 2012), http://eventdata.psu.edu/papers.dir/automated.html.

[98]The GDELT Project, "The GDELT Project," The GDELT Project, 2019, https://www.gdeltproject.org/.

use.[99] "TERRIER" (Temporally Extended, Regular, Reproducible International Events Records) includes events from 1979 to 2016 that are derived from the LexisNexis Complete Collection.[100] All of the above-mentioned US-based[101] event datasets use different white-lists, different coding engines, and various other (important) technical details such as deduplication algorithms, actor dictionaries, etc. But they do all use the same event coding scheme for actors and events (CAMEO)[102]. This enables researchers to compare and analyze their findings. Usefully for our topic under examination, the CAMEO event coding scheme contains a "Threaten" root code. Figure 29 shows the trend over time (from January 2014 until May 2021), of Russia (as the source actor) threatening (based on that CAMEO event code) all other countries (as target countries) in five different automated event datasets. As an example, we have annotated the first peak of April 2014 (Crimea), in which GDELT English automatically extracted 353 events (or 0.4% of all international dyadic events in that dataset) in which Russia threatened another country. The raw numbers vary across the datasets, but we observe that three out of the five all show that same peak. Threatening a country is not the same as deterring. But the fact that events can be extracted from sources algorithmically and analyzed more systematically still demonstrates that we should also be able to extract deterrent events and build a dedicated dataset that is immeasurably more granular that the ones we surveyed in this section. The question remains why, on a topic of such enormous importance, this has not been done.

[99] Elizabeth Boschee et al., "ICEWS Automated Daily Event Data (IN TESTING)," October 11, 2018, https://doi.org/10.7910/DVN/QI2T9A.
[100]OU Event Data, "Terrier," August 28, 2018, http://terrierdata.org/#about.
[101] There are some European efforts as well (most notably the event extraction tools that the European Union Joint Research Centre uses in their European Media Monitor), but they do not tend to be ontology-based.
[102] Philip A. Schrodt, "Cameo: Conflict and Mediation Event Observations Event and Actor Codebook," 2012, http://eventdata.parusanalytics.com/data.dir/cameo.html.

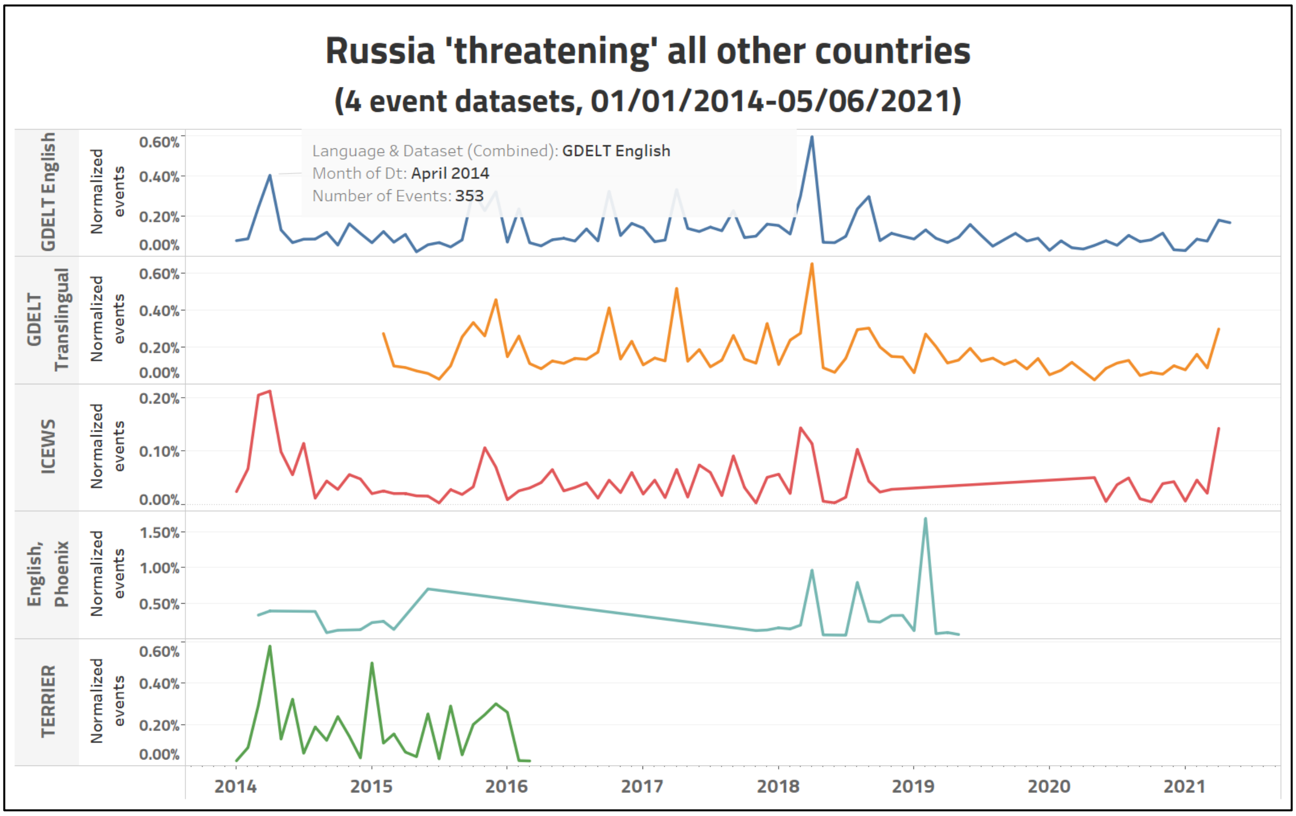


Figure 29: The 'Threaten' CAMEO-code in the automated event datasets.

What do the numbers say?

The first, more 'technical' part of our 'MRI-scan' of the publicly available body of knowledge about (also Russian) deterrence that has been encoded into texts by various scholars over the past few decades has revealed a number of potentially quite consequential flaws. Based on the evidence we have adduced, we feel quite confident in asserting that the volume, velocity, collegiality and uptake of publicly available scientific insights into 'security deterrence' remain decidedly suboptimal.

We consider our claims about the scientific thoroughness of this field compelling, but less conclusive. We do not (yet) have the 'trustworthiness' metrics that would allow us to claim with confidence that the literature 'recall' problems that we did lay bare also constitute a genuine epistemic problem that should be of concern to decision-makers who consult this body of knowledge – let alone base decisions on it. We find the (in our view) unbearable lightness of the empirical evidence on international security deterrence far more damaging. Much of the literature[103] on Russian deterrence, for instance, reads more like an at times almost scholastic exegesis

[103]Our team is still working on finding (semi-supervised-NLP-based) ways to identify which fraction of publications also take into account actual pieces of 'real-life' evidence of Russian deterrence – i.e. in the purely military realm evidence like funding, acquisition, deployment, fueling, alert-status, etc. choices.

of Russian writings on these issues than like an attempt to systematically collate, parse and validate all available stated and revealed Russian preferences/choices.

With respect to the 'exegesis' that towers so dominantly over the deterrence-IS literature itself we can only observer and regret that we do not – as yet – possess the tools to assess the logical (deductive) rigor, coherence and consistency of the main precepts of deterrence theory. We can only observe that the initial cocky confidence in the 'logic' of deterrence from the 50s and 60s of the previous century and that seemed so unassailable in Cold War writings has made place for a much more qualified and tentative attitude in the post-Cold War period[104].

In most disciplines, and especially in those that have humans in the loop and that have real-life policy implications, deductive logic is typically deemed necessary but not sufficient. The evidence-based turn in policy analysis can arguably partially be seen as a recognition that some of the supposedly iron-clad deductive 'laws' that were to guide presumably Pareto-superior policy improvements proved to be considerably more tenuous than initially posited. In the field of economics, for instance, the 'Chicago School' neoclassical theoretical paradigm that Ronald Reagan in the US, or Margaret Thatcher in the UK so forcefully propagated and implemented 'in real life' did deliver a number of fairly undeniable structural successes. But that neoliberal set of policies also played an important role in the events leading up to the 2008-9 financial-economic crisis and the subsequent (and ongoing) populist backlash against 'laissez-faire' globalization, the consequences of which we – in many ways – are still suffering to this day. The more pragmatic, experiential, evidence-based turn in policy analysis ("let's just find out, based on honest data, where market mechanisms perform better, and where government intervention proves to be more effective"[105]) now also seems to be leading to different – often more intellectually 'modest' – types of approaches towards public policy formulation, implementation and evaluation[106] in different fields. One can only hope that we will not have to wait for similar cataclysms in the international security realm before we start addressing some of the both theoretical and empirical problems that our analysis of the more 'technical' aspects of the literature on deterrence has revealed.

---

[104]For a great example, see George P. Shultz et al., "A World Free of Nuclear Weapons," *Wall Street Journal*, January 4, 2007, sec. Opinion, https://www.wsj.com/articles/SB116787515251566636.

[105]One of the world's best examples of this more pragmatically eclectic, experiential approach to policy is probably the city-state of Singapore, which merges a still extremely 'laissez faire' approach in many economic areas (international trade, economic competition, etc.) with an equally extremist 'dirigiste' approach to many other social and policy issues (educational excellence, urban housing, law and order, (homeland) security, etc.). For an overview of the country's public administration system, see Jon S. T. Quah, *Public Administration Singapore-Style*, Research in Public Policy Analysis and Management 19 (Bingley: Emerald, 2010), 247.

[106]On some of these new approaches in a defense and security context, see Stephan De Spiegeleire, Peter Wijninga, and Tim Sweijs, *Designing Future Stabilization Efforts* (The Hague, The Netherlands: The Hague Centre for Strategic Studies, 2014); Stephan De Spiegeleire et al., "Implementing Defence Policy: A Benchmark-'Lite,'" *Defense & Security Analysis*, February 2019, 1–23, https://doi.org/10/gft4wf.

4. AN 'MRI'-SCAN OF THE FIELD: THE CONTENT

Whereas the previous section focused on the more *technical* aspects of the body of written scholarly work on (different layers of) deterrence, this section presents some of the *substantive* findings of our epistemic 'MRI-scan'. We want to emphasize that this section is best seen (and read) as an early 'beta' version of a more extensive treatment we will produce at a later stage.

'Absorbing' the literature – Main available approaches

There are currently two main ways for humans to 'absorb' the accumulated knowledge that rests encoded in written text corpora. The first way is based on a skill that has been around for (still only!) a few thousands of years and that our educational systems spend enormous efforts instilling and perfecting in the brains of more and more children: to read and try to comprehend these texts[107]. The second way has only been around for a mere few decades, and only a very small fraction of humans have so far (consciously) been exposed to it. It used to be known as 'text-mining' – using computers to 'mine' texts for (also epistemic) nuggets; but it is currently mostly referred to as 'Natural Language Processing' (NLP): using computers to systematically 'process' – if not quite yet 'understand' – much larger amounts of (also scholarly) publications than the human brain would be able to. Given the plethora of structural problems we have laid bare in our 'technical' bibliometric analysis of the deterrence-IS literature in terms of recall, volume, uptake, velocity, etc. – we submit that this second mode of knowledge absorption deserves far more attention than it has received thus far.

Where we come from: Human-only interpretation

Humans, reading and sensemaking – a match not made in heaven?

Humans have only been encoding knowledge in text for a few thousands of years – a mere trifle, as Stanislas Dehaene, a leading French neuroscientist, reminds us[108], in the 8 million years of evolution of hominids or even in the 300,000 year-evolution of homo sapiens. "Evolution thus did not have the time to develop specialized reading circuits in Homo sapiens. Our brain is built on the genetic blueprint that allowed our hunter-gatherer ancestors to survive… Nothing in our evolution could have prepared us to absorb language through vision. Yet brain imaging demonstrates that the adult brain contains fixed circuitry exquisitely attuned to reading."

Current thinking in cognitive neuroscience suggests that this cognitive ability to read and understand, which seems to be unique to humans, essentially repurposed part of the circuitry through

[107] Stanislas Dehaene, *Reading in the Brain: The Science and Evolution of a Human Invention* (New York: Viking, 2009), http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1115790.
[108]Dehaene.

which our brain processes *general visual* cues and stores the thus acquired information into long-term memory in order to enable it to encode and decode *textual* information/knowledge as well. This aptitude does not come naturally or easily to humans. Instead it requires a significant amount of sustained long-term effort from parents, educators, and myriad others; and of course first and foremost from reading children and adults themselves. One of the fascinating findings of cognitive neuroscience is that many of the human natural language processing that take place between our retina and our brain when we read are surprisingly akin to the processes that machine-based natural language processing uses when it acquires and processes text[109]. Another robust finding from this literature is that our cognitive ability to 'read' (ingest), 'comprehend' (digest)[110] and 'share' (egest) knowledge accurately is severely constrained by some of the well-known limitations of the brain circuitry that we inherited from our primate evolution. This is a sobering finding for those of us who want to make strategic decision-making increasingly evidence- and knowledge-based. The silver lining in this dark finding, however, may be that some of the natural language technologies that are mushrooming all around us could start to compensate for some of these evolutionary human limits. If we will prove to be open to that eventuality.

Human reading and sensemaking of the deterrence-IS literature

The task of reading, categorizing and making sense of the substantive content embedded in the relevant literature on deterrence in the 60s and 70s of the previous century – that saw the emergence of 'classical deterrence theory' in IR/IS – was still quite manageable for the very simple reason that the literature was limited. Scholars could spend a few weeks reading and absorbing the few tens of scholarly publications that had been written on deterrence 'in the open[111]' – and the knowledge they contained. Even then, there were clearly already 'recall' issues[112], but by and large the field remained cognitively digestible.

Today, as we have illustrated in the previous section, the production of substantive contributions to the knowledge landscape in all fields – even this one – has exploded in depth *and* in breadth (if not always in quality). This means that the task of categorizing, let alone absorbing it simply

[109]This includes things like using 'windows' of text to focus on; tokenizing it into smaller chunks; normalizing/lemmatizing the main chunks; doing parts-of-speech and dependency analysis; linking up the semantic elements in the text with existing knowledge; etc. etc. See also the more recent Stanislas Dehaene, *How We Learn: Why Brains Learn Better Than Any Machine ... for Now.*, 2020, https://www.overdrive.com/search?q=92024A08-400B-4F71-9982-5564CC370435.

[110]On this, one of the most interesting (and devastating) overviews of human cognitive frailties can be found in the cognitive bias codex on Wikipedia, "List of Cognitive Biases," in *Wikipedia*, April 24, 2021, https://en.wikipedia.org/w/index.php?title=List_of_cognitive_biases&oldid=1019556972. See also Dehaene, *How We Learn*; Dan Ariely, *Predictably Irrational: The Hidden Forces That Shape Our Decisions*, Rev. and expanded ed., 3. [print] (New York, NY: Harper Collins Publ, 2009); Daniel Kahneman, *Thinking, Fast and Slow*, 1st ed (New York: Farrar, Straus and Giroux, 2011); Jonathan Baron, *Thinking and Deciding*, 4th ed (New York: Cambridge University Press, 2008).

[111]And the same was almost certainly also true for 'cleared' analysts in the classified area.

[112]The difference in the uptake of more 'technical' game-theoretical analyses like Daniel Ellsberg's work in 1958-1959 (but only published in 1968) vs more discursive analyses like Tom Schelling's is quite revealing here. Daniel Ellsberg, "The Theory and Practice of Blackmail," RAND Paper (Santa Monica, CA: RAND Corporation, July 1968); Thomas C. Schelling, *The Strategy of Conflict* (Harvard University Press, 1960); Thomas C Schelling, *Arms and Influence* (New Haven and London: Yale University Press, 1966).

exceeds humans' innate cognitive aptitude. Most scholars therefore resort to various cognitive 'shortcuts' like piggybacking on existing 'literature reviews' or schemas for classifying/synthesizing the extant literature. To give an example: one of the currently increasingly popular ways of structuring the deterrence-IS literature is to talk about 5 consecutive 'waves' in deterrence theory[113]. This particular schema essentially conflates[114] two taxonomic principles – a temporal one and a substantive one – to introduce readers to this body of literature. While undoubtedly useful, any scholar who has read even part of the Deterrence-IS literature realizes that such a two-dimensional representation is heroically reductionist.

There are, indeed, myriad other ways in which this body of literature can be sliced and diced. Possible other taxonomic dimensions include: a publication's prime purpose (e.g. theoretical or applied; positive, normative or interpretive; etc.); methodological approach (e.g. descriptive/historical, game theoretic, agent-based-modeling, etc.); theoretical 'school of thought' (e.g. constructivist, liberal or realist); level of analysis (e.g. the individual, the group, the state or the 'system'); 'quadic' focus (monadic, dyadic, triadic, etc.); geographical focus (e.g. US, Russian, Chinese deterrence vs Russia vs the West); temporal focus (the past, the present, or future), its author's overall attitude towards deterrence (e.g. positive, negative, or neutral), the definition of deterrence she operates under (see the section on manual coding in this paper) etc. Any publication could be positioned in a particular location in that multi-dimensional space[115]. The ability to position a piece of knowledge within that space – and only the ability to do so – would allow scholars/analysts to put claims made by different authors across this space more in their proper perspective. Unfortunately, humans experience extreme difficulties thinking in higher-dimensional spaces, however, and we therefore typically end up thinking (and writing and visualizing) in lower-dimensional ones. The question as to which price we pay for this cognitive dimensionality reduction in terms of real epistemic understanding remains an open one for now.

Where we are heading: Machines to the rescue

From analogue to digital

Painfully aware of the limitations inherent in the more traditional, 'analogue'[116], low-dimensional

---

[113]Frans Osinga and Tim Sweijs, *NL ARMS Netherlands Annual Review of Military Studies 2020 : Deterrence in the 21st Century—Insights from Theory and Practice* (Springer Nature, 2021), https://library.oapen.org/handle/20.500.12657/47298; Tim Sweijs and Frans Osinga, "Conclusion: Insights from Theory and Practice," in *NL ARMS Netherlands Annual Review of Military Studies 2020: Deterrence in the 21st Century—Insights from Theory and Practice*, ed. Frans Osinga and Tim Sweijs, NL ARMS (The Hague: T.M.C. Asser Press, 2021), 503–30, https://doi.org/10.1007/978-94-6265-419-8_26.

[114]In this sense, such mono- (or oligo-) dimensional literature classification schemes may very well be based on some 'dimensionality reduction' that takes place in the human brain. That is to say: any author most likely has (consciously or subconsciously) some 'intuitive' – and essentially untraceable – mental awareness of these multiple dimensions, but she ultimately typically ends up with a few 'schools of thought' along one or two dimension in her literature review.

[115]An exercise that is always fascinating – since it can also reveal possible epistemic 'holes'.

[116]The authors are grateful to CNA's Michael Kofman for this 'analogue' vs 'digital' analogy.

way of approaching literature reviews or reviewing/structuring the literature in a field of research[117], a number of scientific disciplines – with the life sciences, as usual, leading the way – have been moving towards more 'digital' and data-savvy ways of generating 'big picture' overviews of their field of inquiry. This trend is known under different names depending on one's discipline or research angle. We have already mentioned the popularity – especially also in some policy-analysis communities – of 'meta-analyses'[118] in various more empirically-inclined fields. Another ascendant term in this context is 'literature-based discovery' (LBD) which "generates discoveries, or hypotheses, by combining what is already known in the literature[119]". Over the past two decades, LBD has become a widely used tool in – again – the biomedical field, but it is clearly starting to spread beyond that. A third related term *en vogue* is that of a 'systematic review' – "a particular type of literature review that is characterized by a methodical, replicable, and transparent approach"[120]. Many of these terms can also be subsumed under the broader (also methodological) rejuvenation of 'meta-science' – the scientific study of science itself[121].

One of the fascinating shared aspects of the scientific efforts behind all of these terms is that they do not merely 'digitize' the old 'analogue' way of doing literature reviews, but also open up entirely new avenues of knowledge discovery and validation. This section of our paper therefore has a double aim. One important aim is to showcase some of the new research techniques that have emerged to do this more semantic (meaning-related) 'big picture' analysis. The second goal is to illustrate what some of these various methodological developments can produce by presenting a few different ('beta') mappings of 'what we know' about deterrence. A key role in all of is played by an exciting new set of tools that have emerged at the intersection of Artificial Intelligence and Linguistics.

Natural language processing

Natural language processing is a set of machine-learning-based algorithms that allow computers to process texts written by humans. It ranks among the cutting-edge technologies that have been making vertiginous strides in just the past few years[122]. In contrast to previous eras in artificial

[117]On the difference between these two, see Andy P. Siddaway, Alex M. Wood, and Larry V. Hedges, "How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses," *Annual Review of Psychology* 70, no. 1 (January 4, 2019): 747–70, https://doi.org/10.1146/annurev-psych-010418-102803.
[118]For more background, recommendations and a number of caveats about 'meta-analyses' in different disciplines, see Borenstein, *Introduction to Meta-Analysis*; Bafeta et al., "Analysis of the Systematic Reviews Process in Reports of Network Meta-Analyses"; Wallace, Nwosu, and Clarke, "Barriers to the Uptake of Evidence from Systematic Reviews and Meta-Analyses"; Kugley et al., "Searching for Studies"; Petrokofsky, "Guidelines and Standards for Evidence Synthesis in Environmental Management"; Christensen, Freese, and Miguel, *Transparent and Reproducible Social Science Research*; Higgins and Cochrane Collaboration, *Cochrane Handbook for Systematic Reviews of Interventions*.
[119]Dimitar Hristovski et al., "Constructing a Graph Database for Semantic Literature-Based Discovery," *Studies in Health Technology and Informatics* 216 (2015): 1094.
[120]For an excellent introduction and guidelines, see Siddaway, Wood, and Hedges, "How to Do a Systematic Review," 751.
[121]David Faust and Paul E. Meehl, "Using Meta-Scientific Studies to Clarify or Resolve Questions in the Philosophy and History of Science," *Philosophy of Science* 69, no. S3 (September 2002): S185–96, https://doi.org/10.1086/341845.
[122]For a good overview of recent developments in this area, see Ming Zhou et al., "Progress in Neural NLP: Modeling, Learning, and Reasoning," *Engineering* 6, no. 3 (March 1, 2020): 275–90, https://doi.org/10.1016/j.eng.2019.12.014.

intelligence (and language analysis) that were driven by the public sector (and often by defense organizations[123]), this time around private-sector behemoths like Amazon, Baidu, Facebook, Google, Microsoft, Tencent, etc. have made massive investments that are leading to breakthroughs on an almost weekly basis[124].

One of the biggest breakthroughs in applied machine learning took place in the teens of this century in the subfield of visual image recognition with the appearance in 2009 of *ImageNet*. This was a dataset containing more than 14 million images that had been hand-annotated by humans to indicate what objects were pictured, often with bounding boxes. ImageNet contained about 100 images for each of some 20,000 synsets – synonymous words or word phrases that identify a concept (e.g. a type of bird, a cat, a piece of furniture, etc.). By being able to use these (presumably) reliably human-annotated images, machine learning algorithms could then apply various neural network algorithms to that dataset to 'learn' from the tagged images. These validated pre-trained patterns could then be transferred to new images to be finetuned.

In the text realm, natural language processing had already been making steady progress since the development around 2001 of neural language models that could be used for tasks like spelling auto-correction or virtual keyboards like Swipe. Another major algorithmic step forward came with the appearance around 2013 of word embeddings – vector representations of words in corpora that convert text to numbers that can processed by existing neural network algorithms to solve NLP tasks like document clustering (which documents are related), topic modeling (what are the key topics to emerge out of a corpus), etc.

The big 'ImageNet-moment' for text, however, arrived only around 2018. April of that year saw the publication of a paper co-authored by collaborators from the University of Washington and DeepMind, the Alphabet-owned artificial intelligence company. It introduced a benchmark called GLUE (General Language Understanding Evaluation) consisting of nine reading/comprehension tasks for computers built on established existing datasets and selected to cover a diverse range of dataset sizes, text genres, and degrees of difficulty. Examples of these tasks include question answering, sentiment analysis, text similarity, etc.

Initially, even state-of-the-art neural networks scored quite poorly: none of them scored higher than 69 out of 100 on the GLUE benchmarks across all nine tasks. But Google was already working on a set of pre-trained language models that were trained on massive text corpora and that outperformed the then state-of-the-art models on most NLP tasks by large margins (Google's BERT produced an overall GLUE score of 80.5). Google open-sourced a number of BERT-language models which were able to analyze words (or sentences or paragraphs) in their

---

[123]G.M. White, "Speech Recognition: A Tutorial Overview," *Computer* 9, no. 5 (May 1976): 40–53, https://doi.org/10.1109/C-M.1976.218586.

[124] Stephan De Spiegeleire, Matthijs Maas, and Tim Sweijs, *Artificial Intelligence and the Future of Defense: Strategic Implications for a Small Force Provider* (The Hague, The Netherlands: HCSS, 2017), http://hcss.nl/report/artificial-intelligence-and-future-defense.

context, allowing them to disambiguate – just like humans do – the usage of the word 'bank' in a financial context or in a 'river' context based on a an analysis of a context window of a certain amount of words to its left or right.

The largest language model of BERT was trained on text corpora (Google's Book corpus and Wikipedia) containing 2500 million words with a model that contained 340 million parameters and could analyze inputs that were 512 tokens (in this case wordpieces) wide. OpenAI's GPT-3, which came out in May 2020, was trained on the Common Crawl dataset containing nearly a trillion words in multiple languages with a model containing 175 billion (!) parameters and based on inputs that could be up to 2048 tokens wide. And then in July 2020, Google released BigBird, a once again disruptively superior new set of models that can process input sequences of 4096 tokens and is allegedly even capable of going up to 16k+ tokens. This allows these models to get much more insight in the broader context of concepts, entities, relations, etc.

BERT represented such a dramatic improvement on an established (if still decidedly imperfect) benchmark, that the team behind GLUE developed a human benchmark to compare machine vs human performance[125]. Already in June of 2020, a BERT-based model from Microsoft (with an all-Chinese team of authors) surpassed human performance with an overall average GLUE score of 87.6 vs. 87.1[126]. At that point, however, there were still many tests on which the AI-models scored significantly worse than humans. As of the time of this writing (March 27, 2021) the top model on the (improved) 'SuperGLUE' leaderboard – Baidu's Ernie – scores 90.9 overall, and also shows superhuman performance on 8 of the (now) 11 tasks, with quite small (and declining) differences on the final 3[127] and growing margins on most of the others.

These achievements, as astounding as they are, should still be put in their proper perspective. It is indisputably not (yet) the case that computer algorithms are smarter, more 'intelligent', etc. than humans across the board. The day that they will surpass humans across all GLUE tasks seems near. But these GLUE (even SuperGLUE) tasks remain a far cry from the level of (especially semantic) 'understanding' of textually encoded knowledge that humans have evolutionarily developed and that are required for the type of sense-making that would be trustworthy for decision-making support. We would still be well advised, also as researchers working in the field of international security, to bear in mind, however, that algorithmic evolution moves orders of magnitude faster than its human counterpart.

---

[125]Nikita Nangia and Samuel R. Bowman, "Human vs. Muppet: A Conservative Estimate of Human Performance on the GLUE Benchmark," *ArXiv:1905.10425 [Cs]*, June 1, 2019, http://arxiv.org/abs/1905.10425.

[126]kexugit, "Microsoft's MT-DNN Achieves Human Performance Estimate on General Language Understanding Evaluation (GLUE) Benchmark," June 20, 2019, https://docs.microsoft.com/en-us/archive/blogs/stevengu/microsoft-achieves-human-performance-estimate-on-glue-benchmark.

[127]Alex Wang et al., "GLUE Benchmark Leaderboard," GLUE Benchmark Leaderboard, 2021, https://gluebenchmark.com/.

Where we stand: This report's approach

This section of our report essentially sets out to showcase a few illustrations of where the field of NLP currently stands and what that could mean for the study of key issues in international security like deterrence. Our own overall assessment is that the current state of NLP is significantly more impressive (and useful!) than the majority of scholars working in our field currently realize; that it is at the same time still far more modest than what the current-day hype suggests[128], but that – most importantly – the gap between these two is closing precipitously quickly. This is why we decided to put some NLP-tools 'through the wringer' to find out which of them may hold more promise. This section of this report is therefore not intended to provide the 'final' word on the topics it addresses. Instead it can best be read as a few early illustrations based on these new NLP-tools and embedded in an invitation to our colleagues to join us in a – hopefully more collaborative – effort to explore the promise *and* the peril in this (exceptionally quickly moving) field.

At this juncture in time, we have access to two main pathways through which machines can help us to explore the knowledge landscape encoded within or across larger sets of scholarly publications. The first one goes once again – as in the previous section of this report – through only the *metadata* that are available in bibliographical/bibliometric databases, but it focuses on the substantive ("what is being argued here?") instead of the more technical/sociological metadata fields (who, when, where, with whom, etc.). The second pathway primarily analyzes the *full texts* of all relevant publications that are available in academic databases, while still taking advantage of their metadata as well to map trends over time, etc. Both approaches have advantages and disadvantages.

Parameter	**Metadata-only approach**	**Full-text approach**
Text data available for analysis	Title, abstract, keywords (TAK)	TAK + full text
Focus	Sharp – succinct summary of the main points made	Broader – more explanation, argumentation, data, digressions
Richness	Relatively poor	Much higher
Ease of use	Relatively easy (established software with graphical user interfaces)	Harder (mostly – not yet? – integrated or with GUIs; require some knowledge of Python or R)
Cost	Free	Mostly free
Ability to map key topics	Yes	Yes

[128]In this sense we tend to concur with Gartner, the global research and advisory company, who have positioned NLP as one of their 'At the Peak (of inflated expectations)' technologies, meaning that it will still have to go through what they expressively label the 'trough of disillusionment' and the 'slope of enlightenment' before reaching the 'plateau of productivity'. Gartner Research, "Hype Cycle for Natural Language Technologies, 2020," Gartner, July 6, 2020, https://www.gartner.com/en/documents/3987154/hype-cycle-for-natural-language-technologies-2020. But we hasten to add that homo sapiens' self-assessment of her own cognitive prowess, since at least the day when Swedish botanist and taxonomist Carl Linnaeus hybristically decided to name our species 'sapiens' in his broader zoological taxonomy of species, shows signs of a more structural and arguably worrisome – also from a policy decision-support point of view – case of 'inflated expectations', as the current comparison between human and artificial intelligence NLP benchmarks quite clearly shows.

Ability to identify key concepts	Yes	Yes
Ability to track changes over time	Yes, easy to implement	Yes, but requires additional efforts
Ability for 'human labelling'	Not reasonable due to the scarcity of data	Yes
Ability for 'active learning'	Not reasonable due to the scarcity of data	Yes

Table 4: Comparing bibliometric and full-text datasets/corpora

As we already saw, scholarly publications' metadata also contain some more 'substantive' fields that do not just contain the more 'technical' details of a publication like its author, publication source (e.g. journal), date, number of pages etc., but also some more substantively 'meaty' fields like the abstract, the keywords, etc. As abstracts are supposed to be a more succinct restating by the author(s) of the central epistemic contributions of publications, they allow scholars/analysts using NLP to just analyze their individual *and* collective quintessence. Applying bibliometric software tools, which are starting to include more and more NLP-elements in their functionality, to these substantive metadata fields can therefore already provide a uniquely useful first-cut analysis. The full texts of these publications, on the other hand, enable researchers to pick up on a number of more in-depth, more thoroughly explained, argumented and contextualized, possibly more tangential but still important (maybe also in different contexts) epistemic claims that would get lost if they were only to focus on titles, abstracts and/or keywords. Full text analysis – sometimes referred to as 'corpus analytics' – can unleash the full array of NLP tools upon this much richer source of textual data.

There are different ways of categorizing various types of NLP analysis, one of which is illustrated in the following diagram. The categorization is based on the degree of human intervention, and all three categories will be explored in our analysis.

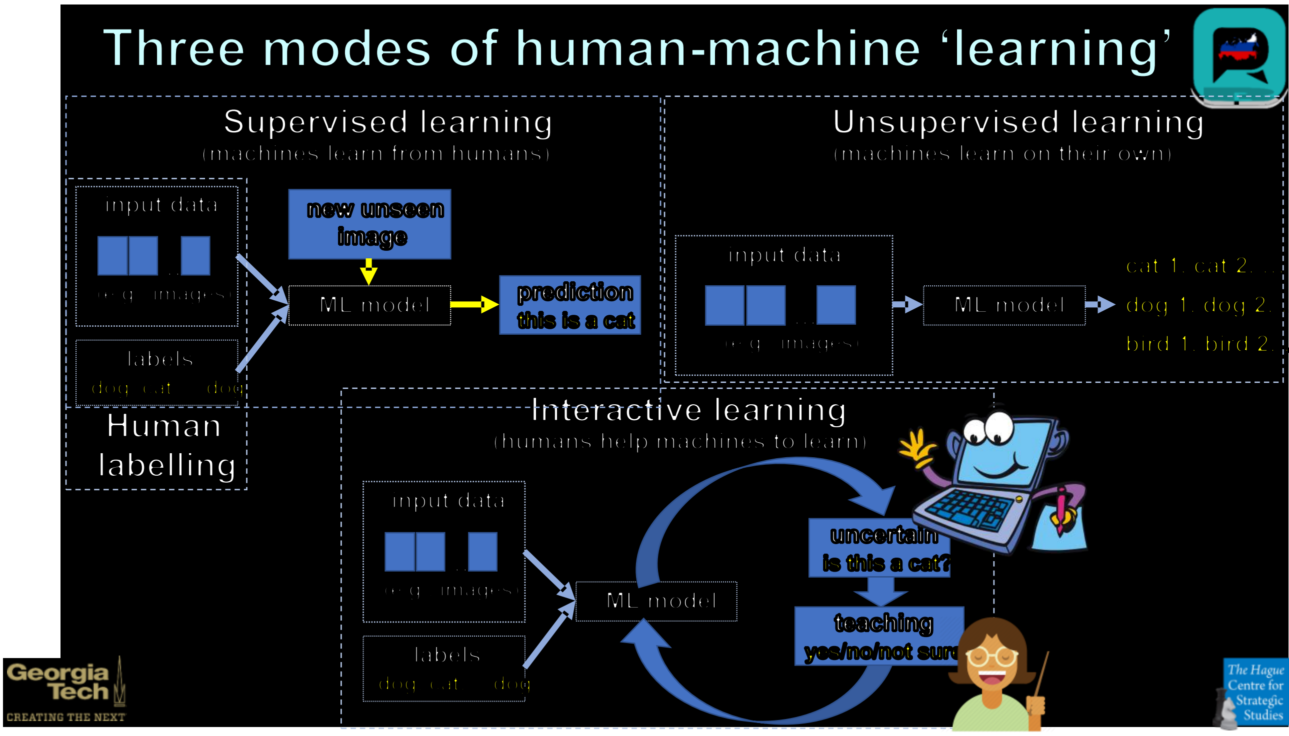


Figure 30: Three modes of human-machine co-'learning', one of human-only labelling

All categories are anchored in full-text data, which, in the absence of rich empirical or synthetic dataset (as we argued in the first section), remain the mainstay of what we (think we) know about deterrence. Since this slide only shows computer-assisted modes of textual analysis, it omits the currently most dominant ('analogue') mode in which scholars just ingest (read: read), digest (read: think through), and egest (read: publish) knowledge.

We coined the first mode of textual analysis that is both human and machine-assisted as '*human labelling*'. This is the mode in which humans use their trained biological intelligence to systematically and transparently annotate textually-encoded (presumed) knowledge in software programs without any 'AI-smarts'. It is important to realize that every researcher actually 'codes'. When we read texts, we select context windows that we 'make a mental note of' in our brain, that we connect to some element(s) in our coding scheme(s), that may lead us to update that coding scheme or the knowledge it contains, etc. All of these countless coding choices, however, that we made on every page of every publication we read are made with us realizing it, are not documented, are essentially impossible to trace, etc. The main difference with human labelling is that in the latter at least some of these choices are made explicit, are annotated, stored and (in the best case) documented. They can therefore be contested, updated, improved etc. based on the normal scientific canon. Without any such careful human annotation an author may claim, for instance, that the Russian literature on deterrence pays more or less, or different attention to a particular form of deterrence (e.g. deterrence by denial). Readers would have to accept that claim at face value – based on the author's authority or persuasiveness. With systematic coding of all text spans in a well-specified corpus that deal with definitions of deterrence, readers *are* able to verify whether the corpus really does use all relevant sources, whether all relevant text spans were found and used, how exactly these specific text spans were coded, etc.

Any transparently traceable, systematic human annotation effort of the available literature on this topic – maybe even with some level of collaborative mapping of the different readings of various excerpts – therefore already represents a significant step forward for scholars and decision-makers/supporters alike, even without any recourse to any artificial intelligence. In the deterrence-IS field, we have argued that the CNA-team that coded a few 100s of Russian military publications on escalation management based on a transparent coding scheme has brought this field much closer to this ambition level[129]. Our own efforts in this report to collect a much larger corpus of full-text publications on deterrence (in English and in Russian), to select a subcorpus that deals specifically with definitions and typologies of deterrence; and to then systematically and traceably code relevant text-spans in a software program based on a rich coding scheme is another example of his 'human labeling' approach.

In the (impressive) CNA report, this human labeling approach represented the final stage of their analysis. In our own approach, we see human labelling as the first step in an incremental research protocol that increasingly draws in machine intelligence to enrich human intelligence. In the slide, we position *'supervised' (human-machine) learning* as the attempt to have machines learn from humans. Regrettably (some may even say inexcusably), we must acknowledge as a community[130], that we do not, as yet, have a sufficiently reliably annotated 'gold corpus' on Deterrence-IS (let alone Deterrence-IS-Russia) to build an ML model that would allow us (and others) to embark on this 'supervised learning' track. We want to stress that we very much deplore that this is the only research avenue that we are currently not in a position to pursue, and that we consider this to be one of the extremely promising options to produce more trustworthy knowledge on deterrence.

The other two 'modes' of 'human-machine' analytical modes of corpus-analysis – 'unsupervised' AND (what we now call) 'interactive' learning – are ones we fortunately have been able to pursue in this project. *Unsupervised machine learning* tools – also in NLP – just take data (in the case of NLP – large (full-text) corpora) to algorithmically identify patterns, trends, topics, bursts, etc. in the literature. We will present different examples of topic-modeling and one other way to visually explore a corpus that might help scholars and analysts in their sense-making efforts.

The final – and in our opinion (for now by far) most promising – research avenue for exploring text corpora more systematically is one that we label *'(inter)active learning'*. 'Active learning' is the

[129]Michael Kofman, Anya Fink, and Jeffrey Edmonds, "Russian Strategy for Escalation Management: Evolution of Key Concepts" (Arlington, VA: Center for Naval Analysis, April 13, 2020), https://www.cna.org/CNA_files/PDF/DRM-2019-U-022455-1Rev.pdf.

[130]And we, frankly, doubt that the Western decision-making (or funding) community is fully aware of this.

more widely used term for this[131], but in the context of categorizing different machine-learning algorithms, we thought this term more felicitously captured what this approach is all about: to let humans interactively help machines to 'learn' from text corpora. This research avenue, very much like the (aspirational) supervised learning approach, starts from a smaller number (that is significantly smaller than the datasets that would traditionally be required for supervised learning approaches – we started with a few 100s) of sentences that have already been labeled by subject-matter experts as belonging to one or more categories. This labeled dataset is then used by a software program (in our case Prodigy) that can leverage various pre-trained NLP language models to identify other 'relevant' sentences that seem similar to the already (humanly) labeled ones. The way it does this is by training a model on the initial human-annotated dataset, and to then 'serve' human experts classification cases that the initial model feels most uncertain about. The human can then either accept the model's classification suggestion, decline it, or admit that she also does not know. Based on that human feedback, the algorithm can then keep tweaking the ML model until it achieves optimal performance.

Summing up, this section will first look at what we can still squeeze out of our bibliometric datasets from a substantive point of view ("what is this literature really about?"). It will then turn its attention to what unsupervised, interactive-learning, *and* 'manual' NLP approaches can tell us about this substance.

Using metadata (bibliometric-substantive)

To map the epistemic landscape in the Deterrence-IS field bibliometrically, we proceed in three steps. First, we survey the field substantively *in toto*. Secondly, we use the Deterrence-IS dataset as a benchmark to map English publications on specifically Russian deterrence (Deterrence-IS-Russia/English) and to compare their content to the overall field of Deterrence-IS. Our third step is to contour Russian studies on deterrence (Deterrence-IS-Russia/Russian) and to compare them to the broader international datasets (Deterrence-IS, Deterrence-IS-Russia/English), paying special attention to the similarities and the differences between the international and the Russian approaches to deterrence-IS.

For the first two steps, we decided to primarily use the Deterrence-IS (Scopus) dataset, as it has a larger number of documents than Deterrence-IS (WoS) – 810 vs 695 docs; as well as better coverage in terms of abstracts and keywords than Deterrence-IS (Lens) – 63.7% of papers have keywords vs 0.02% for Lens; for abstracts the numbers are 84% vs 70%. For the third step, we used the Deterrence-IS-Russia/Russian (WoS:RSCI) dataset collection as it is the most complete Russian dataset available to us. Finally we compared Deterrence-IS-Russia/English (Scopus) and

[131]Christopher Schröder and Andreas Niekler, "A Survey of Active Learning for Text Classification Using Deep Neural Networks," *ArXiv:2008.07267 [Cs]*, August 17, 2020, http://arxiv.org/abs/2008.07267.

Deterrence-IS-Russia/Russian (WoS:RSCI) in the last step in our substantive bibliometric analysis.

In terms of software, we opted for the powerful and (in our assessment) more substance-centric[132] free software program CiteSpace to extract noun-phrases from the more substantive metadata (Title/Abstract/Keyword) and to then build and visualize networks with these noun-phrases as the main nodes to discover and compare various substantively interesting aspects of these literatures. We will provide additional details about CiteSpace's algorithms as we build up our analysis throughout this section and also in Annex B – List of configurations for CiteSpace.

Deterrence-IS

One of the useful functions of CiteSpace is the automatic extraction of noun-phrases (hereafter: 'terms') from the datasets' titles, abstracts, and keywords. To find these, CiteSpace first uses natural language processing techniques to tag every word in a dataset as a specific 'part of speech' (a noun, adjective, article, etc.). It then goes out to find 'noun phrases' – i.e. combinations of more than one word that contain a noun (e.g. deterrence, weapons, etc.) alongside some other parts of speech (e.g. an adjective – nuclear, extended, etc.). Finally, it computes a numeric value to each of these terms that indicates how relatively important that term is across all publications. For this sake, CiteSpace applies the *term frequency-inverse document frequency algorithm* (tf*idf). As the name suggests, this algorithm multiplies term frequency (tf) by inverse document frequency (idf).

Term frequency refers to the raw count of occurrences of a noun phrase in a dataset. Inverse document frequency measures how common or rare a term is for a given corpus/dataset. It is counted by dividing the total number of documents by the number of documents that contain a term. If the term is very common, the idf value approaches 0; if it is unique only for specific documents, its value approaches 1. By multiplying tf by idf, the algorithm, firstly, allows excluding general words, like "the", "be", "to" or "a", which are common for all documents. It, secondly, also puts most the semantically relevant terms that reflect the content of a given dataset at the top of the list. The higher the tf*idf index, the more semantically important a certain term is assumed to be.

Table 5 lists the top 10 terms in our Deterrence-IS (Scopus) dataset as extracted, computed and ranked by CiteSpace. One important term that is not mentioned in the table is *nuclear deterrence*. It has a higher tf (449) than the sum of the top five terms in the table. However, it is so common in deterrence literature that its idf equals 0 (the same value as for "the", "be", or "to").

[132]In our own experiments with various bibliometric software programs (Bibliometrix, VantagePoint, Vosviewer, etc.) we have found CiteSpace providing analysts with far more functionality, more latitude to specify and experiment with various parameters, and to be overall more useful for what we call 'substantive' bibliometric analysis. Unfortunately, the program is not (yet) documented as well as one would hope, but we are willing to share our own (quite extensive) documentation with fellow scholars who might be interested in trying it out.

tf	idf	tf*idf	term
215	1.1	236.2	nuclear weapons
90	2.08	187.15	cold war
41	2.94	120.72	conventional deterrence
39	3	116.83	south asia
38	3.04	115.69	soviet union
37	3.04	112.65	nuclear war
37	3.04	112.65	strategic stability
36	3.09	111.28	cyberdeterrence
36	3.09	111.28	international relation
36	3.09	111.28	nuclear disarmament

Table 5: Top-10 most salient terms in deterrence-IS

An overview of key noun-phrases in a bibliometric dataset provides us with a general 'first-cut' appreciation of its content that is already slightly more intelligent than just 'beancounting' words or terms. Mapping the connections between these noun-graphs in publications' titles/abstracts/keywords as networks, however, allows us to gain additional insights into which of these noun-phrases tend to 'travel together' throughout the dataset and thus constituting its main substantive topics, as well as into those topics' relative importance and centrality. Any network consists of 'nodes' and of connections between nodes – called 'edges' in graph theory, the discipline that studies networks[133]. Bibliometric software allows analysts to build up networks based on different types of nodes[134]:

- ***References*** – i.e. the footnotes or bibliographical references cited in and/or at the end of a publication, which are connected in the network whenever one or more other publications cite them together (these are called co-citation networks);
- ***Terms*** – i.e. noun-phrases that co-occur within titles, abstracts, or keywords (called term co-occurrence networks);
- ***Countries*** – i.e. the countries mentioned in co-authors' affiliations: if a publication has one author from the US, and one from Russia, the network will mark that as one connection between these two countries (country-networks); and
- other nodes like institutions, cited authors, cited journals, etc.

Bibliometric networks are typically weighted networks with the edges connecting nodes not only indicating whether there is a relation between these nodes or not but also the strength of this relation. Using various measures of network centrality, the mathematics behind graph theory allows us to calculate the importance of any given node in a network and to cluster the nodes

[133]For a great introduction to this field, see Albert-László Barabási and Márton Pósfai, *Network Science* (Cambridge, United Kingdom: Cambridge University Press, 2016); Albert-László Barabási, *Linked: The New Science of Networks*, 2002, http://site.ebrary.com/id/10460889.
[134]For an excellent succinct overview of this, see Nees Jan van Eck and Ludo Waltman, "Visualizing Bibliometric Networks," in *Measuring Scholarly Impact*, ed. Ying Ding, Ronald Rousseau, and Dietmar Wolfram (Cham: Springer International Publishing, 2014), 285–320, https://doi.org/10.1007/978-3-319-10377-8_13.

that have more connections between themselves than with others in meaningful ways.

Specifically in this case, we used the list of terms (noun-phrases) generated by CiteSpace (not only the top 10 we just listed, but the entire list of 1253 unique terms extracted from the Deterrence-IS (Scopus) dataset) to build the network that is visualized in Figure 31.

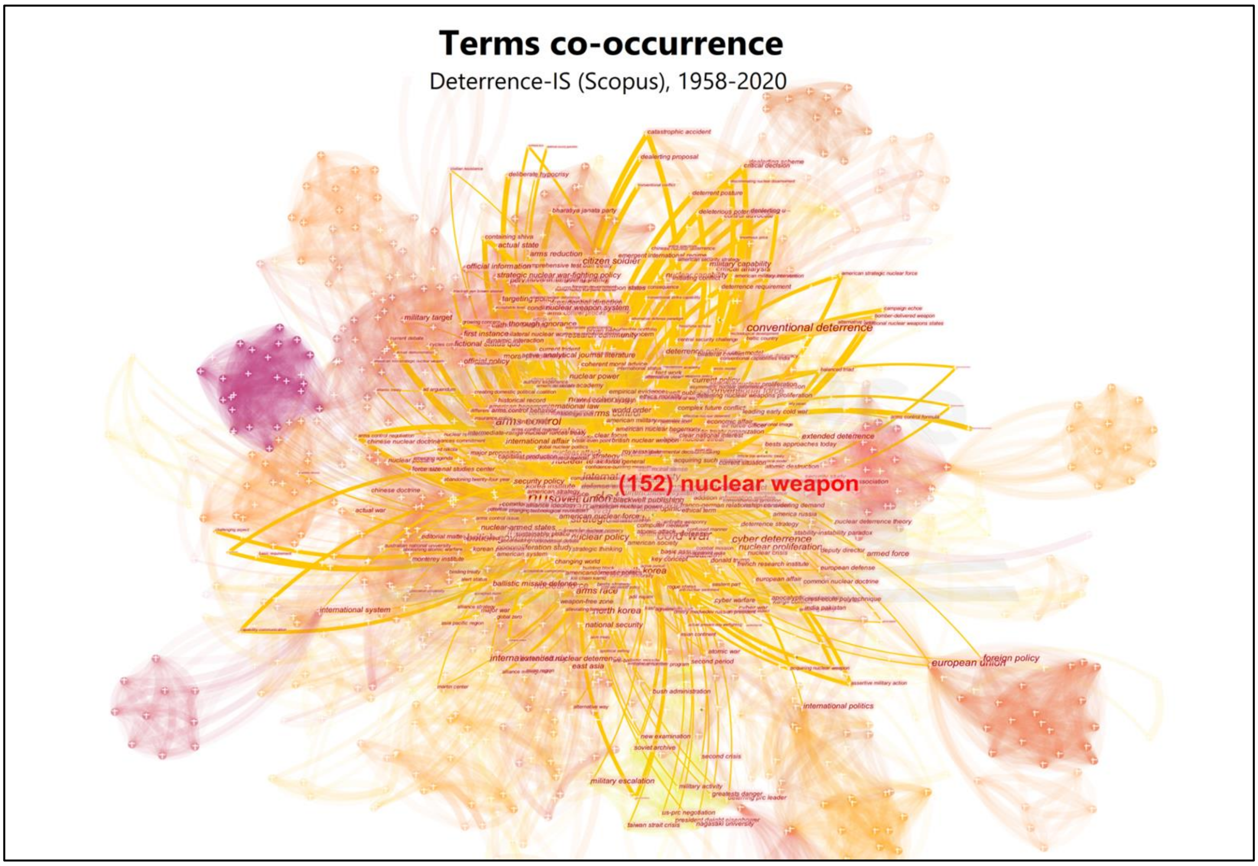


Figure 31: Term co-occurrence network - deterrence-IS[135]

CiteSpace supports different clustering algorithms that allow to link up the nodes with the strongest connections together into clusters. Since terms that most commonly co-occur in various publications' titles, abstracts, or keywords are likely to pertain to a similar substantive topic, these clusters are one way of visualizing the most important research topics within any given dataset. We want to point out that these graph visualizations are – by nature – dynamic and interactive, and that they allow analysts to single out and/or zoom in on clusters that are of interest to them. In this report, we can only reproduce a very small subset of all visuals we produced and examined and we can only do so in a static way, but we invite our readers to experiment and explore our (or their own) datasets for themselves[136].

---

[135]For the sake of clarity, the image shows only some of the terms. Each of the nodes on the image, however, represents a term extracted from the Deterrence-IS (Scopus) dataset by Citespace.
[136] "HCSS-StratBase/HCSS-Deterrence-Datasets," GitHub, accessed April 28, 2021, https://github.com/HCSS-StratBase/HCSS-Deterrence-datasets.

Having generated algorithmically clustered terms, the next task is to find out what these clusters are actually about substantively. Labeling automatically discovered clusters algorithmically in ways that make sense to humans remains one of the major challenges in any form of unsupervised clustering or topic modeling.[137] CiteSpace allows the analyst to label these clusters – again algorithmically – based on terms that the clusters' citing publications share in their titles, abstracts, or keywords. CiteSpace labels use a user's input parameters to discover and assign a 'real' label to a cluster – e.g. the most occurring terms across all abstracts. In some cases, however, these labels can still end up being meaningless or even misleading for human perception. To address this problem, our team's analysts investigated each cluster more deeply with the help of CiteSpace's *cluster explorer* (for an example, see Figure 32). *Cluster explorer* gives users more detailed information about each of the clusters: the terms it contains, how 'old' the publications in which they co-occur are, the terms' frequency and centrality, etc. Researchers can then use all of this more detailed information to come up with and assign that cluster with the label they deem to be the most meaningful one.

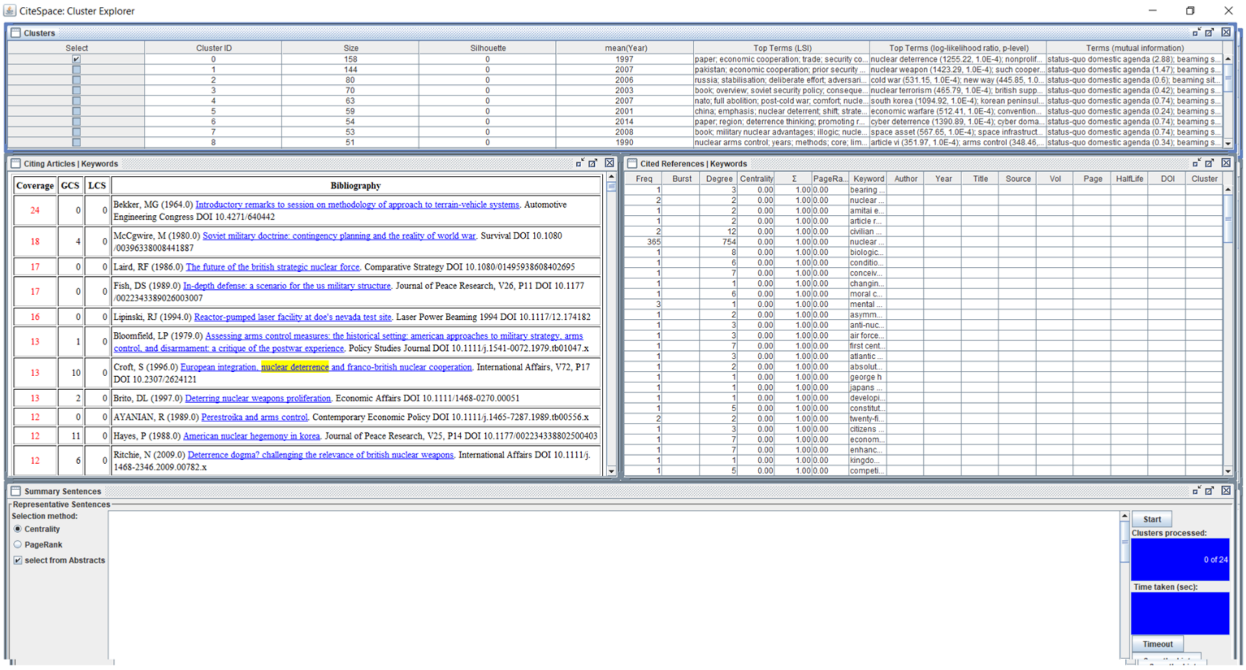


Figure 32: CiteSpace's cluster explorer

The following visual shows the same network of terms co-occurrence that emerges out of our Deterrence-IS (Scopus) bibliometric dataset, but already properly clustered and labeled. The visualization shows a number of key aspects of the dataset:

---

[137]Ngurah Agus Sanjaya et al., "Harnessing Truth Discovery Algorithms On The Topic Labelling Problem," in *Proceedings of the 20th International Conference on Information Integration and Web-Based Applications & Services*, IiWAS2018 (New York, NY, USA: Association for Computing Machinery, 2018), 8–14, https://doi.org/10.1145/3282373.3282390.

- The ***clusters*** themselves. We have already described how these are generated based on co-occurrence, but they are always displayed in the visuals as dots ('nodes') and lines between them ('edges'). The clusters are labeled here based on a log-likelihood algorithm[138] and are numbered in declining importance with topic #0 (on the center-left, labeled 'nuclear deterrence) being the most prominent one, and the 'minimum deterrence' one (#20 on the bottom left) as the least prominent one of the clusters that are being displayed here.
- ***Substantive proximity***. The visualization algorithm also tries to position clusters in the visualization based on the nodes – in this case, co-cited references – that they have in common. So, for instance, clusters 0 ('nuclear deterrence'), 2 ('cold war'), 3 (nuclear terrorism) apparently share a number of important terms, are therefore thought to be dealing with similar substantive topics and are visualized close to each other.
- ***Age***. The colors in this visual also visualize how 'old' the citing articles that share references with one another are on average, with purple ones being old clusters and yellow ones being more recent ones (as per the color-coding bar on the right). So the 'mean' year of topic #12 ('European Union) on the bottom right is relatively old, whereas topic #6 (cyber deterrence) in the middle) is more recent.
- ***Centrality***. Clusters are linked to each other through nodes. But some clusters have more links than others; these clusters have higher *centrality value*. CiteSpace places the nodes with the higher centrality value closer to the center, and those with the lower centrality closer to the edges of the network. The centrality value matters because it is an indicator of the importance of the cluster within the broader field. For example, #1 ('nuclear weapon') is the most central cluster in the deterrence field. Few other clusters avoid direct or indirect engagement with this theme. On the contrary, #12 ('European Union') is one of the least central clusters. It is rarely linked to discussions in other clusters.

[138]The log-likelihood algorithm identifies a unique combination of words (in this case words from abstracts, keywords, and titles) that are more likely to be used together than not. E.g. "witch hunt" is more probable than the use of "witch" or "hunt" independently or with other words.

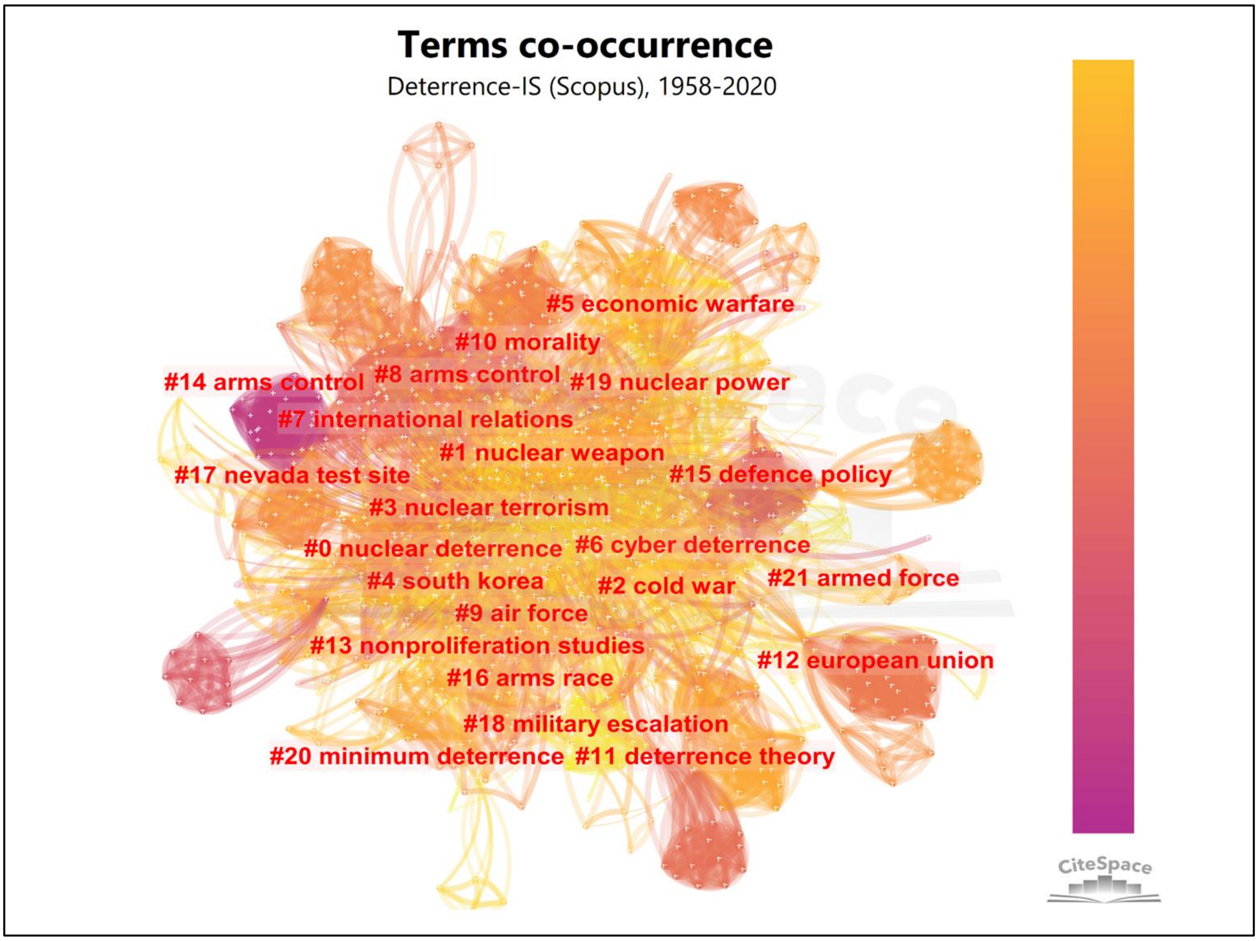


Figure 33: Term co-occurrence network: deterrence-IS, with labeled clusters

CiteSpace, however, allows us to go even beyond this initial 'big picture' analysis of the field. Scholars (as well as policy makers) are, for instance, often interested in particularly active areas of research or in emerging new concepts or insights. These can now also be algorithmically identified through citation or term bursts[139], which essentially represent the actual 'substance' behind the more 'technical' entropy metric we presented in the previous technical section (Research entropy). Such bursts can be detected within the CiteSpace software program based on the metadata it finds in the bibliometric datasets. A *citation* burst, for instance, provides evidence that a certain publication has evidently attracted an unusually high degree of attention from the scholarly community and may be turning into a new emerging 'giant' that new publications should probably at least be cognizant of. A *term* burst shows a similar comparative peak in attention but in this case based on the increased usage and centrality of a particular term in titles, abstracts, or keywords[140]. Figure 34 shows the top 10 term bursts in our bibliometric Deterrence-IS (Scopus)

[139]Jon Kleinberg, "Bursty and Hierarchical Structure in Streams," *Data Mining and Knowledge Discovery* 7, no. 4 (2003): 373–97; Chaomei Chen, *How to Use CiteSpace* (Leanpub, 2019), https://leanpub.com/howtousecitespace.

[140]In terms of *precision*, using bibliometric datasets therefore may still possess certain advantages over using full-text corpora, because the occurrence of the word "Russia" in a title or an abstract or in the keywords provides a more reliable indicator that Russia really is an important focus of that publication and not just a (potentially tangential) reference somewhere in the full-text of a publication. In terms of *recall*, however, we remain interested in at least being able to retrieve such tangential references

dataset as detected by CiteSpace. We observe that the most recent burst terms in this dataset include cyberdeterrence, nuclear war, North Korea, and strategic stability.

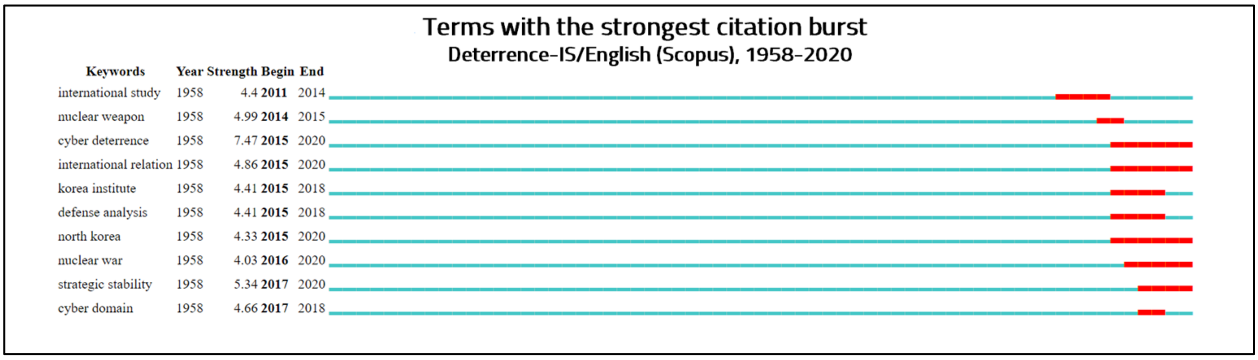


Figure 34: Terms with the strongest citation burst - deterrence-IS

Deterrence-IS-Russia/English

Our second step was to map our Russia-specific Deterrence-IS-Russia/English (Scopus) dataset and to compare it with the broader Deterrence-IS one. In order to investigate these datasets, we use the same research approach as before. We first extract and rank noun-phrases from titles, abstracts, and keywords. Secondly, we use CiteSpace to plot the noun-phrases and their connections in term co-occurrence networks. Finally, we provide an initial human interpretation of these findings.

Table 6 displays the 10-top noun-phrases according to the tf*idf algorithm. As was the case with Deterrence-IS (Scopus), nuclear deterrence has the highest frequency (tf 51) but also scores very low on inverse document frequency and is therefore excluded from this list.

tf	idf	tf*idf	term
41	0.69	28.42	nuclear weapons
18	1.39	24.95	cold war
13	1.79	23.29	strategic stability
11	2.08	22.87	nuclear disarmament
7	2.48	17.39	arms control
7	2.48	17.39	international security
7	2.48	17.39	north korea
7	2.48	17.39	soviet union
6	2.64	15.83	missile defenses
6	2.64	15.83	nuclear arms control

Table 6: Most salient terms - deterrence-IS-Russia/English

To compare the relative weight of terms in our Russia-specific Deterrence-IS-Russian/English

---

in publications whose main focus may lie elsewhere, but that may (after some creative human sense-making) provide a different – and possibly stimulating – insight into Russian deterrence.

(Scopus) and our broader Deterrence-IS (Scopus) dataset, we rescaled all of our extracted noun phrases' tf*idf values to a scale from 0 to 100. The most important term within each dataset thus receives the value '100', the lowest one '1', and all other values are distributed proportionally between these two extremes to yield a comparative metric of their relative importance. This yields the following 'slope chart' .

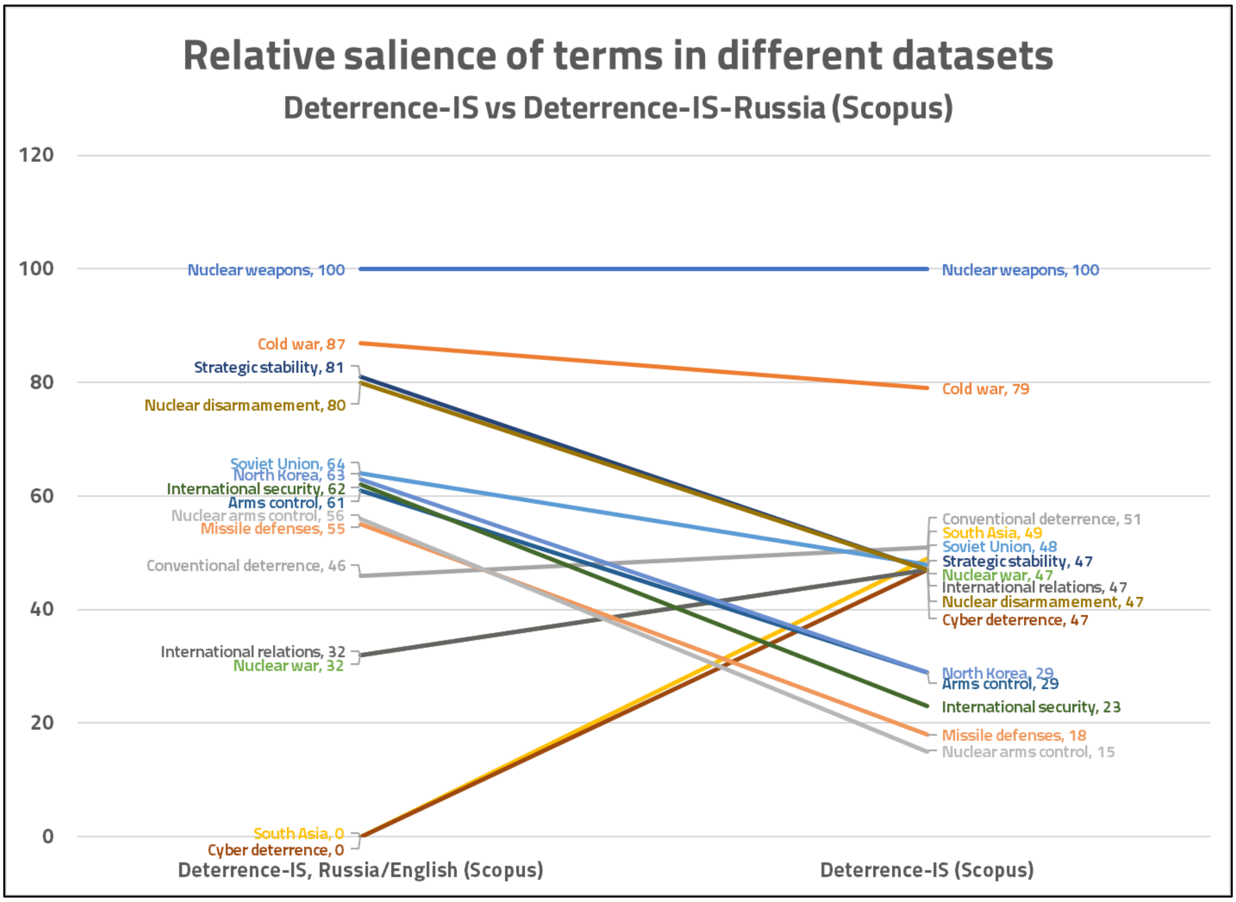


Figure 35: Comparing the most salient terms: deterrence-IS vs deterrence-IS-Russia/English

We want to draw attention to at least a few interesting findings from this visual:

- Nuclear weapons dominate both datasets, which confirms the overall importance of nuclear-related deterrence. Cyberdeterrence, however, plays an insignificant role in the Russia-specific dataset;
- In terms of the Russia-specific dataset, scholars pay significantly more attention to issues of arms control, nuclear disarmament, and nuclear arms control; and
- Overall scholars seem much more interested in missile defense in the Russia-specific dataset.

Connecting the noun-phrases into a term co-occurrence network yields the following major themes and their evolution over time.

- As was the case with the non-Russia-specific international deterrence studies, nuclear deterrence and nuclear weapons are among the most central discussions. However, the themes of missile defense, Arctic region, and Armed Force play a more important role in the discussion about specifically Russia-related deterrence than in the broader deterrence-IS dataset.
- While some clusters seem fairly obvious – #0 ('missile defense') and #1 ('nuclear deterrence'), #3 ('Cold war') and #1 ('nuclear deterrence'); we still find a few that are somewhat counter-intuitive. Discussion of #7 ('missile defense') is closely linked to the #14 ('Arctic region') and #5 ('armed force'), while cluster #11 ('nuclear navy') is related to #6 ('crisis management') and #0 ('missile defenses).
- In terms of historical evolution, Clusters #6 ('crisis management) and #10 ('critical analysis'), #11 ('nuclear navy'), and #8 ('Russian government') are the oldest themes. The most recent ones are #4 ('conventional deterrence'), #2 ('nuclear weapon'), and #12 ('international relations').

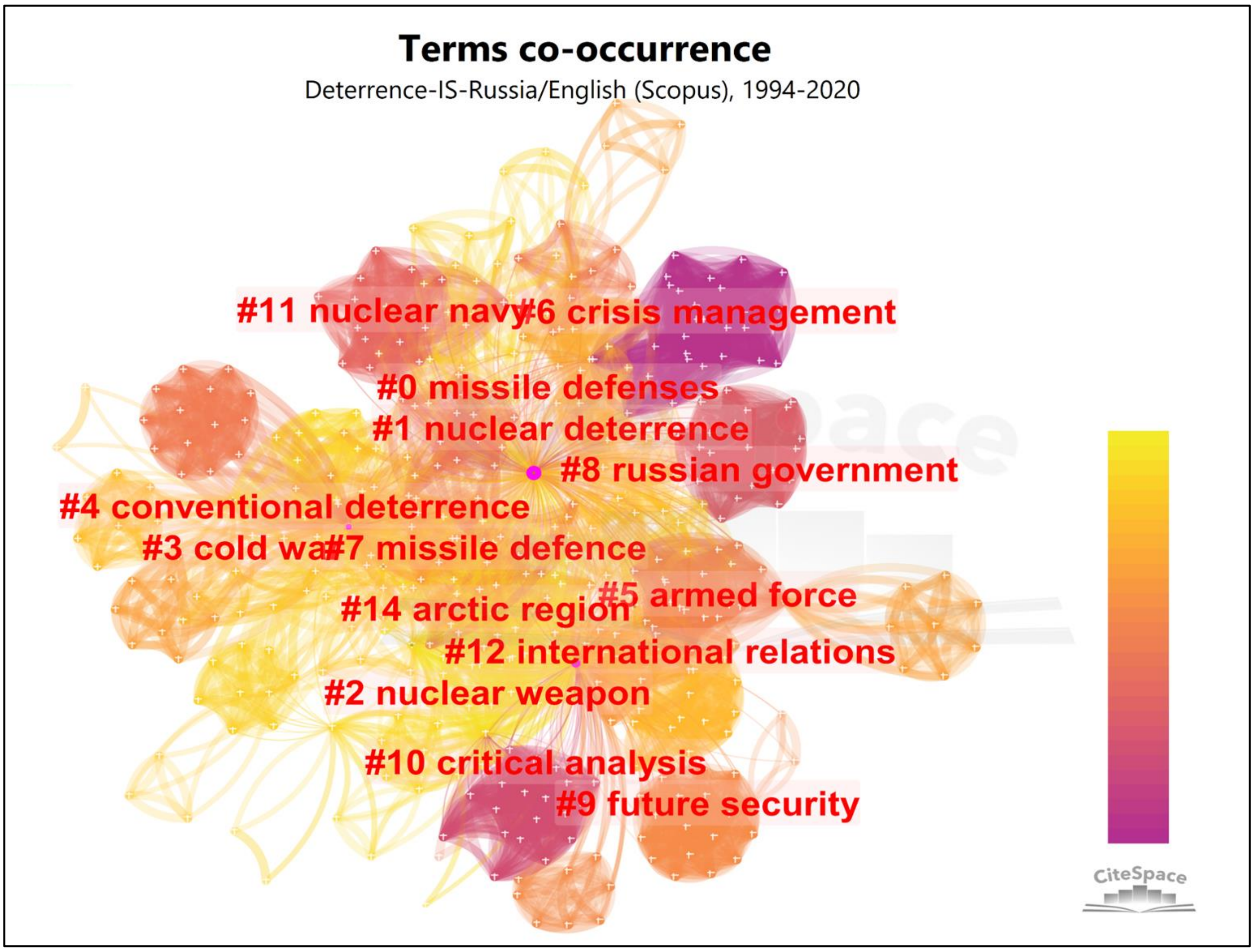


Figure 36: Term co-occurrence network: deterrence-IS-Russia/English, with labeled clusters

CiteSpace also allows users to visualize and explore clusters' evolution on a timeline. In the case of terms co-occurrence analysis, the placement of a node on the timeline is calculated based on the average year of publication of documents that contain the keyword in their abstract. The timeline with nodes laid out by the average year of publication allows researchers to investigate the evolution of each cluster from a historical perspective.

Figure 37 displays a timeline layout for the Deterrence-IS-Russia/English (Scopus) dataset. Each thick horizontal line represents a cluster. They are labeled on the right side of the timeline with the same labels as the graph visualization layout. The nodes, which are represented by noun-phrases from titles, abstracts, and keywords of documents in the dataset, are placed on each of the horizontal lines according to the affiliated cluster. The purple dates on top mark the average year of publication.

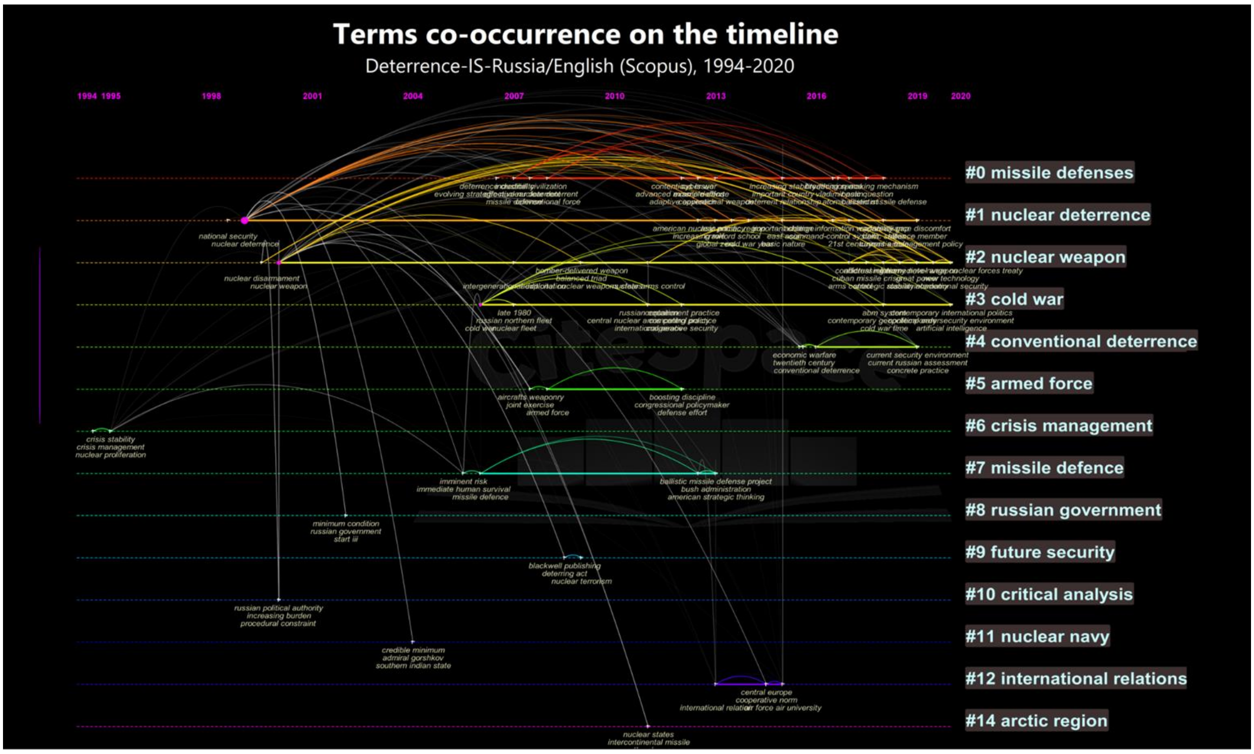


Figure 37: Term co-occurrence per cluster over time - deterrence-IS-Russia/English[141]

The timeline in Figure 37 can be used to draw a variety of preliminary conclusions with regards to the evolution of each of the themes. Let us, however, highlight a few insights that struck us.

- In 2007, the #0 ('missile defense') cluster was discussed in the context of deterrence credibility, strategic nuclear deterrence, effective nuclear deterrence. Since 2016, it is increasingly focused on personalities (Vladimir Putin) and decision-making mechanisms.
- #1 ('nuclear deterrence) was all about national security in 1998. In between 2013 and 2016, the main focus shifted to American nuclear primacy, the Asia-Pacific region, and East

[141]Unfortunately, the timeline can be unclear in some places due to the number of noun-phrases used and the overlay between them.

Asia. Presently, the cluster is all about capability gap, Baltic states, alliance discomfort, alliance members, and crisis management policy.

- #3 ('Cold war) was linked to the late 1980, nuclear fleet, and the Russian Northern fleet. The cluster currently is investigated in terms of contemporary international politics, security environment, and artificial intelligence. Apparently, scholars attempt to connect the Cold war experience to the realities of modern security and international politics.

The information displayed in the CiteSpace timeline layout is one useful way to investigate the evolution of themes. Yet the clusters that do not have a high frequency or overall citation rate but were still parts of major discussions might get lost among other nodes. To mark such themes, CiteSpace provides a list of terms with the strongest citation burst. We have already described the principle behind the algorithm when we discussed bursts in the context of our analysis of the Deterrence-IS/English (Scopus) dataset. Figure 38 uses the same algorithm to mark the strongest term bursts in the English-language Russia-specific dataset between 1994 and 2020. The most recent ones are related to arms control, nuclear disarmament, and nuclear weapons. 'Arms control' recently experienced the strongest term burst (2.54) among all the terms since 1994.

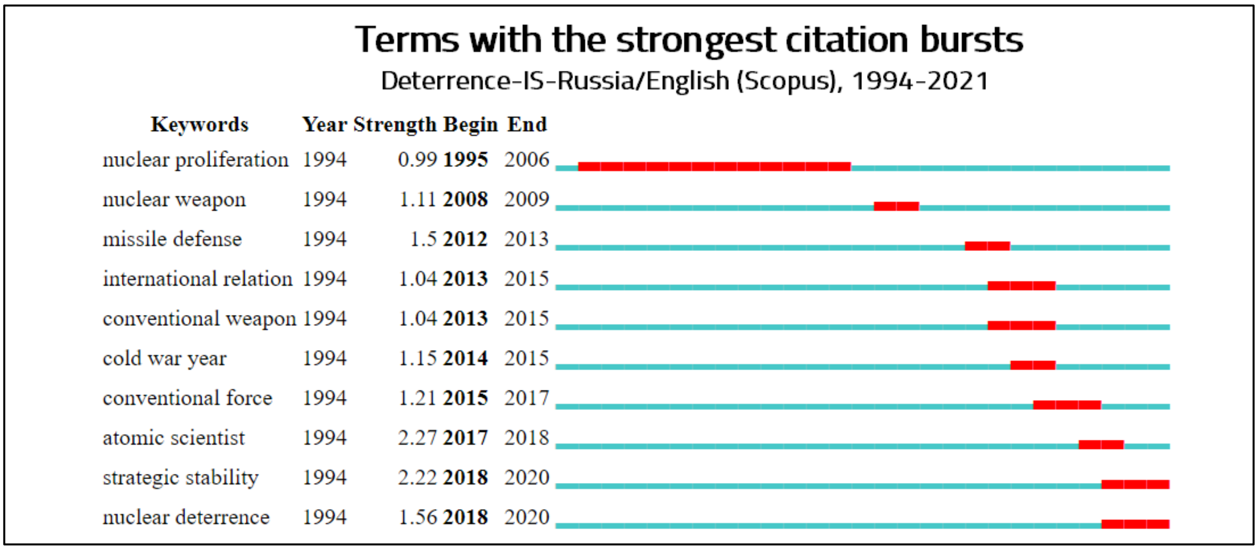


Figure 38: Terms with the strongest bursts - deterrence-IS-Russia/English

Deterrence-IS-Russia/Russian

The previous section analyzed the substance of international studies on (Russian) Deterrence-IS. But what do Russian scholars themselves write about Russia and deterrence? To investigate this, we can only use the Deterrence-IS-Russia/Russian (WoS-RSCI) dataset. Although these publications are all in Russian, the WoS dataset also contains abstracts, titles, and keywords in English. We use the English versions of these publications' metadata in order to make the analysis more accessible for an international audience. The following table highlights the 10 most important terms for the Russian dataset.

term	tf	idf	tf*idf
nuclear weapons	11	1.39	15.25
nuclear deterrence	9	1.61	14.48
international relation	7	1.79	12.54
cold war	6	1.95	11.68
nuclear arsenal	5	2.2	10.99
strategic stability	5	2.2	10.99
asia-pacific region	4	2.4	9.59
inf treaty	4	2.4	9.59
21st century	3	2.71	8.12
arms control	3	2.71	8.12

Table 7: Most salient terms - deterrence-IS-Russia/English

Comparing the English (IS-Deterrence-Russia/English (Scopus)) and Russian datasets (IS-Deterrence-Russia/Russian(WoS)), we can draw a few conclusions:

- Both datasets are dominated by the nuclear deterrence and nuclear weapon terms. Conventional deterrence is less important in the Russian dataset. Cyberdeterrence is insignificant in both datasets;
- Arms control is common to both datasets. But nuclear disarmament and nuclear arms control appears to be less important terms for Russian scholars.

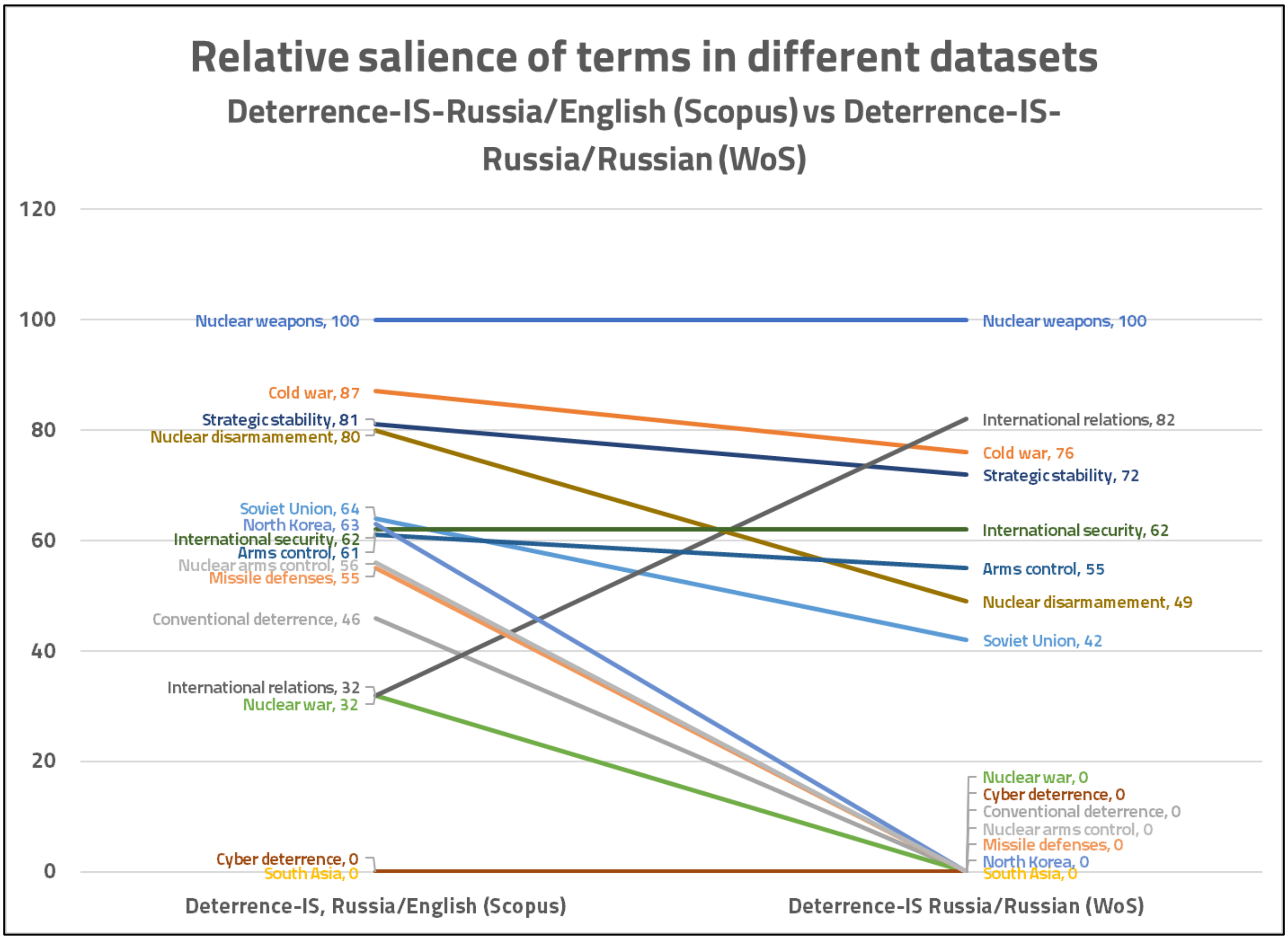


Figure 39: Comparing the most salient terms: deterrence-IS-Russia/English vs deterrence-IS-Russia/Russian

The term co-occurrence network reveals the following main clusters in the Deterrence-IS-Russia/Russian (WoS-RSCI) dataset:

- The most central themes for the Russian dataset are #1 ('the US') and #0 ("nuclear deterrence')
- #2 ('China'), #1 (the US), and #0 ('nuclear deterrence') are strongly related topics in the Russian dataset. #7 ('Trump administration') is linked to themes about US and nuclear deterrence, but interestingly there are few links with the #2 ('China'), suggesting that Russian scholars appear to be much less interested in Trump's policy toward China. #4 ('conventional deterrence') has weak links with #2 ('China'), but it is mostly studied separately from the rest of the topics.
- Clusters #0 ('nuclear deterrence'), #1 ('US'), #2 ('China'), and #7 ('Trump administration') are the most recent themes. #3 ('Cold War) and #4 ('conventional deterrence') are relatively old clusters.

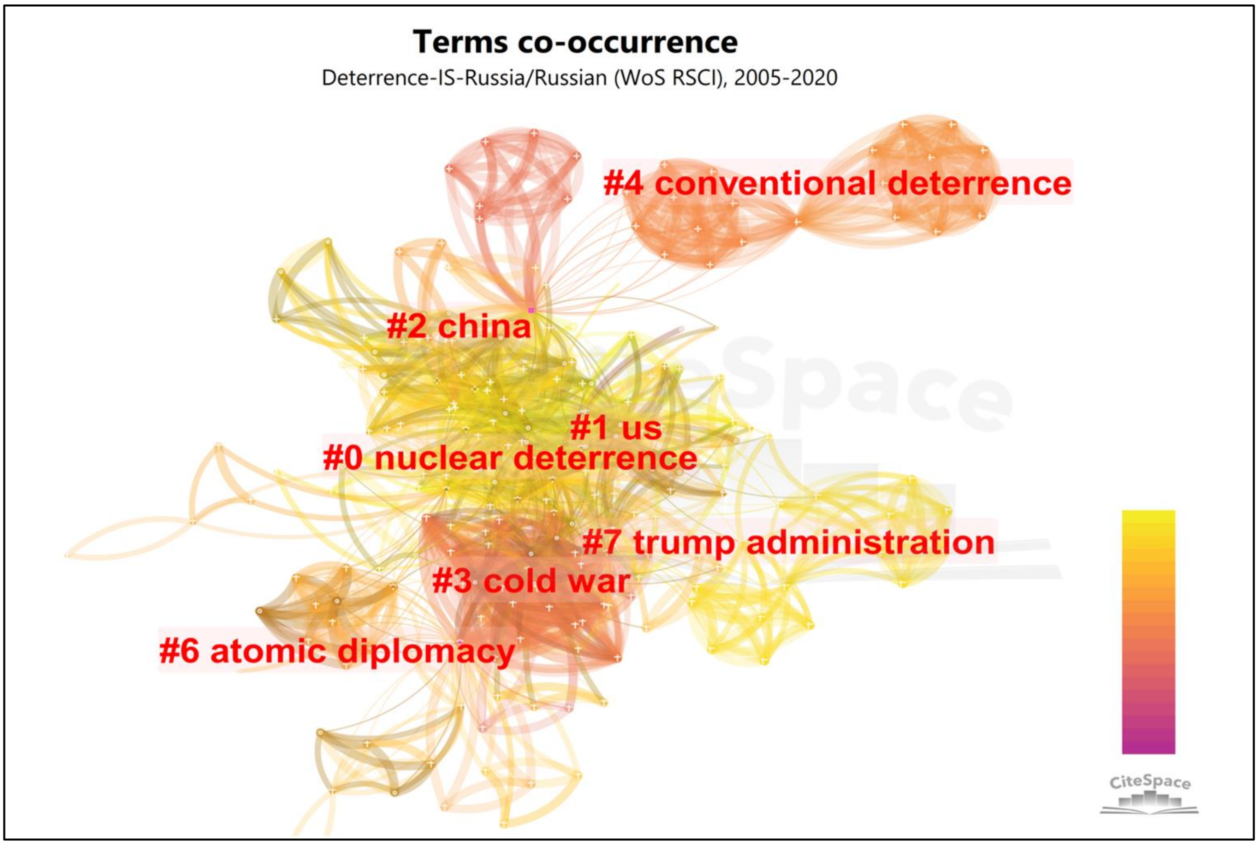


Figure 40: Term co-occurrence network - deterrence-IS-Russia/Russian

The evolution of these clusters also yields some interesting nuggets:

- #0 ('nuclear deterrence') cluster was investigated in terms of complete annihilation, military conflict, and nuclear weapon in 2013-2016. More recently, the focus of this cluster shifted to disarmament issues. The cluster's main terms are international security, nuclear disarmament, and the INF (intermediate-range nuclear forces) treaty.
- In 2016, the main focus in the #1 ('the US') cluster was on foreign policy, security, and achieving strategic parity. Since 2020, the cluster is more about Arctic politics and space activity.

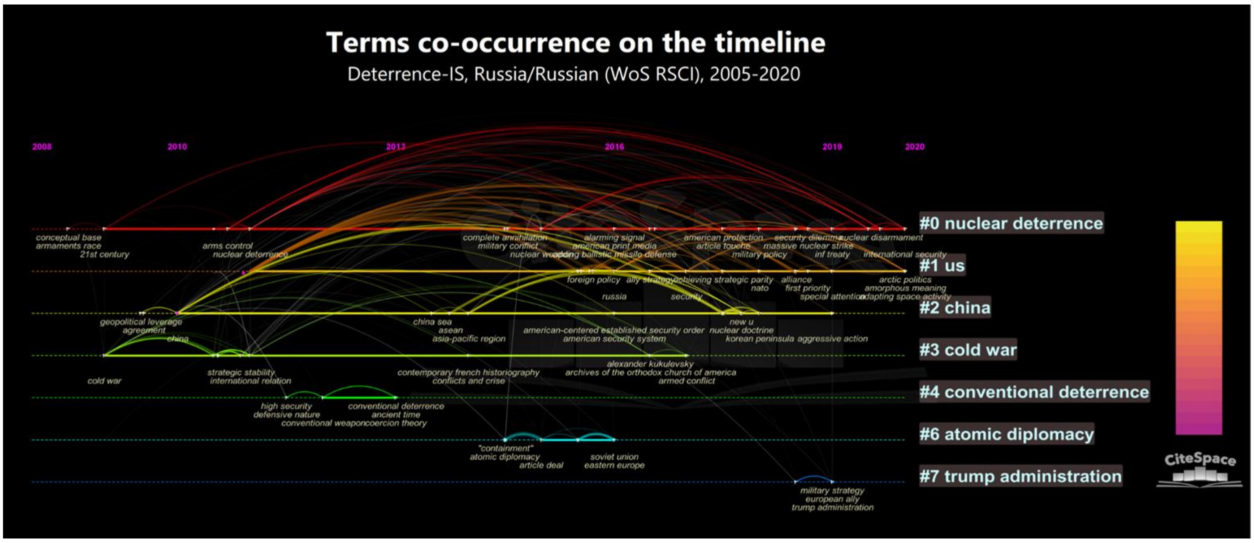
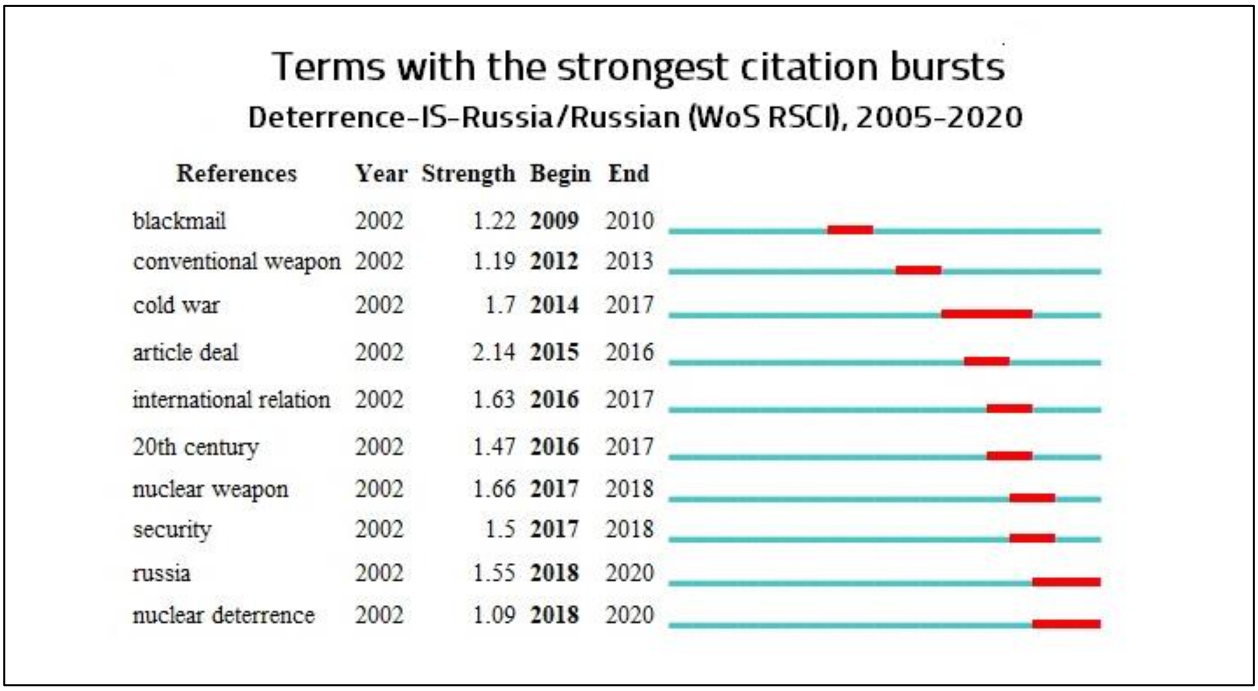


Figure 41: Term co-occurrence per cluster over time - deterrence-IS-Russia/Russian

Particularly important among those terms are those that experience citation bursts. Figure 42 below lists of these themes and shows the period of burst. As we can see, the most recent bursts were related to nuclear deterrence and Russia. The strongest burst was related to the 'nuclear weapon' term in between 2017 and 2018.


Terms with the strongest citation bursts

Deterrence-IS-Russia/Russian (WoS RSCI), 2005-2020

References	Year	Strength	Begin	End
blackmail	2002	1.22	**2009**	2010
conventional weapon	2002	1.19	**2012**	2013
cold war	2002	1.7	**2014**	2017
article deal	2002	2.14	**2015**	2016
international relation	2002	1.63	**2016**	2017
20th century	2002	1.47	**2016**	2017
nuclear weapon	2002	1.66	**2017**	2018
security	2002	1.5	**2017**	2018
russia	2002	1.55	**2018**	2020
nuclear deterrence	2002	1.09	**2018**	2020

Figure 42: Terms with the strongest bursts - deterrence-IS-Russia/Russian

Comparing English and Russian datasets, we note certain commonalities and differences:

- According to our noun-phrase analysis and our investigation of co-occurrence graphs, Russia-related deterrence is investigated predominantly in terms of nuclear deterrence by

both Russian and English-speaking scholars. Conventional deterrence has a less important place, and cyberdeterrence is still not present among the key themes at all.
- Both datasets still pay increased attention to the issues of disarmament – an interesting observation at a time that arms control and disarmament seemed to become much less dominant in the policy discourse. In the case of the Russia dataset, it is the main focus of the nuclear deterrence cluster, while for the English-language dataset arms control and nuclear disarmament are the terms that experienced the strongest citation bursts recently.
- While the English dataset contains mostly conceptual and theoretical issues (nuclear deterrence, nuclear weapon, strategic stability…), the Russian dataset devotes a significant parts of its academic efforts to specific countries and governments (US, China, Trump Administration). Russian deterrence publications therefore appear to be more applied/focused.

Using full-text

Sources: Full-text databases

As we pointed out, most of the world's epistemic patrimony rests encoded in scholarly publications. While bibliographical metadata can already provide us with some useful high-level insights, applying NLP to the full texts of these publications opens up far more and richer research avenues. Our team, therefore, collated a number of extensive full-text scholarly corpora on security deterrence (*Deterrence-IS*) from a wide array of publicly available English *and* Russian sources (see Table 8). Unfortunately, access to the full-text versions of these scholarly (and many other) publications remains even more difficult (and expensive) than access to bibliometric meta-data (see Bibliometric databases – the state of the field).

The international market for scholarly publishing clearly finds itself in full disruption as it moves from a 'closed' market structure to what looks likely to become a much more diversified one in which open access publications are already starting to play a more important role[142]. At this point in time, however, scholars' access to the world's epistemic patrimony is still gated in a number of commercial databases that contain full-text academic publications. These online databases are, in

[142]For some recent overviews, see Alexander Grossmann and Björn Brembs, "Current Market Rates for Scholarly Publishing Services," *F1000Research* 10 (January 12, 2021): 20, https://doi.org/10.12688/f1000research.27468.1; Julia Frankland and Margaret A. Ray, "Traditional versus Open Access Scholarly Journal Publishing: An Economic Perspective," *Journal of Scholarly Publishing* 49, no. 1 (October 1, 2017): 5–25, https://doi.org/10.3138/jsp.49.1.5; Vincent Larivière, Stefanie Haustein, and Philippe Mongeon, "The Oligopoly of Academic Publishers in the Digital Era," *PLOS ONE* 10, no. 6 (June 10, 2015): e0127502, https://doi.org/10.1371/journal.pone.0127502; David W. Lewis, "Is Scholarly Publishing Like Rock and Roll?," *Journal of Librarianship and Scholarly Communication* 8, no. 1 (November 9, 2020): 2333, https://doi.org/10.7710/2162-3309.2333; Jaime A. Teixeira da Silva et al., "Predatory and Exploitative Behaviour in Academic Publishing: An Assessment," *The Journal of Academic Librarianship* 45, no. 6 (November 1, 2019): 102071, https://doi.org/10.1016/j.acalib.2019.102071; Cassidy R. Sugimoto and Vincent Larivière, *Measuring Research: What Everyone Needs to Know* (Oxford University Press, 2018); National Academies of Sciences Medicine Engineering, and et al., *Open Science by Design: Realizing a Vision for 21st Century Research* (National Academies Press, 2018).

the English language domain, *in certain cases* developed, curated, and marketed by content copyright-holders (*academic publishing houses* like Reed-Elsevier, Springer, Taylor & Francis, Wiley-Blackwell, etc.; *university presses* such as Cambridge Core, The MIT Press, National Academies Press, Oxford Academic Journals, etc.; but also *learned societies*, *international organizations*, specialized commercial companies like Jane's Information Group, etc.); *in other cases* by academic 'aggregators' like EBSCO, Gale Academic, JSTOR, ProQuest, etc. As we saw was the case with the bibliometric databases we used in the first section of this paper, access to full-text scholarly publications has been improving these past few years. New and disruptive (and legal) major entrants in this field include Semantic Scholar Open Research Corpus, an open-source general-purpose corpus managed by the Allen Institute for Artificial Intelligence (ai2[143]) with about 15 million scientific papers from different academic disciplines to be used for NLP and text mining purposes[144] and especially Unpaywall[145], an open database of some 30 million free scholarly articles offered by *Our Research*, a nonprofit dedicated to making scholarship more accessible to everyone[146].

Despite undisputed progress, however, the overwhelming majority of academic (often publicly funded) text-encoded knowledge remains behind paywalls. If we take into consideration that Microsoft Academic Services contains 250 million publications[147], for instance, we realize that these open-access improvements still have a long way to go. The following visual, extracted from Lens.org, shows the percentage of all scholarly articles contained in our deterrence-related datasets that are 'open-access'. We see that – also on this metric – the situation is significantly worse than in other comparable fields: only 13% of academic pubs in the Deterrence-IS field is publicly available, compared to more than double that amount in the Deterrence-broad field (30%) and also worse than the overall 'international security' field, where 18% of publications are open access. And even within these open-access sources, the 'gold' (where the content is published in fully open access journals) and 'green' ones (where some version of the content is publicly available, even if the final fully formatted pdf remains behind the publisher's paywall) are significantly more numerous in the Deterrence-broad dataset[148].

[143]Allen Institute for AI, "Allen Institute for AI," Allen Institute for AI, 2012, https://allenai.org/.
[144]Allen Institute for AI, *Allenai/S2orc*, Python (2019; repr., AI2, 2021), https://github.com/allenai/s2orc.
[145] Holly Else, "How Unpaywall Is Transforming Open Science," *Nature* 560, no. 7718 (August 15, 2018): 290–91, https://doi.org/10.1038/d41586-018-05968-3.
[146]Heather Piwowar, Jason Priem, and Richard Orr, "Our Research," Our Research, 2021, https://ourresearch.org/.
[147]Microsoft, "Home | Microsoft Academic," Microsoft Academic, 2021, https://academic.microsoft.com/home.
[148]We urge our readers to bear in mind that these scholarly databases do not include the so-called 'grey literature', which includes the – in this subject area not unimportant – publications by think tanks, international organizations, etc., most of which tend to be in the 'diamond' OA-category: publications by not-for-profit, non-commercial organizations, associations or networks that are made available online in digital format, are free of charge for readers and authors and do not allow commercial and for-profit re-use. "A non-profit academic publishing model that makes academic knowledge a common good". Christian Fuchs and Marisol Sandoval, "The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious," *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society* 11, no. 2 (September 9, 2013): 428–43, https://doi.org/10.31269/triplec.v11i2.502.

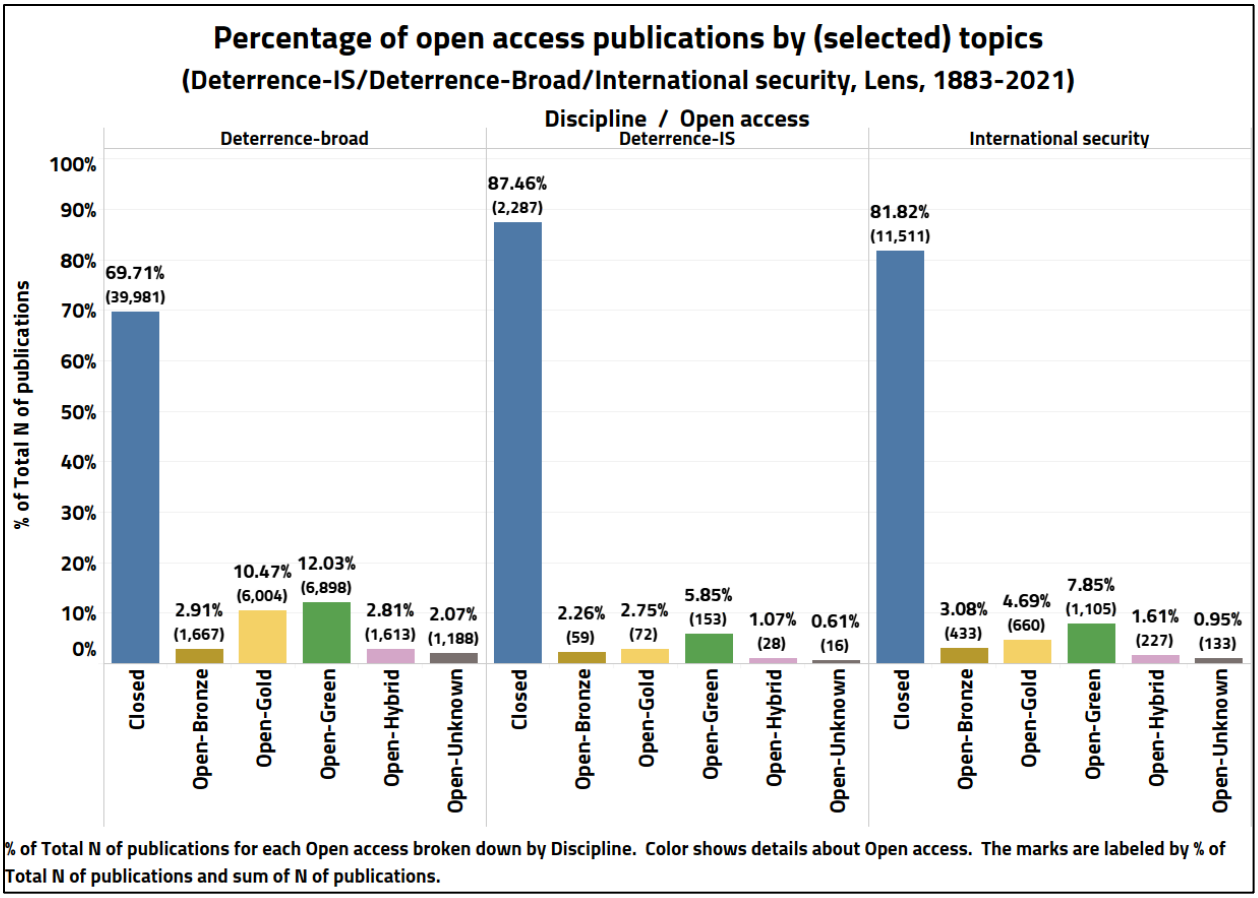


Figure 43: Percentage of open-access publications by dataset

Building full-text corpora

The following table provides a synoptic overview of the datasets we collected on (Russian) deterrence, including the final tally of the full-text documents we retrieved.

Available data → Data sources ↓	Corpora	Years covered[149]	Number of documents
EVP	Deterrence-IS/Russian	2010-2020	5,609
Integrum	Deterrence-IS/Russian	2010-2020	30,869
Militera	Deterrence-IS/Russian	2010-2019	1,026
Russian/Soviet Military Books	Deterrence-IS/Russian	2010-2019	76
"Seminal" works	Deterrence-IS/Russian	2010-2019	36
ProQuest Central	Deterrence-IS/English	1939-2020	47,999

[149] The time coverage is based on the earliest (typically the oldest one in the database) and latest available (typically the most recent one at the time of downloading) papers in a given dataset.

***ProQuest Central**	Deterrence-IS/Russia/English	1958-2020	7,436
EBSCO Ultimate Academic	-	1945-2020	695
Google Scholar	Deterrence-IS	1990-2020	44,888
***Google Scholar-Academic**	Deterrence-IS-Academic	1990-2020	16,881
***Google Scholar**	Deterrence-IS-Russia	1990-2020	1,887
***Google Scholar**	Deterrence-IS-Think tanks	1990-2020	550

Table 8: Collated Russian and English full-text datasets

To build these corpora our team was required to apply slightly different deterrence-related queries for different databases, because most of them – regrettably – use different search query syntaxes. We also created some subcorpora by filtering out documents based on additional queries which contain subcorpora-related publications and terms like in the case of the Russia-specific subcorpora of ProQuest Central. The following table provides an overview of the search queries we used for the different databases.

<table>
<tr><th>Available data →
Data sources ↓</th><th>Search queries</th></tr>
<tr><td>EVP</td><td rowspan="4">(сдерживан* OR устрашен* OR запугиван* OR шантаж*) AND (ядерн* OR атомн* OR кримин* OR преступ* OR оруж* OR вооруж* OR воен*)</td></tr>
<tr><td>Militera</td></tr>
<tr><td>Russian/Soviet Military Books</td></tr>
<tr><td>“Seminal” works</td></tr>
<tr><td>Integrum</td><td>(сдерживание ИЛИ устрашение ИЛИ запугивание ИЛИ шантаж) И (неядерный ИЛИ предъядерный ИЛИ реалистический ИЛИ стратегический ИЛИ нестратегический ИЛИ взаимный ИЛИ региональный ИЛИ силовой ИЛИ несиловой ИЛИ минимальный ИЛИ оборонительный ИЛИ наступательный ИЛИ двойной ИЛИ конвенциональный ИЛИ коалиционный ИЛИ многосторонный[150] ИЛИ двухсторонний ИЛИ глобальный ИЛИ региональный ИЛИ punishment ИЛИ denial) /п1 И (безопасность ИЛИ военный ИЛИ оборона</td></tr>
<tr><td>ProQuest Central</td><td rowspan="2">“nuclear deterrence" OR "military deterrence" OR "conventional deterrence" OR "hybrid deterrence" OR "cyber deterrence" OR cyberdeterrence OR "cross-domain deterrence</td></tr>
<tr><td>*ProQuest Central</td></tr>
</table>

[150] This spelling mistake was discovered too late in the game to be rectified.

EBSCO Ultimate Academic	“nuclear deterrence" OR "military deterrence" OR "conventional deterrence" OR "hybrid deterrence" OR "cyber deterrence" OR cyberdeterrence OR "cross-domain deterrence
Google Scholar	"nuclear deterrence" OR "military deterrence" OR "conventional deterrence" OR "hybrid deterrence" OR "cyber deterrence" OR cyberdeterrence OR "cross-domain deterrence"

Table 9: Search queries for English and Russian full-text datasets

Our team subsequently ‘cleaned’ and consolidated these full-text datasets, which may sound like a straightforward task but decidedly is anything but.

Available data → Corpora used ↓	Sources	Years covered[151]	Number of documents
Deterrence-IS/Russian	Deterrence-IS/Russian	2010-2020	37,616
Deterrence-IS/Russia/Russian	Deterrence-IS/Russian	2010-2020	6,509
Deterrence-IS/Russian/Academic	EVP	2010-2020	5,609
Deterrence-IS/English	PQ Central	1939-2020	47,999

Table 10: Final corpora that were used in the analysis

For our English language corpus we used ProQuest’s central database – a collection of thousands of full-text scholarly journals, trade and professional titles, newspapers, magazines, dissertations, working papers, case studies[152] – as our main source. [We also built corpora from EBSCO Ultimate Academic[153] and Google Scholar[154], but did not use them for further analysis based on manual coding and unsupervised topic modeling for reasons of time economy]. For our topic modeling effort, we also extracted from our general PQ corpus a narrower one on deterrence ***and*** Russia by retaining only excerpts that contained Russian-related (russi*) tokens, which resulted in 7,436 texts.

For Russian sources, this paper relies essentially on two major commercial aggregators of Russian full-text content: Eastview Press[155], a US-based aggregator providing access to much academic,

151The year coverage is based on the earliest (typically the oldest one in the database) and latest available (typically the most recent one at the time of downloading) papers in a given dataset.
152“ProQuest | Databases, EBooks and Technology for Research,” 2021, http://about.proquest.com/.
153EBSCO, “Home Page | EBSCO,” EBSCO Information Services, Inc. | www.ebsco.com, 2021, https://www.ebsco.com/.
154 Google Scholar, “About Google Scholar,” Google Scholar, May 6, 2021, https://scholar.google.com/intl/en/scholar/about.html.
155 “East View Press,” East View Press, 2021, https://www.eastviewpress.com/.

military, other periodicals, etc. full-text content[156]; and Integrum[157], a Russia-based aggregator that has more extensive coverage on (central and regional) media articles and official publications. Our entire Deterrence-IS/Russian corpus was furthermore expanded with relevant texts from a collection of Russian military books, Militera library publications[158] and number of additional 'seminal' papers[159]. For our final full-text Russian corpus we renamed the metadata columns of text in a consistent, unified way and then merged the datasets from EVP and Integrum databases in python. As a result, we obtained a corpus with 37,616 relevant Russian full-text publications on Deterrence-IS.

We want to point out that our full-text corpora are quite different from the more purely 'scholarly' bibliometric datasets we used in our first section but also from the academic AND official/military dataset that we coded manually and that we will report on in our Mnaula NLP section (Manual NLP – Human coding of 'points of interest' in corpora). Our Russian consolidated corpus, for instance, also contains a significant number of newspaper articles on Deterrence-IS from the 'central' printed media that are covered in the Integrum database. Although the search queries were identical or at least similar, the results, as we shall see, were quite different.

Getting the corpora ready for processing

The datasets that were exported from these different databases still required a number of preparatory steps before they could be processed and analyzed by NLP-tools.

The English and Russian corpora were initially converted into the JSON file format, maintaining essentially the same metadata structure[160]. In terms of pre-processing, we first split the full-text field into paragraphs for better text processing, with each paragraph inheriting the metadata from its parent document in the JSON schema. For one of these metadata fields, the date of publication, we created and applied a script that converted all the different date formats encountered throughout the corpus into one consistent date format, to ensure that we would be able to analyze the content of the corpus over time as well. We then extracted those paragraphs that contained deterrence-related text excerpts using python's 'regular expression operations' ('re') module. We subsequently tokenized the text into relevant words/phrases and lemmatized those (a process of stripping a word from all of its inflectional grammatical elements down to its 'lemma'-form) using

---

[156]The following EVP's UDB-databases were used: Higher School of Economics Academic Journals, Moscow University Press, Russian Military and Security Periodicals, Russian Social Sciences & Humanities, St.Petersburg University Press, Russian Central Newspapers, Russian Governmental Publications, Russian Regional Newspapers.
[157]"Integrum World Wide," 2021 1998, http://www.integrumworld.com/.
[158]"Militera: Military Literature," 2001, http://militera.lib.ru/.
[159]Those papers included, for instance, references from Dima Adamsky's – admirably copiously referenced – works on deterrence that we did not yet have in our EVP or Integrum corpora and that we therefore manually added.
[160]"HCSS-StratBase/HCSS-Deterrence-Datasets."

NLTK[161] and Gensim[162] Python libraries for English and SpaCy[163] for Russian corpus lemmatization[164]. Our team is more than willing to make our scripts and notebooks available to interested colleagues as a shared resource.

The resulting JSON-files were then subjected to different combinations of human and machine algorithms to 'x-ray' the content of the deterrence studies and the place of Russia in them. This is what the following sections will report on.

Unsupervised NLP – Topic modelling

We already saw how substantive bibliometric analysis can assist researchers in mapping the 'knowledge landscape' in a field or a subfield. In this bibliometric approach, the substantive clusters are algorithmically identified based on their ***bibliographic metadata*** and not on the publications' actual full-text content. In this part of our report we want to explore how useful various computer algorithms can be in identifying the main substantive topics in a corpus of scholarly publications based on their ***full-text content***.

Topic modeling[165] is one of the most popular NLP applications that allows us to do precisely that: to identify topics that best describe a set of documents. Technically speaking, topic modeling refers to a collection of unsupervised statistical learning methods that aim to discover latent topics in larger text corpora. They first try to identify the key terms – individual nouns like 'deterrence', or nouns phrases like 'nuclear deterrence' or 'nuclear weapon proliferation' – that tend to travel together through a text corpus without any prior knowledge about the subject and without human 'supervision'. Those key terms that tend to co-occur in different publications (or in selected excerpts within them) are then clustered in a 'topic'. The algorithms calculate the probability distributions of topics over publications within a corpus (e.g. this publication in this corpus is likely to be 30% about topic X, 20% about topic Y, etc.) ***and*** the distribution of terms over topics (e.g. the most 'influential' term in topic 1 is term A, the second one is term B, etc.).

There are many different topic modeling algorithms. Older ones (like LDA, Latent Dirichlet Allocation) are based on 'bag-of-words' feature representations of publications which look for terms that co-occur anywhere in a document and that also do not reflect any semantic similarity (e.g. they do recognize that 'small', 'smaller' and 'smallest' have something in common because

---

[161]Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing with Python*, 1st ed (Beijing ; Cambridge [Mass.]: O'Reilly, 2009).
[162]Radim Řehůřek and Petr Sojka, "Software Framework for Topic Modelling with Large Corpora," in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (Valletta, Malta: ELRA, 2010), 45–50.
[163] Matthew Honnibal et al., *SpaCy: Industrial-Strength Natural Language Processing in Python* (Zenodo, 2020), https://doi.org/10.5281/zenodo.1212303.
[164]Bird, Klein, and Loper, Natural Language Processing with Python.
[165]For a useful overview of what topic modelling is and how it is used, see Jordan Boyd-Graber, Yuening Hu, and David Mimno, "Applications of Topic Models," *Foundations and Trends® in Information Retrieval* 11, no. 2–3 (2017): 143–296, https://doi.org/10.1561/1500000030.

they share the same 'stem' or 'lemma' – small; but they do not 'know' that 'big' and 'small' also have something in common). More recent implementations of topic modeling like top2vec, (D)ETM, BERTopic, etc. take advantage of the newer (post-2018) more context-aware language models that also take positional and semantic proximity into account. In this section, we present some illustrative findings based on the application of some of the aforementioned modeling algorithms on our English and Russian full-text corpora.

Latent Dirichlet Allocation (LDA)

LDA is one of the earliest and still most popular topic modeling methods. It uses Bayesian statistics and Dirichlet distributions through an iterative process to identify coherent latent topics in a text corpus[166]. Running the model requires specifying two hyperparameters: alpha and beta. The alpha parameter represents the document-topic density – 'with a higher alpha, documents are assumed to be made up of more topics and result in more specific topic distribution per document'. The beta parameter is essentially the same but represents the topic-word density, – 'with high beta, topics are assumed to made of up most of the words and result in a more specific word distribution per topic'[167].

Our LDA model was generated by applying the Gensim LDA package in Python to our preprocessed corpora. Gensim LDA topic modeling requires two main inputs: a preprocessed corpus (converted into a bag of words format) and a dictionary (an object topic model used to reference terms). Also, LDA requires setting the hyperparameters described above and defining the number of estimated topics. All models were generated by setting the hyperparameter alpha = 0.2 and beta = 0.01. Another important decision to be made by the researcher in any topic modelling effort is the optimal number of topics to be identified. In this case, this was estimated in the case of English corpora by applying Gensim coherence-perplexity metrics. It helped to estimate the coherence topics score as 0.51 and the estimation of the optimal number of topics as 23. In the case of Russian corpora, these metrics were not used due to inconsistent results and the optimal number of topics was chosen manually by substantively assessing the results from different preset numbers and then selecting the one that made the most sense to our team.

---

[166]See David M. Blei and John D. Lafferty, "Dynamic Topic Models," in *Proceedings of the 23rd International Conference on Machine Learning – ICML '06* (the 23rd international conference, Pittsburgh, Pennsylvania: ACM Press, 2006), 113–20, https://doi.org/10/bjkmq9; Sergey I. Nikolenko, Sergei Koltcov, and Olessia Koltsova, "Topic Modelling for Qualitative Studies," *Journal of Information Science* 43, no. 1 (February 2017): 88–102, https://doi.org/10/gfkpvz; Carina Jacobi, Wouter Van Atteveldt, and Kasper Welbers, "Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling," *Digital Journalism* 4, no. 1 (2016): 89–106, https://doi.org/10/f3s2sg. For some more intuitive explanations of LDA, see Thushan Ganegedara, "Intuitive Guide to Latent Dirichlet Allocation," Medium, December 5, 2020, https://towardsdatascience.com/light-on-math-machine-learning-intuitive-guide-to-latent-dirichlet-allocation-437c81220158; Giri Rabindranath, "Topic Modeling with LDA: An Intuitive Explanation," HDS, October 19, 2020, https://highdemandskills.com/topic-modeling-intuitive/; Pratik Barhate, "Latent Dirichlet Allocation for Beginners: A High Level Intuition," Medium, February 5, 2019, https://medium.com/@pratikbarhate/latent-dirichlet-allocation-for-beginners-a-high-level-intuition-23f8a5cbad71; Luis Serrano, *Latent Dirichlet Allocation*, 2020, https://www.youtube.com/watch?v=T05t-SqKArY.

[167]Shashank Kapadia, "Topic Modeling in Python: Latent Dirichlet Allocation (LDA)," Medium, December 29, 2020, https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0.

The generated results were displayed via the pyLDAvis visualization package. It displays the relative size of the topics (by percent of tokens, – lemmatized terms in each topic), the distance between them and the top-30 tokens that appear in the corpus. In the interactive visualization, a researcher can hover the cursor over a certain topic and LDAvis will display the terms that are included in the topic. The pyLDAvis visualization also contains a lambda-slider for adjustment of relevance metric: setting lambda to 1 ($\lambda = 1$) visualizes the most overall salient terms within the topic, but moving the slider ($\lambda = 0$) allows the researcher to also view the most distinctive terms in that topic. We have found it very useful for any substantive interpretation of the essence of each topic to be able to see not only the most frequent but also the most unique terms. Based on this algorithmically derived information our team then proceeded to a human sense-making effort, in which we also paid attention to the size of the topic ("how important is it in this corpus"?) and the distances between topics ("how closely related is topic X to topic Y?") to generate some interesting insights from a number of the full-text corpora we collected.

Deterrence-IS/Russian

What are the main topics in our corpus of Russian publications on deterrence in an international security context (Deterrence-IS/Russian corpus – ~38k publications)? To find out, we used the LDA algorithm with the hyperparameters alpha = 0.2, beta = 0.01 and with 20 topics[168], visualized in LDAvis. The interactive LDA visualization can be consulted on HCSS' Github account[169], but we present some static representations of it here in the paper.

[168]The number of 20 topics was chosen by our team as yielding the most substantively meaningful topics within this corpus.
[169] The Hague Centre for Strategic Studies, *HCSS-StratBase/Deterrence-IS-Russia-Russian_LDA*, HTML, 2021, https://github.com/HCSS-StratBase/Deterrence-IS-Russia-Russian_LDA.

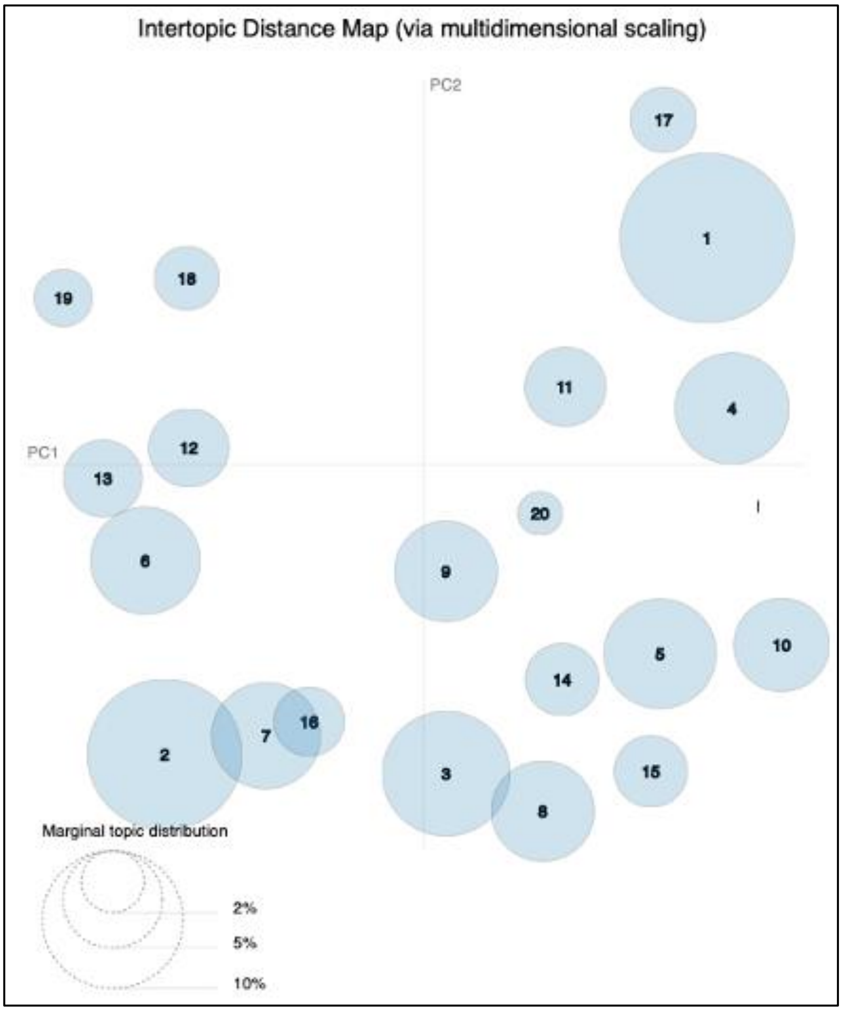


Figure 44: Intertopic distance map - LDA: Deterrence-IS/Russian

Topic #	**Topic description**	**% of tokens**
1	Post-Soviet countries I	15.2%
2	Nuclear deterrence	12%
3	General terms on IR and international security	8.2%
4	Post-Soviet countries II	6.5%
5	Economy and finance	6.3%
6	Military development and operational activity	6%
7	Russia/US-NATO deterrence in the military but not in nuclear terms	5.9%
8	Indo-Pacific region security	5.3%
9	Russia/US-NATO deterrence in the Europe	5.3%
10	Ideas and role of Russia in the world	4.6%
11	Russian internal policy process	3.3%
12	Naval nuclear issues	3.2%
13	Deterrence discourse in the official state documents	3.1%
14	MENA region security	2.7%

15	Soviet/USA Cold War security issues	2.7%
16	Air/space nuclear issues	2.5%
17	President's involvement into negotiations and foreign policy	2.2%
18	Russian Ministry of Defense and military development	2.1%
19	President's involvement in military planning and strategy	1.7%
20	Other/unidentifiable	1%

Table 11: HCSS-assigned topic labels - LDA: Deterrence-IS/Russian

LDAvis shows the frequency of term mentions within each topic which can help us researchers to interpret the topic's substantive meaning. For example, topic 1 shows us terms such as 'Украина' ('Ukraine'), 'Грузия' ('Georgia'), 'Киев' ('Kyiv'), 'Белоруссия' ('Belarus'), 'Молдавия' ('Moldova') which led us to interpret this topic as focusing on Post-Soviet countries and to label it as such. Table 11 shows the labels our team assigned to each topic after exploring their most salient terms.

Some of these topics (like 'nuclear deterrence') overlap with the topics/clusters that our substantive bibliometric analysis also revealed in the academic literature. But some of them still stand out – like the (quite dominant) post-Soviet ones or the 'financial and economic' one. A list of the most salient words in this Russian corpus (Table 12) confirms the importance of such 'real-life' issues and countries. This is an interesting finding in its own right which may in first instance surprise scholars working in this field but will become more comprehensible when they realize this dataset also covers newspaper publications. Our own take on this is that this findings once again highlights the striking gap between the theoretical debates on deterrence and its 'real-life' manifestations.

Top-30 Most salient terms		
Term in Russian	**Term translation in English**	**Frequency of mentions**
Сдерживание	Deterrence	31k
Украина	Ukraine	30k
Ядерный	Nuclear	27k
Военный	Military	27k
США	USA	23k
Стратегический	Strategic	19k
Российский	Russian	13k
РФ	RF (Russian Federation)	10k
Президент	President	10k
Система	System	10k
Безопасность	Security	10k
Нато	NATO	9k
Москва	Moscow	9k

Средство	Means	7k
Грузия	Georgia	7k
Политика	Politics	7k
Американский	American	7k
Ядерный оружие	Nuclear weapon	6k
Белоруссия	Belarus	6k
Китай	China	5k
Путин	Putin	5k
Российский-федерация	Russian Federation	5k
Янукович	Yanukovych	5k
Заявить	Declare	5k
Киев	Kyiv	5k
Тимошенко	Tymoshenko	4k
Ющенко	Yushchenko	4k
Приднестровье	Transnistria	4k
Противник	Adversary	4k
Европа	Europe	4k

Table 12: Top-30 Most salient terms - LDA: Deterrence-IS-Russia

Deterrence-IS-Russia/Russian

If we zoom in on the subset of the previous dataset that deals with Russia itself (Deterrence-IS/Russia/Russian – ~5600 publications) using an LDA model with alpha = 0.2, beta = 0.01, and specifying 13 topics), we obtain the following visualization[170].

[170] *The Hague* Centre for Strategic Studies.

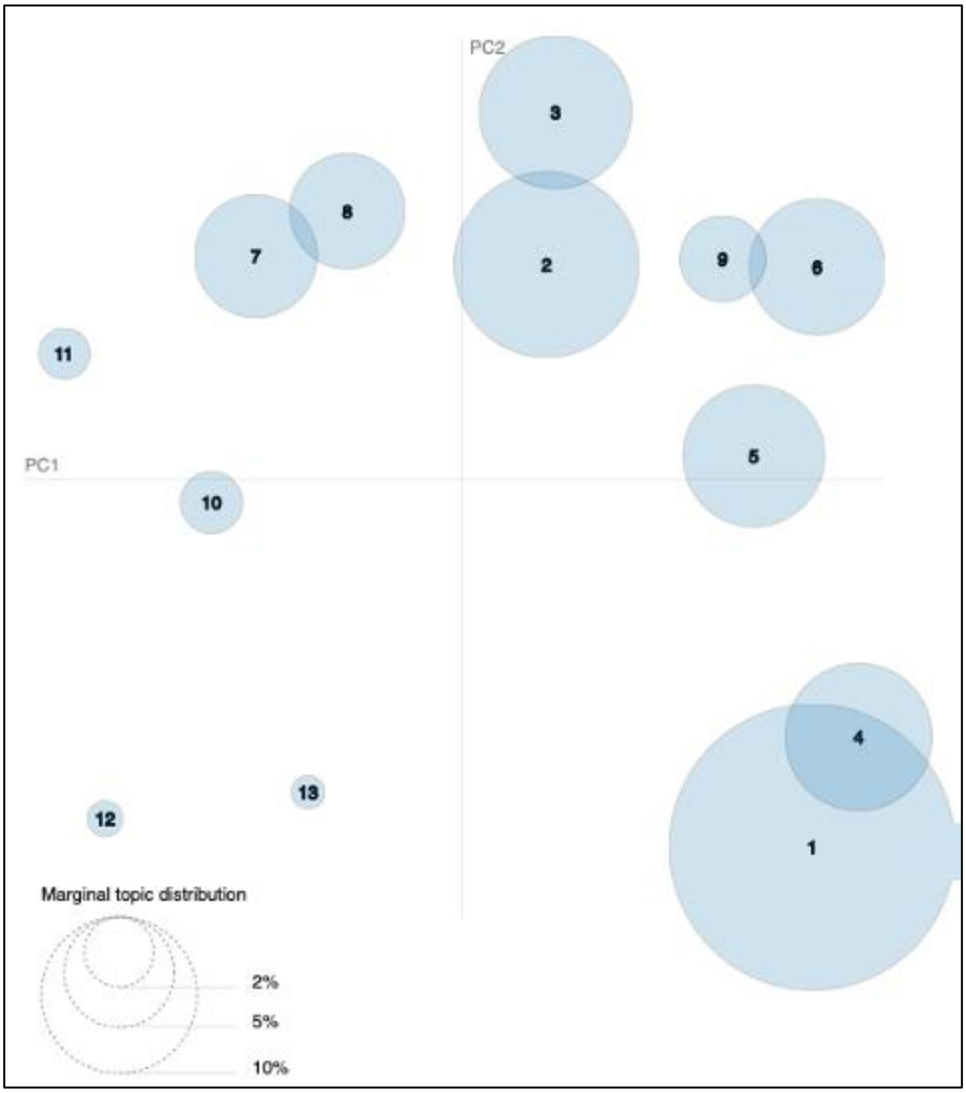


Figure 45: Intertopic distance map - LDA: Deterrence-IS-Russia/Russian

Table 13 displays the labels our team assigned to the LDA-topics.

Topic #	Topic interpretation	% of tokens
1	Post-Soviet countries I	33.6%
2	Deterrence strategy	14.1%
3	Military development	9.6%
4	Post-Soviet countries II	8.9%
5	World regions issues	8.2%
6	Russian internal policy development	7.5%
7	Nuclear deterrence	6.2%
8	Nuclear and military deterrence in offensive terms	5.4%
9	Other/Unidentifiable	3%
10	International treaties and collective security	1.6%
11	Military nuclear technology and operational activity	1.1%

12	East Europe and Asia security issues	0.5%
13	General terms on deterrence in Eurasia	0.4%

Table 13: Table 12 HCSS-assigned topic labels - LDA: Deterrence-IS-Russia/Russian

The next table (Table 14) shows this corpus' 30 most salient terms.

Top-30 Most salient terms		
Terms in Russian	**Terms' translation in English**	**Frequency of mentions**
Украина	Ukraine	20k
Военный	Military	12k
Российский	Russian	10k
Сша	USA	8k
Ядерный	Nuclear	7k
Сдерживание	Deterrence	7k
Президент	President	7k
Стратегический	Strategic	6k
Москва	Moscow	6k
Российский_федерация	Russian Federation	5k
Нато	NATO	5k
Отношение	Relation	5k
Белоруссия	Belarus	5k
Грузия	Georgia	5k
Безопасность	Security	4k
Государство	State	4k
Путин	Putin	4k
Система	System	4k
Цель	Goal	3k
Сотрудничество	Cooperation	3k
Европа	Europe	3k
Американский	American	3k
Международный	International	3k
Договор	Treaty	3k
Действие	Action	3k
Китай	China	2k
Запад	West	2k
Армия	Army	2k
Ядерный_оружие	Nuclear weapon	2k
Организация	Organization	2k

Table 14: Top-30 Most salient terms - LDA: Deterrence-IS-Russia/Russian

Regarding the generated topics we observe that certain regional topics that appeared in the general Russian corpus on deterrence are absent in its Russia-specific subset. For example, there are no distinct topics for the MENA or the Indo-Pacific region. This could be due to the general lower amount of topics we selected (13 vs 20) but the topic 'World regions issues' which unites both MENA, Indo-Pacific and Europe-related terms contains less percent of tokens (8.2%) than the sum of the topics 'Indo-Pacific region security', 'Russia/US-NATO deterrence in the Europe', 'MENA region security' in the whole Russian corpus (13.3%). The percent of tokens on 'Post-Soviet countries' topics in the precisely Russia-specific Russian corpus (42.5%) is higher than in the whole Russian corpus (21.7%). Generally, this highlights the continued prominence in Russian discussions about deterrence of the post-Soviet space vis-a-vis the 'far abroad' (including the US).

Top2vec

Top2Vec is a word embedding technique that illustrates a context-aware approach to topic modeling. Instead of using 'bag-of-words' feature representations like LDA does, top2vec relies on distributed representations of words and documents that take semantic similarity of words into account. To give an example: the latter "recognize[s] the similarity of words like big and large" while the former does not as those words "do not share a word stem"[171].

The model uses Gensim's Doc2Vec algorithm to generate a semantic space – jointly embedded document and word vectors that place similar (in a semantic and not a lexicographical sense) items 'close together in the embedding space'[172]. The main idea here comes from the distributional hypothesis stating that 'words with similar meanings are used in similar contexts'[173], and so should be placed closely together. But the Top2vec model also adds an *additional topic vector* that is 'calculated from dense areas of document vectors'[174], meaning that clustered documents outline a continuous representation of topics. To find clusters – and thus topics – the model utilizes UMAP (for dimensionality reduction) and the density-based clustering algorithm HDB-SCAN (used for finding dense clusters of documents in reduced dimensions)[175]. The model creates topic vectors by calculating their centroids, or 'arithmetic mean of all the document vectors in the same dense cluster'[176].

How we trained the model

We trained our model mostly using default hyperparameters[177], although we did make some changes. Firstly, the Top2Vec model was trained using both fast-learn and deep-learn training

[171] Dimo Angelov, "Top2Vec: Distributed Representations of Topics," *ArXiv:2008.09470 [Cs, Stat]*, August 19, 2020, http://arxiv.org/abs/2008.09470.
[172] Angelov.
[173] Angelov.
[174] Angelov.
[175] Angelov.
[176] Angelov.
[177] Angelov.

speeds (as opposed to the default 'learn' mode). The 'fast-learn' mode produced lower quality vectors but also took the least amount of time – around 5 hours (this might be a good option for those having hardware limitations). We finally opted for 'deep-learn' as its embeddings proved to be of better quality, at the cost of spending 12 hours to complete. We, secondly, also experimented with UMAP hyperparameters, reducing 'n_components' from 5 to 2 to produce 2D visualizations of embeddings. In terms of the corpus used, we trained our model on our preprocessed ProQuest Central deterrence-IS corpus containing 26,525 non-empty documents. We tested two different approaches – (a) using the same preprocessing steps as described in the 'Building full-text corpora' section, i.e. using raw text, and (b) feeding noun phrases to the model.

What we found

After initial training we ended up with 229 automatically generated topics (each unique document was assigned to a single, most similar topic). As noted in the paper, the initial number of topics 'can be hierarchically reduced to any number of topics' by 'merging the smallest topic into its most semantically similar topic'[178]. Therefore, we utilized this method to reduce the number of topics to 20.[179]

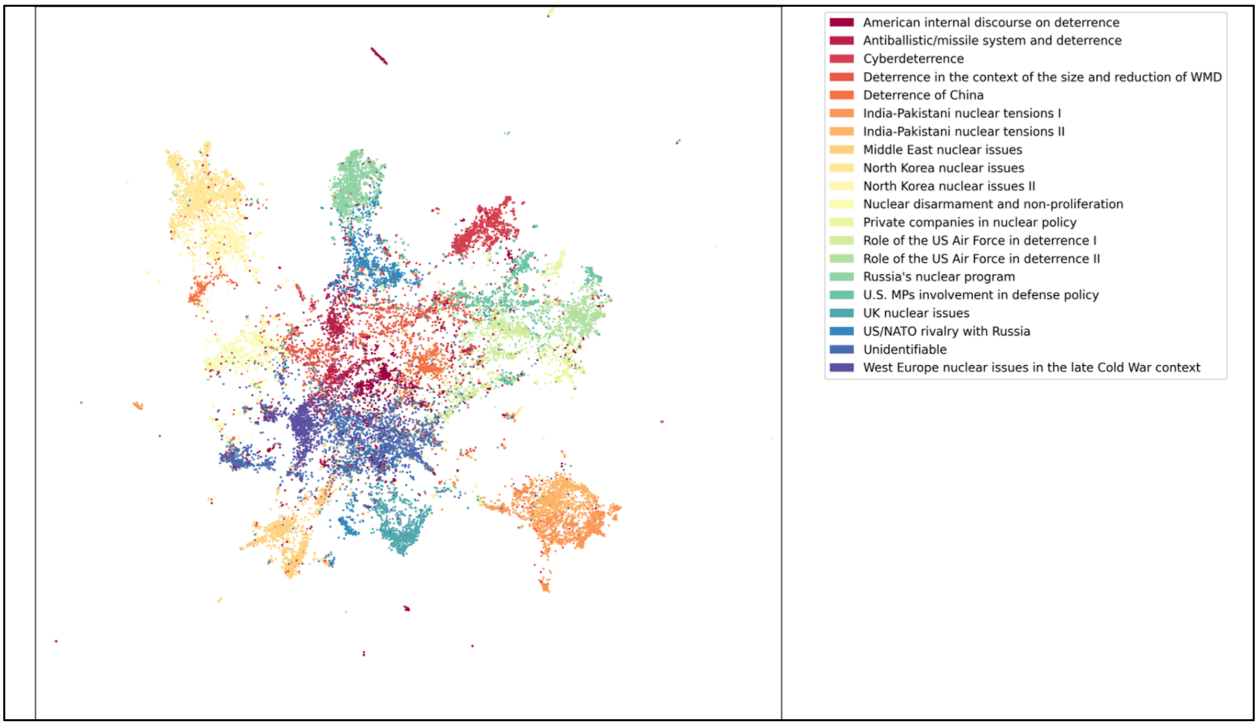


Figure 46: Top2Vec topic modelling - Deterrence-IS

[178]Angelov.
[179] To determine the optimal number of topics, we first plotted UMAP projections of the model with reduced topics (ranging from 2 to 20, with 5 most common words representing each topic), and then interpreted the reduced topics (top 25 documents of each topic for each iteration) to see which configuration made the most sense based on full texts – we also did this for noun phrases (based on this dashboard (see https://share.streamlit.io/hcss-utils/streamlit-topic-reduction/main/app/main.py) but ultimately decided to use raw text instead.

While discussing LDA results, we paid close attention to the most salient terms (like 'deterrence') that would define a topic ('terms' were our unit of observation). With the Top2Vec model we concentrate on both terms and documents making use of the model's built-in methods. The model provides us with attributes and methods that help define a specific topic from both – words and documents – perspectives.

Words	Scores
nato	0.64
russia	0.58
enlargement	0.57
moscow	0.53
europe	0.53
putin	0.50
ukraine	0.50
kremlin	0.49

Firstly, the model stores words and their vector scores for each topic as attributes, making it easier to understand what the topic is about. For example, *US/NATO rivalry with Russia* topic is best described by the following top words:

Secondly, the model is designed to search documents by topic, returning id, document score, and its text. For example, we could see which of the documents have the highest semantic similarity to a specific latent topic. The distribution of documents across topics is shown in Table 15 below.

Topic	Number of documents
Broad/Unidentifiable	2740
North Korea nuclear issues	1920
India-Pakistani nuclear tensions I	1583
Middle East nuclear issues	1490
Russia's nuclear program	1426
Deterrence in the context of the size and reduction of WMD	1376
US/NATO rivalry with Russia	1322
American internal discourse on deterrence	1303
West Europe nuclear issues in the late Cold War context	1275
North Korea nuclear issues II	1228
Nuclear disarmament and non-proliferation	1184
Role of the US Air Force in deterrence I	1165
India-Pakistani nuclear tensions II	1162
Cyberdeterrence	1159
UK nuclear issues	1098
U.S. MPs involvement in defense policy	1069
Antiballistic/missile system and deterrence	1066
Private companies in nuclear policy	1017
Deterrence of China	981
Role of the US Air Force in deterrence II	961

Table 15: Top2vec topics labelled by HCSS

So what?

The results we have shown align with what the author considers top2vec's main advantage over LDA: the model 'automatically finds the number of topics' and allows for topic reduction, and also understands word semantics which 'bag-of-words' approach ignores. It also simplifies the analysis as handles text preprocessing under the hood and does not require stopwords.

The display of the topics shows the relative distance of the topics from each other. For example, we can see that topic on 'U.S. and maintaining of the world order' is at the center of the map and borders with a lot of regional (MENA region security issues, Asia-Pacific region security issues, Transatlantic cooperation and NATO-Russia rivalry in Europe) and issue topics (Non-Proliferation, Anti-Ballistic missile defense). Also, the other similar topics that are connected to each other like topics 'U.S. Air Force' and 'U.S. Defense budgeting and policy-making' are overlapping. The general view of the map also shows that some topics are quite distanced from the core cluster of documents. There are three key clusters of documents that are partially separated from the other. 'India-Pakistani nuclear issues' topic in the bottom-right part of the map is separated from the general cluster which highlights the general uniqueness of the case as the rivalry between non-western actors in which the other great powers (like China, U.S./NATO, Russia, EU) are rather observers than active participants. The topic on 'cyberdeterrence' is also separated from the core cluster and located at the bottom-left part of the map, which highlights the distinctiveness of the cyber domain from the other issues and regional topics. Also, blotches of the 'cyberdeterrence' topic appear in the other regional topics, which imply the relative presence of this domain of competition in most international security rivalries. The third separated topic is the 'Korean peninsula security/nuclear issues' located at the top-right part of the map. The topic overlaps with the Asia-Pacific region security but still constitutes a distinct cluster from the other topics, which highlights the possible uniqueness and distinction of the DPRK deterrence from the other regional deterrence-related topics.

Interactive Learning – Identifying deterrence-relevant text spans

So far, the full-text NLP algorithms that we used were overwhelmingly 'machine-centric', even if they still required a significant amount of human sense-making – e.g. to select (in some cases) the optimal number of topics; or to give algorithmically discovered topics more meaningful labels that truly allow human researchers to understand what these topics are about. In this section, we turn our attention to another type of NLP in which humans play a more 'active' role in helping machines discover meaningful features in text.

The previous substantive analyses in this paper were based on corpora and text spans within them that were essentially discovered through human keyword-based search queries. Intelligently constructed human search queries can certainly generate useful bibliometric datasets and/or full-text corpora. In the latter, they can also help in identifying the main passages in these publications that an analyst/scholar should pay particular attention to. From a 'recall' perspective, however, such 'human filters' may still leave 'holes': relevant documents that may not have been discovered/retrieved; but also relevant excerpts within the retrieved documents that may be overlooked.

The basic intuition

Interactive machine learning[180] offers an approach that can help us overcome the so-called 'long tail' challenge[181]. Long tail problems stem from a growing number of academic publications. Since the number of publications are far beyond researchers' processing capabilities, only a small amount of publications are referenced, while some academic documents are not used at all. Essentially, we change the current paradigm from "keyword-based search" to "discover about a topic"[182] in which we loosely describe what we are looking for and then ask an algorithm to use already 'learned' lexical, grammatical and even semantic text features to find similar texts or parts of texts.

To do this, we start with a random sample of the full corpus at hand and evaluate whether its elements (in our case sentences) are about the topic we are interested in or not (in our case deterrence). After collecting a few dozen sentences identified as 'about deterrence' ('positive' sentences), the model has already learned which sentences are about deterrence and which are definitely not about deterrence. The model then starts to present human annotators with a series of sentences it is mostly uncertain about, asks them to classify them based on their subject-matter expertise ("Yes, this is about deterrence", "No, this is not about deterrence", or "Ignore") and then updates itself with the information the annotators provide. This way humans improve the model's performance and can eventually automate the process by making a full-fledged model that will be able to assess the relevance of any sentence from any corpus, based on the annotations that were made. For this purpose, we use very intuitive software, called Prodigy – a user-friendly interface that allows anyone without knowledge of NLP or advanced data science to make annotations. The model itself is calculated using the SpaCy software[183]. The final stage is to apply sentence-based models to the entire full-text corpus.

Our deterrence-specific application

To give a better understanding of the method and how to replicate it, we will accompany the description of our deterrence-specific application with some technical details.

---

[180] Maxime, "Introduction to Semi-Supervised Learning and Adversarial Training," Medium, May 24, 2019, https://medium.com/inside-machine-learning/placeholder-3557ebb3d470.
[181] Kate McKenzie, "What Is the Long Tail of Publishing? – Publishing Trendsetter," October 4, 2013, http://publishingtrendsetter.com/industryinsight/long-tail-publishing/.
[182]This paradigm is also used by the Amazon recommendation system "you may like".
[183]"Software · Explosion," Explosion, 2020, https://explosion.ai/software#spacy.

The final outcome of our efforts should be a model that can evaluate any corpus/document/sentence on its relevance to the deterrence topic. To get there, however, we first need to build a model for a topic-specific corpus that will allow us to tag enough 'positive' annotations in a relatively short time. In our case, we used our IS-Deterrence/English corpus with 47,999 documents (7,777 of them contained full text), built based on a keyword search. First of all, the corpus was cleaned – it had to be converted into continuous, uninterrupted text without tables, captions, endnotes, footnotes, in-line references, hyperlinks, and sub-sentences between brackets. After cleaning we split the text documents into sentences (as a result, the corpus consisted of some 520,000 sentences) and randomized these. This guarantees that the annotations are made from a representative part of the corpus. We used the first 50k sentences of the randomized corpus for annotations. We annotated a number of sentences with the label of "about deterrence" and "not about deterrence". The concept 'deterrence' was used quite loosely (although still always in an IR/IS context), and we treated any form, or degree, of deterrence as a 'positive'.

In total, 1800 annotations were made by domain experts using the Prodigy software[184]. The advantage of Prodigy is not only in its interface but also in its 'active learning' approach. After each annotation, the model re-evaluates its state and asks the user to make a decision on the sentences it is most uncertain about. This approach is much more effective than random annotations. The annotators face a binary choice yes/no[185], which allows for making very rapid annotations; evaluating a single sentence takes under 15 seconds. The total annotation effort was below 6 hours.

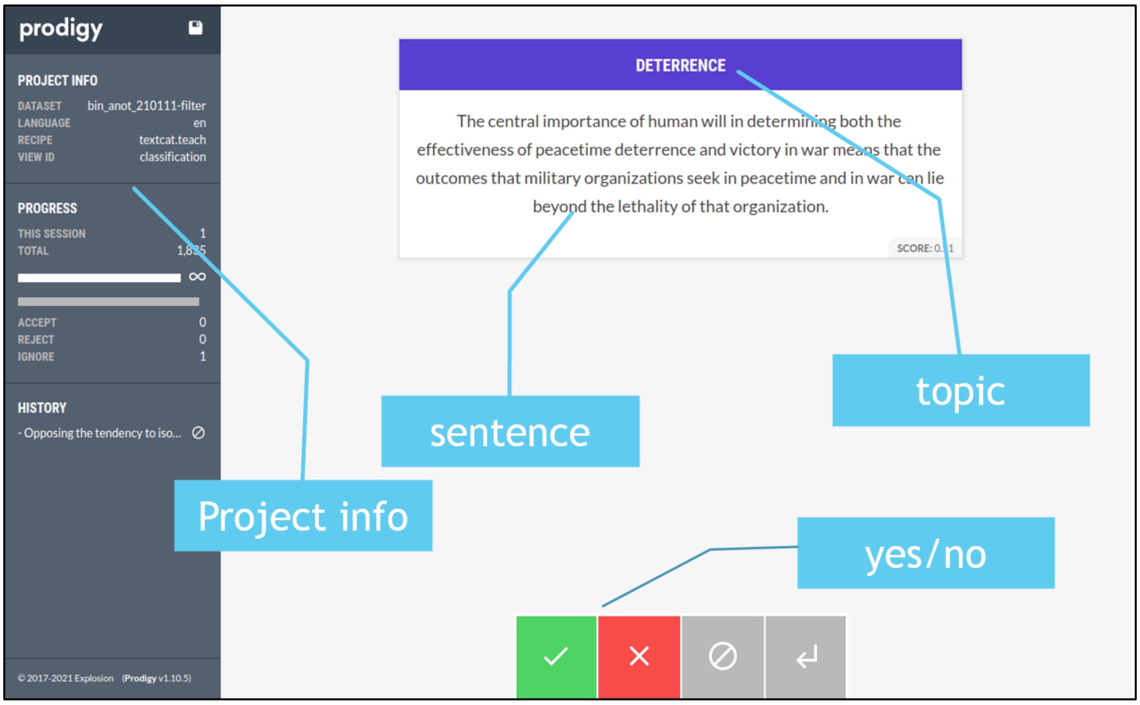


Figure 47: The Prodigy 'active learning' interface

[184]"Software · Explosion," Explosion, 2020, https://explosion.ai/software#prodigy.
[185]There is a third option – "Ignore". It is used for "junk sentences" – text in other wording, text that does not make sense, poorly cut parts of texts (e.g. only a few words that do not convey a real meaning), etc.

As we mentioned, after 'feeding'(/priming') the model some annotations, we can switch into the mode, where the model shows its evaluation of a given sentence ('score'). A low score (0-0.4) means that the model considers the sentence not to be about deterrence (the lower the score, the more 'certain' the model is about this), high score (0.6-1) means that the model thinks the sentence is on deterrence. Scores close to 0.5 mean that the model is undecided about how to evaluate a sentence.

An important aspect of this modeling approach is that it can provide users with evaluation metrics showing how well the model performs in its classification task. The main metric used for it is called the Area under the Curve (AuC). It is a measure of the quality of the model. An AuC of 1.0 means that the model always gives the correct answer; an AuC of 0.5 means the answer is a random 50/50 guess, and an AuC of 0.0 means the answer is always wrong. For example, our model, based on 526 "positive" annotations, has an AuC of 0.859, which means that the model should be accurate[186].

Finally, to get a good model, the annotations should be consistent. If several people are involved in the annotation process, they should agree beforehand on what they consider to be deterrence or not. Even if one person does all the work, she should treat similar cases the same way. Since the model is based on the statistical properties of all sentences, it is not critical to have no mistakes at all during the annotation process.

After we built a model, it was important to verify the model's performance. Two annotators assessed 40 random sentences, 10 for the ranges of 0.6-0.7, 0.7-0.8, 0.8-0.9, 0.9-1 without knowing anything about the degree of deterrence that the model assigned to these sentences. This blind review revealed that the human annotators agreed with the model in all instances with the highest degree while disagreeing in some instances, where the model was less certain. The downward trend suggests that the model was built correctly, making the safe bet that sentences with a degree >0.9 are most certainly about deterrence.

Degree range	0.60 – 0.70	0.7 – 0.8	0.8 – 0.9	0.9 – 1.0
# correct	3	1	5	10

The next step was to go from sentence-level to document-level, as this is the main unit of information in the corpus. In our effort, we used two criteria for this: (1) the length of the document and (2) the average 'degree of deterrence' of the top-N sentences. So if a document contained at least 10 sentences and the average 'degree of deterrence' of the top-5 sentences was above 0.9 (although the cutoff threshold could be changed), it was considered as being about deterrence[187].

[186]Overall, 500 "positive" annotations is a minimum number, required for building a decent binary classification model.
[187]To determine a reasonable cut-off we calculate the number of false-positives and false-negatives for each option. In this use case a false-positive is a sentence that we believe is about deterrence, but is not. Analyzing such a sentence we identify as

Degree cutoff	0.70	0.75	0.80	0.85	0.90
Percentage of documents kept	52.1%	45.4%	38.9%	31.6%	23.7%

As a result, we are now able to identify all documents in our corpus that are really relevant to the deterrence discourse and do not just mention the concept once or twice. At the same time, we are also able to identify all sentences – even in documents that are not particularly 'deterrence(/IS)-heavy' – that we may want to take a closer look at. From this point, we can start a real analysis without bothering ourselves with the many false positives keyword-based searches typically yield.

To put this in perspective – training such classification models based on large corpora and in ways that not only the evaluation algorithms but also most subject matter experts agree yields excellent results would represent a major leap towards 'gold corpora'. It would allow any researcher to collect higher-quality corpora, but also to identify the relevant excerpts within them: and not only the deterrence-relevant ones, but also any other ones that she might like to rain a model for. To give an example – our team is currently starting to set up a pipeline to train a model to identify sentences that deal with the nuclear threshold issue: whether Russia is indeed, as some scholars claim, prepared to resort to use nuclear weapons earlier in the escalation ladder. Once the 'escalate to de-escalate'-relevant sentences can be identified, the annotators can then also train additional models to identify which of those sentences can be read as "this sentence suggests that Russia may be lowering the nuclear threshold" or not. A move towards more collaborative modes of cumulative knowledge building might even see some leading scholars working together in this '(inter-)active learning mode.

In the current-day, scholar-centric model, debates about these issues do take place in that different researchers, often consecutively, present their arguments and counter-arguments based on their own (human) reading of the corpora they happen to have access to and/or are willing to consult – arguments that are then 'backed up' by text spans that are again hand-picked by them, that are interpreted based on their (unrevealed) priors and are often buried in selected cited references that force diligent readers to go look for the adduced piece of evidence in the cited source, sometimes even without the benefit of a page number. More collaborative, corpus-centric modes of knowledge elicitation would instead start from a much more complete corpus, could then call in the assistance of both the various machine-based NLP-tools *and* of human experts working in this field to build more validated findings that can be traced back to the actual text-spans that would have been transparently coded one way or another through human-machine interaction.

positive but is not about deterrence is wasted effort; irritating, but not a serious problem. More problematic are false-negatives, sentences that discuss deterrence but are not identified as such. The number of false-negatives increases with increased cutoff range. The # lines indicates considered to be about deterrence decreases with increasing cutoff value.

Using the results for navigating corpora

The binary classification model we trained on a training set and then applied to the full corpus can also be used for other purposes, one of which is enable scholars to explore the full-text content of a corpus in novel ways.

A simple wordcount can sometimes be surprisingly instructive. Once sentences in a corpus are categorized into two dimensions (deterrence/not deterrence) one can do a simple word count of how often a word occurs in one or the other category. This yields characteristic words for each class (scattertext – the equivalent of what the popular scatterplot is for numeric data – for 'text as data').

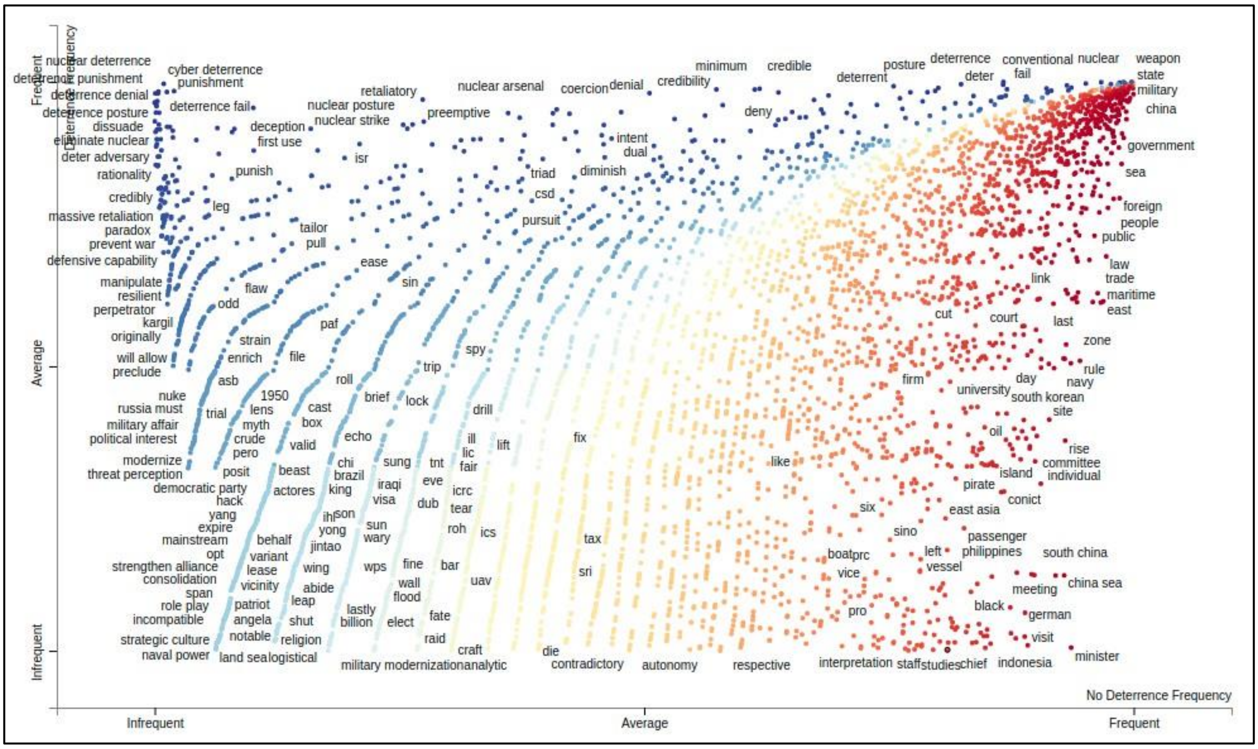


Figure 48: Scattertext visualization - Deterrence-IS

On Figure 48 blue dots represent words that occur in the class "deterrence", red ones – words in the class "non-deterrence". Color intensity is a measure for occurrence in each class. Words in the top left are characteristic for deterrence, but rare in the non-deterrence class. The opposite is true for words in the lower-right corner. Words on the diagonal are "neutral" words, not characteristic for either class. The scatterplot has an interactive version in which, by hovering over a point, the word is shown. By selecting the word one navigates to a table where the usage of the word in context is shown. For example, the word "counterforce" is mostly used in the context of "counterforce strike" and "counterforce targeting". As we can see, the words mostly characteristic for the 'pure' deterrence discourse (the top left) are different types of deterrence – nuclear deterrence, cyber deterrence, deterrence by denial/by punishment. One also can identify concepts widely used in the explanation of the phenomenon – deterrence fail(ure), rationality, retaliation, credibility, preemptive (strike).

Let us identify a few initial[188] observations to emerge from this scattertext. Most deterrence types are found quite away from "strategic culture" and "threat perception", for instance. To be more specific, they are situated at the diametrically opposite poles within the frequency of usage. This indicates that deterrence seems to be considered more via the prism of (neo)realism than critical theories, particularly constructivism. The presence of verbs such as "manipulate" indicates the cognitive nature of deterrence, which is more manifested in mind games and confirms the strong psychological component of deterrence. Regarding the mode of deterrence, the discourse appears skewed towards a defensive approach, which correlates with the cognitive nature of deterrence. Numerous mentions of conventional deterrence point out the importance of hard security issues such as military-related threats. The frequency of cyber deterrence helps realize that there is much more to security than the physical domain. Approximately the same level of concentration on the aforementioned elements demonstrates that the deterrence discourse is pinned around a deep understanding of both hard and soft security issues as well as reasons underpinning threats and means of their elimination.

Manual NLP – Human coding of 'points of interest' in corpora

The final 'NLP' section of this paper leverages purely human intelligence to try and 'map' elements of knowledge about deterrence in an international security context. We already mentioned that much of the deterrence literature is far more conceptual (read: scholastic/exegetical) than evidence-based. In many ways, that remains very much the name of the game today. Researchers ingest/digest/egest the epistemic nuggets their – constrained (as cognitive neuroscientists, as we have seen, tell us) – brains 'mine' from the extant literature based on their arguably rich but ultimately inevitably primed[189], non-traceable and biased priors.

In this final section of the paper, we will present some findings from our attempt to systematically 'code' some texts on Deterrence-IS-Russia. Before we delve into this, we want to recall our readers that every single scholar 'codes' the literature she reads and absorbs. Most do so implicitly. This section tries to do so explicitly.

Definitions of deterrence

There seems to be an implicit assumption in the deterrence literature that most scholars (and

[188]It is visually impossible to accommodate all terms in the scattertext visualization. One would therefore have to dig deeper into the matrix to obtain a more complete overview.
[189]'Priming' occurs when A's exposure to certain previous stimuli influences her response to subsequent stimuli in ways that she is not even cognizant of.

decision-makers) share a common understanding of what 'deterrence' actually means. Our own previous[190] research suggests that this may indeed be the case at some very basic level. In general, writers of deterrence agree that the goal of deterrence is to make another actor ***not*** pursue a certain course of action. Most deterrence theorists also put the concept of 'fear' (of retaliatory actions, damage, high costs, etc.) front and center in their definitions of deterrence. Thomas Schelling talked about "to turn aside or discourage through fear; hence to prevent action by fear of consequences[191]"; Paul Huth and Bruce Russett defined deterrence as "dissuasion by means of threat.[192]" From this shared low-level understanding, however, different schools of thought, different professional communities. different time-periods but also different countries take that basic shared concept in quite distinct directions. As Dima Adamsky has put it, for instance: "Emerging in a specific cultural context, deterrence conceptualisation is not universal and varies across strategic communities.[193]" We would add that similar contextual differences can be found not only across national strategic communities but also within them[194]. We therefore decided to take a closer look at these definitional differences by systematically mapping the various definitional nuances in our full-text Russian corpora.

What we used

The results of the manual coding that we present in this paper are based on a part of the Russian corpus, collected during an earlier iteration – Eastview Press until 12/11/2019, Russian military books, Militera, and 'seminal' papers (6,639 docs overall). To enable the 'slicing and dicing' of the publications in our corpus through manual coding we added metadata for all the documents that were coded. In particular, we added the gender of the authors (male, female, male/female), the type of publication (civilian, military), and the authors' background (civilian, military, civilian/military). We considered a publication to be military if the military sphere is its main focus (e.g. Военно-промышленный курьер) or if it is owned by some military entity (e.g. Военная мысль). An author was considered to have a military background if s/he had a military rank of officer or higher. If a publication was authored by both military and civilian authors, then the authors were considered to have civilian/military backgrounds.

What we did

To gain a better understanding of definitional differences in this corpus, our team systematically

[190]As yet unpublished but available upon request.
[191]Schelling, Arms and Influence, 71.
[192]Paul Huth and Bruce Russett, "Deterrence Failure and Crisis Escalation," *International Studies Quarterly* 32, no. 1 (March 1988): 30, https://doi.org/10.2307/2600411.
[193]Adamsky, "From Moscow with Coercion: Russian Deterrence Theory and Strategic Culture," 56.
[194]For a classic example of such strategic 'cultural' differences even within just the US military , see Carl H. Builder, *The Masks of War: American Military Styles in Strategy and Analysis*, A Rand Corporation Research Study (Baltimore: Johns Hopkins University Press, 1989).

scanned part of our Deterrence-IS/Russian corpora[195] to first identify and then 'dissect' the definitions that were offered in them. The guiding intuition behind this dissection effort is that every definition consists of a number of conceptual 'building blocks'[196]. Imagine, if you will, all of these possible definitional building blocks binned in a number of categories: one containing the actors ***who*** are doing the deterring , one with the means being used (***with what***), one with the timing of the deterrent actions (***when***), etc. Scholars formulating a definition of deterrence could then be thought of as picking and choosing some of these building blocks from different bins and putting them together in a specific mix to arrive at their particular definition of deterrence. Some, for instance, may specify the means that 'deterrence' uses, with some restricting it to military (or physical) means and others just opening the concept up much more widely (like the recently popular 'cross-domain' deterrence). But some may also decide not to specify any means. Our dissection effort endeavored to 'reverse-engineer' these decisions by authors by clinically dissecting these definitions into their constituent conceptual components.

The definitional 'dissection' of our corpora was conducted inductively and in two iterations. Throughout the first iteration, the relevant text-spans containing definitional 'building blocks' were identified by our team, highlighted, and labeled as higher-level codes. The latter were also already sorted into the top-level conceptual 'bins' which were formulated in advance as 'wh*'-questions for simplicity. During the second iteration, the coding team analyzed these definition-relevant text-spans to identify the scope of constitutive elements of the 'building blocks', translate them into meaningful lower-level codes, and then apply those labels to the corresponding parts of the text-spans selected during the first iteration. As we built up and then applied and refined the coding scheme, every code was first discussed and documented in our coding scheme: what that code really means, coding examples, coding notes, etc.[197]

The following mind map shows the entire treemap of the definitional categories (/'bins') we used with the various components (/'building blocks') they contain.

[195]The corpus comprised 6639 items overall. 5 501 items were downloaded from the EVP military and academic databases on the basis of the query *(сдерживан* OR устрашен* OR запугиван* OR шантаж*) AND (ядерн* OR атомн* OR кримин* OR преступ* OR оруж* OR вооруж* OR воен*)* within the timeframe from 1 January 2010 to 12 November 2019. The rest of the items were collected from the militera.lib.ru database as well as several collections of Russian/Soviet military books (1 026 and 76 items respectively), supplemented by 36 seminal works in the field.

[196]For another illustration of this approach, see Stephan De Spiegeleire, Peter Wijninga, and Tim Sweijs, *Designing Future Stabilization Efforts* (The Hague, The Netherlands: The Hague Centre for Strategic Studies, 2014), 7–10.

[197]The team is grateful for the repeated and extensive feedback and encouragement it received from Dr. Dima Adamsky on our coding scheme.

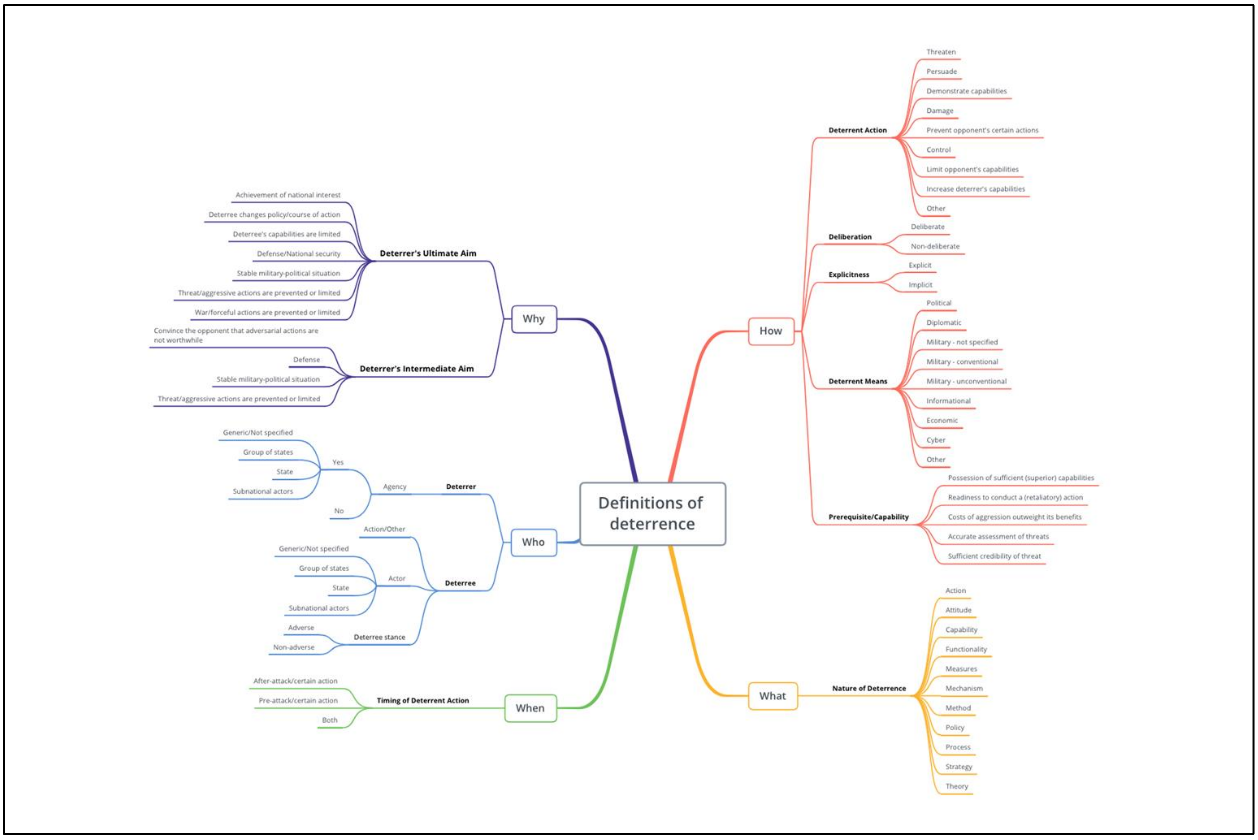


Figure 49: The RuBase human coding scheme for definitions of deterrence

The following breakout picture shows in some more detail the example of one of the branches ('bins') that deals with the different 'means' that could be part of a definition of deterrence.

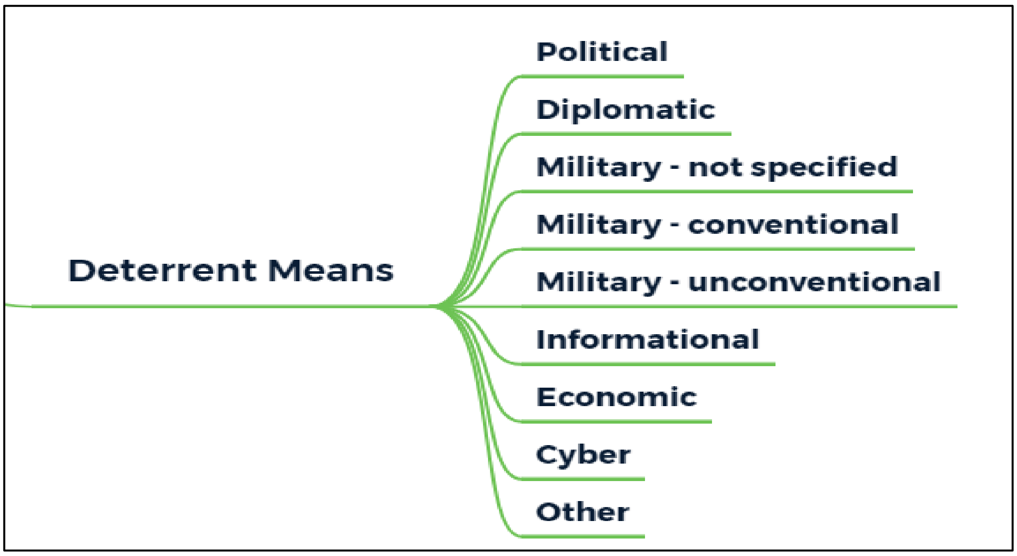


Figure 50: The RuBase human coding scheme for definitions of deterrence - a one-branch-example

Whenever our team of annotators would find one of these 'means' mentioned in a definition in any of the corpora's documents, the relevant text span would be highlighted and coded as such in a collaborative, computer-assisted, cloud-based, and qualitative data analysis[198] software program called atlas.ti[199].

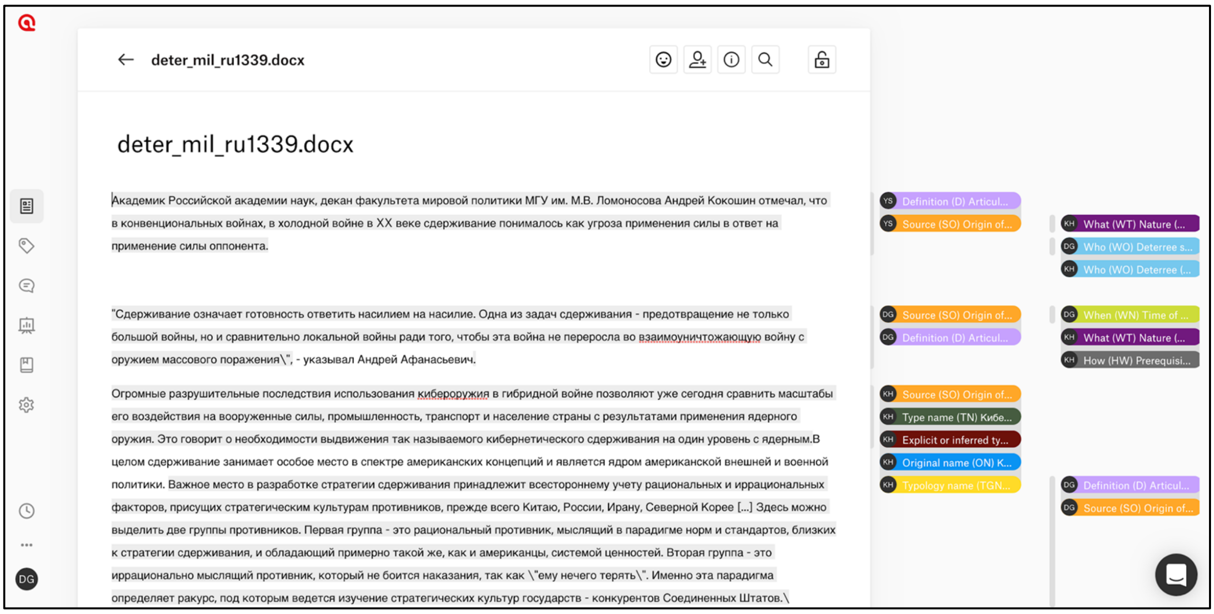


Figure 51: Coding example in atlas.ti

The results of this human-only coding process were subsequently visualized in a widely used 'business intelligence' software program, Tableau[200].

Additionally, our team also added a number of additional metadata to the bibliographical metadata our full-text databases had already provided us (like author, publication year, etc.). Among those we were mostly interested in the type of the publication (military or civilian)[201], the gender of the author(s), and the background (military or civilian) of the author(s)[202]. For each coded definition, we also captured whether the author(s) of a definitional excerpt presented their view or a translation of/quotation from a non-Russian source. The definitional excerpts were thus coded as "Russian" and "Non-Russian" accordingly.

---

[198]Susanne Friese, *Qualitative Data Analysis with Atlas.Ti* (London: Sage, 2012).
[199]Johnny Saldaña, *The Coding Manual for Qualitative Researchers*, 2nd ed (Los Angeles: SAGE, 2013).
[200]Tableau, "Tableau: Business Intelligence and Analytics Software," Tableau, 2021, https://www.tableau.com/.
[201]A publication was considered military if: 1) it was collected from the East View military database; 2) its publisher (or journal it was published in) belonged to the Russian Ministry of Defense or the Russian Armed Forces; 3) its publisher (or journal it was published in) only contained non-fictional publications on military topics. If none of these criteria were met, the publication was considered as civilian.
[202]An author was considered military if he/she: 1) had a military rank of at least an officer; 2) was a military journalist. If none of these criteria were met, an author's background was considered civilian.

What we found

An exhaustive analysis of all of our findings from this research strand would transcend the scope of this report. We, therefore, limit ourselves – as we also did in the other sections of this report- to a number of findings that struck us as being interesting and/or important.

Figure 52 provides a big-picture overview of which top-level (who, how, what, etc.) and second level codes (e.g. for who: anybody, only states, also non-state actors, etc.) were applied how many times in two different subsets of our data: the top pane shows the number of coded texts segments where Russian authors present their own or a/'the' Russian point of view, and the bottom one a translation/quotation of a non-Russian view. Even just a cursory glance at this visual already illustrates the broad definitional diversity in the Russian corpus.

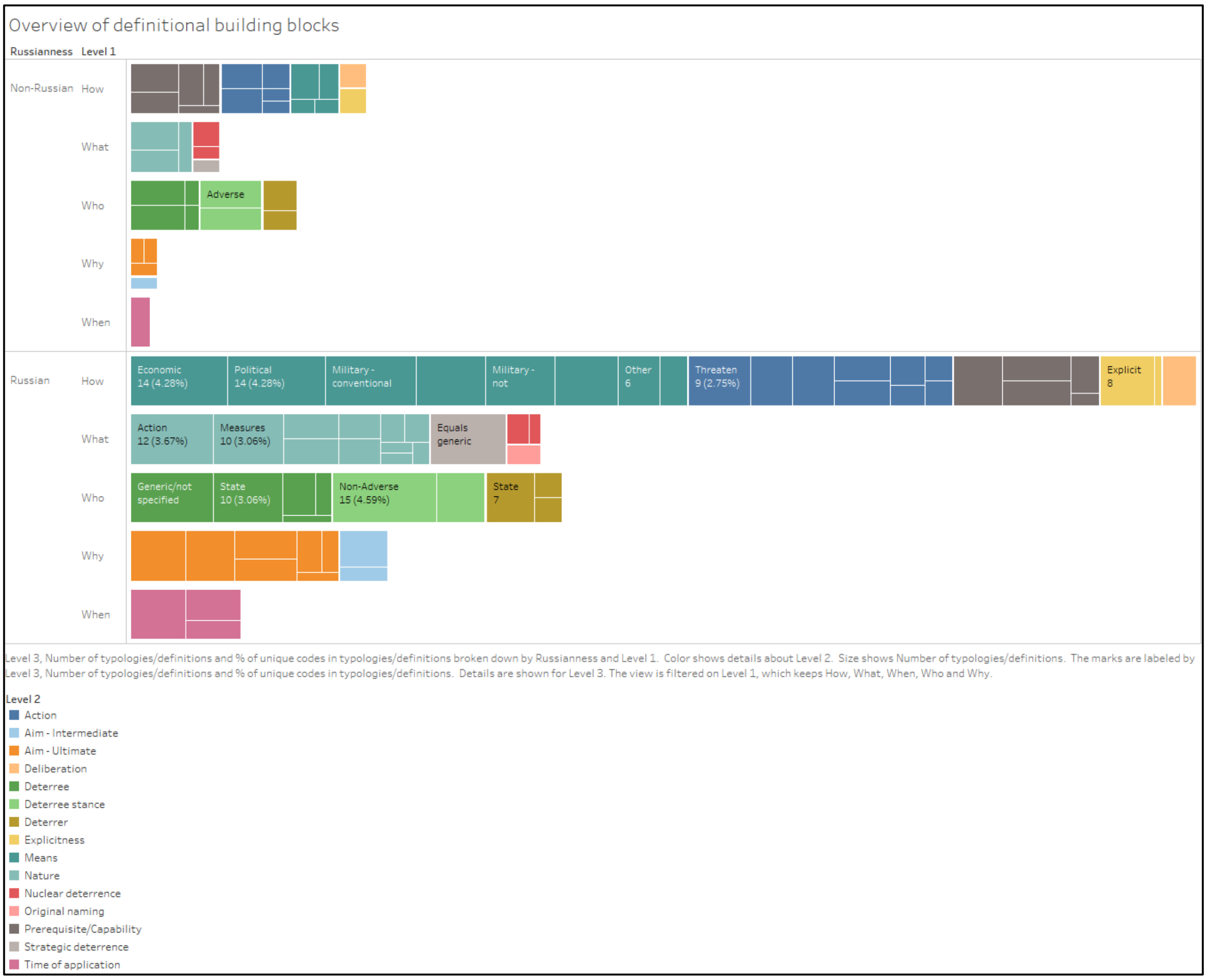


Figure 52: Definitional building blocks – overview

Within the Russian deterrence-definitional space, we observe a significant amount of disagreement on what deterrence actually is. The very nature of deterrence, as captured by the coding team, varies from a well-developed strategy or policy to simply a capability or an attitude. The most shared view is that deterrence is an action, followed with a visible margin by understanding deterrence as a set of measures. Whereas the former view is equally present across military and civilian publications, the latter is dominant only in the military sources. In civilian publications, the second most shared view on deterrence is that it is, in essence, a strategy. Understanding deterrence as a mechanism and functionality is seen only in Russian military publications, whereas deterrence is seen as a strategy, capability, method, or process exclusively in civilian sources.

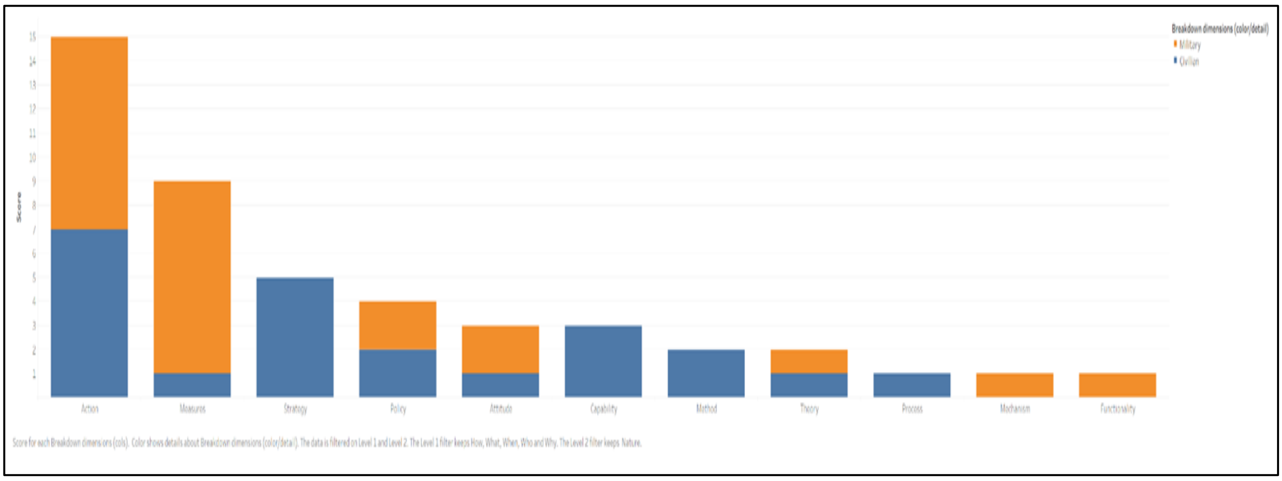


Figure 53: Definitional "what" building blocks - military vs civilian

This lack of consensus on the nature of deterrence also feeds into a significant amount of conceptual confusion regarding the ***operational dimension*** of deterrence and its aims, on the one hand, as well as the ***intermediate and ultimate aims*** of enacting deterrence, on the other. For instance, some authors consider limiting a deterree's capabilities as the deterrer's ultimate goal of deterrence, whereas for others, limiting a deterree's capabilities is just a step undertaken by a deterrer in order to achieve a more overarching goal like preventing or limiting a specific threat or war. Furthermore, some authors see preventing a specific threat or war as an intermediate step towards an even more overarching goal of deterrence, like, for instance, stabilizing the political-military situation either in the context of particular actors engaged in deterrence or on an international scale.

The Russian definitional landscape proves more homogeneous in how authors conceptualize deterrence in terms of means: the particular types of instruments that a deterrer uses to enact deterrence. Among the various categories of deterrent means identified – military, economic, political, informational, and diplomatic – the different military ones collectively receive substantially more attention than the other types, whereas the references to diplomatic means are least frequent. Further specification of military means shows that in defining deterrence, conventional military means appear more often than unconventional ones; however, there are also many instances when the specification of what constitutes military means is not given.

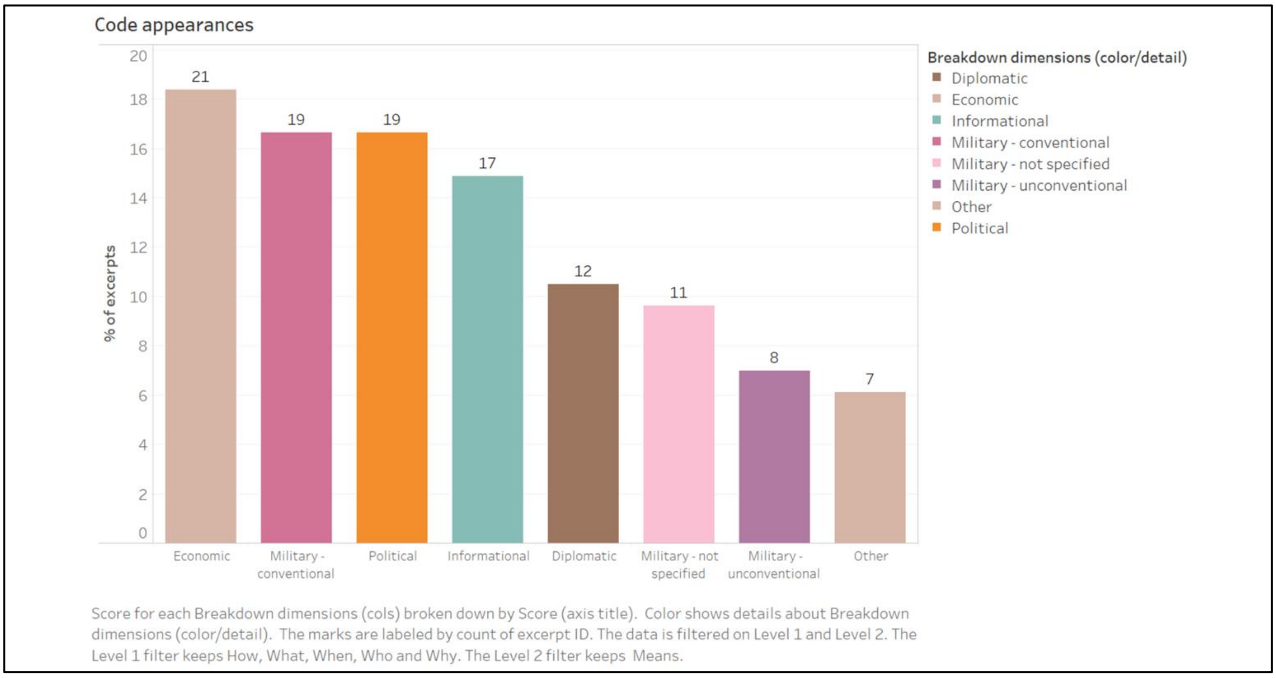


Figure 54: Definitional "with what?" (means) building blocks – overview

It is noteworthy that definitional building blocks of non-Russian origin appear only in reference to unconventional military means. Also, except for the nuclear, no explicit definitional mentions of other types of weapons of mass destruction were found in the corpus. Combining these two findings may point to a tendency among Russian authors to employ foreign-authored definitions when the nuclear component in the deterrence framework is explicitly mentioned in the definition. Another way of looking at this is that many Russian definitional discussions about deterrence may automatically assume that they are self-evidently about nuclear deterrence, which therefore does not have to be explicitly specified.

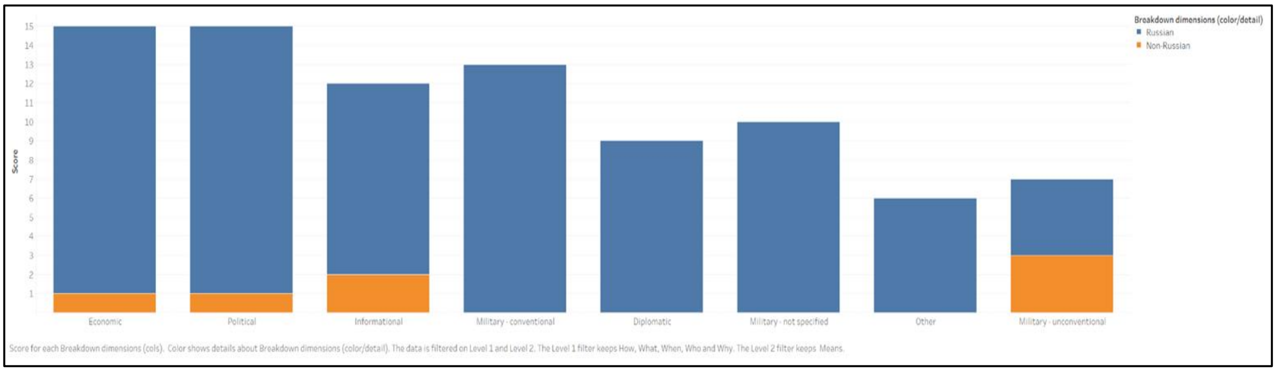


Figure 55: Definitional "with what?" (means) building blocks - military vs civilian

The Russian definitional landscape also offers a chequered picture with regard to the actors that are involved in deterrence. The role of deterrer is seen to be performed by either a state or a group of states, with the latter view expressed exclusively by military authors of Russian origin.

As to deterrees, in addition to a state or a group of states, subnational actors and even specific actions are mentioned as possible deterrence targets. We note that only authors with a military background see subnational actors as a target of deterrence.

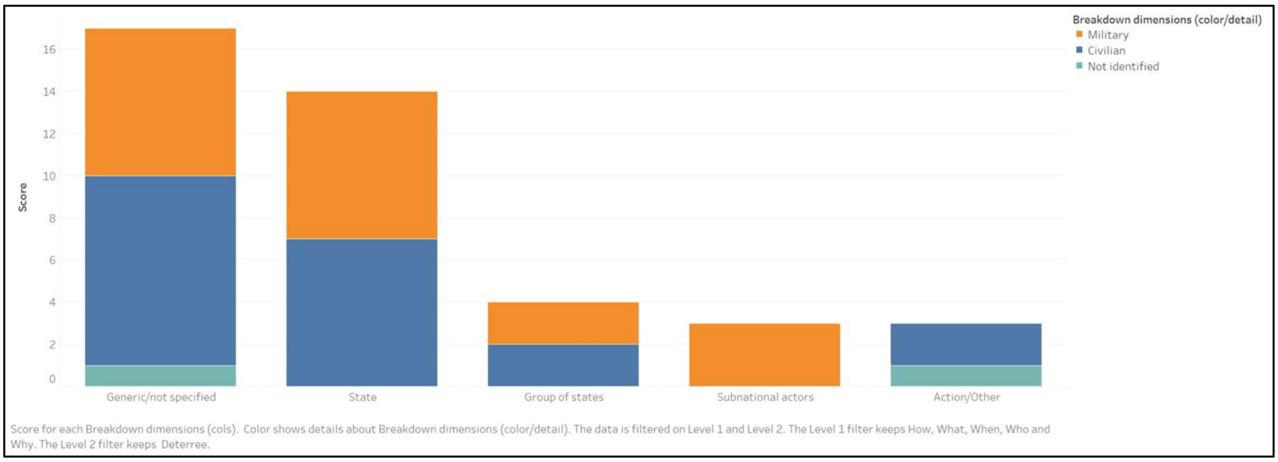


Figure 56: Definitional "at whom?" ('target) building blocks - military vs civilian

Besides the types of actors that can be deterred, the coding team also captured, depending on the semantic differences in the language used to describe a deterree, whether a deterree is framed as an adverse or non-adverse actor. For instance, when a deterree was described with the help of generic and neutral words like "target" or "state", the corresponding building block was labeled as "non-adverse"; when in a definition a deterree was framed as "aggressor" or "opponent" instead, the code "adverse" was assigned. As Figure 57 shows, deterring non-adverse actors is more widespread within the Russian definitional space. This finding can be interpreted in different ways, which will be one of the points of attention in our subsequent work, where we will have subject matter experts work with our dataset.

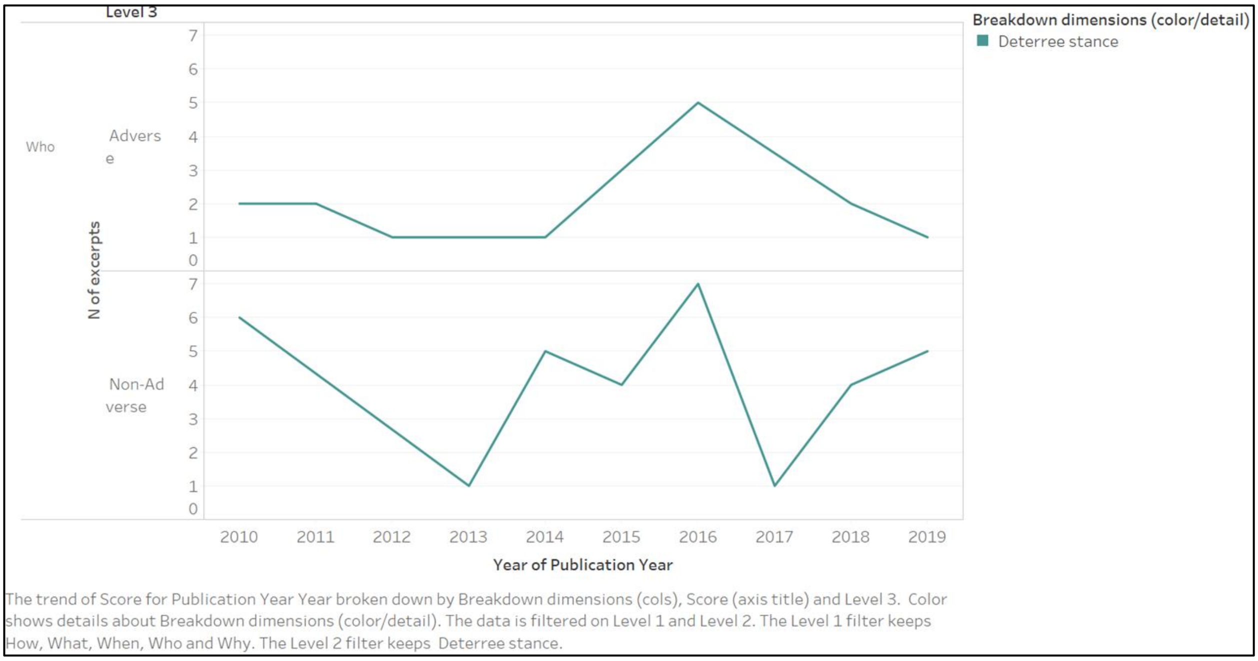


Figure 57: Definitional "at whom?" (target) building blocks – adversariness over time

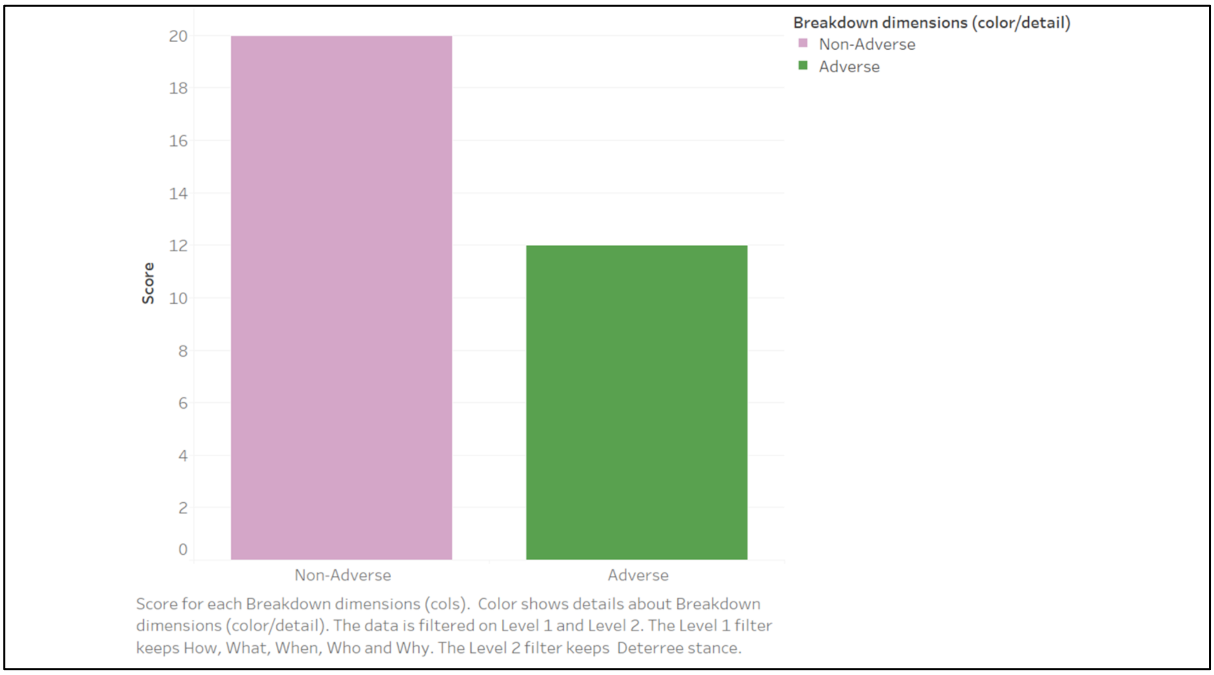


Figure 58: Definitional "at whom?" (target) building blocks – adversariness

This reading to some extent correlates with the finding that most authors see deterrence as preceding a certain action – not necessarily an aggressive one – which may be undertaken by a deterree.

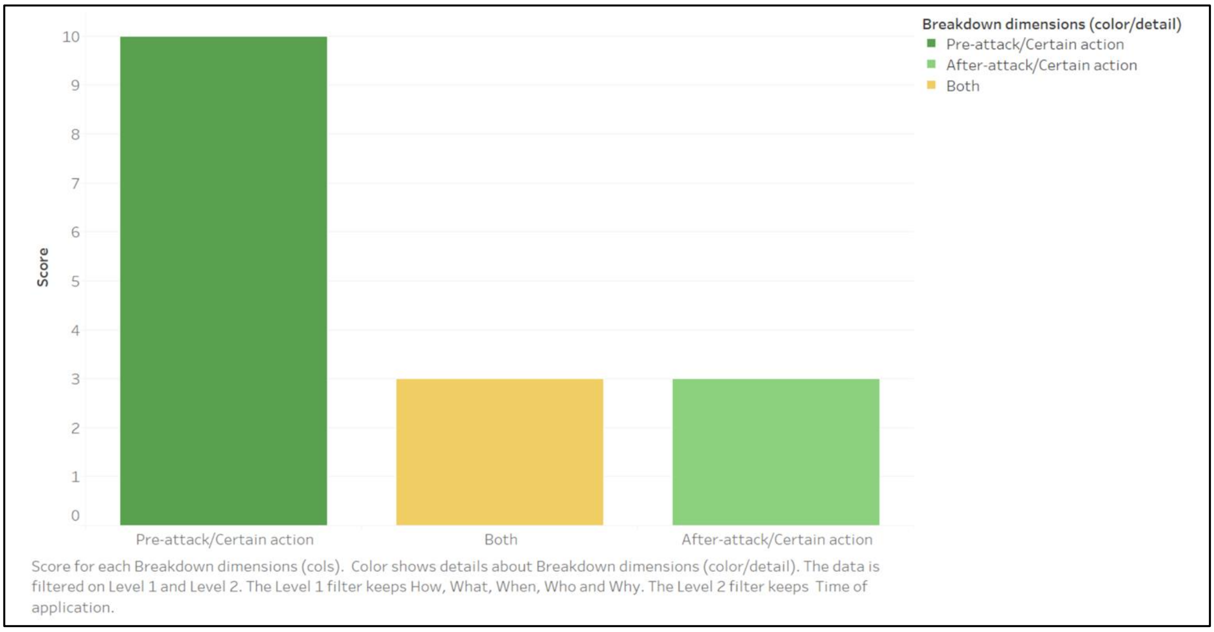


Figure 59: Definitional "when? " (target) building blocks – timing vis-a-vis adversary action/attack

Types of deterrence

Similar to the definitional diversity, the Russian corpus also contains a varied palette of deterrence ***types***.

What we did

During the first phase of corpus analysis, each mention of a deterrence type (with the explanation provided by the author, where available) was identified and, again with the help of the atlas.ti software program, coded as either an explicit or implicit part of a deterrence typology. By 'explicit' we mean instances, where the author him/herself indicated that deterrence can be categorized into several types, e.g. nuclear and conventional. On the contrary, a typology was coded as implicit, when the author did not indicate any particular set of types but mentioned several names that could be interpreted as types (cyberdeterrence, existential deterrence).

In the next phase, our team grouped these types either following the logic of already existing typologies found in the corpus (suggested by the authors themselves) or based on a particular shared characteristic (e.g. regional deterrence and global deterrence can be seen as members of one typology, given that both these types refer to the geographical scope of deterrence). Based on such characteristics each typology was given a unique name and every mention of a deterrence type was assigned to a corresponding typology. Given that in some cases, authors referred to essentially the same deterrence types using slightly different wording (e.g. "cyberdeterrence" and “deterrence in cyberspace”), the team treated such instances as a single type and coded it accordingly, while still preserving the original names alongside the additional code.

What we found

Our findings are based on 143 Russian documents and 662 English documents that spell out different types of deterrence. The timeframe of the majority of these documents encompasses publications from 2010 until 2019 including several publications from previous decades for the Russian part and from 1956 to 2020 for the English part.

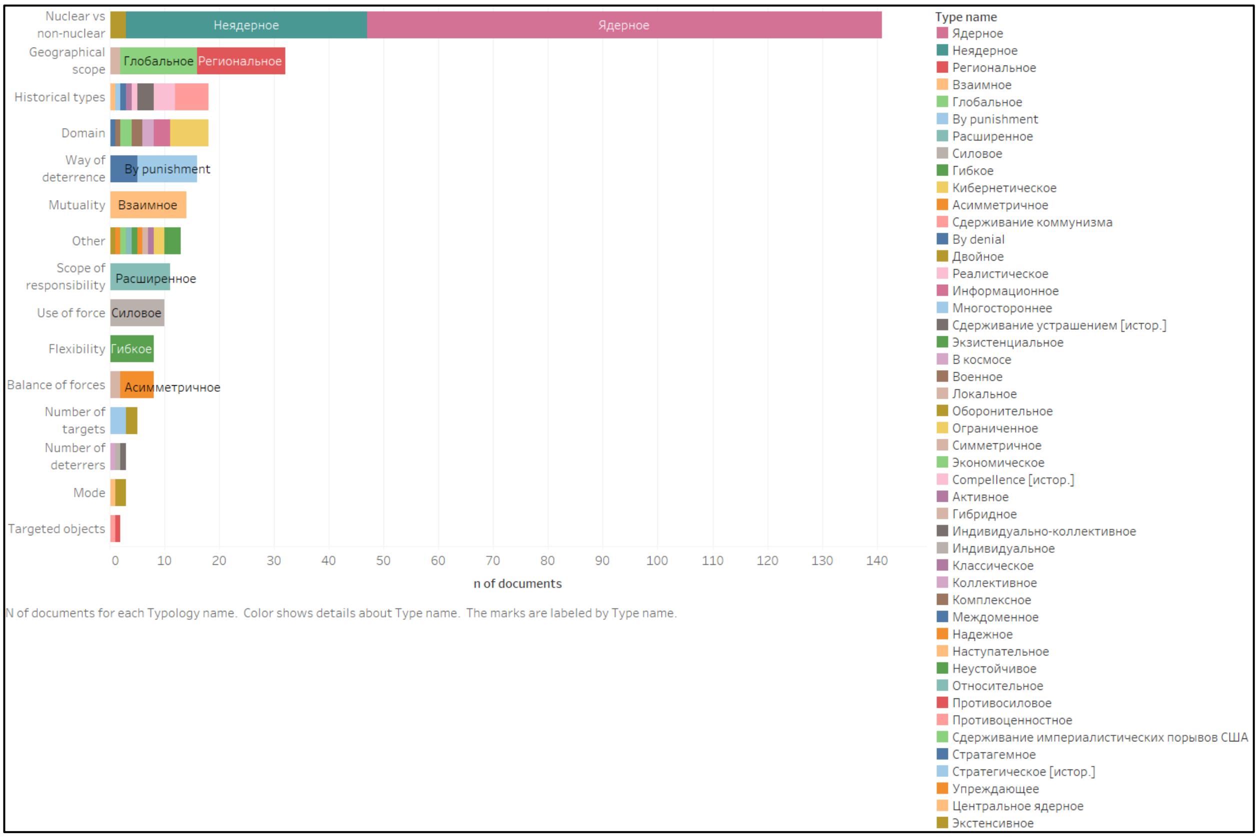


Figure 60: Russian deterrence types – overview

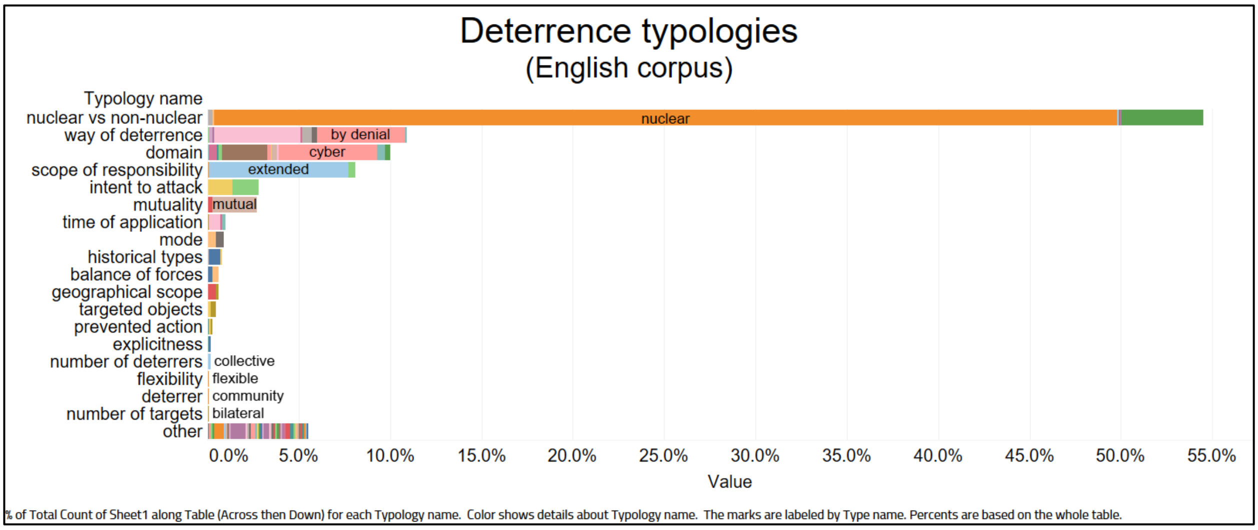


Figure 61: English deterrence typologies – overview

Figure 60 and Figure 61 show that the nuclear/non-nuclear typology dominates the deterrence discourse. We find more mentions of just nuclear deterrence than of any other type. This is true for both the Russian and the English corpus. We also identified a few instances in the Russian corpus of a type that involves both nuclear and non-nuclear deterrence – in every instance it goes under three different names: hybrid, dual and general. While the English corpus did not contain such a type, we did find a couple of interesting alternatives to nuclear and conventional deterrence – e.g. weaponless and technological deterrence. The latter usually is applied to Japan, which allegedly possesses technological capabilities to develop nuclear weapons swiftly, which serves as a deterrent to potential aggressors.

The classical Western dichotomy of deterrence by punishment vs by denial (which turned out to be the second most popular typology in our English corpus) also comes back in the Russian corpus but in smaller numbers. Our coding effort detected 5 instances of it under 4 different wordings (отрицание; лишение доступа; убеждение что атака бесполезна; отрицание доступности). Most of those are suggested by civilian authors all of whom come from the Russian Academy of Science. On the contrary, mentions of deterrence by punishment occur equally frequently in works by military authors as by civilian ones.

Relatively popular concepts in the Russian corpus are mutual deterrence, extended deterrence, and flexible deterrence. Although logically they could be expected to be accompanied by other isonyms (e.g. mutual vs unilateral, extended vs direct, flexible vs rigid), these 'counterparts' do not appear in this corpus. At the same time, the English corpus contains those isonyms for mutual (unilateral) and extended (basic, direct) deterrence. Besides, extended deterrence draws much more attention in English discourse than in the Russian one.

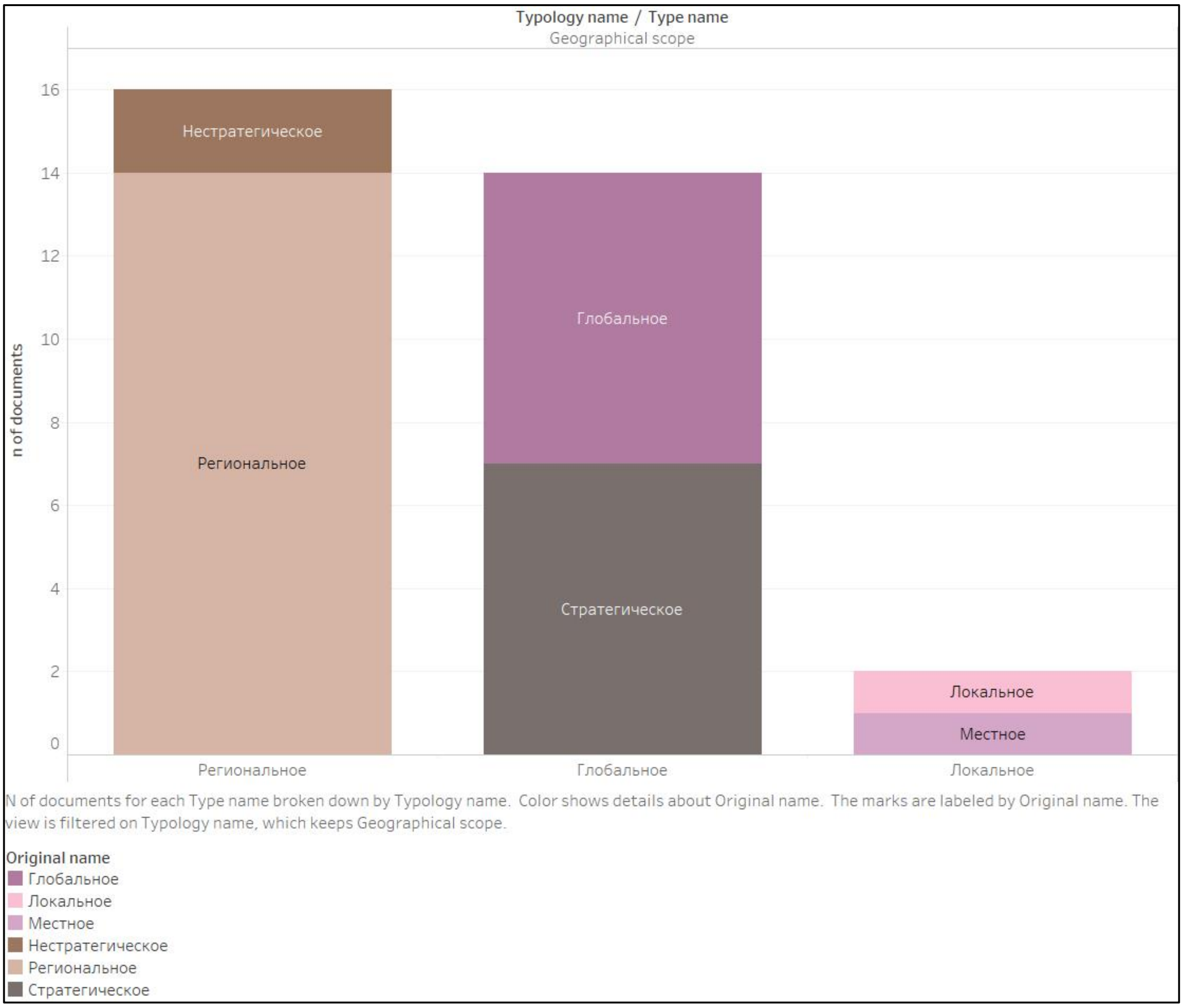


Figure 62: Russian deterrence types - geographical scope

Another popular typology in the Russian corpus pertains to deterrence's geographical scope. The authors break it down into not only global and regional, but also local deterrence. Also noticeable is the fact that global deterrence is defined equally often through two different words – global and strategic (which in other instances is often equal to the generic concept of deterrence). Unexpectedly, this dimension of deterrence is barely present in the English corpus.

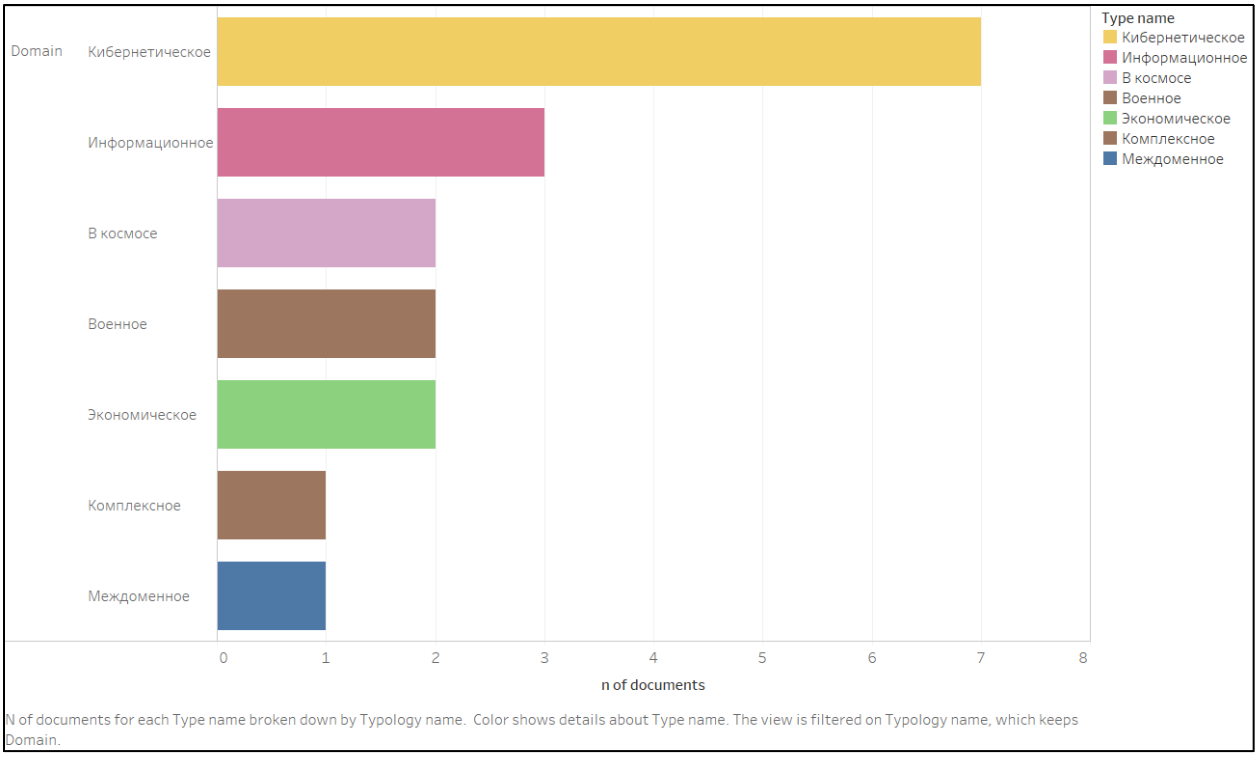


Figure 63: Russian deterrence types – domain

Often the types of deterrence are defined and differentiated by the means used for its purpose. As we explained above traditional military deterrence, defined through nuclear and conventional types (alongside a more generic military deterrence presented on this visualization) is still the most popular type across both corpora – because of its significance we analyzed it as a separate typology. Alternative domain types are much less present, but cyberdeterrence is the most conferred one among the rest, while informational, economic, and space deterrence lag behind in both corpora. So we may conclude that the focus of deterrence discourse still sticks to the realm of hard security, although cyberdeterrence is quickly climbing the ladder, especially in the English-language discussions of deterrence.

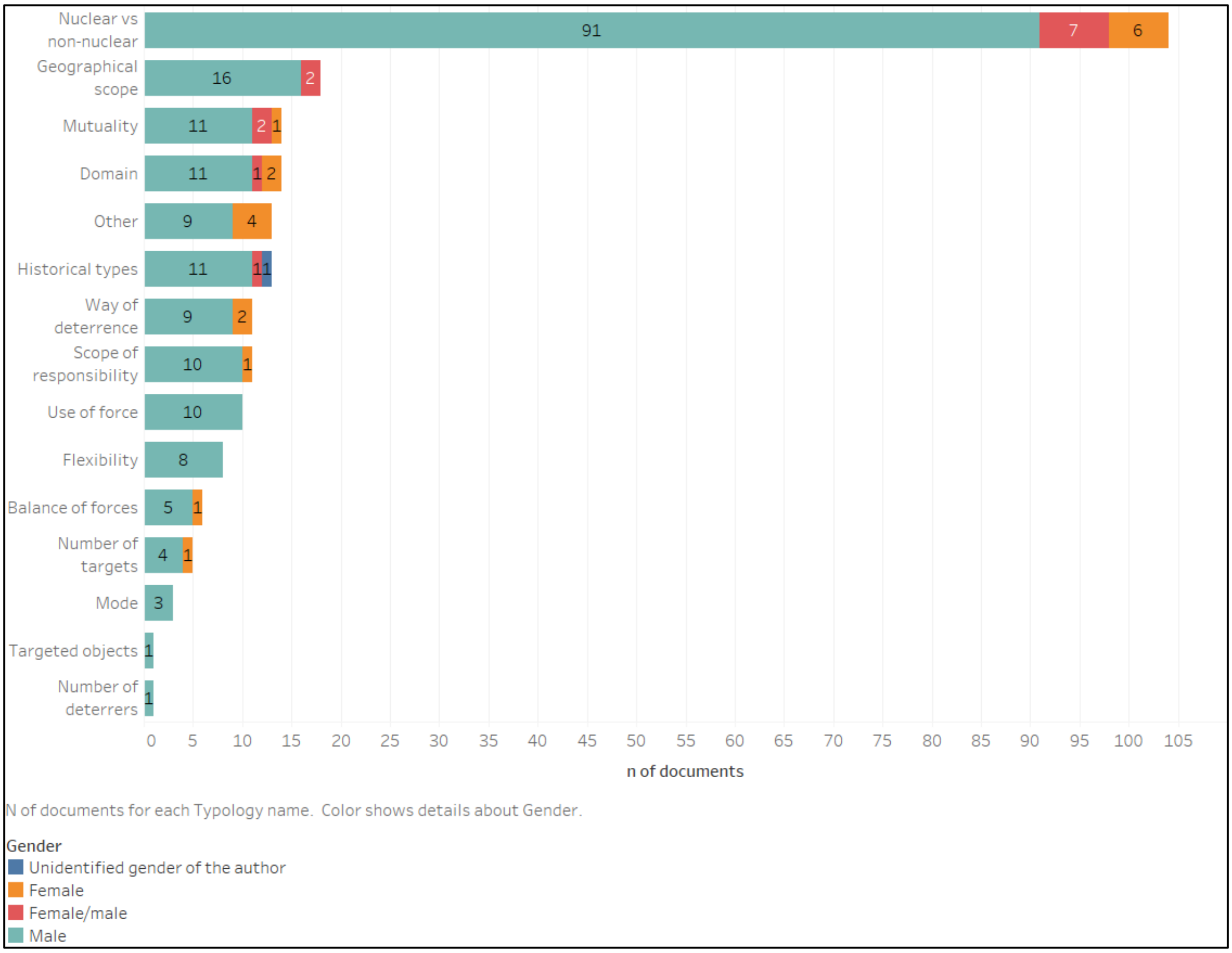


Figure 64: Russian deterrence typologies - overview by authors' gender

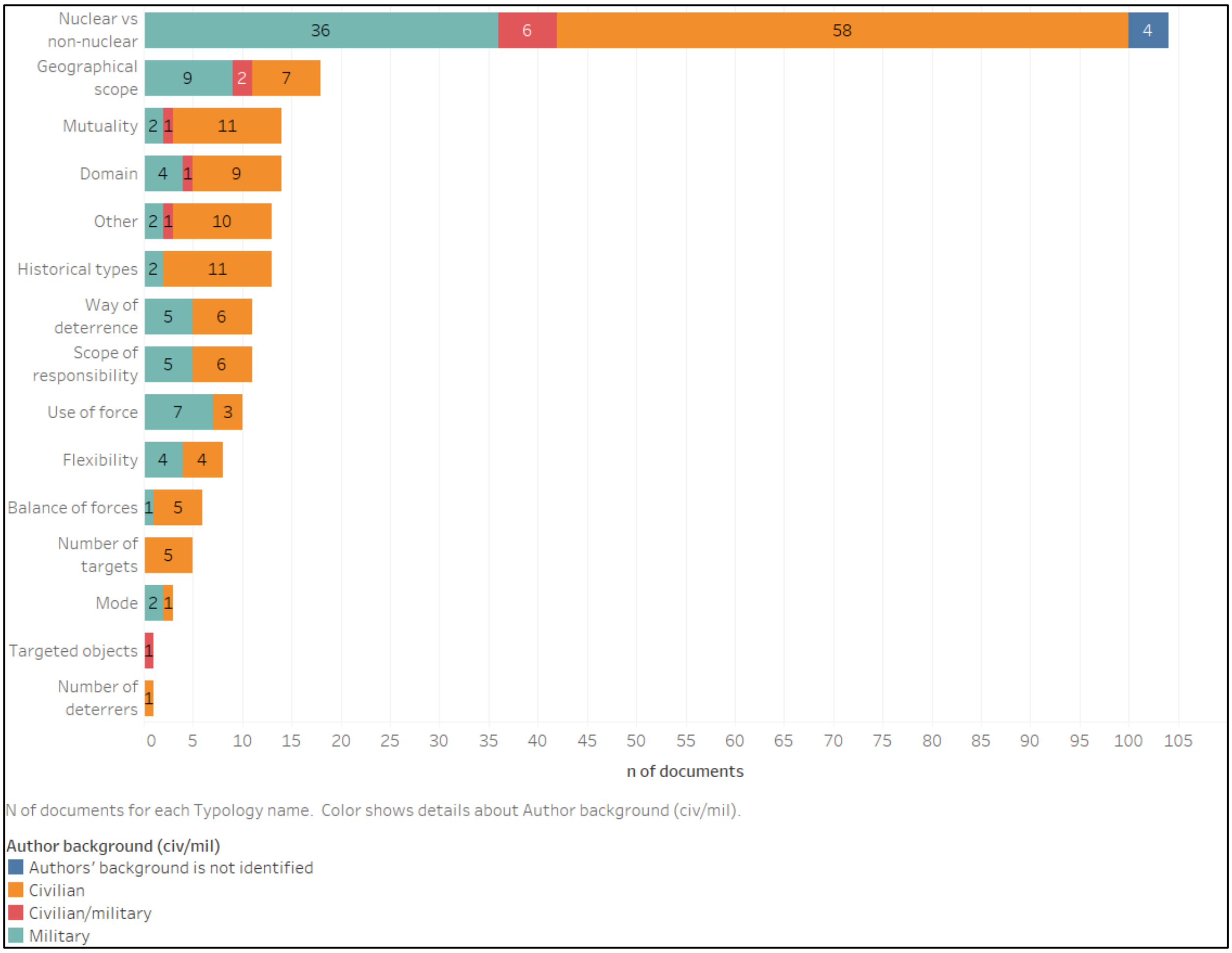


Figure 65: Russian deterrence typologies - overview by military vs civilian authors

The lion's share of documents across all years has been composed by male authors. Simultaneously, the majority of publications are authored by civilians, while authors with military backgrounds dedicate relatively more attention to the Nuclear vs non-nuclear typology, Geographical scope, and Use of force typologies.

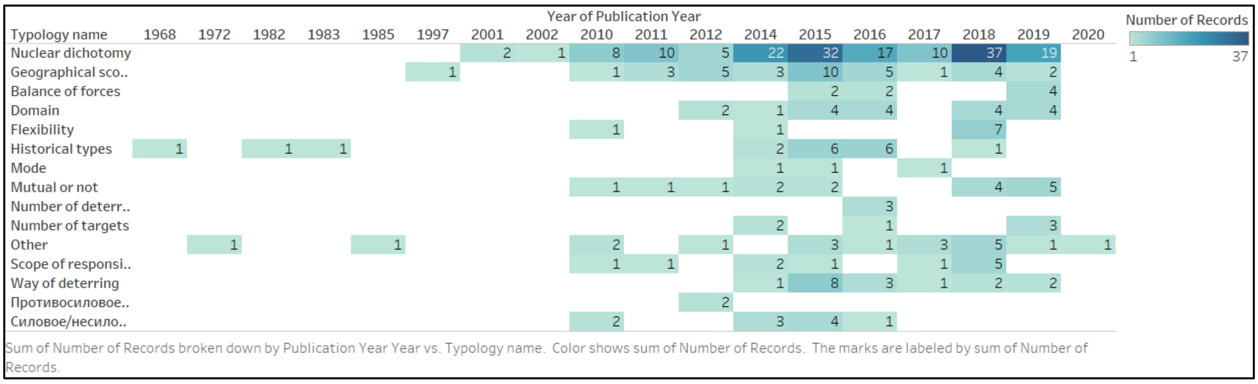


Figure 66: Russian deterrence typologies over time

Finally, the longitudinal analysis reveals that since 2014 the literature on definitions and types of deterrence has expanded significantly, while the dominance of the nuclear/non-nuclear discourse has rebounded firmly after a drop in 2012-2013.

Can NLP add something to our knowledge in this field?

On balance, the field of deterrence studies – and we would add: of most international security-related fields – remains surprisingly self-confident that the almost exclusively 'analogue' methods ('read-think-write') that still overwhelmingly dominate the field remain in and of themselves sufficient to generate reliable knowledge. This report raises some serious questions about this assumption. It also, however, showcases a few new datasets and -tools that we suspect may offer this epistemic community unprecedented new opportunities to raise its game. Center-front among all of these stands NLP – humans' ability to increasingly team up with machines to make sense of homo sapiens' epistemic patrimony that has been encoded in 'text'.

Critics of NLP often question some of the basic choices it forces users to make: the documents it processes ("garbage in, garbage out"); the various (classification, clustering, visualization, etc.) algorithms it applies ("how do you know these are the 'right' ones, etc."); the inevitable remaining human biases in the human sensemaking process that is still required to interpret what comes out of the machine algorithms ("why did you code that text span or interpret this topic like this and not like that"), etc.

We have always been mystified about how immeasurably more tolerant homo sapiens has so far proved to be towards the abundantly documented frailties in her own biological intelligence than towards the probably equally well-chronicled imperfections in the nascent sparks of artificial intelligence[203]. In the particular case of natural language processing, for instance, she typically downplays the (in our opinion) far more questionable analogous steps human analysts/scholars working in this field typically go through.

With respect to the ***ingestion*** of textual data, 'garbage-in, garbage-out' applies to human source selection just as much as to more algorithmic machine-based source selection. We have argued that the jury is out on whether humans really do trump machines on this. For the record – we suspect that still to be the case. But we also found that, for instance, Schelling's towering dominance in this field is also easily discovered by purely bibliometric means. We discovered that we can do much better on recall by allowing machines to help us in mapping much more of the scholarly epistemic landscape and in revealing some of its previously unknown characteristics like entropy, 'bursts', structural holes, etc. And we also emphasized that at least with machine-based approaches, the entire ingestion process – if properly documented – becomes transparent, which

[203]Whereby some of those AI imperfections are still primarily reflections of human biases that were encoded in the human-annotated training datasets that AI processed – e.g. the widely debated issue of AI 'being biased' in terms of race, gender, etc. "Debiasing humans is harder than debiasing AI systems." Will Knight, "AI Is Biased. Here's How Scientists Are Trying to Fix It," *Wired*, December 19, 2019, https://www.wired.com/story/ai-biased-how-scientists-trying-fix/ .

cannot be said about its human counterparts. Especially given the troubling 'recall' issues we have documented the Deterrence-IS field being demonstrably and particularly vulnerable to in the first main section of this paper, we would like to submit that NLP may very well embody this field's single biggest research promise to 'mine' the entire scholarly record in search of trustworthy and useful nuggets.

The same applies, we would submit, to the ***digestion*** part. There can be no doubt – as we have shown – that using different machine algorithms or even different parameters within the same machine algorithms (e.g. in topic modelling) on the same corpus can yield different results. For quite a few of these algorithms, even running the very same algorithm with the very same parameters on the very same corpus may yield (often slightly) different results just due to their stochastic nature. But how is that substantially different from the human algorithm that determines which relevant documents a scholar publishing in this field has (and has not) truly consulted – let alone 'absorbed' – when she produces a scholarly publication? How do we ascertain which relevant text spans she has (and has not) processed? How do we find out how exactly she has interpreted those text spans that she did process? As we pointed out already, the fact of the matter is that both humans and machines 'code'. The – in our opinion – single biggest difference is that machines code (mostly) transparently and systematically, whereas humans (mostly) do not. We would also argue that, at the current juncture in time, the 'dream team' for making (reliable, actionable) human sense of semantically rich textual corpora would see humans work alongside machines to mitigate each other's (well known) weaknesses and leverage their respective strengths.

When it comes to ***egestion***, we come to a similar conclusion. The overwhelming majority of the current (presumed) deterrence knowledge consists of raw text in articles, reports and books. These represent the 'output' of scholars' (mostly undocumented, untraceable, non-replicable) ruminations on the topic. The (human *and* 'hybrid' human-machine) algorithms used in this paper, on the other hand, yielded mathematically constructed visuals, in which human-written texts were first algorithmically processed and visualized, but were then still (re-)interpreted by human analysts. It was decidedly not our intention, to come up with the 'definitive' exposition of what 'we' know about deterrence. We also want to emphasize in no uncertain terms that we are not claiming that our findings would in any way be 'superior' to the extant human-elicited knowledge base. Instead, our primary ambition was to showcase some more 'digital' tools and techniques that we strongly suspect could add value to our currently mostly 'analogue' way of making sense of the scholarly body of literature. Our main hope is that we can nudge at least some of our colleagues to accompany us on this journey to find a better balance between 'analogue' and 'digital' (NLP) modes of analysis.

We hope to have illustrated some of the promise *and* the peril of this new more digital mode of analysis of (broader) text corpora. We once again want to emphasize that we consider especially our full-text NLP-findings very early beta-versions of what we hope – and plan – to grow into a more solid (and more collaborative) knowledge base. We have documented most of what we have learned in the process on the datasets and -tools as well as on the findings we obtained from them on our collaborative knowledge management platform Rizzoma. We would be more than happy to grant interested colleagues access to any and all of this documentation. We very much look forward to the input and feedback on the many 'bugs' that our beta-version undoubtedly still contains. But, strengthened in our resolve by our findings on the striking lack of collaboration in this field, we still decided to already share this early beta-version to move this process forward. Our own short-term future focus in this area will lie in what we have called the 'interactive learning' mode of NLP. We think this probably represents this community's single best hope to start building more comprehensive, more granular, ***and*** more high-fidelity datasets ('gold standard' corpora/'ground truth') about the various manifestations of deterrence in international relations – that we could then start 'learning' from. We would also submit, however, that these very same tools could also grant more 'traditional' (analogue) analysts/scholars more expansive access to relevant documents/text spans across the scholarly record that they could then work with in more analogue ways. Our RuBase team is, therefore, determined to leverage various forms of NLP much more across all of our international-security-related knowledge-building efforts, and to also do this in more collaboratively ways. We sincerely and warmly encourage anybody who is interested or even just intrigued about any of this to reach out.

5. CONCLUSION

Many of our nations currently continue to (even increasingly) wager significant blood and bullion on the assumption (typically positioned as a self-evident axioma) that 'deterrent' 'ways' of achieving one's strategic objectives ('ends') provide good value for money.

This study set out to conduct a clinical epistemic equivalent of an MRI-scan of what we actually know, based on currently publicly available scholarly knowledge, about 'deterrence' in an international security context. Our starting premise was that strategic decision-making should be based on (validated) knowledge. We availed ourselves of some of the more advanced datasets as well as software tools and algorithms that have come online over the past few years. These datasets and -tools increasingly allow us, humans, to also leverage the exponentially increasing aptitude of 'machines' (read: algorithms) to at least process – if not yet fully 'understand' – human natural language, the main vehicle through which we humans encode our knowledge about the world, including about deterrence. We explicitly (and atypically) paid special attention to this field's 'recall' predicament: have we really taken optimal advantage of all the available (narrow AND broad) sources of (inductive AND deductive) knowledge that can currently be accessed? Our findings are highly discomforting to us in a variety of different ways.

The single most pressing concern our research effort leaves us with pertains to the ***unbearable empirical lightness*** of the scholarly record on deterrence-IS. We discern that lightness in the absence of serious granular panel (cross-sectional *and* longitudinal) datasets on how deterrence works in real life, at the macro, meso and micro level – especially also compared to other 'ways' of achieving one's goals. We know that most – if not all – state- *and* non-state actors continue to pursue deterrence-based courses of actions on a daily, even hourly, basis and across strategic, operational and tactical levels and domains[204]. And yet international security is arguably the single deterrence-related discipline that has not deemed it necessary to carefully construct meticulous empirical datasets to explore the dynamics and the outcomes of those courses of action.

We also find the empirical lightness reflected in the perplexing paucity of more realistic modelling efforts of complex dynamic interactions involving deterrence. Also here this weakness is all the more glaring since the other 'layers' of (human) deterrence seem to have been quite a bit more diligent about (and successful in) these endeavors. But we also notice this lightness in the totally different topics that emerged from our purely academic datasets as opposed to the corpora that also blend in newspapers, magazines, military and official publications. We especially want to single out what we see as the irresponsible disregard for the real-life costs of deterrent 'ways' of achieving one's objectives – both in financial 'value for money' terms (an issue most scholarly articles do not even address), but also in terms of the broader second- or N^{th}-order effects of deterrent courses of action[205].

Our 'MRI-scan' therefore essentially reveals – with a few rare exceptions – a field that is quite 'heavy' on theoretical ('scholastic') debates about concepts, whilst being remarkably 'light' on diligent attempts to match these concepts with real-live actions and decisions (acquisition choices, targeting strategies, fueling options, deployment patterns, military exercises; but also the coercive use of threats in diplomacy, strategic balance of investment issues, etc.).

Our second most worrying conclusion is that the substantive evidence we gathered for this paper paints a bleak picture of various aspects of the state of our 'knowledge' about deterrence in general, and about Russia's deterrence-related thinking and acting in specific. We documented ample evidence of ***substantive confusion and –*** maybe even more strikingly – ***unknowns of various ilk***.

To start on a more positive note here, it is true that we suspect (although we have not (yet) been able to develop metrics to ascertain this systematically) that there is a broad conceptual agreement on the fact that:

- strategic deterrence belongs to the '**ways**' component in the (in)famous – and still surprisingly underexplored – 'ends, ways and means' triad;
- it is about **suasion:** one (or more) actors attempting to influence the future behavior of (an)other actor(s);
- it is not about random suasion, but about suasion with the particular aim of **making somebody NOT do something** (dissuade/deter); and that

[204] Tim Sweijs and Samo Zilincik, *Cross Domain Deterrence and Hybrid Conflict* (The Hague: The Hague Centre for Strategic Studies, 2019), https://hcss.nl/sites/default/files/files/reports/Cross%20Domain%20Deterrence%20-%20Final_0.pdf.

[205] We have not (yet) specifically trained an NLP model to explore this in detail, but we are currently doing that for (what we see as) another 'strategic function': influencing.

- that suasion is based on an element of **fear.**

Beyond these very basic elements of agreement, however, our analysis reveals a surprising amount of disagreement, epistemic holes and – maybe even more than anything – conceptual confusion and ignorance in the broader literature on deterrence in international security in general, but even more so in the literature on Russian deterrence in specific.

The only two topics that we really delved into with an extensive human coding effort (definitions and types of deterrence) revealed a conceptual cacophony. Much has been made about alleged differences *between* Russian and Western definitions of deterrence. Our analysis of the very 'building blocks' of the definition of deterrence confirms some (but not all) of those differences. But it also reveals much less appreciated and yet gaping cleavages *within* these two language domains – including within the Russian one.

There are myriad other key deterrence-related issues in the literature on (also Russian) deterrence that we did not (yet) analyze with the same intensity. Our main substantive finding here is probably that on most of these, we just do not 'know' because we have – at least in the public domain[206] – not yet done our homework. The debate on 'escalate to de-escalate' may serve as a nice example here[207]. It remains a highly politicized topic (especially, but not only) in the United States and one that thus deserves – in our opinion – researchers' utmost attention. The initial claim that certain Russian official statements were increasingly 'lowering the nuclear threshold' were based on a partial reading of the official Russian (mostly stated) record. But also some of the subsequent refutations of this claim were based on a non-exhaustive reading of the available ((textually) stated AND (empirically) revealed) evidence. We suggested that our more 'corpus/NLP-centric' approach may help us to parse the available multi-source evidence in a far more systematic way to obtain a more transparent, traceable and (hopefully) accurate reading of what we now (think we) know. And this 'escalate to de-escalate' example is (arguably) just one 'known unknown' that deserves more systematic 'forensic' investigation. There are – undoubtedly – many others, in all quadrants of the Rumsfeldian 2x2 matrix.

So ***how do we move forward on this***? The main premise of this paper (and of the underlying – and broader – RuBase effort) is not and never was to criticize or deplore the current state of the field. It was even less to blame scholars who have been contributing to the reservoir of knowledge

[206]And irrespective of whether this homework has been done (satisfactorily) in the classified realm or not, a good argument can be made that the realm in which we really want/need it, is the public one.

[207] For some exemplars of this debate, see Nikolai Sokov, "The New, 2010 Russian Military Doctrine: The Nuclear Angle," *James Martin Center for Nonproliferation Studies* 5 (2010); Stephen J. Blank, ed., *Russian Nuclear Weapons: Past, Present, and Future:* (Carlisle, PA: U.S. Army War College Strategic Studies Institute, 2011), https://doi.org/10.21236/ADA555143. Katarzyna Zysk, "Escalation and Nuclear Weapons in Russia's Military Strategy," *The RUSI Journal* 163, no. 2 (2018): 4–15, https://doi.org/10/gj53fv; Kristin Ven Bruusgaard, "Russian Nuclear Strategy and Conventional Inferiority," *Journal of Strategic Studies* 44, no. 1 (2021): 3–35, https://doi.org/10/ghhgmx; Bruno Tertrais, "Russia's Nuclear Policy: Worrying for the Wrong Reasons," *Survival* 60, no. 2 (2018): 33–44, https://doi.org/10/gfsv46; Elbridge Colby, "Russia's Evolving Nuclear Doctrine and Its Implications," 2016.

on deterrence that we have partially x-rayed with our various machine-augmented tools in this paper. Instead, our main ambition always has been and very much remains to galvanize this scholarly community into taking more advantage of recent technological developments to make much faster, greater, and more trustworthy progress than our analysis suggests we have made to date. Until just a few years ago, access to sources was extremely difficult. Only a (very) few fortunate scholars had (and to this day still have) access to the 'raw materials' that are so essential for top-notch (also policy-relevant) research. Access to the best scholarly and other specialized literature (Jane's, IISS' Military Balance, Integrum/Eastview Press, etc.) as well as other commercial datasets, remained locked up behind paywalls. Tools to synthesize, explore, visualize, etc, this literature barely existed. Those that did were either financially out of reach to most researchers or required a very high level of 'quantitative' expertise.

In our experience, most people working in this field are only dimly aware of the dramatic changes that have taken over just the past few years (and are continuing to take place!). We therefore have no doubt that many of our colleagues will remain skeptical about some (or even most) of the tools we used in this paper and/or about some (or even most) of our interpretations of the findings to come out of them. We want to emphasize that we too follow these developments with a healthily critical attitude. But we still strongly suspect that very few of our colleagues would disagree with the proposition that better access to the relevant (English *and* Russian) literature would benefit all, irrespective of whether they would then want to explore more 'quantitative' approaches to analyze that literature, or not.

With that latter group, we will be more than happy to share any and all of the open-access corpora (and datasets) we have collected, as well as the tools that would allow them to update these (or any other corpora of their choosing) whenever they would want to. For those colleagues, however, who might be more open to explore some of the more 'quantitative' tools we used throughout this paper, the generous multi-year grant from the Carnegie Corporation of New York allows us to additionally also share the various – mostly open-source – tools we used, the various scripts and notebooks which we made ourselves to obtain all the results we showcased in this paper, and – in our view most importantly – our own documentation of all of this on Rizzoma, our collaborative knowledge-sharing platform.

This brings us to our final, and from our point of view most important, take-away. It is related to collaboration – the lack thereof in our field, but especially also the benefits thereof for ourselves as knowledge-providers, but also for the users of that knowledge, include the broader policy community. As we have seen in our bibliometric analysis, the main knowledge providers in this field hail from two fairly different backgrounds: the purely 'academic' community; and the 'policy analytical' community (think tanks, government analysts, international organizations, etc.). The incentive structures in these two communities differ quite significantly. But what they do have in common, is the pervasive disincentivization of collaboration. The overwhelming 'publish-or-perish' incentive in academia pushes individual academic scholars to regularly publish relatively shorter pieces of work that they can finish (alongside their teaching and/or administrative load) within a few months and to do so preferably on their own, as that diminishes transaction costs and maximizes their own accomplishment recognition (as single authored pieces still seem to be valued more highly than multi-authored ones). The overwhelming 'procure-or-participate-in- paid-projects-or-perish' incentive in the think tank world has a similarly pernicious (and – some might argue – anti-competitive) effect on collaboration. If a think tank obtains funding from a (typically government) customer to conduct a project on Russian deterrence, that project leader's incentive is to staff that project with in-house experts – even if there are better (and/or cheaper) experts on that specific topic available elsewhere. Our own assessment is that the instances that really fund this work (in both cases more often than not 'public' customers) are only dimly aware of this. The undisputable net result, however, is that progress tends to be stovepiped, slow, and qualitatively decidedly suboptimal.

There is, alas, not that much that we as a community of scholars living and working in these environments can do in the short term about our own respective incentive structures. The main author of this paper has argued for quite some time that epistemic communities (also in the defense and security policy analysis world) tend to underestimate their actual 'strategic leverage'. It has been encouraging to see (data-intensive) virological and epidemiological communities empowered in the ubiquitous public policy debates around Covid-19. We remain convinced that we, as a community, could and should do much more to convey to our 'customers' what an enormous opportunity cost they incur by not providing us with more incentives to collaborate. But even in the short term, we cannot see any compelling reasons that would prevent us, even within our current incentivization confines, from collaborating more on such ('open access' and unclassified) topics. We sincerely hope our colleagues will heed our call.

ANNEX A – LIST OF THINK TANKS INCLUDED IN THE ANALYSIS

Domain Name	Number of publications
rand.org	76
csis.org	56
carnegieendowment.org	54
belfercenter.org	47
heritage.org	40
inss.org.il	40
ifri.org	31
brookings.edu	22
ndu.edu	19
sipri.org	15
frstrategie.org	13
rsis.edu.sg	13
rusi.org	13
cna.org	12
ffi-publikasjoner.archive.knowledgearc.net	12
chathamhouse.org	10
cfr.org	8
csbaonline.org	8
swp-berlin.org	8
cnas	7
carnegiecouncil.org	5
clingendael.org	5
hcss.nl	5
inss.ndu.edu	5
atlanticcouncil.org	4
cnas.org	3
carnegie.org	2
cato.org	2
edam.org.tr	2
ndu.edu.tw	1
atlantkommitten.se	1
baselpeaceoffice.org	1
biiss.org	1
carnegie.ru	1
cms.polsci.ku.dk	1
egmontinstitute.be	1
foi.se	1
inss.re.kr	1
iris-france.org	1
paxchristi.org.au	1
swisspeace.ch	1

ANNEX B – LIST OF CONFIGURATIONS FOR CITESPACE

Project properties	
Alias List	on
Export Space	off
Save Merged Slice	off
Noun Phrase: Maximum words	4
Maximum GML Node Label Length	8
Include GP	off
Node Degree Weighted	true
Link Retaining Factor	-1
Maximum Link Per Node	-1
Filter Refs By Intrinsic Citations	on
Use Authors' Full Names	on
Normalize Citations	off
Global Check	off
Exclusion list	on
Export abstracts	on
Enable JDIC	on
Noun Phrase: Minimum Words	2
Burst Terms Threshold	0.00
CTSA	1
Include ED	off
Look Back Years	-1
Percentage of Nodes to Label	100
TopN	0.00001
Time slicing	
Years Per Slice	1
From; To	Depends on a dataset. See 'Our bibliometric datasets' section
Text processing	
Terms source	Title, Abstracts, Author Keywords, Keywords Plus
Terms type	Noun Phrases
Nodes	
Node types	Terms, keywords
Strength	Cosine
Scope	Within Slices
Selection Criteria	
G-index	k = 25
Pruning	
Pathfinder	off
Minimum Spanning Tree	off
Pruning slices networks	off
Pruning merged networks	off
Visualization	
Cluster View	Static
Show Networks by Time Slices	off
Show Merged Networks	on

BIBLIOGRAPHY

Abelson, Donald E. *Do Think Tanks Matter? Assessing the Impact of Public Policy Institutes*. Third edition, Revised and Expanded. Montreal: McGill-Queen's University Press, 2018.

Adamsky, D. "From Moscow with Coercion: Russian Deterrence Theory and Strategic Culture." *Journal of Strategic Studies*, 2018. https://doi.org/10/gdxs6v.

Allen Institute for AI. "Allen Institute for AI." Allen Institute for AI, 2012. https://allenai.org/.

———. *Allenai/S2orc*. Python. 2019. Reprint, AI2, 2021. https://github.com/allenai/s2orc.

Angelov, Dimo. "Top2Vec: Distributed Representations of Topics." *ArXiv:2008.09470 [Cs, Stat]*, August 19, 2020. http://arxiv.org/abs/2008.09470.

Ariely, Dan. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. Rev. and Expanded ed., 3. [print]. New York, NY: Harper Collins Publ, 2009.

Bafeta, Aïda, Ludovic Trinquart, Raphaèle Seror, and Philippe Ravaud. "Analysis of the Systematic Reviews Process in Reports of Network Meta-Analyses: Methodological Systematic Review." *BmJ* 347 (2013). https://doi.org/10/gb3shj.

Baldwin, David A. *Economic Statecraft*. New edition. Princeton: Princeton University Press, 2020.

Barabási, Albert-László. *Linked: The New Science of Networks*, 2002. http://site.ebrary.com/id/10460889.

Barabási, Albert-László, and Márton Pósfai. *Network Science*. Cambridge, United Kingdom: Cambridge University Press, 2016.

Barhate, Pratik. "Latent Dirichlet Allocation for Beginners: A High Level Intuition." Medium, February 5, 2019. https://medium.com/@pratikbarhate/latent-dirichlet-allocation-for-beginners-a-high-level-intuition-23f8a5cbad71.

Baron, Jonathan. *Thinking and Deciding*. 4th ed. New York: Cambridge University Press, 2008.

Baum, Seth, and Anthony Barrett. "A Model for the Impacts of Nuclear War." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 3, 2018. https://doi.org/10.2139/ssrn.3155983.

Bello, Luciano. *Scholarly-Python-Package/Scholarly*. Python. 2014. Reprint, scholarly-python-package, 2021. https://github.com/scholarly-python-package/scholarly.

Bernstein, Douglas A., ed. *Psychology*. 9th Ed. Belmont, CA: Wadsworth, Cengage Learning, 2012.

Bird, Steven, Ewan Klein, and Edward Loper. *Natural Language Processing with Python*. 1st ed. Beijing ; Cambridge [Mass.]: O'Reilly, 2009.

Björk, Bo-Christer, and David Solomon. "The Publishing Delay in Scholarly Peer-Reviewed Journals." *Journal of Informetrics* 7, no. 4 (October 1, 2013): 914–23. https://doi.org/10.1016/j.joi.2013.09.001.

Blank, Stephen J., ed. *Russian Nuclear Weapons: Past, Present, and Future:* Carlisle, PA: U.S. Army War College,Strategic Studies Institute, 2011. https://doi.org/10.21236/ADA555143.

Blei, David M., and John D. Lafferty. "Dynamic Topic Models." In *Proceedings of the 23rd International Conference on Machine Learning – ICML '06*, 113–20. Pittsburgh, Pennsylvania: ACM Press, 2006. https://doi.org/10/bjkmq9.

Borenstein, Michael, ed. *Introduction to Meta-Analysis*. Chichester, U.K: John Wiley & Sons, 2009.

Bornmann, Lutz, Raf Guns, Michael Thelwall, and Dietmar Wolfram. "Which Aspects of the Open Science Agenda Are Most Relevant to Scientometric Research and Publishing? An Opinion Paper." *Quantitative Science Studies*, February 10, 2021, 1–31. https://doi.org/10.1162/qss_e_00121.

Boschee, Elizabeth, Jennifer Lautenschlager, Sean O'Brien, Steve Shellman, and James Starz. "ICEWS Automated Daily Event Data (IN TESTING)," October 11, 2018. https://doi.org/10.7910/DVN/QI2T9A.

Boyd-Graber, Jordan, Yuening Hu, and David Mimno. "Applications of Topic Models." *Foundations and Trends® in Information Retrieval* 11, no. 2–3 (2017): 143–296. https://doi.org/10.1561/1500000030.

Brainard, Jeffrey. "What a Massive Database of Retracted Papers Reveals About Science Publishing's 'Death Penalty.'" *Science*, October 25, 2018. https://doi.org/10.1126/science.aav8384.

Builder, Carl H. *The Masks of War: American Military Styles in Strategy and Analysis*. A Rand Corporation Research Study. Baltimore: Johns Hopkins University Press, 1989.

Cambridge Review of International Affairs. "Submit a Manuscript," January 25, 2016. https://www.cria.polis.cam.ac.uk/Submissions/submit-a-manuscript.

Chen, Chaomei. *How to Use CiteSpace*. Leanpub, 2019. https://leanpub.com/howtouseCiteSpace.

———. "Science Mapping: A Systematic Review of the Literature." *Journal of Data and Information Science* 2, no. 2 (March 21, 2017): 1–40. https://doi.org/10/gfc4wk.

Chen, Chaomei, and Min Song. *Representing Scientific Knowledge: The Role of Uncertainty*. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-62543-0.

———. *Representing Scientific Knowledge: The Role of Uncertainty*. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-62543-0.

Christensen, Garret S., Jeremy Freese, and Edward Miguel. *Transparent and Reproducible Social Science Research: How to Do Open Science*. Oakland, California: University of California Press, 2019.

Clarivate Analytics. "Web of Science." *Web of Science Group* (blog), 2019. https://clarivate.com/webofsciencegroup/solutions/web-of-science/.

Collier, David, James Mahoney, and Jason Seawright. "Claiming Too Much: Warnings About Selection Bias." *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, 2004, 85–102.

Committee on Reproducibility and Replicability in Science, Board on Behavioral, Cognitive, and Sensory Sciences, Committee on National Statistics, Division of Behavioral and Social Sciences and Education, Nuclear and Radiation Studies Board, Division on Earth and Life Studies, Board on Mathematical Sciences and Analytics, et al. *Reproducibility and Replicability in Science*. Washington, D.C.: National Academies Press, 2019. https://doi.org/10.17226/25303.

Cooper, Harris M., Larry V. Hedges, and Jeff C. Valentine, eds. *The Handbook of Research Synthesis and Meta-Analysis*. 2nd ed. New York: Russell Sage Foundation, 2009.

Cooper, Robert. *The Post-Modern State and the World Order*. Paper / Demos, no. 19. London: Demos, 1996.

Covington, Stephen R. "The Culture of Strategic Thought Behind Russia's Modern Approaches to Warfare." Paper. Harvard Kennedy School, Belfer Center for Science and International Affairs, October 2016.

Croitoru, Madalina and GKR, eds. *Graph Structures for Knowledge Representation and Reasoning*. Vol. 5th International Workshop, Gkr 2017, Melbourne, Vic, Australia, August 21, 2017: Revised Selected Papers. Berlin Heidelberg: Springer, 2018.

Dario Taraborelli. "A Good Progress Bar." Tweet. *@ReaderMeter* (blog), January 20, 2021. https://twitter.com/ReaderMeter/status/1351943177293369349.

Davis, Paul K. *Simple Models to Explore Deterrence and More General Influence in the War with Al-Qaeda*. RAND Occasional Paper Series. Santa Monica, Calif.: RAND National Defense Research Inst, 2010.

De Spiegeleire, Stephan, Khrystyna Holynska, Yar Batoh, and Tim Sweijs. *Reimagining Deterrence: Towards Strategic (Dis)Suasion Design*. The Hague, The Netherlands: The Hague Centre for Strategic Studies (HCSS), 2020.

De Spiegeleire, Stephan, Karlijn Jans, Mischa Sibbel, Khrystyna Holynska, and Deborah Lassche. "Implementing Defence Policy: A Benchmark-'Lite.'" *Defense & Security Analysis*, February 2019, 1–23. https://doi.org/10/gft4wf.

De Spiegeleire, Stephan, Matthijs Maas, and Tim Sweijs. *Artificial Intelligence and the Future of Defense: Strategic Implications for a Small Force Provider*. The Hague, The Netherlands: HCSS, 2017. http://hcss.nl/report/artificial-intelligence-and-future-defense.

De Spiegeleire, Stephan, Peter Wijninga, and Tim Sweijs. *Designing Future Stabilization Efforts*. The Hague, The Netherlands: The Hague Centre for Strategic Studies, 2014.

Dehaene, Stanislas. *How We Learn: Why Brains Learn Better Than Any Machine ... for Now.*, 2020. https://www.overdrive.com/search?q=92024A08-400B-4F71-9982-5564CC370435.

———. *Reading in the Brain: The Science and Evolution of a Human Invention*. New York: Viking, 2009. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1115790.

Dong, Xin, Evgeniy Gabrilovich, Geremy Heitz, Wilko Horn, Ni Lao, Kevin Murphy, Thomas Strohmann, Shaohua Sun, and Wei Zhang. "Knowledge Vault: A Web-Scale Approach to Probabilistic Knowledge Fusion." In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 601–10. KDD '14. New York, NY, USA: ACM, 2014. https://doi.org/10/gfvn2s.

Dong, Xin Luna, Evgeniy Gabrilovich, Kevin Murphy, Van Dang, Wilko Horn, Camillo Lugaresi, Shaohua Sun, and Wei Zhang. "Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources," February 12, 2015. https://arxiv.org/abs/1502.03519v1.

Dong, Xin Luna, and Divesh Srivastava. "Knowledge Curation and Knowledge Fusion: Challenges, Models and Applications," 2063–66. ACM Press, 2015. https://doi.org/10/gghvvh.

Du, Jianfeng, Jeff Z. Pan, Sylvia Wang, Kunxun Qi, Yuming Shen, and Yu Deng. "Validation of Growing Knowledge Graphs by Abductive Text Evidences." *Proceedings of the AAAI Conference on Artificial Intelligence* 33 (July 17, 2019): 2784–91. https://doi.org/10.1609/aaai.v33i01.33012784.

East View Press. "East View Press," 2021. https://www.eastviewpress.com/.

EBSCO. "Home Page | EBSCO." EBSCO Information Services, Inc. | www.ebsco.com, 2021. https://www.ebsco.com/.

Eisner, D. A. "Reproducibility of Science: Fraud, Impact Factors and Carelessness." *Journal of Molecular and Cellular Cardiology* 114 (January 1, 2018): 364–68. https://doi.org/10.1016/j.yjmcc.2017.10.009.

Ellsberg, Daniel. "The Theory and Practice of Blackmail." RAND Paper. Santa Monica, CA: RAND Corporation, July 1968.

Else, Holly. "How I Scraped Data from Google Scholar." *Nature*, April 11, 2018. https://doi.org/10.1038/d41586-018-04190-5.

———. "How Unpaywall Is Transforming Open Science." *Nature* 560, no. 7718 (August 15, 2018): 290–91. https://doi.org/10.1038/d41586-018-05968-3.

Elsevier. "Scopus," 2019. https://www.scopus.com/standard/marketing.uri.

European Commission. Directorate General for Research and Innovation, Lisbon Council, CWTS, ESADE, and Elsevier. *Open Metadata of Scholarly Publications: Open Science Monitor Case Study*. LU: Publications Office, 2019. https://data.europa.eu/doi/10.2777/132318.

Explosion. "Software · Explosion," 2020. https://explosion.ai/software#spacy.

Explosion. "Software · Explosion," 2020. https://explosion.ai/software#prodigy.

Faust, David, and Paul E. Meehl. "Using Meta-Scientific Studies to Clarify or Resolve Questions in the Philosophy and History of Science." *Philosophy of Science* 69, no. S3 (September 2002): S185–96. https://doi.org/10.1086/341845.

Fedus, William, Barret Zoph, and Noam Shazeer. "Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity." *ArXiv:2101.03961 [Cs]*, January 11, 2021. http://arxiv.org/abs/2101.03961.

Ferguson, Cat, Adam Marcus, and Ivan Oransky. "Publishing: The Peer-Review Scam." *Nature News* 515, no. 7528 (November 27, 2014): 480. https://doi.org/10.1038/515480a.

Foote, Alyssa. "Darpa Wants to Solve Science's Reproducibility Crisis With AI." *Wired*, February 15, 2019. https://www.wired.com/story/darpa-wants-to-solve-sciences-replication-crisis-with-robots/.

Frankland, Julia, and Margaret A. Ray. "Traditional versus Open Access Scholarly Journal Publishing: An Economic Perspective." *Journal of Scholarly Publishing* 49, no. 1 (October 1, 2017): 5–25. https://doi.org/10.3138/jsp.49.1.5.

Friese, Susanne. Qualitative Data Analysis with Atlas.Ti. London: Sage, 2012.

Fuchs, Christian, and Marisol Sandoval. "The Diamond Model of Open Access Publishing: Why Policy Makers, Scholars, Universities, Libraries, Labour Unions and the Publishing World Need to Take Non-Commercial, Non-Profit Open Access Serious." *TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society* 11, no. 2 (September 9, 2013): 428–43. https://doi.org/10.31269/triplec.v11i2.502.

Ganegedara, Thushan. "Intuitive Guide to Latent Dirichlet Allocation." Medium, December 5, 2020. https://towardsdatascience.com/light-on-math-machine-learning-intuitive-guide-to-latent-dirichlet-allocation-437c81220158.

Garfield, Eugene. "Can Citation Indexing Be Automated ?" In *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*, edited by Laurence B. Heilprin, Laurence R. Heilprin, Mary Elizabeth Stevens, and Vincent Edward, 189–93. Washington D.C.: U.S. Government Printing Office, 1966.

———. “Letter of May 21, 1959 from Dr. E. Garfield to Dr. J. Lederberg,” May 21, 1959. http://www.garfield.library.upenn.edu/lederberg/052159.html.

Gartner Research. “Hype Cycle for Natural Language Technologies, 2020.” Gartner, July 6, 2020. https://www.gartner.com/en/documents/3987154/hype-cycle-for-natural-language-technologies-2020.

Geddes, Barbara. “How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics.” *Political Analysis*, 1990, 131–50.

Google. “Google Scholar,” 2021. https://scholar.google.com/intl/en/scholar/about.html.

Greenberg, S. A. “How Citation Distortions Create Unfounded Authority: Analysis of a Citation Network.” *BMJ* 339, no. jul20 3 (July 23, 2009): b2680–b2680. https://doi.org/10.1136/bmj.b2680.

Grennan, Mark, Martin Schibel, Andrew Collins, and Joeran Beel. “GIANT: The 1-Billion Annotated Synthetic Bibliographic-Reference-String Dataset for Deep Citation Parsing.” In *AICS*, 260–71, 2019.

Grossmann, Alexander, and Björn Brembs. “Current Market Rates for Scholarly Publishing Services.” *F1000Research* 10 (January 12, 2021): 20. https://doi.org/10.12688/f1000research.27468.1.

Hao, Jia, Qiangfu Zhao, Yan Yan, and Guoxin Wang. “A Review of Tacit Knowledge: Current Situation and the Direction to Go.” *International Journal of Software Engineering and Knowledge Engineering* 27, no. 05 (June 2017): 727–48. https://doi.org/10.1142/S0218194017500279.

Harris, Dave. *Literature Review and Research Design: A Guide to Effective Research Practice*. 1st ed. Routledge, 2019. https://doi.org/10.4324/9780429285660.

Harzing, Anne-Wil. “Reflections on the H-Index.” Harzing.com, April 23, 2008. https://harzing.com/publications/white-papers/reflections-on-the-h-index.

GitHub. “HCSS-StratBase/HCSS-Deterrence-Datasets.” Accessed April 28, 2021. https://github.com/HCSS-StratBase/HCSS-Deterrence-datasets.

Hendricks, Ginny, Dominika Tkaczyk, Jennifer Lin, and Patricia Feeney. “Crossref: The Sustainable Source of Community-Owned Scholarly Metadata.” *Quantitative Science Studies* 1, no. 1 (February 1, 2020): 414–27. https://doi.org/10.1162/qss_a_00022.

Higgins, Julian P. T., and Cochrane Collaboration, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Second edition. Cochrane Book Series. Hoboken, NJ: Wiley-Blackwell, 2020.

Hodson, Hal. “Google Wants to Rank Websites Based on Facts Not Links.” New Scientist, February 25, 2015. https://www.newscientist.com/article/mg22530102-600-google-wants-to-rank-websites-based-on-facts-not-links/.

Holynska, Khrystyna, Yevhen Sapolovych, Mikhail Akimov, and Stephan De Spiegeleire. “Events Datasets and Strategic Monitoring.” The Hague, 2020.

Honnibal, Matthew, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. *SpaCy: Industrial-Strength Natural Language Processing in Python*. Zenodo, 2020. https://doi.org/10.5281/zenodo.1212303.

Horbach, Serge P. J. M. “Pandemic Publishing: Medical Journals Strongly Speed up Their Publication Process for COVID-19.” *Quantitative Science Studies* 1, no. 3 (August 1, 2020): 1056–67. https://doi.org/10.1162/qss_a_00076.

Hristovski, Dimitar, Andrej Kastrin, Dejan Dinevski, and Thomas C. Rindflesch. "Constructing a Graph Database for Semantic Literature-Based Discovery." *Studies in Health Technology and Informatics* 216 (2015): 1094.

Huth, Paul, Christopher Gelpi, and D. Scott Bennett. "Escalation of Great Power Disputes: Deterrence Versus Structural Realism, 1816-1984: Version 1." ICPSR – Interuniversity Consortium for Political and Social Research, 1995. https://doi.org/10.3886/ICPSR06355.V1.

———. "The Escalation of Great Power Militarized Disputes: Testing Rational Deterrence Theory and Structural Realism." *The American Political Science Review* 87, no. 3 (1993): 609–23. https://doi.org/10.2307/2938739.

Huth, Paul K. "Extended Deterrence and the Outbreak of War." *The American Political Science Review* 82, no. 2 (1988): 423–43. https://doi.org/10.2307/1957394.

Huth, Paul, and Bruce Russett. "Deterrence Failure and Crisis Escalation." *International Studies Quarterly* 32, no. 1 (March 1988): 29–45. https://doi.org/10.2307/2600411.

I4OC. "I4OC: Initiative for Open Citations." I4OC, 2021. https://i4oc.org/.

"Integrum World Wide," 2021 1998. http://www.integrumworld.com/.

Irawan, Dasapta Erwin. "Berlawan Dengan Prinsip Umum Para Pembeli Barang [Contrary to the Principle of Buying Goods in General]." Medium, January 19, 2021. https://medium.com/open-science-indonesia/berlawan-dengan-prinsip-umum-para-pembeli-barang-5c4893fa362a.

Jacobi, Carina, Wouter Van Atteveldt, and Kasper Welbers. "Quantitative Analysis of Large Amounts of Journalistic Texts Using Topic Modelling." *Digital Journalism* 4, no. 1 (2016): 89–106. https://doi.org/10/f3s2sg.

Johnson, Jesse C., Brett Ashley Leeds, and Ahra Wu. "Capability, Credibility, and Extended General Deterrence." *International Interactions* 41, no. 2 (March 15, 2015): 309–36. https://doi.org/10.1080/03050629.2015.982115.

Johnson, Jesse, Brett Ashley Leeds, and Ahra Wu. "Capability, Credibility, and Extended General Deterrence." Harvard Dataverse, May 22, 2015. https://doi.org/10.7910/DVN/LT9TPB.

Jun Zhao, Frank van Harmelen, Jie Tang, Xianpei Han, Quan Wang, and Xianyong Li, eds. *Knowledge Graph and Semantic Computing. Knowledge Computing and Language Understanding*. New York, NY: Springer Berlin Heidelberg, 2019.

Kahneman, Daniel. *Thinking, Fast and Slow*. 1st ed. New York: Farrar, Straus and Giroux, 2011.

Kapadia, Shashank. "Topic Modeling in Python: Latent Dirichlet Allocation (LDA)." Medium, December 29, 2020. https://towardsdatascience.com/end-to-end-topic-modeling-in-python-latent-dirichlet-allocation-lda-35ce4ed6b3e0.

kexugit. "Microsoft's MT-DNN Achieves Human Performance Estimate on General Language Understanding Evaluation (GLUE) Benchmark," June 20, 2019. https://docs.microsoft.com/en-us/archive/blogs/stevengu/microsoft-achieves-human-performance-estimate-on-glue-benchmark.

King, Gary, Robert O. Keohane, and Sidney Verba. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, N.J: Princeton University Press, 1994.

Kleinberg, Jon. "Bursty and Hierarchical Structure in Streams." *Data Mining and Knowledge Discovery* 7, no. 4 (2003): 373–97.

Knight, Will. "AI Is Biased. Here's How Scientists Are Trying to Fix It." *Wired*, December 19, 2019. https://www.wired.com/story/ai-biased-how-scientists-trying-fix/.

Kofman, Michael, Anya Fink, and Jeffrey Edmonds. "Russian Strategy for Escalation Management: Evolution of Key Concepts." Arlington, VA: Center for Naval Analysis, April 13, 2020. https://www.cna.org/CNA_files/PDF/DRM-2019-U-022455-1Rev.pdf.

Kozak, Marcin, Lutz Bornmann, and Loet Leydesdorff. "How Have the Eastern European Countries of the Former Warsaw Pact Developed Since 1990? A Bibliometric Study." *Scientometrics* 102, no. 2 (February 1, 2015): 1101–17. https://doi.org/10.1007/s11192-014-1439-8.

Kugley, Shannon, Anne Wade, James Thomas, Quenby Mahood, Anne-Marie Klint Jørgensen, Karianne Hammerstrøm, and Nila Sathe. "Searching for Studies: A Guide to Information Retrieval for Campbell Systematic Reviews." *Campbell Systematic Reviews* 13, no. 1 (2017): 1–73. https://doi.org/10.4073/cmg.2016.1.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago press, 2012.

Larivière, Vincent, Stefanie Haustein, and Philippe Mongeon. "The Oligopoly of Academic Publishers in the Digital Era." *PLOS ONE* 10, no. 6 (June 10, 2015): e0127502. https://doi.org/10.1371/journal.pone.0127502.

Letrud, Kare, and Sigbjorn Hernes. "Affirmative Citation Bias in Scientific Myth Debunking: A Three-in-One Case Study." *PloS One* 14, no. 9 (2019): e0222213–e0222213. https://doi.org/10.1371/journal.pone.0222213.

Lewis, David W. "Is Scholarly Publishing Like Rock and Roll?" *Journal of Librarianship and Scholarly Communication* 8, no. 1 (November 9, 2020): 2333. https://doi.org/10.7710/2162-3309.2333.

Lopez, Patrice. "Introduction – GROBID Documentation." GROBID Documentation, 2020. https://grobid.readthedocs.io/en/latest/Introduction/.

Luis Serrano. *Latent Dirichlet Allocation*, 2020. https://www.youtube.com/watch?v=T05t-SqKArY.

Mahala, Avijit, Gayatri Dwivedi, and Manorama Tripathi. "A Bibliometric Study of Grey Literature (2007-2019)." *Collection and Curation* ahead-of-print, no. ahead-of-print (January 1, 2020). https://doi.org/10.1108/CC-12-2019-0043.

Maisonobe, Marion, Laurent Jégou, Nikita Yakimovich, and Guillaume Cabanac. "NETSCITY: A Geospatial Application to Analyse and Map World Scale Production and Collaboration Data between Cities," 12. Rome, Italy, 2019.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*. Reprinted. Cambridge: Cambridge Univ. Press, 2009.

Mark, Melvin M., and Charles S. Reichardt. "Quasi-Experimentation." In *Handbook of Applied Social Research Methods*, edited by Leonard Bickman and Debra J. Rog. SAGE, 1998.

Martín-Martín, Alberto, Mike Thelwall, Enrique Orduna-Malea, and Emilio Delgado López-Cózar. "Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A Multidisciplinary Comparison of Coverage via Citations." *Scientometrics*, 2020, 1–36.

Maxime. "Introduction to Semi-Supervised Learning and Adversarial Training." Medium, May 24, 2019. https://medium.com/inside-machine-learning/placeholder-3557ebb3d470.

Mazarr, Michael, Arthur Chan, Alyssa Demus, Bryan Frederick, Alireza Nader, Stephanie Pezard, Julia Thompson, and Elina Treyger. *What Deters and Why: Exploring Requirements for Effective Deterrence of Interstate Aggression*. RAND Corporation, 2018. https://doi.org/10.7249/RR2451.

McGann, James G. "2019 Global Go To Think Tank Index Report," 2020. https://repository.upenn.edu/cgi/viewcontent.cgi?article=1018&context=think_tanks.

———. *Think Tanks and Policy Advice in the United States: Academics, Advisors and Advocates*. Routledge Research in American Politics 1. Abingdon [England] ; New York: Routledge, 2007.

McKenzie, Kate. "What Is the Long Tail of Publishing? – Publishing Trendsetter," October 4, 2013. http://publishingtrendsetter.com/industryinsight/long-tail-publishing/.

Medicine, National Academies of Sciences, Engineering, and, Policy and Global Affairs, Board on Research Data and Information, and Committee on Toward an Open Science Enterprise. *Open Science by Design: Realizing a Vision for 21st Century Research*. National Academies Press, 2018.

Menard, Alvaro de. "What's Wrong with Social Science and How to Fix It: Reflections After Reading 2578 Papers." Fantastic Anachronism, September 11, 2020. https://fantasticanachronism.com/2020/09/11/whats-wrong-with-social-science-and-how-to-fix-it/index.html.

Microsoft. "Home | Microsoft Academic." Microsoft Academic, 2021. https://academic.microsoft.com/home.

Mielke, K., and A. K. Hornidge. "Area Studies at the Crossroads: Knowledge Production After the Mobility Turn." books.google.com, 2017. https://books.google.com/books?hl=en&lr=&id=OLEDgAAQBAJ&oi=fnd&pg=PR8&dq=%22area+studies%22&ots=P4el-y61uT&sig=vBy-HcbiSsw6Dh9NMs-7T6b6grWM.

"Militera: Military Literature," 2001. http://militera.lib.ru/.

Mindeli, L. E., and V. A. Markusova. "Bibliometric Studies of Scientific Collaboration: International Trends." *Automatic Documentation and Mathematical Linguistics* 49, no. 2 (March 1, 2015): 59–64. https://doi.org/10.3103/S0005105515020065.

Nangia, Nikita, and Samuel R. Bowman. "Human vs. Muppet: A Conservative Estimate of Human Performance on the GLUE Benchmark." *ArXiv:1905.10425 [Cs]*, June 1, 2019. http://arxiv.org/abs/1905.10425.

Nicholson, J.M., M. Mordaunt, P. Lopez, A. Uppala, D. Rosati, N.P. Rodrigues, P. Grabitz, and S.C. Rife. "Scite: A Smart Citation Index That Displays the Context of Citations and Classifies Their Intent Using Deep Learning." Preprint. Scientific Communication and Education, March 16, 2021. https://doi.org/10.1101/2021.03.15.435418.

Nicholson, Josh. "Scite_," February 16, 2021.

Nikolenko, Sergey I., Sergei Koltcov, and Olessia Koltsova. "Topic Modelling for Qualitative Studies." *Journal of Information Science* 43, no. 1 (February 2017): 88–102. https://doi.org/10/gfkpvz.

Open Science Collaboration. "Estimating the Reproducibility of Psychological Science." *Science* 349, no. 6251 (August 28, 2015): aac4716–aac4716. https://doi.org/10.1126/science.aac4716.

Osinga, Frans, and Tim Sweijs. NL ARMS Netherlands Annual Review of Military Studies 2020 : Deterrence in the 21st Century—Insights from Theory and Practice. Springer Nature, 2021. https://library.oapen.org/handle/20.500.12657/47298.

OU Event Data. "Terrier," August 28, 2018. http://terrierdata.org/#about.

Pan, Jeff Z., Guido Vetere, Jose Manuel Gomez-Perez, and Honghan Wu, eds. *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. Springer International Publishing, 2017. //www.springer.com/us/book/9783319456522.

Penders, Bart. "Ten Simple Rules for Responsible Referencing." *PLOS Computational Biology* 14, no. 4 (2018): e1006036. https://doi.org/10.1371/journal.pcbi.1006036.

Petrokofsky, Gillian. "Guidelines and Standards for Evidence Synthesis in Environmental Management: Version 5.0." Monograph. Collaboration for Environmental Evidence, 2018. https://eprints.soton.ac.uk/420195/.

Piwowar, Heather, Jason Priem, and Richard Orr. "Our Research." Our Research, 2021. https://ourresearch.org/.

Polanyi, Michael. *Personal Knowledge: Towards a Post-Critical Philosophy*, 2005. https://ebookcentral.proquest.com/lib/uqac-ebooks/detail.action?docID=179903.

"ProQuest | Databases, EBooks and Technology for Research," 2021. http://about.proquest.com/.

Prasad, Animesh, Manpreet Kaur, and Min-Yen Kan. "Neural ParsCit: A Deep Learning-Based Reference String Parser." *International Journal on Digital Libraries* 19, no. 4 (2018): 323–37.

Quah, Jon S. T. *Public Administration Singapore-Style*. Research in Public Policy Analysis and Management 19. Bingley: Emerald, 2010.

Rabindranath, Giri. "Topic Modeling with LDA: An Intuitive Explanation." HDS, October 19, 2020. https://highdemandskills.com/topic-modeling-intuitive/.

Řehůřek, Radim, and Petr Sojka. "Software Framework for Topic Modelling with Large Corpora." In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50. Valletta, Malta: ELRA, 2010.

Rich, Andrew. *Think Tanks, Public Policy, and the Politics of Expertise*. Cambridge: Cambridge University Press, 2004. https://doi.org/10.1017/CBO9780511509889.

Ritchie, Stuart. *Science Fictions: How Fraud, Bias, Negligence, and Hype Undermine the Search for Truth*. First edition. New York: Metropolitan Books ; Henry Holt and Company, 2020.

Rodriguez, Luisa. "How Many People Would Be Killed as a Direct Result of a Us-Russia Nuclear Exchange?" Rethink Priorities, June 29, 2019. https://www.rethinkpriorities.org/blog/2020/6/19/how-many-people-killed-nuclear-war.

Roemer, Robin Chin, and Rachel Borchardt. *Meaningful Metrics: A 21st Century Librarian's Guide to Bibliometrics, Altmetrics, and Research Impact*. Chicago: Association of College and Research Libraries, A division of the American Library Association, 2015.

Rogers, Adam. "Darpa Wants to Build a BS Detector for Science." *Wired*, July 30, 2017. https://www.wired.com/story/darpa-bs-detector-science/.

Romary, Laurent, and Patrice Lopez. "Grobid-Information Extraction from Scientific Publications." *ERCIM News* 100 (2015).

Saldaña, Johnny. *The Coding Manual for Qualitative Researchers*. 2nd ed. Los Angeles: SAGE, 2013.

Sanjaya, Ngurah Agus, Mouhamadou Lamine Ba, Talel Abdessalem, and Stéphane Bressan. "Harnessing Truth Discovery Algorithms On The Topic Labelling Problem." In *Proceedings of the 20th International Conference on Information Integration and Web-Based Applications & Services*, 8–14.

IiWAS2018. New York, NY, USA: Association for Computing Machinery, 2018. https://doi.org/10.1145/3282373.3282390.

Schelling, Thomas C. *Arms and Influence*. New Haven and London: Yale University Press, 1966.

Schelling, Thomas C. *The Strategy of Conflict*. Harvard University Press, 1960.

Schröder, Christopher, and Andreas Niekler. "A Survey of Active Learning for Text Classification Using Deep Neural Networks." *ArXiv:2008.07267 [Cs]*, August 17, 2020. http://arxiv.org/abs/2008.07267.

Schrodt, Philip A. "Analyzing International Event Data: A Handbook of Computer-Based Techniques," March 22, 2012. http://eventdata.psu.edu/papers.dir/automated.html.

Schrodt, Philip A. "Cameo: Conflict and Mediation Event Observations Event and Actor Codebook," 2012. http://eventdata.parusanalytics.com/data.dir/cameo.html.

Sechser, Todd S. "Militarized Compellent Threats, 1918–2001." *Conflict Management and Peace Science* 28, no. 4 (September 1, 2011): 377–401. https://doi.org/10.1177/0738894211413066.

———. "Replication Data for: Militarized Compellent Threats, 1918-2001." Harvard Dataverse, January 2, 2013. https://doi.org/10.7910/DVN/VDJQ1E.

Shannon, Claude Elwood. "A Mathematical Theory of Communication." *ACM SIGMOBILE Mobile Computing and Communications Review* 5, no. 1 (2001): 3–55.

Shotton, David. "Elsevier Endorses DORA and Opens Its Journal Article Reference Lists." *OpenCitations Blog* (blog), December 20, 2020. https://opencitations.wordpress.com/2020/12/20/elsevier-opens-references/.

Shultz, George P., William J. Perry, Henry A. Kissinger, and Sam Nunn. "A World Free of Nuclear Weapons." *Wall Street Journal*, January 4, 2007, sec. Opinion. https://www.wsj.com/articles/SB116787515251566636.

Siddaway, Andy P., Alex M. Wood, and Larry V. Hedges. "How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews, Meta-Analyses, and Meta-Syntheses." *Annual Review of Psychology* 70, no. 1 (January 4, 2019): 747–70. https://doi.org/10.1146/annurev-psych-010418-102803.

Singhal, Amit. "Introducing the Knowledge Graph: Things, Not Strings." Google, May 16, 2012. Knowledge graph construction techniques.

Sokov, Nikolai. "The New, 2010 Russian Military Doctrine: The Nuclear Angle." *James Martin Center for Nonproliferation Studies* 5 (2010).

Springer Nature. "Formatting Guide | Nature," 2021. https://www.nature.com/nature/for-authors/formatting-guide.

Staddon, J. E. R. *Scientific Method: How Science Works, Fails to Work, and Pretends to Work*. New York, NY: Routledge/Taylor & Francis Group, 2018.

statista.com. "Coronavirus Deaths Worldwide by Country." Statista. Accessed February 20, 2021. https://www.statista.com/statistics/1093256/novel-coronavirus-2019ncov-deaths-worldwide-by-country/..

Sugimoto, Cassidy R., and Vincent Larivière. *Measuring Research: What Everyone Needs to Know*. Oxford University Press, 2018.

Sweijs, Tim, and Frans Osinga. "Conclusion: Insights from Theory and Practice." In *NL ARMS Netherlands Annual Review of Military Studies 2020: Deterrence in the 21st Century—Insights from Theory and Practice*, edited by Frans Osinga and Tim Sweijs, 503–30. NL ARMS. The Hague: T.M.C. Asser Press, 2021. https://doi.org/10.1007/978-94-6265-419-8_26.

Sweijs, Tim, and Samo Zilincik. *Cross Domain Deterrence and Hybrid Conflict*. The Hague: The Hague Centre for Strategic Studies, 2019. https://hcss.nl/sites/default/files/files/reports/Cross%20Domain%20Deterrence%20-%20Final_0.pdf.

Tableau. "Tableau: Business Intelligence and Analytics Software." Tableau, 2021. https://www.tableau.com/.

Tertrais, Bruno. "Russia's Nuclear Policy: Worrying for the Wrong Reasons." *Survival* 60, no. 2 (2018): 33–44. https://doi.org/10/gfsv46.

Teixeira da Silva, Jaime A., Judit Dobránszki, Panagiotis Tsigaris, and Aceil Al-Khatib. "Predatory and Exploitative Behaviour in Academic Publishing: An Assessment." *The Journal of Academic Librarianship* 45, no. 6 (November 1, 2019): 102071. https://doi.org/10.1016/j.acalib.2019.102071.

The GDELT Project. "The GDELT Project." The GDELT Project, 2019. https://www.gdeltproject.org/.

The Hague Centre for Strategic Studies. *HCSS-StratBase/Deterrence-IS-Russia-Russian_LDA*. HTML, 2021. https://github.com/HCSS-StratBase/Deterrence-IS-Russia-Russian_LDA.

TilmannR, design: John Manoogian IIIcategories and descriptions: Buster Bensonimplementation: *List of 188 Cognitive Biases*. June 6, 2018. This file was derived from: The Cognitive Bias Codex – 180+ biases, designed by John Manoogian III (jm3).png: https://commons.wikimedia.org/wiki/File:Cognitive_bias_codex_en.svg.

Tkaczyk, Dominika, Andrew Collins, Paraic Sheridan, and Joeran Beel. "Machine Learning vs. Rules and Out-of-the-Box vs. Retrained: An Evaluation of Open-Source Bibliographic Reference and Citation Parsers." In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*, 99–108. JCDL '18. New York, NY, USA: Association for Computing Machinery, 2018. https://doi.org/10.1145/3197026.3197048.

Tkaczyk, Dominika, Pawe\l Szostek, Mateusz Fedoryszak, Piotr Jan Dendek, and \Lukasz Bolikowski. "CERMINE: Automatic Extraction of Structured Metadata from Scientific Literature." *International Journal on Document Analysis and Recognition (IJDAR)* 18, no. 4 (2015): 317–35.

Trowbridge, C W, and Jan K Sykulski. "On the Shoulders of Giants," 2013. https://eprints.soton.ac.uk/355119/1/COMPUMAG%25202013%2520CTW%252BJKS.pdf.

Tsao, J.Y., K.W. Boyack, M.E. Coltrin, J.G. Turnley, and W.B. Gauster. "Galileo's Stream: A Framework for Understanding Knowledge Production." *Research Policy* 37, no. 2 (March 2008): 330–52. https://doi.org/10.1016/j.respol.2007.10.004.

Eck, Nees Jan van, and Ludo Waltman. "Visualizing Bibliometric Networks." In *Measuring Scholarly Impact*, edited by Ying Ding, Ronald Rousseau, and Dietmar Wolfram, 285–320. Cham: Springer International Publishing, 2014. https://doi.org/10.1007/978-3-319-10377-8_13.

Ven Bruusgaard, Kristin. "Russian Nuclear Strategy and Conventional Inferiority." *Journal of Strategic Studies* 44, no. 1 (2021): 3–35. https://doi.org/10/ghhgmx.

Visser, Martijn, Nees Jan van Eck, and Ludo Waltman. "Large-Scale Comparison of Bibliographic Data Sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic." *ArXiv:2005.10732 [Cs]*, January 17, 2021. http://arxiv.org/abs/2005.10732.

Wallace, John, Bosah Nwosu, and Mike Clarke. "Barriers to the Uptake of Evidence from Systematic Reviews and Meta-Analyses: A Systematic Review of Decision Makers' Perceptions." *BMJ Open* 2, no. 5 (2012).

Waltman, Ludo. "Q&A about Elsevier's Decision to Open Its Citations." *Leiden Madtrics* (blog), December 22, 2020. https://leidenmadtrics.nl/articles/q-a-about-elseviers-decision-to-open-its-citation.

Waltman, Ludo, Robert J.W. Tijssen, and Nees Jan van Eck. "Globalisation of Science in Kilometres." *Journal of Informetrics* 5, no. 4 (October 2011): 574–82. https://doi.org/10.1016/j.joi.2011.05.003.

Wang, Alex, Amanpreet Singh, Julian Michael, Felix Hill, Omer Levy, and Samuel Bowman. "GLUE Benchmark Leaderboard." GLUE Benchmark Leaderboard, 2021. https://gluebenchmark.com/ .

Wang, Xiaoyu, Wang Zhao, Yongzhong Wang, Qin Zhao, Xuejie Yang, Kaixiang Su, and Dongxiao Gu. "Visual Analysis on Information Theory and Science of Complexity Approaches in Healthcare Research." *Entropy* 22, no. 1 (January 2020): 109. https://doi.org/10.3390/e22010109.

Watts, Duncan J. "Should Social Science Be More Solution-Oriented?" *Nature Human Behaviour* 1, no. 1 (January 10, 2017): 1–5. https://doi.org/10.1038/s41562-016-0015.

Weede, Erich. "Extended Deterrence, Superpower Control, and Militarized Interstate Disputes, 1962-76." *Journal of Peace Research* 26, no. 1 (1989): 7–17.

Weitz, Stefan. *Search: How the Data Explosion Makes Us Smarter*. London, UNITED KINGDOM: Taylor & Francis Group, 2016. http://ebookcentral.proquest.com/lib/gatech/detail.action?docID=4732811.

White, G.M. "Speech Recognition: A Tutorial Overview." *Computer* 9, no. 5 (May 1976): 40–53. https://doi.org/10.1109/C-M.1976.218586.

Wikipedia. "Johnny We Hardly Knew Ye." In *Wikipedia*, March 25, 2020. https://en.wikipedia.org/w/index.php?title=Johnny_We_Hardly_Knew_Ye&oldid=947351758.

———. "List of Cognitive Biases." In *Wikipedia*, April 24, 2021. https://en.wikipedia.org/w/index.php?title=List_of_cognitive_biases&oldid=1019556972.

Xin Luna Dong. "Leaving No Valuable Data Behind: The Crazy Ideas and the Business." Presented at the MLConf, Seattle, Wash., May 2017. https://www.slideshare.net/SessionsEvents/luna-dong-principal-scientist-amazon-at-mlconf-seattle-2017.

Xue, Linting, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, and Colin Raffel. "Mt5: A Massively Multilingual Pre-Trained Ext-to-Text Transformer." *ArXiv:2010.11934 [Cs]*, March 11, 2021. http://arxiv.org/abs/2010.11934.

Zhou, Ming, Nan Duan, Shujie Liu, and Heung-Yeung Shum. "Progress in Neural NLP: Modeling, Learning, and Reasoning." *Engineering* 6, no. 3 (March 1, 2020): 275–90. https://doi.org/10.1016/j.eng.2019.12.014.

Zysk, Katarzyna. "Escalation and Nuclear Weapons in Russia's Military Strategy." *The RUSI Journal* 163, no. 2 (2018): 4–15. https://doi.org/10/gj53fv.

TABLE OF FIGURES

TABLE OF TABLES