DHQ: Digital Humanities Quarterly

2025
Volume 19 Number 3

Keywords in Digital Humanities – A critical assessment of computational techniques for mapping security and freedom in historical debates

Tobias Blanke <t_dot_blanke_at_uva_dot_nl>, University of Amsterdam https://orcid.org/0000-0003-0956-987X
Chloe Papadopoulou <c_dot_papadopoulou_at_uva_dot_nl>, University of Amsterdam https://orcid.org/0009-0000-4226-9639

DOI: https://doi.org/10.63744/3gct5wexdcgu

Abstract

The article investigates the continued importance of keywords in digital humanities and especially their relation to recent machine-learning approaches. Different research practices related to digital humanities agree on the importance of keywords to present issues and/or provide the baseline for new stories about the past, about literature or media. Keywords are useful to target ideas that have no clear definitions and productive in describing contested categorisations that are key to humanities scholarship. At the same time, keywords are often employed used without considering specific contexts and how they are generated. To understand the diversity of keywords in digital humanities, we consider three approaches to computationally generating keywords, from traditional and established ones to state-of-the-art language modelling. With these three approaches, we analyse a case where keywords should be especially powerful, as underlying considerations are uncertain. We cover the relation between security, human-rights and freedom according to discussions in United Nations documents. Finally, we present a number of approaches to productively use keywords to tell a different story about the relation of security and freedom according to UN discussions.

1. Introduction –Keywords Methods in the Digital Humanities

Keywords have become essential instruments in the digital humanities, bridging the gap between algorithmic processing and human interpretation. Yet, the methods by which these words are identified and the assumptions behind their significance demand closer scrutiny. Far from being neutral markers, keywords are the product of specific methodological and interpretative choices that shape how texts are understood, categorized, and discussed. A *keyword* is not an intrinsic property of a text, but a term that has been assigned a special status as a clue to the text's significance. 1

Words are always present in a text, but we produce *keywords* through our methods of selecting and analysing them. This practice has a long history with many inflections in the digital humanities, from metadata for discovery (e.g. keyword search) to the focus of this article: methods that computationally identify words as being especially significant for determining a text’s overall meaning, genre or classification. The central question of this article is not about finding keywords but interrogating the methodological choices we make to select certain words and treat them as keywords. This paper argues that computationally-generated keywords should not be treated as objective indicators of textual meaning, but as interpretive tools. 2

Using computational methods to navigate complex texts is a by now a well-established approach in the digital humanities [Jockers and Underwood 2015]. Ted Underwood has shown how text mining can reveal literature genre not as a fixed category but as a “set of relations between works” [Underwood 2019, 41]. Machine learning can be attractive to humanities as it helps describe cultural things as heterogeneous, their descriptions as “blurry” and their distinctions 3 as “brittle”. Similarly, Richard Jean So employs predictive analytics to deconstruct racial inequality in postwar fiction, using computation as an engine to challenge certainty rather than create it [So 2020]. He is interested in categories that are not “ontologically coherent” and uses a range of “textual features” like narration style or lexical density as inputs to his models. In our own work, machine learning has been employed to decode categories that escape clear definitions. We worked on leadership in a political scientific organisation [Blanke et al. 2024].

While computational methods can identify complex patterns that might be missed otherwise, we still need a way to interpret the results. For digital humanities, this interpretation is overwhelmingly mediated by designating certain words or phrases as significant. We can see this interpretative bridge in action in Underwood’s work. To understand a model that predicts a book's success, he translates its statistical logic into a human-readable narrative. He does this by highlighting the words the model has weighted most heavily, treating them as keywords that signal the model's reading of a text. He writes that the model favours “dreams and death... cold, but not hot; fear, but not joy” [Underwood 2019, 83]. The work of the digital humanist, then, is to weave a narrative from the words a machine has been trained to prioritize. 4

[So 2020] performs similar acts of translation when he discusses the outputs of models by focusing on what he calls *signals*, which are words or textual features his models identify as highly predictive. He explains: “If we skim the words that tend to appear in bestsellers [...], a signal arises: ‘divorced’, ‘airport’, and ‘hour’.” He then interprets these signals: “‘divorced’ supports the claim that bestsellers turned toward stories of private life...; ‘airport’ echoes the argument that bestsellers were chiefly interested in postwar affluence” [So 2020, 123–124]. For both scholars, the computational process identifies statistically important words, and the researcher then uses these words as keywords or *signals* to build their argument. 5

This practice of designating words as analytically significant also extends to the digital methods of Internet Studies. Scholars like Richard Rogers focus on methodologies “native” to the web to study society [Rogers 2013] and not cultural heritage as Underwood and So. For Rogers, an Internet search ranking does not only reflect information relevance but also societal importance. Scraping Google-news search results allows researchers to understand attention and airtime given to, for example, climate-change sceptics. Keywords as indicators of significance and keys to further computation are a central component of these “methods of the medium” and appear repeatedly in Roger’s book *Digital Methods* to trace “partisanship and issue resonance” online [Rogers 2013, 113]. [Burgess and Matamoros-Fernández 2016] are a typical example inspired by keywords as Internet research methods. They explicitly build their analysis by treating the word “gamergate” as their keyword to define their dataset of Tweets, a conscious methodological choice. They decide which word to treat as a keyword to structure the inquiry. 6

Beyond digital humanities and new media research, the practice of designating certain words as analytical lenses has a rich intellectual history in cultural analysis, most famously articulated by the cultural theorist Raymond Williams [Williams 1973]. Williams deliberately selected a vocabulary of words, his keywords, not to resolve their meanings, but to use them as prisms through which to view cultural debates and controversies. Since then, we have seen many adaptations of keywords for the critical study of (digital) cultures [Bennett et al. 2013] [Striphas 2015] [Peters 2016] [AI Now Institute 2021] [Thylstrup et al. 2021]. While Williams's selection was manual and interpretive, and ours is computational and automated, the underlying principle is the same: Both approaches treat certain words as having a special status for analysing cultural phenomena that resist easy definition. 7

This transition from interpretive concept to computational datum carries profound consequences for both research methodology and the nature of textual analysis itself. The significance of keywords is amplified in the digital era, because their very function has been transformed. Once a concept for human-led qualitative analysis, keywords are now systematically operationalized by computational methods. The opportunities for using them as special kinds of maps have therefore changed. This article examines computational production of keywords for the digital humanities and the productive role they could play in both analysing text computationally as well as further interpreting computational results. 8

Thus, we understand *keywords* not as pre-existing entities but as an analytical construct. This construct serves a dual 9 purpose: First, it is the output of a method for condensing texts and identifying their most statistically salient features. Second, it serves as the input for interpretation, allowing researchers to integrate computational findings into their own narratives. Perhaps their most important function has not changed since Raymond Williams described his own project as a “special kind of map” [Willaims 1963, XI]. While his selection was manual and ours is computational, the underlying principle holds: Designating certain words as analytically significant is a way of drawing an interpretive map of complex cultural relations. Our research investigates how different cultural map-making tools produce different kinds of maps.

While the practice of using keywords to interpret computational results is widespread, the analytical consequences of how we generate them are less understood. This paper moves beyond the general claim that keywords are important and instead poses a more targeted set of questions: 10

1. How do different computational methods construct the *keyword* differently?
2. What are the analytical trade-offs between these methods in terms of interpretability, semantic nuance, and conceptual bias?
3. How can we understand computationally generated *keywords* as map-making techniques of cultural phenomena?

This paper presents a computational investigation designed to answer these questions. We study how three different methods for identifying keywords produce different analytical lenses: TF-IDF, KeyBERT, and a custom attention model. TF-IDF offers a simple, interpretable baseline; KeyBERT leverages pre-trained language models for semantic understanding; and the custom attention model provides a more transparent deep-learning approach, enabling closer scrutiny of keyword relevance. 11

We explore the impact of these three keyword lenses on a complex, high-stakes case. In the next section, we introduce the historical and contradictory relationship between security and freedom as one such complex cultural terrain that can benefit from a comparative analysis of these map-making techniques. Our choice of a collection for this experiment is not arbitrary. We require a domain where the concepts are ambiguous, contested and politically charged, to understand the power of keywords and the epistemic and social importance of selecting the right ones. For this reason, we focus on the historical debates surrounding “security” and “freedom” in the United Nations General Assembly. Both security and freedom are core concepts in political discourses whose meanings are constantly negotiated and thus well suited for a keyword analysis. Furthermore, the distinction between them is often semantically blurry and historically contingent. Finally, the way these terms are defined and prioritized has real-world consequences. This case allows us to move from a technical comparison of tools to a critical assessment of how our methodological choices can shape our understanding of history. By analysing how each method frames the relationship between freedom and security, we can see directly how the choice of a keyword-generation technique influences the stories we can tell. 12

2. A Special Kind of Map of Freedom and Security

In this section, we delve into the discursive history of security and freedom as debated within the United Nations General Assembly. These terms are not static ideals but cornerstone principles of political life, perpetually in flux as their meanings are negotiated and redefined. This inherent dynamism makes their evolution a particularly fertile ground for a (computational) keyword analysis, revealing shifts in international priorities and power. The distinction between security on the one hand and freedom and rights on the other has been a key debate in culture and society where it is difficult to determine clear boundaries. Both terms are themselves not easy to define and make for even more conceptual difficulties when brought together. [Hobbes 2017] is often cited for a tradition of thought that wants security first and at high cost. In his political world, we give up individual freedoms to achieve security for all. His is a freedom from fear and not the freedom to do something. This kind of freedom, where security comes first, has since been strongly criticized as fundamentally lacking and reducing the human experience. Freedom has also been defended in its full experience in Hannah Arendt’s work. For her, there is no compromising: “Men can only be free with reference to one another, only, that is, in the fields of politics and of the things they do; it is only in these spheres that they come to realize that freedom is something positive” [Arendt 1961, 191]. As freedom is linked to the fact of being in a political community, Arendt has introduced the human “right to have rights” (for a discussion see [Oman 2010]). 13

Despite such high-profile contributions, we are still no closer to resolving the relation of freedom and security, which makes it attractive for a machine-learned keyword analysis. The sociologist Zygmunt Bauman sees no resolve for the tension between freedom and security: “[W]e cannot be human without both security and freedom; but we cannot have both at the same time and both in quantities that we find fully satisfactory” [Bauman 2000, 5]. The current debate on freedom and security is hardly ever as radical as Hobbes and is happy to state that tensions cannot be resolved. With machine learning, we can work out these tensions using public discussions about freedom and security. In a computational version of Williams’s special kind of map, keywords will help us describe the key ideas of freedom and security without having to fix them. 14

A favourite metaphor of politicians is to “balance” security and freedom in the work on rights for their citizens. A former UK prime minister once put this at the centre of a new Bill of Rights [Cameron 2006]. According to critical security studies scholar Didier Bigo who investigated the language used to justify the Global War on Terror, this kind of language is the “way to diminish the value of freedom. […] The balance metaphor masks the imbalance that exists between the two dimensions and silences the capacity of political judgement” [Bigo 2011, 398]. He observes that there are many different words and languages of freedom, as it is reduced to individual levels. Freedom becomes fragmented while security becomes the global unified goal of safety and protection. 15

As freedom and security and their relations escape a clear definition, keywords provide a way to outline their semantic complications and historical debates. To describe freedom and security relations through keywords, we rely on how they are talked about in political debates. What are their associations and how do these change over time? We have decided to focus on the discussions on political freedom and security within an international organisation that has been founded to provide global security after the Second World War and enable the global developments of rights and freedom. The newly established United Nations adopted the Universal Declaration of Human Rights (UDHR) in 1948, to protect the rights of all individuals around the world. Amnesty International thus calls it a “global road map for freedom and equality” [Amnesty 2023]. The UDHR manifests for the first time that all people deserve to live freely. It contains thirty rights and freedoms that all humans have, which include freedom of expression and the right to asylum. But the United Nations (UN) assembly is as much about the issue of security, as it is about the issue freedom and rights. The United Nations Charter, the foundational treaty of the United Nations from 1945, is equally committed to “further international peace and security” as well as to “encourage respect for human rights and for fundamental freedoms for all without distinction as to race, sex, language, or religion” [United Nations 1945]. The keyword “security” dominates the text with over 150 appearances compared to only 6 for “freedom.” 16

The UN Charter already mentions the United Nations Security Council (UNSC), which has seen renewed attention through the war in Ukraine. It has almost since its beginnings been criticized for its ineffectiveness, recently criticized even by the UN Secretary General Guterres [BBC 2022]. It is, however, still one of the few permanent places where ideas of security are compared and debated internationally. UNSC has five permanent members in China, France, Russia (formerly the Soviet Union), United Kingdom, and the United States as well as several non-permanent members, selected for a period of two years [Council on Foreign Relations 2022]. It is one of six major organs of the UN and aims to provide international peace and security and friendly international relations. Members of the UN are supposed to carry out the decisions of UNSC, which is the only UN organ that has such direct powers. UNSC can also take direct action to ensure peace by undertaking mediation, through special envoys or dedicated missions. 17

To discover the debates on security in the United Nations, we downloaded the UNSC resolutions as PDFs from http://unscr.com/en, which offers additional metadata and bibliographic references. The earlier resolutions had to be recreated as JPEG images and then OCR’ed with Tesseract [Smith 2007]. For future use, we have also downloaded the metadata of the resolutions like related resolutions, cited resolutions, and UNSC composition, although we do not use these in this investigation. In the end, we have a corpus of ~2,600 documents from 1946 until the end of 2021. The first resolution discussed membership in UNSC in 1946, and the last one from 2021, the situation in Libya. 18

Whereas UNSC has been a body of the UN from the beginning, stable groupings for monitoring and promoting freedom and rights have followed later. They are also more diverse. To compare the language of security in the UNSC with the UN’s language of rights and freedoms, we have downloaded documents published by the United Nations Human Rights 19

Council (UNHRC) – its predecessor – and related organisations in the same timeframe from 1946 to end 2021. Since its establishment in 2005, UNHRC has aimed to promote and protect human rights, and it is formed by 47 members selected for a term of three years. It has replaced the United Nations Commission on Human Rights (UNCHR) and works closely with the Office of the High Commissioner for Human Rights (OHCHR). Like the resolutions, the human-rights documents are PDFs, but they are more varied. The first one in our collection is a yearbook from 1946, while the last one for 2021 is about “Promoting accountability in the Democratic People’s Republic of Korea” and a report of the United Nations High Commissioner for Human Rights. Whereas the UNSC has been criticized for its ineffectiveness and partisan politics, the UNHRC has been highly controversial at times for including members with a record of human rights abuses.

Overall, we have downloaded about 2,600 resolutions and about 2,550 human-rights related documents from the UN, which we merged into a common dataset and then removed documents that were either in the bottom 10% quantile or the top 10% quantile of document lengths. We ended up with 4,321 documents in a mixed corpus of UN-documents related to security and human rights. Our corpus contained about 11.5 million words with an average of about 2,700 words per document and about 380,000 unique tokens. For our further work with keywords, we decided to retain only adjectives, nouns and proper nouns, as these parts of speech are the most semantically informative and relevant for capturing the thematic content of the documents. In the earlier years of the UN, official records frequently included a mix of languages, reflecting its multilingual reporting standard. This led to non-English terms being interspersed throughout the text. To maintain linguistic consistency and improve keyword accuracy, we applied language filtering techniques to exclude non-English words from our analysis. This was particularly important for early documents where parallel translations were often merged in a single source. In later years, however, the UN began releasing separate language versions of documents, allowing us to focus exclusively on the English-language texts for those periods without loss of data fidelity. 20

3. Methodologies: A Comparative Approach to Keyword Extraction

Keywords are foundational to computational analysis in the digital humanities, yet the methods for deriving them are often applied with little critical reflection. To provide more critical reflection for mapping security and freedom, we will compare three distinct methods for keyword extraction: Term Frequency-Inverse Document Frequency (TF-IDF), KeyBERT, and a custom attention model. We chose these three specifically because they represent a significant historical path in text analysis, allowing us to examine the trade-offs between traditional statistical approaches, modern pre-trained language models, and more interpretable “white-box” deep-learning techniques. To explain our rationale, we will first discuss each method, situating it in relation to the broader evolution of these technologies before detailing our specific implementation. 21

Our first method, TF-IDF, is a classic and highly influential statistical approach. It is our baseline and builds on the work of [Luhn 1957], who assumed that a word's frequency (Term Frequency (TF)) correlates with its importance to a document. To offset the prominence of common words (like “the” or “and”), [Spärck Jones 1972] introduced Inverse Document Frequency (IDF), which measures a term's specificity by penalizing words that appear in many documents across a collection. The combination, TF-IDF, remains a reliable and widely used baseline in text processing. In the digital humanities research cited earlier, term frequencies are a dominant method, but the IDF component is generally ignored. By employing the full TF-IDF in our experiments, we establish a robust, interpretable baseline that accounts for both term frequency within a document and term specificity across our entire corpus. In terms of TF-IDF, keywords are those words that have a high TF-IDF value based on a pre-defined cut-off value. They are thus keywords within the context of a document and the collection the document is part of. 22

Our second method is based on the paradigm shift in text analysis with the rise of language and foundation models. Unlike TF-IDF, which treats words as independent units, these models learn from the sequence and context of words, capturing semantic relationships. This is achieved through transfer learning, where a model pre-trained on a massive text corpus can transfer its general “knowledge” of language to a specific task, such as keyword extraction. 23

For our second method, we use KeyBERT [Grootendorst 2022], a state-of-the-art tool that leverages transfer learning. 24

KeyBERT works by creating vector representations (embeddings) for the document and for candidate keywords using BERT [Devlin et al. 2019]. The candidates whose vectors are most similar to the document's overall meaning are selected as keywords. This allows KeyBERT to find keywords that reflect the context and semantics of the text, rather than just relying on frequency. Again, a cut-off value determines which words become keywords. A limitation, however, is that this process does not inherently relate keywords to a broader collection. To mitigate this and enable a better comparison with TF-IDF, we implement a two-step “guided KeyBERT” process. First, we run KeyBERT on the entire human rights and security collections to generate a smaller master list of salient keywords for each domain. Second, we use this list to “seed” the keyword extraction for each individual document, ensuring the resulting keywords are relevant not only to the document itself but also to its parent collection.

While powerful, foundation models like BERT are often criticized for their “black-box” nature; they provide results, but the internal processing is opaque [Dobson 2020]. To gain a better understanding of how a model determines importance, our third method deconstructs the language model and focuses on its core component: the attention mechanism. At the centre of the seminal paper “Attention is all you need” that introduces Transformers [Vaswani et al. 2017], this mechanism allows a model to weigh the importance of different words in a sequence when performing a task. 25

Specifically, we use a Hierarchical Attention Network (HAN) [Yang et al. 2016] to determine whether a document belongs to the “security” or “human rights” collection. During training, the attention mechanism learns to assign high weights to the words most decisive for making the correct classification of security- or human rights-related. Instead of treating every word equally, attention gives more weight to the words that are more relevant to the task. For example, suppose we want to classify the sentence, “The persecuted population is in danger” as related to freedom or security. The model learns to give higher attention to words like “persecuted” and “danger”, because those strongly indicate security problems, while giving less attention to filler words like “is” or “the.” By focusing on the most important words, attention helps the model get the classification more often right and even helps humans understand why the model made a certain decision. But researchers have debated whether attention weights really tell us how the model is doing [Jain and Wallace 2019] [Wiegreffe and Pinter 2019]. Our experiments will investigate this for digital humanities. 26

Attention is our third way of identifying the most discriminating keywords in comparison to the other two methods which are based on document and collection contexts. While we implemented this model ourselves, we built upon publicly available code repositories for core components. We use a bidirectional Gated Recurrent Network (GRU) to capture word context from both directions. On “Papers with Code”, HANs still rank number 4 for document classification tasks and are surpassed only by models that employ foundation language models, which is something we explicitly do not want to do because we are interested in deconstructing the work of language models [Papers With Code 2023]. 27

Using these three distinct methods, a statistical baseline (TF-IDF), a state-of-the-art pre-trained model (KeyBERT), and a purpose-built interpretable deep learning model (attention), we will generate three different sets of keywords for our corpora. In Section 5, we will analyse and compare these sets to understand how each method's architectural biases shape its conception of “important” terms. Subsequently, in Sections 6 and 7, we will use these computed keywords to analyse historical trends in security and freedom debates through prediction and topic modelling. 28

4. TF-IDF, KeyBERT and Attention Compared

We ran the three keyword experiments, TF-IDF, KeyBERT, and attention to extract the top 500 and top 1,000 most important keywords in the whole collection according to their median values. We chose median to balance out outliers. The example keywords presented in the tables below are illustrative samples selected from the highest-ranked terms for each method. As summarized in Table 1, all three methods differ heavily in their top keywords. 29

Metric	Top 500 Keywords	Top 1,000 Keywords
Keywords Unique to One Method	> 90%	~84%
Keywords Found by All Three Methods	1 ('treasurer')	1 ('treasurer')
Highest Pairwise Agreement	~60% (Attention & KeyBERT)	~58% (Attention & KeyBERT)

Table 1. Overview Top Keywords

Table 2 describes overall characteristics of the top keywords each method identified across all documents. 30

Method	Identified Top Keywords	General Character of Keywords
TF-IDF	wy, sie, tx, magnificent, happy, nail	Focuses on statistically rare words, including OCR errors like “wy”, unexpected emotions in official documents and other non-topical terms.
KeyBERT	malawi, reforms, syria, mali, zambia, recruitment, initiatives	Identifies recurring topics, countries of concern and required actions. Provides a good thematic summary.
Attention	birmingham, melbourne, veterans, published, charged, michigan, belfast	Pinpoints specific locations (cities) and words related to particular events, but lacks a clear link to the organization's overall concerns.

Table 2. Top Keywords Per Method

Table 3 details how each method performed when analyzing two specific UN documents. The UNSC resolution 80 (S/1469)[1] from 1953 deals with the relationship between India and Pakistan as well as the question of Kashmir and a cease fire along a specific line where Fleet Admiral Chester W. Nimitz played a mediating role. The UNHRC document “Working Group on the Universal Periodic Review” (A/HRC/WG.6/2/FRA/1) discusses the human rights situation in France in 2008. 31

Document	Method	Example Keywords	Interpretation/Effectiveness
UNSC Resolution India and Pakistan	TF-IDF	chester, implementing, fleet, administrator, sir	Zooms in on a specific, non-representative words like “fleet” admiral. Fails to summarize the document's main topic, with “India” missing.
	KeyBERT	procedure, soviet, disposition, council, pakistan	Provides a good summary of diplomatic procedure and an official document but misses the core border conflict.
	Attention	line, final, reason, order, fundamental democratic	Is more specific, identifying the “cease-fire line” as a key concept. The more generic words of diplomatic procedure to establish this order can be only found later in the top keywords.
UNHRC Report about Human Rights in France	TF-IDF	promptly, crafts, duration, convention, bible	Picks up on unusual phrases like a “prompt” reaction and specific details but does not provide a coherent summary.
	KeyBERT	compliance, legislation, discrimination, territories, humanitarian, refugees	Provides a strong summary of a human rights country report, capturing key themes as well as some specifics about French oversea territories.
	Attention	origin, principle, women, grant, courts, ethnic	Identifies specific, granular issues like discrimination by “origin”, rights for “women” and inability to action rights with “courts”.

Table 3. Analysing Example Documents

The differences in outputs illustrate the core logic of each method. TF-IDF operates on a purely statistical principle of rarity. It has no semantic understanding and elevates words that are infrequent across the entire UN corpus. This leads to highlighting rare terms without semantic understanding and elevating OCR errors (e.g., “wy”, “sie”) and unusual words (“magnificent”, “happy”) in Table 2. In Table 3, Admiral Chester W. Nimitz's first name, Chester, is rare across the collection, granting it a high inverted document frequency. While Chester can be found in an English dictionary, the name Nimitz has been filtered out because it does not appear there. TF-IDF finds what is syntactically surprising, not necessarily what is thematically important. 32

KeyBERT is semantic, because it uses embeddings to capture a document’s overall meaning, producing thematic keywords like “malawi”, “reforms”, and “humanitarian” in Table 2. This works well for summaries but can miss specific details, such as the cease-fire line in the India-Pakistan resolution (Table 3), because it favours the general idea. This is a specific detail that is crucial to the conflict but may be less representative of the document's overall diplomatic 33 language.

Attention models finally are task-specific and emphasize discriminative terms that best predict a document's status. Its keywords are not necessarily rare or thematic, but discriminative. This explains the prevalence of specific place names in Table 2 like "birmingham" and "belfast", or event-specific words like "veterans" and "charged". For the UNSC Resolution in Table 3, identifying line is especially relevant, since "cease-fire line" distinguishes security-related documents. The trade-off is that attention keywords might yield fragmented, event-specific signals rather than cohesive summaries. 34

"Treasurer" is the only keyword all three methods identify, which suggests that it is structural term. It is common enough for KeyBERT, rare enough for TF-IDF, and distinctive enough for the attention model. Nevertheless, we were surprized to not find more common terms. In summary, all three methods differ strongly in what they consider to be significant keywords, which emphasizes the need for more critical reflection about which method we choose. TF-IDF tends to overfocus both for the overall summary of the collection as well as specific documents, while KeyBERT provides a good summary but can miss out on specifics. The attention model is better at identifying relevant specifics. Next, we use keywords first to develop a relational view on freedom and security and then to retell historical stories with keywords through clustering and topic models. 35

5. Relational View on Freedom and Security

Predictive analytics can be used to deconstruct ontological assumptions by developing a relational view between its inputs and outputs [Blanke 2018]. In this section, predictive analytics helps us develop a relational view of the ideas of security and human rights, respectively. To this end, we have created a simple decision-tree algorithm to predict whether a document is about human rights or security. Decision trees are used as a tried-and-tested method for classification and are interpretable [Rudin 2019]. They learn simple decision rules based on the features of the data, which we will define as the keywords. Based on the features, the data is divided into branches of a tree where the outcomes of the predictions can be found in the bottom leaf nodes. In our case, we employed "Gini impurity", which evaluates possible splits of a tree into branches in terms of how "impurely" each feature/keyword would divide the data at a certain split. The split happens for the feature with the lowest Gini-value. 36

In our first set of experiments, we first selected the top 500 keywords for TF-IDF, BERT and attention. We then created three different tokenized versions of the collection, one each for the top 500 keywords of TF-IDF, BERT and attention. Finally, we ran three decision trees, one per method, to predict whether a document is a security or human-rights related one. 37

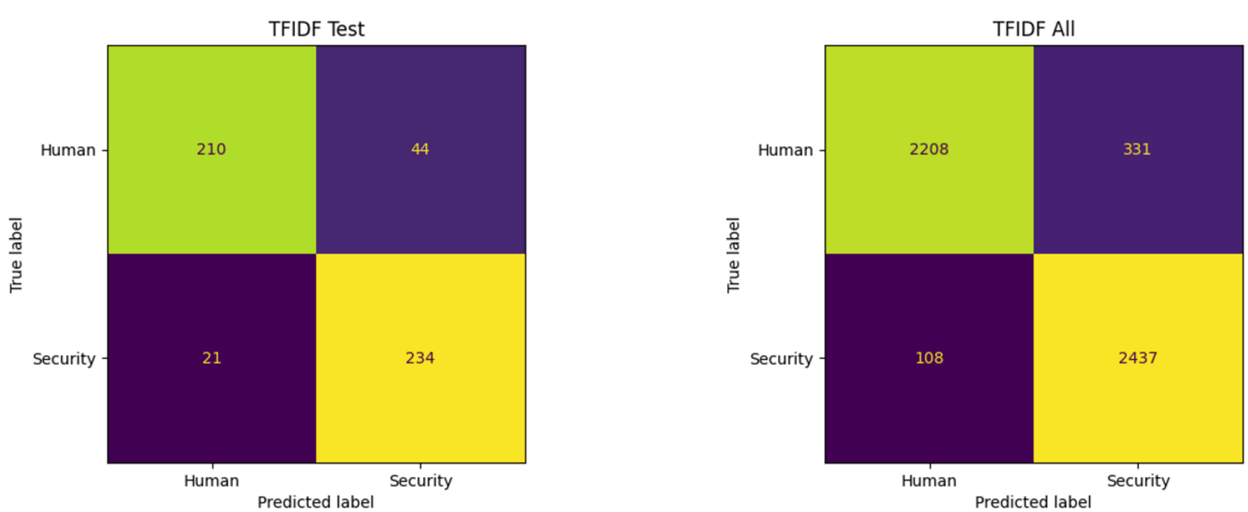

Figure 1. TF-IDF Confusion Matrix

Despite limiting the information to 500 keywords, we have achieved a very good 87% accuracy for the TF-IDF corpus 38
with an out-of-bag test collection of documents (see Figure 1). This shows the power of this simple, tried-and-tested method. But the model was slightly overfitting with 92% accuracy for the training collection and 91% accuracy when we ran the model on the whole collection. The model seems to struggle more with the human-rights documents. For the whole collection, 329 were misclassified as security-related with only 112 security documents wrongly assigned as human-rights related. That means that over 10% of the human-rights documents are missed. Overall, we are still surprized by such good results, given that the overall corpus contained over 252,000 features. It shows the power of simple, tried-and-tested analytics.

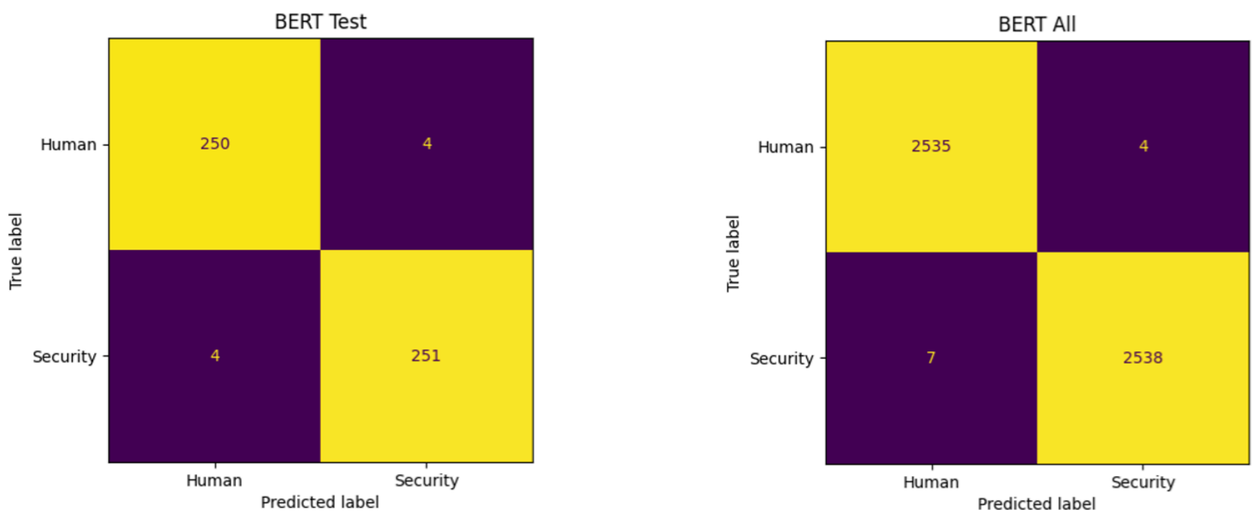

Figure 2. KeyBERT Confusion Matrix

If the results for TF-IDF are good, the results for KeyBERT have exceeded our expectations by far (see Figure 2). With 39
the out-of-bag test collection, we have reached an outstanding 98% accuracy with our simple decision tree. There is no notable difference in the model's ability to relate a document to either security or human rights. For the whole collection, only 4 human-rights related documents are misidentified and 7 security-related ones.

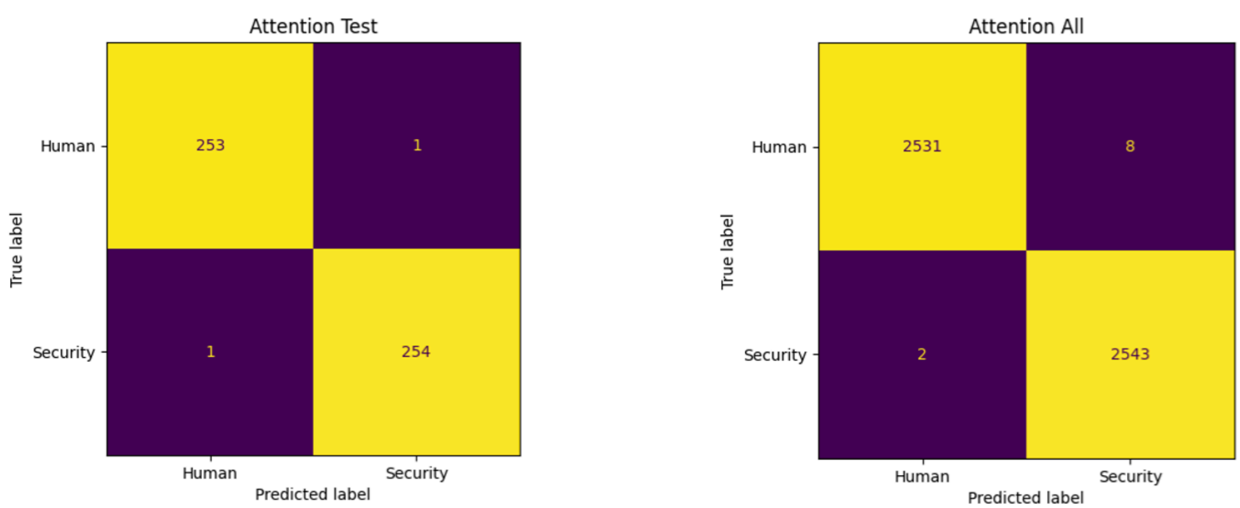

Figure 3. Attention Confusion Matrix

The results are a little more mixed for the attention-based keywords in Figure 3. Overall, we have also achieved near 40
perfect performance with only 10 documents missed in the total collection, which is about the same as for KeyBERT.

However, out of these, 8 human-rights-related documents are wrongly classified, while 2 security-related ones are also misclassified.

Looking at the missed predictions for KeyBERT and attention, it is first of all interesting that none of the documents missed by attention have also been misclassified by KeyBERT and vice versa. On further inspection, the attention model has done better than the initial numbers suggest. Seven of the eight misclassified human-rights-related documents are almost entirely illegible due to OCR and other character-translation issues. The remaining misclassified one is for some reason truncated. The attention model has shown the limitations of our own attempts to clean the collection and demonstrates how hard it is to operate at scale across such an international, complex set of historical documents. The two misclassified UNSC-resolutions on the other hand, are very close to the corpus of human rights. Resolution 876 (S/RES/876) from 1993 deals with the human rights violations in Georgia and humanitarian assistance. Resolution 568 (S/RES/568) from 1985 covers sanctioning the South African apartheid regime and dealing with refugees. The errors demonstrate how close the work across the UN can be. 41

The misclassifications with KeyBERT's keywords confirm this close relationship between the UN bodies. For instance, it believes wrongly that the "Report of the Secretary General on the human rights in the occupied Syrian Golan" (A/HRC/22/36) from 2012 to be about security, which is not surprising given that it deals with the direct resolution of the occupation of foreign territory by the Israeli military. UNSC resolution 1293 (S/RES/1293) is seen as about human rights, because it covers money in Iraq that can be used to improve local economic infrastructure. KeyBERT, however, does not pick up on any of the documents that are illegible, as it only picks up particular words and not their context. 42

In the future, we need to think about better methods to filter out problematic documents before we start our computational analysis. Nevertheless, both KeyBERT and the attention model perform well in detecting the often not well-defined distinction between human-rights and security language at the UN. They both also demonstrate – particularly through their misclassifications and other prediction struggles – how hard it is to separate out the ideas of freedom and security in the first place, supporting a relational view of both. 43

6. Keyword Stories

In this section, we want to turn the task of keywords around from identifying to employing them for in-depth historical and semantic analysis. For our final experiment, we want to find out what kind of keywords describe common issues that could define research narratives about the relation of freedom and security. How can we improve approaches by researchers to narrate the results of their modelling by combining keywords into narratives? To this end, we have chosen the attention model's keywords and two types of clustering in order to describe their joined appearance in human-rights and security documents. First, we create with the attention-model's keywords a simple frequency matrix for all the documents, which we have normalized with a z-score to provide an input matrix to the clustering. 44

Mean-shift clustering [Wu and Yang 2007] is our first method for this part. First, it calculates the distance between all the data points. Then, weights are added to penalize points further away. The data points are iteratively updated so that they are pushed close to each other. Mean-shift has an important advantage over other better-known clustering variants, as it does not require selecting the number of clusters in advance. Instead, it operates with a parameter called bandwidth, which can be automatically selected. This means we can use mean-shift to examine the diversity of the attention keywords. The larger the number of clusters, the more diverse the keywords. 45

For the UNSC resolutions, we have found over 180 clusters, while for the human-rights-related documents there are over 120. This is already interesting, as we would have expected the opposite to be the case as the security documents have been collected from a single source. Furthermore, according to (Bigo 2011), security is supposed to be a more closely defined issue than freedom. But this result should also not surprize us, as we have already seen that the security documents from the UNSC are about a wide range of issues. After all, it is the only UN body that meets regularly for a long time. That it is more diverse in terms of clusters thus rather confirms the insights from critical security studies [Bigo 2011] that security becomes a unifying element, bringing together diverse political issues. Most of the clusters only consist of one feature and should be seen as outliers, which should not be surprising given that both bodies deal with highly diverse global issues and hard to define ideas of freedom and security. 46

Each feature or keyword can also be assigned to one cluster using maximum likelihood. We can thus invert the keyword-generation and go back from the keywords to the documents. For instance, the following attention words are clustered together: “previously”, “form”, “accordingly”, “contain”, “measure”, and “yes”. Together they identify, among others, UNSC resolution 1454 (S/RES/1454) and UNSC resolution 2083 (S/RES/2083). Resolution 1454 from 30 December 2002 covers the sovereignty of Kuwait and its relationship to Iraq, while resolution 2083 is about threats to international peace and security caused by terrorist acts. However, these links are generally based on only very few keywords, because the attention-based keywords target specific information. 47

Another way of narrating the stories across UN documents using keywords is through topic models, which relate words automatically with each other without relying on existing structures. Topic models offer a way to automatically identify keyword relationships within documents, an approach we have previously shown to be effective [Blanke and Wilson 2017]. A common method, Latent Dirichlet Allocation (LDA), assumes each document is a mix of topics, but interpreting these topics can be subjective, like “reading tea leaves” [Chang et al. 2009]. As an alternative, seeded LDA allows researchers to guide the model by providing it with pre-defined “seed” keywords for certain topics [Jagarlamudi et al. 2012]. For instance, we could bias the term “data” toward a specific topic. 48

With seeded topic models, we can retell the history in our collection through a visualisation as in Figure 1. To this end, we have divided the top attention keywords into four topics: locations, actions, attributes, and issues. The first covers a range of locations and related activities that have particular importance for attention. The second is about actions or verbs that are stressed, while the attributes topic cluster stands for closer descriptions. The final topic on issues covers discussion areas that the model has paid particular attention to. Each topic consists of seeded attention keywords as well as words the topic model associates with them. Tuning the topic model, we get: 49

- *Locations*: pennsylvania, michigan, kingston, melbourne, birmingham, seattle, belfast, para, right, child, see, committee, recommend, convention, report
- *Actions*: international, right, human, published, entitled, recognized, included, charged, permitted, obtained, changed, trained, faced, allocated, declined
- *Issues*: everyone, females, veterans, stocks, radar, euros, genre, enquiries, security, council, resolution, government, report, arm, support
- *Attributes*: together, regularly, significantly, properly, plus, dozens, shall, article, state, person, act, may, right, law, court

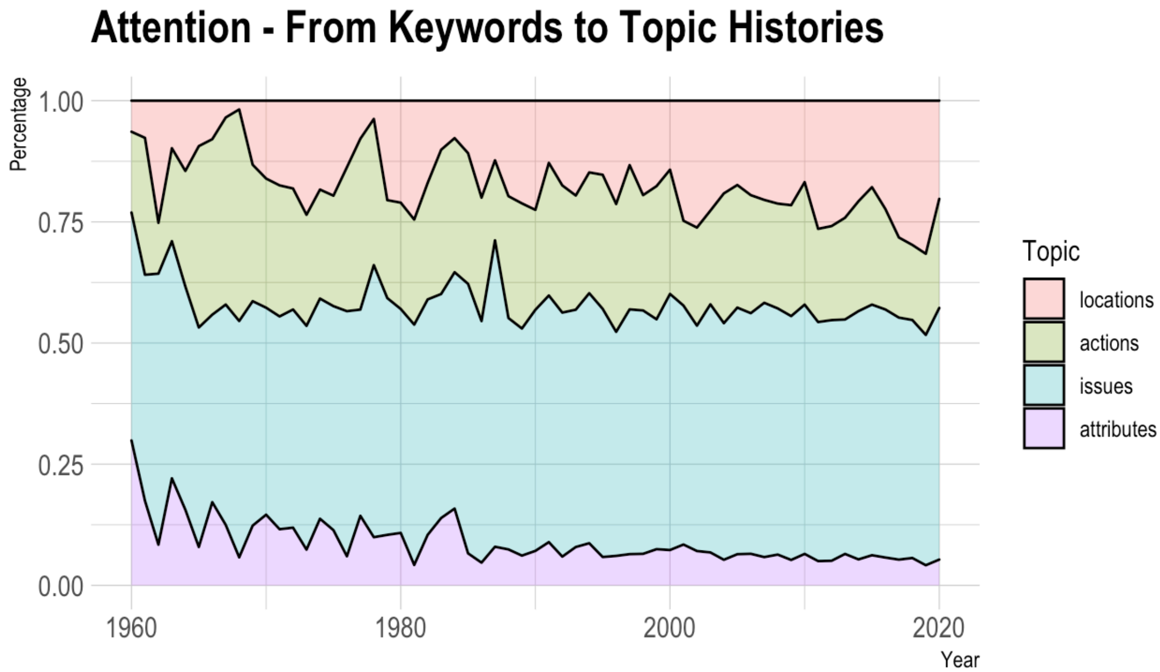

Figure 4. From Keywords to Topic Histories

In Figure 4, we see how locations grow in importance over the years, probably because more attention is paid to diverse global issues. Over time, smaller regions also receive attention. Attributes lose importance, but only marginally, while the topic issues stay fairly constant. Overall, the topic distributions seem to be evenly distributed over the years, providing a good visual summary of UN discussions. The topics are well chosen to summarize the UN-related discussions on security and human rights. 50

7. Conclusion

This article argues that the computational production of keywords should be further critically reflected upon. They should not be taken for granted. Their significance is threefold: They are essential not just for feeding computational models with textual data but also for analyzing their results. Furthermore, they align with the long-standing tradition in the critical humanities of using keywords to describe complex historical-cultural phenomena that escape clear definitions. These three dimensions should motivate a renewed critical focus on the role of keywords. 51

Despite different research practices agreeing on their importance for analyzing cultural issues and providing a baseline for new stories about the past, the origins of these keywords are often neglected, perhaps simply because they have been around for so long. Keywords have offered an easy, basic information unit for computational text analysis, with methods like TF-IDF being in popular use since the 1970s. This paper moves beyond this established technique to also explore how keywords are formed and used with two more recent methods: KeyBERT and attention models, which are a core element of the Transformer architecture that has revolutionized natural language processing. 52

The power of keywords as a special kind of map remains relevant for computational analysis, where they are increasingly seen as a tool for describing the contested definitions and categorizations central to humanities scholarship. In our case, we apply this approach to the complex and often fraught relationship between security, human rights, and freedom. To this end, we downloaded a corpus of United Nations documents from its human rights bodies and the Security Council. Our analysis demonstrates how extracting keywords with TF-IDF, BERT, and attention models allows us to describe the intricate intersections of these concepts at the UN. Crucially, these complex relations are also 53 expressed in the errors and struggles of the models themselves. Where our keyword methods struggle to cleanly differentiate a document, it is not a sign of technical failure. Instead, these moments reveal the conceptual closeness of security and human rights issues within the UN's discourse and the inherent difficulty of discriminating between them.

We tested three keyword extraction methods on the UN documents: TF-IDF, KeyBERT and attention-based. Their performance varied significantly by task. For summarizing content, TF-IDF proved least effective, often highlighting terms not central to the documents' main topics. In contrast, KeyBERT generated useful general summaries, while the attention model excelled at identifying specific relevant issues. This suggests TF-IDF (and its TF-only version) should no longer be the default method in digital humanities. However, when we used these keywords to classify documents as "security" or "human rights" with a decision tree, all three methods performed exceptionally well, despite selecting largely different keywords. KeyBERT and the attention model were almost perfect classifiers. This strong performance validates our intuition that keywords remain a powerful tool for digital humanities research, enabling us to effectively re-narrate the history of the UN. 54

Our research has made it increasingly clear that we must engage more critically with the concept of keywords and how they are generated, interpreted and used in computational analysis. There are numerous methods for both creating and reading keywords, ranging from statistical to deep-learning approaches, and each carries its own assumptions, strengths and limitations. These choices significantly shape what machines "read" from texts and, in turn, what we as researchers interpret as meaningful output. While this diversity can be challenging, it also presents an opportunity. The plurality of methods opens new avenues for inquiry, allowing us to ask different questions and explore texts from fresh perspectives. 55

Importantly, working with keywords allows us to integrate computational models into broader humanities research in a way that supports diverse narrative-formation, visualization and interpretive analysis. But we must be cautious. Deep-learning models, while powerful and promising, often introduce layers of complexity that hinder interpretability. It is essential that we retain the ability to critically engage with and explain these models and their outputs. Furthermore, the uneven availability of language models privileges certain data, and particularly English. Other languages are sidelined, including those with fewer resources or those that are not traditionally written in the Western sense. This article has aimed to interrogate the assumptions underlying the keyword arts, urging a more reflective stance on what we do with keywords, how we use them, when they are appropriate and just as importantly when they are not. 56

Acknowledgements 57

This research was partly funded by the European Research Council project *Deep Culture-Living with Difference in the Age of Deep Learning* (Grant no. 101141330). 58

Notes

[1] All UN documents are cited with their unique document symbol.

Works Cited

AI Now Institute 2021 *A New AI Lexicon*. Available at: https://medium.com/a-new-ai-lexicon (Accessed: 11 February 2023).

Amnesty 2023 Amnesty (2023) *Universal Declaration of Human Rights*, *Amnesty International*. Available at: https://www.amnesty.org/en/what-we-do/universal-declaration-of-human-rights/ (Accessed: 14 September 2023).

Arendt 1961 Arendt, H. (1961) "Freedom and politics", in A. Hunold (ed.) *Freedom and Serfdom: An Anthology of Western Thought*. Dordrecht: Springer Netherlands, pp. 191–217. Available at: https://doi.org/10.1007/978-94-010-3665-8_11 .

BBC 2022 BBC, (2022) "Ukraine: UN security council failed, says Guterres", *BBC News*. Available at: https://www.bbc.com/news/av/world-61262756 (Accessed: 26 May 2022).

Bauman 2000 Bauman, Z. (2000) *Community: Seeking Safety in an Insecure World*. Oxford: Polity.

Bennett et al. 2013 Bennett, T., et al. (2013) *New Keywords: A Revised Vocabulary of Culture and Society*. Hoboken, NJ:

John Wiley & Sons.

Bigo 2011 Bigo, D. (2011) “Delivering liberty and security? the reframing of freedom when associated with security”, in D. Bigo and R. Walker (eds) *Europe’s 21st Century Challenge*. London: Routledge.

Blanke 2018 Blanke, T. (2018) “Predicting the past”, *Digital Humanities Quarterly*, 12(2). Available at: https://dhq.digitalhumanities.org/vol/12/2/000377/000377.html.

Blanke and Wilson 2017 Blanke, T. and Wilson, J. (2017) “Identifying epochs in text archives”, in *2017 IEEE International Conference on Big Data (Big Data)*, pp. 2219–2224. Available at: https://doi.org/10.1109/BigData.2017.8258172.

Blanke et al. 2024 Blanke, T., et al. (2024) “A peek inside two black boxes-an experiment with explainable artificial intelligence and IPCC leadership”, International Journal Digital Humanities 6, pp. 45–69. Available at: https://doi.org/10.1007/s42803-023-00080-z

Burgess and Matamoros-Fernández 2016 Burgess, J. and Matamoros-Fernández, A. (2016) “Mapping sociocultural controversies across digital media platforms: One week of #gamergate on Twitter, Youtube, and Tumblr”, *Communication Research and Practice* [Preprint]. Available at: https://doi.org/10.1080/22041451.2016.1155338 (Accessed: 13 September 2023).

Cameron 2006 Cameron, D. (2006) “Balancing freedom and security - A modern British Bill of Rights”, *The Guardian*, 26 June. Available at: https://www.theguardian.com/politics/2006/jun/26/conservatives.constitution (Accessed: 14 September 2023).

Chang et al. 2009 Chang, J. et al. (2009) “Reading tea leaves: How humans interpret topic models”, in *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper_files/paper/2009/hash/f92586a25bb3145facd64ab20fd554ff-Abstract.html (Accessed: 23 July 2023).

Council on Foreign Relations 2022 Council on Foreign Relations, (2022) *The UN Security Council*, *Council on Foreign Relations*. Available at: https://www.cfr.org/backgrounder/un-security-council (Accessed: 26 May 2022).

Devlin et al. 2019 Devlin, J., et al. (2019) “BERT: Pre-training of deep bidirectional transformers for language understanding”, in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. NAACL-HLT 2019, Minneapolis, Minnesota: Association for Computational Linguistics, pp. 4171–4186. Available at: https://doi.org/10.18653/v1/N19-1423.

Dobson 2020 Dobson, J. (2020) “Interpretable outputs: Criteria for machine learning in the humanities”, *Digital Humanities Quarterly*, 15(2). Available at: https://dhq.digitalhumanities.org/vol/15/2/000555/000555.html.

Grootendorst 2022 Grootendorst, M. (2022) *KeyBERT*. Available at: https://maartengr.github.io/KeyBERT/ (Accessed: 14 September 2023).

Hobbes 2017 Hobbes, T. (2017) *Leviathan*. Edited by C. Brooke. Penguin Classics.

Jagarlamudi et al. 2012 Jagarlamudi, J., et al. (2012) “Incorporating lexical priors into topic models”, in *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. USA: Association for Computational Linguistics (EACL ’12), pp. 204–213.

Jain and Wallace 2019 Jain, S. and Wallace, B.C. (2019) “Attention is not explanation”, arXiv. Available at: https://doi.org/10.48550/arXiv.1902.10186.

Jockers and Underwood 2015 Jockers, M.L. and Underwood, T. (2015) “Text-mining the humanities”, in S. Schreibman, R. Siemens, and J. Unsworth (eds.) *A New Companion to Digital Humanities*. John Wiley & Sons, Ltd, pp. 291–306. Available at: https://doi.org/10.1002/9781118680605.ch20.

Luhn 1957 Luhn, H.P. (1957) “A Statistical approach to mechanized encoding and searching of literary information”, *IBM Journal of Research and Development*, 1(4), pp. 309–317. Available at: https://doi.org/10.1147/rd.14.0309.

Oman 2010 Oman, N. (2010) “Hannah Arendt”s “Right to Have Rights”: A philosophical context for human security’, *Journal of Human Rights*, 9(3), pp. 279–302. Available at: https://doi.org/10.1080/14754835.2010.501262.

Papers With Code 2023 Papers With Code (2023) *Hierarchical Attention Networks for Document Classification*. Available at: https://paperswithcode.com/paper/hierarchical-attention-networks-for-document (Accessed: 15 September 2023).

Peters 2016 Peters, B. (2016) *Digital Keywords: A Vocabulary of Information Society and Culture*. Princeton: Princeton University Press.

Rogers 2013 Rogers, R. (2013) *Digital Methods*. Cambridge, MA: MIT Press.

Rogers 2023 Rogers, R. (2023) *Doing Digital Methods*. London: SAGE Publications Ltd.

Rudin 2019 Rudin, C. (2019) “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead”, *Nature Machine Intelligence*, 1(5), pp. 206–215. Available at: https://doi.org/10.1038/s42256-019-0048-x.

Smith 2007 Smith, R. (2007) “An overview of the Tesseract OCR Engine”, in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, pp. 629–633. Available at: https://doi.org/10.1109/ICDAR.2007.4376991.

So 2020 So, R.J. (2020) *Redlining Culture: A Data History of Racial Inequality and Postwar Fiction*. New York, NY: Columbia University Press.

Spärck Jones 1972 Spärck Jones, K. (1972) “A statistical interpretation of term specificity and its application in retrieval”, *Journal of Documentation*, 28(1), pp. 11–21. Available at: https://doi.org/10.1108/eb026526.

Striphas 2015 Striphas, T. (2015) “Algorithmic culture”, *European Journal of Cultural Studies*, 18(4–5), pp. 395–412. Available at: https://doi.org/10.1177/1367549415577392.

Thylstrup et al. 2021 Thylstrup, N.B., et al. (2021) *Uncertain Archives: Critical Keywords for Big Data*. Available at: https://doi.org/10.7551/mitpress/12236.001.0001.

Underwood 2019 Underwood, T. (2019) *Distant Horizons – Digital Evidence and Literary Change*. Illustrated edition. Chicago: University of Chicago Press

United Nations 1945 United Nations (1945) *United Nations Charter*. United Nations. Available at: https://www.un.org/en/about-us/un-charter/full-text (Accessed: 14 September 2023).

Vaswani et al. 2017 Vaswani, A., et al. (2017) “Attention is all you need”, in *Advances in Neural Information Processing Systems*. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html (Accessed: 11 February 2023).

Wiegreffe and Pinter 2019 Wiegreffe, S. and Pinter, Y. (2019) “Attention is not not explanation”. arXiv. Available at: https://doi.org/10.48550/arXiv.1908.04626.

Willaims 1963 Williams, R. (1963) *Culture and society, 1780-1950*. Harmondsworth: Penguin (Pelican books).

Williams 1973 Williams, R. (1973) *Keywords: A Vocabulary of Culture and Society*. New York, NY: Oxford University Press.

Wu and Yang 2007 Wu, K.L. and Yang, M.S. (2007) “Mean shift-based clustering”, *Pattern Recognition*, 40(11), pp. 3035–3052. Available at: https://doi.org/10.1016/j.patcog.2007.02.006.

Yang et al. 2016 Yang, Z., et al. (2016) “Hierarchical attention networks for document classification”, in *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. NAACL-HLT 2016, San Diego, California: Association for Computational Linguistics, pp. 1480–1489. Available at: https://doi.org/10.18653/v1/N16-1174.

Recommendations

DHQ is testing out three new article recommendation methods! Please explore the links below to find articles that are related in different ways to the one you just read. We are interested in how these methods work for readers—if you would like to share feedback with us, please complete our short evaluation survey. You can also visit our documentation for these recommendation methods to learn more.

SPECTER Recommendations

Below are article recommendations generated by the SPECTER model:

1. Developing Computational Models for Formalizing Concepts in the British Colonial India Corpus, 2023, Shanmugapriya T, University of Toronto Scarborough
2. Introduction: The Politics and Ethics of Naming the Names of Enslaved People in Digital Humanities Projects, 2025, Walter Hawthorne, Michigan State University; Richard Roberts, Stanford University; Fatoumata Seck, Stanford University; Rebecca Wall, Loyola Marymount University
3. Welcome to Digital Humanities Quarterly, 2007, Julia Flanders, Brown University; Wendell Piez, Mulberry Technologies, Inc.; Melissa Terras, University College London

4. Digital Humanities: On Finding the Proper Balance between Qualitative and Quantitative Ways of Doing Research in the Humanities, 2013, Helle Porsdam, University of Copenhagen
5. Circling around texts and language: towards pragmatic modelling in Digital Humanities, 2016, Arianna Ciula, University of Roehampton; Cristina Marras, Istituto per il Lessico Intellettuale Europeo e Storia delle Idee, Consiglio Nazionale delle Ricerche

DHQ Keyword Recommendations

Below are article recommendations generated by DHQ Keywords:

1. Developing Geographically Oriented NLP Approaches to Sixteenth–Century Historical Documents: Digging into Early Colonial Mexico, 2020, Diego Jiménez–Badillo, Museo del Templo Mayor, Instituto Nacional de Antropología e Historia; Patricia Murrieta–Flores, Digital Humanities Hub–History Department, Lancaster University; Bruno Martins, The Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento em Lisboa, INESC–ID, University of Lisbon; Ian Gregory, Digital Humanities Hub–History Department, Lancaster University; Mariana Favila-Vázquez, Museo del Templo Mayor, Instituto Nacional de Antropología e Historia; Raquel Liceras-Garrido, Digital Humanities Hub–History Department, Lancaster University
2. Mining Eighteenth Century Ontologies: Machine Learning and Knowledge Classification in the Encyclopédie, 2009, Russell Horton, Digital Library Development Center, University of Chicago; Robert Morrissey, University of Chicago; Mark Olsen, ARTFL Project, University of Chicago; Glenn Roe, ARTFL Project, University of Chicago; Robert Voyer, Powerset
3. Theorizing Connectivity: Modernism and the Network Narrative, 2011, Wesley Beal, Lyon College; Stacy Lavin, Independent Scholar
4. From close listening to distant listening: Developing tools for Speech-Music discrimination of Danish music radio, 2021, Iben Have, Aarhus University; Kenneth Enevoldsen, Aarhus University
5. Probing Through Iranian Architectural History Within the Framework of an Ontology Development Process, 2021, Dena Shamsizadeh Hayatdavoodi, Shahid Beheshti University; Niloofar Razavi, Faculty of Architecture & Urban Planning, Shahid Beheshti University; Mehrdad Qayyoomi Bidhendi, Faculty of Architecture & Urban Planning, Shahid Beheshti University

TF-IDF Recommendations

Below are article recommendations generated by the TF-IDF Model:

1. A Text Network Analysis of Discursive Changes in German, Austrian and Swiss New Year’s Speeches 2000-2021, 2022, Kimmo Elo, University of Turku
2. Unveiling the Critical Nexus of Data Preprocessing and Transparent Documentation for Result Quality and Reproducibility in Digital History, 2025, Clodomir Santana, Tadeusz Manteuffel Institute of History, Polish Academy of Sciences, Poland and University of California, Davis, USA; Demival Vasques Filho, Centre for Contemporary and Digital History, University of Luxembourg, Luxembourg; Michał Bojanowski, Department of Quantitative Methods and Information Technology, Kozminski University, Poland and Department of Anthropology, Autonomous University of Barcelona, Spain; Agata Błoch, Tadeusz Manteuffel Institute of History, Polish Academy of Sciences, Poland
3. Exploratory Search Through Visual Analysis of Topic Models, 2017, Patrick Jähnichen, Machine Learning Group, Humboldt-Universität zu Berlin; Patrick Oesterling, Image and Signal Processing Group, Leipzig University, Germany; Gerhard Heyer, Natural Language Processing Group, Leipzig University, Germany; Tom Liebmann, Image and Signal Processing Group, Leipzig University, Germany; Gerik Scheuermann, Image and Signal Processing Group, Leipzig University, Germany; Christoph Kuras, Natural Language Processing Group, Leipzig University, Germany
4. Making Sense of the Emergence of Manslaughter in British Criminal Justice, 2026, Tim Hitchcock, Professor Emeritus of Digital History, University of Sussex; William J. Turkel, Professor of History, The University of Western Ontario
5. Expertise vs. statistics. A qualitative evaluation of three keyness measures (logarithmic Zeta, Welch’s t-test, and Log-likelihood ratio test) applied to subgenres of the French novel, 2025, Julia Röttgermann, University Trier, Germany; Keli Du, University of Trier, Germany; Julia Havrylash, University of Trier, Germany; Christof Schöch, University of Trier, Germany