

The semantic taxonomy of vulnerability: a bibliometric study using natural language processing

Received: 13 January 2026

Accepted: 22 July 2026

Published online: 31 July 2026

Cite this article as: Vidal Marciel, J., Berástegui Pedro-Viejo, A., De Nicolás, V.L. The semantic taxonomy of vulnerability: a bibliometric study using natural language processing. *Humanit Soc Sci Commun* (2026). <https://doi.org/10.1057/s41599-026-08513-0>

Javier Vidal Marciel, Ana Berástegui Pedro-Viejo & Víctor Luis De Nicolás

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

The semantic taxonomy of vulnerability: a bibliometric study using natural language processing

Abstract

Vulnerability has become a central but conceptually diffuse term across the human and social sciences, used to designate heterogeneous risks, conditions, and causal processes. This conceptual dispersion hinders cumulative theorizing and limits the research's capacity to inform policy and intervention. This article presents a reproducible, multi-resolution mapping of how vulnerability is organized in the scholarly literature, based on a Natural Language Processing pipeline that combines contextual keyword extraction, static keyword aggregation, data-driven dimensionality reduction, and hierarchical agglomerative clustering with Ward linkage. A Scopus search conducted in May 2025 identified 12,450 English-language journal articles in social science, psychology, arts and humanities, economics, and business with “vulnerability” in the title. The resulting dendrogram is reported across the full range $k = [2, 12]$, and a further analysis inspects the specific $k = 7$ cut, at which seven semantic clusters emerge: psychopathology, health, rights and violence, climate and rural livelihood, floods and natural disasters, security and infrastructure, economic poverty, and pandemic and migration. Cluster quality is assessed through a multi-metric validation envelope that combines geometric fit with topic coherence. Three VosViewer keyword co-occurrence maps trace how the lexical landscape evolves across historically meaningful periods. Since 2000, climate- and disaster-related vulnerability has expanded sharply, while psychopathology-oriented research has lost relative prominence. The contribution is positioned as a methodologically transparent instrument for critical inquiry into how a contested concept is organized and how that organization changes over time.

Keywords: vulnerability, bibliometric analysis, natural language processing, word embeddings

1. Introduction

The term vulnerability is an elusive concept encompassing multiple connotations and meanings across a wide range of disciplines. In the scope of human and social science, the term is currently used to represent adverse phenomena, conditions, or causal processes without clear theoretical boundaries (Brown et al., 2017; Gilodi et al., 2022; Mackenzie, 2013). This high level of abstraction and divergence obstructs building cumulative knowledge, sharing findings across areas, and reaching an actionable consensus. Consequently, it hinders the capacity of research to guide policymaking, intervention design, and program evaluation (Brown, 2011). Given the exponentially increasing size of the literature available, traditional efforts to bring conceptual synthesis through manual reviews face feasibility issues. While bibliometric analyses that leverage data-driven methods are available, current methodological alternatives fail to mitigate key threats to scientific validity. This paper does not attempt to adjudicate between the competing definitions of vulnerability, which can be read in the conceptual literatures cited through this Introduction. Instead, it proposes a reproducible, multi-resolution mapping of how vulnerability is organized across fields that use the term, built on a Natural Language Processing (NLP) pipeline and intended as an inquiry into interdisciplinary conceptual interaction and meaning transformations.

1.1. Conceptual ambiguity around vulnerability.

The concept of vulnerability has gained considerable attention in recent times, and its use has spread both in academia, politics, and popular culture (Kuhlicke et al., 2023; Nungesser and Schirgi, 2024). Being vulnerable has become a broad category that includes infinite types of risks, adversities, and expressions of suffering. Depending on the discipline, facing a vulnerability can range from experiencing the adverse effects of climatic disasters (Ishtiaque, 2022) to mental health problems (Lajoie et al., 2020), medical conditions (De Groot, 2019), or socioeconomic exclusion (Skoglund, 2024). Similarly, it has been linked with a plethora of other contested concepts such as resilience (Lindbom et al., 2015; Proag, 2014), coping or adaptation (Hofmann et al., 2011; Schoorman, 2025), poverty (Aristegui et al., 2017; Moran et al., 2013), or marginalization (Adger, 2006). While an overarching theory of vulnerability may simply not be possible given its interdisciplinary spread, a threat exists of the term becoming empty (Virokannas et al., 2020).

Initial divergence can be seen around the question of who is vulnerable. In some views, it is understood as an inherent human attribute that all individuals face, to varying extents, in relation to omnipresent hazards (Fineman 2008; Kottow, 2005). Other authors, however, interpret the term as a category that can only be attributed to marginalized or disadvantaged groups in a society. Hence, “the vulnerable” is a specific group of people that can fluctuate depending on which type of adversity is

observed. This perspective is frequently held in government policy and aid prioritization systems (Gilodi et al., 2022; Turner, 2025).

Vulnerability is also ambiguous in relation to how it is operationalized in causal models. Most literature treats the concept in a predictive sense, inherently tied to the idea of a potential risk. This view adopts an equational approach, where vulnerability is the product of the severity of an adversity, combined with the degree of exposure to it, and adaptive capacity to its consequences (Berrouet et al., 2018; Schneiderbauer et al., 2017). There is great confusion here on how adaptability interacts with other generic terms such as capability, resilience, coping, fragility, and sensitivity (Jackson et al., 2012; Miller et al., 2010; Valdés, 2021). Other perspectives differ, merging vulnerability with *ex post* phenomena such as poverty and social exclusion (Arellano y Sánchez, 2019; McIlwaine, 2002; Timmer 2021). While all these views are not necessarily incompatible, they certainly prompt confusion around what being vulnerable is.

Global institutions, policymakers, and social organizations are increasingly using this term to anchor the purpose of programs and social initiatives. Vagueness around how to identify who is vulnerable or not, and what problems they face, creates the risk of making such efforts inefficient. Effectively mapping and classifying different views on vulnerability across the literature could bring clarity to the conceptual discussion. The diffusion of vulnerability across disciplines evidences an active debate about how multidisciplinary expert knowledge travels into policymaking, scientific advising, and public communication (Bijker et al., 2009; Resnyansky, 2019). The term has been mobilized in distinctive and partly incompatible ways across migration and social work (Brun, 2025; Leboeuf et al., 2025; Virokannas et al., 2020), social policy (Brown et al., 2017), ethics (Mackenzie et al., 2014), health and health-emergency response (de Groot et al., 2019; Hammer et al., 2019), disaster studies and disaster-management practice (Gibb, 2018; Uekusa et al., 2025), and responses to climate change (Oliver-Smith, 2016). The disciplinary diversity surfaced by these literatures is the phenomenon to be described rather than a deficit to be resolved: the present paper proposes a method for surfacing this diversity at scale, leaving its theoretical adjudication to the substantive disciplines.

1.2. Research synthesis and bibliometric analysis.

Efforts that review and consolidate trends across research are crucial for building coherent theories that lead to relevant conclusions. Research synthesis methods include a wide array of procedures that systematize the gathering, comparison, and summarization of existing literature (Maynard, 2025). Each type of review adapts to the nature of the phenomenon or concept analyzed, the aim intended by the authors, or the contextual constraints during research (Ghosh and Choudhury, 2025). Some approaches, like systematic reviews, are directed towards integrating evidence (Uttley et al., 2023), while others, such as

scoping reviews, attempt to map the structure and nature of emerging bodies of knowledge (Mak and Thomas, 2022).

An exponential growth in the number of peer-reviewed publications has made it substantially harder for authors to make sense of the literature (Hanson et al., 2024). A Scopus search for sources in social science and psychology about vulnerability reveals that, on average, in 2005, a paper was published approximately every three days, while in 2013 this had risen to an article per day. In 2025, four papers with the term “vulnerability” on their title are indexed each day on Scopus. The number of sources published since 2020 now surpasses the total from the preceding 56 years. Although research synthesis methods have gained prominence and aided scientists in navigating the overwhelming evidence, traditional procedures can’t possibly cope with the pace of publication (Affengruber et al., 2024; Tóth et al., 2024). This is especially true for those qualitative synthesis methods designed for clarifying contested concepts such as vulnerability. Narrative, scoping, and concept reviews heavily rely on manual selection and screening of sources. Additionally, identifying common or diverging patterns in the literature and mapping alternative views on a similar idea is, in most cases, a subjective task (Jaakkola, 2020; Pollock et al., 2024; Sukhera, 2022). Features such as these make research reviews particularly labor-intensive, as well as barring replicability and methodological transparency. While in more compact and restrained bodies of knowledge, these issues may not arise, wider areas of research with soaring rates of publication can overcome the scalability of such methods.

In response to this widespread reality, statistical and data-driven alternatives have emerged as complementary tools able to bring clarity to the overabundance of literature (Iqbar, 2023; Kozłowski, 2021). Bibliometric analysis (BA) comprises a set of quantitative techniques that model the thematic content and publication dynamics of large quantities of bibliographical data (Donthu et al., 2021b). To consolidate what steps, variables, and hypotheses are usually employed in BA, Passas (2024) outlines two main approaches: performance analysis to evaluate the output, impact, and level of collaboration between authors and institutions; and science mapping techniques that seek to understand the thematic structure of a literature database.

Although bibliometric analysis has a relatively long history (De Solla, 1965; Kessler, 1963), present conditions bring new incentives and create a more urgent demand. The structured digitalization of most scientific knowledge, the expansion of information systems, as well as the creation of specific software solutions (Aria and Cuccurullo, 2017; Chen, 2006; Kara et al., 2025; Pereira et al., 2023; Ruiz-Rosero et al., 2019; Van Eck and Waltman, 2010), have fueled the growth of BA papers.

Thematic or semantic modelling requires text as raw material for analysis. In bibliometric methodological designs, author keywords predominate as the appropriate variable to represent the content of each source (Chen and Xiao, 2016; Chen et al., 2024). Most examples of BA perform statistical calculations on keywords to infer thematic trends, similarities, or disparities between sources, or to propose content clusters (Alam et al., 2021; Baker et al., 2020; Comerio and Strozzi, 2019; Donthu et al., 2021; Kim and Kim, 2021; Kumar et al., 2021; Tun and Khairuddin, 2024). This method, however, brings substantial limitations that represent a threat to empirical validity. First, there are no universal guidelines that structure the choice of what keywords better represent a peer-reviewed article. Consequently, the selection process is purely subjective and has no transparency or replicability. Second, internet browsers are currently the main tool for accessing scientific knowledge. Search engines use proprietary and non-accessible algorithms to arrange website recommendations, and optimization techniques can be used to appear among the first results. Because being placed on top can substantially boost article views and citations, hence contributing to the author's career and reputation, a conflict of interest is present in keyword selection. There is a considerable incentive for selecting optimized terms rather than those that better represent the content of the paper. (Brouwer and Hollenbach, 2022; Corrin et al., 2022). Finally, even if an objective and harmonious criteria for determining keywords is employed, manually classifying or grouping terms to compose topics remains a subjective task. Without the proper tools, assumptions on how a topic differs or can be associated with another can't be measured. Therefore, no scientific conclusions on the intellectual structure of a scientific field can be adequately put forward.

Because these issues are not exclusive to BA and transcend many other tasks that require text classification, it has been a traditional use case for many Natural Language Processing (NLP) techniques. For the issue here presented, a method for linguistic representation and modelling called word embeddings is introduced.

1.3. Word embeddings for bibliometric analysis.

Natural Language Processing is a scientific field that develops computational methods for modelling unstructured text or speech data to encode, model, and generate language (Cheng et al., 2025). All human language has, according to traditional assumptions (Bloom and Lahey, 1978), three main components: content, form, and use. Formal aspects like syntax or morphology can be defined by an objective and determined set of rules, and therefore, pioneer NLP models worked with such structural properties (Weizenbaum, 1966; Winograd, 1972). Semantics, however, is much harder to code into a definite program. Human symbols are created, transformed, and diluted through emergent collective processes without an explicit and deterministic behavior. Word embeddings address this core issue in computational linguistics.

The central idea proposed, tracing back to the 1950s (Firth, 1957; Harris, 1954; Osgood et al., 1957), is that words with similar meaning tend to appear in similar contexts. Therefore, a probability distribution of how often two different words occur in a similar group of sentences (ie, context) allows a numerical estimation of how semantically close they are (Jurafsky and Martin, 2025; Satvika et al., 2021). Additionally, this value for lexical proximity can be understood geometrically as a distance, positioning words in a space where those with similar meaning are “closer” (Kornai, 2023). Vector semantics is the most common approach to geometrical meaning representation, and word embeddings (WE) are the multidimensional vectors that position language in this space. Vectors are usually extracted by training the model with massive amounts of text data, making the probabilistic distributions sufficiently adequate to mimic human meaning (Arseniev-Koehler, 2024). A more intuitive explanation and visualization of WE can be found in Hessani (2025).

Using WE to model the semantic structure of large amounts of scientific literature has substantial potential for bibliometric analysis. Existing work spans several distinct tasks, each of which differs in important ways from the pipeline proposed here. Thesaurus construction from biomedical corpora aggregates terms into a controlled vocabulary but does not produce a hierarchical map of how a contested concept is structured (Kawamura et al., 2016). Document-level similarity methods such as Paper2vec (Tian and Zhuo, 2017) embed full articles for retrieval and recommendation, rather than identifying internal thematic structure. Citation-context embedding methods (Berger et al., 2017; Hourrane et al., 2018; Singhal et al., 2021; Thijs, 2020) map papers by the contexts in which they are cited, addressing literature search rather than topical organization. Combinations of WE with flat clustering produce thematic groups but typically pre-specify the number of clusters and provide no nested structure (Chen et al., 2021; Onan, 2019; Zhang et al., 2018). Citation-intent classification (Roman et al., 2020) and novelty measurement (Shibayama et al., 2021) operate on within-paper content. Author-trajectory analysis (Lombardo et al., 2022) and temporal trend analysis (Fijačko et al., 2024; Odden, 2024) target authorship dynamics and chronological evolution rather than the structural organization of the literature. The contribution of the present paper is the integration of contextual keyword extraction, static keyword aggregation, hierarchical clustering with explicit cluster-transition tracing, and multi-metric validation, applied to a contested social-scientific concept rather than to a domain-specific or task-specific corpus.

Applications of WE methods to bibliometric analysis remain scarce relative to the potential that such methods open. The present paper applies WE to a contested social-scientific concept (vulnerability) and uses the resulting structure to ask how the concept is organized and how that organization has changed over time. Two specific research questions guide the analysis:

1. What themes in the vulnerability literature can be identified through a combination of WE with clustering methods?
2. Can historic trends be identified in how vulnerability is configured across the modelled corpus?

2. Methods

To gather a sufficient sample of bibliographical records, a query was executed on Scopus in May 2025 using the following search string: ‘title ("vulnerability") and (limit-to (subjarea, "soci") or limit-to (subjarea, "psyc") or limit-to (subjarea, "arts") or limit-to (subjarea, "econ") or limit-to (subjarea, "busi")) and (limit-to (doctype, "ar")) and (limit-to (language, "english"))’.

A total of 12,911 sources were exported and processed using Python 3. The sample was reduced to 12,450 after dropping records without a registered abstract. The pipeline was executed in Google Colaboratory, a cloud-based Python notebook environment, with GPU acceleration (NVIDIA T4) used for the keyword-extraction and embedding stages at no cost. The workflow consists of four main stages: (a) keyword extraction, (b) keyword embedding, (c) dimensionality reduction, and (d) hierarchical agglomerative clustering; followed by a multi-metric validation layer. Fig. 1 provides a diagram of the complete process.

2.1. Keyword extraction

Author-supplied keywords have been argued unreliable for bibliometric validity for three reasons: (a) there is no universal guideline for how authors select keywords for an article, so the selection is largely opaque; (b) keyword selection is influenced by indexing-strategy and search-engine optimization incentives that are not aligned with conceptual representativeness; and (c) the number of keywords supplied per article is inconsistent across journals and authors. The pipeline therefore replaces author-supplied keywords with automatically extracted ones.

Keyword extraction is performed with KeyBERT (Grootendorst, 2020), which selects keywords by computing cosine similarities between candidate-term embeddings and the embedding of the surrounding text. The sentence-embedding underlying model is "all-mpnet-base-v2" (Song et al., 2020), a contextual transformer (Vaswani et al., 2017) trained for sentence-level semantic comparison. The contextual character of this model is what makes the extraction sensitive to abstract content: candidates are evaluated against the meaning of the abstract in which they appear, not against a static lexical neighborhood.

Five keywords are retained per abstract. This number balances two pressures: enough terms to characterize an abstract beyond the idiosyncrasy of any single keyword, but few enough to keep each article's lexical profile concentrated rather than diffuse. The five terms are five semantically representative

terms under the cosine-similarity criterion. Variants of the focal concept ("vulnerability", "vulnerable", "vulnerabilities") are included in a custom stopwords list at the extraction stage, so that they cannot occupy keyword slots that would otherwise be informative. This filtering is implemented upstream rather than as a post-hoc display rule.

2.2. Keyword embedding

Each of the five keywords per article is embedded using a pretrained GloVe model (Pennington et al., 2014) trained on the Dolma 2024 corpus (Soldaini et al., 2024), which contains approximately 220 billion tokens of contemporary text (Carlson et al., 2025). The embedding dimensionality is 300. Each article is then represented as the mean of its five keyword vectors.

The architectural choice to use a contextual transformer (MPNet) for keyword extraction and a static word-level model (GloVe Dolma) for keyword aggregation is deliberate. The two stages perform different computational tasks. At the extraction stage, the operative unit is the abstract, and a contextual model is optimal because the comprehensive meaning of the abstract should shape which terms are selected. At the aggregation stage, the computational object is no longer a passage but a bag of lexical items, and the goal is to position that bag in a stable space in which articles with overlapping vocabulary end up near each other. A static embedding offers three properties that contextual models do not: (a) each keyword has a single vector regardless of the article in which it was selected (a precondition for the "mean of keyword vectors" operation to be meaningful), (b) the resulting vectors are embedded in a Euclidean space (matching the assumptions of Ward's agglomerative criterion below), and (c) the representation is deterministic and auditable. The polysemy of the focal concept is therefore not resolved by the embedding model, but revealed by the clustering, surfacing the distinct lexical neighborhoods in which "vulnerability" is embedded across the corpus. Vocabulary coverage of the corpus's keywords against the Dolma model's vocabulary is 96.27%. The remaining 3.73% are represented by the zero vector at aggregation, consistent with standard practice (Mikolov et al., 2013).

Dimensionality is reduced before clustering using a data-driven procedure. Vectors are first standardized (zero mean, unit variance). A Principal Component Analysis is then fit to the standardized matrix purely as a diagnostic, and the smallest number of components whose cumulative explained variance reaches 80% is selected automatically as the target dimensionality. As a robustness check, the elbow of the cumulative-variance curve is also located independently using the Kneedle algorithm (Satopää et al., 2011). For the present corpus, both diagnostic methods agree on 60 components, which is the dimensionality used for clustering throughout the paper. UMAP (McInnes et al., 2018) is then applied to project the standardized matrix into this 60-dimensional space.

2.3. Clustering

Articles are grouped through hierarchical agglomerative clustering. The choice is justified along three axes. First, hierarchical rather than partitional clustering is preferred because the analytic interest is not in a single grouping but in how groupings nest at increasing levels of granularity, which a partitional method like KMeans cannot expose. Second, agglomerative rather than divisive hierarchical clustering is used because agglomerative methods are computationally tractable on a corpus of this size and because the bottom-up merge sequence is more naturally interpretable as a record of which articles share keyword vocabulary. Third, Ward linkage is preferred over single, complete, or average linkage because Ward minimizes within-cluster variance at each merge, producing compact and comparably sized clusters that are well-suited to the Euclidean structure of the GloVe space and to the variance-based validation metrics used below. The dendrogram is computed across the full range $k = [2, 12]$, with each article represented by the mean of its five keyword embeddings.

The resulting structure can be read as a taxonomy in the descriptive sense, illustrating a map of how articles in the corpus group together by keyword vocabulary at increasing levels of granularity. It is not a normative definite classification of vulnerability research, and the hierarchy reflects geometric proximity in the embedding space, itself a function of (a) the keyword extraction model, (b) the static embedding space, and (c) the corpus retrieved for the search query.

2.4. Cluster validation

Four geometric goodness of fit metrics are computed across the full range $k = [2, 12]$. The silhouette score measures the relative compactness and separation of clusters (Rousseeuw, 1987). The Davies–Bouldin index (Davies and Bouldin, 1979) measures the ratio of within-cluster dispersion to between-cluster separation. The Caliński–Harabasz index (Caliński and Harabasz, 1974) is the ratio of between-cluster to within-cluster variance. The cophenetic correlation (Sokal and Rohlf, 1962) is reported once per linkage matrix rather than per k , because it measures how faithfully the dendrogram preserves the original pairwise distances between observations and is therefore the appropriate indicator for the hierarchical structure as a whole (Saraçlı et al., 2013).

A second layer evaluates topic coherence. C_v and C_{npmi} (Röder et al., 2015) are computed over the top keywords of each cluster against actual co-occurrence patterns in the abstracts. Geometric metrics evaluate whether clusters are well-separated as point sets in vector space; coherence metrics evaluate whether the keywords associated with each cluster make linguistic sense together.

3. Results

Cluster contents at each level k are characterized by the seven most frequent keywords in the articles assigned to each cluster. Two methodological caveats apply throughout this section. First, the hierarchy reflects geometric proximity in the embedding space and is therefore a function of the keyword extraction model, the static embedding space, and the corpus retrieved for the search query. Under different conditions the surface organization could in some parts differ. Second, “taxonomy” in this paper refers to the data structure produced by Ward linkage, not to a normative classification that future researchers are being asked to adopt.

3.1. Dendrogram

The full dendrogram from $k = 2$ to $k = 12$ is presented in Fig. 2. Cluster size in the figure is scaled to the total number of citations accumulated by articles in each cluster, so that the thematic prominence in the literature is visible alongside the thematic structure. The labels shown for each cluster are the seven most frequent words. Fig. 2 is generated as the deterministic output of a rendering script that takes the post-clustering data as its only input.

3.2. Validation metrics

The validation phase reports six indices, each evaluating a different aspect of cluster quality in two complementary layers. The four geometric indices and the cophenetic correlation are reported in Fig. 3.

Across the range $k = [2, 12]$, silhouette scores remain within a moderate band of $[0.31, 0.45]$, with the highest value at $k = 2$ (0.4458) reflecting the trivial separation of a coarse two-way split. From $k = 4$ onward the silhouette stabilizes between 0.31 and 0.35, indicating moderate but consistent within-cluster compactness across granularities. The Davies-Bouldin index, where lower values indicate better separation, fluctuates between 0.85 and 1.27 across the range, with a local minimum at $k = 10$. The Caliński-Harabasz index, where higher is better, decreases monotonically from 14,388.0 at $k = 2$ to 8,539.2 at $k = 12$, a pattern consistent with the agglomerative dilution of variance ratios as the partition becomes finer. The cophenetic correlation, computed once on the linkage matrix, is 0.7087. Cophenetic correlation is the appropriate goodness-of-fit indicator for the dendrogram as a whole, because it measures how faithfully the tree distances preserve the original pairwise distances between observations. Values above 0.7 usually indicate that the hierarchical structure is a faithful representation of the underlying distance matrix (Sokal and Rohlf, 1962; Saraçlı et al., 2013). The obtained value of 0.7087 sits at this threshold, which is consistent with the finding that the corpus does have hierarchical lexical structure but with substantial within-cluster heterogeneity.

Topic coherence is computed separately at each cut from $k = 2$ to $k = 12$ by averaging cluster-level C_v and C_{npmi} over the clusters present at that cut. Coherence indices evaluate whether the keywords associated with each cluster co-occur in the abstracts in linguistically meaningful ways. The C_v trajectory rises modestly with k and stabilizes at approximately 0.68 from $k = 9$ onward. C_{npmi} follows the same pattern more compressed ranging from 0.046 to 0.083 across the corpus. These patterns are plotted in Fig. 4.

The shape of the trajectory is a substantive finding: disaggregating beyond $k = 5$ produces clusters whose top keywords genuinely become more linguistically coherent with the abstract they describe.

Taken together, the geometric indices indicate moderate cluster quality consistent with the heterogeneity of an interdisciplinary corpus. The cophenetic correlation indicates that the hierarchical structure is a faithful representation of the underlying distance matrix; and the coherence trajectories indicate that the lexical content of the clusters becomes meaningfully tighter as granularity increases up to and beyond $k = 7$. These computational metrics are treated as essential for objective validation but not sufficient. They are reported alongside the qualitative inspection of cluster contents, which remains a meaningful validation step.

3.3. Cluster characterization

Since a specific cross-section of the dendrogram is necessary for cluster content inspection, the remainder of this section focuses on the cut at $k = 7$. This is a curated choice rather than a formal-criterion choice, justified along three converging lines. First, $k = 7$ separates substantially distinct sub-vocabularies that lower cuts collapse, while no cluster falls below approximately six hundred articles. Second, the validation metrics reported in Section 3.2 do not identify a single global optimum, but place $k = 7$ within the supported range across all indices. Third, $k = 7$ remains interpretively legible. Higher values of k disaggregate further but at a cost in narrative coherence and plot readability. At $k = 7$ the silhouette score is 0.3168, the Davies-Bouldin index is 1.2466, and the Caliński-Harabasz index is 9,829. Regarding topic coherence at this dendrogram cut, C_v has a value of 0.6559 and C_{npmi} of 0.0717.

Cluster contents at $k = 7$ are summarized in Table 1. Cluster names are descriptive labels suggested by the top-frequency keywords for interpretability, not analytic categories with objective validity. Annual growth is computed as a smoothed compound annual growth using three-year averages at the 2000 and 2024 endpoints to reduce single-year noise.

Table 1. Cluster characterization at $k = 7$

Clusters	Total article count	Total citations	Total ratio citations/articles	Average annual growth (article count)	Top 7 keywords
Psychopathology	2,156	83,218	38.6	6.6%	depression; anxiety; stress; cognitive; psychological; psychopathology
Health, Rights & Violence	2,333	45,259	19.4	11.5%	hiv; feminist; health; violence; ethics; aids; crime
Climate & Rural Livelihood	1,949	78,683	40.37	14.92%	climate; livelihood; livelihoods; adaptation; farmers; agricultural; households
Floods & Natural Disasters	2,701	75,558	27.97	16.16%	flood; climate; disasters; flooding; floods; gis; resilience
Security & Infrastructure	1,322	32,388	24.5	15.71%	security; risk; resilience; infrastructure; consumers; transportation; consumer
Economic Poverty	1,357	30,136	22.21	10.95%	poverty; households; economic; financial; income; economy; welfare
Pandemic & Migration	633	10,573	16.37	23.04%	pandemic; covid; migrants; migrant; refugees; refugee; migration

The largest cluster by article count is Floods & Natural Disasters (C4, 2,701 articles), which is also the third-most-cited (75,558 citations) and one of the fastest-growing branches over 2000-2024. Psychopathology is the second largest by article count and the most-cited overall (83,218 citations) but the slowest-growing of the seven clusters; its citation impact reflects an established literature whose share of new production has been contracting. Health, Rights & Violence is internally heterogeneous, bundling HIV-and-public-health vocabulary with feminist, ethics, violence, and crime vocabulary. This cluster emerges as a recognizable strand only at this level, since lower cuts merge it with Psychopathology. Climate & Rural Livelihood and Floods & Natural Disasters (C4) sit on adjacent but distinct branches at $k = 7$, a separation driven by the keyword vocabulary of the corpus and is itself a finding. Security & Infrastructure, Economic

Poverty and Pandemic & Migration round out the seven-cluster picture. This last cluster is the smallest and most temporally concentrated, reflecting the post-2020 emergence of pandemic-era vocabulary alongside the established migration literature.

3.4. Historical analysis

Two complementary views of the temporal evolution of the literature are reported. Fig. 5 shows the yearly cluster share at $k = 7$ from 2000 to 2024, computed as a percentage stacked-area chart on the article counts per cluster.

The most pronounced trajectories are the contraction of the Psychopathology share (from 39.06% in 2000 to 9% in 2024) and the expansion of the Floods & Natural Disasters share, from 6.25% to 26.29% over the same window. Climate & Rural Livelihood similarly expands from 4.69% to 15.8%. Corresponding to the COVID-19 publication wave, the Pandemic & Migration cluster spikes during 2020-2022. A cut at a higher k level that separates Migration from Pandemic-related sources would probably show how the former has advanced through the decades as an isolated subject. Health, Rights & Violence and Economic Poverty show comparatively flat trajectories, fluctuating around 18% and 11% respectively. Security and Infrastructure surges at the beginning of the 2000s and then stabilizes around 12% over the rest of the period examined.

The structural view in Fig. 5 is complemented by lexical views in Figs. 6-8, three VosViewer maps of the full corpus across historically meaningful periods. VosViewer is an open-source bibliometric software for visualizing scientific literature by building keyword co-occurrence networks. The input to VosViewer is the same set of KeyBERT-extracted keywords used throughout the pipeline. Three chronological windows are presented: 1964 to 1999 (604 records), 2000 to 2015 (3,358 records), and 2016 to 2025 (8,488 sources). Different cutoff points were selected for keyword co-occurrence (the number of minimum co-occurrences), since the imbalance of sample sizes would make the latter maps unreadable.

The maps make several patterns visible that are obscured at the cluster-share level. Fig. 6 (1964-1999) is dominated by psychopathology vocabulary (depression, depressive, stress, anxiety, psychiatric, schizophrenia) with HIV and AIDS terminology emerging on the periphery as a yellow-colored late-period addition. Fig. 7 (2000-2015) shows the first substantive emergence of climate-and-disaster vocabulary as a coherent lexical region (climate, disasters, environmental disasters) alongside the still-dominant psychopathology vocabulary. Both regions are spatially distinct on the map, which corresponds to their geometric separation in the dendrogram. Fig. 8 (2016-2025) consolidates the climate-related sources as the dominant theme in the literature and shows the abrupt insertion of the pandemic and COVID vocabulary.

4. Discussion

This paper shows a multi-resolution map of how vulnerability is organized across an interdisciplinary scholarly corpus, and the validation metrics indicate moderate cluster quality on the resulting structure. The discussion that follows addresses the two research questions in conceptual rather than descriptive terms.

Two preliminary clarifications are in order. First, the analysis does not propose that machine-extracted keywords replace author-supplied keywords as a general criterion. Key-BERT-derived keywords are presented as a reproducible alternative for the specific task of large-scale semantic clustering, where author keywording is known to be inconsistent in quantity and biased. Second, the workflow is reproducible and cost-accessible, addressing concerns raised by Affengruber et al. (2024) and Tóth et al. (2024) about the scalability of research synthesis methods in the face of exponentially growing publication volumes.

4.1. What themes in the vulnerability literature can be identified through a combination of WE with clustering methods?

Top keywords inspected at $k = 7$ group together along lines that, while geometrically driven by keyword vocabulary, are interpretable as substantively distinct framings of vulnerability. The group (“climate”, “livelihood”, “livelihoods”, “adaptation”, “farmers”) shares a recognizable thematic register that differs from the group (“depression”, “anxiety”, “stress”, “depressive”, “cognitive”). This two-way separation is a primary finding of the analysis and is preserved across cuts from $k = 2$ onwards. Hierarchical clustering produces a nested structure of partitions, which differs from the static partitions obtained from flat clustering algorithms used in related work (Chen et al., 2021; Onan, 2019; Zhang et al., 2018) and which permits the audit of alternative cuts on the same underlying linkage matrix.

The dendrogram should technically be interpreted from bottom ($k = 12$) to top ($k = 2$), since the algorithm iteratively merges adjacent observations until only one cluster remains. However, because the analytic interest here is in how a contested concept is organized at progressively finer granularities, a top-down reading of the structure is more intuitive. Clusters that only merge at the top of the hierarchy, such as Security & Infrastructure, contain articles whose keyword vocabulary is more idiosyncratic relative to the rest of the corpus. Three broad framings of vulnerability can be distinguished in the $k = 7$ cut.

A first framing addresses vulnerability at the systemic level and studies how human and natural environments are exposed, impacted, and adapt to climatic and geophysical hazards. At $k = 7$ this framing spans Floods & Natural Disasters, Climate & Rural Livelihood, and Security & Infrastructure. Highly cited papers in this branch include Adger (2006), Cutter et al. (2003), and Smit and Wandel (2006). The operative

model is equational: vulnerability is the joint product of hazard severity, the degree of exposure, and the system's adaptive capacity. The unit of analysis is the human–environment interaction, and comparisons across communities are central.

A second framing is embodied by literature on psychological distress and disorders. At $k = 7$ this is the Psychopathology cluster (C1), which maintains internal coherence at this cut and only splits at higher k into more specialized sub-clusters on epidemiology and family psychology. Top-cited sources here include Vansteenkiste and Ryan (2013) and Charles (2010). In this branch, vulnerability is operationalized as a within-person constellation of biological, cognitive, and affective factors that elevate the probability of mental-health pathology, with particular attention to developmental transitions. Although psychological phenomena may be studied in social contexts (Hankin and Abramson, 2001), the unit of analysis is the individual.

A third framing positions vulnerability as an attribute of marginalized or differently positioned social groups, where ethnic, sociodemographic, and structural factors shape exposure and harm distribution. At $k = 7$ this framing is distributed across three clusters: Health, Rights & Violence, which bundles HIV-and-public-health vocabulary with feminist, ethics, violence, and crime vocabulary; Economic Poverty, centered on household economic precarity; and Pandemic & Migration, where post-2020 pandemic vocabulary co-clusters with the established migration-and-displacement literature. Highly cited examples include Quesada et al. (2011) and Kim and Bostwick (2020) on migration; Briguglio et al. (2009) on socioeconomic vulnerability; Chakraborti and Garland (2012) on victimization; and Middlemiss and Gillard (2015) on living conditions. The marker that distinguishes this framing from the climate-systemic one is that vulnerability is not a function of natural-hazard exposure plus adaptive capacity but a function of structural social position. The marker that distinguishes it from the psychopathology framing is that the unit of analysis is groups rather than individuals.

A tentative interpretive synthesis emerges from the cluster structure. The Psychopathology cluster sits on a branch that merges only at higher cuts with the structurally oriented clusters. This geometric isolation suggests that the literatures focusing on individual expressions of vulnerability and those focusing on the structural production of vulnerability share comparatively little keyword vocabulary. Such a separation invites comparison with what the broader social-scientific literature describes as the Micro-Macro Divide (Eckart et al., 2019) or the Structure-vs-Agency Debate (Krause, 2013), although we present this as a tentative reading of the cluster structure rather than as a direct empirical result. Whether the lexical separation surfaced here reflects a paradigmatic disconnection in vulnerability research, or whether it reflects vocabulary differences that obscure underlying conceptual continuities, is a question that the present method cannot conclude.

4.2. Can historical trends be identified to discern changes in how vulnerability is understood?

The analysis at $k = 7$ from 2000 to 2024 (Fig. 5) shows two pronounced and several minor temporal trajectories. In 2000, articles on Psychopathology accounted for 39.06% of the year's production. By 2024 this share had contracted to 9%. Over the same window, Floods & Natural Disasters expanded from 6.25% to 26.29%, and Climate & Rural Livelihood from 4.69% to 15.80%. Pandemic & Migration rose from 1.56% to 6.45% with a visible peak during 2020–2022 corresponding to the COVID-19 publication wave. Health, Rights & Violence and Economic Poverty show comparatively flat trajectories, fluctuating around 18% and 11% respectively. The directional reading is robust: the contraction of psychopathology and the expansion of the climate-disaster strands are visible across all reasonable cut levels, although the exact numerical shares are conditional on the cut. Additionally, the VosViewer maps show a consistent transition, with climate-related vocabulary displacing psychological terms as the center of the co-occurrence networks.

The contraction of psychopathology challenges the assumption, present in some bibliometric and conceptual literatures, that vulnerability research originated primarily in relation to climate change and disaster studies during the 1980s (Alexander, 2002; Biswas and Nautiyal, 2023; Frerks et al., 2011). The corpus suggests instead that the term was first predominantly mobilized in psychological and psychiatric registers and later expanded into environmental ones. The oldest sources in the corpus are consistent with this: Nielsen (1964) on social vulnerability in cerebral-palsied children, Hansen (1965) on family vulnerability to stress, and Garnezy (1971) on children vulnerable to the development of psychopathology in adulthood are all psychological in framing. The mid-1990s onward expansion of the climate-disaster strand likely reflects the international advance of the climate-and-sustainability paradigm and is consistent with findings from adjacent bibliometric studies that use different methods (Kim et al., 2021; Li and Wang, 2022; Lu et al., 2023).

4.3. Limitations

Several limitations bound the claims made above and identify productive directions for follow-up work. First, the 12,450 records derive from Scopus only, restricted to English-language journal articles in five subject areas. Grey literature, policy documents, non-English scholarship, and conference proceedings are by design out of scope. The mapping therefore reflects how the academic literature in these subject areas constructs vulnerability, not how the concept circulates between scholarly and non-scholarly domains. Application of the pipeline to grey literature and to ideologically and culturally loaded discourses is identified as a particularly productive future direction.

Second, the validation reported combines geometric metrics, dendrogram-level fit, and topic coherence. It does not include human-in-the-loop validation against a domain-expert gold-standard subset. Constructing such a subset within the present revision was considered outside its scope: it requires recruiting specialists from each disciplinary field represented in the corpus, designing a coding protocol that does not privilege any single disciplinary register, computing inter-rater reliability across coders with different theoretical commitments, and adjudicating disagreements.

Third, the pipeline is applied rather than methodologically comparative. A systematic benchmark study against contextual-embedding aggregation, BERTopic, and LLM-based zero-shot extractors is beyond the scope of a paper whose primary contribution is the integrative application of the pipeline to a contested social-scientific concept and would definitely yield a productive direction.

Finally, the pipeline does not provide a systematic classification of definitions of vulnerability. It surfaces an array of lexical frames within which the concept circulates, each carrying different connotations and uses. A complete synthesis of what vulnerability is would require complementing this work with a contrast of definitions in the conceptual literature. The present method offers raw material for that synthesis, not a substitute.

5. Conclusion

This paper has presented a reproducible, multi-resolution mapping of how vulnerability is organized across an interdisciplinary scholarly corpus, leveraging an NLP pipeline that combines contextual keyword extraction, static keyword aggregation, data-driven dimensionality reduction, hierarchical agglomerative clustering, and a multi-metric validation envelope. The contribution is methodologically transparent and cost-accessible, and it is positioned as an instrument for critical inquiry into how a contested concept is configured. The primary audience for the present work is the interdisciplinary social-scientific community for whom such a tool can support critical engagement and scalable literature synthesis.

6. References

- Adger, W. N. (2006). Vulnerability. *Global Environmental Change*, 16, 268-281. <https://doi.org/10.1016%2Fj.gloenvcha.2006.02.006>.
- Affengruber, L., van der Maten, M. M., Spiero, I., Nussbaumer-Streit, B., Mahmić-Kaknjo, M., Ellen, M. E., Goossen, K., Kantorova, L., Hooft, L., Riva, N., Poulentzas, G., Lalagkas, P. N., Silva, A. G., Sassano, M., Sfetcu, R., Marqués, M. E., Friessova, T., Baladia, E., Pezzullo, A. M., ... Spijker, R. (2024). An exploration of available methods and tools to improve the efficiency of systematic review production: a scoping review. *BMC Medical Research Methodology*, 24(210). <https://doi.org/10.1186/s12874-024-02320-4>.
- Alam, S. M. S., Chowdhury, M. A. M., and Razak, D. B. A. (2021). Research evolution in banking performance: a bibliometric analysis. *Future Business Journal*, 7(1). <https://doi.org/10.1186/s43093-021-00111-7>.
- Alexander, D. (2002). The Study of Natural Disasters, 1977-1997: Some Reflections on a Changing Field of Knowledge. *Disasters*, 21(4), 284-304. <https://doi.org/10.1111/1467-7717.00064>.
- Arellano, J. R., y Sánchez, A. (2019). El alcance de los enfoques de vulnerabilidad y pobreza para la definición de la población objetivo en programas sociales. *Gestión y política pública*, 28(2). <https://doi.org/10.29265/gypp.v28i2.623>.
- Aria, M., and Cucurullo, C. (2017). *Bibliometrix*: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>.
- Aristegui, I., Beloki, U., Díez, A., and Silvestre, M. (2017). Vulnerabilidad social percibida en contexto de crisis económica. *Revista Española de Sociología*, 26(3 supl.), 17-39. <https://doi.org/10.22325%2Ffes%2Fres.2017.33>.
- Arseniev-Koehler, A. (2024). Theoretical Foundations and Limits of Word Embeddings: What Types of Meaning can They Capture? *Sociological Methods & Research*, 53(4), 1753-1793. <https://doi.org/10.1177/00491241221140142>.
- Baker, H. K., Pandey, N., Kumar, S., and Haldar, A. (2020). A bibliometric analysis of board diversity: Current status, development, and future research directions. *Journal of Business Research*, 108, 232-246. <https://doi.org/10.1016/j.jbusres.2019.11.025>.

- Berger, M., McDonough, K., and Seversky, L. M. (2017). Cite2vec: Citation-Driven Document Exploration via Word Embeddings. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 691-700. <https://doi.org/10.1109/TVCG.2016.2598667>.
- Berrouet, L. M., Machado, J., and Villegas-Palacio, C. (2018). Vulnerability of socio-ecological systems: A conceptual Framework. *Ecological Indicators*, 84, 632-647. <http://dx.doi.org/10.1016/j.ecolind.2017.07.051>.
- Bijker, W. E., Bal, R., and Hendriks, R. (2009). *The Paradox of Scientific Authority: The Role of Scientific Advice in Democracies*. MIT Press. <https://doi.org/10.7551/mitpress/9780262026581.001.0001>.
- Biswas, S., and Nautiyal, S. (2023). A review of socio-economic vulnerability: The emergence of its theoretical concepts, models and methodologies. *Natural Hazards Research*, 3(3), 563-571. <https://doi.org/10.1016/j.nhres.2023.05.005>.
- Bloom, L., and Lahey, M. (1978). *Language Development and Language Disorders*. John Wiley and Sons, Inc. <https://doi.org/10.7916/D8QZ2GQ5>.
- Briguglio, L., Cordina, G., Farrugia, N., and Vella, S. (2009). Economic Vulnerability and Resilience: Concepts and Measurements. *Oxford Developmental Studies*, 37(3), 229-247. <https://doi.org/10.1080/13600810903089893>.
- Brouwer, W. P., and Hollenbach, M. (2022). Search engine optimization for scientific publications: How one can find your needle in the haystack. *United European Gastroenterol. J.*, 10(8), 906-908. <https://doi.org/10.1002/ueg2.12311>.
- Brown, K. (2011). "Vulnerability": Handle with Care. *Ethics and Social Welfare*, 5(3), 313-321. <http://dx.doi.org/10.1080/17496535.2011.597165>.
- Brown, K., Ecclestone, K., and Emmel, N. (2017). The Many Faces of Vulnerability. *Social Policy & Society*, 16(3), 497-510. <https://doi.org/10.1017/S1474746416000610>.
- Brun, C. (2025). "Vulnerability": The Trouble with Categorical Definitions in Institutional Ethical Reviews, Forced Migration Research & Humanitarian Practice. *Daedalus*, 154(2), 114-131. https://doi.org/10.1162/daed_a_02143.
- Calinski, T., and Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics*, 3(1), 1-27. <https://doi.org/10.1080/03610927408827101>.

- Carlson, R., Bauer, J., and Manning, C. D. (2025). A New Pair of Gloves. *arXiv*. <https://doi.org/10.48550/arXiv.2507.18103>.
- Chakraborti, N. and Garland, J. (2012). Reconceptualizing hate crime victimization through the lens of vulnerability and “difference”. *Theoretical Criminology*, 16(4), 499-514. <https://doi.org/10.1177/1362480612439432>.
- Charles, S. T. (2010). Strength and Vulnerability Integration (SAVI): A Model of Emotional Well-Being Across Adulthood. *Psychol Bull.*, 136(6), 1068-1091. <https://doi.org/10.1037/a0021232>.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377. <https://doi.org/10.1002/asi.20317>.
- Chen, G., and Xiao, L. (2016). Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. *Journal of Informetrics*, 10, 212-223. <http://dx.doi.org/10.1016/j.joi.2016.01.006>.
- Chen, G., Hong, S., Du, C., Wang, P., Yang, Z., and Xiao, L. (2024). Comparing semantic representation methods for keyword analysis in bibliometric research. *Journal of Informetrics*, 18, 101529. <https://doi.org/10.1016/j.joi.2024.101529>.
- Chen, T., Li, G. P., Deng, Q. P., and Wang, X. M. (2021). Using Network Embedding to Obtain a Richer and More Stable Network Layout for a Large Scale Bibliometric Network. *Journal of Data and Information Science*, 6(1), 154-177. <https://doi.org/10.2478%2Fjdis-2021-0006>.
- Cheng, X., Nanjundan, P., and George, J. P. (2025). *Introduction to Natural Language Processing: Exploring Techniques, Applications, and Challenges*. River Publishers. ISBN 978-87-7004-831-6.
- Comerio, N., and Strozzi, F. (2019). Tourism and its economic impact: A literature review using bibliometric tools. *Tourism Economics*, 25(1), 109-131. <https://doi.org/10.1177/1354816618793762>.
- Corrin, L., Thompson, K., Hwang, G., and Lodge, J. M. (2022). The importance of choosing the right keywords for educational technology publications. *Australasian Journal of Educational Technology*, 38(2), 1-8. <https://doi.org/10.14742/ajet.8087>.
- Cutter, S. L., Boruff, B. J., and Shirley, W. L. (2003). Social Vulnerability to Environmental Hazards. *Social Science Quarterly*, 84(2), 242-261. <https://doi.org/10.1111/1540-6237.8402002>.

- Davies, D. L., and Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(1), 224-227. <https://doi.org/10.1109/TPAMI.1979.4766909>.
- De Groot, N., Bonsel, G. J., Birnie, E., and Valentine, N. B. (2019). Towards a universal concept of vulnerability: Broadening the evidence form the elderly to perinatal health using Delphi approach. *PLoS ONE* 14(2), e0212633. <https://doi.org/10.1371/journal.pone.0212633>.
- De Solla, D. J. (1965). Networks of Scientific Papers: The pattern of bibliographical references indicates the nature of the scientific research front. *Science*, 149(3683), 510-515. <https://doi.org/10.1126/science.149.3683.510>.
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., and Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>.
- Donthu, N., Kumar, S., Pandey, N., Pandey, N., and Mishra, A. (2021b). Mapping the electronic word-of-mouth (eWOM) research: A systematic review and bibliometric analysis. *Journal of Business Research*, 135, 758-773. <https://doi.org/10.1016/j.jbusres.2021.07.015>.
- Eckart, R., Crocker, A., Ahn, Y., Floyd, S. W., Boyd, B. K., Hodgkinson, G. P., Kozlowski, S. W. J., Moliterno, T. P., y Starbuck, W. H. (2019). Reflections on the micro-macro divide: Ideas from the trenches and moving forward. *Strategic Organization*, 17(3), 385-402. <https://doi.org/10.1177/1476127018770048>.
- Fijačko, N., Masterson, R., Abella, B. S., Kocbek, P., Metličar, S., Greif, R., and Šriglic, G. (2024). Using generative artificial intelligence in bibliometric analysis: 10 years of research trends from the European Resuscitation Congresses. *Resuscitation Plus*, 18, 100584. <https://doi.org/10.1016/j.resplu.2024.100584>.
- Fineman, M. A. (2008). The Vulnerable Subject: Anchoring Equality in the Human Condition. *Yale Journal of Law & Feminism*, 20(1), 8-40. <https://papers.ssrn.com/abstract=1131407>.
- Firth, J. (1957). A Synopsis of Linguistic Theory, 1930-1955. In *Studies in Linguistic Analysis* (pp. 1-31). Special Volume of the Philological Society. Oxford: Blackwell.
- Frerks, G., Warner, J., and Weijs, B. (2011). The politics of vulnerability and resilience. *Ambient Soc.* 14(2). 105-121. <https://doi.org/10.1590/S1414-753X2011000200008>.
- Garnezy, N. (1971). Vulnerability and the issue of primary prevention. *American Journal of Orthopsychiatry*, 41(1), 101-116. <https://psycnet.apa.org/doi/10.1111/j.1939-0025.1971.tb01111.x>.

- Ghosh, A., and Choudhury, S. (2025). Understanding different types of review articles: A primer for early career researchers. *Indian J Psychiatry*, 67(5), 535-541. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_373_25.
- Gibb, C. (2018). A critical analysis of vulnerability. *International Journal of Disaster Risk Reduction*, 28, 327-334. <https://doi.org/10.1016/j.ijdrr.2017.11.007>.
- Gilodi, A., Albert, I., and Nienaber, B. (2022). Vulnerability in the Context of Migration: A Critical Overview and a New Conceptual Model. *Hu Arenas*, 7, 620-640. <https://doi.org/10.1007/s42087-022-00288-5>.
- Grootendorst, M. (2020). *KeyBERT: Minimal keyword extraction with BERT* (version 0.3.0) [Language Model]. Zenodo. <https://doi.org/10.5281/zenodo.4461265>.
- Hankin, B. L., and Abramson, L. Y. (2001). Development of gender differences in depression: an elaborated cognitive vulnerability-transactional stress theory. *Psychol Bull.* 127(6), 773-796. <https://doi.org/10.1037/0033-2909.127.6.773>.
- Hansen, D. A. (1965). Personal and Positional Influence in Formal Groups: Propositions and Theory for Research on Family Vulnerability to Stress. *Social Forces*, 44(2), 202-210. <https://doi.org/10.2307/2575628>.
- Hanson, M. A., Gómez, P., Crosetto, P., and Brockington, D. (2024). The strain on scientific publishing. *Quantitative Science Studies*, 5(4), 823-843. https://doi.org/10.1162/qss_a_00327.
- Harris, Z. S. (1954). Distributional Structure. *Word*, 10(2-3), 146-162. <https://doi.org/10.1080/00437956.1954.11659520>.
- Hessani, H. S. (2025). *LLM Embeddings Explained: A Visual and Intuitive Guide*. HuggingFace. https://hesamaton-primer-llm-embedding.static.hf.space/assets/pdf/pdf_version.pdf.
- Hofmann, M. E., Hinkel, J., and Wrobel, M. (2011). Classifying knowledge on climate change impacts, adaptation, and vulnerability in Europe for informing adaptation research and decision-making: A conceptual meta-analysis. *Global Environmental Change*, 21, 1106-1116. <http://dx.doi.org/10.1016/j.gloenvcha.2011.03.011>.
- Hammer, C. C., Brainard, J., Innes, A., and Hunter, P. R. (2019). (Re-) conceptualising vulnerability as a part of risk in global health emergency response: updating the pressure and release model for global health emergencies. *Emerging Themes in Epidemiology*, 16(2). <https://doi.org/10.1186/s12982-019-0084-3>.

- Hourrane, O., Mifrah, S., Benlahmar, E. H., Bouhriz, N., and Rachdi, M. (2018). Using Deep Learning Word Embeddings for Citations Similarity in Academic Papers. In Y. Tabii, M., Lazaar, M., Al Achhab, and N. Enneya (Eds.), *Big Data, Cloud and Applications. BDCA 2018. Communications in Computer and Information Science*, vol 872. Springer. https://doi.org/10.1007/978-3-319-96292-4_15.
- Iqbar, A., Shahid, A., Roman, M., Afzal, M. T., and Yahya, M. (2023). Exploiting Contextual Word Embeddings for Identification of Important Citations: Incorporating Section-Wise Citation Counts and Metadata Features. *IEEE Access*, 11, 114044-114060. <https://doi.org/10.1109/ACCESS.2023.3320038>.
- Ishtiaque, A., Estoque, R. C., Eakin, H., Parajuli, J., and Rabby, Y. W. (2022). IPCC's current conceptualization of "vulnerability" needs more classification for climate change vulnerability assessments. *Journal of Environmental Management*, 303(1), 114246. <https://doi.org/10.1016/j.jenvman.2021.114246>.
- Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review*, 10, 18-26. <https://doi.org/10.1007/s13162-020-00161-0>.
- Jackson, D., Hayter, M., Carter, B., and Nyamathi, A. (2012). Revisiting the concept of vulnerability: Recognising strength and resilience in the context of risk and susceptibility. *Contemporary Nurse*, 42(2), 142-144. <http://dx.doi.org/10.1080/10376178.2012.11002642>.
- Jurafsky, D., and Martin, J. H. (2025). *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models* (3rd Ed). Online manuscript released August 24, 2025. <https://web.stanford.edu/~jurafsky/slp3>.
- Kara, B. C., Şahin, A., and Dirsehan, T. (2025). BibexPy: Harmonizing the bibliometric symphony of Scopus and Web of Science. *SoftwareX*, 30, 102098. <https://doi.org/10.1016/j.softx.2025.102098>.
- Kawamura, T., Kozaki, K., Kushida, T., Watanabe, K., and Matsumura, K. (2016, November). *Expanding Science and Technology Thesauri from Bibliographic Datasets using Word Embedding*. 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI). San Jose, United States of America. <https://doi.org/10.1109/ICTAI.2016.0133>.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10-25. <https://doi.org/10.1002/asi.5090140103>.

- Kim, B. J., Jeong, S., and Chung, J. (2021). Research trends in vulnerability studies from 2000 to 2019: Findings from a bibliometric analysis. *International Journal of Disaster Risk Reduction*, 56. <https://doi.org/10.1016/j.ijdrr.2021.102141>.
- Kim, J. S., and Kim, S. (2021). Thirty years of the Journal of Derivatives and Quantitative Studies: a bibliometric analysis. *Journal of Derivatives and Quantitative Studies*, 29(4), 258-279. <https://doi.org/10.1108/JDQS-08-2021-0020>.
- Kim, S. J., and Bostwick, W. (2020). Social Vulnerability and Racial Inequality in COVID-19 Deaths in Chicago. *Health Educ Behav*. 1-5. <https://doi.org/10.1177/1090198120929677>.
- Kornai, A. (2023). Foundations of non-compositionality. In *Vector Semantics*. Cognitive Technologies. Springer. https://doi.org/10.1007/978-981-19-5607-2_1.
- Kottow, M. H. (2005). Vulnerability: what kind of principle is it? *Medical Health Care Philosophy*, 7(3), 281-287. <https://doi.org/10.1007/s11019-004-6857-6>.
- Kozlowski, D., Dusdal, J., Pang, J., and Zilian, A. (2021). Semantic and relational spaces in science of science: deep learning models for article vectorization. *Scientometrics*, 126, 5881-5910. <https://doi.org/10.1007/s11192-021-03984-1>.
- Krause, M. (2013). Recombining micro/macro: The grammar of theoretical innovation. *European Journal of Social Theory*, 16(2), 139-152. <http://est.sagepub.com/content/16/2/139>.
- Kuhlicke, C., de Brito, M. M., Bartkowski, B., Botzen, W., Doğulu, C., Han, S., Hudson, P., Karanci, A. N., Klassert, C. J., Otto, D., Scolobig, A., Soares, T. M., and Rufat, S. (2023). Spinning in circles? A systematic review on the role of theory in social vulnerability, resilience and adaptation research. *Global Environmental Change*, 80, 102672. <https://doi.org/10.1016/j.gloenvcha.2023.102672>.
- Kumar, S., Lim, W. M., Pandey, N., and Westland, J. C. (2021). 20 years of Electronic Commerce Research. *Electronic Commerce Research*, 21, 1-40. <https://doi.org/10.1007/s10660-021-09464-1>.
- Lajoie, C., Poleksic, J., Bracken-Roche, D., MacDonald, M. E., and Racine, E. (2020). The Concept of Vulnerability in Mental Health Research: A Mixed Methods Study on Researcher Perspectives. *Journal of Empirical Research on Human Research Ethics*, 15(3), 128-142. <https://doi.org/10.1177/1556264620902657>.
- Leboeuf, L., Brun, C., and Marchetti, S. (2025). Introduction. Between Protection and Harm. Negotiated Vulnerabilities in Asylum Laws and Bureaucracies. In L. Leboeuf, C. Brun, H. Lidén, S. Marchetti,

- D. Nakache, and S. Sarolea (Eds.), *Between Protection and Harm*. IMISCOE Research Series. Springer. https://doi.org/10.1007/978-3-031-69808-8_1.
- Lombardo, G., Tomaiuolo, M., Moddonini, M., Codeluppi, G., and Poggi, A. (2022). Mobility in Unsupervised Word Embeddings for Knowledge Extraction –The Scholars’ Trajectories across Research Topics. *Future Internet*, 14(25). <https://doi.org/10.3390/fi14010025>.
- Li, H., and Wang, W. (2022). Knowledge Domain and Emerging Trends of Social Vulnerability Research: A Bibliometric Analysis (1991-2021). *Int. J. Environ. Res. Public Health*, 19(14), 8342. <https://doi.org/10.3390/ijerph19148342>.
- Lindbom, H., Tehler, H., Eriksson, K., and Aven, T. (2015). The capability concept – On how to define and describe capability in relation to risk, vulnerability and resilience. *Reliability Engineering and System Safety*, 135, 45-54. <https://doi.org/10.1016/j.res.2014.11.007>.
- Lu, S., Huang, J., and Wu, J. (2023). Knowledge Domain and Development Trend on Urban Food Vulnerability Research: A Bibliometric Analysis. *Water*, 15(10). <https://doi.org/10.3390/w15101865>.
- Mackenzie, C. (2013). The Importance of Relational Autonomy and Capabilities for an Ethics of Vulnerability. In C. Mackenzie, W. Rogers, and S. Dodds (Eds), *Vulnerability: New Essays in Ethics and Feminist Philosophy*, Studies in Feminist Philosophy. Oxford Academic. <https://doi.org/10.1093/acprof:oso/9780199316649.003.0002>.
- Mackenzie, C., Rogers, W., and Dodds, S. (Eds.). (2014). *Vulnerability: New Essays in Ethics and Feminist Philosophy*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199316649.001.0001>.
- Mak, S., and Thomas, A. (2022). Steps for Conducting a Scoping Review. *J Grad Med Educ.*, 14(5), 565-567. <https://doi.org/10.4300/JGME-D-22-00621.1>.
- Maynard, B. R. (2025). Research Synthesis Methods: A Guide for Conducting Rigorous and Relevant Research. *Research on Social Work Practice*, 35(6), 607-620. <https://doi.org/10.1177/10497315251334150>.
- McIlwaine, C. (2002). Perspectives on poverty, vulnerability and exclusion. In K. Willis and C. McIlwaine (Eds.) *Challenges and Change in Middle America*. Routledge. <https://doi.org/10.4324/9781315838441>.

- McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv*, 1802.03426. <https://doi.org/10.48550/arXiv.1802.03426>.
- Middlemiss, L., and Gillard, R. (2015). Fuel poverty from the bottom-up: Characterising household energy vulnerability through the lived experience of the fuel poor. *Energy Research and Social Science*, 6, 146-154. <https://doi.org/10.1016/j.erss.2015.02.001>.
- Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv*, 1301.3781. <https://doi.org/10.48550/arXiv.1301.3781>.
- Miller, F., Osbahr, H., Boyd, E., Thomalla, F., Bharwani, S., Ziervogel, G., Walker, B., Birkmann, J., Van der Leeuw, S., Rockström, J., Hinkel, J., Downing, T., Folke, C., and Nelson, D. (2010). Resilience and Vulnerability: Complementary or Conflicting Concepts? *Ecology and Society*, 15(3). <http://www.ecologyandsociety.org/vol15/iss3/art11/>.
- Moran, R., Hollenbeck, J., and Phoenix, C. (2013). Structural vulnerability and narrative: sensitising concepts for understanding the health impacts of climate change. *Advances in Medical Sociology*, 15, 109-124. [https://doi.org/10.1108%2FS1057-6290\(2013\)0000015009](https://doi.org/10.1108%2FS1057-6290(2013)0000015009).
- Nielsen, H. H. (1964). Social Vulnerability and Adjustment of Cerebral Palsied Children. *Scandinavian Journal of Psychology*, 5(1), 26-32. <https://doi.org/10.1111/j.1467-9450.1964.tb01405.x>.
- Nungesser, F., and Schirgi, A. (2024). Debating the Vulnerability Zeitgeist: Introduction to an Interdisciplinary Trialogue. *Human Studies*, 47, 251-260. <https://doi.org/10.1007/s10746-024-09742-5>.
- Odden, T. O. B., Tyseng, H., Mjaaland, J. T., Kreutzer, M. F., and Malthe-Sørensen, A. (2024). Using text embeddings for deductive qualitative research at scale in physics education. *Physical Review Physics Education Research*, 20, 020151. <https://doi.org/10.1103/PhysRevPhysEducRes.20.020151>.
- Oliver-Smith, A. (2016). Disaster risk reduction and applied anthropology. *Annals of Anthropological Practice*, 40(1), 73-85. <https://doi.org/10.1111/napa.12089>.
- Onan, A. (2019). Two-Stage Topic Extraction Model for Bibliometric Data Analysis Based on Word Embeddings and Clustering. *IEEE Access*, 7, 145614-145633. <https://doi.org/10.1109/ACCESS.2019.2945911>.
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. (1957). *The measurement of meaning*. University of Illinois Press.

- Passas, I. (2024). Bibliometric Analysis: The Main Steps. *Encyclopaedia*, 4, 1014-1025. <https://doi.org/10.3390/encyclopedia4020065>.
- Pennington, J., Socher, R., and Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. In A. Moschitti, B. Pang, and W. Daelemans (Eds) *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532-1543). Association for Computational Linguistics. <https://doi.org/10.3115/v1/D14-1162>.
- Pereira, V., Pereira, M., Tarjano, C H. (2023). pyBix –A Python Library for Biliometric and Scientometric Analysis Powered with Artificial Intelligence Tools. *arXiv*, 2304.14516. <https://doi.org/10.48550/arXiv.2304.14516>.
- Pollock, D., Evans, C., Menghao, R., Alexander, L., Pieper, D., Brandão de Moraes, E., Peters, M. D. J., Tricco, A. C., Khalil, H., Godfrey, C. M., Saran, A., Campbell, F., and Munn, Z. (2024). “How-to”: scoping review? *Journal of Clinical Epidemiology*, 176, 111572. <https://doi.org/10.1016/j.jclinepi.2024.111572>.
- Proag, V. (2014). The Concept of Vulnerability and Resilience. *Procedia Economics and Finance*, 18, 369-376. [https://doi.org/10.1016/S2212-5671\(14\)00952-6](https://doi.org/10.1016/S2212-5671(14)00952-6).
- Quesada, J., Hart, L. K., and Bourgois, P. (2011). Structural Vulnerability and Health: Latino Migrant Laborers in the United States. *Medical Anthropology*, 30(4), 339-362. <https://doi.org/10.1080/01459740.2011.576725>.
- Resnyansky, L. (2019). Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge. *Big Data & Society*, 6(1). <https://doi.org/10.1177/2053951718823815>.
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*, 399–408. <https://doi.org/10.1145/2684822.2685324>.
- Roman, M., Shahid, A., Khan, S., Koubaa, A., and Yu, L. (2020). Citation Intent Classification Using Word Embedding. *IEEE Access*, 9. <https://ieeexplore.ieee.org/iel7/6287639/9312710/09319154.pdf>.
- Rousseuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20, 53-65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).

- Ruiz-Rosero, J., Ramirez-Gonzalez, G., and Viveros-Delgado, J. (2019). Software survey: ScientoPy, a scientometric tool for topics trend analysis in scientific publications. *Scientometrics*, 121(2), 1165-1188. <https://doi.org/10.1007/s11192-019-03213-w>.
- Saraçlı, S., Dogan, N., and Dogan, I. (2013). Comparison of hierarchical cluster analysis methods by cophenetic correlation. *Journal of Inequalities and Applications*, 2013. <https://doi.org/10.1186/1029-242X-2013-203>.
- Satvika, Thada, V., and Singh, J. (2021). A Primer on Word Embedding. In I. J. Jacob, S. K., Shanmugam, S., Piramuthu, and P. Falkowski-Gilski (Eds.), *Data Intelligence and Cognitive Informatics. Algorithms for Intelligent Systems*. Springer. https://doi.org/10.1007/978-981-15-8530-2_42.
- Satopaa, V., Albrecht, J., Irwin, D., & Raghavan, B. (2011). Finding a "kneedle" in a haystack: Detecting knee points in system behavior. 2011 *31st International Conference on Distributed Computing Systems Workshops*, 166–171. <https://doi.org/10.1109/ICDCSW.2011.20>.
- Schneiderbauer, S., Calliari, E., Eidsvig, U., and Hagenlocher, M. (2017). The most recent view of vulnerability. In M. Marin, K. Polijanšek, I. Clarck, and T. de Groeve (Eds.) *Science for disaster risk management 2017*. Publications Office of the European Union. <https://op.europa.eu/en/publication-detail/-/publication/4bc0e055-3712-11e7-a08e-01aa75ed71a1/language-en#>.
- Schoorman, F. D., Ballinger, G. A., and Sharma, K. (2025). The big circle (conceptual space) of vulnerability. *Journal of Trust Research*, 15(1), 115-121. <https://doi.org/10.1080/21515581.2025.2460186>.
- Shibayama, S., Yin, D., and Matsumoto, K. (2021). Measuring novelty in science with word embedding. *PLoS ONE*, 16(7), e0254034. <https://doi.org/10.1371/journal.pone.0254034>.
- Singhal, T., Liu, J., Blessing, L. T. M., and Lim, K. H. (2021, December). *Analyzing Scientific Publications using Domain-Specific Word Embedding and Topic Modelling*. 2021 IEEE International Conference on Big Data (Big Data), United States of America. <https://doi.org/10.1109/BigData52589.2021.9671598>.
- Skoglund, J., and Thörnblad, R. (2024). "Vulnerability" and Its Unintended Consequences. *Social Sciences*, 13(5), 258. <https://doi.org/10.3390/socsci13050258>.
- Smit, B., and Wandel, J. (2006). Adaptation, adaptive capacity and vulnerability. *Global Environmental Change*, 16(3), 282-292. <https://doi.org/10.1016/j.gloenvcha.2006.03.008>.

- Sokal, R. R., and Rohlf, F. J. (1962). The comparison of dendrograms by objective methods. *TAXON*, 11(2), 33-40. <https://doi.org/10.2307/1217208>.
- Soldaini, L., Kinney, R., Bhagia, A., Schwenk, D., Atkinson, D., Authur, R., Bogin, B., Chandu, K., Dumas, J., Elazar, Y., Hofmann, V., Jha, A. H., Kumar, S., Lucy, L., Lyu, X., Lambert, N., Magnusson, I., Morrison, J., Muennighoff, N., . . . Lo, K. (2024). *Dolma: An open corpus of three trillion tokens for language model pretraining research*. arXiv. <https://doi.org/10.48550/arXiv.2402.00159>.
- Song, K., Tan, X., Qin, T., Lu, J., and Liu, T. (2020). MPNet: Masked and Permuted Pre-training for Language Understanding. *arXiv*, 2004.09297. <https://doi.org/10.48550/arXiv.2004.09297>.
- Sukhera, J. (2022). Narrative Reviews: Flexible, Rigorous, and Practical. *J Grad Med Educ*, 14(4), 414-417. <https://doi.org/10.4300/JGME-D-22-00480.1>.
- Tian, H., and Zhuo, H. H. (2017). Paper2vec: Citation-Context Based Document Distributed Representation for Scholar Recommendation. *arXiv*, 1703.06587. <https://doi.org/10.48550/arXiv.1703.06587>.
- Timmer, A., Baumgärtel, M., Kotzé, L., and Sligenberg, L. (2021). The potential and pitfalls of the vulnerability concept for human rights. *Netherlands Quarterly of Human Rights*, 39(3), 190-197. <https://doi.org/10.1177%2F09240519211048009>.
- Thijs, B. (2020). Using neural-network based paragraph embeddings for the calculation of within and between document similarities. *Scientometrics*, 125, 835-849. <https://doi.org/10.1007/s11192-020-03583-6>.
- Tóth, B., Berek, L., Gulácsi, L., Péntek, M., and Zrubka, Z. (2024). Automation of systematic reviews of biomedical literature: a scoping review of studies indexed in PubMed. *Systematic Reviews*, 13(174). <https://doi.org/10.1186/s13643-024-02592-3>.
- Tun, N., and Khairuddin, Z. (2024). Exploring 50 years of research on students' identities in higher education context: a bibliometric analysis. *Journal of Nusantara Studies*, 9(1), 368-394. <https://doi.org/10.24200%2Fjonus.vol9iss1pp368-394>.
- Turner, L. (2025). Defining, operationalising and translating “vulnerability” in humanitarian work in Jordan. *Journal of Ethnic and Migration Studies*, 51(7), 1647-1666. <https://doi.org/10.1080/1369183X.2024.2359683>.
- Uekusa, S., Wynyard, M., and Matthewman, S. (2025). Reinvestigating social vulnerability from the perspective of Critical Disaster Studies (CDS): directions, opportunities and challenges in Aotearoa

- disaster research. *Kōtuitui: New Zealand Journal of Social Sciences Online*, 20(1), 159-176. <https://doi.org/10.1080/1177083X.2024.2305639>.
- Uttley, L., Quintana, D. S., Montgomery, P., Carroll, C., Page, M. J., Falzon, L., Sutton, A., and Moher, D. (2023). The problems with systematic reviews: a living systematic review. *Journal of Clinical Epidemiology*, 156, 30-41. <https://doi.org/10.1016/j.jclinepi.2023.01.011>.
- Valdés, M. (2021). Vulnerabilidad social, genealogía del concepto. *Gazeta de Antropología*, 37(1). ISSN 2340-2792.
- Van Eck, N. J., and Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84, 523-538. <https://doi.org/10.1007/s11192-009-0146-3>.
- Vansteenkiste, M., and Ryan, R. M. (2013). On psychological growth and vulnerability: basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration*, 23(3), 263-280. <https://psycnet.apa.org/doi/10.1037/a0032359>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention Is All You Need. *arXiv*, 1706.03762. <https://doi.org/10.48550/arXiv.1706.03762>.
- Virokannas, E., Liuski, S., and Kuronen, M. (2020). The contested concept of vulnerability –a literature review. *European Journal of Social Work*, 23(2), 327-339. <https://doi.org/10.1080/13691457.2018.1508001>.
- Weizenbaum, J. (1966). Computational Linguistics. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>.
- Winograd, T. (1972). Understanding natural language. *Cognitive Psychology*, 3(1), 1-191. [https://doi.org/10.1016/0010-0285\(72\)90002-3](https://doi.org/10.1016/0010-0285(72)90002-3).
- Zhang, Y., Liu, J., Liu, F., Liu, Q., Porter, A., Chen, H., and Zhang, G. (2018). Does deep learning help topic extraction? A kernel k-means clustering method with word embedding. *Journal of Informetrics*, 12, 1099-1117. <https://doi.org/10.1016/j.joi.2018.09.004>.

List of figures

Fig. 1 Diagram of embedding and clustering procedure.

Fig. 2 Dendrogram of vulnerability corpus, $k = [2, 12]$. Ribbon width is scaled to the total number of citations accumulated by articles in each cluster. Cluster labels show the seven most frequent keywords.

Fig. 3 Cluster validation metrics: silhouette, Davies-Bouldin, and Caliński-Harabasz indices across $k = [2, 12]$ (cophenetic correlation = 0.709).

Fig. 4 Mean C_v and C_{npmi} across $k = [2, 12]$.

Fig. 5 Cluster share of article count per year, $k = 7$, 2000-2024.

Fig. 6 VOSViewer map of the vulnerability corpus, 1964-1999.

Fig. 7 VOSViewer map of the vulnerability corpus, 2000-2015.

Fig. 8 VOSViewer map the of vulnerability corpus, 2016-2025.

Statements

Funding

The authors did not receive specific funding from any organization for the submitted work.

Competing interests

The authors have no competing interests to declare that are relevant to the content of this article.

Ethical Approval

The study is a bibliometric analysis of aggregated bibliographic records retrieved from the Scopus database. It does not involve human participants, human tissue, or animals, and does not use any personal or identifiable data. Accordingly, no ethical approval from an institutional board was required. The research was conducted in accordance with the ethical principles for research set out in the Declaration of Helsinki, insofar as the applicable to studies that do not involve human participants.

Informed Consent

This study did not involve human participants and relied exclusively on aggregated publicly available bibliographic metadata. No individual persons or their personal data were involved at any stage of the research.

Data availability

All data used for this analysis, together with the code necessary for exact reproduction of this research, are openly available in a public Zenodo repository: <https://doi.org/10.5281/zenodo.21214869>.

The repository contains the following files:

1. “scopus_export_database.csv”: the bibliographic data exactly as exported from Scopus for the bibliometric study.
2. “The Semantic Taxonomy of Vulnerability_pythoncode.html”: the full method described in the manuscript, executed with descriptions of each step and the corresponding code output.
3. Additional output files:
 - a. “vulnerability_extractedkeywords.csv”: all studied references with the keywords extracted using MPNet and KeyBERT.
 - b. “vulnerability_300d_dataframe.csv”: all extracted keywords with their corresponding 300-dimensional vector embeddings.
 - c. “df_with_clusters.csv”: all references and the clusters to which they are assigned for every dendrogram configuration.
 - d. “hierarchical_clusters.json”: each cluster extracted for all configurations, with the top 7 most frequent keywords in each.

Keyword

Sample: 12,450 indexed articles about

mpnet model embeds each article

Keyword

Each keyword is assigned a 300d vector,

Dimensions

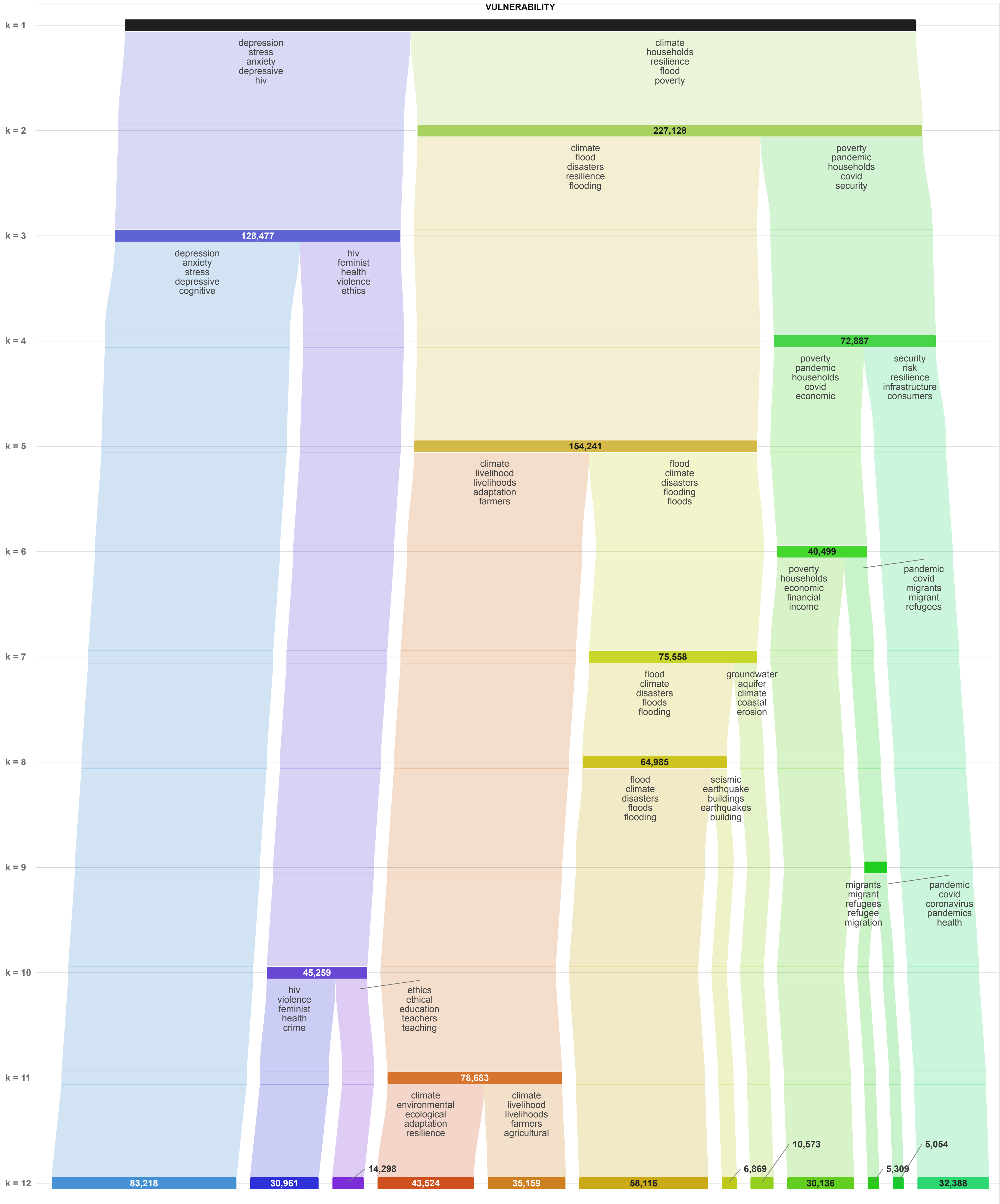
Dimensions are reduced from 300d to

Clusters

Each article's embedded keywords are

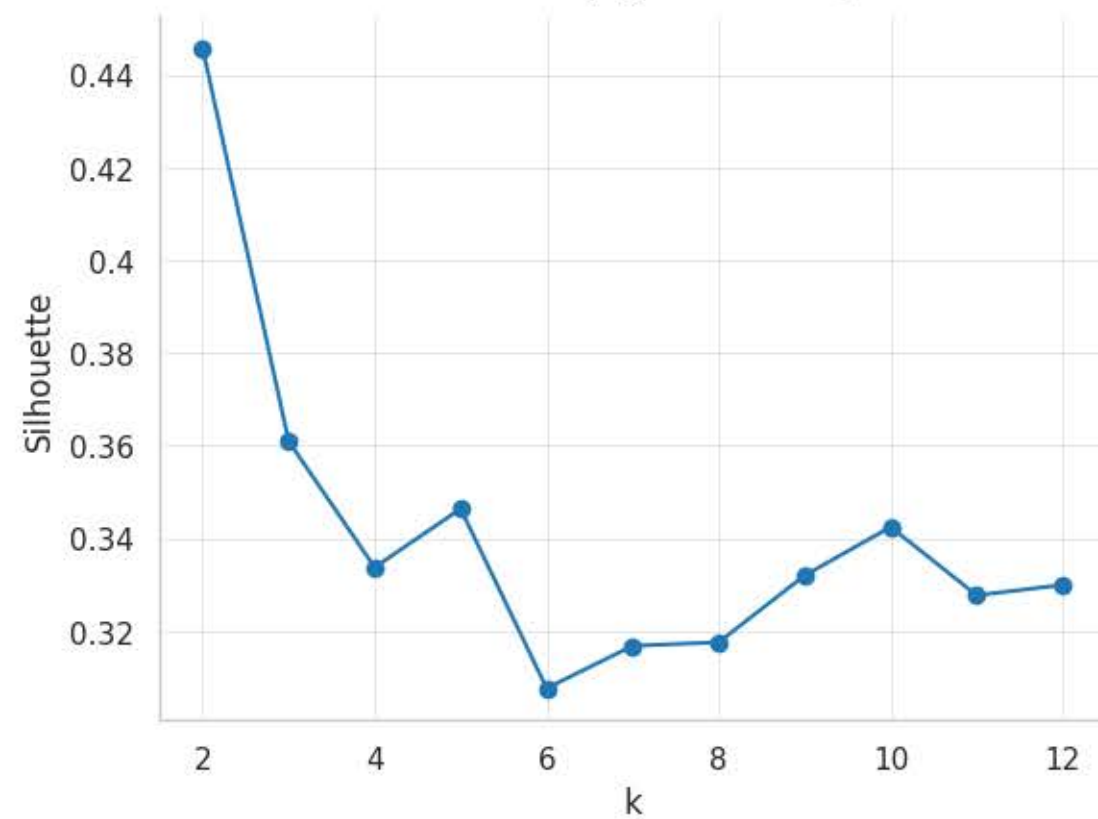
Articles are clustered using
Hierarchical Clustering to get a dendrogram of

Ribbon width = cites

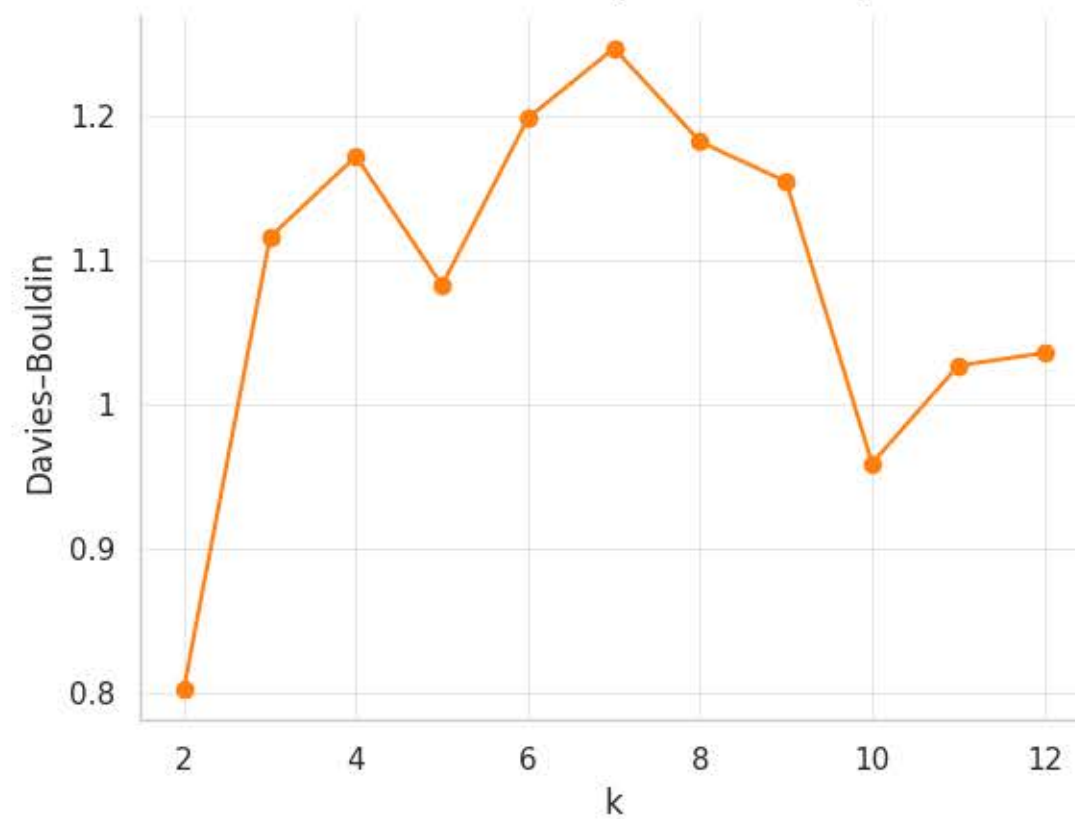


a

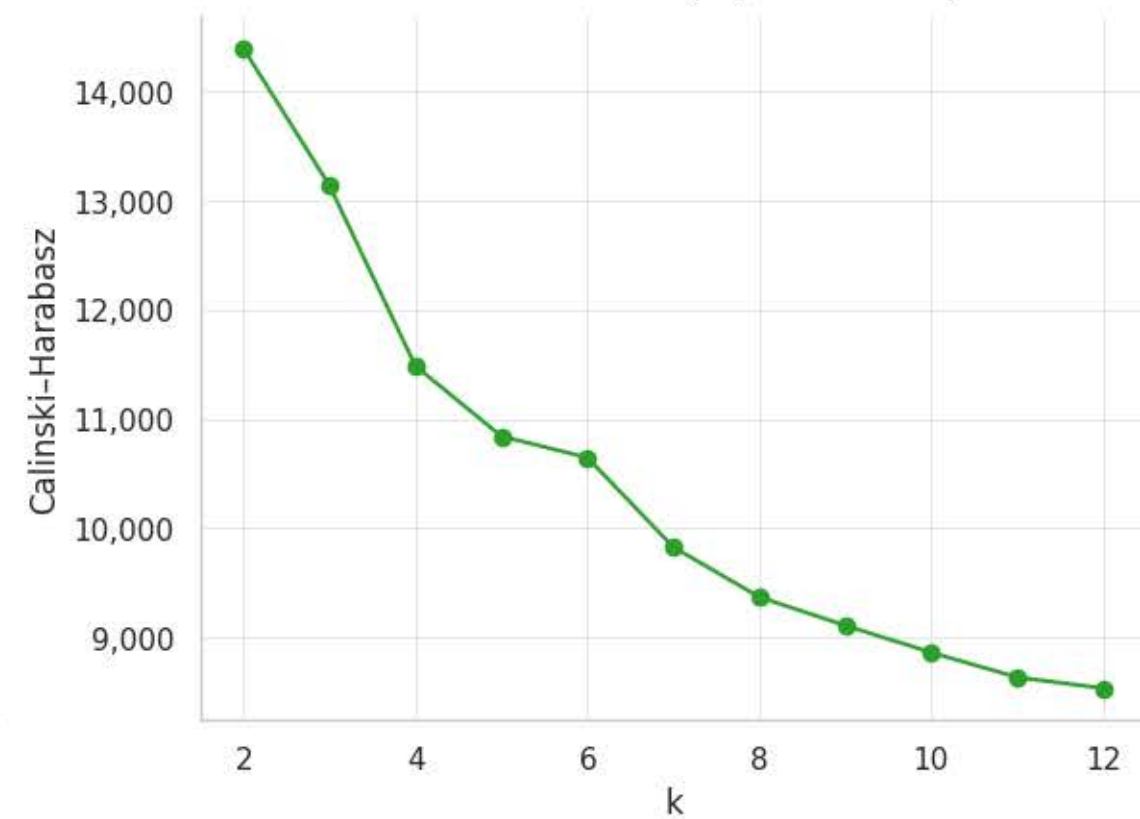
Silhouette (higher is better)

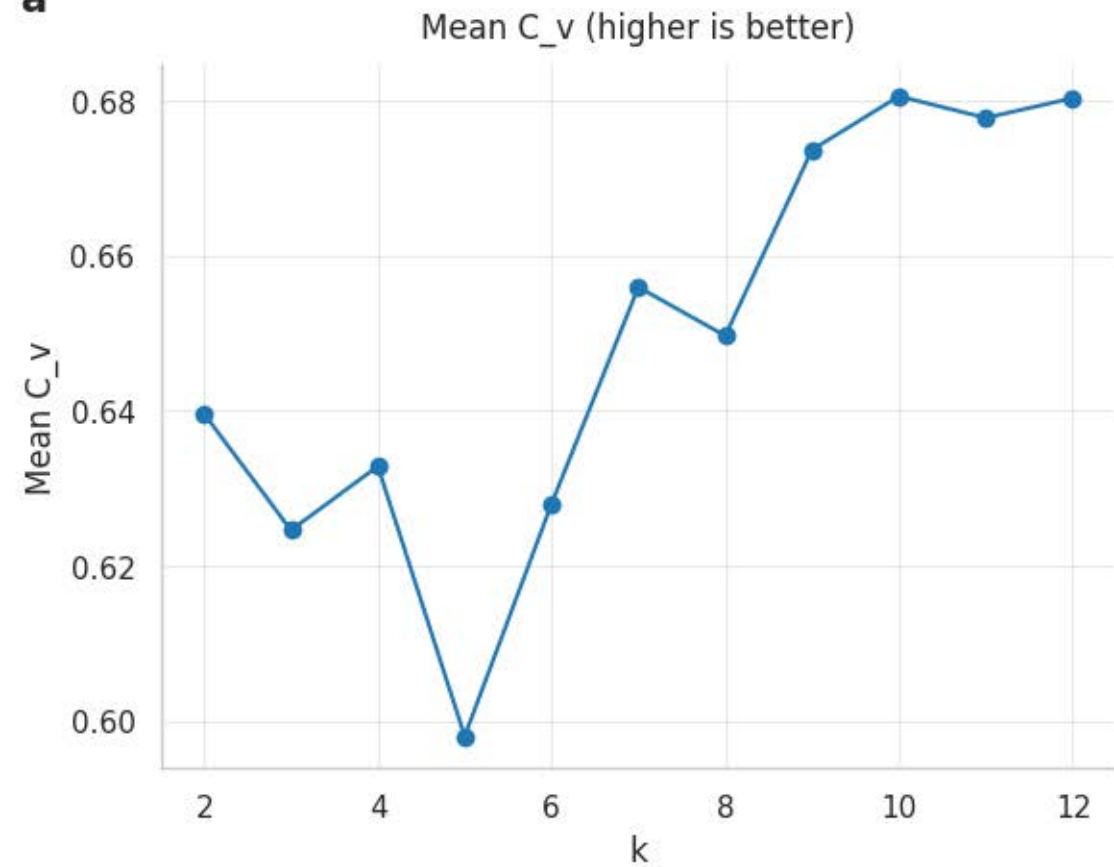
**b**

Davies-Bouldin (lower is better)

**c**

Calinski-Harabasz (higher is better)



a**b**