



Exploring emerging knowledge landscapes through preprints: A DIKW-based computational framework for transforming data into actionable scholarly insights

Tzu-Yu Lin and Pei-Chun Lee

DOI: <https://doi.org/10.47989/ir31ISIC65160>

Abstract

Introduction. Preprints offer a critical vantage point for tracing emerging knowledge trajectories but remain underexamined. Within the data–information–knowledge–wisdom (DIKW) framework, preprints are positioned as raw data requiring computational transformation into interpretable knowledge.

Method. Using explainable artificial intelligence (XAI) preprints as a case domain, this study applied a computational knowledge transformation framework. After a two-stage quality filter ensuring preprints’ metadata completeness and institutional validity, latent Dirichlet allocation (LDA) extracted thematic structures, Word2Vec analysed semantic relations, and content analysis refined interpretive coherence.

Analysis. The framework operationalised the DIKW hierarchy by tracing how unstructured preprint texts evolved into structured information and interconnected knowledge. LDA identified thematic clusters, while Word2Vec and content analysis revealed semantic proximity and conceptual cohesion, forming an interpretable knowledge network.

Results. Two thematic trajectories were identified: (1) algorithmic development and model optimisation, highlighting methodological innovation; and (2) user-centred evaluation and performance assessment, emphasising interpretability and user interaction.

Conclusion. With appropriate quality control and computational modelling, preprints can serve as an early-warning system for emerging knowledge and technologies, offering valuable insights for scholars and data scientists.

Introduction

In contemporary research environments characterised by rapid evolution and intense competition, the ability to access, interpret, and utilise up-to-date information has become a decisive factor in generating new insights and maintaining scholarly relevance. This demand is particularly salient for researchers seeking to detect and interpret emerging intellectual trends (Pontis et al., 2017). While the traditional peer-reviewed publication system remains fundamental for ensuring academic rigor and reliability, its protracted review and publication cycles frequently impede timely dissemination at the research frontier (Alfonso & Crea, 2023). In fast-developing domains, such as explainable artificial intelligence (XAI)—which has rapidly emerged to address the opacity of artificial intelligence models (Arrieta et al., 2020; Castelvechi, 2016; Jacovi, 2023)—research topics are expanding both in diversity and in pace. Within such contexts, the academic community increasingly relies on preprints as critical platforms for rapid scholarly communication. Preprints enable the dissemination of research outputs prior to formal publication, providing scholars with the opportunity to engage with cutting-edge developments at the earliest stage of knowledge production (Aviv-Reuven et al., 2024; Wang et al., 2020).

However, despite these advantages, preprints also present substantial challenges: they often lack peer review, exhibit structural inconsistencies, and vary considerably in quality (Kaiser, 2017; Kodvanj et al., 2022). Although prior studies indicate that the textual content of preprints frequently remains stable through to final publication (Kaiser, 2017; Klein et al., 2016), inconsistencies in formatting and metadata presentation continue to undermine their utility as reliable scholarly resources.

Problem statement

Within the framework of the data–information–knowledge–wisdom (DIKW) hierarchy (Baskarada & Koronios, 2013; Peters et al., 2024; Rowley, 2007), preprints are more appropriately conceptualised as raw data—fragmented, insufficiently validated, and structurally inconsistent—rather than as processed information capable of directly supporting information seeking or knowledge construction (Kaiser, 2017; Kodvanj et al., 2022). While data possesses latent value, its utility depends on analytical transformation, reflecting the epistemic shift from ‘data as the new oil’ to ‘analysis as the new oil’ (Dhande, 2021; Peters et al., 2024).

From an analytical standpoint, preprints constitute a vast corpus of unstructured textual material whose usefulness depends on their conversion into coherent and interpretable informational representations—a process enabled through computational and data-driven methodologies (Assaf & Srouf, 2021; Hussain et al., 2021; Yoon, 2023). Despite their widespread adoption across disciplines—particularly during the COVID-19 pandemic—as accelerators of scholarly communication and innovation (Berg et al., 2016; Fraser et al., 2020; Sarabipour et al., 2019), empirical studies rarely approach preprints as big data resources. Few investigations have examined how they can function as analytical datasets for detecting emerging research trajectories and generating usable insights through computational processing (Langnickel et al., 2021).

Research purpose

Building on the identified research gap, this study seeks to establish an analytical framework for transforming preprints from unstructured textual data into structured and interpretable information. By integrating topic modelling, semantic analysis, and content analysis, the framework systematically identifies salient themes and core concepts, enhances interpretive depth through semantic association mapping, and contextualises findings via qualitative synthesis. Using XAI preprints as an empirical case, this study demonstrates how the convergence of computational and interpretive methodologies can operationalise the transition from data to actionable information within the DIKW paradigm.

Literature review

The data–information–knowledge–wisdom (DIKW) hierarchy: From data to actionable knowledge

The data–information–knowledge–wisdom (DIKW) hierarchy delineates the epistemological transformation of data into information, knowledge, and ultimately wisdom. In this model, data represents discrete observations; information embodies structured and contextualised data; and knowledge arises through interpretation and synthesis (Baskarada & Koronios, 2013; Peters et al., 2024; Rowley, 2007).

Although the model has been critiqued for its linearity and technocentric orientation (Frické, 2009; Peters et al., 2024), the DIKW framework remains salient in articulating the instrumental role of computational mediation in the data-to-knowledge transformation. Specifically, it foregrounds the necessity of algorithmic, analytical, and visualisation techniques to extract meaning and operational value from unstructured or semi-structured data.

Preprints as research data in the scholarly knowledge ecosystem: Opportunities and limitations

Preprints provide an efficient channel for researchers to disseminate their findings prior to formal peer review, supporting principles of transparency, openness, and early-stage scholarly communication (Aviv-Reuven et al., 2024; Fraser et al., 2021; Stuart et al., 2020). The COVID-19 pandemic further accelerated their adoption, positioning preprints as vital infrastructures for the rapid exchange of scientific knowledge (Cohen et al., 2020; Fraser et al., 2021).

Nevertheless, preprints often exhibit incomplete metadata and inconsistent quality due to the lack of rigorous peer-review mechanisms (Kaiser, 2017; Kodvanj et al., 2022). Within the DIKW framework, such characteristics position preprints predominantly at the data layer—unrefined, fragmented, and lacking contextual integrity. Their scholarly potential resides not in their immediate form but in their capacity for analytical transformation.

Preprints provide timely access to emerging scientific developments and serve as early indicators of innovation trajectories (Fraser et al., 2021). From a data-centric perspective, they can be viewed as data-rich artefacts of scientific discourse, whose informational potential can be systematically extracted through computational techniques such as topic modelling, semantic analysis, and text mining (Benkassioui et al., 2022; Larsen et al., 2008; Zadgaonkar & Agrawal, 2024).

Furthermore, the analytical utility of preprints depends heavily on the implementation of appropriate quality control and filtering mechanisms (Alfonso & Crea, 2023; Chaleplioglou & Koulouris, 2023). Given their heterogeneous quality and inconsistent metadata, establishing robust curation processes is essential to ensure data integrity (Alfonso & Crea, 2023). When properly curated and computationally processed, preprints can be transformed from unstructured text into semantically enriched information, enhancing their role in facilitating analytical insight and accelerating scientific discovery.

Uncovering semantic structures of textual data through topic modelling

Topic modelling has emerged as a powerful unsupervised technique for uncovering latent semantic structures within large collections of unstructured text. Among various topic modelling approaches, latent Dirichlet allocation (LDA) remains the most widely adopted and influential model (Celard et al., 2020; Di Gennaro et al., 2021; Rieger et al., 2024). By analysing word co-occurrence patterns across documents, LDA enables the identification of coherent topics, providing valuable insights for computational social science, digital humanities, and bibliometric research.

Latent Dirichlet allocation (LDA): Applications, strengths, and limitations

Proposed by Blei et al. (2003), LDA is a generative probabilistic model that represents documents as mixtures of latent topics, each defined by word distributions. It has been widely applied in various domains, including research abstracts (Fang et al., 2018) and scientific papers (Bo & Gang, 2017) to extract meaningful thematic structures from textual data (Lei et al., 2023).

Despite its interpretability, LDA's reliance on statistical word distributions limits its capacity to capture deeper semantic relations among conceptually related terms, often reducing topic coherence (Rieger et al., 2024; Roberts et al., 2013).

Extending LDA through word embeddings: Toward semantically rich topic models

With advances in neural language modelling, hybrid approaches integrating LDA and word embeddings have gained prominence for enhancing topic coherence and semantic depth (Dieng et al., 2020; Mikolov et al., 2013; Rieger et al., 2024). Word embeddings represent words in continuous vector spaces, effectively capturing syntactic and semantic relations beyond co-occurrence patterns (Dos Santos & Gatti, 2014; Peng et al., 2017). For instance, Zhang et al. (2019) employed sentence-LDA to extract topic vectors from news titles and used *FastText* to learn word representations. Limwattana and Prom-on (2021) developed a neural topic model guided by pre-trained embeddings, while Mu et al. (2022) proposed ClusTop, which applies community detection for topic discovery. Recently, many Word2Vec-based approaches have emerged in natural language processing. Applications include technology forecasting using patent data (Kwon et al., 2022), enhance recommendation systems based on user's historical data (Lei et al., 2020), topic evolution modelling (Hu et al., 2019), and life-cycle analysis of thematic development in mental models fields (Ma et al., 2023).

These studies demonstrate the potential of combining LDA with Word2Vec to improve semantic depth. However, few have applied such integrated models to preprints, which offer rich, time-sensitive data for identifying emerging research trends. Exploring this integration may provide an effective strategy for analysing innovation frontiers in fast-evolving domains.

Methodology

This study aims to develop a computational framework for transforming unstructured preprint texts into structured and interpretable knowledge. As illustrated in Figure 1, the framework integrates topic modelling, semantic similarity analysis, and content analysis to operationalise the process through which raw textual data are organised, semantically linked, and contextually interpreted. This integrated approach enables the systematic exploration of emerging thematic and conceptual structures within dynamic research domains such as XAI.

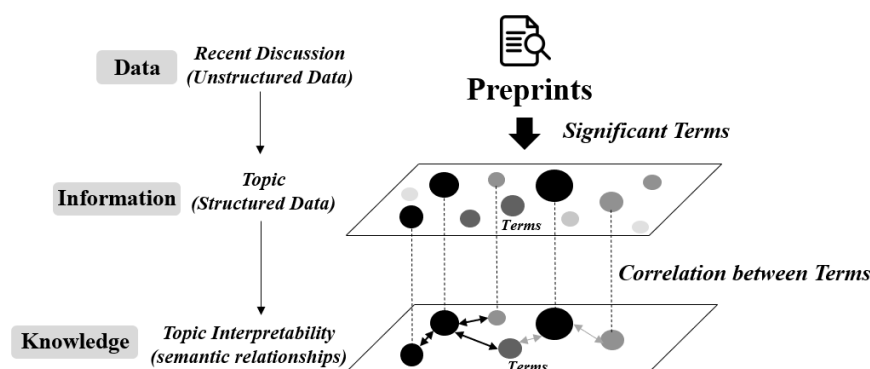


Figure 1. Research framework.

The research process, as Figure 2 shown, consists of four main stages: (1) preprints collection, (2) data preprocessing, and (3) topic modelling, and (4) semantic and content analysis.

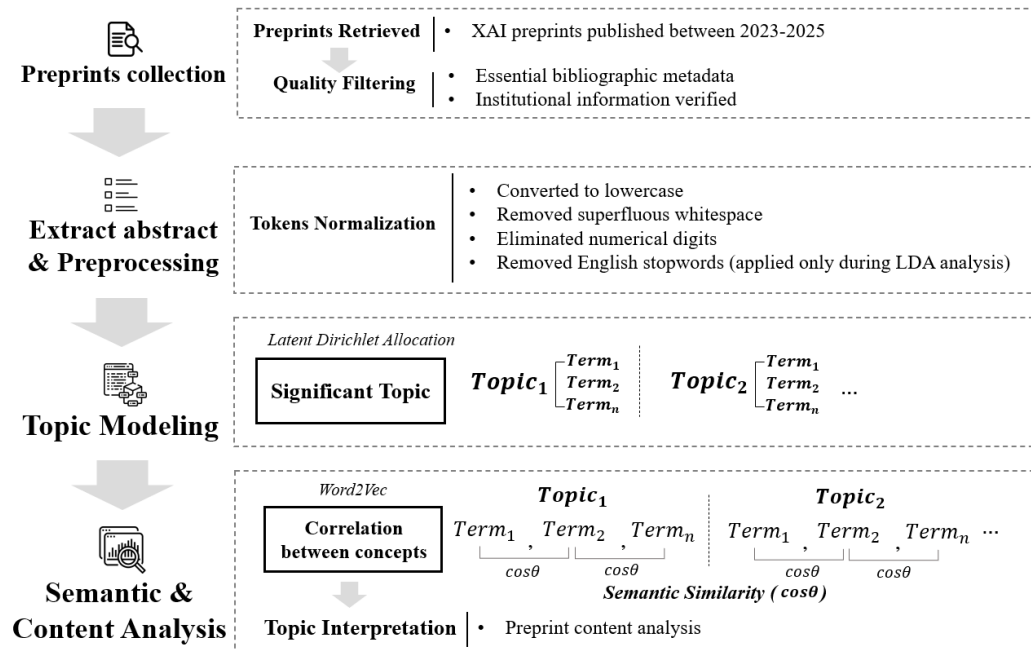


Figure 2. Research process.

Preprints collection

To ensure the currency and relevance of the dataset, preprints pertaining to XAI published between 2023 and 2025 were retrieved from the *Web of Science Preprint Citation Index*, an integrated platform aggregating preprints from leading repositories with standardised metadata.

Given the conceptual and terminological diversity of XAI, a keyword-based retrieval strategy was employed following Jacovi (2023). A refined list of domain-relevant terms was iteratively developed and applied to both titles and abstracts. To enhance precision, only records containing at least two keywords were retained, achieving a 99% precision rate in a random validation of 100 samples. The retrieve query was constructed as follows:

TS=("xai" OR "hcxai" OR "explainability" OR "interpretability" OR "explainable ai" OR "explainable artificial intelligence" OR "interpretable ml" OR "interpretable machine learning" OR "interpretable model" OR "feature attribution" OR "feature importance" OR "global explanation" OR "local explanation" OR "local interpretation" OR "global interpretation" OR "model explanation" OR "model interpretation" OR "saliency" OR "counterfactual explanation")

Although most preprints undergo basic screening prior to publication, the collected records still exhibited issues of incomplete or inconsistent bibliographic metadata. To address these deficiencies, a two-stage quality filtering protocol was implemented.

- (1) Preprints lacking essential bibliographic information—such as title, abstract, or institutional affiliation—were excluded.
- (2) Institutional identities were manually verified through authoritative directories. Records associated with unverifiable or fictitious institutions were subsequently removed to ensure the overall reliability and integrity of the corpus.

Applying these strategies, a total of 7,394 preprint records were identified and collected for analysis. The abstract of each preprint was then extracted as the primary text unit for subsequent processing.

Data preprocessing

To mitigate noise, the raw corpus—comprising abstract texts—was subjected to a rigorous normalisation process. All tokens were lowercased and numerical digits and cleansed of superfluous whitespace. Additionally, stopwords from the standard English lexicon were removed to eliminate high-frequency, low-informational content. These preprocessing operations were implemented using the *tm* and *stringr* packages in R. The resulting cleaned corpus served as the input for both topic modelling and word embedding procedures.

Topic modelling

To identify central themes inherent in unstructured textual data, LDA was applied as the primary topic modelling technique. LDA assumes a predefined number of topics (K) and models each document as a mixture of topics, with each topic represented by a distribution over words.

To determine an appropriate K -value, we evaluated model performance using perplexity, a standard metric for topic coherence. Based on this evaluation and interpretability considerations, as Figure 3 illustrated, the number of topics was set to $K = 2$. For each topic, the model identified the top 50 terms with the highest posterior probabilities, serving as representative indicators of thematic significance.

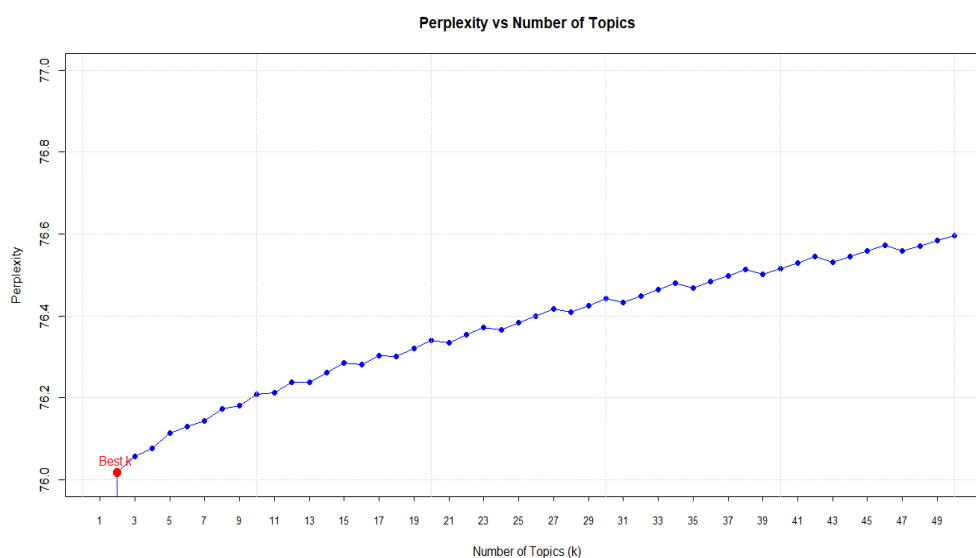


Figure 3. Evaluation of model perplexity across topic counts (K).

Semantic and content analysis

To quantify semantic associations among key terms, this study employed Word2Vec, a neural embedding model that maps words into a continuous multi-dimensional vector space. Within this representation, the spatial proximity of vectors reflects contextual and semantic similarity, forming the computational basis for calculating pairwise correlations among significant terms.

Based on the Word2Vec embedding model, each term was projected into a continuous multi-dimensional vector space to capture its semantic context. To enhance the interpretability of the identified topics and uncover the relationships among key concepts within each thematic cluster, a pairwise similarity analysis was conducted using the top 30 high-probability terms obtained from the topic modelling results.

Semantic proximity between term vectors was quantified using *cosine similarity*, a standard metric that measures the angular closeness between two n-dimensional vectors (Lai et al., 2016). A similarity threshold of 0.5 was applied to define edges in the semantic network, indicating moderate to strong conceptual associations among topic-relevant terms. The cosine similarity formula is as follows:

$$\text{similarity} = \cos\theta = \frac{\bar{x} \cdot \bar{y}}{\|\bar{x}\| \|\bar{y}\|} \quad \text{Equation (1)}$$

where:

$\bar{x} \cdot \bar{y}$: Vector dot product from x and y . $\sum_{k=1}^n x_i y_k$

$\|\bar{x}\|$: Long vector x . $\sum_{k=1}^n x_k^2$

$\|\bar{y}\|$: Long vector y . $\sum_{k=1}^n y_k^2$

Subsequently, content analysis was conducted to reinforce the interpretive validity of the computational findings. Preprints corresponding to each topic were revisited to contextualise and substantiate the semantic relationships identified through modelling, ensuring that algorithmically derived themes align with the substantive contours of emerging XAI scholarship.

This multi-stage process illustrates how computational techniques can systematically transform preprints—from raw, unstructured textual data into structured informational entities and ultimately into interpretable knowledge—thereby advancing the understanding of knowledge emergence and scholarly communication within fast-evolving research domains.

Results

The proposed computational-epistemic framework yielded a dual-layered analytical outcome: (1) Topic modelling results revealed the informational architecture of preprint discourse; (2) Semantic similarity mapping and content analysis elucidated the relational structures that constitute knowledge-level understanding.

Together, these findings empirically demonstrate how the DIKW model can be operationalised through natural language processing to convert unstructured academic text into structured, interpretable knowledge.

Topic modelling results: The informational structure of XAI preprints

As shown in Table 1, the LDA analysis identified two major thematic structures, each characterised by the fifty most probable terms. The probability values represent each term's relative importance within a topic. Collectively, these topics outline the conceptual landscape of XAI preprint research.

A comparative examination of the thematic compositions indicates that *model*, *method*, and *interpretability* emerge as shared core terms, underscoring the centrality of methodological innovation aimed at enhancing model interpretability across the XAI research landscape. Nonetheless, the two topics diverge in both epistemological focus and research orientation.

Topic 1 – Algorithmic development and model optimisation in XAI

Topic 1 is characterized by high-frequency terms such as *learn*, *method*, *feature*, *dataset*, and *approach*. These terms converge around discussions of model construction, algorithmic learning, feature extraction, and data processing. This thematic cluster represents a technology-driven research trajectory, emphasising methodological innovation, model optimisation, and performance improvement.

Topic 2 – user-centered evaluation and model performance assessment

In contrast, Topic 2 encompasses high-frequency terms such as *datum*, *network*, *explanation*, *performance*, *accuracy*, *evaluation*, and *user*. This topic foregrounds evaluation-oriented and user-

centric approaches in XAI, focusing on how models can deliver transparent, accurate, and understandable outputs for end-users.

Topic 1			Topic 2		
Rank	Terms	Probability	Rank	Terms	Probability
1	model	0.036	1	model	0.016
2	learn	0.020	2	datum	0.011
3	method	0.015	3	network	0.009
4	use	0.014	4	much	0.009
5	feature	0.014	5	ai	0.008
6	base	0.011	6	image	0.008
7	approach	0.009	7	interpretability	0.008
8	dataset	0.009	8	method	0.007
9	datum	0.008	9	framework	0.007
10	demonstrate	0.008	10	explanation	0.007
11	task	0.007	11	performance	0.007
12	can	0.007	12	can	0.007
13	propose	0.007	13	prediction	0.007
14	high	0.007	14	propose	0.006
15	train	0.006	15	good	0.006
16	improve	0.005	16	accuracy	0.006
17	neural	0.005	17	base	0.006
18	result	0.005	18	show	0.005
19	work	0.005	19	provide	0.005
20	machine	0.005	20	result	0.005
21	identify	0.005	21	enhance	0.005
22	study	0.005	22	system	0.005
23	deep	0.005	23	also	0.004
24	interpretability	0.005	24	different	0.004
25	paper	0.005	25	use	0.004
26	provide	0.004	26	explainability	0.004
27	novel	0.004	27	study	0.004
28	analysis	0.004	28	state	0.004
29	human	0.004	29	knowledge	0.004
30	time	0.004	30	challenge	0.004
31	performance	0.004	31	graph	0.004
32	process	0.004	32	evaluation	0.004
33	however	0.004	33	language	0.004
34	challenge	0.004	34	interpretable	0.004
35	representation	0.004	35	process	0.004
36	classification	0.003	36	train	0.004
37	develop	0.003	37	information	0.004
38	system	0.003	38	make	0.004
39	prediction	0.003	39	analysis	0.004
40	explainable	0.003	40	xai	0.004
41	technique	0.003	41	however	0.004
42	decision	0.003	42	approach	0.004
43	application	0.003	43	large	0.003
44	design	0.003	44	achieve	0.003
45	score	0.003	45	decision	0.003
46	make	0.003	46	significant	0.003
47	apply	0.003	47	understand	0.003
48	present	0.003	48	structure	0.003

49	exist	0.003	49	address	0.003
50	due	0.003	50	user	0.003

Table 1. Extracted informational elements representing latent topics in XAI preprints.

Conceptual–semantic relationships within thematic structures in emerging XAI research identified from preprints

To extend the informational structures identified through topic modelling, this phase integrates semantic similarity analysis with content analysis to uncover the conceptual relationships underlying XAI preprints. Word2Vec quantitatively models lexical proximities, while content analysis interprets these associations within their discursive context.

Topic 1 — Algorithmic foundations and model optimization in XAI: Integrating the technical core and applied dimension

Semantic similarity analysis revealed that emerging research in algorithmic development and model optimization within XAI exhibits two primary regions of semantic concentration (see Figure 4). The upper-left cluster represents the theme of enhancing model reliability through data-driven and deep learning techniques, whereas the lower-right cluster reflects the semantic domain of model deployment and performance validation. Together, these regions delineate the dual orientation of XAI research—bridging methodological innovation with applied model optimisation.

The upper-left semantic cluster demonstrates a pronounced technical and methodological orientation. High inter-term semantic similarity values—such as *approach–model* (*similarity*=0.587), *technique–method* (*similarity*=0.836), *dataset–datum* (*similarity*=0.518), *dataset–task* (*similarity*=0.360), and *deep–machine* (*similarity*=0.783)—indicate a tight coupling between methodological innovation and computational implementation. Content analysis of preprints corroborates this finding, showing that these terms frequently co-occur in discussions on algorithmic pipelines (Liao et al., 2023), neural architecture refinement (Cao et al., 2023; Li et al., 2024; Tokmak et al., 2024), and the strategic deployment of task-oriented datasets (Cheng et al., 2023; Guo et al., 2024).

The use of domain-specific data is shown to facilitate precise model calibration, improve contextual relevance, and enhance interpretability within distinct analytical settings (Carrion et al., 2023; Talukdar & Biswas, 2024). This alignment between algorithmic innovation and task-adaptive data design underscores the epistemic foundation of this cluster, which represents the technical knowledge nucleus of XAI preprints—where methodological rigor, data specificity, and computational precision collectively drive explainable model advancement.

Conversely, the lower-right semantic cluster is centred on model implementation and performance validation. The observed associations—*propose–introduce* (*similarity*=0.879), *apply–use*, *propose–novel* (*similarity*=0.339), *novel–explainable* (*similarity*=0.327), and *result–performance* (*similarity*=0.431)—suggest that researchers focus on the translation of methodological innovation into empirical application. Content analysis reinforces this interpretation, revealing that authors in this group often describe the adaptation of existing models, the validation of proposed frameworks (Hu et al., 2024), and the extension of prior techniques to new problem domains (Hsiao et al., 2024; Y. Li et al., 2023; Liu et al., 2024).

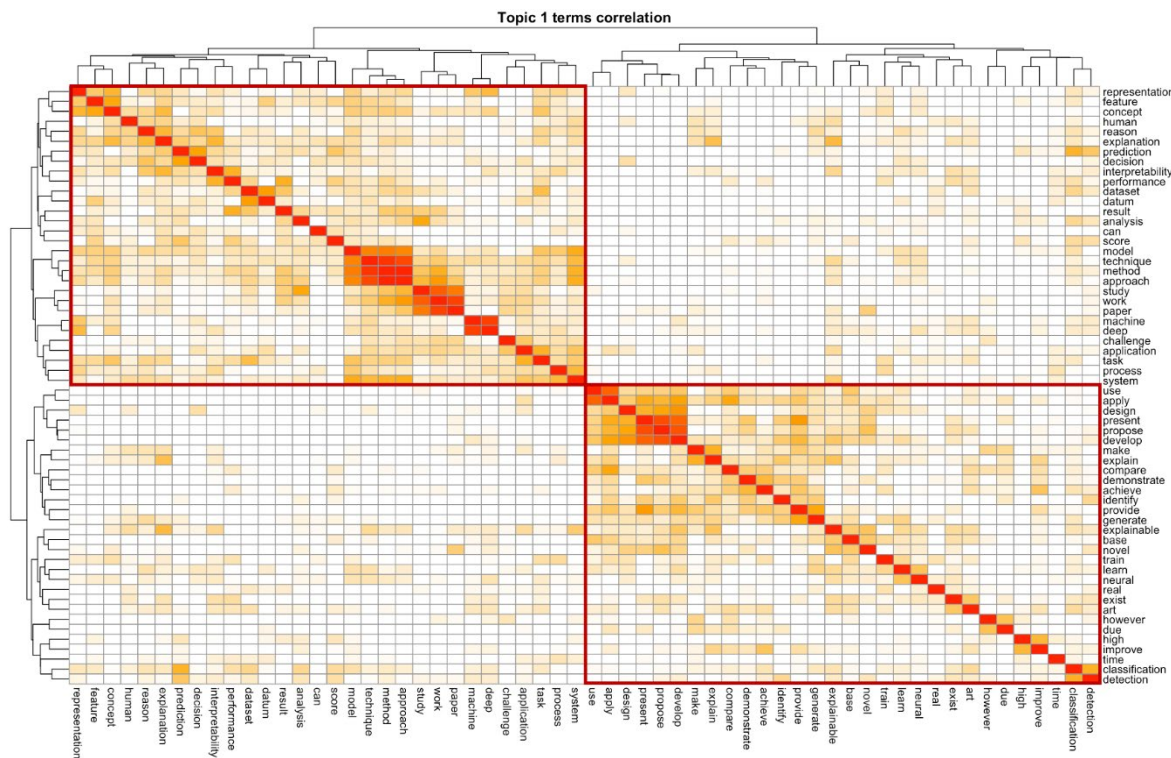


Figure 4. Semantic similarity heatmap of key terms in algorithmic development and model optimization in XAI

Topic 2 — User-centred evaluation and model performance assessment: Integrating constructive actions, interpretability, and human–artificial intelligence interaction

In contrast to the technically grounded orientation of Topic 1, Topic 2 encapsulates a distinct research trajectory centred on user-centred evaluation and performance assessment. The semantic heatmap (Figure 5) reveals two areas of concentrated meaning: the upper-left cluster, which corresponds to procedural and operational activities in model construction, and the lower-right cluster, which centres on interpretability, performance metrics, and human–artificial intelligence interaction—together constituting a semantic progression from technical implementation to user-centred understanding.

The upper-left cluster represents procedural actions underpinning model development. Although these verbs carry lower internal similarity values, their frequent co-occurrence in preprints signals the operational phase preceding interpretive evaluation. Content analysis shows that these terms typically appear in methodological sections, where authors outline model-building steps or data augmentation processes—forming the pragmatic groundwork for subsequent evaluative discourse.

The lower-right cluster, by contrast, constitutes the semantic nucleus of Topic 2. The inter-term associations—*framework–algorithm* (similarity=0.599), *performance–accuracy* (similarity=0.820), *evaluation–result* (similarity=0.372), *information–knowledge* (similarity=0.485), *user–ai* (similarity=0.354), *user–explanation* (similarity=0.347), and *explainability–interpretability* (similarity=0.422)—demonstrate the integration of computational precision with human-centred interpretation. Content analysis of this cluster reveals that many preprints explicitly connect model interpretability to evaluation frameworks (Chen et al., 2024), trustworthiness metrics (Cunnington et al., 2022; Geng et al., 2025), and user-oriented design principles (Bhargava et al., 2023; J. Li et al., 2023; Lichtenberg et al., 2024).

Notably, a substantial portion of these studies involve industry participation, reflecting active collaboration between academia and enterprise in developing human-centric evaluation systems. This cluster thus situates interpretability as both a technical metric and a human-centered communicative process within applied XAI contexts.

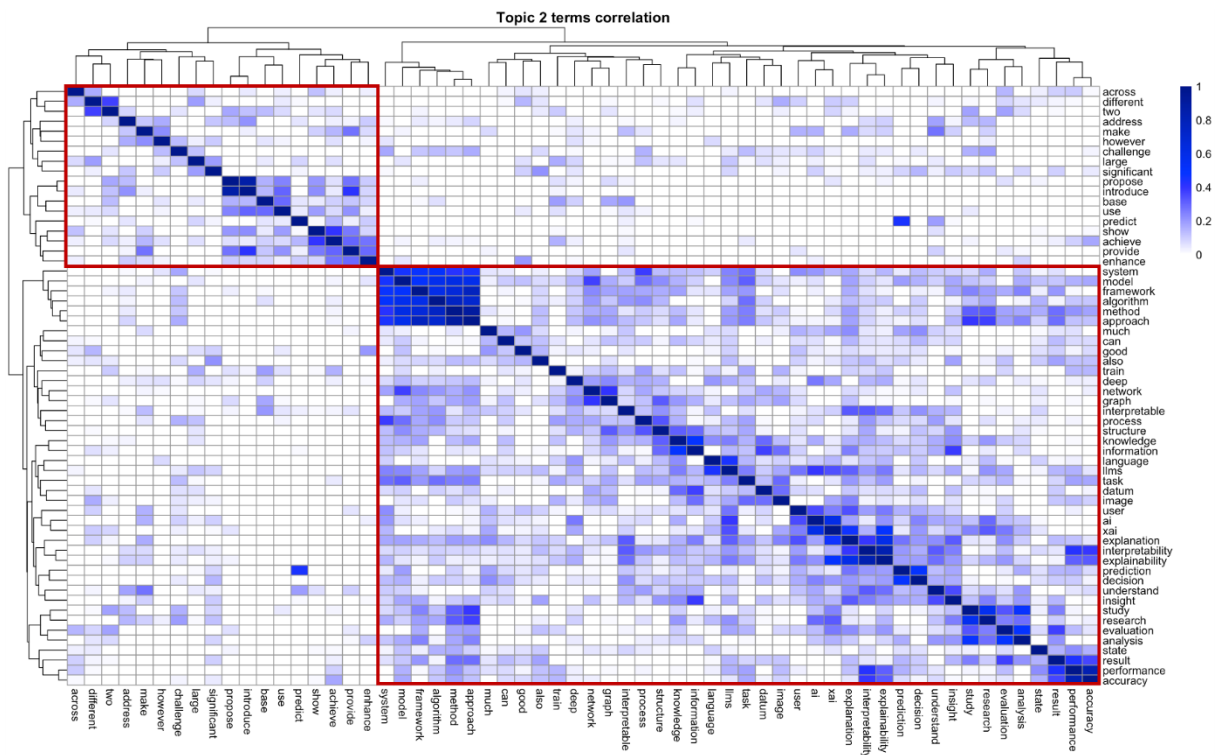


Figure 5. Semantic similarity heatmap of key terms in user-centred evaluation and model performance assessment in XAI

From semantic association to knowledge formation

The integration of semantic association and content analysis represents the epistemic culmination of the DIKW transformation process. Topic modelling delineated the informational scaffolding of preprint discourse, while this phase advances toward semantic coherence and interpretive synthesis. By aligning computationally derived clusters with contextual textual evidence, the analysis constructs a semantic knowledge structure in which conceptual relationships become explicit and epistemically meaningful.

This synthesis illustrates how unstructured textual data evolve into semantically coherent knowledge representations, reflecting how the XAI research community conceptualises algorithmic innovation, interpretability, and human trust in the broader context of emerging scientific knowledge.

Discussion and conclusion

Research findings and practical insights

Grounded in the DIKW theoretical framework, this study examined how preprints can be conceptualised as early-stage information objects and systematically transformed into actionable scholarly insights. From an information science perspective, the study positions preprints within an information creation–transformation–use continuum, demonstrating their value for sense-making and early assessment of emerging research directions. Recognising the heterogeneity and uneven quality of preprint repositories, a two-stage quality filtering mechanism was implemented

to ensure a baseline level of scholarly validity and informational integrity, rather than treating preprints as undifferentiated textual data.

The proposed computational knowledge transformation framework—integrating latent Dirichlet allocation (LDA), Word2Vec semantic modelling, and content analysis—illustrates how unstructured textual data can be progressively organised into structured information and synthesised into meaningful knowledge through a mixed-methods approach. Applied to explainable artificial intelligence (XAI) preprints, the framework identified two dominant thematic trajectories: algorithmic development and model optimisation and user-centred evaluation and performance assessment. These trajectories reflect not only patterns of technical innovation but also differences in how information is interpreted, evaluated, and used across scholarly and applied contexts, highlighting the dual emphasis in XAI on methodological advancement and human-centred interpretation.

Collectively, the findings advance both methodological and epistemological insights. Methodologically, the study demonstrates the value of integrating quantitative modeling with interpretive analysis to support large-scale knowledge transformation. Epistemologically, it shows that when subjected to systematic filtering and semantic computation, preprints can function as real-time observatories of knowledge emergence rather than merely preliminary communications. By operationalising DIKW transformations computationally, this research underscores the role of preprints as early indicators of research frontiers and as data infrastructures for understanding the diffusion, interpretation, and use of scientific information.

Limitations and future work

While this study demonstrates the feasibility of using preprints for computational knowledge transformation and technological trend monitoring, several limitations should be acknowledged. First, the dataset was limited to English-language XAI preprints published between 2023 and 2025, which may constrain the cross-disciplinary and cross-linguistic generalisability of the findings. Second, although the two-stage quality filtering mechanism enhanced data reliability, elements of manual verification inevitably involved subjective judgment. Finally, while Word2Vec enabled the identification of quantitative semantic associations, the subsequent interpretive synthesis relied primarily on researcher interpretation, which may introduce epistemic and interpretive bias.

Future research may extend this framework in several directions. Applying it across multiple disciplines and languages would allow for systematic evaluation of its robustness and scalability. Incorporating institutional- and author-level indicators could further refine quality assurance in preprint-based analytics. Importantly, future work may integrate expert validation mechanisms—such as Delphi-based or panel-based approaches—to complement computational modelling. Such integration would enrich semantic interpretation, strengthen epistemic validity, and provide a structured form of human-machine collaboration in large-scale knowledge analysis.

About the authors

Tzu-Yu Lin is a PhD student in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University. She also serves as a researcher at the Science & Technology Policy Research and Information Center, National Institutes of Applied Research. Her research interests include information science, data governance, and bibliometrics. She can be contacted at tzyulin@nlar.org.tw.

Pei-Chun Lee is an Associate Professor at the Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Taiwan. She received her PhD in Technology Management. Her research interests include service innovation, innovation systems, patent analytics, and sustainability and resilience. She can be contacted at pcllee@nccu.edu.tw.

References

- Alfonso, F., & Crea, F. (2023). Preprints: a game changer in scientific publications? In: Oxford University Press US. <https://doi.org/10.1093/eurheartj/ehac665>
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., & Benjamins, R. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Assaf, S., & Srour, I. (2021). Using a data driven neural network approach to forecast building occupant complaints. *Building and Environment*, 200, 107972. <https://doi.org/10.1016/j.buildenv.2021.107972>
- Aviv-Reuven, S., Bronstein, J., & Rosenfeld, A. (2024). Exploring scholarly perceptions of preprint servers. *Information Research an international electronic journal*, 29(2), 173–178. <https://doi.org/10.47989/ir292820>
- Baskarada, S., & Koronios, A. (2013). Data, information, knowledge, wisdom (DIKW): A semiotic theoretical and empirical exploration of the hierarchy and its quality dimension. *Australasian Journal of Information Systems*, 18(1), 5–19.
- Benkassioui, B., Kharmoum, N., Hadi, M. Y., & Ezziyyani, M. (2022). NLP methods' information extraction for textual data: an analytical study. *International Conference on Advanced Intelligent Systems for Sustainable Development*,
- Berg, J. M., Bhalla, N., Bourne, P. E., Chalfie, M., Drubin, D. G., Fraser, J. S., Greider, C. W., Hendricks, M., Jones, C., & Kiley, R. (2016). Preprints for the life sciences. *Science*, 352(6288), 899–901. <https://doi.org/10.12815/kits.2026.25.1.75>
- Bhargava, S., Dhoot, A., Jonsson, I.-M., Nguyen, H. L., Patel, A., Yu, H., & Renkens, V. (2023). Referring to Screen Texts with Voice Assistants. arXiv:2306.07298. Retrieved June 01, 2023, from <https://ui.adsabs.harvard.edu/abs/2023arXiv230607298B>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993–1022.
- Bo, Y., & Gang, W. (2017, 2017/03). A Hybrid HDP-ME-LDA Model for Sentiment Analysis. *Proceedings of the 2017 2nd International Conference on Automation, Mechanical Control and Computational Engineering (AMCCE 2017)*. <https://doi.org/10.2991/amcce-17.2017.114>
- Cao, D., Chen, G., Jiang, J., Yu, J., Zhang, R., Chen, M., Zhang, W., Chen, L., Zhong, F., & Zhang, Y. (2023). EquiScore: A generic protein-ligand interaction scoring method integrating physical prior knowledge with data augmentation modeling. *BioRxiv*, 2023.2006.2018.545464. <https://doi.org/10.1101/2023.06.18.545464>
- Carrion, J., Nandakumar, R., Shi, X., Gu, H., Kim, Y., Raskind, W. H., Peter, B., & Dinu, V. (2023). A data-fusion approach to identifying developmental dyslexia from multi-omics datasets. *BioRxiv*, 2023.2002.2027.530280. <https://doi.org/10.1101/2023.02.27.530280>
- Castelvecchi, D. (2016). Can we open the black box of AI? *Nature News*, 538(7623), 20.
- Celard, P., Vieira, A. S., Iglesias, E., & Borrajo, L. (2020). LDA filter: A Latent Dirichlet Allocation preprocess method for Weka. *Plos one*, 15(11), e0241701. <https://doi.org/10.1371/journal.pone.0241701>

- Chaleplioglou, A., & Koulouris, A. (2023). Preprint paper platforms in the academic scholarly communication environment. *Journal of Librarianship and Information Science*, 55(1), 43–56. <https://doi.org/10.1177/09610006211058908>
- Chen, Y., Ren, K., Song, K., Wang, Y., Wang, Y., Li, D., & Qiu, L. (2024). EEGFormer: Towards Transferable and Interpretable Large-Scale EEG Foundation Model. arXiv:2401.10278. Retrieved January 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv240110278C>
- Cheng, H., Jafari, R., Russell, A., Klopfer, R., Lu, E., Striner, B., & Gormley, M. R. (2023). MDACE: MIMIC Documents Annotated with Code Evidence. arXiv:2307.03859. Retrieved July 01, 2023, from <https://ui.adsabs.harvard.edu/abs/2023arXiv230703859C>
- Cohen, Y., Madhavan, M., Slomka, P., Iskandrian, A., & Einstein, A. (2020). Preprint manuscripts and servers in the era of coronavirus disease 2019. *Journal of Evaluation in Clinical Practice*, 27(1), 16–21. <https://doi.org/10.1111/jep.13498>
- Cunnington, D., Law, M., Lobo, J., & Russo, A. (2022). Neuro-symbolic learning of answer set programs from raw data. arXiv:2205.12735. Retrieved May 01, 2022, from <https://ui.adsabs.harvard.edu/abs/2022arXiv220512735C>
- Dhande, M. (2021). Data was, Analytics is, the New Oil. Geospatial World. Retrieved October 2 from <https://www.geospatialworld.net/prime/business-and-industry-trends/data-was-analytics-is-the-new-oil/>
- Di Gennaro, G., Buonanno, A., & Palmieri, F. A. (2021). Considerations about learning Word2Vec. *The Journal of Supercomputing*, 77(11), 12320–12335. <https://doi.org/10.1007/s11227-021-03743-2>
- Dieng, A. B., Ruiz, F. J., & Blei, D. M. (2020). Topic modeling in embedding spaces. *Transactions of the Association for Computational Linguistics*, 8, 439–453. https://doi.org/10.1162/tacl_a_00325
- Dos Santos, C., & Gatti, M. (2014). Deep convolutional neural networks for sentiment analysis of short texts. Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers,
- Fang, D., Yang, H., Gao, B., & Li, X. (2018). Discovering research topics from library electronic references using latent Dirichlet allocation. *Library Hi Tech*, 36(3), 400–410. <https://doi.org/10.1108/LHT-06-2017-0132>
- Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálffy, M., Nanni, F., & Coates, J. A. (2020). Preprinting the COVID-19 pandemic. *BioRxiv*, 2020.2005. 2022.111294. <https://doi.org/10.1101/2020.05.22.111294>
- Fraser, N., Brierley, L., Dey, G., Polka, J. K., Pálffy, M., Nanni, F., & Coates, J. A. (2021). The evolving role of preprints in the dissemination of COVID-19 research and their impact on the science communication landscape. *PLoS biology*, 19(4), e3000959. <https://doi.org/10.1371/journal.pbio.3000959>
- Frické, M. (2009). The knowledge pyramid: a critique of the DIKW hierarchy. *Journal of information science*, 35(2), 131–142. <https://doi.org/10.1177/0165551508094050>
- Geng, C., Xu, X., Xing, A., Zhao, Z., & Si, X. (2025). From neural representations to interpretable logic rules. arXiv:2501.08281. Retrieved January 01, 2025, from <https://ui.adsabs.harvard.edu/abs/2025arXiv250108281G>

- Guo, R., Chen, B., Zhang, Q., Zhao, Y., Wang, X., & Zhu, Z. (2024). A data-driven crowd simulation framework integrating physics-informed machine learning with navigation potential fields. *arXiv:2410.16132*. Retrieved October 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv241016132G>
- Hsiao, C.-H., Kumar, K., & Rathje, E. (2024). Explainable AI models for predicting liquefaction-induced lateral spreading. *arXiv:2404.15959*. Retrieved April 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv240415959H>
- Hu, K., Luo, Q., Qi, K., Yang, S., Mao, J., Fu, X., Zheng, J., Wu, H., Guo, Y., & Zhu, Q. (2019). Understanding the topic evolution of scientific literatures like an evolving city: Using Google Word2Vec model and spatial autocorrelation analysis. *Information Processing & Management*, 56(4), 1185–1203. <https://doi.org/10.1016/j.ipm.2019.02.014>
- Hu, Z., Zheng, T., & Huang, H. (2024). A Bayesian approach to harnessing the power of LLMs in authorship attribution. *arXiv:2410.21716*. Retrieved October 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv241021716H>
- Hussain, M., Satti, F. A., Ali, S. I., Hussain, J., Ali, T., Kim, H.-S., Yoon, K.-H., Chung, T., & Lee, S. (2021). Intelligent knowledge consolidation: From data to wisdom. *Knowledge-Based Systems*, 234, 107578. <https://doi.org/10.1016/j.knosys.2021.107578>
- Jacovi, A. (2023). Trends in explainable AI (XAI) literature. *arXiv preprint arXiv:2301.05433*. <https://doi.org/10.48550/arXiv.2301.05433>
- Kaiser, J. (2017). The preprint dilemma. *Science* 357, 1344–1347. <https://doi.org/10.1126/science.357.6358.1344>
- Klein, M., Broadwell, P., Farb, S. E., & Grappone, T. (2016). Comparing published scientific journal articles to their pre-print versions. *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*. <https://doi.org/10.1145/2910896.2910909>
- Kodvanj, I., Homolak, J., Virag, D., & Trkulja, V. (2022). Publishing of COVID-19 preprints in peer-reviewed journals, preprinting trends, public discussion and quality issues. *Scientometrics*, 127(3), 1339–1352. <https://doi.org/10.1007/s11192-021-04249-7>
- Kwon, K., Jun, S., Lee, Y.-J., Choi, S., & Lee, C. (2022). Logistics technology forecasting framework using patent analysis for technology roadmap. *Sustainability*, 14(9), 5430. <https://doi.org/10.3390/su14095430>
- Lai, S., Liu, K., He, S., & Zhao, J. (2016). How to generate a good word embedding. *IEEE Intelligent Systems*, 31(6), 5–14. <https://doi.org/10.1109/MIS.2016.45>
- Langnickel, L., Baum, R., Darms, J., Madan, S., & Fluck, J. (2021). COVID-19 preVIEW: Semantic search to explore COVID-19 research preprints. In *Public Health and Informatics* (pp. 78–82). IOS Press. <https://doi.org/10.3233/SHTI210124>
- Larsen, K. R., Monarchi, D. E., Hovorka, D. S., & Bailey, C. N. (2008). Analyzing unstructured text data: Using latent categorization to identify intellectual communities in information systems. *Decision Support Systems*, 45(4), 884–896. <https://doi.org/10.1016/j.dss.2008.02.009>
- Lei, C., Dai, H., Yu, Z., & Li, R. (2020). A service recommendation algorithm with the transfer learning based matrix factorization to improve cloud security. *Information Sciences*, 513, 98–111. <https://doi.org/10.1016/j.ins.2019.10.004>

- Lei, L., Deng, Y., & Liu, D. (2023). Examining research topics with a dependency-based noun phrase extraction method: a case in accounting. *Library Hi Tech*, 41(2), 570–582. <https://doi.org/10.1108/LHT-12-2019-0247>
- Li, J., Zhang, W., Wang, T., Xiong, G., Lu, A., & Medioni, G. (2023). GPT4Rec: A generative framework for personalized recommendation and user interests interpretation. *arXiv:2304.03879*. Retrieved April 01, 2023, from <https://ui.adsabs.harvard.edu/abs/2023arXiv230403879L>
- Li, M., Peng, P., Zhang, J., Wang, H., & Shen, W. (2024). SCCAM: Supervised Contrastive Convolutional Attention Mechanism for ante-hoc interpretable fault diagnosis with limited fault samples. *IEEE Transactions on Neural Networks and Learning Systems*, 35(5), 6194–6205. <https://doi.org/10.1109/TNNLS.2023.3313728>
- Li, Y., Wang, J., Dai, X., Wang, L., Yeh, C.-C. M., Zheng, Y., Zhang, W., & Ma, K.-L. (2023). How does attention work in vision transformers? A visual analytics attempt. *arXiv:2303.13731*. Retrieved March 01, 2023, from <https://ui.adsabs.harvard.edu/abs/2023arXiv230313731L>
- Liao, W., Porte-Agel, F., Fang, J., Bak-Jensen, B., Yang, Z., & Zhang, G. (2023). Improving the accuracy and interpretability of neural networks for wind power forecasting. *arXiv:2312.15741*. Retrieved December 01, 2023, from <https://ui.adsabs.harvard.edu/abs/2023arXiv231215741L>
- Lichtenberg, J. M., Di Benedetto, G., & Ruffini, M. (2024). Ranking across different content types: The robust beauty of multinomial blending. *arXiv:2408.09168*. Retrieved August 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv240809168L>
- Limwattana, S., & Prom-on, S. (2021). Topic modeling enhancement using word embeddings. 2021 18th International Joint Conference on Computer Science and Software Engineering (JCSSE). <https://doi.org/10.1109/JCSSE53117.2021.9493816>
- Liu, J., Wang, Y., Du, J., Tianyi Zhou, J., & Liu, Z. (2024). MedCoT: Medical Chain of Thought via hierarchical expert. *arXiv:2412.13736*. Retrieved December 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv241213736L>
- Ma, J., Wang, L., Zhang, Y.-R., Yuan, W., & Guo, W. (2023). An integrated latent Dirichlet allocation and Word2vec method for generating the topic evolution of mental models from global to local. *Expert Systems with Applications*, 212, 118695. <https://doi.org/10.1016/j.eswa.2022.118695>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26. <https://proceedings.neurips.cc/paper/2013/file/9aa42b31882ec039965f3c4923ce901b-Paper.pdf>
- Mu, W., Lim, K. H., Liu, J., Karunasekera, S., Falzon, L., & Harwood, A. (2022). A clustering-based topic model using word networks and word embeddings. *Journal of big data*, 9(1), 38. <https://doi.org/10.1186/s40537-022-00585-4>
- Peng, M., Gao, W., Wang, H., Zhang, Y., Huang, J., Xie, Q., Hu, G., & Tian, G. (2017). Parallelization of massive textstream compression based on compressed sensing. *ACM Transactions on Information Systems (TOIS)*, 36(2), 1–18. <https://doi.org/10.1145/3086702>
- Peters, M. A., Jandrić, P., & Green, B. J. (2024). The DIKW model in the age of artificial intelligence. *Postdigital Science and Education*. <https://doi.org/10.1007/s42438-024-00462-8>

- Pontis, S., Blandford, A., Greifeneder, E., Attalla, H., & Neal, D. (2017). Keeping up to date: An academic researcher's information journey. *Journal of the Association for Information Science and Technology*, 68(1), 22–35. <https://doi.org/https://doi.org/10.1002/asi.23623>
- Rieger, J., Jentsch, C., & Rahnenführer, J. (2024). LDAPprototype: A model selection algorithm to improve reliability of latent Dirichlet allocation. *PeerJ Computer Science*, 10, e2279. <https://doi.org/10.7717/peerj-cs.2279>
- Roberts, M. E., Stewart, B. M., Tingley, D., & Airoldi, E. M. (2013). The structural topic model and applied social science. *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*, <https://mimno.infosci.cornell.edu/nips2013ws/slides/stm.pdf>
- Rowley, J. (2007). The wisdom hierarchy: Representations of the DIKW hierarchy. *Journal of information science*, 33(2), 163–180. <https://doi.org/10.1177/0165551506070706>
- Sarabipour, S., Debat, H. J., Emmott, E., Burgess, S. J., Schwessinger, B., & Hensel, Z. (2019). On the value of preprints: An early career researcher perspective. *PLoS biology*, 17(2), e3000151. <https://doi.org/10.1371/journal.pbio.3000151>
- Talukdar, W., & Biswas, A. (2024). Improving Large Language Model (LLM) fidelity through context-aware grounding: A systematic approach to reliability and veracity. arXiv:2408.04023. Retrieved August 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv240804023T>
- Tokmak, A., Schön, T. B., & Baumann, D. (2024). PACSBO: Probably approximately correct safe Bayesian optimization. arXiv:2409.01163. Retrieved September 01, 2024, from <https://ui.adsabs.harvard.edu/abs/2024arXiv240901163T>
- Wang, Z., Glänzel, W., & Chen, Y. (2020). The impact of preprints in library and information science: an analysis of citations, usage and social attention indicators. *Scientometrics*, 125(2), 1403–1423. <https://doi.org/10.1007/s11192-020-03612-4>
- Yoon, S. (2023). Building digital twinning: Data, information, and models. *Journal of Building Engineering*, 76, 107021. <https://doi.org/10.1016/j.jobe.2023.107021>
- Zadgaonkar, A., & Agrawal, A. J. (2024). An approach for analyzing unstructured text data using topic modeling techniques for efficient information extraction. *New Generation Computing*, 42(1), 109–134. <https://doi.org/10.1007/s00354-023-00230-5>
- Zhang, F., Gao, W., & Fang, Y. (2019). News title classification based on sentence-LDA model and word embedding. 2019 *International Conference on Machine Learning, Big Data and Business Intelligence (MLBDI)*. <https://doi.org/10.1109/MLBDI48998.2019.00053>

© [CC-BY-NC 4.0](#) The Authors. For more information, see our [Open Access Policy](#).