Analysis of the evolution of key terms in scientific publications on the topic of "Artificial intelligence"*

Oleh Dmytrenko[1,*], Anton Astakhov[1,†], Oksana Vlasenko[1,†], Volodymyr Tsyhanok[1,†] and Serhii Patuk[1,†]

[1] *Institute for Information Recording of the National Academy of Sciences of Ukraine, Kyiv, Ukraine*

Abstract

This paper proposes a methodology for analyzing the evolution of key terms in scientific publications on artificial intelligence during 2015–2024 based on Directed Weighted Networks of Terms (DWNT). The study aims to identify structural and semantic changes in the Artificial Intelligence terminological space and to reveal emerging research trends. The experimental dataset is formed from arXiv publications in the Computer Science – Artificial Intelligence (AI) category. The methodology includes linguistic preprocessing (tokenization, lemmatization, and part-of-speech tagging), extraction of multi-word terms using syntactic templates, computation of Global Term Frequency (GTF), and construction of yearly DWNTs, where nodes represent terms and edges encode semantic inclusion and co-occurrence relationships. Separate DWNTs are built for each year, enabling longitudinal analysis of term dynamics and network structure. The top-10 terms by GTF are identified annually, while newly emerging terms and changes in network connectivity are tracked over time. The results demonstrate stable growth of foundational concepts such as model and learning, gradual decline of generic terms such as problem, and a sharp rise of large language models (LLM) after 2022, reflecting the paradigm shift toward generative and foundation models. The obtained results confirm the effectiveness of DWNT for capturing paradigm shifts in scientific discourse and demonstrate its applicability for scientometric monitoring and trend detection in artificial intelligence research.

Keywords

artificial intelligence, term evolution, directed weighted network of terms, semantic networks, scientometrics, natural language processing, language models, global term frequency, network dynamics

1. Introduction

The rapid development of artificial intelligence (AI) technologies over the past decade has led to the emergence of new research directions, methods, and terms describing their essence [1–3]. The dynamics of new concept formation and transformation of established terminology reflects changes in scientific discourse and research priorities, particularly the transition from classical machine learning methods to deep and generative models [2, 4]. However, traditional linguistic and bibliometric analysis methods do not provide a comprehensive understanding of semantic relationships between terms and their temporal evolution [5–7].

An actual scientific task is identifying regularities in the development of the AI terminological space, in particular studying how the meaning and usage context of key terms changes in scientific publications [8, 9]. This requires approaches combining natural language processing (NLP) [10, 11], statistical analysis of thematic structures [12], and network modeling of relationships between terms and scientific communities [13].

*ITS-2025: Information Technologies and Security, December 11, 2025, Kyiv, Ukraine

* Corresponding author.

† These authors contributed equally.

dmytrenko.o@gmail.com (O. Dmytrenko); astakhov.anton97@gmail.com (A. Astakhov); vo12062007@gmail.com (O. Vlasenko); vvtsyg@gmail.com (V. Tsyhanok); filtronsua@gmail.com (S. Patuk)

0000-0001-8501-5313 (O. Dmytrenko); 0000-0002-9630-1101 (A. Astakhov); 0000-0003-4292-6218 (O. Vlasenko); 0009-0005-0717-1077 (V. Tsyhanok); 0009-0009-7313-402X (S. Patuk)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

published 2026-11-26

CEUR-WS.org/Vol-4337/paper20.pdf

1.1. Related work

Textual data analysis in scientific publications has been widely studied through keyword extraction, topic modeling, and semantic network construction. Topic modeling approaches such as Latent Semantic Analysis, Probabilistic Latent Semantic Analysis, and Latent Dirichlet Allocation have been extensively reviewed [14], demonstrating their effectiveness in identifying latent thematic structures and document-level topic distributions.

Word co-occurrence networks demonstrated the effectiveness of graph-based representations where nodes correspond to words and edges represent contextual co-occurrence frequencies [15]. Such graphs enable identification of central concepts, topic clusters, and latent relationships [16], allowing researchers to uncover structural properties of scientific discourse, detect emerging themes, and analyze semantic proximity between concepts. Semantic networks further support modeling conceptual relations and integration with ontologies [17], facilitating knowledge representation and enabling interoperability with domain-specific semantic frameworks.

Recent studies have also explored hybrid approaches combining statistical topic models with network-based representations, aiming to improve interpretability and capture higher-order semantic structures [18, 19]. Graph-enhanced topic modeling frameworks and semantic co-occurrence networks enable identification of latent conceptual communities and improve topic coherence by incorporating structural constraints into probabilistic models. However, despite these advances, most existing approaches primarily focus on static corpora and aggregate statistics, without explicitly modeling temporal dynamics of terminology [20]. This limitation significantly restricts their capability to analyze long-term conceptual evolution, paradigm shifts, and the emergence of novel scientific constructs. While diachronic word embeddings provide partial insight into semantic drift [21], they lack explicit representation of relational semantics and hierarchical term expansion, which are essential for tracing conceptual reorganization in rapidly evolving domains such as artificial intelligence.

In recent years, Dynamic Topic Modeling (DTM) has been proposed to capture temporal variations in topic distributions across document collections by introducing time-dependent latent variables that model how topics and their vocabularies evolve over successive periods [22, 23]. These models extend classical probabilistic topic models such as Latent Dirichlet Allocation by allowing topic parameters to drift over time, thus enabling tracking of shifts in macro-level thematic trends and vocabulary usage in longitudinal corpora. Although DTM provides valuable insights into high-level topical evolution, it typically operates at the level of latent topic distributions and does not explicitly represent fine-grained semantic relationships between individual terms. Moreover, the probabilistic nature of topic evolution limits interpretability of exact term-to-term transformations and hinders direct characterization of semantic dependencies between emerging concepts.

Directed Weighted Networks of Terms (DWNT) overcome these limitations by explicitly representing key terms as nodes and their semantic-syntactic interactions as directed weighted edges [24–26]. Unlike dynamic topic models and embedding-based approaches, DWNT preserves linguistic structure while incorporating quantitative importance measures, enabling direct observation of term expansion, contraction, and integration into evolving semantic cores. This network-based formalism allows simultaneous analysis of local term dynamics and global structural reconfiguration, making it particularly effective for detecting paradigm shifts, identifying newly central concepts, and characterizing long-term evolution of scientific terminology.

1.2. Research relevance

Between 2015 and 2024, the field of artificial intelligence underwent profound transformations, evolving from classical machine learning paradigms toward deep learning architectures and, more

recently, generative AI systems [27, 28]. During the first half of this period, major breakthroughs in deep neural networks significantly improved performance in computer vision and natural language processing tasks, establishing deep learning as the dominant methodological paradigm [1, 4, 27, 28].

These technological shifts were accompanied not only by the emergence of novel algorithms and training strategies but also by the rapid formation of new terminological structures, particularly within generative artificial intelligence. Concepts such as *foundation models*, *transformers*, and *large language models* gradually became central elements of contemporary AI discourse, serving as conceptual foundations for a wide range of generative applications [27, 28]. Generative models now encompass systems capable of synthesizing text, images, audio, video, and multimodal content from large-scale training data, opening unprecedented opportunities for creative, scientific, and industrial use cases.

The appearance of terms such as *foundation models*, *transformers*, *diffusion models*, and *large language models* can therefore be interpreted as a direct reflection of evolving data processing strategies, learning paradigms, and model scaling principles. In this context, terminology acts as an observable proxy for deeper methodological transitions, including the shift from task-specific architectures toward general-purpose pretrained models and from supervised learning toward self-supervised and generative frameworks.

Consequently, the development of new methodologies for analyzing terminological evolution in scientific publications is both timely and necessary, addressing several critical objectives:

- tracking disciplinary trends by identifying emerging concepts and monitoring their frequency dynamics over time;
- detecting shifts in research priorities by comparing longitudinal terminological patterns, revealing how novel techniques gradually displace earlier approaches;
- forecasting future developments by leveraging term dynamics to anticipate emerging research directions, technological innovations, and application domains.

Thus, terminological analysis not only documents the historical trajectory and current state of artificial intelligence but also functions as a strategic instrument for researchers, developers, and industry practitioners seeking to remain at the forefront of technological progress. The development of automated methodologies capable of tracking changes in key terms and the structure of their semantic relationships represents an essential step toward deeper understanding of the semantic dynamics underlying scientific knowledge production.

1.3. Objective and research tasks

The objective of this study is to develop and validate a methodology for analyzing the evolution of key terms in artificial intelligence scientific publications by constructing Directed Weighted Networks of Terms (DWNT) over the period 2015–2024. The proposed approach aims to quantitatively capture semantic dynamics, structural transformations, and emerging conceptual trends within the AI research domain through longitudinal network modeling.

To achieve this objective, the following research tasks are addressed:

1. To construct a longitudinal corpus of artificial intelligence research articles published between 2015 and 2024 using open-access sources, primarily arXiv, with filtering by relevant subject categories.
2. To design and implement a DWNT construction algorithm for each year, incorporating linguistic preprocessing (tokenization, part-of-speech tagging, lemmatization), syntactic key-term extraction, Global Term Frequency estimation, and directed edge formation based on term inclusion and positional relationships.

3. To analyze terminological evolution by identifying yearly top-10 key concepts and investigating variations in inter-term connectivity, emergence of novel terms, and integration of new concepts into the semantic core.
4. To perform temporal DWNT analysis by generating yearly semantic graphs and examining changes in network topology, including node centrality, connectivity patterns, and structural cohesion.
5. To evaluate the effectiveness of the proposed methodology for detecting terminological trends, paradigm shifts, and conceptual consolidation in artificial intelligence research.

Through these tasks, the study seeks to demonstrate that DWNT-based modeling provides a robust framework for longitudinal scientometric analysis and offers enhanced interpretability of semantic evolution compared to traditional frequency-based or topic-modeling approaches.

2. Research methodology

The proposed methodology for analyzing terminological evolution in artificial intelligence research is based on the construction and examination of Directed Weighted Networks of Terms (DWNT), which represent semantic relationships between key textual concepts through directed and weighted graph structures [24–26]. DWNT enables explicit modeling of both local term interactions and global semantic organization, providing a quantitative framework for longitudinal analysis of scientific discourse.

The methodology consists of four main stages:

1. Corpus formation,
2. DWNT construction,
3. Analysis of network structures,
4. Analysis of terminological evolution.

Each stage is designed to capture complementary aspects of semantic dynamics, from linguistic preprocessing to network-level trend detection.

2.1. Corpus formation

The study employs a longitudinal corpus of scientific publications related to Artificial Intelligence covering the period from 2015 to 2024. The corpus was obtained from arXiv, an open-access repository of scientific preprints that provides comprehensive coverage of contemporary AI research.

Specifically, the publicly available arXiv Scientific Research Papers Dataset [27] was utilized, containing metadata and textual abstracts of publications across multiple scientific domains, including the Computer Science – Artificial Intelligence (cs.AI) category. Articles were filtered by publication year and subject classification to ensure thematic consistency and temporal continuity.

For each year, a dedicated sub-corpus was constructed, enabling independent DWNT generation and facilitating comparative temporal analysis. Only English-language abstracts were retained to ensure uniformity of linguistic preprocessing. Duplicate records and incomplete entries were removed during data cleaning.

This corpus formation strategy ensures representative sampling of AI research outputs while maintaining scalability for large-scale semantic analysis.

2.2. DWNT construction

Key terms were identified using syntactic phrase templates that reflect typical structural patterns observed in scientific discourse. The following n-gram configurations were employed:

- Bigrams (2-grams): ***ADJ_NOUN***
- Trigrams (3-grams): ***NOUN_CCONJ_NOUN***, ***ADJ_ADJ_NOUN***

These templates enable extraction of semantically meaningful noun phrases while preserving grammatical coherence. Only tokens tagged as nouns (***NOUN***), adjectives (***ADJ***), and coordinating conjunctions (***CCONJ***) were considered. Proper nouns (***PROPN***) were mapped to ***NOUN*** to ensure uniform treatment of domain-specific concepts.

For each extracted term, a tuple consisting of three elements was constructed:

- the term itself (single word or syntactically formed phrase);
- the corresponding part-of-speech tag or composite template label;
- a numerical Global Term Frequency (GTF) value.

Global Term Frequency was used as a corpus-level importance measure and is defined as:

$$GTF = \frac{n_i}{\sum_k n_k}, \quad (1)$$

where n_i denotes the number of occurrences of term i within the corpus, and $\sum_k n_k$ represents the total number of extracted terms across all documents.

Unlike TF-IDF, which downweights globally frequent terms, GTF emphasizes domain-relevant concepts that consistently appear throughout the corpus. This property is particularly advantageous for specialized scientific collections, where key terms are expected to recur across multiple publications.

Consequently, GTF provides a robust estimate of terminological significance for subsequent network construction.

The ordered sequence of extracted terms within each document was preserved to support structural graph generation. Longer multi-word expressions were prioritized over their constituent subterms, ensuring hierarchical consistency during edge formation.

The construction of the Directed Weighted Network of Terms was performed in three sequential stages designed to capture both local syntactic dependencies and global semantic structure:

1. Stage 1: Undirected Network Generation via Horizontal Visibility Graph
2. Stage 2: Direction Assignment of Edges
3. Stage 3: Edge Weight Computation and Aggregation

2.2.1. Stage 1: Undirected network generation via Horizontal Visibility Graph

At the initial stage, an undirected term network was constructed using the Horizontal Visibility Graph (HVG) algorithm [30]. Originally proposed for transforming time series into complex networks, HVG was adapted in this study to operate on ordered sequences of extracted terms, where each term corresponds to a node and its Global Term Frequency (GTF) value is treated as a vertical height.

Within each sentence, terms were placed along the horizontal axis according to their original textual order, while GTF values defined their vertical coordinates. Two nodes t_i and t_j were considered horizontally visible – and thus connected by an undirected edge – if and only if the following condition holds:

$$x_k < min(x_i, x_j), \forall\, t_k : t_i < t_k < t_j, \quad (2)$$

where x_i, x_j and x_k denote the GTF values of the corresponding terms.

This criterion ensures that semantic connections are established only between terms that are not obscured by intermediate terms of higher importance, effectively modeling prominence-based

relational visibility. The HVG approach provides several advantages, including invariance to affine transformations and computational efficiency, while preserving structural characteristics of the original term sequence.

Applying HVG at the sentence level allows capturing fine-grained contextual interactions, which are subsequently aggregated across the entire corpus.

2.2.2. Stage 2: Direction assignment of edges

After constructing the undirected visibility graph, semantic directionality was introduced to transform the network into a directed structure.

Edge orientation was determined according to two complementary rules:

1. If a shorter term was contained within a longer multi-word expression, the direction was assigned from the shorter term toward its extended form, reflecting semantic specialization.
2. For all remaining connections, directionality was assigned from left to right based on the original word order in the sentence, preserving discourse flow and syntactic precedence.

This hybrid orientation strategy enables representation of both hierarchical conceptual expansion and sequential semantic dependencies, thereby encoding asymmetric relationships between terms.

2.2.3. Stage 3: Edge weight computation and aggregation

Edge weights were computed by merging identical term nodes across all sentences and documents. The weight of a directed edge corresponds to the number of times the same ordered term pair appeared throughout the corpus.

Formally, the weight w_{ij} between nodes i and j is defined as:

$$w_{ij} = \sum_{s=1}^{S} I_{ij}^{(s)}, \tag{3}$$

where $I_{ij}^{(s)}$ is an indicator function equal to 1 if a directed edge from i to j exists in sentence s, and 0 otherwise.

This aggregation process transforms sentence-level graphs into a global Directed Weighted Network of Terms, capturing statistically significant semantic relationships across the entire corpus.

To eliminate self-loops introduced during HVG aggregation, diagonal elements of the adjacency matrix were set to zero. The resulting DWNT is thus represented by a square weighted adjacency matrix without self-connections, enabling subsequent computation of network metrics such as node degree, weighted centrality, and clustering coefficients.

The proposed DWNT construction procedure integrates linguistic structure, term importance, and contextual ordering into a unified graph model. Unlike traditional co-occurrence networks, this approach simultaneously accounts for syntactic hierarchy, semantic prominence, and directional dependencies, making it particularly suitable for longitudinal analysis of scientific terminology.

3. Results

For each year in the period 2015–2024, Global Term Frequency (GTF) values were computed for all extracted terms. Based on these values, the top-10 terms with the highest GTF scores were identified annually (Table 1), representing the most concepts characterizing the scientific discourse of each respective year.

This ranking procedure provides a quantitative snapshot of terminological prominence and allows direct comparison of dominant concepts across different temporal intervals. The resulting top-10 lists reveal both persistent core terminology and rapidly emerging concepts, reflecting shifts in research focus and methodological priorities within the AI community.

Table 1
the top-10 terms with the highest GTF scores

Terms	2015	2016	2017	2018	...	2021	2022	2023	2024
model	0,012	0,0139	0,0122	0,0127	...	0,0138	0,0137	0,0141	0,0153
learning	0,0056	0,007	0,009	0,0107	...	0,0114	0,0115	0,0114	0,0114
problem	0,0165	0,0127	0,0126	0,0123	...	0,0118	0,0114	0,0112	0,0106
method	0,01	0,0088	0,0083	0,0086	...	0,0083	0,0089	0,0092	0,0094
datum	0,0112	0,0109	0,0094	0,0086	...	0,0084	0,0088	0,0093	0,0091
approach	0,0087	0,0097	0,0087	0,0095	...	0,0099	0,0095	0,0093	0,009
algorithm	0,0115	0,0115	0,0122	0,0117	...	0,0104	0,0098	0,0093	0,0087
knowledge	0,0068	0,0073	0,0072	0,0067	...	0,0069	0,0068	0,0069	0,007
agent	0,0021	0,0034	0,0058	0,006	...	0,0065	0,0063	0,0066	0,0069
task	0,0016	0,0026	0,0032	0,0041	...	0,0049	0,0052	0,0058	0,0058

To further investigate temporal dynamics, yearly DWNT graphs were compared to detect terms that appeared for the first time in a given year as well as terms that persisted or disappeared in subsequent periods (Table 2). Newly introduced terms were interpreted as indicators of emerging research directions, whereas declining or vanishing terms often corresponded to concepts whose relevance diminished over time.

Table 2
GTF dynamics of *artificial intelligence*, *llm*, and *neural network*

Terms	2015	2016	2017	2018	...	2021	2022	2023	2024
artificial intelligence	0,0015	0,0015	0,0018	0,0016	...	0,002	0,0021	0,0021	0,002
LLM	0	0	0	0	...	0	0,0002	0,0002	0,0019
neural network	0,0003	0,0005	0,001	0,0009	...	0,0016	0,002	0,0018	0,0018

This longitudinal comparison enabled identification of macro-level trends in the AI terminological space. In particular, stable terms such as *model* and *learning* formed the semantic backbone of the networks across the entire observation period, while terms related to generative architectures, including *large language model* exhibited sharp increases after 2020, signaling a paradigm shift toward generative artificial intelligence.

The Table 3 presents aggregated topological characteristics of the DWNT constructed for each year from 2015 to 2024, including the number of nodes, number of edges, average degree, average clustering coefficient, modularity, and number of communities. These indicators enable a structural –dynamic analysis of the evolution of the AI terminological space over the last decade.

Table 3
Aggregated topological characteristics of the DWNT, 2015–2024

Characteristic	2015	2016	2017	2018	...	2021	2022	2023	2024
Number of nodes	2462	4334	5918	7606	...	12095	13742	15333	17095
Number of edges	8411	16241	23770	32078	...	55424	64696	73330	82681
Average degree	3,416	3,747	4,017	4,217	...	4,582	4,708	4,782	4,837
Average clustering coefficient	0,301	0,314	0,324	0,336	...	0,362	0,37	0,376	0,381
Modularity	0,522	0,469	0,435	0,414	...	0,383	0,374	0,367	0,365
Number of communities	23	27	28	29	...	32	28	31	30

Taken together, these results demonstrate that DWNT-based analysis captures not only changes in term frequency but also deeper structural transformations in semantic relationships. The combined evaluation of GTF rankings, term emergence, and network topology provides a comprehensive perspective on the evolution of scientific terminology, enabling identification of both incremental developments and disruptive conceptual shifts.

4. Analysis of results

Within this study, time series of importance values were analyzed for ten key terms over the period 2015–2024 (Table 1).

4.1. Analysis of AI terminological space evolution

The term *model* exhibits a clear and steady increase in importance, particularly after 2020 (Figure 1). This trend can be associated with the rapid expansion of large-scale architectures, including Transformers, GPT-based systems, and diffusion models. The observed dynamics indicate an evolution of the concept of "model" from classical machine learning paradigms toward large language models and foundation models. Notably, *model* represents the most stable term with a consistent upward trajectory, while its post-2020 amplification correlates with the transition from traditional ML approaches to model-centric and foundation-model frameworks.

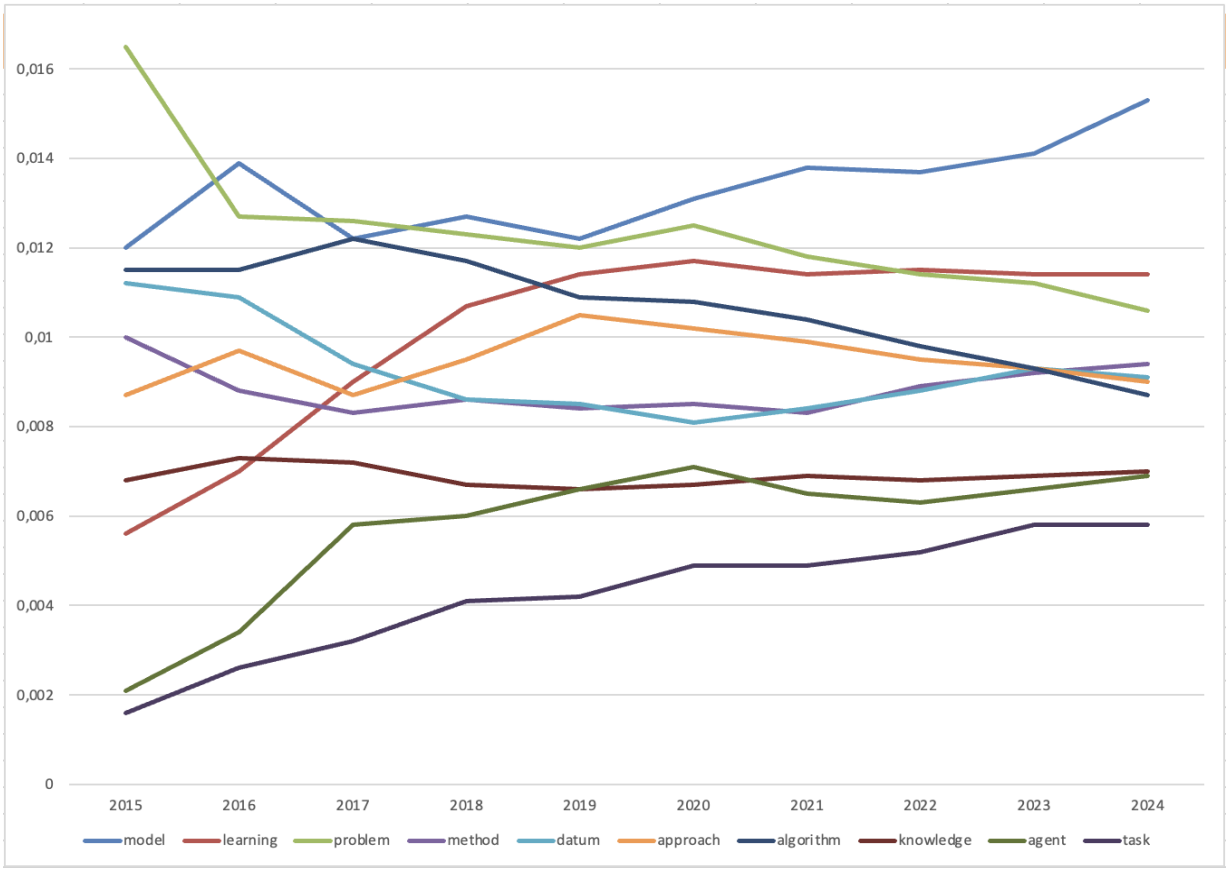


Figure 1: Evolution of ten key terms in Artificial Intelligence research.

The term *learning* demonstrates rapid growth until 2019, followed by stabilization after 2020. During 2015–2019 its importance nearly doubled, reflecting the "golden age" of deep learning. Subsequently, a plateau is observed, suggesting thematic saturation and increasing specialization toward subfields such as deep learning, representation learning, and active learning.

Conversely, the term *problem* shows a persistent declining trend, with decreasing importance almost every year. This pattern likely reflects a shift in AI publications away from problem formulation toward methodological innovation, model development, and data-driven pipelines. The decreasing prominence of this less technical and less specific term highlights an evolution in scientific writing style within AI research.

The term *method* exhibits oscillatory dynamics with renewed growth after 2020. Although characterized by instability, a moderate upward trend is evident during 2021–2024, potentially driven by emerging methodologies in reinforcement learning, explainable AI, and multimodal learning.

An especially interesting pattern is observed for the term *datum*, which displays a U-shaped trajectory: declining during 2015–2020 and increasing during 2021–2023. The early decrease corresponds to the dominance of model-centric AI, whereas the subsequent recovery reflects the paradigm shift toward data-centric AI.

In addition, the temporal evolution of three foundational terms – *artificial intelligence*, *neural network*, and *llm* (large language models) – was analyzed over 2015–2024 (Table 2, Figure 2). For each term, statistical indicators including linear trend, value range (max–min), and volatility (standard deviation) were computed to characterize their semantic dynamics.

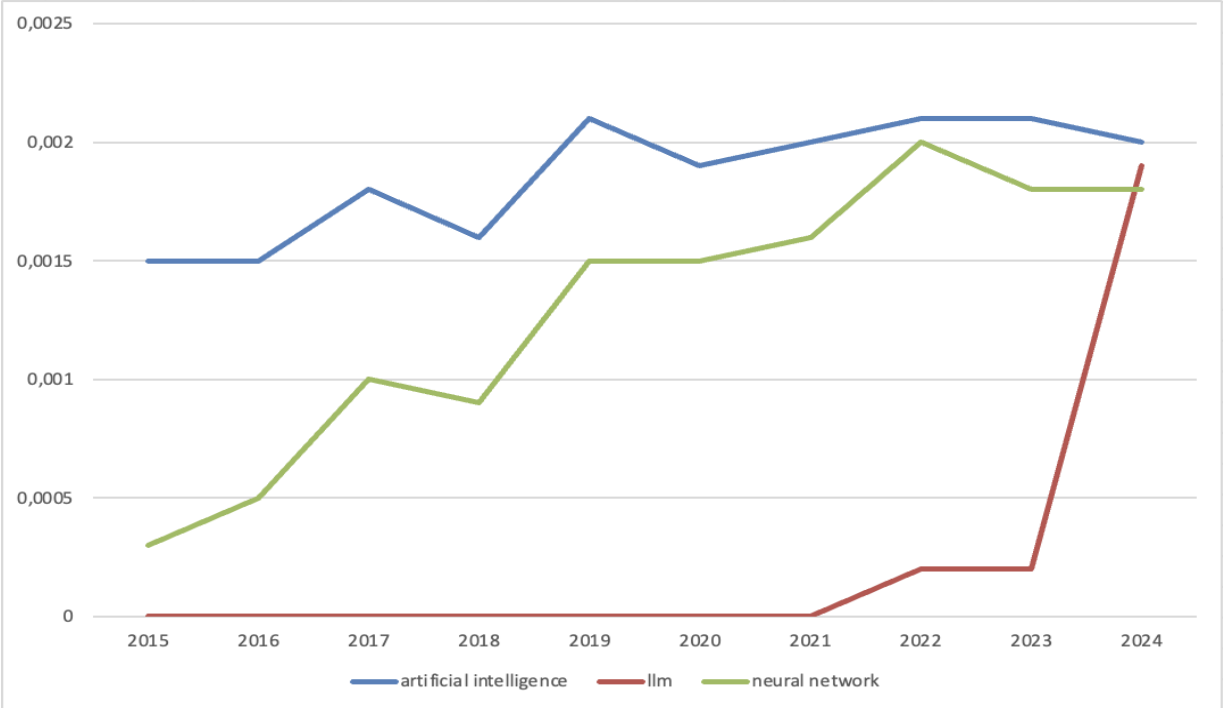


Figure 2: Evolution of the terms *artificial intelligence*, *llm*, and *neural network*.

Both *artificial intelligence* and *neural network* demonstrate long-term presence in the publication corpus and relatively low volatility. Specifically, *artificial intelligence* shows moderate growth (trend +0,000068) with a small fluctuation range (0,0006), confirming its role as a stable umbrella concept maintaining consistent semantic load throughout the study period.

In contrast, *neural network* exhibits the strongest positive trend (+0,00018), with gradual growth from 2015 and a peak in 2022. This trajectory reflects the institutionalization of deep learning and the transition of neural networks from experimental tools to the foundational technological backbone of modern AI systems. Together, these terms form the conceptual core of the AI terminological space, ensuring its structural continuity.

Unlike these baseline concepts, *llm* displays an abrupt change in dynamics after 2022. Prior to this point its importance was nearly negligible; however, during 2023–2024 a sharp increase is

observed, accompanied by the largest value range (0,0019) and highest volatility (0,00059) among all analyzed terms. This behavior corresponds to the classical emergence pattern of an innovation-driven term and aligns with the widespread adoption of generative language models and general-purpose foundation models.

The rapid rise of large language models reflects a fundamental shift in research priorities and the formation of a new paradigm in artificial intelligence. The obtained results indicate a substantial transformation of scientific discourse after 2022 – from classical neural architectures toward generative language systems. These findings further validate the suitability of network-based approaches, particularly Directed Weighted Networks of Terms, for capturing not only frequency variations but also structural reorganization within the AI terminological space.

4.2. Growth of the terminological space

The most pronounced positive dynamics are observed for the number of nodes and number of edges.

The number of nodes increases from 2,462 in 2015 to 17,095 in 2024, representing more than a sevenfold growth. Similarly, the number of edges rises from 8,411 to 82,681, corresponding to nearly a tenfold increase.

This exponential expansion indicates:

- rapid growth of the AI terminological vocabulary;
- continuous emergence of new concepts;
- increasing complexity of semantic relationships among terms.

Such dynamics directly reflect the accelerated development of artificial intelligence research, particularly after 2018, coinciding with the widespread adoption of transformer architectures, large language models, and generative approaches. A marked acceleration after 2019 can be interpreted as a paradigm shift toward foundation models

4.3. Network densification and increasing local coherence

The average degree steadily increases from 3,42 to 4,84, implying that each term becomes connected to a growing number of related concepts. This trend indicates progressive integration of the semantic space.

In parallel, the average clustering coefficient rises from 0,301 to 0,381, demonstrating strengthening local connectivity. This reflects the formation of dense thematic clusters, increased triadic closure, and tighter conceptual grouping of terms.

Together, these trends suggest a transition from a fragmented terminological landscape toward a more structured knowledge system, where newly introduced concepts are rapidly embedded into existing semantic frameworks.

4.4. Modularity dynamics

In contrast to the aforementioned indicators, modularity exhibits a consistent decline, decreasing from 0,522 in 2015 to 0,365 in 2024.

This trend is particularly significant. High modularity in early years indicates the presence of relatively isolated thematic subfields. Over time, the decreasing modularity reflects increasing interconnection among these domains, signaling interdisciplinary convergence.

The reduction in modularity suggests that previously independent areas such as computer vision, natural language processing, and reinforcement learning are gradually merging into a unified conceptual ecosystem – a characteristic feature of a maturing scientific paradigm.

4.5. Evolution of community structure

The number of communities increases moderately from 23 to approximately 30–32, without a

strictly monotonic pattern. This behavior indicates simultaneous processes of differentiation and integration: new thematic directions emerge while existing communities partially merge.

Such dynamics highlight continuous structural reorganization of the terminological network, where diversification and consolidation occur concurrently.

4.6. Interpretation from the perspective of knowledge evolution

Overall, three major phases can be distinguished:

1. 2015–2017 – Structural formation phase: rapid growth in nodes and edges combined with high modularity.
2. 2018–2020 – Expansion phase: sharp increase in network scale accompanied by the onset of inter-domain integration.
3. 2021–2024 – Consolidation phase: densification of connections, declining modularity, and stabilization of community numbers.

These findings demonstrate that the AI terminological space evolves from a loosely connected collection of topics into a densely integrated conceptual framework.

The DWNT-based analysis reveals that advances in artificial intelligence are accompanied not only by an increasing number of terms but also by profound restructuring of their semantic relationships. Rising network density and decreasing modularity indicate a transition toward an integrated scientific paradigm, in which emerging concepts are rapidly incorporated into an established knowledge structure.

5. Conclusion

This study proposed and empirically validated a methodology for constructing and analyzing Directed Weighted Networks of Terms to investigate the evolution of scientific terminology in artificial intelligence over the period 2015–2024. The approach integrates linguistic preprocessing, syntactic key-term extraction, Global Term Frequency estimation, and network-based modeling to capture both frequency-driven trends and structural transformations of semantic relationships.

A longitudinal corpus of AI-related publications was assembled from arXiv and organized by publication year, providing a representative foundation for analyzing long-term terminological dynamics. For each year, DWNT graphs were constructed, enabling quantitative assessment of inter-term connectivity, node centrality, and network cohesion. This framework facilitated systematic identification of dominant concepts, emerging terminology, and shifts in semantic structure.

The results reveal several key patterns. First, foundational terms such as *artificial intelligence* and *neural network* exhibit long-term stability and low volatility, forming a persistent conceptual backbone of the field. Second, core methodological concepts including *model* and *learning* demonstrate sustained growth and consolidation, reflecting the transition toward large-scale and foundation-model paradigms. Third, innovation-driven terms, most notably *llm*, display abrupt post-2022 emergence characterized by high volatility and rapid integration into the network core, indicating a paradigm shift toward generative AI and large language models.

Beyond frequency-based observations, temporal DWNT analysis exposed deeper structural changes: increasing network density, declining modularity, and rising node centralization over time. These trends signify progressive conceptual integration, moving from fragmented subfields in early years toward a more unified semantic space dominated by deep learning and generative modeling frameworks. The rapid transition of newly introduced terms from peripheral positions to highly connected hubs further highlights accelerated knowledge diffusion in contemporary AI research.

Taken together, these findings demonstrate that DWNT-based analysis effectively captures both incremental thematic development and disruptive conceptual transitions. The combined evaluation

of GTF rankings, term emergence, centrality dynamics, and global network topology provides a comprehensive perspective on the evolution of scientific discourse.

The proposed methodology thus offers a robust and scalable framework for monitoring disciplinary evolution, detecting paradigm shifts, and identifying early signals of emerging research directions. Its applicability extends beyond artificial intelligence and may be generalized to other rapidly developing scientific domains, supporting strategic research planning, trend forecasting, and large-scale semantic analysis of scholarly knowledge.

Consequently, Directed Weighted Networks of Terms provide a powerful quantitative instrument for capturing macroscopic changes in scientific discourse and for analyzing long-term terminological evolution.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Y. Bengio, Y. LeCun, G. Hinton, Deep learning, Nature 521 (2015) 436–444. doi:10.1038/nature14539.

[2] A. Vaswani, N. Shazeer, N. Parmar, et al., Attention is all you need, Advances in Neural Information Processing Systems 30 (2017). doi:10.48550/arXiv.1706.03762.

[3] M. I. Jordan, T. M. Mitchell, Machine learning: Trends, perspectives, and prospects, Science 349 (6245) (2015) 255–260. doi:10.1126/science.aaa8415.

[4] T. B. Brown, B. Mann, N. Ryder, et al., Language models are few-shot learners, Advances in Neural Information Processing Systems 33 (2020). doi:10.48550/arXiv.2005.14165.

[5] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent Dirichlet Allocation, Journal of Machine Learning Research 3 (2003) 993–1022. doi:10.5555/944919.944937.

[6] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, Advances in Neural Information Processing Systems 26 (2013). doi:10.48550/arXiv.1310.4546.

[7] M. E. J. Newman, Networks: An Introduction, Oxford University Press, Oxford, 2010. doi:10.1093/acprof:oso/9780199206650.001.0001.

[8] N. Tahmasebi, L. Borin, A. Jatowt, Survey of computational approaches to lexical semantic change, Journal of Artificial Intelligence Research 63 (2018) 1–46. doi:10.1613/jair.1.11261.

[9] W. L. Hamilton, J. Leskovec, D. Jurafsky, Diachronic word embeddings reveal statistical laws of semantic change, Transactions of the Association for Computational Linguistics 4 (2016) 31–45. doi:10.1162/tacl_a_00099.

[10] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT), Association for Computational Linguistics, 2019. doi:10.18653/v1/N19-1423.

[11] J. Pennington, R. Socher, C. D. Manning, GloVe: Global vectors for word representation, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, 2014. doi:10.3115/v1/D14-1162.

[12] J. Grimmer, B. M. Stewart, Text as data: The promise and pitfalls of automatic content analysis methods for political texts, American Journal of Political Science 57 (2) (2013) 267–281. doi:10.1111/ajps.12026.

[13] M. E. J. Newman, The structure and function of complex networks, SIAM Review 45 (2) (2003) 167–256. doi:10.1137/S003614450342480.

[14] R. Alghamdi, K. Alfalqi, A survey of topic modeling in text mining, International Journal of Advanced Computer Science and Applications 6 (2015). doi:10.14569/IJACSA.2015.060121.
[15] S. Lozano, L. Calzada-Infante, B. Adenso-Díaz, S. García, Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature, Scientometrics 120 (2) (2019) 609–629. doi:10.1007/s11192-019-03132-w.
[16] M. Garg, P. Kumar, A. Gupta, Identifying influential segments from word co-occurrence networks, Computer Speech & Language 44 (2017) 28–45. doi:10.1016/j.csl.2017.01.003.
[17] E. Segev (Ed.), Semantic Network Analysis in Social Sciences, Routledge, London, 2020. doi:10.4324/9780429469711.
[18] J. Diesner, From texts to networks: Detecting and analyzing socio-semantic relationships, in: Proceedings of the IEEE International Conference on Social Computing, IEEE, 2012, pp. 1–8. doi:10.1109/SocialCom.2012.87.
[19] A. Gerlach, F. Font-Clos, E. G. Altmann, Similarity of symbol frequency distributions with heavy tails, Physical Review X 6 (2) (2016) 021009. doi:10.1103/PhysRevX.6.021009.
[20] X. Wang, A. McCallum, Topics over time: A non-Markov continuous-time model of topical trends, in: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2006, pp. 424–433. doi:10.1145/1150402.1150429.
[21] H. Dubossarsky, D. Weinshall, E. Grossman, Outta control: Laws of semantic change and inherent biases in word representation models, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, 2017, pp. 1136–1145. doi:10.18653/v1/D17-1127.
[22] D. M. Blei, J. D. Lafferty, Dynamic topic models, in: Proceedings of the 23rd International Conference on Machine Learning (ICML), ACM, 2006, pp. 113–120. doi:10.1145/1143844.1143859.
[23] X. Wang, A. McCallum, X. Wei, Topical n-grams: Phrase and topic discovery, with an application to information retrieval, in: Proceedings of the IEEE International Conference on Data Mining (ICDM), IEEE, 2007, pp. 697–702. doi:10.1109/ICDM.2007.85.
[24] D. Lande, O. Dmytrenko, Creating directed weighted network of terms based on analysis of text corpora, in: Proceedings of the 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing, SAIC 2020, IEEE, 2020, pp. 123–130. doi:10.1109/SAIC51560.2020.00025
[25] D. V. Lande, O. O. Dmytrenko, Methodology for extracting of key words and phrases and building directed weighted networks of terms with using part-of-speech tagging, in: Proceedings of the International Conference on Intelligent Technologies and Systems (ITS 2020), pp. 168–177, 2020.
[26] D. V. Lande, O. O. Dmytrenko, Using part-of-speech tagging for building networks of terms in legal sphere, in: Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems, COLINS 2021, Kharkiv, Ukraine, April 22–23, 2021, CEUR Workshop Proceedings, Vol. 2870, CEUR-WS.org, 2021, pp. 87–97, ISSN 1613-0073.
[27] R. He, J. Cao, T. Tan, Generative artificial intelligence: a historical perspective, National Science Review 12 (2025) 1–15. doi:10.1093/nsr/nwaf050.
[28] M.Á. Rodríguez Ortiz, P.C. Santana Mancilla, L.E. Anido Rifón, Machine learning and generative AI in learning analytics for higher education: A systematic review of models, trends, and challenges, Applied Sciences 15 (2025) 8679. doi:10.3390/app15158679.
[29] Sumit Mishra, arXiv Scientific Research Papers Dataset: A Curated Dataset of arXiv Papers with Titles, Summaries, Categories & Metadata. 2025. URL: https://www.kaggle.com/datasets/sumitm004/arxiv-scientific-research-papers-dataset.
[30] B. Luque, L. Lacasa, F. Ballesteros, A. Luque, Horizontal visibility graphs: Exact results for random time series, Physical Review E 80 (2009) 046103. doi:10.1103/PhysRevE.80.046103.