Submission ID: 23164

Attention-Based Concept Mining and Community Detection in AI Research

AOM Division: Research Methods

ABSTRACT

Artificial Intelligence (AI) research has grown at a fast pace in recent years, introducing a variety of new methods (e.g., Large Language Models), specialized datasets (e.g., ImageNet), and emergent tasks (e.g., multi-modal text generation). Traditional bibliometric analyses shed light on publication volumes and co-authorship patterns but often fail to capture the specific concepts that underlie these collaborations. To address this gap, we propose a two-tiered approach that merges attention-based concept mining with co-authorship network analysis. First, we integrate frequency-based and dictionary-based techniques, guided by anchor-phrases, to train a neural network that identifies not only high-frequency AI terms but also rare, high-impact concepts. By leveraging attention maps from a pre-trained transformer, we reduce noise—distinguishing genuine new ideas like "Vision Transformer" from random token clusters. Second, we situate these concepts in temporal co-authorship networks, applying community detection to reveal where innovations emerge, how they travel across research subfields, and what collaboration patterns accelerate or inhibit their diffusion. This synthesis provides a robust framework for research methods, enabling richer inquiries into how AI knowledge diffuses and reshapes scientific collaboration.

Attention-Based Concept Mining and Community Detection in AI Research

INTRODUCTION

The field of Artificial Intelligence (AI) has experienced unprecedented growth in recent years, introducing advanced methods (Convolutional Neural Networks, Large Language Models, etc.), specialized datasets (CIFAR-10, ImageNet, etc.) and emerging tasks (image classification, text generation, etc.). This expansion has transcended traditional computer science boundaries, with numerous scientific disciplines integrating AI methodologies to enhance research outcomes and foster innovation (Hajkowicz et al., 2023). The magnitude of this progression is evident in the substantial increase in AI-related publications and collaborative authorships over the past decade. However, traditional bibliometric approaches track citations or co-authorship networks but often overlook the specific concepts driving these collaborations.

To address this gap, we propose an attention-based concept mining pipeline fueled by well-known anchor-phrases. This requires a careful composition of AI-related publications as well as a range of well-known methods, tasks, datasets and descriptions for training purposes. It will result in a neural-network driven framework for mining phrases and identifying valid AI concepts from large text corpora.

Building upon the extracted concepts, we will construct temporal co-authorship networks to analyze the evolution of collaborations and the dissemination of innovations within the AI research community. This approach enables us to identify how emerging concepts influence the formation of new collaborations and the restructuring of existing research communities. By employing advanced community detection algorithms, such as the Louvain method (Blondel et al., 2008), we can uncover the modular structure of these networks, revealing clusters of researchers who collectively contribute to specific AI advancements.

The dynamic nature of these networks allows us to observe concept drift—the evolution of research themes and methodologies over time—and its impact on collaborative patterns. Analyzing these shifts provides insights into how novel innovations catalyze new partnerships and interdisciplinary ventures, thereby shaping the trajectory of AI research. Furthermore, understanding the structure and dynamics of co-authorship networks offers a quantitative framework to assess the diffusion of knowledge and the emergence of influential research communities (Barabasi et al., 2002).

In summary, mapping temporal co-authorship networks and tracking concept evolution through community detection methodologies provide a robust framework for understanding the interplay between innovation and collaboration in AI research. This analysis not only clarifies the pathways through which new ideas propagate but also highlights the pivotal role of collaborative networks in advancing the field.

CONCEPT MINING FROM LARGE TEXT CORPORA

A concept represents an abstract unit of knowledge defined by shared properties or contextual significance (Stock, 2010), such as 'transformer-based reinforcement learning' in AI. Concepts are identified through their intension (e.g., tasks, methods, or datasets associated with the term) and extension (instances of papers or authors using the term).

To gather AI-related concepts, we construct a database containing metadata and full-texts from PapersWithCode[1](PwC), a repository of around 500k AI papers tagged with tasks, methods, datasets, conferences, and code repositories. These tags stem from advanced extraction algorithms[2] and community contributions. Data from OpenAlex (Priem et al., 2022) and SemanticScholar (Kinney et al., 2023) further enriches this database with disambiguated authors and fulltexts, creating an extensive knowledge graph on AI research (see Figure 1).

Insert Figure 1 about here

Traditional concept-extraction methods, such as n-gram frequency thresholds, Part-of-Speech tags, or dictionary lookups, often fail to capture impactful low-frequency terms or avoid false positives like short tokens or acronyms. Recent deep learning-based strategies (Shang

[1]paperswithcode.com
[2]github.com/paperswithcode/sota-extractor

et al., 2017; Meng et al., 2021; Choi et al., 2023) address these issues. For example, UCPhrase (Shen et al., 2022) leverages attention maps of sentence-level embeddings to identify relevant phrases. Inspired by this, we adopt a similar approach with anchor-phrases serving as a baseline dictionary. Attention maps from pre-trained transformers add contextual understanding, highlighting meaningful connections among tokens while filtering noise. For instance, in identifying "GAN," attention patterns often focus on nearby terms like "generator" or "adversarial training," validating the term as a coherent concept. Our framework combines three signals to build the training set:

1. **Dictionary-based**: Known tasks, methods, datasets, and phrases from sources like PapersWithCode, the Computer Science Ontology (Salatino et al., 2020), or textbook indices (Goodfellow et al., 2016; Murphy, 2022; Prince, 2023).
2. **Frequency-based**: Repeated n-grams surpassing specific thresholds while avoiding trivial phrases.
3. **Attention-based**: Subword alignment filters concept candidates through their attention patterns, skipping short or numeric tokens.

This approach is especially powerful for spotting emergent phrases—like "Vision Transformer" or "causal NLP" that might be infrequent but still recognized by learned attention signals. Likewise, the dictionary helps identify known concepts even when they appear only once or twice. By cross-checking these hits with the model's attention, we discard ambiguous or uninformative tokens (e.g., "0," "u").

AI DIFFUSION AND COLLABORATION

While the previous section focused on *how* we mine AI concepts, an equally critical question is **where** these concepts emerge and spread. To address this, we construct temporal co-authorship networks, linking authors who publish AI-related papers each year. Figure 2 shows snapshots of these networks, where nodes represent authors, and edges signify co-authorships. By overlaying the concept space mined in the last Section, we could highlight which authors dominate specific tasks, methods, or datasets (see Figure 3). This integration of concept mining and network analysis maps knowledge diffusion, linking concept adoption—like the use of "Vision Transformer"—to patterns of collaboration. Unlike citation analyses, this approach connects idea propagation to shared projects, offering a richer understanding of how AI innovations spread (Barabasi et al., 2002).

Insert Figure 2 about here

Insert Figure 3 about here

To analyze AI diffusion, we examine yearly co-authorship networks through metrics highlighting individual and structural growth (Figure 4). Rising author counts (a) and co-author edges (b) reflect the rapid expansion of AI research and collaboration. Metrics like network density (d) and clustering coefficients (Watts and Strogatz, 1998) (f) reveal how research communities form: higher clustering often signals localized collaborations, while an expanding largest connected component shows increasing global integration and idea exchange.

Metrics such as the degree of centrality (c) track average collaborative intensity per author, while the size of the largest connected component (e) indicates dominant collaborative hubs. Widely-adopted concepts, like "transformers," often begin in smaller clusters before reshaping broader co-authorship networks. By linking these network dynamics to concept adoption identified through attention-based mining, we uncover how innovations diffuse through collaborative bridges, emphasizing the interplay of tightly-knit clusters and expansive networks in shaping AI research.

Insert Figure 4 about here

CONCLUSION

This work introduces a dual-layered framework combining attention-based concept mining with co-authorship network analysis to capture how AI innovations emerge, diffuse, and reshape collaborative structures in scientific communities. By leveraging frequency-driven and dictionary-based anchor-phrases, our approach uses attention maps from transformer models to refine concept identification, even for low-frequency but influential terms. These extracted concepts serve as pivots for analyzing temporal co-authorship networks, revealing how new AI ideas flow within and across disciplinary clusters.

In research methods, the approach demonstrates how attention-based concept mining goes beyond standard bibliometrics by focusing on concrete knowledge elements rather than mere citation links. Researchers can adopt this pipeline to examine how particular tasks (e.g., “few-shot reinforcement learning”) arise in close-knit subfields, travel across communities, or combine with other methods. Because the pipeline highlights both the local (clustering) and global (connected-component) structure of co-authorship, it offers an empirical basis for understanding where fresh ideas are incubated and how they gain traction in wider domains.

As AI continues to permeate disparate fields, the need to pinpoint where breakthroughs occur—and how they diffuse across the broader ecosystem—grows ever more pressing. By combining concept mining with network analysis, our framework delivers a granular view of the real-time evolution of AI research. Academics and industry practitioners alike can leverage these insights to foster targeted collaborations, identify underexplored avenues of innovation, and shape research agendas that respond effectively to emerging trends.

REFERENCES

Barabasi, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., and Vicsek, T. 2002. Evolution of the social network of scientific collaborations. ***Physica A: Statistical Mechanics and its Applications***, 311(3-4): 590–614. arXiv:cond-mat/0104162.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. 2008. Fast unfolding of communities in large networks. ***Journal of Statistical Mechanics: Theory and Experiment***, 2008(10): P10008. arXiv:0803.0476 [physics].

Choi, M., Gwak, C., Kim, S., Kim, S., and Choo, J. 2023. SimCKP: Simple Contrastive Learning of Keyphrase Representations. In Bouamor, H., Pino, J., and Bali, K.(Eds.), ***Findings of the Association for Computational Linguistics: EMNLP 2023***, 3003–3015, Singapore. Association for Computational Linguistics.

Goodfellow, I., Bengio, Y., and Courville, A. 2016. ***Deep Learning***. MIT Press. Google-Books-ID: omivDQAAQBAJ.

Hajkowicz, S., Sanderson, C., Karimi, S., Bratanova, A., and Naughtin, C. 2023. The Diffusion of Artificial Intelligence Technology Across Research Fields: A Bibliometric Analysis of Scholarly Publications from 1960-2021.

Kinney, R., Anastasiades, C., Authur, R., Beltagy, I., Bragg, J., Buraczynski, A., Cachola, I., Candra, S., Chandrasekhar, Y., Cohan, A., Crawford, M., Downey, D., Dunkelberger, J., Etzioni, O., Evans, R., Feldman, S., Gorney, J., Graham, D., Hu, F., Huff, R., King, D., Kohlmeier, S., Kuehl, B., Langan, M., Lin, D., Liu, H., Lo, K., Lochner, J., MacMillan, K., Murray, T., Newell, C., Rao, S., Rohatgi, S., Sayre, P., Shen, Z., Singh, A., Soldaini, L., Subramanian, S., Tanaka, A., Wade, A. D., Wagner, L., Wang, L. L., Wilhelm, C., Wu, C., Yang, J., Zamarron, A., Van Zuylen, M., and Weld, D. S. 2023. The Semantic Scholar Open Data Platform. arXiv:2301.10140 [cs].

Meng, R., Zhao, S., Han, S., He, D., Brusilovsky, P., and Chi, Y. 2021. Deep Keyphrase Generation. arXiv:1704.06879 [cs].

Murphy, K. P. 2022. ***Probabilistic machine learning: An introduction***. MIT Press.

Priem, J., Piwowar, H., and Orr, R. 2022. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. arXiv:2205.01833 [cs].

Prince, S. J. 2023. ***Understanding deep learning***. MIT Press.

Salatino, A. A., Thanapalasingam, T., Mannocci, A., Birukou, A., Osborne, F., and Motta, E. 2020. The Computer Science Ontology: A Comprehensive Automatically-Generated Taxonomy of Research Areas. ***Data Intelligence***, 2(3): 379–416.

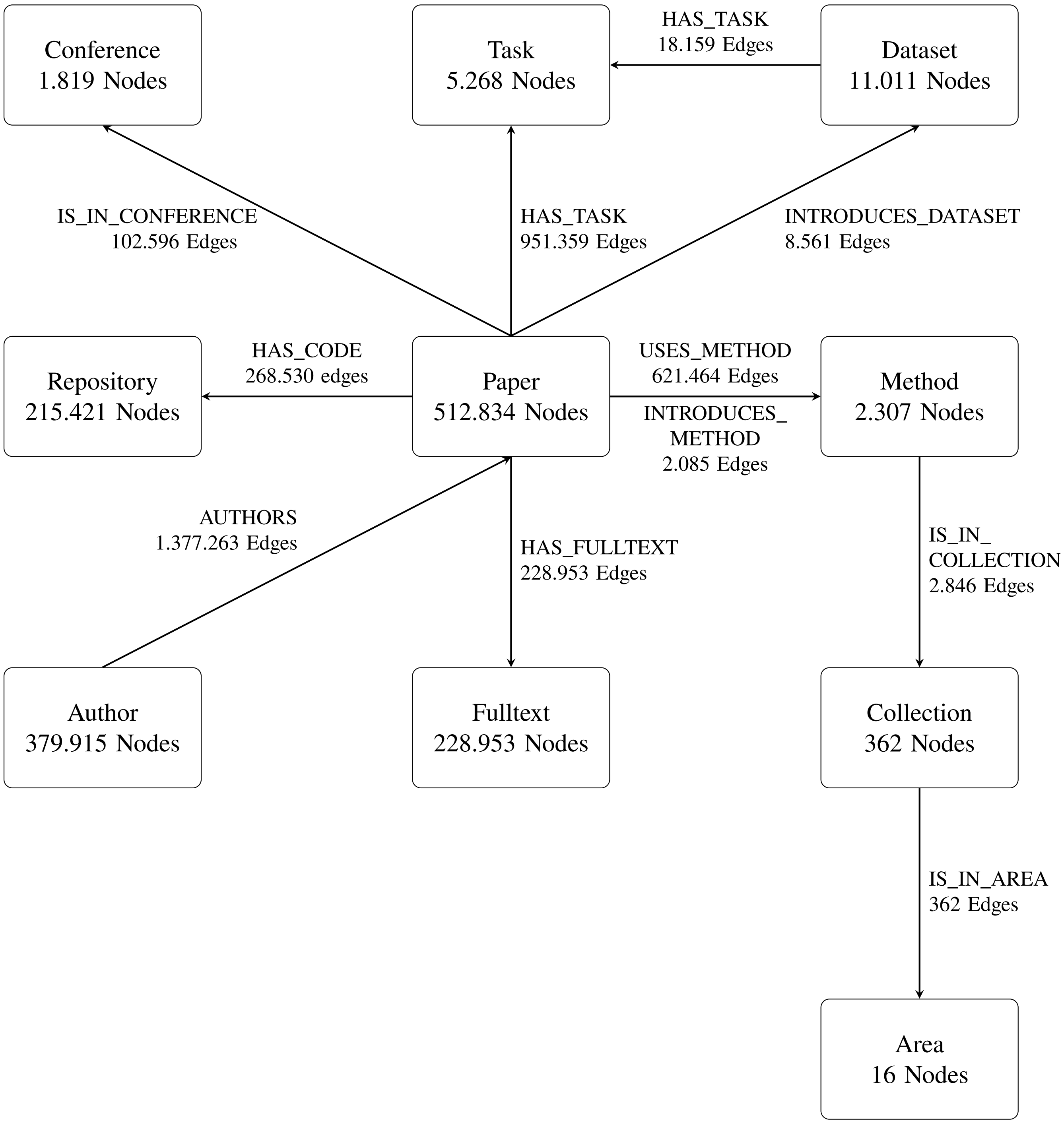
Shang, J., Liu, J., Jiang, M., Ren, X., Voss, C. R., and Han, J. 2017. Automated Phrase Mining from Massive Text Corpora. arXiv:1702.04457 [cs].

Shen, X., Wang, Y., Meng, R., and Shang, J. 2022. Unsupervised Deep Keyphrase Generation. ***Proceedings of the AAAI Conference on Artificial Intelligence***, 36(10): 11303–11311.

Stock, W. G. 2010. Concepts and semantic relations in information science. ***Journal of the American Society for Information Science and Technology***, 61(10): 1951–1969.

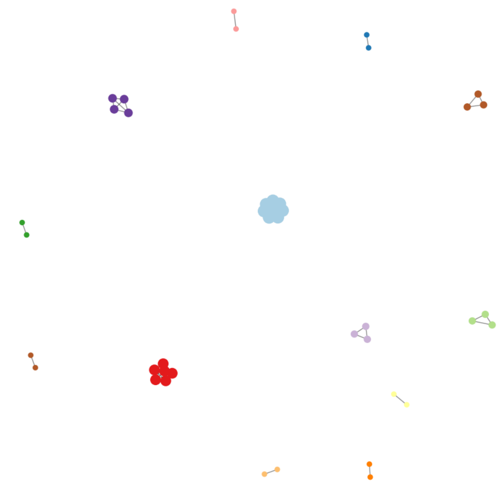
Watts, D. J. and Strogatz, S. H. 1998. Collective dynamics of 'small-world' networks. ***Nature***, 393(6684): 440–442. Publisher: Nature Publishing Group.

FIGURE 1
AI Papers Knowledge Graph

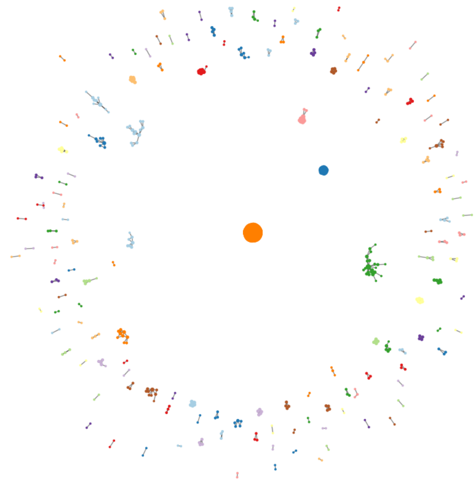


Notes: A knowledge graph database created for this publication.

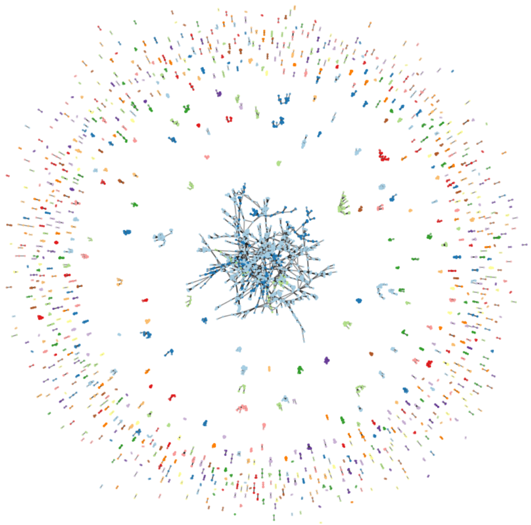
FIGURE 2
Sociograms of Co-Authorships in AI-related publications over time



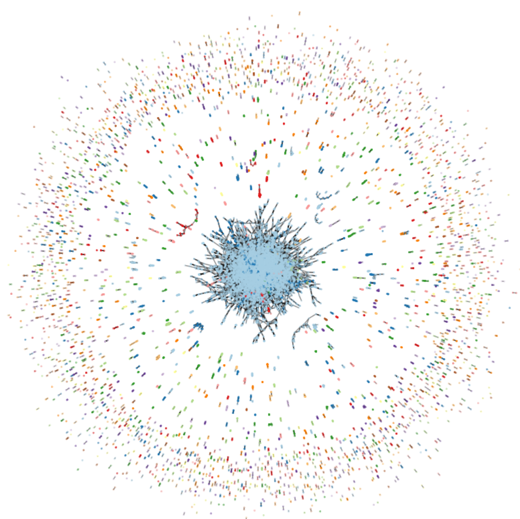
Sociogram in 2004



Sociogram in 2008



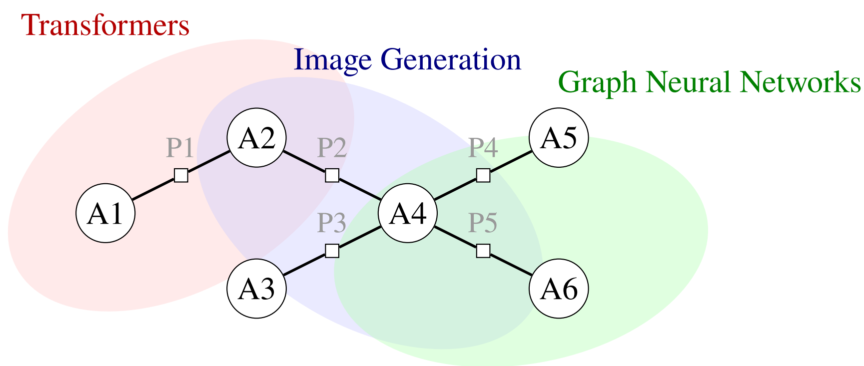
Sociogram in 2012



Sociogram in 2016

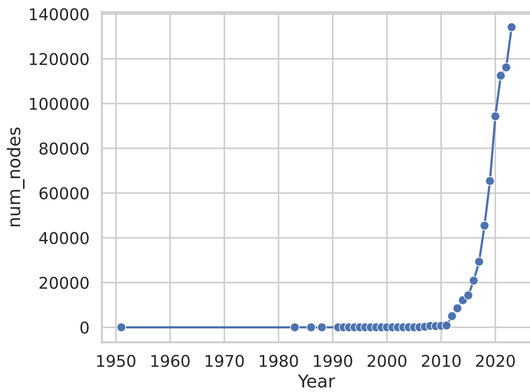
Notes: This figure illustrates the evolution of collaboration between authors in AI-related publications. With an increasing number of papers, authors tend to work closer together.

FIGURE 3
Co-Authorship Network with Concept-Clusters

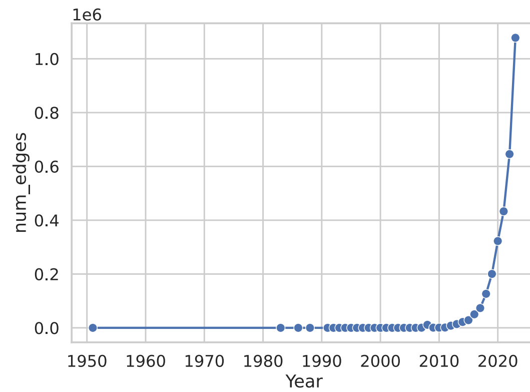


Notes: Authors (*Ax* Nodes) are connected via Co-Authorships through Papers (*Px* Edges) which in turn are labeled with concepts (Background-Ellipses).

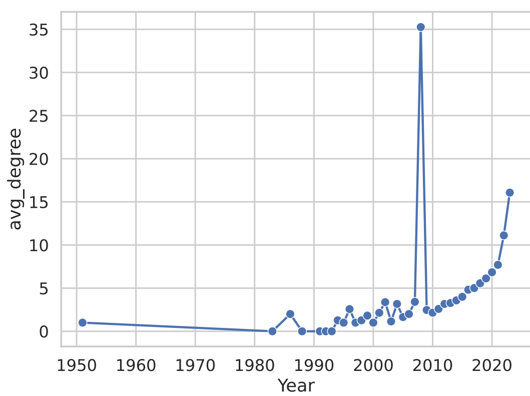
FIGURE 4
Evolution of AI Co-Authorship Network Over Time



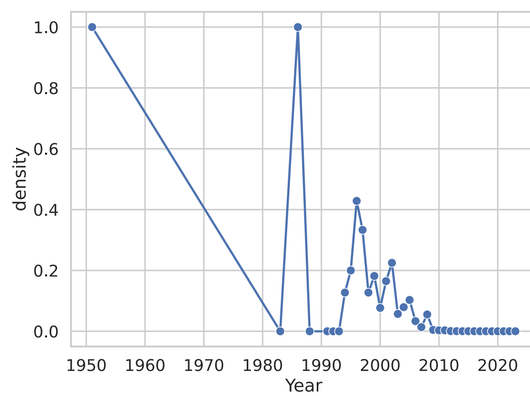
(a) Number of Authors



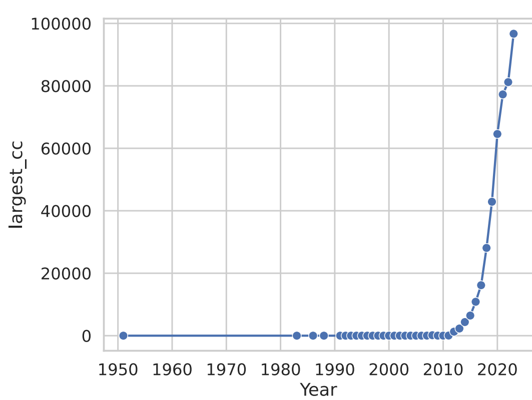
(b) Number of Co-Author Edges



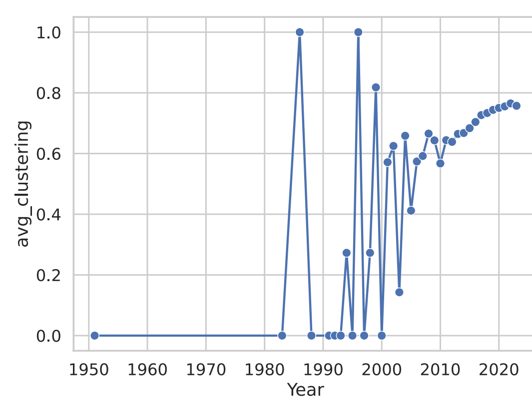
(c) Average Degree of Centrality



(d) Network Density



(e) Largest Connected Component Size



(f) Average Clustering Coefficient

Notes: This figure illustrates key metrics tracking the evolution of the AI co-authorship network. It includes the annual growth in the number of authors and collaborations, network density, average connections per author, the size of the largest collaboration cluster, and local cohesiveness through clustering coefficients. Together, these metrics highlight the expanding participation, increasing collaboration, and structural changes in the field over time.