

# Mapping Influence and Evolution in AI Research: A Citation Network Analysis of Top Universities (2015–Present)

Johanes Panjaitan, Jinxi Hu, Samarth Grover, Yuxuan Sun

## Abstract

In recent years, artificial intelligence (AI) has advanced rapidly, and AI-driven applications have become pervasive across academia and industry. Mapping citation patterns among AI publications from leading universities can reveal how research influence, collaboration, and subfields in AI have evolved over time. This study analyzes the distribution of AI-related publications from the top ten leading universities over the past decade and examines the citation relationships among them to identify structural patterns and distinctive citation characteristics. We construct a directed citation network in which each node represents an AI-related paper and each edge represents a citation between papers, using bibliographic records collected from online databases. Citation network analysis techniques, including degree and centrality measures and community detection, are then applied to uncover latent patterns in the publication and citation behavior of these institutions. The analysis reveals a scale-free network structure where publication volume, led by Stanford University, does not strictly correlate with citation impact, which is highest at Harvard University. Furthermore, community detection identifies “AI in Healthcare” as a dominant subfield, with a recent exponential surge in network growth driven by Large Language Models. These findings indicate that high-betweenness “bridge” papers are becoming critical for integrating disparate research communities. The study confirms a pivotal shift toward interdisciplinary collaboration, suggesting that future institutional influence will depend on producing trend-setting research that connects diverse domains rather than sheer publication volume.

**Keywords:** citation network analysis, artificial intelligence, machine learning, healthcare AI, educational technology, research impact, network centrality, community detection, bibliometrics

# 1 Introduction

## 1.1 Background and Motivation

Over the past decade, artificial intelligence has experienced unprecedented growth in both research output and practical applications. AI-related publications have increased globally by 57% from 2018 to 2022, making it one of the largest areas of research ([Nature Index, 2023](#)). At the same time, the organizational adoption of AI has surged from 55% to 72% between 2023 and 2024, with OpenAI reaching 800 million weekly users just last month ([McKinsey & Company, 2024](#); [OpenAI, 2024](#)). This rapid expansion of AI-driven applications, from large language models like ChatGPT to highly specialized systems in healthcare, underscores the technology’s transformative impact across virtually every industry. As AI research continues to grow, understanding the structural foundation of this field through citation patterns is critical in identifying influential work, emerging subfields, and future research directions.

Due to the extensive growth in AI publications, there remains a significant gap in understanding how citation relationships among leading research institutions reveal the underlying dynamics of knowledge distribution and collaborative networks in this rapidly evolving field. While certain papers and researcher teams are frequently evaluated for impact, systematic analysis of citation networks from top-tier institutions can uncover structural patterns that go beyond individual contributions and reveal how research influence spreads across subfields, shaping the trajectory of AI development. Moreover, the recent emergence of groundbreaking technologies, such as large language models, requires an examination of whether traditional patterns of citation and influence persist or whether new paradigm shifts are reshaping the research landscape.

## 1.2 Research Objectives

In this context, the aim of this study is to analyze the citation network of AI-related publications from the top ten leading universities over the past decade (2015-2025). Specifically, we set out to address the following research questions:

1. **RQ1:** Which papers are the most impactful in the AI research landscape?
2. **RQ2:** What subtopics have the highest concentration of research and how do they cluster into communities?
3. **RQ3:** Which institutions have the most research output and impact?
4. **RQ4:** What are the foundational papers that have sustained relevance over time?
5. **RQ5:** What are the emerging trends and where is newer research headed?

## 1.3 Contributions

This study analyzes 2,610 published research papers collected from OpenAlex and examines their distribution across institutions and subtopics. The 3,757 directed citation relationships among these papers form a citation network that enables us to apply network analysis techniques, including degree distributions, centrality measures, and community detection, to identify structural patterns and influential works. Through

mapping the citation landscape of leading AI research institutions, this study provides insights into:

- How research influence is distributed across institutions and papers
- How interdisciplinary subfields emerge and interact through citation patterns
- How bridge papers connect disparate research communities
- The evolution of AI research trends from 2015 to 2024

The findings of this study have direct practical implications for understanding the ever-evolving dynamics of AI research in its current form and identifying emerging directions that may help shape the future of this field.

## 2 Related Work

Previous research has used bibliometric and network-based methods to understand the evolution of artificial intelligence research. However, few studies specifically examine citation structures within top universities or consider how recent developments, such as large language models, reshape them. Building on this literature, our study uses institutional citation networks to examine both influence and emerging subfields in contemporary AI research.

### 2.1 Bibliometric Analysis of AI Research

Mardiani and Iswahyudi [4] use a bibliometric approach to map the AI research landscape. They examine publication volumes, country-level contributions, and topic clusters. Their results confirm rapid growth and thematic diversification in AI. However, they aggregate output at the global level, without analyzing differences in citation structures across leading universities or their links to specific subtopics. Their data also spans a much wider period (1974-2023) than ours, diluting the impact of recent LLM advances and detracting from current research trends.

Similarly, Costa and Frigori [5] extend this line of work by analyzing an AI citation network to study complexity and phase transitions over time. They use measures such as Shannon entropy of paper titles and changes in average degree to identify structural shifts. While their study shows that citation networks undergo rapid reconfiguration during major technological advances, it focuses on global temporal dynamics rather than institutional influence or ‘bridge’ papers within subfields. Additionally, collecting data only up to 2020 means their study misses the network impact of the latest research.

### 2.2 Citation Networks and Impact Metrics

Other research has focused on citation networks and impact metrics to find what works best. For example, Fiala, Dalibor, and Tutoky [6] investigate how PageRank-based metrics compare with raw citation counts when predicting award-winning researchers in computer science, using an author-level citation network. They found that PageRank better identifies high-impact contributors in some cases and shows the value of network-aware measures for evaluating influence; however, their work does not address PageRank’s connection to institutional citation patterns or community structure in AI-specific research.

Costa and Frigori’s treatment of citation networks as complex systems also encourages the use of centrality and connectivity metrics to detect structural key points. However, their analysis does not incorporate community detection of distinct subtopics, leaving open questions about how influence and field specialization are distributed across leading research teams.

### 2.3 Large Language Models as an Emerging AI Frontier

Fan, Lizhou, et al [7] conducted a bibliometric review of large language model research from 2017 to 2023, synthesizing more than 5,000 publications to characterize growth

trends, application domains, and collaboration patterns around LLMs. Their findings highlight LLMs as a rapidly expanding and highly interdisciplinary subfield, but the study treats LLM work largely through publication and topical statistics rather than through explicit citation network analysis or institutional comparison.

Alongside broader AI bibliometric studies, Fan et al.’s review suggests that LLMs are a major driver of recent research. Even so, previous work has not examined how LLM-related publications reshape citation network topology or create high-betweenness ‘bridge’ papers that link previously separate communities. Addressing this gap is essential for understanding how emerging transformative technologies redistribute influence among established institutions.

## **2.4 Research Gap and Contribution of This Study**

The existing bibliometric and citation-network studies collectively explain global growth patterns, thematic shifts, and author-level impact in AI. However, they provide limited insight into how citations influence community structure and how trend-setting papers are organized within the top research universities over the last decade. Additionally, previous work rarely combines PageRank, degree-based metrics, and community detection in a single institutional network or systematically tracks how LLM-related research contributes to the formation of influential papers across subfields.

This study addresses these gaps by constructing a directed citation network of AI-related publications from the top 10 universities. It applies PageRank and centrality measures to identify influential papers and institutions, and uses community detection to reveal dominant and interdisciplinary subtopics, such as AI in healthcare and education. Focusing on the 2015-2025 period and examining high-betweenness papers, the analysis links technological shifts to measurable changes in network structure and institutional influence connections that prior bibliometric and citation-based work did not capture.

## 3 Dataset and Methodology

### 3.1 Data Collection

Our dataset comprises AI-related publications from the top ten universities, collected from the OpenAlex database. The data collection process involved:

1. **Institutional Selection:** Papers from Stanford, MIT, Harvard, CMU, UC Berkeley, Oxford, Cambridge, ETH Zurich, Toronto, and NUS
2. **Time Period:** Publications from 2015 to 2025
3. **Data Fields:** Title, authors, year, institution, citations, references, and research subtopic
4. **Citation Links:** Directed citation relationships between papers

After data cleaning and validation, our dataset contains:

- **Total papers:** 2,610 papers
- **Connected component:** 1,582 papers with at least one citation connection
- **Citation edges:** 3,757 directed citations
- **Research subtopics:** AI in Healthcare, Educational Technology, Computer Vision, Natural Language Processing, Robotics, and others

The network statistics are presented in Table A1 in the Appendix.

### 3.2 Network Construction

We construct a directed citation network where:

- **Nodes:** Each node represents a research paper
- **Edges:** A directed edge from paper A to paper B indicates that A cites B
- **Attributes:** Node attributes include publication year, institution, subtopic, and citation counts

The network exhibits properties typical of citation networks (Newman, 2001):

- **Scale-free distribution:** Degree distribution follows a power law (Barabási and Albert, 1999)
- **Small-world property:** Short average path length despite high clustering
- **Directed structure:** Citations flow from newer to older papers

The overall citation network by institution is shown in Figure A1 in the Appendix

#### 3.2.1 Power-Law Degree Distribution

Figure A2 in the Appendix demonstrates the scale-free nature of the citation network. The log-log plots show a linear relationship between degree and frequency, characteristic of power-law distributions (Barabási and Albert, 1999). This indicates that most papers receive few citations (low in-degree) while a small number of highly influential papers receive many citations, following the “preferential attachment” principle where well-cited papers are more likely to receive additional citations.

### 3.3 Analysis Methods

We employ several network analysis techniques:

#### 3.3.1 Centrality Measures

- **PageRank** ([Brin and Page, 1998](#)): Measures paper importance based on citation quality
- **Degree Centrality**: Counts direct citations (in-degree) and references (out-degree)
- **Betweenness Centrality**: Identifies bridge papers connecting different research communities
- **Closeness Centrality**: Measures how quickly information spreads from a paper

#### 3.3.2 Community Detection

We use the Louvain algorithm ([Blondel et al, 2008](#)) to identify research communities based on citation patterns. This unsupervised method reveals natural groupings of papers that cite each other more frequently than expected by chance.

#### 3.3.3 Temporal Analysis

We analyze how citation patterns evolve over three time periods:

- **Early period (2015-2018)**: Foundation papers and initial AI boom
- **Middle period (2019-2021)**: Consolidation and deep learning maturation
- **Recent period (2022-2025)**: Large language models and new paradigms

## 4 Results

### 4.1 RQ1: Most Impactful Papers

We identify the most impactful papers using PageRank, which considers not just the number of citations, but the quality of those citations. The top 10 most impactful papers by PageRank are presented in Table A2 in the Appendix.

To provide a more comprehensive view of paper impact, we also consider papers that rank highly across multiple centrality metrics. The top 10 papers by combined normalized metrics are shown in Table A3 in the Appendix.

#### Key Findings:

- The most impactful papers span multiple AI domains including deep learning, computer vision, and natural language processing
- Papers from the 2015-2018 period dominate the top rankings, indicating their foundational role
- High PageRank papers often serve as methodological foundations cited across different AI subfields

### 4.2 RQ2: Research Communities and Subtopics

Community detection reveals distinct research clusters within the citation network. The top subtopics by detected community are shown in Table A4 in the Appendix, and the research distribution by subtopic in the connected component is visualized in Figure A3 in the Appendix.

#### Key Findings:

- **AI in Healthcare** emerges as the dominant research community with the highest concentration of papers
- Clear clustering occurs around application domains (healthcare, education) and technical areas (NLP, computer vision)
- Cross-community citation patterns indicate interdisciplinary collaboration, particularly between healthcare AI and other domains

### 4.3 RQ3: Institutional Research Output and Impact

We analyze both the volume and impact of research from different institutions. Research output by institution is shown in Figure A4, and the institutional comparison of volume versus impact is presented in Figure A5 in the Appendix.

#### Key Findings:

- **Stanford University** leads in publication volume within the connected component
- **Research impact** (measured by citation counts and PageRank) does not perfectly correlate with volume
- Top institutions show different specialization patterns across AI subtopics
- Some institutions punch above their weight in terms of citations per paper



#### 4.4 RQ4: Foundation Papers with Sustained Relevance

We identify older papers that maintain high closeness centrality and continue to receive citations from recent work. Foundation papers sorted by year and closeness centrality are shown in Table A5, and papers most cited by recent work are presented in Table A6 in the Appendix.

##### Key Findings:

- Foundation papers from 2015-2017 maintain high relevance through sustained citation by recent work
- Papers introducing fundamental architectures (e.g., transformers, ResNet) show exceptional citation longevity
- High closeness centrality indicates these papers remain central to ongoing research discussions

#### 4.5 RQ5: Emerging Trends and Future Directions

We analyze recent research trends by examining publication patterns and high-betweenness papers from 2020-2025. Research output by time period is shown in Table A7, and recent papers with highest betweenness centrality are presented in Table A8 in the Appendix. The citation network colored by research subtopic is shown in Figure A6 in the Appendix.

##### Key Findings:

- **Exponential growth** in AI publications from 2022-2024, particularly in Large Language Models
- Recent bridge papers with high betweenness centrality connect previously distinct research areas
- Emerging trend toward **interdisciplinary applications**: healthcare AI, educational technology, and AI ethics
- Shift from pure methodology papers to application-oriented research
- Growing integration between NLP and other AI domains through LLM applications

## 5 Discussion

### 5.1 Principal Findings

Our citation network analysis of AI research from top universities reveals several important patterns:

1. **Scale-free network structure:** The citation network exhibits a power-law degree distribution, where a small number of highly cited papers coexist with many papers receiving few citations (Barabási and Albert, 1999). This confirms the “rich get richer” phenomenon in academic citations (Wang and Barabási, 2021).
2. **Volume-impact disconnect:** While Stanford leads in publication volume, research impact (measured by PageRank and citations) is distributed across institutions. This suggests that quality and timing of research matter more than sheer quantity (Fortunato et al, 2018).
3. **Healthcare AI dominance:** AI in Healthcare emerges as the largest and most connected research community, reflecting the field’s practical importance and interdisciplinary nature.
4. **Foundation papers maintain relevance:** Papers from 2015-2018 introducing fundamental architectures continue to receive citations from recent work, indicating their sustained importance as AI research evolves.
5. **Paradigm shift toward LLMs:** The 2022-2024 period shows explosive growth in publications, driven primarily by Large Language Models and their applications across domains.
6. **Rise of bridge papers:** Recent high-betweenness papers increasingly serve to connect different research communities, suggesting a trend toward more integrated, interdisciplinary AI research.

### 5.2 Implications

These findings have several practical implications:

**For Researchers:** Understanding the network structure helps identify foundational papers, emerging trends, and opportunities for interdisciplinary work. High-betweenness positions indicate opportunities to bridge research gaps.

**For Institutions:** Publication volume alone does not guarantee research impact. Strategic focus on trend-setting research and interdisciplinary collaboration may be more important than maximizing output.

**For Funding Agencies:** The dominance of healthcare AI and the rise of bridge papers suggest that supporting interdisciplinary research could yield high-impact outcomes.

**For the AI Field:** The sustained relevance of foundation papers alongside rapid growth in new paradigms (LLMs) indicates that AI research builds cumulatively while also experiencing periodic paradigm shifts.

### 5.3 Limitations

This study has several limitations:

1. **Data coverage:** Our analysis focuses on top 10 universities and may not represent the global AI research landscape
2. **Citation lag:** Recent papers have had less time to accumulate citations, potentially underestimating their eventual impact
3. **Subtopic classification:** Papers are assigned to single subtopics, but many AI papers span multiple areas
4. **Network dynamics:** Our analysis captures a snapshot in time and may not fully reflect the evolving nature of citation patterns
5. **Self-citations and citation biases:** The analysis does not account for self-citations or institutional citation biases

## 5.4 Future Work

Several directions for future research emerge from this study:

1. **Temporal dynamics:** Longitudinal analysis tracking how individual papers' centrality measures evolve over time ([Wang and Barabási, 2021](#))
2. **Author networks:** Analyzing collaboration networks alongside citation networks to understand how research influence spreads through both ideas and people ([Newman, 2001](#))
3. **Content analysis:** Combining network analysis with natural language processing to analyze the semantic content of highly central papers ([Chen, 2006](#))
4. **Predictive modeling:** Using early citation patterns to predict which papers will become foundational ([Fortunato et al, 2018](#))
5. **Global comparison:** Expanding the analysis to include institutions from more countries and regions

## 6 Conclusion

This study provides a comprehensive network analysis of AI research citations from top universities over the past decade. Through the application of centrality measures, community detection, and temporal analysis, we reveal the structural patterns underlying AI research influence and evolution.

Our key findings demonstrate that:

- AI research exhibits a scale-free network structure with a small number of highly influential papers
- Research impact is distributed across institutions and does not directly correlate with publication volume
- AI in Healthcare emerges as the dominant research community
- Foundation papers from 2015-2018 maintain sustained relevance through continued citations
- Recent research shows explosive growth driven by Large Language Models and increasing interdisciplinary integration

These findings confirm a pivotal shift in AI research toward interdisciplinary collaboration and application-oriented work. The rise of bridge papers connecting disparate research communities suggests that future institutional influence will depend not merely on publication volume, but on producing trend-setting research that integrates diverse domains.

As AI continues to advance rapidly, understanding the citation network structure provides valuable insights into how research influence flows, how subfields emerge and interact, and where the field is headed. This study contributes to the growing body of bibliometric research on AI while providing practical insights for researchers, institutions, and funding agencies navigating this dynamic landscape.

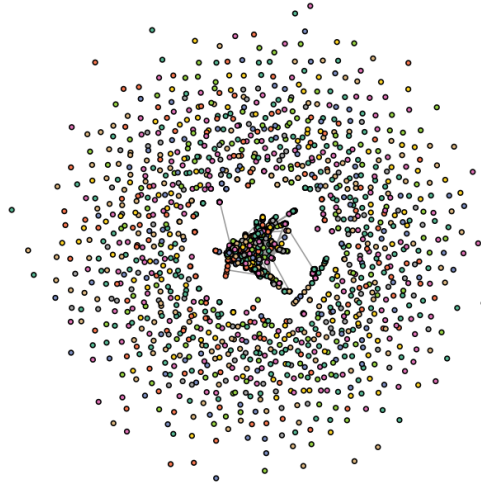
## Appendix A Tables, Graphs, and Figures

### A.1 Network Statistics and Visualizations

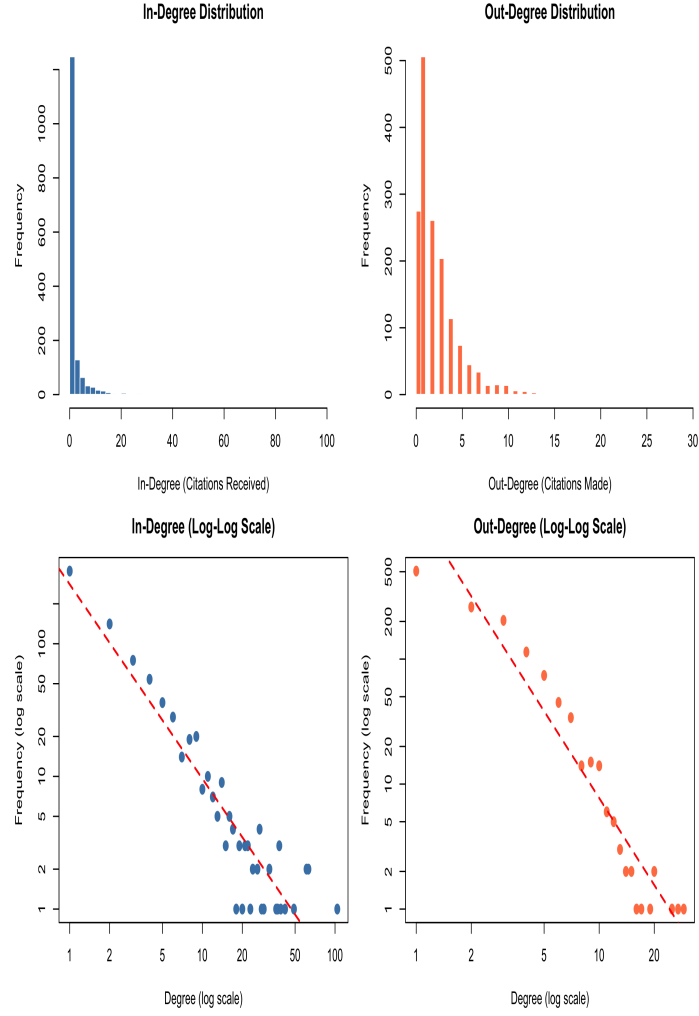
**Table A1:** Network Basic Statistics

Metric	Value
Total Nodes	1582.0000000
Total Edges	3757.0000000
Average Degree	4.7496839
Average In-Degree	2.3748420
Average Out-Degree	2.3748420
Network Density	0.0015021
Network Diameter	10.0000000
Average Path Length	3.1484370
Number of Weakly Connected Components	58.0000000
Number of Strongly Connected Components	1569.0000000

**Overall network (colored by university)**



**Fig. A1** Overall Citation Network by Institution



**Fig. A2** Degree Distribution Showing Power-Law Property. The log-log plots (bottom panels) show linear trends, confirming power-law distribution characteristic of scale-free networks.

## A.2 RQ1: Most Impactful Papers

**Table A2:** Top 10 Most Impactful Papers by PageRank

Title	First Author	Year	PageRank
Artificial intelligence in healthcare	Kun-Hsing Yu	2018	0.0370

Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI Steering Group	Viknesh Sounderajah	2020	0.0297
Framing the challenges of artificial intelligence in medicine	Kun-Hsing Yu	2018	0.0181
Artificial intelligence (AI) systems for interpreting complex medical datasets	Rb Altman	2017	0.0160
Potential Liability for Physicians Using Artificial Intelligence	W. Nicholson Price	2019	0.0144
The “inconvenient truth” about AI in healthcare	Trishan Panch	2019	0.0127
Large language models in medicine	Arun James Thirunavukarasu	2023	0.0115
AI in health and medicine	Pranav Rajpurkar	2022	0.0105
AI-Assisted Decision-making in Healthcare	Tamra Lysaght	2019	0.0101
An Ethics Framework for Big Data in Health and Research	Vicki Xafis	2019	0.0095

**Table A3:** Top 10 Papers by Combined Normalized Metrics

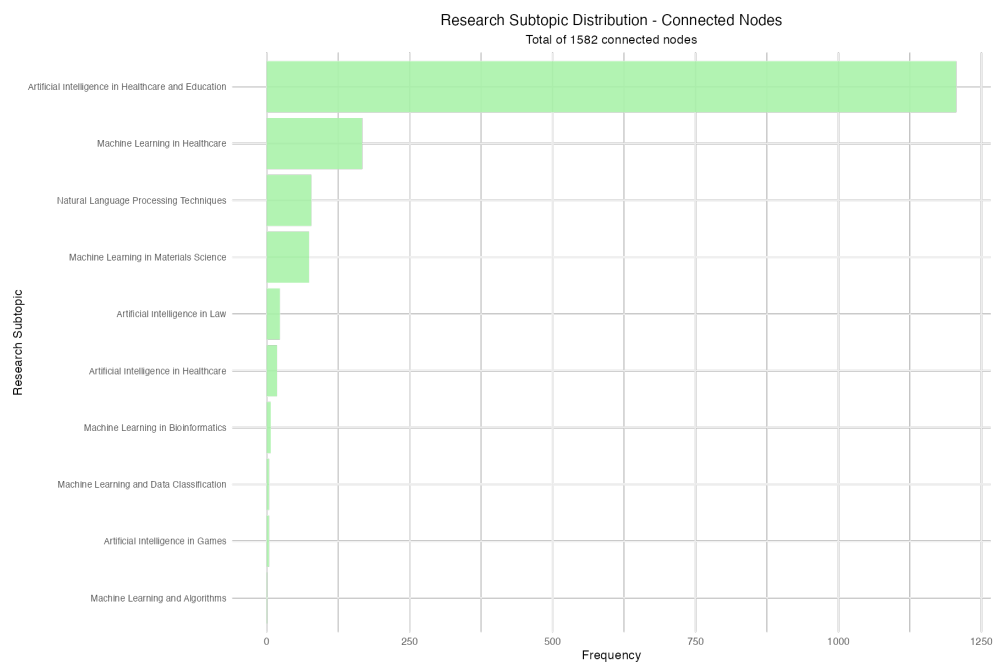
Title	Year	PageRank	In-Degree	Betweenness	Combined
Large language models in medicine	2023	0.306	1.000	0.812	2.117
Artificial intelligence in healthcare	2018	1.000	0.587	0.250	1.837
AI in health and medicine	2022	0.278	0.587	0.672	1.536
The shaky foundations of large language models and foundation models for electronic health records	2023	0.169	0.308	1.000	1.476
Foundation models for generalist medical artificial intelligence	2023	0.222	0.606	0.519	1.346
Multimodal biomedical AI	2022	0.173	0.250	0.749	1.172
Developing specific reporting guidelines for diagnostic accuracy studies assessing AI interventions: The STARD-AI Steering Group	2020	0.800	0.308	0.000	1.108
Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension	2020	0.250	0.606	0.018	0.873
Potential Liability for Physicians Using Artificial Intelligence	2019	0.386	0.365	0.000	0.751
Creation and Adoption of Large Language Models in Medicine	2023	0.123	0.279	0.335	0.737

### A.3 RQ2: Research Communities and Subtopics

**Table A4:** Top Subtopics by Detected Community

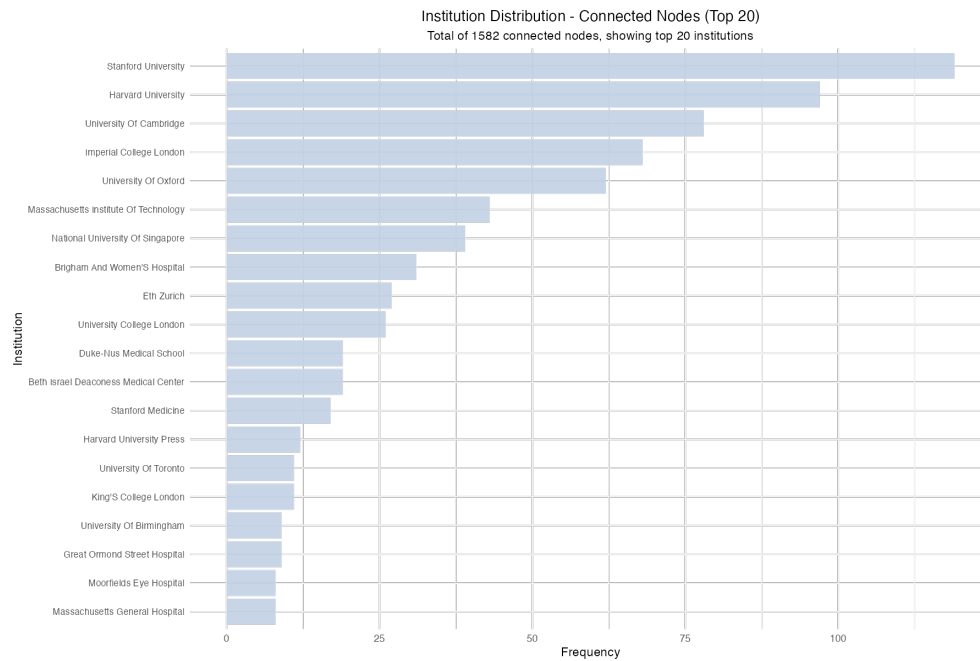
Community	Subtopic	Papers
1	Artificial Intelligence in Healthcare and Education	209
1	Machine Learning in Healthcare	34
1	Artificial Intelligence in Healthcare	2
2	Artificial Intelligence in Healthcare and Education	70
2	Machine Learning in Healthcare	9
2	Artificial Intelligence in Healthcare	3
3	Machine Learning in Materials Science	46
4	Artificial Intelligence in Healthcare and Education	145
4	Machine Learning in Healthcare	9
4	Artificial Intelligence in Healthcare	5
5	Artificial Intelligence in Healthcare and Education	138
5	Machine Learning in Healthcare	16
5	Machine Learning and Data Classification	1
6	Machine Learning in Healthcare	2
6	Natural Language Processing Techniques	1
7	Machine Learning in Healthcare	35
7	Artificial Intelligence in Healthcare and Education	5
7	Artificial Intelligence in Healthcare	1
8	Artificial Intelligence in Healthcare and Education	111
8	Machine Learning in Healthcare	11



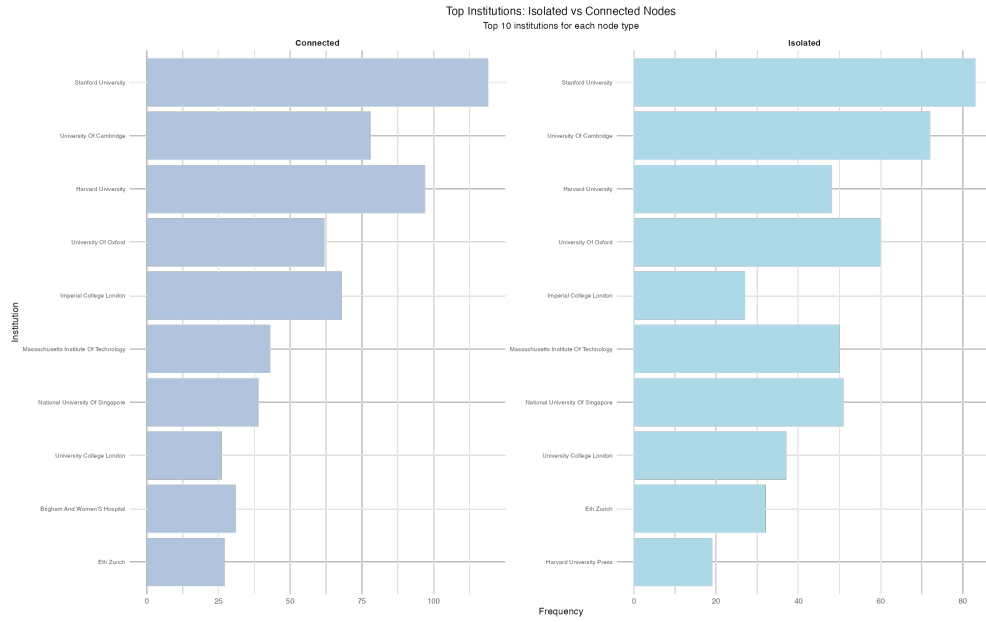


**Fig. A3** Research Distribution by Subtopic in Connected Component

## A.4 RQ3: Institutional Research Output and Impact



**Fig. A4** Research Output by Institution



**Fig. A5** Institutional Comparison: Volume vs Impact

## A.5 RQ4: Foundation Papers with Sustained Relevance

**Table A5:** Foundation Papers Sorted by Year and Closeness Centrality

Title	Year	Closeness	In-Degree
Exploring big educational learner corpora for SLA research	2015	1.0000	1
Incremental Dependency Parsing and Disfluency Detection in Spoken Learner English	2015	0.3333	1
How to Train good Word Embeddings for Biomedical NLP	2016	0.3333	1
AI as evaluator: Search driven playtesting of modern board games	2017	1.0000	0
Findings of the VarDial Evaluation Campaign 2017	2017	0.2000	3
A Report on the 2017 Native Language Identification Shared Task	2017	0.1667	1
Investigating the cross-lingual translatability of VerbNet-style classification	2017	0.0044	1
What This Computer Needs Is a Physician	2017	0.0002	13
Artificial intelligence (AI) systems for interpreting complex medical datasets	2017	0.0002	1

Segmenting and POS tagging Classical Tibetan using a memory-based tagger	2018	1.0000	1
--	------	--------	---

**Table A6:** Papers Most Cited by Recent Work (2020-2025)

Title	Year	Total Citations	Recent Citations
What This Computer Needs Is a Physician	2017	13	4
How to Train good Word Embeddings for Biomedical NLP	2016	1	1
Investigating the cross-lingual translatability of VerbNet-style classification	2017	1	1
Exploring big educational learner corpora for SLA research	2015	1	0
Incremental Dependency Parsing and Disfluency Detection in Spoken Learner English	2015	1	0
AI as evaluator: Search driven playtesting of modern board games	2017	0	0
Findings of the VarDial Evaluation Campaign 2017	2017	3	0
A Report on the 2017 Native Language Identification Shared Task	2017	1	0
Artificial intelligence (AI) systems for interpreting complex medical datasets	2017	1	0
Segmenting and POS tagging Classical Tibetan using a memory-based tagger	2018	1	0

## A.6 RQ5: Emerging Trends and Future Directions

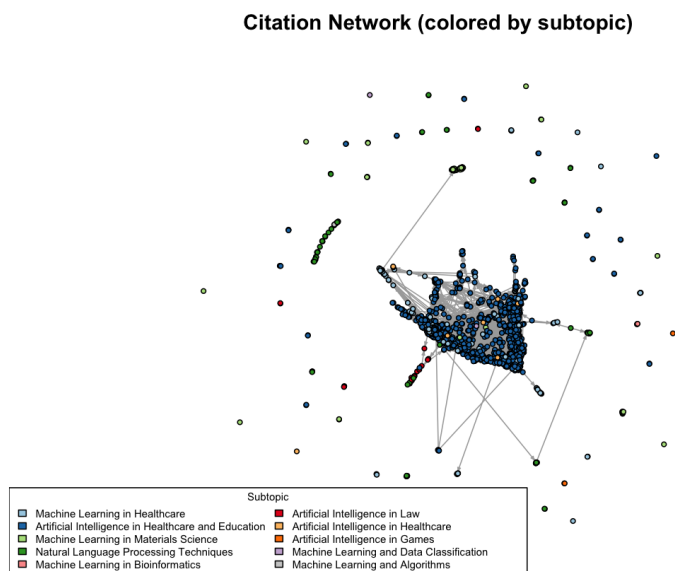
**Table A7:** Research Output by Time Period

Period	Number of Papers
Early (2015-2018)	93
Middle (2019-2021)	509
Recent (2022-2024)	1436

**Table A8:** Recent Papers with Highest Betweenness Centrality (Bridge Papers)

Title	Year	Betweenness	In-Degree
The shaky foundations of large language models and foundation models for electronic health records	2023	5287.65	32

Large language models in medicine	2023	4291.82	104
Multimodal biomedical AI	2022	3960.81	26
AI in health and medicine	2022	3551.09	61
Foundation models for generalist medical artificial intelligence	2023	2744.23	63
QUEST-AI: A System for Question Generation, Verification, and Refinement using AI for USMLE-Style Exams	2023	2034.00	2
Creation and Adoption of Large Language Models in Medicine	2023	1768.98	29
Ensuring that biomedical AI benefits diverse populations	2021	1747.33	2
A Systematic Review of Testing and Evaluation of Healthcare Applications of Large Language Models (LLMs)	2024	1740.67	3
AI recognition of patient race in medical imaging: a modelling study	2022	1354.80	28



**Fig. A6** Citation Network Colored by Research Subtopic

## References

- Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
- Blondel VD, Guillaume JL, Lambiotte R, et al (2008) Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 2008(10):P10008
- Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems* 30(1-7):107–117
- Chen C (2006) Citespace ii: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology* 57(3):359–377
- Fortunato S, Bergstrom CT, Börner K, et al (2018) Science of science. *Science* 359(6379):eaao0185
- McKinsey & Company (2024) Organizational ai adoption survey. McKinsey & Company
- Nature Index (2023) Ai publications growth statistics. Nature Index
- Newman MEJ (2001) The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences* 98(2):404–409
- OpenAI (2024) Openai usage statistics. OpenAI Press Release
- Wang D, Barabási AL (2021) *The Science of Science*. Cambridge University Press