

【学术前沿动态】人工智能生成内容（AIGC）主题论文分析

在当前数字世界和物理世界加速融合的大背景下，人工智能生成内容（Artificial Intelligence Generated Content，简称 AIGC）正在悄然引导着一场深刻的变革，重塑甚至颠覆数字内容的生产方式和消费模式，也为各个行业与学科领域带来了巨大的机会与挑战。本期学术前沿动态即以 AIGC 为主题，梳理相关文献发展脉络，探究各国研究主题分布，分析国内外机构发文表现，以期为相关科研人员提供参考。

本报告检索用数据库为 Web of Science (WOS) 核心合集 (SCIE、SSCI、AHCI、CPCI-S、CPCI-SSH) 和中国知网 (CNKI)，文献类型包含期刊论文与会议论文（含 Early Access 和网络首发论文），文献时间窗为 2019/1/1-2024/5/13，检索日期为 2024 年 5 月 14 日。因本文论文集未包含预印本文献，故研究热点（特别是技术类）等结论有所滞后。

一、论文发文趋势

2019 年至今，CNKI 检出 6105 篇（期刊论文 5907 篇、会议论文 198 篇），WOS 检出 AIGC 相关论文 15122 篇，其中 SCI 论文 7261 篇、SSCI 论文 2600 篇、A&HCI 论文 262 篇、CPCI-S 论文 6300 篇、CPCI-SSH 论文 982 篇。图 1 反映了 2019-2024 年 AIGC 相关论文数量分布情况，可以看出，CNKI 中 AIGC 相关论文数量在 2019-2022 年间并无显著增长，2023 年发文量则由 2022 年的 335 篇增至 3188 篇，呈现爆发式增长；WOS 论文数显著多于 CNKI 发文量，且呈持续增长趋势，其中自然科学领域论文 (SCI 和 CPCI-S) 显著多于人文社会与艺术类论文 (SSCI、CPCI-SSH 和 A&HCI)，2019-2022 年间 SCI 相关发文增幅始终高于 SSCI 发文增幅，2023 年则呈现相近增幅，考虑到 2022 年 11 月 30 日 OpenAI 发布 ChatGPT 引发 AIGC 新浪潮，可推测 2023 年中外期刊论文数量激增与此有关。

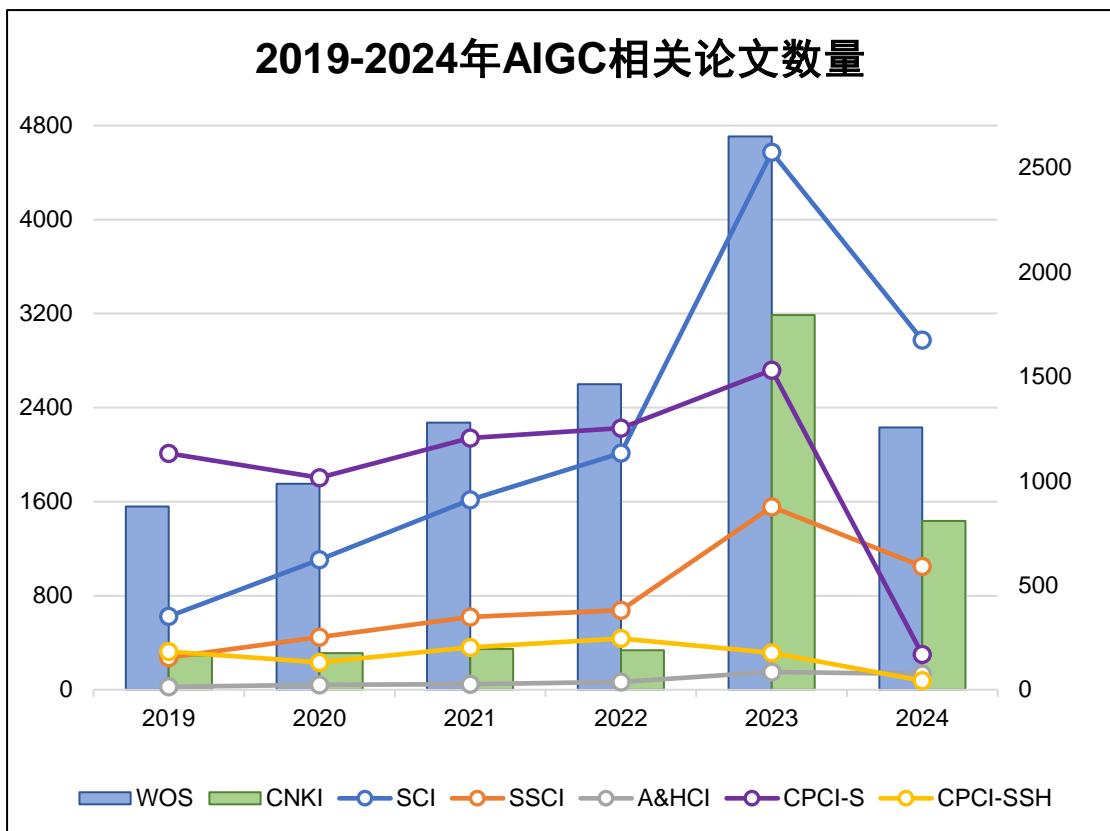


图 1 2019-2024 年 AIGC 相关论文数量

注：由于论文来源存在多重收录现象，如某一期刊可能同时被 SCI 和 SSCI 收录，故各库论文数量加和不等于论文总数。

二、学科领域分布

图 2 展示了 CNKI 与 WOS 发文量排名前 10 的学科领域及相应论文数量。可知，CNKI 论文中，数量最大的为自动化技术与计算机软件及计算机应用，两学科发文总数约占全部论文数量的 70%，其余发文主要集中于人文社会科学领域，如新闻与传媒、信息经济与邮政经济、教育理论与教育管理等；计算机科学是 WOS 发文最集中的领域，其次为电子电气工程、电信等，影像科学和照相技术、语言学等学科亦有较多发文。

AIGC 相关 SCI 和 SSCI 论文数量排名前 10 的学科领域见图 3。数据显示，SCI 论文中比重最大的 WOS 学科为电气和电子工程，其次为人工智能、信息系统等计算机科学，电信、卫生保健科学和服务、医学信息学及应用物理学等排名其后；教育和教学研究是 AIGC 相关 SSCI 发文量最多的学科领域，其次为商业、信息科学与图书馆科学等，卫生保健科学和服务、医学信息学以及部分计算机科学在 SSCI 发文中同样占据较大比重，管理学、人体工程学、心理学等学科对 AIGC 亦有较多关注。



图 2 AIGC 相关中外文献 Top10 学科

注：一篇论文可能被同时划分至不同学科领域中，故不同学科论文数量加和不等于论文总数。

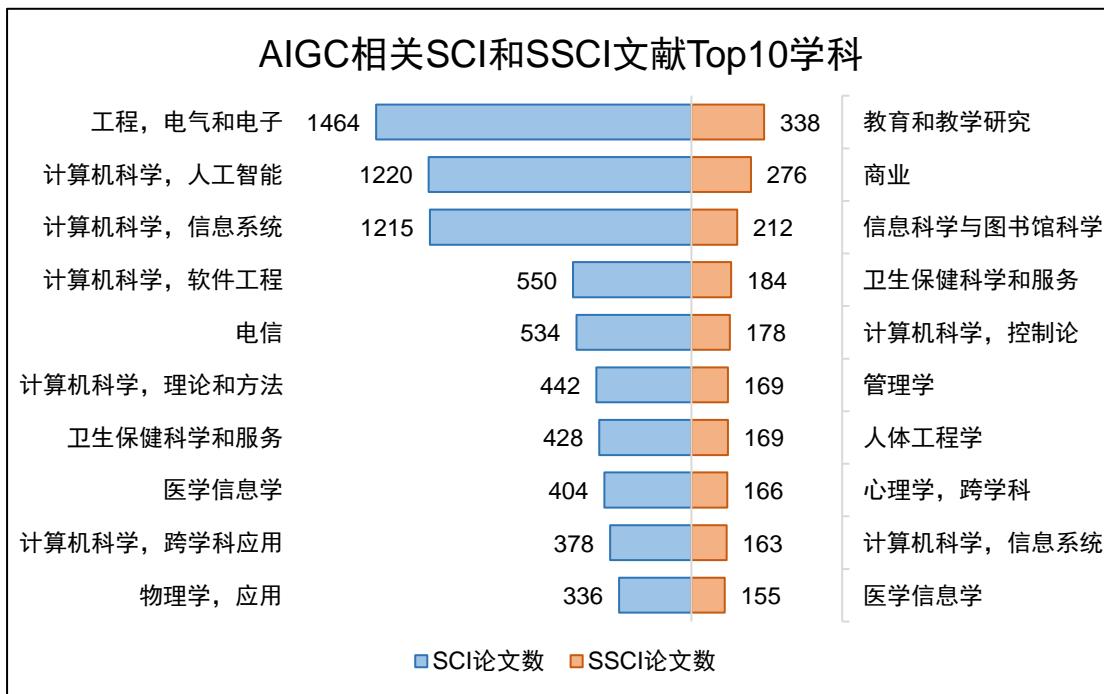


图 3 AIGC 相关 SCI 和 SSCI 文献 Top10 学科

对 2019-2024 年 AIGC 相关的 CNKI 论文进行关键词共现分析（图 4）发现：2021 年前后，国内论文主要聚焦 AIGC 相关基础技术（如深度学习、自然语言处理、注意力机制等）以及著作权、知识产权等相关主题；经过 2022 年对人工智能的爆发式讨论后，AIGC 成为

热点话题，2023 年始则重点围绕 ChatGPT 和生成式 AI 展开研究，内容涉及大语言模型、人机协同、SORA、教育数字化、知识生产、智慧图书馆、编辑出版、伦理风险、技术治理、法律规制等多个领域。可以看出，国内论文在分析 AIGC 相关应用的同时对其带来的法律与道德风险予以了较多关注。

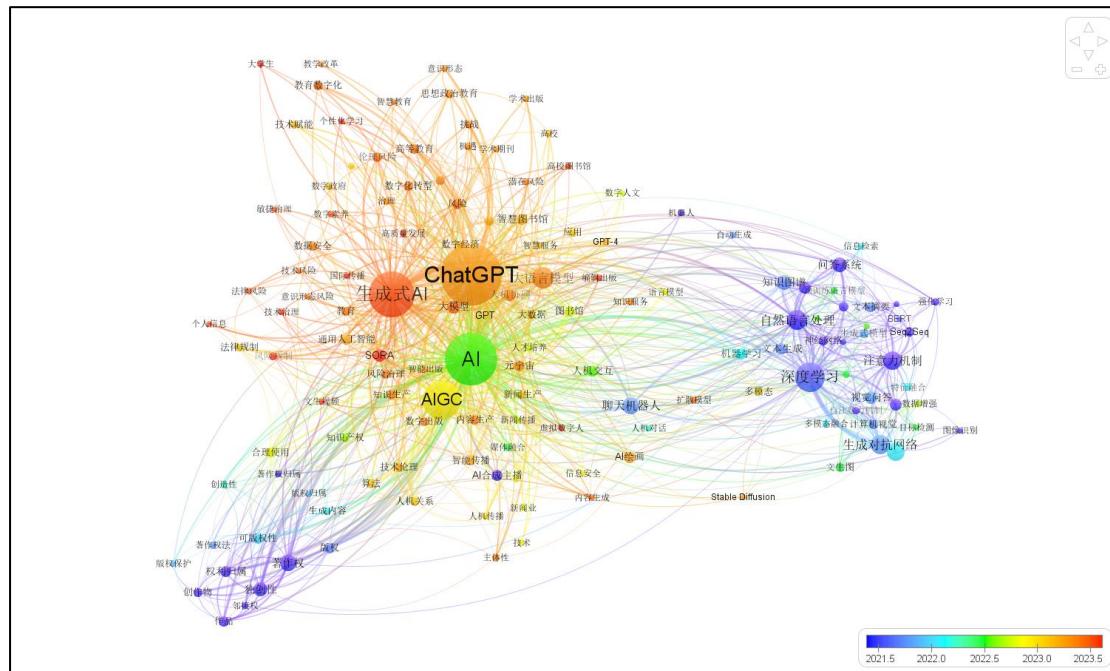


图 4 AIGC 相关 CNKI 论文关键词共现图

WOS 相关论文关键词共现如图 5 所示。可知, 2020 年以前, 研究热点主要集中在 Conversational Agents (对话代理), 2021 年转变为 Chatbot (聊天机器人); 随着 2022 年新一代人工智能的崛起, ChatGPT、Generative AI (生成式 AI)、Large Language Models (大语言模型) 成为新的研究热潮, 与此相关的研究主题除 Transformer、GPT (Generative Pre-trained Transformer, 生成式预训练 Transformer 模型)、prompt engineering (提示词工程)、language models (语言模型) 等技术讨论外, 还有 (higher) education (教育/高等教育)、medical education (医学教育)、diagnosis (诊断)、ethic (伦理)、bias (偏见)、quality (质量)、creativity (创新性)、challenges (挑战)、security (安全) 等。可见, 伴随 AIGC 的发展, 还引发了大量对教育变革、医学发展、伦理道德、内容生成质量与安全等相关思考。

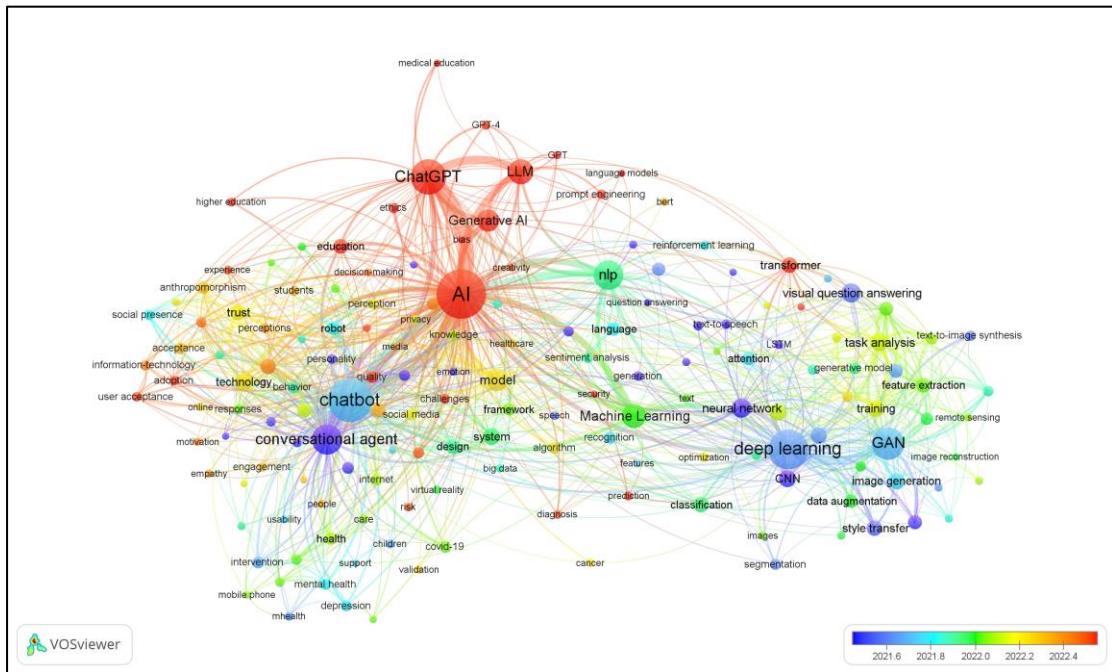


图 5 AIGC 相关 WOS 论文关键词共现图

三、国家/地区分布

图 6 为 AIGC 相关 WOS 发文前 10 国家/地区的 Top10 学科分布图。从发文量来看, WOS 所收录的 AIGC 领域论文中, 美国位居第一, 中国排名第二, 我国也是排名前五的国家/地区中唯一一个亚洲国家。人工智能等计算机科学子类学科以及电子电气工程是各个国家发文的重点学科, 我国对电信、影像科学和照相技术领域也有较多涉猎。从学科发文相对全球平均水平的影响力来看, 计算机科学领域中影响力较高的主要为美国、中国和澳大利亚, 特别是澳大利亚的计算机理论和方法、计算机控制论、人工智能以及我国的计算机理论和方法领域表现突出; 德国在 AIGC 与影像科学和照相技术, 放射学、核医学和医学成像, 以及管理学等相关的研究中具有较高的学术影响力; 卫生保健科学和服务、医学信息学领域论文影响力则以英国/英格兰见长; 英国/英格兰、韩国、印度在 AIGC 与商业的相关研究中表现较为突出; 我国 Top10 学科影响力均超过或接近全球平均水平, 在计算机科学的理论和方法、人工智能、跨学科应用、控制论以及电信科学领域表现较优, 但计算机硬件和体系结构、影像科学和照相技术领域的学术影响力略低于全球平均水平。

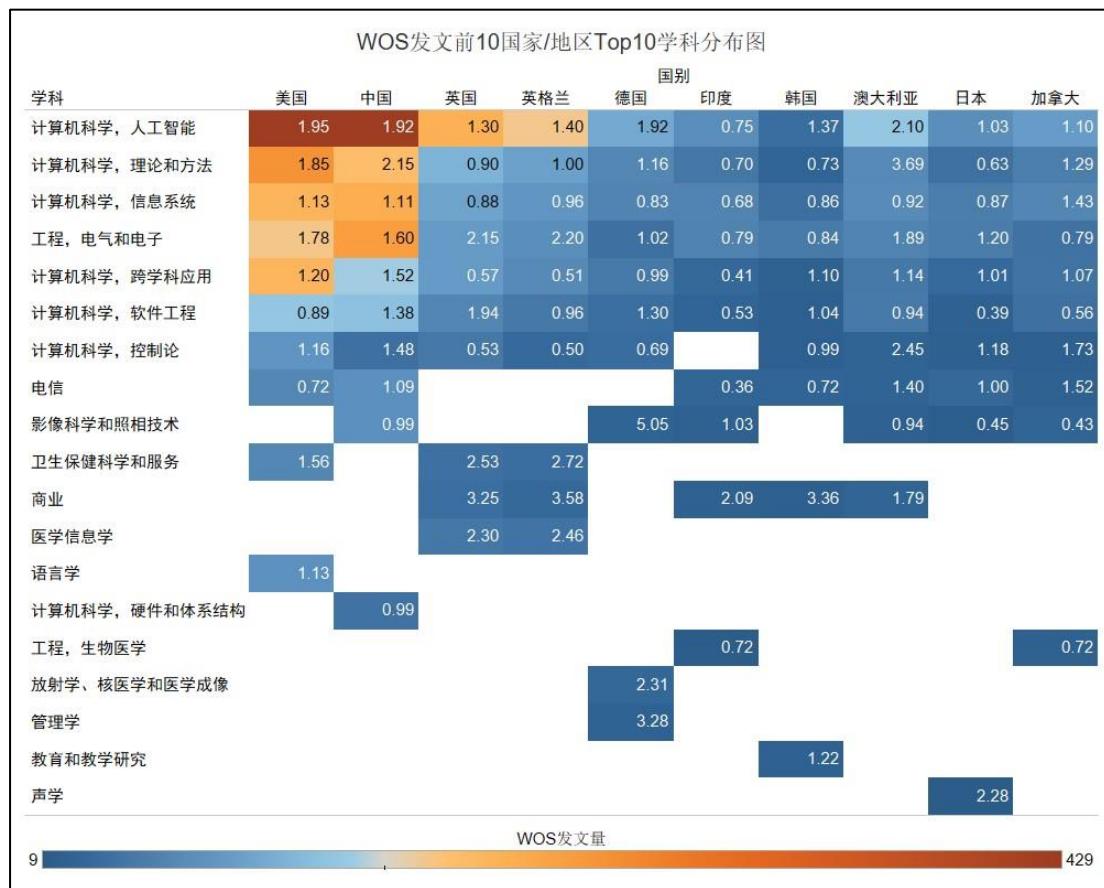


图 6 WOS 发文前 10 国家/地区 Top10 学科分布图

注：1. 色调越暖表示发文量越多，色块中标签数字为该国在该学科发文相对于全球平均水平的影响力，数字大于 1 则表明其影响力超过全球平均水平；图片仅展示 10 个国家发文量前 10 的学科，无色块与标签数字表明该学科未上榜该国发文量前 10 名。

2. 数据来源为 InCites 数据库，其中英国发文量包含英格兰地区发文量和其他英属地发文量，图 7 同。

图 7 为 WOS 发文前 10 国家/地区的 Top5 微观研究主题分布图。可以看出，10 个国家/地区的 AIGC 相关论文 Top5 微观研究主题集中在深度学习、自然语言处理、人机交互、医疗保健中的人工智能、西布曲明（一种药品）、用户满意度、语音识别、剽窃等 8 种中，其中西布曲明主题相关论文主要为聊天机器人等在医疗健康（特别是心理健康治疗）中的应用研究。我国发文量最多的两大主题是深度学习和自然语言处理，近五年发文数量分别为 371 篇和 138 篇，其中深度学习领域有 12 篇入选 ESI 高被引论文。而从论文质量和影响力来看，我国优势领域在于用户满意度主题，被引次数排名前 1% 的论文比例为 31.82%，相对于全球平均水平的影响力高达 6.61，而医疗保健中的人工智能领域论文影响力相比欧美国家仍有较大差距。

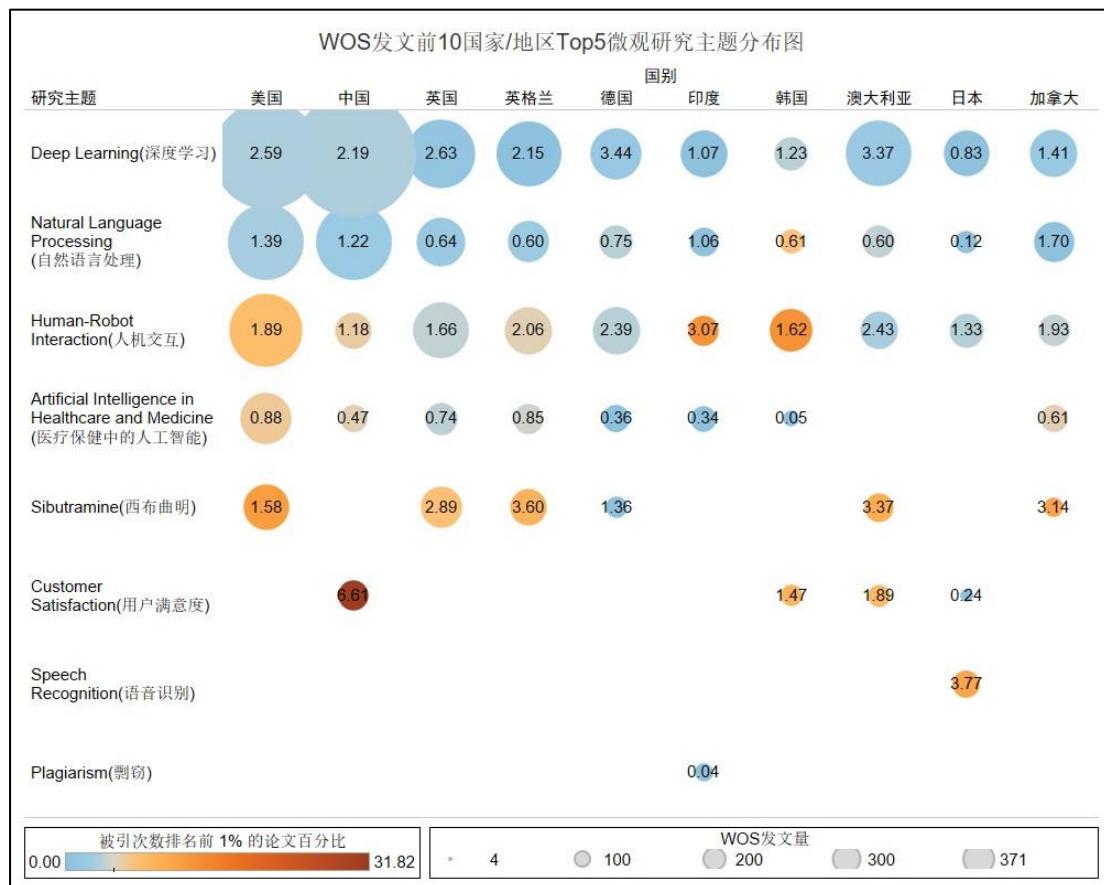


图 7 WOS 发文前 10 国家/地区 Top5 微观研究主题分布图

注：气泡越大表示发文量越大，气泡色调越暖表示被引次数排名前 1% 的论文比例越大，气泡中的数字标签为相对于全球平均水平的影响力，数字大于 1 则表明其影响力超过全球平均水平，无气泡则表明该研究主题未上榜该国家/地区发文量 Top5 主题。

四、机构分布

AIGC 相关 WOS 论文数排名前 20 的机构及其论文表现如表 1 所示。由图可知，发文最多的为中国科学院，其次是加利福尼亚大学系统（University of California System）、清华大学、伦敦大学和中国科学院大学等，排名前五的机构中我国有三所，其中中国科学院不仅拥有最高的发文量，高被引论文数量同样排名第一；美国是上榜机构最多的国家（9 所），占前 20 名机构的 45%，我国共有 6 所机构上榜，除排名前五的三所机构外还有浙江大学、北京大学和上海交通大学，分别位列第 8 名、第 14 名和第 17 名；微软、斯坦福大学、佛罗里达州立大学系统和佐治亚大学系统发文量虽未能跻身前五名，但却拥有较高的论文质量和影响力。

AIGC 相关 WOS 论文数排名前 20 的国内机构及其论文表现如表 2 所示。数据显示，我校近五年相关论文数量为 70 篇，位列第 14 名，被引次数排名前 1% 的论文占比排名第五（8.57%），表明发文中高水平论文占比较高。电子科技大学虽仅有 69 篇 AIGC 相关论文，但被引频次高达 2221 次，在论文质量和影响力方面表现突出。

图 8 展示了 CNKI 发文量 Top20 机构及其被引情况。图表显示，无论是 CNKI 论文数还是总被引频次，北京师范大学均有最优表现，陕西师范大学拥有最高的篇均被引次数；我校相关论文量为 79 篇（第 5 名），总被引 913 次（第 3 名），篇均被引频次为 11.56（第 2 名），整体表现位于上游水平。

表 1 WOS 发文量 Top20 机构及其论文表现

机构名称	国别	WOS论文数	被引频次	高被引论文	被引次数排名前 1% 的论文百分比	学科规范化的引文影响力	相对于全球平均水平的影响力
中国科学院	中国	394	4847	11	5.84	2.39	1.34
加利福尼亚大学系统	美国	336	3229	5	8.93	5.11	1.05
清华大学	中国	193	3285	6	8.29	3.48	1.85
伦敦大学	英国	185	1747	6	7.03	3.00	1.03
中国科学院大学	中国	181	2398	7	6.08	2.26	1.44
微软	美国	159	3605	5	13.21	6.39	2.47
印度理工学院	印度	153	1317	1	1.96	4.22	0.94
浙江大学	中国	146	1844	2	5.48	2.13	1.37
斯坦福大学	美国	144	2324	5	15.28	4.94	1.76
法国国家科学研究中心	法国	141	1830	2	4.96	2.12	1.41
哈佛大学	美国	138	2610	8	13.04	4.73	2.06
佛罗里达州立大学系统	美国	138	6845	6	9.42	9.21	5.40
新加坡国立大学	新加坡	136	1962	8	10.29	7.08	1.57
北京大学	中国	132	2052	1	5.3	3.06	1.69
得克萨斯大学系统	美国	124	1306	3	8.06	2.86	1.15
卡耐基梅隆大学	美国	123	1326	2	8.13	3.90	1.17
上海交通大学	中国	121	1202	0	4.96	1.88	1.08
南洋理工大学	新加坡	114	968	0	3.51	2.22	0.92
麻省理工学院	美国	109	1651	6	8.26	3.39	1.65
佐治亚大学系统	美国	104	7699	4	12.5	7.37	8.05

注：1. 数据背景颜色越深表示该机构的指标表现越优，表 2 同。

2. 中国科学院包含中国科学院大学、中国科学院各院所以及中国科学技术大学等，表 2 同。

表 2 WOS 发文量 Top20 中国机构及其论文表现

机构名称	WOS论文数	被引频次	高被引论文	被引次数排名前 1% 的论文百分比	学科规范化引文影响力	相对于全球平均水平的影响力
中国科学院	394	4847	11	5.84	2.39	1.34
清华大学	193	3285	6	8.29	3.48	1.85
中国科学院大学	181	2398	7	6.08	2.26	1.44
浙江大学	146	1844	2	5.48	2.13	1.37
北京大学	132	2052	1	5.3	3.06	1.69
上海交通大学	121	1202	0	4.96	1.88	1.08
中国科学院自动化研究所	100	1031	4	10	3.22	1.12
中山大学	99	1370	6	12.12	3.19	1.51
中国科学技术大学	92	1480	2	6.52	2.64	1.75
北京航空航天大学	82	856	0	1.22	1.48	1.14
哈尔滨工业大学	74	512	2	2.7	2.28	0.75
北京理工大学	71	812	1	5.63	1.71	1.24
阿里巴巴团队	71	658	1	2.82	2.26	1.01
武汉大学	70	690	1	8.57	2.69	1.07
电子科技大学	69	2221	6	13.04	5.04	3.50
北京邮电大学	67	338	1	4.48	1.65	0.55
复旦大学	59	710	2	5.08	2.45	1.31
合肥工业大学	58	468	1	3.45	1.74	0.88
腾讯	54	652	1	1.85	2.20	1.31
四川大学	53	594	4	9.43	3.06	1.22

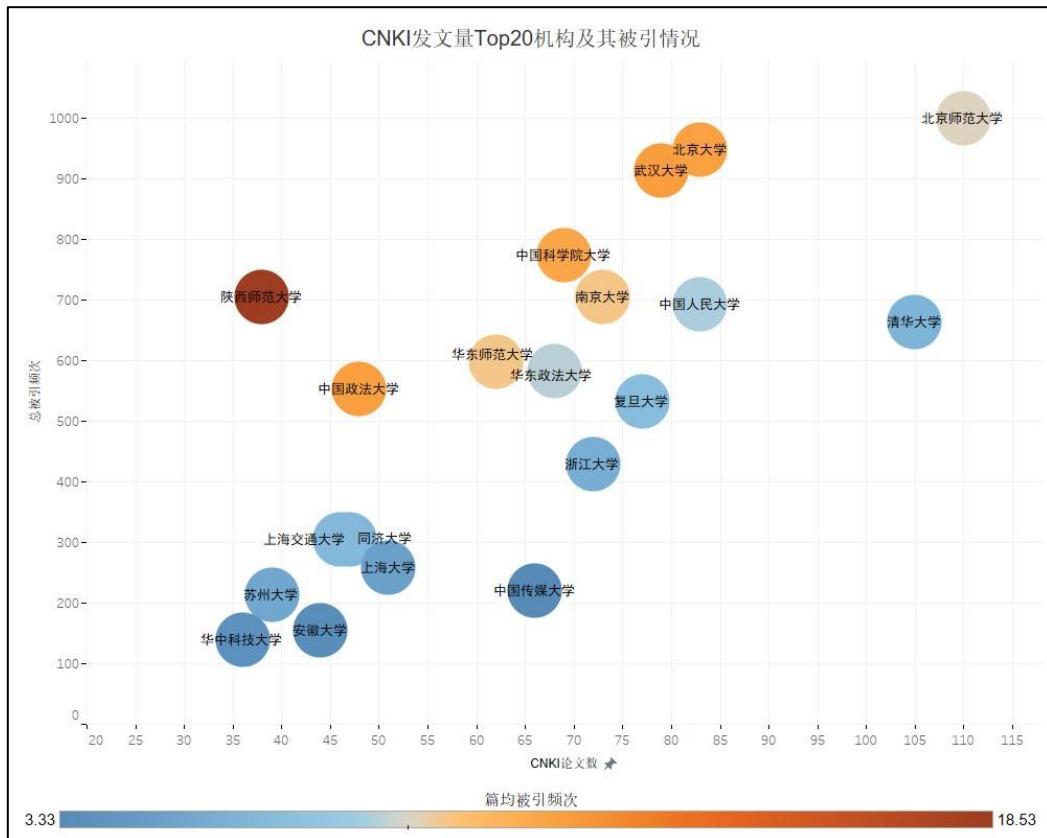


图 8 CNKI 发文量 Top20 机构及其被引情况

注：气泡色调越暖表示篇均被引频次越高。

五、高影响力文献

本部分选取了 2023 年以来部分高影响力中外文献，其中包含计算机科学、医疗、教育、金融、法律、出版、新闻等多个研究领域以供参考。

1. 外文论文

- [1] Dwivedi Y K, Kshetri N, Hughes L, et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy[J]. International Journal of Information Management, 2023,71.
- [2] Sallam M. ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns[J]. Healthcare, 2023,11(6).
- [3] Lee P, Bubeck S, Petro J. Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine[J]. New England Journal of Medicine, 2023,388(13):1233-1239.
- [4] Ayers J W, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum[J]. JAMA Internal Medicine, 2023,183(6):589-596.
- [5] Ji Z W, Lee N, Frieske R, et al. Survey of hallucination in natural language generation[J]. ACM Computing Surveys, 2023,55(12).
- [6] Cotton D, Cotton P A, Shipway J R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT[J]. Innovations in Education and Teaching International, 2024,61(2):228-239.
- [7] Casella M, Montomoli J, Bellini V, et al. Evaluating the feasibility of ChatGPT in healthcare: An analysis of multiple clinical and research scenarios[J]. Journal of Medical Systems, 2023,47(1).
- [8] Salvagno M, Taccone F S, Gerli A G. Can artificial intelligence help for scientific writing?[J]. Critical Care, 2023,27(1).
- [9] Haug C J, Drazen J M. Artificial intelligence and machine learning in clinical medicine, 2023[J]. New England Journal of Medicine, 2023,388(13):1201-1208.
- [10] Thirunavukarasu A J, Ting D, Elangovan K, et al. Large language models in medicine[J]. Nature Medicine, 2023,29(8):1930-1940.
- [11] Lu S Y, Liu M Z, Yin L R, et al. The multi-modal fusion in visual question answering: A review of attention mechanisms[J]. PeerJ Computer Science, 2023,9.
- [12] Lu S Y, Ding Y M, Liu M Z, et al. Multiscale feature extraction and fusion of image and text in VQA[J]. International Journal of Computational Intelligence Systems, 2023,16(1).
- [13] Khan R A, Jawaid M, Khan A R, et al. ChatGPT- Reshaping medical education and clinical management[J]. Pakistan Journal of Medical Sciences, 2023,39(2):605-607.

[14] Farrokhnia M, Banihashem S K, Noroozi O, et al. A SWOT analysis of ChatGPT: Implications for educational practice and research[J]. Innovations in Education and Teaching International, 2024,61(3):460-474.

[15] Cooper G. Examining science education in ChatGPT: An exploratory study of generative artificial intelligence[J]. Journal of Science Education and Technology, 2023,32(3):444-452.

[16] Lund B D, Wang T, Mannuru N R, et al. ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing[J]. Journal of the Association for Information Science and Technology, 2023,74(5):570-581.

[17] Yeo Y H, Samaan J S, Ng W H, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma[J]. Clinical and Molecular Hepatology, 2023,29(3):721-732.

[18] Guo M H, Liu Z N, Mu T J, et al. Beyond self-attention: External attention using two linear layers for visual tasks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023,45(5):5436-5447.

[19] Gao C A, Howard F M, Markov N S, et al. Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers[J]. npj Digital Medicine, 2023,6(1).

[20] Dowling M, Lucey B. ChatGPT for (Finance) research: The Bananarama Conjecture[J]. Finance Research Letters, 2023,53.

2. 中文论文

[1] 车万翔,窦志成,冯岩松,等. 大模型时代的自然语言处理:挑战、机遇与发展 [J]. 中国科学:信息科学, 2023, 53 (09): 1645-1687.

[2] 桑基韬,于剑. 从 ChatGPT 看 AI 未来趋势和挑战 [J]. 计算机研究与发展, 2023, 60 (06): 1191-1201.

[3] 卢经纬,郭超,戴星原,等. 问答 ChatGPT 之后:超大预训练模型的机遇和挑战 [J]. 自动化学报, 2023, 49 (04): 705-717.

[4] 孙凯丽,罗旭东,罗有容. 预训练语言模型的应用综述 [J]. 计算机科学, 2023, 50 (01): 176-184.

[5] 朱光辉,王喜文. ChatGPT 的运行模式、关键技术及未来图景 [J]. 新疆师范大学学报 (哲学社会科学版), 2023, 44 (04): 113-122.

[6] 李白杨,白云,詹希旎,等. 人工智能生成内容 (AIGC) 的技术特征与形态演进 [J]. 图书情报知识, 2023, 40 (01): 66-74.

[7] 卢宇,余京蕾,陈鹏鹤,等. 生成式人工智能的教育应用与展望——以 ChatGPT 系统为

例 [J]. 中国远程教育, 2023, 43 (04): 24-31+51.

[8] 张凌寒. 深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成型人工智能治理的中国路径 [J]. 法律科学(西北政法大学学报), 2023, 41 (03): 38-51.

[9] 方卿, 丁靖佳. 人工智能生成内容 (AIGC) 的三个出版学议题 [J]. 出版科学, 2023, 31 (02): 5-10.

[10] 喻国明, 李钒. 内容范式的革命: 生成式 AI 浪潮下内容生产的生态级演进 [J]. 新闻界, 2023, (07): 23-30.

纰误之处, 敬请批评指正! 同时, 我们面向全校师生征集关注的领域和专题, 欢迎提出宝贵建议。联系方式: 68754550, Email:guoying49@whu.edu.cn。

(编辑: 郭颖 审稿: 刘霞、刘颖)