聊天机器人与 ChatGPT：文献计量分析与系统综述——Web of Science 和 Scopus 数据库中的出版物

Hamed Khosravi[1,*], Mohammad Reza Shafie[2], Morteza Hajiabadi[3], Ahmed Shoyeb Raihan[4], Imtiaz Ahmed[5]

[1]西弗吉尼亚大学工业与管理系统工程系，摩根敦，WV 26505，hk00024@mix.wvu.edu

[2]伊朗科技大学电气工程系，德黑兰，
伊朗，Mr.shafie7731@gmail.com

[3]学院　计算机工程，伊朗科技大学，德黑兰　　，
伊朗，hajiabadi1377@gmail.com

[4]西弗吉尼亚大学工业与管理系统工程系，摩根敦，西弗吉尼亚州 26505，
ar00065@mix.wvu.edu

[5]工业与管理系统工程系，西弗吉尼亚大学，摩根敦，WV 26505，
imtiaz.ahmed@mail.wvu.edu

[*]通讯作者：hk00024@mix.wvu.edu

摘要

本文对与聊天机器人相关的科学文献进行了文献计量分析，并重点关注 ChatGPT。近年来，聊天机器人日益受到关注，在 Web of Sciences（WoS）和 Scopus 上的年增长率分别为 19.16% 和 27.19%。本研究通过分析 Scopus 和 Web of Science 数据库中的数据，探讨了该领域的结构、概念演变和发展趋势。研究分为两个阶段：（i）分析聊天机器人相关文献；（ii）全面综述有关 ChatGPT 的科学文献。第一阶段针对 Scopus（5839 条）和 WoS（2531 条）数据库中 1998 年至 2023 年期间发表的所有聊天机器人相关文献开展文献计量分析，包括期刊文章、书籍章节、会议论文和综述。对来源、国家、作者影响力和关键词的深入分析表明，ChatGPT 是聊天机器人领域的最新趋势。

因此，在第二阶段，我们对 ChatGPT 相关出版物进行了文献计量分析，并根据研究方法、创新性和结论，对 45 项已发表研究进行了深入分析。研究确定的关键兴趣领域可分为三组：人工智能及相关技术、对话代理的设计与评估，以及数字技术与心理健康。总体而言，本研究旨在为研究人员提供指导，帮助他们更有效地开展聊天机器人领域的研究，并特别指出 ChatGPT 未来研究的重要方向。

关键词：聊天机器人、ChatGPT、文献计量学、人工智能、自然语言处理、生成式人工智能

1. 引言

人工智能（AI）是一项革命性技术，能够从数据中学习并执行通常需要人类智能才能完成的任务（McCarthy, 2007）。凭借构建智能系统的能力，人工智能已应用于多个领域（Amara, Hadj Taieb, and Ben Aouicha, 2021）。生成式人工智能（GAI）是人工智能的一种，它会创作新内容，而不是识别或分类已有内容（Jovanovic & Campbell, 2022）。近年来，深度学习这一人工智能的新兴分支不断发展，涵盖了多种网络结构。如今，深度学习已应用于多个领域，包括计算机视觉、语音识别和自然语言处理（NLP）（Chai et al., 2021）。自然语言处理是生成式人工智能的重要组成部分，它使机器能够理解并生成人类语言（Aydin & Erdem, 2022）。在这项技术中，词语被拆解为更小的组成部分，对其关系进行分析，并探索这些组成部分的组合方式，以形成有意义的模式（IBM, 2017; Khodadadi, Ghandiparsi, and Chuah, 2022）。

聊天机器人是模拟人类对话并利用自然语言处理（NLP）自动回答用户询问的人工智能应用（Dwivedi 等，2021）。近年来，聊天机器人日益普及，因为它们能够改善客户服务、缩短响应时间并自动处理重复性任务（Dwivedi 等，2021）。除客户服务外，它们还用于多种用途，例如教育、心理健康支持和财务管理（Ashfaq 等，2020）。值得注意的是，聊天机器人能够全天候提供客户支持，从而提升客户满意度。此外，它们还能同时处理多个对话，进而提高效率（Adamopoulou & Moussiades，2020a）。Siri、Alexa 和 Google Assistant 等聊天机器人是其中最受欢迎的几种（Adamopoulou & Moussiades，2020b）。图 1 按发布日期列出了若干知名聊天机器人。

ChatGPT 是近年来涌现的一款知名聊天机器人。它是一种基于人类反馈强化学习的语言模型，旨在生成对话式文本（Thorp，2023年）。随着时间推移，ChatGPT 凭借其出色的能力，能够围绕广泛主题生成连贯且逼真的回答，因而广受认可和欢迎（Lund & Wang，2023）。这项技术对科学和社会的影响不容忽视，因为 ChatGPT 有潜力影响多个行业和领域（van Dis 等，2023年）。

本文旨在分析 Scopus 和 Web of Science 收录的、涉及聊天机器人和 ChatGPT 主题的所有论文的元数据。本研究范围广泛，面向多元且包容的读者群体，其中包括来自计算机编程、教育、研究、医疗保健及其他相关领域、对聊天机器人和 ChatGPT 有切身利益的利益相关者。作为本研究的一部分，我们使用“bibliometrix”R 软件包（Aria & Cuccurullo，2017年）以及 VOSviewer 软件（van Eck & Waltman，2009）系统分析了以 ChatGPT 为重点的聊天机器人文献；VOSviewer 是一款用于可视化和分析书目信息的软件工具。文献计量学是一门利用书目数据评估并深入了解特定研究领域结构的学科（Keramatfar、Rafiee 和 Amirkhani, 2022）。本文还对重点关注……的领先国家提供了有价值的见解。

关于聊天机器人和 ChatGPT 的研究论文出版情况，以及这些出版物所探讨主题的发展趋势。此外，我们还探讨了分析的其他方面，例如来源、作者和关键词分析。

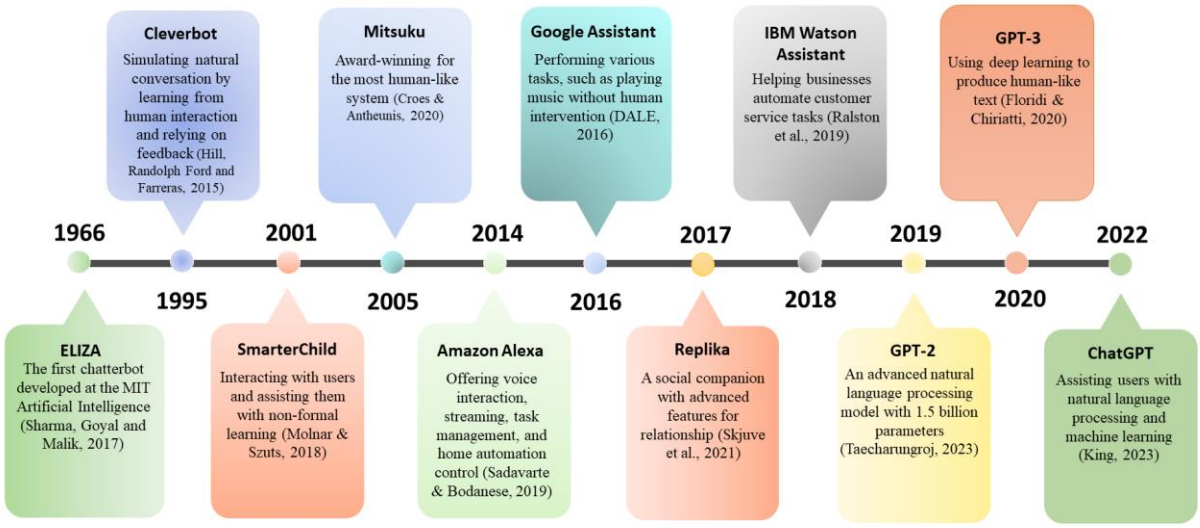


图1. 随时间推移的知名聊天机器人概览

本文其余部分的组织结构如下。第2节概述本研究所采用的方法。第3节介绍我们对出版物进行文献计量分析和综述的结果。第4节讨论我们研究结果中的关键发现。最后，第5节给出本研究的主要结论。

2. 方法论

在科学数据库方面，我们选择了Scopus 和Webof Science，因为两者都收录了计算机科学、统计学、工程学和数学领域最重要的期刊（Martarelli&Nagano,2021）。在我们的文献综述中，我们审阅了广泛的学术来源，包括书籍章节、文章、综述及其他相关出版物。为了确保综述全面，并涵盖所有相关的学术和科学文献，我们没有将检索范围限制在特定来源或出版物类型上。通过采用这种方法，我们能够对该领域现有的研究和知识进行全面评估，从而对这一主题形成全面且细致的理解。总体而言，研究分为两个阶段：

I. 首先，对有关聊天机器人的出版物进行了文献计量分析。分析包括对国家、来源、作者和关键词的评估。研究考察了各国的出版物数量、国家间合作、来源的布拉德福德定律、作者影响力、主题演化、共现分析以及关键词趋势主题。

II. 其次，对聊天机器人领域最新趋势 ChatGPT 进行了文献计量分析。通过共现分析和词云，对该领域使用的关键词进行了审视。为了获得更全面的认识，研究综述了 45 篇有关 ChatGPT 的出版物，其中包括 15 篇期刊文章，并根据所采用的方法、研究发现的创新性以及所得结论对其进行了评估。方法概述见图 2。

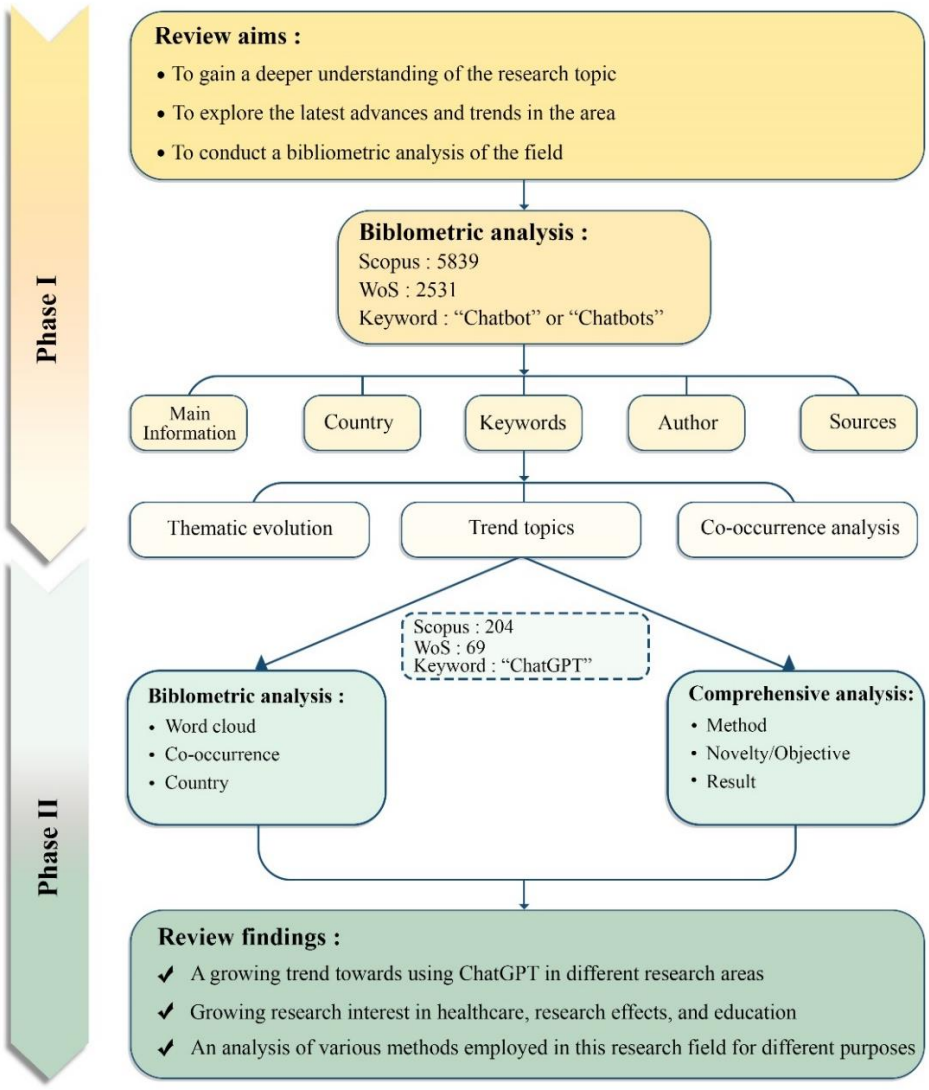


图 2. 研究所采用的方法框架

3. 结果

文献计量分析使用两种工具进行，即 Vosviewer 软件和 R 编程语言中的 bibliometrix（biblioshiny）。本节分两个独立阶段呈现分析结果。

3.1. 第一阶段

本节对有关聊天机器人的科学出版物进行文献计量分析。分析内容包括与这些出版物相关的国家、来源、作者和关键词信息。首先，我们概述这些出版物，并介绍导出文献的详细信息。

3.1.1. 主要信息

如前所述，本研究使用 Scopus 和 Web of Science 数据库开展。导出用于分析的出版物的主要信息见表1。

表1：导出出版物的主要信息

特征	**WoS**	Scopus
时间跨度	1998:2023年	2000:2023年
文献	2531	5839
来源（期刊、书籍等）	1512	4813
作者	7693	14908
期刊文章	1000	1824
书籍章节	16	112
综述	81	179
年增长率 %	19.16	27.19
篇均被引次数	7.567	8.302

根据表1所呈现的数据，可以明显看出，与 Web of Science 数据库相比，Scopus 数据库拥有更多的出版物、作者和书籍章节，以及更高的平均被引次数和年增长率。此外，我们还分析了两个数据库共同发表的文献数量，发现有1946篇出版物同时收录于两个数据库中。因此，分析的文献总数为6424篇。需要注意的是，两个数据库是分别进行分析的，这可能会提供更具实践意义的见解。

我们计算了包含聊天机器人相关研究的 Scopus 和 Web of Science 数据库的传统重叠率（TO）和相对重叠率（RO）。对于数据库 X 和 Y，传统重叠率（TO）定义为这两个数据库之间的重叠百分比（Gluck，1990年）。其计算公式如下：

$$\%TO = 100 \times \frac{|X \cap Y|}{|X \cup Y|}$$

相对重叠率（RO）用于计算数据库 X 相对于另一个数据库 Y 的覆盖重叠部分或覆盖率，反之亦然（Garg、Kumar 和 Singh，2020年）。对于任意两个数据库 X 和 Y，相对重叠率（RO）的计算公式如下。

$$\%Overlap \in X = 100 \times \frac{|X \cap Y|}{|X|}$$

$$\%Overlap \in B = 100 \times \frac{|X \cap Y|}{|Y|}$$

如表1所示，分别从 Web of Science 和 Scopus 数据库中获取了总计2531篇和5839篇文献。其中，6424篇文献出现在 Web of Science 或 Scopus 中，1946篇文献同时出现在这两个数据库中。这两个数据库中文献的传统重叠率（TO）为76.887，计算如下。

$$\%TO = 100 \times \frac{|WoS \cap Scopus|}{|WoS \cup Scopus|} = 100 \times \frac{1946}{6424} = 30.293\%$$

这表明，在 Web of Science 与 Scopus 数据库中，聊天机器人相关研究的相似度为30.293%。同样，Scopus 数据库中的相对重叠率为 33.328%，而 Web of Science 数据库中的相对重叠率为 76.887%。这些数据通过以下计算得出。

$$\%Overlap \in Scopus = 100 \times \frac{|WoS \cap Scopus|}{|Scopus|} = 100 \times \frac{1946}{5839} = 33.328\%$$

$$\%Overlap \in WoS = 100 \times \frac{|WoS \cap Scopus|}{|WoS|} = 100 \times \frac{1946}{2531} = 76.887\%$$

这些结果表明，Web of Science 收录了 Scopus 中 33.328% 的文献，而 Scopus 收录了 Web of Science 中 76.887% 的文献。

3.1.2. 国家分析

截至2023年3月底，Web of Science 数据库收录了来自92个国家/地区的2531篇有关聊天机器人的文献。相比之下，截至同一日期，Scopus 收录了来自116个国家的5839篇与聊天机器人相关的文献。这些发现表明，聊天机器人研究是一个快速发展的领域，在全球范围内具有重要影响力。美国在这一领域遥遥领先，在两个数据库中都拥有最多的出版物（Scopus 和 Web of Science 中分别为897篇和1425篇）。在 Web of Science 中，美国与其他国家的合作最多。英国、德国和意大利也是主要参与者，在 Scopus 中组成了规模最大的合作聚类。

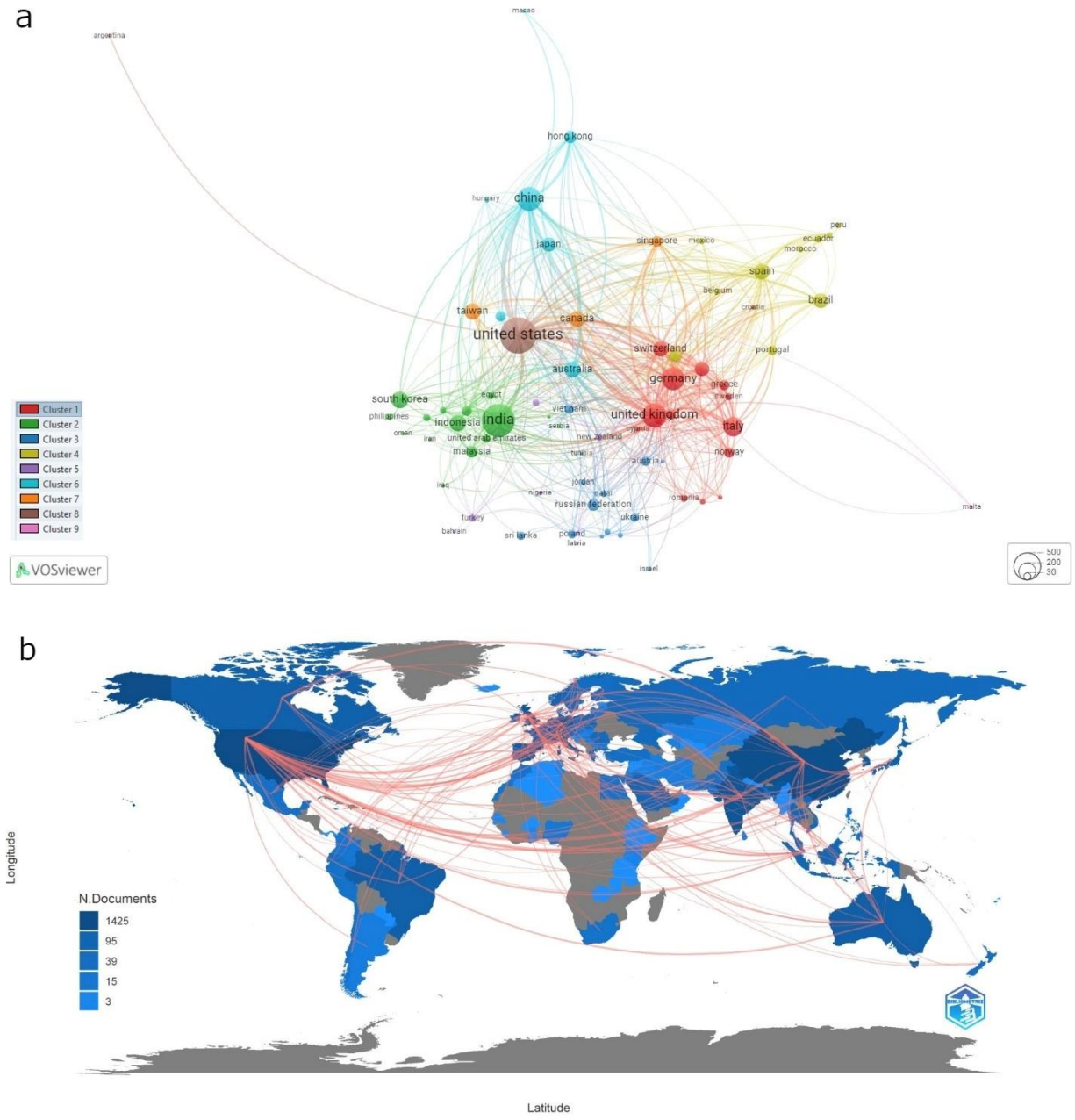


图 3. Scopus 上的国家合作网络（a）和 WoS 上各国出版物的地理分布（b

3.1.3. 来源（布拉德福定律）

根据布拉德福定律（1934年），某一主题的论文数量按照数学函数分布（散布），因此，期刊/信息源的数量必须增加，才能容纳该主题论文数量的增长。产出数量大致相等的期刊组之间的比例约为$1: n: n^2$ ...，其中 n 是布拉德福乘数（Sangam，2015年）。 任何问题或主题领域的前三分之一区域（第1区或核心区），代表该主题文献中被引用最频繁的期刊，因此，这些期刊很可能具有

研究者最感兴趣的内容。中间三分之一（第2区）由平均被引用的期刊组成，而底部三分之一（第3区或尾部）由鲜少被引用、被认为重要性较低的期刊组成（Potter，2010年）。布拉德福德分析有多种潜在应用，包括期刊的遴选与剔除、识别相关出版物的核心集合，以及评估数据收集情况（Nisonger，2008年）。此外，它还可用于预测不完整检索可能遗漏的相关信息范围（Nash-Stewart 等人，2012年）。

3.1.3.1. 语言表述

为检验布拉德福德定律的文字表述，将被引文献按照已发表文献数量降序排列。表2列出了每个来源的排名，以及其对应的文献数量和累计文献数量。32个来源发表了1914篇科学文献，随后是525种期刊中的1972篇文章。其余文献分布在1732个来源中，这意味着每个期刊类别均涵盖总计文献的三分之一。在当前设置下，布拉德福德分布中各区域之间的关系为$32:525:1732$，其中倍数值为54.13。

根据表2，布拉德福定律的计算如下：

$$32:32*54.13:32*(54.13)^2 :: 1:n:n^2$$
$$32:1732.16:93761.82$$

因此，百分比误差表示为：

$$\frac{95525.98-2289}{2289}\times 100 = 40.73\times 10^2$$

表2：各布拉德福德分区的期刊分布与引文数量

区域	数量 of 来源	百分比 来源（%）	文献数量 文献	文献累计数量 文献	布拉德福倍数
1	32	1.40	1914	1914	1
2	525	22.94	1972	3886	16.41
3	1732	75.66	1916	5802	54.13
总计	**2289**	**100**	**5802**		

因此，由于误差百分比较高，无法将数据拟合到布拉德福德定律中。 因此，有必要使用Leimkuhler模型（LEIMKUHLER，1980年）评估数据对布拉德福德分散定律的遵循程度，该模型已被多项先前研究采用（Bhargav、Kishore和Doraswamy，2020年；Qiu等，2017年）。

3.1.3.2. Leimkuhler模型的应用

Leimkuhler模型采用布拉德福分布，首先识别第一个区域中具有不同参考文献的核心期刊。这有助于确定布拉德福倍数（k）。作为一种大小—频率测量方法，Leimkuhler模型随后使用布拉德福倍数的倍数来识别后续区域中的期刊数量。Egghe的数学技术（Egghe，1990年）用于计算Leimkuhler分布的倍数（k）。

Leimkuhler模型根据Bradford的文字表述定义如下（Borgohain等，2021年）。

$$r_0 = \frac{T(k-1)}{(k^p-1)} \tag{1}$$

其中，T 表示期刊总数，k 表示布拉德福倍数，p 表示区域数。在本研究中，文章被划分为三个区域，因此，实现布拉德福德定律的数学表达式可表示如下。

$$k = (e^y \times y_m)^{\frac{1}{p}}$$

其中，y_m 是产出最多的来源中的项目数量。因此，布拉德福的倍数 k 计算如下，

$$y_m = 421$$

$$p = 3$$

$$k = (1.781 \times 421)^{\frac{1}{3}} = 9.08$$

此外，

$$y_0 = \frac{A}{p} = \frac{5802}{3} = 1934$$

其中，A 是文章总数。现在，r_0 按以下方式计算。

$$r_0 = \frac{T(k-1)}{(k^p-1)} = \frac{2289(9.08-1)}{(9.08^3-1)} = 24.97$$

利用 $log（k）$ 的值，我们可以得到以下结果。

$$a = \frac{y_0}{log(k)} = \frac{1934}{log(9.08)} = 2018.61$$

$$b = \frac{k-1}{r_0} = \frac{9.08-1}{24.97} = 0.32$$

因此，

$$r_1 = r_0 \times k = 24.97 \times 9.08 \simeq 225$$

$$r_2 = r_0 \times k^2 = 24.97 \times 9.08^2 \simeq 2025$$

因此，核心期刊的数量为r_0，布拉德福乘数的平均值为k 9.08。因此，Leimkuhler模型中的布拉德福德分布可以表示为：

$$r_0 : r_0 * k : r_0 * k^2$$

因此，我们得到：

$$25 : 25 * 9.08 : 25 * (9.08)^2 :: 1 : n : n^2$$
$$25 : 227 : 2058$$

最后，百分比误差计算如下：

$$\frac{2310 - 2289}{2289} \times 100 = 0.9$$

表3：布拉德福德分布的Leimkuhler模型（分为3个区域）

区域	数量 来源	百分比 of 来源（%）	数量 文献	累计编号 文献数量	布拉德福 倍数
1	25	1.08	1807	1807	1
2	227	9.83	1365	3172	9.08
3	2058	89.09	2630	5802	82.45
总计	**2310**	**100**	**5802**		

计算结果显示，百分比误差极小，仅为0.9%，可视为可忽略不计。表3呈现了使用Leimkuhler模型评估布拉德福分散的分析结果。表中显示，代表三个区域中文献的期刊数量约增加了9倍。对数据进行分区分析后发现，第1区包括25个来源和1807篇文献，第2区包括227个来源和1365篇文献，第3区包括2058个来源和2630篇文献。

3.1.4. 作者

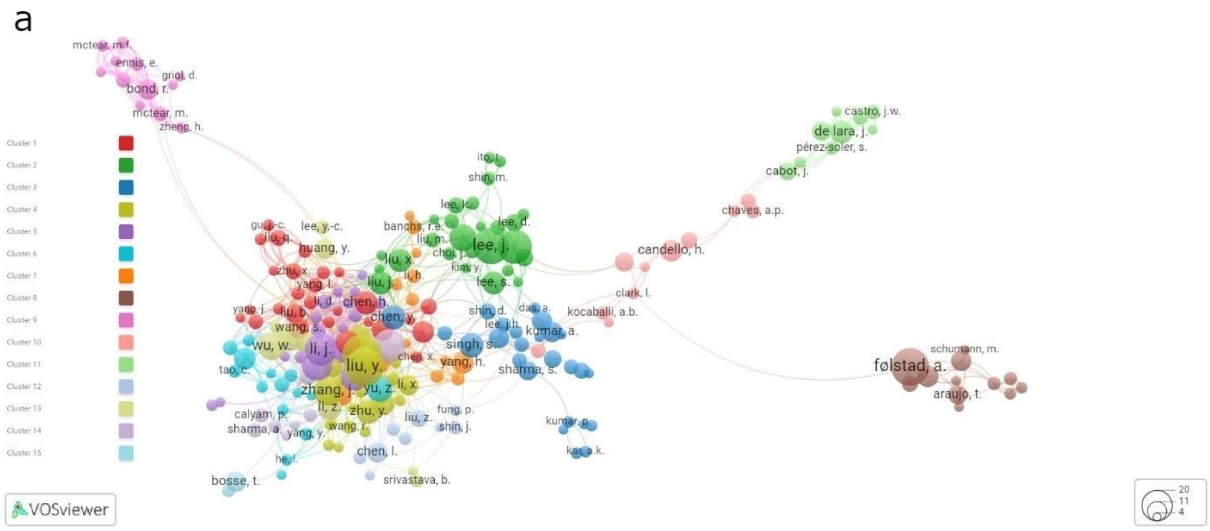
本节根据总引文数对聊天机器人研究领域排名前10位的作者进行排序，并列出这些作者在该领域的出版物数量（N）、h指数、g指数、m指数以及首次发表年份（YS）。随后，展示两个数据库中作者之间的合作情况。

表4：基于总引文数的引用影响力排名前10位的作者

Scopus							**WOS**						
作者	**TC**	h 指数	g_索引	m 索引 N		**YS**	作者	TC h	指数	g_指数	m 指数 N		**YS**
Gao J	**1344**	6	8	0.75	8	2016	Folstad A	**380**	**7**	**13**	**0.395**	**13**	2017
Zhou M	1178	9	13	0.529	13	2007	Sundar SS	375	3	4	0.375	4	**2016**
Littman ML	1170	1	1	0.048	1	**2003**	Li D	345	2	3	0.333	3	2018
Turney PD	1170	1	1	0.048	1	**2003**	Singh D	312	2	3	0.286	3	2017
Galley M	1049	4	4	0.5	4	2016	Batra D	311	2	2	0.286	2	2017

Følstad A	1042	**13**	**25**	**1.857**	**25**	2017	Das A	311	2	2	0.286	2	2017
Wu W	979	11	18	1.375	18	2016	Gupta K	311	2	2	0.286	2	2017
Wu Y	896	10	16	0.625	16	2008	Kottur S	311	2	2	0.286	2	2017
Xing C	803	6	6	0.75	6	2016	Moura JMF	311	2	2	0.286	2	2017
Brandtzaeg PB	771	9	12	1.286	12	2017	Parikh D	311	2	2	0.286	2	2017

根据表4，Folstad A 在 WOS 数据库中的引文数最多，在 Scopus 数据库中则是引文数排名第六的作者。值得注意的是，Folstad A 是唯一一位同时出现在两个数据库引文数最高的前10位作者名单中的作者。此外，Folstad A 在两个数据库中均拥有最高的 h 指数、g 指数和 m 指数，进一步凸显了他对这一研究领域的重要影响。这些文献计量指标表明，Folstad A 对该领域作出了显著贡献，并受到学界其他研究人员的高度认可。数据还显示，Gao J 是 Scopus 中引文数最多的作者，总引文数为1344。尽管只发表了8篇论文，Gao J 的 h 指数仍为6，这意味着他至少有6篇论文各自被引用了6次或以上。图4通过网络分析，识别了两个数据库中的核心作者及其合作网络。分析发现了15个聚类，其中纳入研究的作者近350位（Scopus）和250位（WOS）。



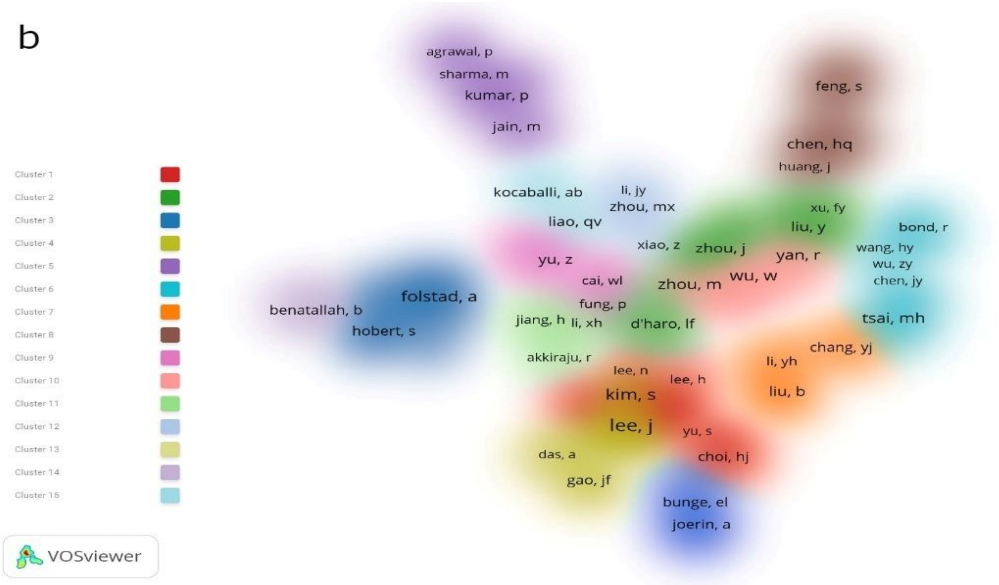


图4. Scopus（a）和 WoS（b）上的作者合作网络

图4采用基于文献权重的关联强度法，分析了 Scopus 和 WoS 数据库中作者之间的合作模式。15个聚类按各组作者人数排序。特定作者（如 Lee J、Kim S 和 Lee D）之间的合作作出了重要贡献（分别对应 Scopus 和 WoS 中的聚类2和聚类4）。这些聚类在 Scopus 中包含2至27位作者，在 WoS 中包含5至16位作者。总体而言，研究结果表明，不同数据库中的作者合作模式存在差异。此外，该图证实了作者 Folstad A 的影响力。

3.1.5. 关键词

研究主题的演变及其相互关系如图5所示，分为三个分析时期。线条的特征表明了关键词之间关系的质量。本研究将 Scopus 的二十三年和 WoS 的二十五年时间段分别划分为四个和五个时期，即 Scopus 的

[2000 2003], [2004 2014], [2015 2017], - - - -, 以 [2018 2023] [1998 2006], [2007 2014], [2015 2017], [2018 2020], [2020 及 WoS 出版物的 - - - - 和 -2023] 。

图5展示了该研究领域中研究主题之间的强互联性。该图表明，研究主题与不同的研究领域相连接。在2015年-2017年期间，深度学习和机器学习成为该领域的重要主题。此外，近期研究表明，自然语言处理（NLP）和心理健康等主题在该研究领域受到了广泛关注。图5b突出显示了近期的信任与任务分析研究领域，表明它们在该领域中的重要性日益提升。这些发现揭示了新兴研究领域的重要性，并为研究主题的结构与演变提供了有价值的信息。接下来，图6展示了相关研究领域中关键词的共现网络。本分析采用的聚类算法为 Walktrap，选取50个关键词进行聚类，并通过关联测度进行归一化处理。

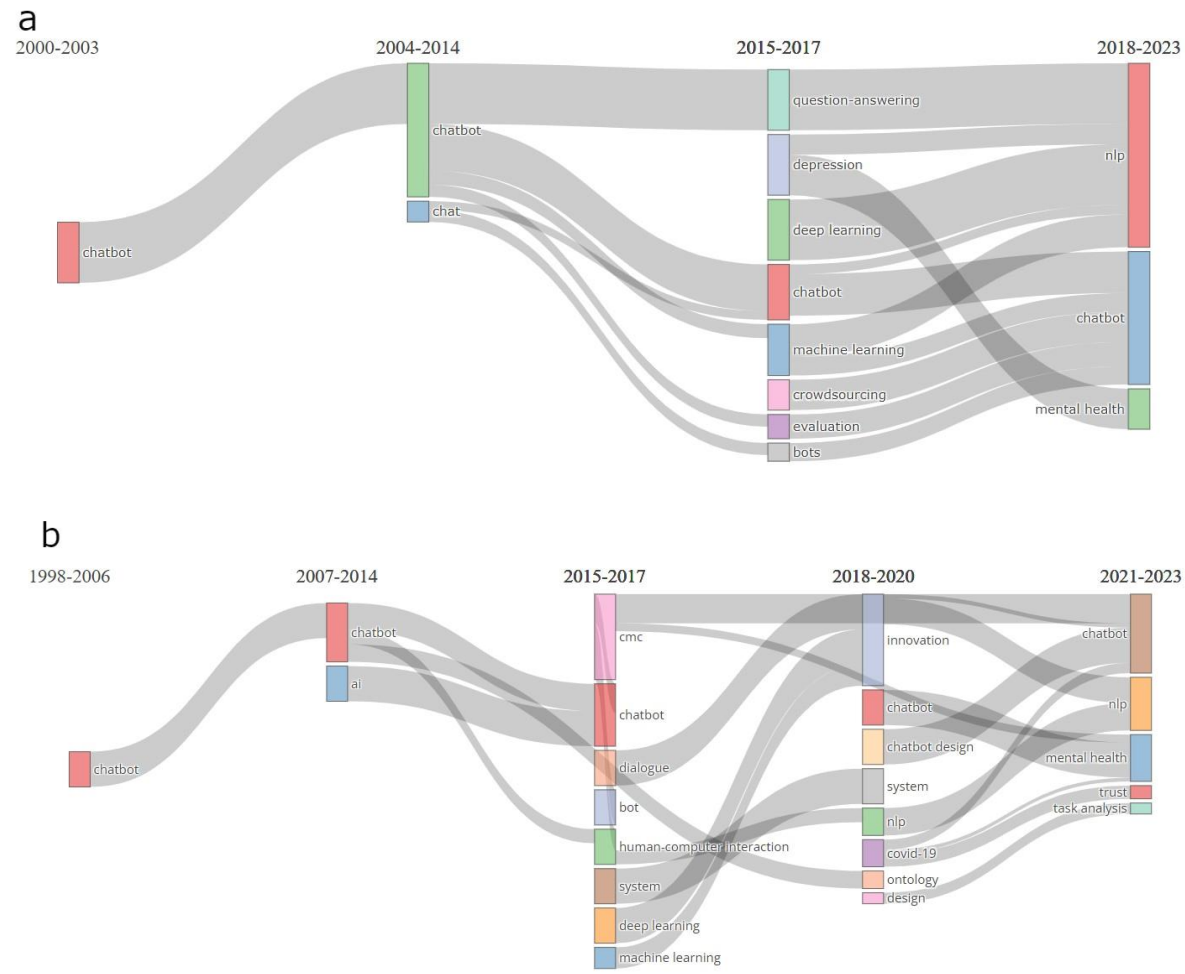
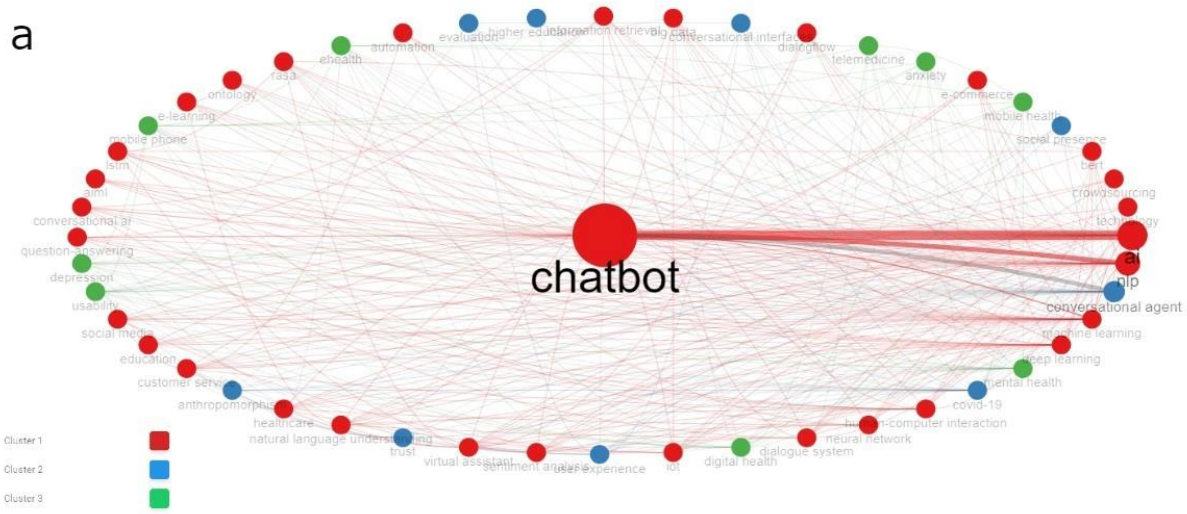


图 5. Scopus（a）和 Web of Science（b）上的关键词主题演化



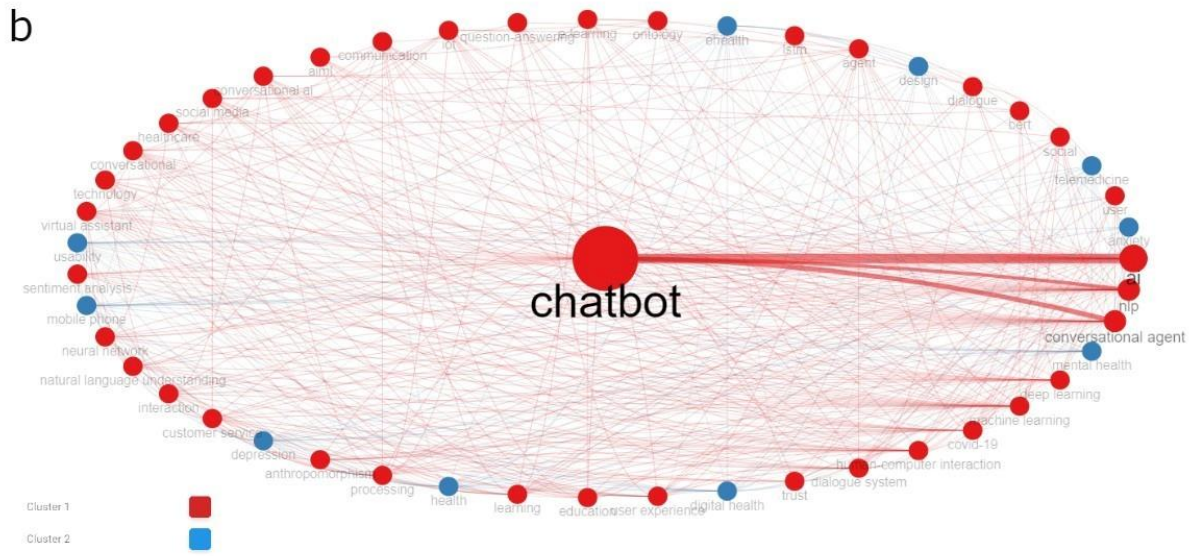


图 6. Scopus (a) 和 Web of Science (b) 上的关键词共现网络

根据图6，基于 Scopus 收录的出版物确定了三个不同的关键词聚类，分别命名为“人工智能及相关技术”“对话代理的设计与评估”以及“数字技术与心理健康”。另一方面，从 Web of Science（WoS）提取的关键词可分为两类：“基于人工智能的聊天机器人和对话代理”以及“数字技术与心理健康”。值得注意的是，Scopus 中的聚类1和聚类2在 WoS 中合并为一个聚类。这些发现表明，这些领域正在开展大量研究，并且在人工智能、数字技术与健康的交叉领域开展更多跨学科研究具有潜力。表5和表6分别列出了 Scopus 和 WoS 数据库中每个聚类的前5个关键词。这为了解各聚类研究的主要议题和主题提供了有价值的见解。

表5：Scopus中每个聚类的前5个关键词

关键词	聚类	中介中心性	接近中心性
聊天机器人	1	663.1045	0.020833
ai	1	65.05425	0.020833
nlp	1	25.86692	0.017544
机器学习	1	8.302174	0.016949
深度学习	1	2.358797	0.014085
人机交互	1	0.621563	0.013158
对话代理	2	27.84586	0.018868
新冠肺炎疫情	2	2.459913	0.014925
用户体验	2	0.324901	0.012987
信任	2	0.066949	0.011628
拟人化	2	0.105056	0.011494

心理健康	3	4.926437	0.014286
数字健康	3	1.037543	0.013158
可用性	3	0.295981	0.012821
抑郁症	3	0.209681	0.012346
手机	3	0.105715	0.012195

表6：WOS各聚类中的前5个关键词

关键词	聚类	中介中心性	接近中心性
聊天机器人	1	572.2521	0.021277
ai	1	38.31512	0.02
nlp	1	22.55187	0.019231
对话代理	1	20.88957	0.018519
深度学习	1	2.363443	0.014286
心理健康	2	2.804667	0.014085
数字健康	2	1.222346	0.014493
健康	2	0.489457	0.014286
抑郁症	2	0.582268	0.013514
手机	2	0.587108	0.013889

如表5和表6所示，其中列出了各关键词的中介中心性和接近中心性。在网络中，中介中心性指一个节点位于其他两个节点（关键词）之间最短距离上的频率。中介中心性较高的节点在连接网络方面发挥着关键作用，对网络的整体运行至关重要。此外，接近中心性指标反映了网络中最短路径的平均长度；节点越接近网络中的其他节点，其在网络中的影响力通常越大，也越容易被其他节点触达（Makkizadeh和Sa'adat，2017年）。 在两个数据库中，聊天机器人的中介中心性和接近中心性均最高，其次是人工智能。对话代理在 Scopus 的聚类2中也具有较高中介中心性，但在 WoS 中属于聚类1，这表明它在 Scopus 中比在 WoS 中是更重要的关键词。这可能是因为 Scopus 涵盖的学科范围比 WoS 更广。在分析的下一部分，我们研究了聊天机器人领域近期最受关注的主题，以帮助研究人员和开发者及时了解该领域的最新进展。这也有助于他们发现新的机会并规避潜在风险。

根据图7，在两个数据库中都观察到，ChatGPT似乎是聊天机器人领域的最新趋势。本分析将关键词的最低出现频次设为5，并将每年的关键词数量设为3。因此，本文第二阶段对这一主题进行了更详细的探讨，包括对已发表文献的全面综述和文献计量分析。

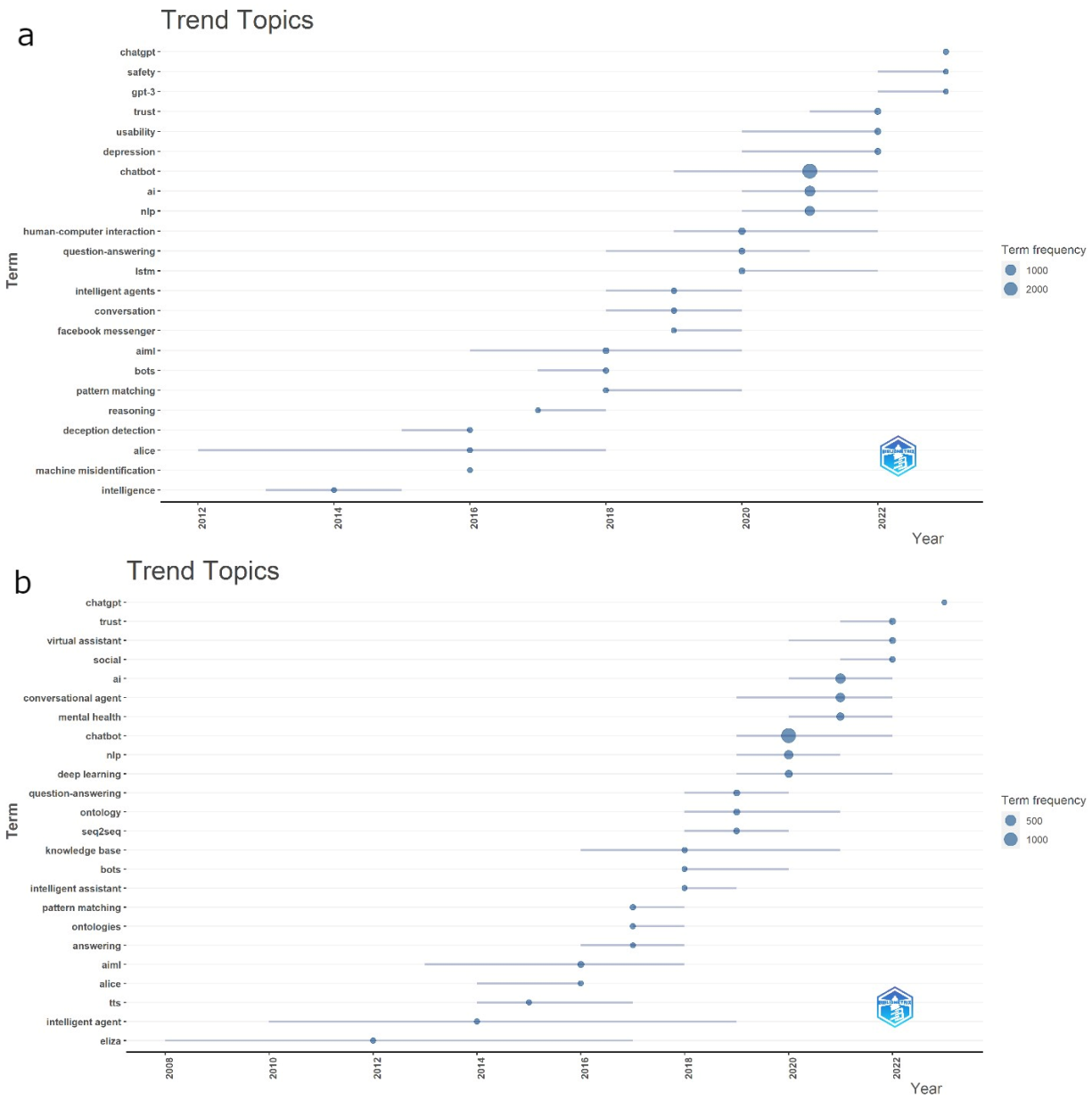


图 7. 基于 Scopus (a) 和 WOS (b) 研究领域关键词的趋势主题

3.2. 第二阶段

在这一阶段，我们全面分析了Scopus和Web of Science数据库中的ChatGPT科学出版物。截至2023年4月，纳入了Scopus中的204篇出版物和Web of Science中的69篇出版物。其中，62条记录同时出现在两个数据库中。因此，我们的分析共纳入211篇科学出版物。分析首先进行初步的文献计量评估，随后对选出的45篇出版物进行详细考察。

3.2.1. 文献计量分析

本节分析了211篇关于ChatGPT的出版物其中62篇被归类为期刊文章。每篇文献的平均被引次数为0.98次，共有644位作者参与了这些出版物的撰写。这些出版物包含334个关键词，其中78篇出版物为单作者文献。本研究考察的文献引用了约5000个来源。首先，关于单国和多国贡献的出版物分布如图8所示。

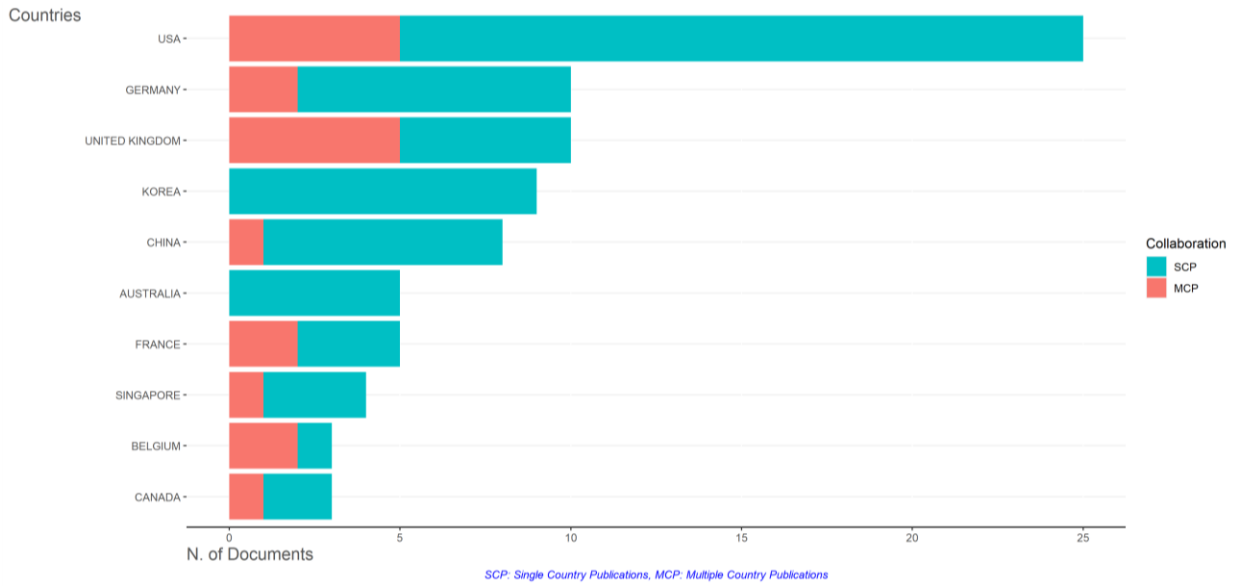


图 8. 文献数量最多的前 10 个国家

根据图 8，美国拥有最多的文献，共有 25 篇出版物。此外，美国和英国来自单个国家的出版物数量最高，各有 5 篇文献。为考察不同来源国家文献的引文模式，图 9 展示了各国获得的引文总数。

如图 9 所示，英国被引用次数最多的文献有 17 篇，平均每篇期刊文章的被引次数为 1.7。该比率高于美国，美国的出版物数量最多（20 篇），平均被引次数为 0.64。爱尔兰的平均被引次数最高，为 3。接下来将分析这一领域的关键词，图 8 展示了关键词词云。词云以可视化方式呈现文本中各词语出现的频率。词语越大，表示其在文本中出现得越频繁。

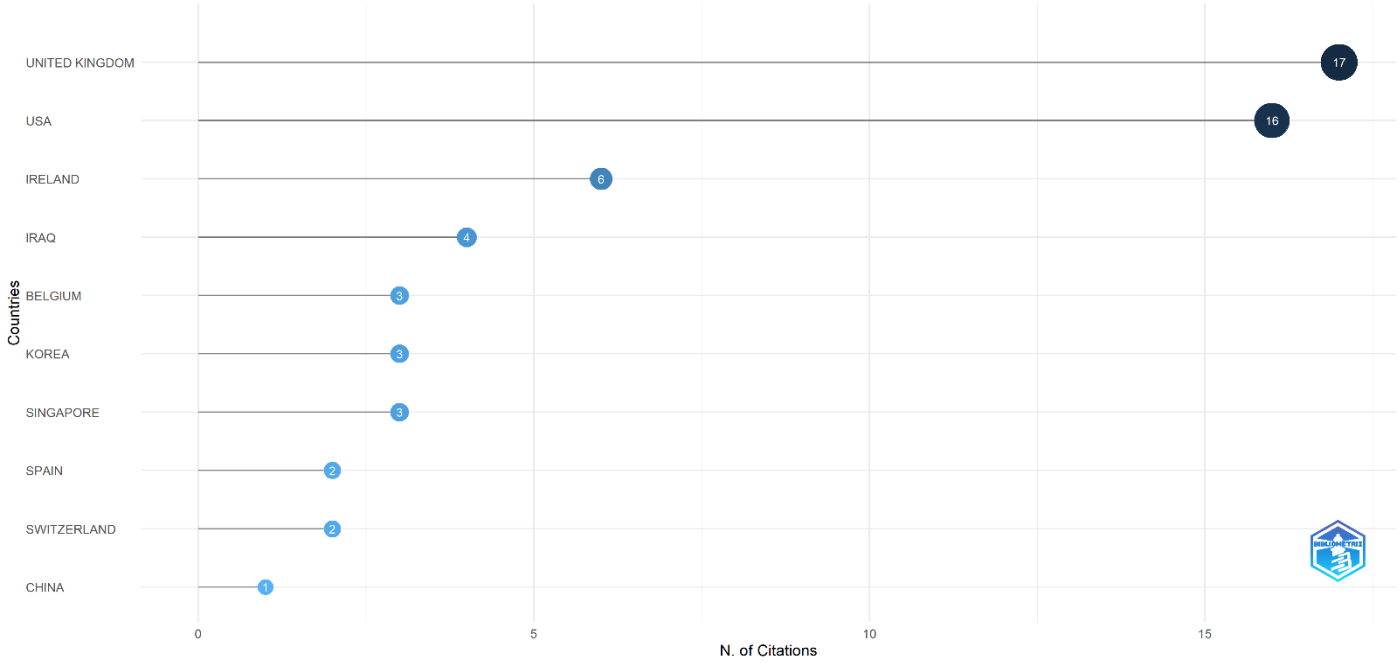


图 9. 引文数量最多的前 10 个国家

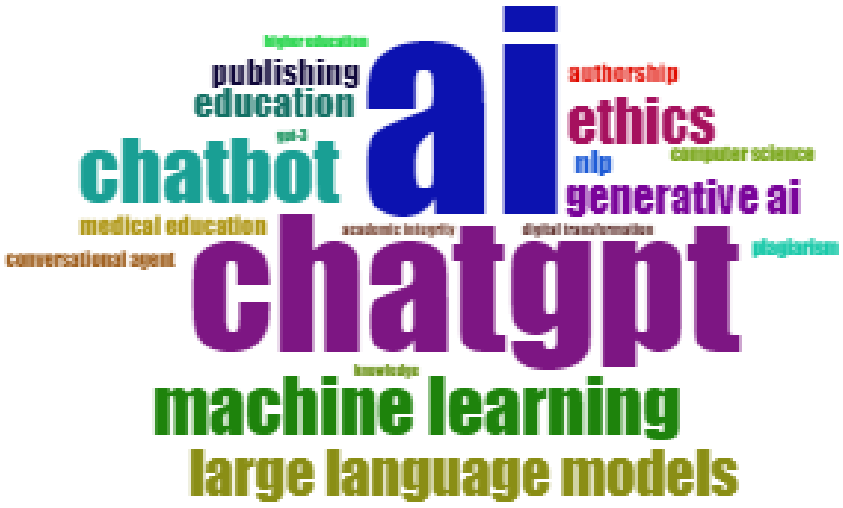


图 10. ChatGPT 相关文献关键词的词云

图 8 中的词云展示了该领域最常用的关键词。最突出的是“人工智能”，其次是“ChatGPT”和“聊天机器人”。这些词分别被用作关键词 55 次、36 次和 17 次。“出版”“剽窃”和“作者身份”表明，研究伦理是这一领域的重要关注点之一。为进一步分析关键词，图 11 展示了关键词的共现情况。

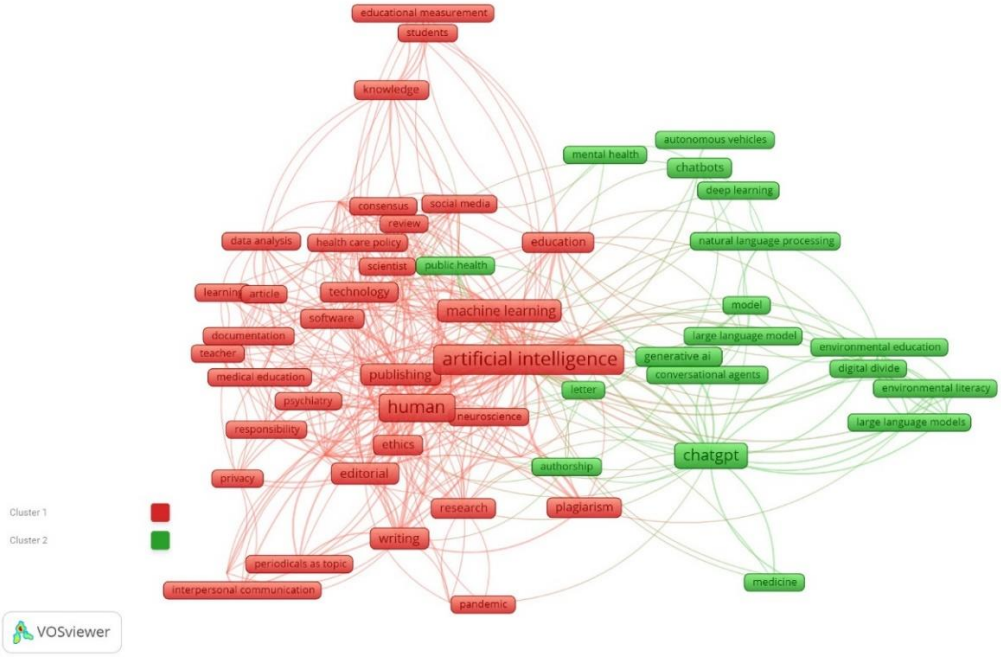


图 11. 关于 ChatGPT 的文献关键词共现

根据图 11，两个聚类聚焦于语言模型、生成式人工智能和对话代理的不同方面。具体而言，聚类 1 强调这些技术在教育和写作领域的应用，而聚类 2 则侧重于它们在医学和医疗保健领域的应用。不过，必须指出，这两个聚类并非互不相关，因为语言模型和对话代理都可以应用于这两个领域。此外，共享关键词“人工智能”表明，这两个聚类都关注在不同领域应用人工智能，因此彼此相关。下一节将对归入这两个聚类的部分论文进行更详细的分析。

3.2.2. 综合综述

本节全面分析了 45 篇已发表文献，其中包括 15 篇关于 ChatGPT 的论文。首先，表 7 概述了根据引用次数、阅读次数及其领域重要性选出的 15 篇期刊文章。随后，将综述其他来源的出版物。

3.2.2.1. 文章

表7：关于 ChatGPT 的 15 篇文章总结

作者	国家	研究领域	创新性/目标	方法	结果
（Nachshon 等，2023年）	ISR	医学	讨论了一例100% 三度烧伤并调查了ChatGPT 在.....方面的效率撰写病例报告	临床评估使用查询	强调 the ChatGPT 无法.....综合该案例的情感或道德方面的通过呈现各个方面仅限事实。
（Rockwell 等，2023年)	USA	医学	撰写了一篇手稿在.....的帮助下ChatGPT，它讨论了.....的诊断以及.....的治疗患有罕见真菌感染的患者。	治疗与全身性抗真菌药物使用详细的查询	生成一篇期刊文章使用 ChatGPT，发现组织胞浆菌病的临床表现可能存在很大差异，且诊断可以是具有挑战性的
（Macdonald 等人，2023年)	UK	医学	评估了 ChatGPT 的加快.....的能力论文撰写，确定疫苗有效性，以及生成代码以进行生存分析并计算风险比。	对.....的描述模拟数据集提供给ChatGPT 和代码已生成通过查询和错误反馈 到允许 ChatGPT 这样做进行自我纠正。	阐释前景ChatGPT 是一项出色的研究辅助工具，可帮助研究人员生成代码、解读数据，以及起草论文
（Salvagno、Taccone 和 Gerli，2023年）	比利时，意大利	医学	探索了.....的能力ChatGPT 在创作.....方面初稿，整理手稿，以及检查错误《重症医学》领域的科学文章《重症医学》领域医学。	文献综述，案例研究和未来应用	阐明.....的必要性在人类干预方面取得重大决策与伦理方面的考量科学方面尽管写作ChatGPT 的科学写作能力
(Mann, 2023年)	USA	医学	讨论了人工智能的作用在转化医学中	开展对.....进行访谈包括 ChatGPT对大数据数据集	解决.....问题的重要性数据相关问题依赖性，复杂性生物系统的，以及伦理影响尽管拥有许多优势在于转化医学包括高效且有效地分析大数据
（Charrois-Durand 等人，2023年）	加拿大	医学	考察了ChatGPT 撰写一篇关于放射性诱发主动脉炎的研究，病例为一名宫颈癌患者	.....的案例研究辐射诱发的主动脉炎	重大贡献ChatGPT 在准备方面的在线a关于辐射的研究-由.....诱发的主动脉炎提供相关且对.....的准确回答向其提出的问题
(Huh, 2023年a)	韩国	医学	评估并比较ChatGPT 的知识截止日期	管理一项 a寄生虫学考试 传统重叠率	表现更好相比于医学学生在.....方面

			和推断技能时 医学生	ChatGPT 和 医学生	知识与推断 寄生虫学领域的能力 考试
（Gilson等，2023年）	USA, 爱尔兰	社会科学	考察了 ChatGPT的表现 在回答问题时 美国的 医学执照 考试	两组 多项选择题 来自 题库的问题： AMBOSS 和 NBME	识别 ChatGPT 作为学习工具的 巨大潜力 用于……的学习工具 由于医学教育中的 其渊博的知识 和逻辑思维 相当于三年级的 医学生
（Lim 等，2023年）	马来西亚, 澳大利亚	社会科学, 商业, 管理与 会计	讨论了 生成式人工智能的应用、优势 和局限性 在教育领域 教育	对现有研究的综述 文献与 时事 与生成式人工智能 和教育相关	生成式人工智能的能力 生成式人工智能能够 变革教育 促进创新性 技术的使用和 挑战生成式人工智能 与……相关的方面 工作的原创性和 剽窃
(Hamdoun 等, 2023年)	USA	社会科学, 工程	探讨了 数字化的影响#+#+#+#+ 心理健康相关的 嵌入式应用 与聊天机器人进行 交流	数字分析 心理健康 从……的应用 技术批判 视角	识别并 讨论数字心理 健康风险和 不足之处
（Lund等，2023年）	USA, 中国, 印度	社会科学	分析了……的影响 ChatGPT 对学术界的影响 以及学术出版	对……的调查 伦理问题	ChatGPT 的巨大潜力 能够显著 对学术界产生影响，并 影响学术研究，此外 还涉及需要考虑的伦理 需要……的问题。 在采用 这项技术时
（Dowling 和 Lucey, 2023年)	爱尔兰, 阿联酋、越南 Nam, 中国	经济学, 计量经济学 与金融学	评估了 ChatGPT 的 协助方面的有效性 金融研究	评估 of 加密货币 研究成果 使用 ChatGPT	凸显 ChatGPT 的卓越 ChatGPT的性能 在三个版本上 金融研究理念 在论文中讨论的内容
(Kirmani, 2022)	法国	化学, 材料 《科学》、能源	考察了……的能力 ChatGPT 创作诗歌	对……能力的分析 ChatGPT 的能力 当要求它……时 写一首诗歌 钙钛矿	显著程度的 勤勉、精准和 所展现出的复杂性 ChatGPT 在撰写一篇 单独的诗歌。
(Mijwil 等，2023年)	伊拉克、美国	计算机 《科学》	讨论了 ……的重要性 网络安全以及 利用 ChatGPT 来 确定应对……的方法 网络犯罪	生成一个章节 期刊文章通过 直接使用 ChatGPT	……的重要性 人工智能在 通过网络安全 写作能力 ChatGPT 的
（Selivanov 等人, 2023年)	俄罗斯	多学科	讨论了自动化 临床图像描述生成 采用整合方式 两个语言模型	综合的 方法 由……组成 Show-Attend-Tell （SAT）和 GPT- 3	开发一种高效的 用于图像描述生成的模型 在对胸部 X 光图像进行……后 在两个数据集上进行测试： MIMIC-CXR 和 MS- COCO

已有多项研究探讨了在医学领域的各个方面使用 ChatGPT的情况。Nachshon 等人的一项研究中，ChatGPT 帮助撰写了一例与三级烧伤面积达 100% 有关的病例描述（Nachshon et al., 2023）。研究发现，尽管 ChatGPT 能够陈述事实，却无法综合分析或评论该病例中的情感或道德层面。Rockwell 等人考察了使用 ChatGPT 撰写的一例罕见真菌感染的诊断和治疗结果（Rockwell et al., 2023）。Macdonald 等人根据模拟数据集的描述，从多个方面评估了 ChatGPT，包括加快论文起草、生成生存分析代码以及计算风险比（Macdonald et al., 2023）。研究结果表明，ChatGPT 在上述领域具有实用价值。另一项研究讨论了 ChatGPT 在总结数据、提出结构建议和查找学术论文等学术任务中的表现。研究发现，在人工监督下，ChatGPT 可以成为一种有用的工具（Salvagno, Taccone and Gerli, 2023）。Mann 的研究探讨了使用 ChatGPT 分析大型转化医学数据集，并指出，由于存在数据依赖、生物系统复杂性和伦理影响等多项限制，这款聊天机器人的效率不高（Mann, 2023）。Charrois-Durand 等人的一项研究报告了一例使用 ChatGPT 并借助影像学检查诊断放射性主动脉炎的病例（2023）。在 Huh 的另一项研究中，作者评估了 ChatGPT 回答医学问题的能力，并将其与学生的表现进行了比较（Huh, 2023a）。ChatGPT 的准确率为 60.8%，而学生的准确率为 90.8%，这表明 ChatGPT 的表现不及医学生。

ChatGPT在社会科学中的应用也引起了研究人员的广泛关注。Gilson等（2023）使用AMBOSS和NBME题目考察了ChatGPT与其他聊天机器人在美国医学执照考试中的效率，结果发现，ChatGPT的表现比InstructGPT和GPT-3高出60%，并且在100%的情况下提供了合乎逻辑的解释。在另一项研究中，Lim等（2023）评估了生成式人工智能（人工智能）在教育中的潜在用途及其缺点。此外，就心理健康管理而言，Hamdoun等（2023）开展的研究表明，通过纳入伦理考量，基于文本或支持语音的对话代理能够发挥有效作用。Lund等（2023）近期开展的一项研究证明，ChatGPT有望提高学术界和学术出版领域的生产力。

其他领域对 ChatGPT 的应用已成为多项研究的主题，包括金融研究、摘要撰写和网络安全。在金融研究方面，Dowling 和 Lucey（2023年）考察了 ChatGPT 生成看似可信的期刊研究的实用性；研究人员的领域专业知识以及获取个人数据的能力决定了生成内容的质量。另一项研究中，Kirmani（2022）指出，ChatGPT 在撰写摘要时，其呈现方式、简洁性和细节丰富程度都颇具前景。此外，ChatGPT 也被用于撰写网络安全相关文章（Mijwil 等，2023年）。研究人员还开发了一种新颖技术，可根据放射影像扫描数据和患者信息自动生成说明文字，并为临床决策提供可解释的结果（Selivanov 等，2023年）。这些研究凸显了生成式人工智能在不同领域应用的潜在益处与挑战。以下综述了30篇关于 ChatGPT 在不同领域应用的科学文献。

3.2.2.2. 其他来源

近期研究表明，ChatGPT 生成的摘要可能难以与真实摘要区分，这引发了人们对科学论文原创性和可信度的担忧。已有多项研究探讨 ChatGPT 在撰写摘要和协助论文发表方面的应用。一个重要的关注点是 ChatGPT 对科学论文写作的影响（González-Padilla，2023年；Gordijn 和 Have，2023年；Huh，2023年b；Lee，2023年；Patel 和 Lam，2023年；Selivanov 等，2023年；Yeo-Teh 和 Tang，2023年）。根据近期对使用 ChatGPT 生成的摘要原创性进行评估的研究，学术审稿人仅识别出63%的虚假摘要 （Thorp，2023年）；另一项研究（Else，2023年）则发现，生成的摘要中仅有68%、真实摘要中仅有86%被正确识别。因此，ChatGPT 能够混入学术出版物，且无法对科学论文的内容或诚信负责（《自然》编辑部，2023年）。许多出版商、作者和预印本审稿人都认同 这种聊天机器人不符合研究作者的标准。另一个原因是，学术手稿作者的正式身份必须与作者作为文献撰写者这一更广泛的概念区分开来（Stokel-Walker，2023年）。将这种聊天机器人纳入学术界存在问题。此外，多项研究（Flanagin 等，2023年；Hosseini、Rasmussen 和 Resnik，2023年；Stokel-Walker，2022；Stokel-Walker 和 Van Noorden，2023年）讨论了 ChatGPT 提供错误和误导性信息的问题，尤其是 Van Dis 等（2023年）发现，当 ChatGPT 面对一系列具体的出版相关技术问题，以及要求深入理解文献的任务时，会生成不准确且具有误导性的回答。此外，进一步调查显示，该模型所用训练数据截至2021年，因此未涵盖《联合国气候变化框架公约》第二十七次缔约方大会（COP27）等较近发生的事件（Chatterjee 和 Dethlefs，2023年）。尽管存在上述问题，ChatGPT 很可能会在未来不断发展，届时能够快速扫描互联网，搜寻数千篇不同主题的开放获取科学文章，并在毫秒内生成附有恰当引文的书面答复（O’Connor 和 ChatGPT，2023年）。

关于ChatGPT，存在大量与医疗保健相关的非期刊文章类科学文献。ChatGPT可用于提供医疗信息和帮助。例如，ChatGPT可以回答医疗问题，或为常见症状提供鉴别诊断 （Arif、Munaf和Ul-Haque，2023年），收集相关临床信息，以中等准确度从手术报告中生成CPT代码，并从临床记录中提取ICD-10代码（DiGiorgio和Ehrenfeld，2023年）。2023年发表的一项研究探讨了相关应用，例如将医学语言翻译成更易理解的文本，或将说明直接翻译成患者的母语（Scerri和Morin，2023年）。ChatGPT还可以通过分析患者记录，协助开发临床决策和支持系统（Dahmen等，2023年；De Angelis等，2023年）。Huang等人研究了ChatGPT对糖尿病技术的贡献（Huang等，2023年）。另一项研究探讨了ChatGPT在外科实践中的作用，包括麻醉和术前外科护理 （Bhattacharya等，2023年）。ChatGPT在医疗保健系统中的另一项应用是生成和改进出院小结；由于其格式在很大程度上已经标准化，这显然是一个合适的应用场景。然而，需要注意的是，该程序的输出在完成之前必须由医生进行人工审核 （Patel和Lam，2023年）。

除了在写作和医疗保健领域的应用外，ChatGPT 也在其他领域引起了关注。一项关于ChatGPT在全球变暖领域潜在用途的研究已经开展，该用途包括为政策决策生成气候情景，同时也探讨了其潜在缺点，例如缺乏情境意识和缺乏问责机制（Biswas，2023年）。另一项研究评估了ChatGPT模拟员工与组织之间互动以及帮助研究人员理解复杂组织动态的能力（Dasborough，2023年）。此外，ChatGPT在旅游业和精神病学领域的功能也已在多项研究中得到讨论（Thornton、D'Souza和Tandon，2023年；Nautiyal等，2023年）。总的来说，尽管ChatGPT是一项相对较新的技术，但它已经进入了各个领域和行业，其中最重要的领域已在本文中提及并进行了讨论。

4. 讨论

近年来，聊天机器人凭借其在医疗保健、客户服务和教育等领域的多样化应用，日益受到欢迎。在这些聊天机器人中，OpenAI 开发的 ChatGPT 凭借基于 GPT-3.5 架构的先进自然语言处理能力脱颖而出。ChatGPT 使研究人员和公众能够提出问题、获得解释，并就不同领域的各种相关主题展开对话。全球研究人员对聊天机器人研究的兴趣日益浓厚，参与研究的人员来自全球约 60% 的国家和地区（图 3）。尽管这是一个相对较新的主题，已有 46 个国家开展了有关 ChatGPT 的研究（图 8）。美国凭借其强大的研究能力，拥有数量最多的科学研究机构，在聊天机器人和 ChatGPT 相关出版物的数量上均位居首位。继美国之后，德国和英国在聊天机器人和 ChatGPT 研究方面处于领先地位（图 3）。它们位列 Scopus 中聊天机器人出版物产出最多的国家之列，在 ChatGPT 出版物数量上分别位居第二和第三。

本研究表明，在 Scopus 和 Web of Science 中，聊天机器人领域的作者均可归为15个聚类。此外，研究从不同角度比较了作者影响力（表4）。在 Scopus 的聊天机器人研究中，Gao J 的引文数最多（1344）；而在 Web of Science 中，FOLSTAD A 的引文数最高。此外，该作者在两个数据库中的 h指数、g指数、m指数和出版物数量均为最高。该作者在 Scopus 中与 Araujo, T 合作密切（图4a），在 WoS 中则与 Hobert, S 合作密切（图4b）。

关键词分析表明，正如图7所示，ChatGPT是聊天机器人领域近期兴起的趋势。此外，在自然语言处理技术的助力下，心理健康已成为聊天机器人开发中日益热门的研究领域（图5）。值得注意的是，正如表5和表6所示，在 Scopus 和 Web of Science 数据库中，心理健康都是相应聚类中的首要关键词。此外，医疗保健也成为 ChatGPT 相关研究的重要关注领域（图11）。研究人员已探索 ChatGPT 在提供个性化医疗保健建议、促进远程患者监测以及协助医疗保健专业人员进行决策方面的潜力。

本研究存在若干局限性。首先，研究仅关注收录于 Scopus 和 Web of Science 的出版物，可能遗漏其他数据库中的相关文献。其次，记录中的某些字段为空，这可能影响了研究结果。这些局限性表明，在外推研究发现时应保持谨慎，并强调有必要开展更多研究以确认这些结果。

5. 结论

本文采用文献计量分析方法，对聊天机器人进行了全面的文献综述，特别聚焦于1998年至2023年的ChatGPT研究，并使用Scopus和Web of Science数据库。聊天机器人研究每年都在稳步增长，Web of Science和Scopus的年增长率分别约为19%和27%。这些论文大多采用深度学习、机器学习和自然语言处理技术。过去，聊天机器人研究主要聚焦于新冠肺炎和本体论，而当前的研究趋势则是心理健康和任务分析。ChatGPT被视为聊天机器人研究中一个很有前景的研究方向，相关文献主要关注其在研究伦理、医学和社会科学等领域的能力与局限性。研究发现，尽管ChatGPT存在一些局限，有时也需要人工协助，但它在多种应用中具有潜力，包括回答医学问题、开发临床决策与支持系统、改进出院小结、协助写作和翻译、模拟员工与组织之间的互动，以及支持政策决策。因此，研究人员正在积极探索ChatGPT框架下的这些研究方向。

资助

本研究未获得公共、商业或非营利部门资助机构提供的任何专项资助。

参考文献

Adamopoulou, E. 与 Moussiades, L.（2020a）。聊天机器人技术概述。《IFIP信息与通信技术进展》，373–383。https://doi.org/10.1007/978-3-030- 49186-4 31_Adamopoulou, E. 与 Moussiades, L.（2020b）。聊天机器人：历史、技术与应用。《机器学习与应用》，2，100006。https://doi.org/10.1016/j.mlwa.2020.100006 Amara, A.、Hadj Taieb, M. A. 与 Ben Aouicha, M.（2021）。网络表示学习系统综述：先驱与当前发展状况。《机器学习与应用》，6，100130 101016 2021 100130 。https://doi.org/ . /j.mlwa. . Aria, M. 与 Cuccurullo, C.（2017）。Bibliometrix：用于全面科学图谱分析的R工具。《信息计量学》，11(4)，959–975。https://doi.org/10.1016/j.joi.2017.08.007 Arif, T. B.、Munaf, U. 与 Ul-Haque, I.（2023）。医学教育与研究的未来：ChatGPT究竟是伪装的福音还是祸害？《医学教育在线》，28(1)。https://doi.org/10.1080/10872981.2023.2181052

Ashfaq, M.、Yun, J.、Yu, S. 与 Loureiro, S. M.（2020）。我，聊天机器人：人工智能驱动的服务代理的用户满意度与持续使用意愿决定因素建模。《远程信息处理与信息学》，54，101473。https://doi.org/10.1016/j.tele.2020.101473 Aydin, N. 与 Erdem, O.（2022）。关于新一代人工智能技术——生成式预训练变换器3的研究。2022年第3届国际信息学与软件工程会议（IISEC）。https://doi.org/10.1109/iisec56263.2022.9998298 Bhargav, P. N. V.、Kishore, A. 与 Doraswamy, M.（2020）。布拉德福定律在生产工程文献领域的应用：博士论文的文献计量分析。《国际图书馆科学研究杂志》，6(2)，24。https://doi.org/10.26761/ijrls.6.2.2020.1330 Bhattacharya, K.、Bhattacharya, A. S.、Bhattacharya, N.、Yagnik, V. D.、Garg, P. 与 Kumar, S.（2023）。外科实践中的ChatGPT——一位新秀。《印度外科杂志》。https://doi.org/10.1007/s12262-023-03727-x Biswas, S. S.（2023）。ChatGPT在全球变暖问题中的潜在应用。《生物医学工程年鉴》。https://doi.org/10.1007/s10439-023-03171-8 Borgohain, D. J.、Verma, M. K.、Nazim, M. 与 Sarkar, M.（2021）。布拉德福定律与Leimkuhler模型在信息科学文献中的应用。《COLLNET科学计量学与信息管理杂志》，15(1)，197–212。https://doi.org/10.1080/09737766.2021.1943041 Chai, J.、Zeng, H.、Li, A. 与 Ngai, E. W. T.（2021）。计算机视觉中的深度学习：新兴技术与应用场景的批判性综述。《机器学习与应用》，6，100134 101016 2021 100134 。https://doi.org/ . /j.mlwa. . Charrois-Durand, C.、Beauchemin, M.-C. 与 Barkati, M.（2023）。一例宫颈癌患者放射性主动脉炎病例。《Cureus》。https://doi.org/10.7759/cureus.35484 Chatterjee, J. 与 Dethlefs, N.（2023）。这一新型会话式人工智能模型可以成为你的朋友、哲学家和向导……甚至是你最可怕的敌人。《Patterns》，4(1)，100676。https://doi.org/10.1016/j.patter.2022.100676 Croes, E. A. 与 Antheunis, M. L.（2020）。我们能与Mitsuku成为朋友吗？人类与社交聊天机器人关系形成过程的纵向研究。《社会与个人关系杂志》，38(1)，279–300。https://doi.org/10.1177/0265407520959463 Dahmen, J.、Kayaalp, M. E.、Ollivier, M.、Pareek, A.、Hirschmann, M. T.、Karlsson, J. 与 Winkler, P. W.（2023）。医学研究中的人工智能机器人ChatGPT：作为双刃剑的潜在游戏改变者。《膝关节外科、运动创伤学与关节镜》，31(4)，1187–1189。https://doi.org/10.1007/s00167-023-07355-6 DALE, R. O. B. E. R. T.（2016）。聊天机器人的回归。《自然语言工程》，22(5)，811– 817。https://doi.org/10.1017/s1351324916000243 Dasborough, M. T.（2023）。人工智能领域令人惊叹的进步：CHATGPT对组织行为领域的影响。《组织行为杂志》，44(2)，177–179。https://doi.org/10.1002/job.2695 De Angelis, L.、Baglivo, F.、Arzilli, G.、Privitera, G. P.、Ferragina, P.、Tozzi, A. E. 与 Rizzo, C.（2023）。ChatGPT与大语言模型的兴起：人工智能驱动的信息疫情对公共卫生构成的新威胁。《SSRN电子期刊》。https://doi.org/10.2139/ssrn.4352931 DiGiorgio, A. M. 与 Ehrenfeld, J. M.（2023）。医学人工智能与CHATGPT：让医生摆脱束缚。《医学系统杂志》，47(1)。https://doi.org/10.1007/s10916-023- 01926-3 Dowling, M. 与 Lucey, B.（2023）。用于金融研究的ChatGPT：Bananarama猜想。《金融研究通讯》，53，103662。https://doi.org/10.1016/j.frl.2023.103662 Dwivedi, Y. K.、Hughes, L.、Ismagilova, E.、Aarts, G.、Coombs, C.、Crick, T.、Duan, Y.、Dwivedi, R.、Edwards, J.、Eirug, A.、Galanos, V.、Ilavarasan, P. V.、Janssen, M.、Jones, P.、Kar, A. K.、Kizgin, H.、Kronemann, B.、Lal, B.、Lucini, B.、……Williams, M. D.（2021）。人工智能（AI）：关于新兴挑战、机遇与研究议程的多学科视角，

实践与政策。《国际信息管理杂志》，57，101994。
https://doi.org/10.1016/j.ijinfomgt.2019.08.002 Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J.,Filieri,R., Jacobson, J., Jain, V., Karjaluoto, H.,Kefi, H., Krishen, A. S.,Kumar, V., Rahman, M. M., Raman, R., Rauschnabel, P. A., Rowley, J., Salo,J., Tran,G. A.,& Wang, Y.（2021年）。确定数字与社交媒体营销研究的未来：观点与研究命题。《国际信息管理杂志》，59，102168。https://doi.org/10.1016/j.ijinfomgt.2020.102168Egghe,L.（1990年）。布拉德福德定律理论在计算莱姆库勒定律及完成书目中的应用。《美国信息科学学会杂志》，41(7)，469–492。
https://doi.org/10.1002/(sici)1097-4571(199010)41:7<469::aid-asi1>3.0.co;2-p Else, H.（2023年）。ChatGPT撰写的摘要愚弄了科学家。《自然》，613(7944)，423–423。
https://doi.org/10.1038/d41586-023-00056-7 Flanagin, A., Bibbins-Domingo, K.,Berkwits, M.,& Christiansen, S. L.（2023年）。非人类“作者”及其对科学出版和医学知识完整性的影响。《美国医学会杂志》，329(8)，637。https://doi.org/10.1001/jama.2023.1344 Floridi,L., & Chiriatti,M.（2020年）。GPT-3：其性质、范围、局限与后果。《心智与机器》，30(4)，681–694。
https://doi.org/10.1007/s11023-020-09548-1Garg,K. C., Kumar, S., &Singh, R.K.（2020年）。图书馆与信息科学领域四种文摘与索引服务所收录期刊覆盖范围及重叠情况的文献计量学研究。《连续出版物图书馆员》，79(1-2)，118–130。https://doi.org/10.1080/0361526x.2019.1704341 Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., &Chartash, D.（2023年）。ChatGPT在美国执业医师资格考试中的表现如何？大语言模型对医学教育和知识评估的影响。《JMIR医学教育》，9。https://doi.org/10.2196/45312Gluck, M.（1990年）。期刊覆盖重叠综述及对重叠定义的扩展。《美国信息科学学会杂志》，41(1)，43–60。
https://doi.org/10.1002/(sici)1097-4571(199001)41:1<43::aid-asi4>3.0.co;2-p González-Padilla,D.A.（2023年）。关于人工智能在手稿写作中潜在风险的担忧。信件。《泌尿外科杂志》，209(4)，682–683。https://doi.org/10.1097/ju.0000000000003131 Gordijn, B.,& Have, H. ten.（2023年）。ChatGPT：演进还是革命？《医学、健康照护与哲学》，26(1)，1–2。
https://doi.org/10.1007/s11019-023-10136-0 Hamdoun, S., Monteleone, R.,Bookman,T., & Michael,K.（2023年）。基于人工智能和数字技术的心理健康应用程序：平衡需求与风险。《IEEE技术与社会杂志》，42(1)，25–36。https://doi.org/10.1109/mts.2023.3241309 Hill, J., RandolphFord, W., & Farreras, I.G.（2015年）。与人工智能进行真实对话：人类—人类在线对话与人类—聊天机器人对话的比较。《人类行为中的计算机》，49，245–250。
https://doi.org/10.1016/j.chb.2015.02.026 Hosseini, M., Rasmussen, L. M.,& Resnik,D.B.（2023年）。使用人工智能撰写学术出版物。《研究问责》，1–9。
https://doi.org/10.1080/08989621.2023.2168535 Huang, J., Yeung, A. M.,Kerr, D., & Klonoff, D.C.（2023年）。使用ChatGPT预测糖尿病技术的未来。《糖尿病科学与技术杂志》，193229682311610。https://doi.org/10.1177/19322968231161095 Huh,S.（2023a）。ChatGPT的知识和理解能力是否可与韩国医学生参加寄生虫学考试时的能力相媲美？：一项描述性研究。《卫生专业人员教育评估杂志》，20，1。https://doi.org/10.3352/jeehp.2023.20.1 Huh, S.（2023b）。医学期刊出版与编辑政策中的近期问题：采用人工智能、预印本、开放式同行评审、模型文本重复使用政策、学术出版最佳实践第4版以及标题中的国家名称。《神经介入》，18(1)，2–8。
https://doi.org/10.5469/neurointr.2022.00493

IBM。（不详）。什么是自然语言处理？https://www.ibm.com/topics/natural-language-processing Jovanovic, M.,&Campbell,M.（2022）。生成式人工智能：趋势与前景。《计算机》，55(10)，107–112。https://doi.org/10.1109/mc.2022.3192720 Keramatfar,A.,Rafiee,M.,& Amirkhani, H.（2022）。图神经网络：文献计量学综述。《机器学习与应用》，10，100401。https://doi.org/10.1016/j.mlwa.2022.100401 Khodadadi, A.,Ghandiparsi, S.,&Chuah, C.-N.（2022）。一种基于自然语言处理和深度学习的模型：利用自由文本客户服务报告进行自动车辆诊断。《机器学习与应用》，10，100424。
https://doi.org/10.1016/j.mlwa.2022.100424King,M. R.（2023年）。关于人工智能、聊天机器人和高等教育中剽窃问题的对话。《细胞与分子生物工程》，16(1)，1–2。
https://doi.org/10.1007/s12195- 022-00754-8Kirmani,A.R.（2022）。人工智能赋能的科学诗歌。《ACS能源快报》，8(1)，574–576。https://doi.org/10.1021/acsenergylett.2c02758《柳叶刀·数字健康》。（2023年）。Chatgpt：朋友还是敌人？《柳叶刀·数字健康》，5(3)。
https://doi.org/10.1016/s2589-7500(23)00023-7Lee, J. Y.（2023年）。人工智能聊天机器人能否成为学术期刊文章的作者？《健康专业教育评价杂志》，20，6。
https://doi.org/10.3352/jeehp.2023.20.6LEIMKUHLER, F. E.R.D.I.N. A.N. D.F.（1980）。布拉德福德定律的精确表述。《文献学报》，36(4)，285–292。https://doi.org/10.1108/eb026699 Lim, W. M., Gunasekara, A., Pallant, J. L.,Pallant, J. I.,& Pechenkina, E.（2023年）。生成式人工智能与教育的未来：诸神黄昏还是改革？来自管理学教育者的悖论视角。《国际管理教育期刊》，21(2)，100790。https://doi.org/10.1016/j.ijme.2023.100790Lund,B.D., & Wang, T.（2023年）。谈论Chatgpt：人工智能和GPT将如何影响学术界与图书馆？《图书馆高科技新闻》。
https://doi.org/10.1108/lhtn-01-2023-0009 Lund, B. D., Wang, T., Mannuru, N. R.,Nie,B., Shimray, S., & Wang,Z.（2023年）。Chatgpt与新的学术现实：人工智能撰写的研究论文以及学术出版中大语言模型的伦理问题。《信息科学与技术协会学报》。
https://doi.org/10.1002/asi.24750 Macdonald, C., Adeloye,D., Sheikh,A., &Rudan, I.（2023年）。Chatgpt能起草研究论文吗？以人群层面的疫苗有效性分析为例。《全球健康杂志》，13。
https://doi.org/10.7189/jogh.13.01003 Makkizadeh, F.,&Sa'adat, F.（2017年）。不孕症领域文章的文献计量学与主题分析（2011-2015）。《国际生殖生物医学杂志》，15(11)，719–728。
https://doi.org/10.29252/ijrm.15.11.719Mann, D.L.（2023年）。人工智能探讨人工智能在转化医学中的作用。《JACC：基础与转化科学》，8(2)，221–223。
https://doi.org/10.1016/j.jacbts.2023.01.001 Martarelli, N. J., & Nagano, M. S.（2021年）。高影响力科学研究如何设计混合数据聚类实验？指导更佳选择的系统性图谱。《机器学习与应用》，5，100056。https://doi.org/10.1016/j.mlwa.2021.100056 McCarthy, J.（2007）。从这里走向人类水平的人工智能。《人工智能》，171(18)，1174–1182。
https://doi.org/10.1016/j.artint.2007.10.009 Mijwil, M.M.,ChatGPT, & Aljanabi, M.（2023年）。迈向基于人工智能的网络安全：应对网络犯罪的实践与CHATGPT生成的方法。《伊拉克计算机科学与数学杂志》，65–70。https://doi.org/10.52866/ijcsm.2023.01.01.0019 Molnar, G., & Szuts, Z.（2018）。聊天机器人在正规教育中的作用。2018 IEEE第16届智能系统与信息学国际研讨会（SISY）。https://doi.org/10.1109/sisy.2018.8524609

Nachshon, A., Batzofin, B., Beil, M., & van Heerden,P.V.（2023年）。姑息治疗可能是管理严重烧伤的唯一选择：一篇借助聊天机器人撰写的病例报告。Cureus。https://doi.org/10.7759/cureus.35649 Nash-Stewart,C.E.,Kruesi,L.M.,&DelMar, C.B.（2012年）。布拉德福德分散定律能否预测Cochrane综述中文献的规模？医学图书馆协会杂志：JMLA，100(2)，135–138。https://doi.org/10.3163/1536-5050.100.2.013《自然》编辑部。（2023年）。诸如ChatGPT之类的工具威胁透明科学；以下是我们使用这些工具的基本规则。《自然》，613(7945)，612–612。https://doi.org/10.1038/d41586-023-00191-1 Nautiyal,R.,Albrecht, J. N., &Nautiyal, A.（2023年）。ChatGPT与旅游学术界。《旅游研究年鉴》，99，103544。https://doi.org/10.1016/j.annals.2023.103544Nisonger, T.E.（2008年）。“80/20法则”与核心期刊。《连续出版物图书馆员》，55(1-2)，62–84。https://doi.org/10.1080/03615260801970774 O’Connor, S., & ChatGPT。（2023年）。护理教育中的开放式人工智能平台：促进学术进步的工具，还是滥用？《护理教育实践》，66，103537。https://doi.org/10.1016/j.nepr.2022.103537 Patel,S.B., &Lam, K.（2023年）。ChatGPT：出院小结的未来？《柳叶刀·数字健康》，5(3)。https://doi.org/10.1016/s2589-7500(23)00021-3 Potter,J.（2010年）。作业治疗文献映射：更新。《医学图书馆协会杂志：JMLA》，98(3)，235–242。https://doi.org/10.3163/1536-5050.98.3.012 Qiu, J., Zhao, R.,Yang,S.,&Dong, K.（2017年）。文献信息的集中与分散分布：布拉德福定律。《信息计量学》，89–119。https://doi.org/10.1007/978-981-10-4032-0_4 Ralston, K., Chen, Y.,Isah, H., & Zulkernine, F.（2019年）。使用IBM Watson的语音交互式多语言学生支持系统。2019年IEEE国际机器学习与应用会议（ICMLA）。https://doi.org/10.1109/icmla.2019.00309 Rockwell, M., Spiller,P., Cuevas-Ocampo,A. K., Malek,A.,& Mankekar, G.（2023年）。外耳道组织胞浆菌病。Cureus。https://doi.org/10.7759/cureus.35644Sadavarte, S. S.,&Bodanese, E.（2019年）。使用Alexa和亚马逊云科技的孕期陪伴聊天机器人。2019年IEEE浦那分会国际会议（PuneCon）。https://doi.org/10.1109/punecon46936.2019.9105762 Salvagno, M., Taccone, F.S., &Gerli, A. G.（2023年）。人工智能能否帮助科学写作？《重症医学》，27(1)。https://doi.org/10.1186/s13054-023-04380-2 Sangam, S.L.（2015年）。布拉德福德经验定律。《图书馆发展杂志》，1(1)，1。https://doi.org/10.32344/jld/v1/i1/2015/1-13Scerri, A., &Morin, K. H.（2023年）。使用类似ChatGPT的聊天机器人支持护理实践。《临床护理杂志》。https://doi.org/10.1111/jocn.16677 Selivanov, A., Rogov, O. Y., Chesakov, D., Shelmanov,A., Fedulova,I.,& Dylov, D. V.（2023年）。通过生成式预训练变换器生成医学图像描述。《科学报告》，13(1)。https://doi.org/10.1038/s41598-023-31223-5 Sharma,V.,Goyal, M.,& Malik, D.（2017年4月1日）。聊天机器人系统展现的智能行为。《新技术与研究国际期刊》。取自 https://www.neliti.com/publications/263312/an-intelligent-behaviour-shown-by-chatbot-system Skjuve, M., Følstad, A.,Fostervold,K. I., & Brandtzaeg, P. B.（2021年）。我的聊天机器人伴侣——一项关于人机聊天机器人关系的研究。《国际人机交互研究期刊》，149，102601。https://doi.org/10.1016/j.ijhcs.2021.102601 Stokel-Walker, C.（2022）。人工智能机器人ChatGPT撰写高质量文章——教授们是否应该担心？《自然》。https://doi.org/10.1038/d41586-022-04397-7 Stokel-Walker,C.（2023年）。ChatGPT被列为研究论文作者：许多科学家对此表示不赞同。《自然》，613(7945)，620–621。https://doi.org/10.1038/d41586-023-00107-z Stokel-Walker, C., & Van Noorden, R.（2023年）。ChatGPT和生成式人工智能对科学意味着什么。《自然》，614(7947)，214–216。https://doi.org/10.1038/d41586-023-00340-6

Taecharungroj, V.（2023年）。“ChatGPT 能做什么？”：分析 Twitter上对创新型人工智能聊天机器人的早期反应。《大数据与认知计算》，7(1)，35。https://doi.org/10.3390/bdcc7010035 Thornton, J., D’Souza,R.,& Tandon,R.（2023年）。人工智能与精神病学研究和实践。《亚洲精神病学杂志》，81，103509。https://doi.org/10.1016/j.ajp.2023.103509 Thorp, H. H.（2023年）。ChatGPT 很有趣，但不是作者。《科学》，379(6630)，313–313。https://doi.org/10.1126/science.adg7879 vanDis, E. A., Bollen, J., Zuidema, W.,van Rooij, R., & Bockting, C. L.（2023年）。ChatGPT：研究的五项优先事项。《自然》，614(7947)，224–226。https://doi.org/10.1038/d41586-023-00288-7 van Eck,N.J.,& Waltman, L.（2009）。软件调查：用于文献计量学映射的计算机程序VOSviewer。《科学计量学》，84(2)，523–538。https://doi.org/10.1007/s11192-009-0146-3 Yeo-Teh,N.S., & Tang,B. L.（2023年）。致编辑的信：诸如 ChatGPT之类的自然语言处理系统不能被列为作者，因为它们无法满足广泛采用的作者资格标准。《研究问责》，1–3。https://doi.org/10.1080/08989621.2023.2177160