**人工智能对信息检索的影响与创新发展**

张正豪

外国语学院 日语系 20222286

**摘要：**

随着人工智能技术的快速发展，信息检索领域正面临着深刻的变革。本综述论文旨在探讨人工智能对信息检索的影响，并介绍其在创新发展方面的应用。通过对相关研究的综合分析，本文将阐述人工智能技术对传统信息检索方法的影响以及新兴应用场景的出现。

**关键词：人工智能；信息检索；影响；创新；应用场景；**

**The Impact and Innovative Development of Artificial Intelligence on Information Retrieval**

Zhang Zheng-hao

**Japanese Language Department, School of Foreign Languages 20222286**

**Abstract:** With the rapid development of artificial intelligence technology, the field of information retrieval is facing profound changes. This review paper aims to explore the impact of artificial intelligence on information retrieval and introduce its applications in innovative development. Through a comprehensive analysis of relevant research, this article will elaborate on the improvement of traditional information retrieval methods by artificial intelligence technology and the emergence of emerging application scenarios.

**Keyword:Artificial intelligence;Information retrieval;Effect;Innovate; Application Scenario;**

**引言：**

在信息爆炸时代，有效地获取所需信息成为人们的重要需求。人工智能的迅速崛起为信息检索带来了前所未有的机遇和挑战。本论文将探讨人工智能对信息检索的影响，并重点关注其在创新发展方面的应用。

**1人工智能在信息检索中的影响**

1.1自然语言处理技术的进步

自然语言处理（NLP）是基于对自然语言理解和自然语言生成信息处理的技术[[[1]](#footnote-0)]。近年来，从深度学习算法被提出到大语言模型不断成熟，从最初只能对简单的语言进行分析处理，遇到逻辑矛盾的情况系统就无法给出正确的结论到如今智能程度不弱于人类的chatgpt之类的大语言模型，甚至在一些方面人工智能比人类的能力更强，大数据时代下人工智能将会在未来越来越智能，应用越来越广泛。自然语言处理作为人工智能的一个分支，从提出以来诞生了一系列如GPT,BERT之类优秀的超级大语言模型，它们以聊天机器人为载体面向大众，聊天机器人与信息检索的关联性在大语言模型可以通过联网进行自主式深度学习过后会大幅上涨，因为在用户交互当中提出的最新消息或着模型数据库中没训练到的知识，语言模型主动链接Internet自我学习，而这就是一个信息检索的过程，并且如2022年10月Microsoft公司推出的New Bing[[[2]](#footnote-1)] 是一个帮助用户检索信息的智能机器人。这点上，说明了目前人工智能的应用场景已经进入到信息检索领域当中，并且New Bing接入chatgpt的API之后投入市场收到反馈良好，为大语言模型深入搜索引擎获得一个较好局面，同时也冲击着传统搜索引擎检索信息的手段。

**2人工智能在信息检索中的创新应用**

2.1智能搜索引擎与个性化推荐

随着Internet网络的飞速发展，Internet已成为人们日常生活中不可缺少的部分，网络作为信息搜寻的主要途径。[[[3]](#footnote-2)]面对巨量的信息空间，社会信息需求量越来越大，传统搜索方式对数据的发现、处理逐渐满足不了用户对高质量、高精确检索的需要，因此更加智能、个性化的搜索方式对满足用户精准需求显得尤为重要。基于机器人的搜索引擎是智能搜索引擎的一个很好应用场景，通过与机器人对话，将需求描述清楚，经过数据投喂的语言模型机器人通过理解和分析用户需求的语句，然后进行网络上的信息收集和检索，智能为用户过滤不需要的信息，并为用户提供更有效的信息，最后直接以URL超链接的方式给予用户，让用户直接点击超链接直达对口信息的网站，有效降低用户的时间成本同时，也提高使用体验。基于目前市面上最先进的检索机器人如Microsoft公司的Bing、Google公司BERT、百度公司文心一言等大语言模型，使用门槛较高，需要用户拥有精确的表达能力和会使用prompt方面的专业指令，才能使用好。用户个性化推荐系统，对用户检索历史分析识别用户喜好，利用时间轴划分，短期用户兴趣学习和长期用户兴趣学习的方法动态识别用户兴趣[[[4]](#footnote-3)]，相对与传统检索，能够提高用户检索信息的满意度和准确度。而这种方式在搜索引擎已经普及的时代下，作用到机器人检索当中会有更广阔的应用场景，更智能的为人类提供服务。

2.2智能问答与对话式检索

智能回答即通过训练过后的智能问答载体对用户阐述需求的描述进行的阅读理解、语义理解后对用户问题做出回答并帮助用户获得所需的信息，这同时也完成了对话是检索的过程，目前的智能回答平台主要以各大搜索引擎推出的如Bing,文心一言为主。该技术实现是基于NLP技术对自然语言处理有了更加深度的精确理解自然语言。

2.3跨语言信息检索

随着人工智能在诸多方面应用场景的诞生，信息检索必定是人工智能进入万家生活的排头兵。而跨语言检索在信息检索中很好的场景拓展。跨语言检索(Cross-language Information Retrieval, CLIR)是指用户通过一种语言提问，检索平台能够跨域该语言的局限检索到另一种或多种有关该信息研究的语言，这将极大拓展用户在Internet获取信息的来源，学习到其他语言文化下对检索信息的研究，拓宽认识。[[[5]](#footnote-4)]NLP技术对自然语言的理解愈深，同语义转化成不同语言的准确率就越高，为用户提供的检索信息就越有效，更好为人类服务，为此提供更好的商用场景。自2022年11月Open Ai公司chatgpt3.5的测试效果展现出极好的文本理解、语义理解能力，随之2023年4月GPT4.0的诞生，模型参数大幅扩大，语言模型升级展现出的效果更是令人吃惊，GPT4.0的稳定性准确性在3.5版本基础上仍有大幅提高，若说GPT3.5遇到数据库中未有的信息还会出现胡编乱造胡言乱语的情况，那么在GPT4.0上几乎看不到该情况。而GPT迭代的速度和测试效果又让跨语言信息检索这一检索场景看到较大可能性，按照NLP的发展速度，或许近几年就能实现跨语言信息检索。

**3人工智能驱动信息检索的未来趋势**

3.1用户个性化和推荐系统

自然语言处理技术能够使机器能更理解用户的搜索，从而更好的服务用户，人工智能下的未来信息检索拥有更广阔的发展的空间，未来人工智能甚至可以通过学习用户的搜索历史、兴趣喜好等，未用户提供更加个性化的服务和更精准的推送。这可以使用户更快速准确的得到他们的所需。

3.2隐私保护与伦理问题的挑战

随着人工智能发展趋势的迅猛，人工智能在方便人类工作生活的同时，也对个人隐私和人工智能与人类之间的伦理问题提出了挑战。一方面，人工智能应用暴露出侵犯数据隐私、制造信息茧房、大数据杀熟、深度伪造(Deep fake[[[6]](#footnote-5)])、过滤气泡(Filter Bubble)以及回声室效应(Echo Chamber[[[7]](#footnote-6)])等各种伦理问题；另一方面，对于人工智能产生自我意识对人类生存的危害可能性的恐慌。此外，对于人工智能落入恐怖主义等反人类组织当中也具有极大的信息泄露危害，由此，对人工智能采取严格管控和制定治理框架具有了必要性，而随之的人工智能安全领域将是一个十分重要的应用场景。AI安全是全世界各国和科技公司都值得关注和注意的问题，从2020年2月欧盟发布《数据战略》，2019年国家新一代人工治理专业委员会发布了《新一代智能治理原则》到2021年发布了更加细化严谨的《伦理规范》等措施，都能看出国家层面对人工智能发展的谨慎，以及对AI安全规范、严谨的重视，这些都是在传递AI安全领域的重要性的信号。

**结论**

人工智能对信息检索领域产生了深远的影响，并催生了创新的应用场景。通过对人工智能在信息检索中的影响与创新发展的综述，我们可以更好地了解当前的研究动态和未来的发展趋势。期待在人工智能与信息检索的融合中取得更多突破，如4G、5G发展那样涌现出全新的应用场景那样，并为用户提供更智能、高效的信息检索及其他服务。

**参考文献：**

1. 王海宁.自然语言处理技术发展[J].中兴通讯技术,2022,28(02):59-64.
2. Anonymous. The Help Desk: Learning About the New Bing[J]. Information Today,2023,40(4).
3. 宋毅,徐志明.个性化搜索用户兴趣更新学习及评价研究[J].计算机技术与发展,2018,28(06):64-66+72.
4. 宋毅,徐志明.个性化搜索用户兴趣更新学习及评价研究[J].计算机技术与发展,2018,28(06):64-66+72.
5. 臧劲松.人工智能在跨语言信息检索中的应用[J].计算机时代,2016(10):29-31+35.DOI:10.16644/j.cnki.cn33-1094/tp.2016.10.009.
6. 张灿.人工智能深度伪造技术的伦理风险与虚无困境[J].山东科技大学学报(社会科学版),2023,25(02):11-19.DOI:10.16452/j.cnki.sdkjsk.2023.02.006.
7. Scheibenzuber Christian,Neagu Laurentiu-Marian,Ruseti Stefan,Artmann Benedikt,Bartsch Carolin,Kubik Montgomery,Dascalu Mihai,Trausan-Matu Stefan,Nistor Nicolae. Dialog in the echo chamber: Fake news framing predicts emotion, argumentation and dialogic social knowledge building in subsequent online discussions[J]. Computers in Human Behavior,2023,140.

1. [↑](#footnote-ref-0)
2. [↑](#footnote-ref-1)
3. [↑](#footnote-ref-2)
4. [↑](#footnote-ref-3)
5. [↑](#footnote-ref-4)
6. [↑](#footnote-ref-5)
7. [↑](#footnote-ref-6)