

基于LDA的“人工智能鸿沟”多维主题分析

李方正¹, 虢毅¹, 陈佩璇², 秦超², 段永恒³

(1. 中南大学生命科学学院, 长沙 410031; 2. 吉林大学公共卫生学院, 长春 130021;
3. 深圳市卫生健康发展研究和数据管理中心, 广东 深圳 518000)

摘要: 随着人工智能技术的发展,人工智能接入不平等,即“人工智能鸿沟”现象逐渐凸显。以Web of Science数据库的文献为语料,基于LDA(latent Dirichlet allocation)模型,识别了“人工智能鸿沟”话题下相关主题,以此为框架,通过人工编码,进行文本分析。最终,识别并归纳智识壁垒、数字教育、医疗卫生和社会效应四个核心研究主题,并揭示技术知识差距、意识差距、数字教育不平等及人工智能在医疗卫生领域的应用不平等现状与差距;同时,人工智能鸿沟的社会效应,包括伦理文化困境、文化趋同和社会不平等问题,也是影响社会结构和个体生活的重要因素。通过解读“人工智能鸿沟”的核心主题,以期深入理解这一命题,助力人工智能增进人类福祉和社会进步。

关键词: 人工智能鸿沟; 数字鸿沟; LDA; 文本分析

中图分类号: G353.1 **文献标志码:** A **文章编号:** 1671-1807(2025)12-0036-08

1990年,托夫勒在《权力的转移》一书中首次提出“数字鸿沟”这一概念,即各国、地区和行业间因信息及通信技术发展应用的不均衡,导致互联网技术与信息拥有、应用及创新能力的显著差异^[1]。1999年,美国国家远程通信和信息管理局定义了“数字鸿沟”,即“在拥有信息时代的工具的人与未曾拥有者之间存在的鸿沟”^[2]。

数字鸿沟问题与数字技术相伴而生,尤其是在当下互联网浪潮冲击与人工智能技术崛起的背景下欲盖弥彰。“人工智能鸿沟”于2018年美国人工智能协会的春季研讨会上首次提出,被定义为“人工智能相关的不平等接入的现象”^[3]。作为数字鸿沟在现实中的投影,人工智能鸿沟既映射出技术、算法使用中的壁垒与差异,又凸显了社会结构、教育资源分配中的不公。

在人工智能与社会变革的交汇点上,人工智能的发展加速了鸿沟生成速度、加深了鸿沟生成程度、加大了鸿沟生成跨度^[4];人工智能或加剧经济和社会发展鸿沟,冲击劳动市场与劳动者^[5],人工智能这一新型文化资本的冲击加剧了社会分化^[6]。具体

而言,人工智能的应用在各行业、各领域构筑了一系列“准入壁垒”。张黎等^[7]审视了教育领域的智能鸿沟,认为其主要表现为学习者的“系统化区隔”“接榫-断裂”的技术反差、“算法茧房”导致的主体性异化困境、学习者之间的数字资本积累差距进一步扩大等;司法领域,黎慈和孟卧杰^[8]认为“鸿沟”问题导致了司法触达不全面、“算法黑箱”造成司法推论不透明、“技术机械”致使司法审判不合理等诸多问题;社会生活领域,贺苗等^[9]洞察到人工智能应用于养老服务所引发的伦理挑战,将其视为加剧老年群体数字鸿沟的重要因素之一。

因此,学界已广泛认识到人工智能鸿沟的具象存在与深远影响。然而,“人工智能鸿沟”作为一个新兴概念,继承了“数字鸿沟”这一母题的多维性与模糊性,当前学界对于理论层面人工智能鸿沟的存在形式、形成机制等层面的深入解构与系统性分析仍显不足。

本文以“人工智能鸿沟”为切入点,基于上述美国人工智能协会的定义,采用文本分析的方法,对人工智能鸿沟相关研究进行整理与分析,着重回答

收稿日期: 2025-01-06

基金项目: 深圳市卫生健康发展研究和数据管理中心项目(HFW202300711);深圳市卫生健康发展研究和数据管理中心项目(sz20230404)

作者简介: 李方正(2001—),男,湖南常德人,硕士研究生,研究方向为文本挖掘、健康信息学;虢毅(1976—),湖南长沙人,博士,教授,研究方向为生物医学信息;陈佩璇(1997—),广东潮阳人,硕士研究生,研究方向为公共卫生;秦超(1999—),浙江衢州人,硕士研究生,研究方向为医学信息学、公共卫生;段永恒(1978—),黑龙江哈尔滨人,博士,助理研究馆员,研究方向为公共卫生、生物信息、医学人工智能。

了两个问题:①人工智能鸿沟研究的热点主题有哪些?这一问题聚焦于研究领域中的重点议题,旨在对研究领域分散而多样的文献进行总结;②这些主题应如何被细分与解析?这一问题着眼于学者对人工智能鸿沟问题的阐述模式与研究过程,旨在揭示不同主题下的研究重点,全面理解该领域的学术进展,最终实现对人工智能鸿沟的研究热点的系统梳理,深化对该现象多维度特征及其社会影响的学术理解,为人工智能鸿沟的理论深化和实践指导提供了新的研究范式。

1 数据来源与研究方法

1.1 数据来源

本文以 Web of Science 核心合集数据库为文献来源,构建检索式 TS=(“artificial intelligence” OR “AI” OR “machine learning” OR “deep learning” OR “neural network” OR “reinforcement learning” OR “genetic algorithm” OR “support vector machine” OR “SVM” OR “decision tree” OR “random forest” OR “unsupervised learning” OR “swarm intelligence” OR “machine vision” OR “expert system” OR “autonomous system” OR “ensemble learning” OR “deep reinforcement learning” OR “DRL”) AND TS=(“digital inclusion” OR “digital divide” OR “digital inequality” OR “digital equality” OR “digital integration” OR “digital exclusion” OR “Information equity” OR “digital access” OR “digital literacy” OR “digital gap” OR “old people” OR “disabled individuals” OR “countryside”),检索时间为 2024 年 10 月 28 日,对检索结果进行筛选,删除新闻、通知等非学术型文献和与主题不相关的无效性文献研究,得到文献 624 篇。

1.2 研究方法

LDA 模型,即潜在狄利克雷分布(latent Dirichlet allocation)模型,是学者 Blei 于 2003 年提出的一种三层贝叶斯概率模型,包含文档、主题和词三层结构,是一种非监督机器学习技术,可以用来识别文档集或语料库中潜藏的主题信息^[10],其结构拓扑如图 1 所示。

以人工智能鸿沟相关的 624 篇文献的标题、摘要与关键词为语料,利用 Python 中 NLTK 库的 corpus、tokenize、stem 等模块,进行分词、去除标点与停用词、词形还原、词干化等预处理后,输入 LDA 模型,输出模型的主题-词层结果,得到“人工智能鸿沟”相关研究的主题-词概率分布,揭示其核心主题

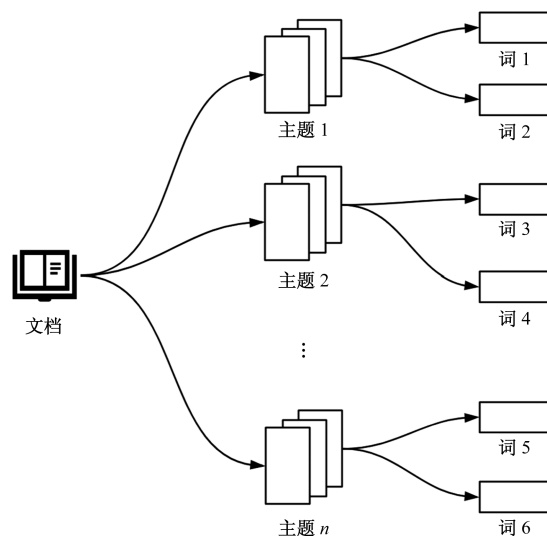


图 1 LDA 模型结构

及其主要词汇,为进一步的主题分析提供坚实的框架。

经参考“术语在线”等相关术语词表,基于主题-词分布,通过人工编码识别“人工智能鸿沟”话题下相关主题。以这些主题为框架,选取并提炼各主题下的代表性文献语料,形成一级编码;再者,比较与关联各一级编码的内容,识别其内在逻辑关系,进行抽象和整合,生成更具概括性的二级编码;最后,对二级编码结果进行挖掘与归纳,形成三级编码,从而实现主题识别与文本分析。具体过程与研究思路如图 2 所示。

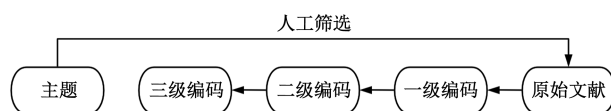


图 2 研究思路

编码过程由具备相关学术背景和丰富专业知识的三位学者共同参与完成,并经过领域内专家的严格审查与评估,确保理论框架具备严谨性与科学性,保证学术质量与可信度。

2 主题识别与文本分析

困惑度(perplexity)是一个用于评估模型质量的度量指标,数值越低表示模型越好。对于 LDA 模型,困惑度反映模型对文档的主题分布的预测准确性。利用 Scikit-learn 库绘制困惑度-主题数曲线(图 3)。由图 3 可知,当主题数为 4 时,困惑度最低,以后随着主题数增加,困惑度上升,且数值逐渐趋于平稳,因此确定最优主题数为 4 个,进而得到主题-词概率分布。以此为依据,对主题提取结果的

4个研究主题进行人工编码,得到“智识壁垒”“数字教育”“医疗卫生”“社会效应”四个主题,具体如表1所示。

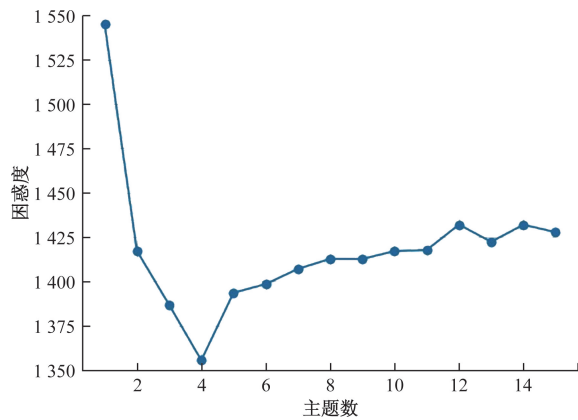


图3 “人工智能鸿沟”相关研究困惑度-主题数曲线

表1 “人工智能鸿沟”研究主题识别结果

序号	主题 (人工编码)	主题词
1	智识壁垒	model; network; predict; result; learn; perform; system; machine; data; control; study; method; algorithm; improve; accuracy
2	数字教育	digital; technology; access; education; develop; learn; data; research; rural; health; intelligent; inform; paper; sustain; industry
3	医疗卫生	health; ai; use; digital; patient; intelligent; healthcare; inform; care; study; develop; educ; result; community; support
4	社会效应	digit; technology; social; educ; use; study; divide; learn; data; develop; research; skill; service; inform; ethic

2.1 智识壁垒

智识,原意为智力、见识和知识。姚娅^[11]认为,相较于知识而言,智识还具有在获得知识后,对社会进行更深入的思考的内涵。主题1的高频关键词有“模型”(model)、“网络”(network)、“系统”(system)、“算法”(algorithm)等,这些词突显了人工智能领域中的技术性与知识素养方面的鸿沟,“学习”(study、learn)也能体现“智识”这一话题。因此,将该主题概括为“智识壁垒”。

表2为该主题下编码结果。结果显示,学者对于该主题的研究集中在技术知识差距与意识差距两方面。

人工智能鸿沟问题的核心在于技术的不平等,技术知识差距催生了人工智能鸿沟的形成。人工智能的技术主导者主要是研究机构与科技企业,资本力量在技术困境中占主导地位^[18];而一般用户由

于缺乏对算法的深入理解、对数据采集与处理方法的不熟悉,以及对于模型训练与优化的知识匮乏等原因,从而难以直接参与人工智能系统的建构与实现。人工智能的使用可及性的不均,体现为不同学历、不同经济阶层、不同年龄人群对于人工智能的“知识差距”,这一差距导致了资源和知识的不对称分布,加剧了人工智能鸿沟的扩大。

表2 智识壁垒编码结果

三级编码	二级编码	一级编码
技术知识 差距	技术壁垒 (A ₁)	大多数中小型企业缺乏利用最新数字创新的知识 ^[12] 和能 ^[12] (a ₁)
		人工智能素养是数字素养的一个子集,涉及批判性评估人工智能技术、与人工智能有效沟通和协作的能力 ^[13] (a ₂)
	使用机会 (A ₂)	并非所有人都能平等获得人工智能技术的好处 ^[13] (a ₃)
		社交媒体暴露和信息精细化是影响人工智能知识差距的重要因素 ^[14] (a ₄)
意识差距	心理接受 (A ₃)	人工智能的可解释性对用户的信任度至关重要 ^[15] (a ₅)
		人工智能的技术采用引发了员工对未来职业的不安全感 ^[16] (a ₆)
	意识态度 (A ₄)	用户对人工智能的态度起着重要作用 ^[17] (a ₇)

技术知识差距固然是构成智识壁垒形成的起点,但智识壁垒并非由技术知识差距单向驱动建构,另一个不容忽视的方面是技术问题所衍生的“意识差距”^[19],即认知、知识和思维等方面的差距。弱势群体对人工智能技术本身的非中立性、最终所导致的对弱势群体的剥削与伤害,以及对优势群体利益的维护,并无明确的认识^[20]。Wang等^[17]认为,首当其冲的弱势群体是“无特长怀疑者”(unskilled skeptics)和“中立无特长者”(neutral unskilled),他们年龄较大、教育水平较低,且缺乏隐私保护技能,对人工智能持消极态度,并认为随着时间的推移,弱势群体可能会持续存在。

总而言之,智识壁垒的形成是一个多维度的复杂过程,其内涵也对应了人工智能鸿沟作为第三层数字鸿沟建构的底层驱动力,即技术知识差距、意识差距以及由此产生的社会不平等现象。

2.2 数字教育

数字教育的兴起意味着教育领域对于技术的需求日益增长。作为实现数字资源平等共享的关键手段,数字教育对于弥合人工智能鸿沟具有重要意义。人工智能作为一种前沿技术,正在渗透到教育的各个方面。然而,这种渗透并非均匀,而是存

在着一定的不平衡。主题 2 中包含“数字”(digital)、“获取”(access)、“学习”(learn)、“教育”(education)等关键词,因此推测该主题为“数字教育”。

2018 年 1 月 17 日,欧盟委员会(European Commission)通过的《数字教育行动计划》中,指出“数字教育”涵盖运用数字技术开展教与学、培养学习者数字能力的发展两方面内容^[21]。因此,在数字教育主题下,选取“数字教学开展”“数字能力培养”两方面为框架与子主题,进行文本分析。

2.2.1 数字教学开展

随着人工智能技术的发展,尤其是生成式人工智能(generative artificial intelligence, GAI)的兴起,教育行业的范式正在迅速转变,人工智能已以不同形式广泛应用于教育行业,如在线智能教育系统、履行教师职责的聊天机器人等^[22]。表 3 为数字教学开展的编码结果,结果显示人工智能对教育的帮助是显著的,既能提高教育的效率,也能优化教学质量,在作业审查、个性化学习体验、辅助学习等方面皆有应用。

与此同时,这种转变亦对建立更具包容性的教育系统带来了风险与挑战。技术的不平等获取和有效使用能力导致难以获得、利用这些智能教育资源的学生进一步被边缘化,进一步扩大了数字教育的不平等与排斥性。正如 Bulathwela 等^[23]指出,人工智能等辅助技术能决定残障人士、留守儿童、低收入者等群体在数字教育中的参与程度,增加其接受公平而高质量的数字教育的机会,但囿于技术的可获得性,人工智能往往会加剧数字鸿沟的深化,而不会使数字教育更加普及。

表 3 数字教学开展编码结果

三级编码	二级编码	一级编码
现状与优势	行政职能(B ₁)	人工智能在审查和评分学生作业等方面提供了显著的帮助 ^[22] (b ₁)
	教学质量(B ₂)	人工智能通过提供个性化的学习体验,增强了教学和学习的质量和效果 ^[22] (b ₂) 生成式人工智能辅助编程任务、完成家庭作业、准备课堂内容等 ^[24] (b ₃)
局限与前瞻	风险挑战(B ₃)	生成式人工智能在教学中可能导致预设的意识形态立场与偏见 ^[25] (b ₄)
		确保技术的公平性和可访问性是人工智能在教育中需要进一步解决的问题 ^[22-23] (b ₅) (人工智能)导致学术不诚信、批判性思维技能可能被侵蚀 ^[24] (b ₆)
	未来构想(B ₄)	增加对数字访问、算法等方面鸿沟的责任和控制可以创造一个良性循环,改善教育的多样性、公平性和包容性 ^[23,26] (b ₇)

2.2.2 数字能力培养

数字能力培养编码结果如表 4 所示,该主题下学者较关注教育中数字技能、学习能力、社会能力、批判思维与自我效能等方面能力的培养。在数字教育的宏观体系与数智赋能的大环境下,重视学生这些数字能力的培养不仅有助于学生的发展,也为弥合人工智能鸿沟提供个体层面的解决方案。

表 4 数字能力培养编码结果

三级编码	二级编码	一级编码
数字行为	数字技能(C ₁)	教学的数字能力调节了信息和通信技术的采用、人工智能和电子学习系统之间的关系 ^[27] (c ₁)
		使用 ChatGPT 可以为学生提供实时的行业信息,提高他们的数字素养技能 ^[28] (c ₂)
	学习能力(C ₂)	人工智能能力对创造力和学习表现有正向影响 ^[29] (c ₃)
数字思维	社会能力(C ₃)	软件·人工智能教学营展示了在积极融入当地社区和承担重要社会角色方面的潜力 ^[30] (c ₄)
		培养批判性思维技能将使个人能够根据准确可靠的信息做出更好的决策,并在各个领域与人工智能有效合作 ^[31] (c ₅)
	自我效能(C ₅)	高等教育机构的人工智能能力显著影响学生的自我效能感和创造力 ^[29] (c ₆)

然而,培养个体的数字能力仅仅是解决问题的一部分,人工智能鸿沟形成的根本原因之一在于不同社会群体在获取和利用人工智能上的不平等。因此,学者们提出了开放教育资源、加强生成式人工智能技术在个性化学习中的应用、建立结果导向的人工智能素养评价框架等建议,致力于在人工智能的背景下构建一个可持续、规模化且具有包容性的数字教育生态系统,从群体层面弥合鸿沟。

2.3 医疗卫生

主题 3 中包含了“健康”(health)、“医疗保健”(healthcare)、“病人”(patient)等关键词,涉及人工智能在医疗卫生行业中的应用,因此将本主题概括为“医疗卫生”。在探讨人工智能在医疗卫生领域的渗透的影响时,可以观察到这一过程呈现显著的二重性,即人工智能技术的应用在促进行业的发展的同时,同样会由于其可及性的不平等,加剧行业的数字鸿沟,间接地延缓了行业的整体发展。从表 5 的编码结果可以看出,学者对该主题的关注集中在具体的技术应用与行业内人工智能鸿沟的形成两方面。

从 21 世纪初医疗数据的数字化和标准化,电子健康记录(electronic health records, EHR)开始普及,到 2010 年代深度学习和强化学习等技术被广泛

应用,人工智能技术开始用于数据分析和预测模型的构建,再到如今人工智能技术在医疗流程中的集成,智能化医疗系统落地,人工智能技术的应用正在遍及整个医疗卫生行业。当下,人工智能被广泛应用于个人助理、临床支持、数据整合等方面,为医疗卫生行业带来了理论、技术与思维的突破。

表 5 医疗卫生编码结果

三级编码	二级编码	一级编码
技术应用	个人助理 (D ₁)	通过整合人工智能驱动的伴侣、虚拟社交环境和远程医疗平台应对孤立老年人的口腔健康管理和社会脱节问题 ^[33] (d ₁)
	临床支持 (D ₂)	(人工智能……)提高筛查过程的效率,建立预测模型,预测发病率 ^[34-35] (d ₂) 人工智能可以用于风险分层、神经诊断支持以及图像识别技术 ^[34,36] (d ₃)
	数据整合 (D ₃)	人工智能可以帮助整合不同和异构的数据领域,促进研究和发现,增加临床工作流程,并可能使医疗保健更加参与性 ^[37] (d ₄) 利用人工智能作为沟通/整合者,改善精准医疗和社会健康决定因素之间的联系 ^[38] (d ₅)
鸿沟形成	伦理问题 (D ₄)	(人工智能导致的)不良医疗事件的责任的程度尚未在实践中体现出来 ^[35] (d ₆)
	数据访问不平等 (D ₅)	如果医疗系统在设计和实施人工智能解决方案时不考虑所有群体的需求,可能会导致某些群体被边缘化 ^[36,39] (d ₇) 牙科数据的可用性和可访问性有限,数据处理和结果验证常常不够可复制和稳健 ^[37] (d ₈)

作为人工智能时代重要的社会现象,人工智能鸿沟在医疗卫生领域同样显著存在。根据编码结果与文本分析,本文总结了医疗卫生行业内人工智能鸿沟递进式形成机制(图 4)。人工智能与医疗卫生行业的交互过程中,数字不平等与医疗不公共同塑造了人工智能鸿沟。Rotz 等^[32]指出,“在数据驱动算法实施的过程中,人工智能往往倾向于强化既有的结构性不平等”。在医疗卫生领域中,这一现象体现为医疗数据的异质性和地域性差异限制了人工智能模型的泛化能力,进而导致算法性能在不同人群中的显著差异。边缘群体的数字排斥构成“数据缺口”,有限的训练集导致人工智能对于边缘群体的医疗普及性与适用性受限,最终形成鸿沟。

人工智能赋能医疗卫生行业实现了从数据整合、临床诊断到个性化治疗的全面革新,提升了医疗服务的精准度和效率。同时,由数据、算法与医疗行业本身所导致的人工智能鸿沟问题仍需重视,确保人工智能技术的公平应用,推动健康信息学、精准医疗和医疗伦理的协同发展,以真正实现医疗卫生服务的智能化和普惠化。

2.4 社会效应

人工智能在数智环境中的泛在化不仅体现在其在各垂直领域内深度的嵌入性,还体现于在多个社会层面中的全面扩散。主题 4 包含“社会”(social)、“鸿沟”(divide)、“发展”(develop)、“伦理”

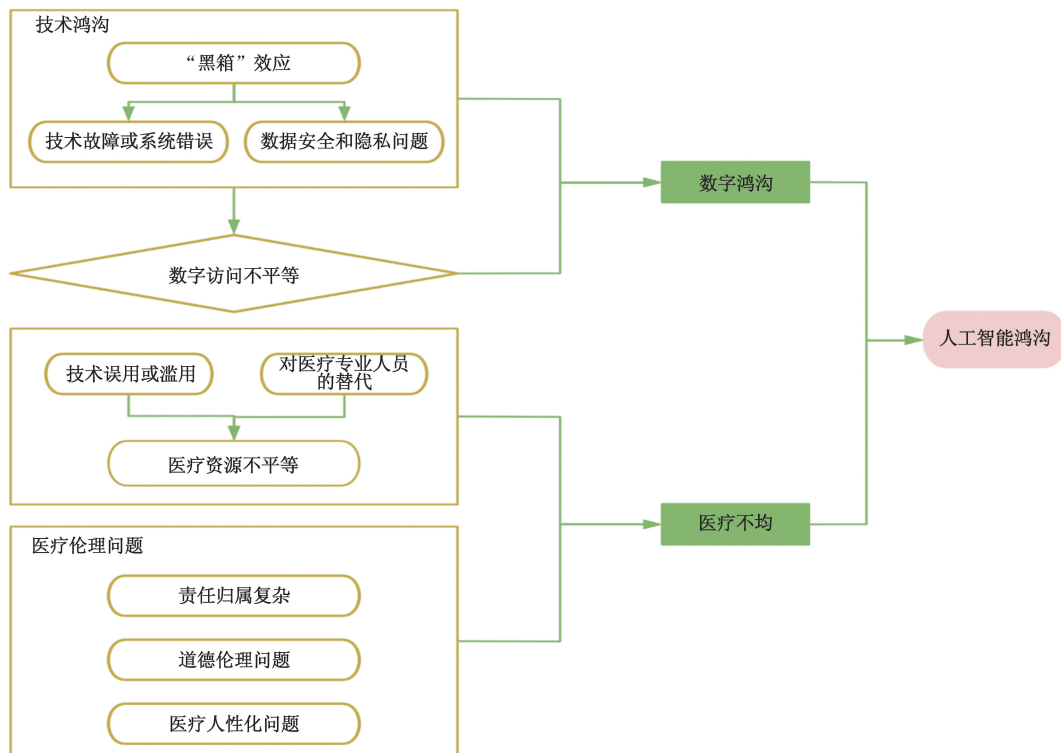


图 4 医疗卫生行业人工智能鸿沟形成机制

(ethic)等词,涉及人工智能对社会的影响,故将此主题总结为“社会效应”,进行编码,结果如表 6 所示,人工智能鸿沟的社会效应体现在伦理文化与社会生活等层面中。

表 6 社会效应编码结果

三级编码	二级编码	一级编码
伦理文化 困境	伦理失范 (E ₁)	强调了人工智能应用中必须考虑的伦理问题,包括数据隐私、算法偏见、透明度和公平性 ^[42] (e ₁)
	文化趋同 (E ₂)	人工智能在文化领域的应用可能因为商业逻辑导致文化产品和表达形式的集中化,对文化多样性产生负面影响 ^[41] (e ₂)
社会 不平等	数字排斥 (E ₃)	数据标准化、基础设施限制、数字素养和算法偏见等是可能阻碍公平获取人工智能益处的挑战 ^[43] (e ₃)
	社会生活 不公(E ₄)	人们越来越意识到人工智能技术的风险和 挑战,尤其是在加剧现有不平等和失衡 以及对人类权利的影响方面 ^[41] (e ₄) 大语言模型可能导致环境影响、不公平的 劳动实践、意外的错误信息或“幻觉”、偏 见的扩散和版权侵犯等 ^[43] (e ₅)

人工智能导致的伦理问题部分源自人机交互过程中主体的实践和认知被淡化^[40]。随着人工智能的自动化与自主性增强,人类在决策过程中的参与度和控制权逐渐被削弱,导致实践主体性和认知主体性的消解,进而引发对算法透明度和可解释性的质疑。主体能动性的边缘化,使得人工智能系统在执行复杂任务时,难以全面考虑伦理规范和社会价值,从而导致伦理困境的产生和扩散。

另外,GAI 重塑了信息的生产与分发机制,在一定程度上影响了社会文化的多样性与复杂性。在信息传播过程中,GAI 的使用面临着商业逻辑导向的风险^[41]——由于主流内容具备更高的市场需求和广告收益,生成式人工智能算法倾向于优先推荐这些内容,以最大化平台流量和盈利。这种偏向进一步通过大量用户数据和互动行为强化了推荐算法,限制了边缘和多样文化内容的曝光机会,从而减少了社会中非主流文化表达形式的可见性与传播,最终导致社会文化生态被破坏。

人工智能鸿沟的社会效应不仅停留在抽象的社会哲学与文化讨论,而是已经具象化为多层次的社会不公,技术决定论的蓝本在人工智能的时代背景下仍然有所体现,如数据获取的不平等、自动化带来的结构性失业、大型数据中心的运行所伴随的巨大的能源消耗和碳排放问题、数据剥削与隐私侵犯导致的人类权利受损等。在人工智能时代,资源

稀缺性成为数字化环境中的关键挑战,这种稀缺性不仅限制了技术的普及性和公平性,也在全球范围内加剧社会的不平等现象。人工智能技术所需的高度集中的数据和计算能力资源深刻影响着不同社会群体的发展轨迹和生活状态,催生了新的社会分化。

3 结论与展望

本文采用 LDA 模型对“人工智能鸿沟”进行了主题分析,揭示了人工智能在不同领域应用的复杂性和多维性。通过文本分析,识别并归纳了四个核心研究主题,即“智识壁垒”“数字教育”“医疗卫生”“社会效应”,揭示了技术知识差距、意识差距、数字教育不平等以及人工智能在医疗卫生领域的应用不公等现状与差距;同时,人工智能鸿沟的社会效应,包括伦理文化困境、文化趋同和社会不平等问题,亦是影响社会结构和个体生活的重要因素。

随着人工智能深度融入社会经济结构,“人工智能鸿沟”揭示了技术与社会系统的复杂互动,同时其成因与危害也逐渐被感知、被重视。面向未来,一方面,技术对社会权力结构的重新定义是人工智能鸿沟的本质问题。在这一语境下,消除智识壁垒、普及数字教育、推动人工智能在各领域的普惠性等,均是实现个体与社会的双重变革,从而解决问题的关键与具体实践路径。另一方面,技术急变可能颠覆社会主体的原生认知机制^[44]。助推这种“意识觉醒”将促进社会走向更具反思性和责任性的社会文化,倡导技术进步与人类价值的协调统一,确保人工智能等新兴技术真正成为人类福祉和社会进步的助力,而不是造成新的不平等和不公正的根源。

参考文献

- [1] [美]阿尔文·托夫勒. 权力的转移[M]. 黄锦桂,译. 北京: 中信出版社, 2018.
- [2] 郭庆光. 传播学教程[M]. 北京: 中国人民大学出版社, 2011.
- [3] 方伟. 人工智能鸿沟: 问题现状与治理思路[J]. 今日科苑, 2020(3): 81-87.
- [4] 杨峰. 生成式人工智能的智能鸿沟生成[J]. 图书馆建设, 2023(4): 22-24.
- [5] 梁桐. 人工智能或加剧社会发展鸿沟[N]. 经济日报, 2024-08-12(004).
- [6] 陈龙, 陈泽. 新型文化资本与智能时代“技术鸿沟”的生成[J]. 暨南学报(哲学社会科学版), 2024, 46(10): 148-165.
- [7] 张黎, 鲍文雨, 赵磊磊. “智能鸿沟”的教育镜像: 教育数

- 数字化转型的底层视角[J]. 现代教育技术, 2024, 34(7): 51-60.
- [8] 黎慈, 孟卧杰. 数字正义视域下数字司法建设面临的障碍与对策研究[J]. 湖北社会科学, 2023(11): 117-126.
- [9] 贺苗, 侯如意, 尹梅. 人工智能养老的伦理困境反思[J]. 医学与哲学, 2024, 45(6): 27-30.
- [10] BLEI D M, NG A Y, JORDAN M I. Latent Dirichlet allocation[J]. *Journal of Machine Learning Research*, 2003, 3(4/5): 993-1022.
- [11] 姚妮. 寓教于器: 民国时期云南展览会与社会智识的传播[D]. 昆明: 云南师范大学, 2024.
- [12] CHOUDRIE J, MANANDHAR N, CASTRO C, et al. Hey siri, Google! can you help me? a qualitative case study of smartphones AI functions in SMEs[J]. *Technological Forecasting and Social Change*, 2023, 189: 122375.
- [13] GOLDENTHAL E, PARK J, LIU S X, et al. Not all AI are equal: exploring the accessibility of AI-mediated communication technology[J]. *Computers in Human Behavior*, 2021, 125: 106975.
- [14] LI W, XU S, ZHENG X, et al. Bridging the knowledge gap in artificial intelligence: the roles of social media exposure and information elaboration[J]. *Science Communication*, 2024, 107: 232352.
- [15] KOBRINSKII B A. Artificial intelligence: problems, solutions and prospects[J]. *Pattern Recognition and Image Analysis*, 2023, 33(3): 217-220.
- [16] MATSUNAGA M. Uncertainty management, transformational leadership and job performance in an AI-powered organizational context[J]. *Communication Monographs*, 2022, 89(1): 118-139.
- [17] WANG C, BOERMAN S C, KROON A C, et al. The artificial intelligence divide: who is the most vulnerable? [J]. *New Media & Society*, 2024, 146: 232345.
- [18] 钟祥铭, 方兴东. 智能鸿沟: 数字鸿沟范式转变[J]. 现代传播(中国传媒大学学报), 2022, 44(4): 133-142.
- [19] ZAROUALI B, BOERMAN S C, VREESE C H. Is this recommended by an algorithm? the development and validation of the algorithmic media content awareness scale (AMCA-Scale) [J]. *Telematics and Informatics*, 2021, 62: 101607.
- [20] O'NEIL C. Weapons of math destruction: how big data increases inequality and threatens democracy[M]. New York: Crown, 2016.
- [21] 董丽丽, 金慧, 李卉萌, 等. 后疫情时代的数字教育新图景: 挑战、行动与思考——欧盟《数字教育行动计划(2021—2027年)》解读[J]. 远程教育杂志, 2021, 39(1): 16-27.
- [22] CHEN L, CHEN P, LIN Z. Artificial intelligence in education: a review[J]. *Ieee Access*, 2020, 8: 75264-75278.
- [23] BULATHWELA S, PEREZ-ORTIZ M, HOLLOWAY C, et al. Artificial intelligence alone will not democratise education: on educational inequality, techno-solutionism and inclusive tools[J]. *Sustainability*, 2024, 16(2): 781.
- [24] WALCZAK K, CELLARY W. Challenges for higher education in the era of widespread access to generative AI[J]. *Economics and Business Review*, 2023, 9(2): 71-100.
- [25] BENDER S M. Awareness of artificial intelligence as an essential digital literacy: ChatGPT and gen-AI in the classroom[J]. *Changing English-Studies in Culture and Education*, 2024, 31(2): 161-174.
- [26] DIETERLE E, DEDE C, WALKER M. The cyclical ethical effects of using artificial intelligence in education [J]. *AI and Society*, 2024, 39(2): 633-643.
- [27] KHAN W, SOHAIL S, ROOMI M A, et al. Opening a new horizon in digitalization for e-learning in malaysia: empirical evidence of covid-19[J]. *Education and Information Technologies*, 2023, 29(8): 9387-9416.
- [28] DALGIC A, YASAR E, DEMIR M. ChatGPT and learning outcomes in tourism education: the role of digital literacy and individualized learning [J]. *Journal of Hospitality Leisure Sport & Tourism Education*, 2024, 34: 100481.
- [29] WANG S, SUN Z, CHEN Y. Effects of higher education institutes' artificial intelligence capability on students' self-efficacy, creativity and learning performance [J]. *Education and Information Technologies*, 2023, 28(5): 4919-4939.
- [30] CHOI E, KIM J, PARK N. A case study of SW•AI education for multicultural students in Jeju, Korea: changes in perception of SW•AI[J]. *Applied Sciences-Basel*, 2023, 13(17): 9844.
- [31] RUSANDI M A, AHMAN, SARIPAH I, et al. No worries with ChatGPT: building bridges between artificial intelligence and education with critical thinking soft Skills[J]. *Journal of Public Health*, 2023, 45(3): 602-603.
- [32] ROTZ S, GRAVELY E, MOSBY I, et al. Automated pastures and the digital divide: how agricultural technologies are shaping labour and rural communities [J]. *Journal of Rural Studies*, 2019, 68: 112-122.
- [33] QI X, WU B. AI's role in improving social connection and oral health for older adults: a synergistic approach [J]. *JDR Clinical & Translational Research*, 2024, 238: 223097.
- [34] KWOK T C, HENRY C, SAFFARAN S, et al. Application and potential of artificial intelligence in neonatal medicine[J]. *Seminars in Fetal and Neonatal Medicine*, 2022, 27(5): 101346.
- [35] RAMESSUR R, RAJA L, KILDUFF C L S, et al. Impact and challenges of integrating artificial intelligence

- and telemedicine into clinical ophthalmology[J]. *Asia-Pacific Journal of Ophthalmology*, 2021, 10(3): 317-327.
- [36] SADLER D, OKWUOSA T, TESKE A J, et al. Cardio oncology: digital innovations, precision medicine and health equity[J]. *Frontiers in Cardiovascular Medicine*, 2022, 9: 951551.
- [37] SCHWENDICKE F, SAMEK W, KROIS J. Artificial intelligence in dentistry: chances and challenges [J]. *Journal of Dental Research*, 2020, 99(7): 769-774.
- [38] KWON W G, KIM S H, MARTIN D. Integrating social determinants of health to precision medicine through digital transformation; an exploratory roadmap [J]. *International Journal of Environmental Research and Public Health*, 2021, 18(9): 5018.
- [39] WILLIAMS D, HORNUNG H, NADIMPALLI A, et al. Deep learning and its application for healthcare delivery in low and middle income countries[J]. *Frontiers in Artificial Intelligence*, 2021, 4: 553987.
- [40] 柏杨, 李晓. 生成式人工智能辅助公共意见传播的空间叙事——伦理约束与路径建构[J]. *华侨大学学报(哲学社会科学版)*, 2024(3): 133-147.
- [41] TSVYK V A, TSVYK I V. Social issues in the development and application of artificial intelligence[J]. *Rudn Journal of Sociology*, 2022, 22(1): 58-69.
- [42] FLORES-VIVAR J M, GARCIA-PENALVO F J. Reflections on the ethics, potential, and challenges of artificial intelligence in the framework of quality education (SDG4)[J]. *Comunicar*, 2023, 31(74): 37-47.
- [43] ONG J C L, SENG B J J, LAW J Z F, et al. Artificial intelligence, ChatGPT and other large language models for social determinants of health: current state and future directions[J]. *Cell Reports Medicine*, 2024, 5(1): 101356.
- [44] 詹希旒, 李白杨. 生成式人工智能背景下的数智可供性与认知带宽调节研究[J]. *图书与情报*, 2024(1): 110-120.

Multi Dimensional Thematic Analysis of the “Artificial Intelligence Divide” Based on LDA

LI Fangzheng¹, GUO Yi¹, CHEN Peixuan², QIN Chao², DUAN Yongheng³

(1. School of Life Sciences, Central South University, Changsha 410031, China;

2. School of Public Health, Jilin University, Changchun 130021, China;

3. Shenzhen Health Development Research and Data Management Center, Shenzhen 518000, Guangdong, China)

Abstract: With the development of artificial intelligence technology, the unequal access to artificial intelligence, known as the “AI divide” phenomenon, is gradually becoming prominent. Using literature from the Web of Science database as corpus and identifies relevant topics under the topic of “Artificial Intelligence Divide” based on LDA model, text analysis was conducted through manual encoding. Finally, four core research themes such as intellectual barriers, digital education, healthcare and social effects are identified and summarized, the current situation and gaps in technological knowledge gaps, awareness gaps, digital education inequality and the unfair application of artificial intelligence in the healthcare field is revealed. Meanwhile, the social effects of the artificial intelligence divide, including ethical and cultural dilemmas, cultural convergence and social inequality, are also important factors affecting social structure and individual lives. This article aims to deepen the understanding of the core theme of the “artificial intelligence divide” and help artificial intelligence enhance human well-being and social progress.

Keywords: artificial intelligence divide; digital divide; LDA; text analysis