

生成式人工智能伦理失范的多维透视

王艺涵

新疆大学马克思主义学院, 新疆 乌鲁木齐

收稿日期: 2026年3月16日; 录用日期: 2026年4月6日; 发布日期: 2026年4月20日

摘要

随着生成式人工智能技术的迅猛发展, 其在诸多领域得到广泛应用, 但同时也引发了一系列复杂的伦理失范。本文从学术造假这一突出问题入手, 深入剖析其如何引发社会信任危机, 并从技术、法律、教育、文化等多维视角对生成式人工智能的伦理失范进行全面审视。研究发现, 生成式人工智能在带来便利与创新的同时, 在学术规范、社会信任、法律责任界定、教育公平与价值观引导以及文化价值观融合等方面面临严峻挑战。为应对这些困境, 本文提出了加强技术研发与伦理约束、完善法律法规、深化教育改革以及促进文化交流与价值观融合等策略, 旨在推动生成式人工智能在遵循伦理原则的基础上实现健康、可持续发展。

关键词

生成式人工智能, 伦理失范, 学术造假, 社会信任危机, 多维透视

Multidimensional Perspective of Generative Artificial Intelligence Ethics

Yihan Wang

School of Marxism, Xinjiang University, Urumqi Xinjiang

Received: March 16, 2026; accepted: April 6, 2026; published: April 20, 2026

Abstract

With the rapid development of generative artificial intelligence technology, it has been widely used in many fields, but it has also caused a series of complex ethical deviations. Starting from the prominent problem of academic fraud, this article deeply analyzes how it causes a social trust crisis, and comprehensively examines the ethical deviation of generative artificial intelligence from the multidimensional perspectives of technology, law, education and culture. The study found that while bringing convenience and innovation, generative artificial intelligence faces serious challenges in terms of

academic norms, social trust, definition of legal responsibilities, educational equity and value guidance, and the integration of cultural values. In order to cope with these difficulties, this article puts forward strategies such as strengthening technological research and development and ethical constraints, improving laws and regulations, deepening educational reform, and promoting cultural exchanges and integration of values, aiming to promote generative artificial intelligence to achieve healthy and sustainable development on the basis of ethical principles.

Keywords

Generative Artificial Intelligence, Ethical Deviation, Academic Fraud, Social Trust Crisis, Multi-Dimensional Perspective

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人工智能也不例外，人工智能作为引领新一轮科技革命和产业变革的战略性技术，正深刻地改变人类生产生活方式。生成式人工智能作为人工智能领域的前沿技术，近年来取得了突破性进展。以 ChatGPT、Sora 等为代表的生成式人工智能模型，能够根据输入的提示生成高质量的文本、图像、视频等内容，在自然语言处理、创意设计、教育、科研等众多领域展现出巨大的应用潜力。然而，随着其广泛应用，一系列伦理失范逐渐浮现，其中学术造假问题尤为突出，并进一步引发了社会信任危机。学术造假不仅损害了学术研究的公正性和严肃性，也破坏了社会对学术成果和科学研究的信任基础。因此，深入研究生成式人工智能引发的伦理失范，特别是从学术造假到社会信任危机的演变路径，并从多维视角进行全面剖析，对于制定有效的应对策略，促进生成式人工智能的健康发展具有重要的理论和实践意义。

2. 问题缘由与研究现状

2.1. 问题缘由

近年来，生成式人工智能在学术活动中也在不断得到广泛运用。调查显示，近一半的剑桥大学学生承认他们在学业中使用 ChatGPT [1]，近四分之三的浙江大学学生在科研活动或课程学习中使用生成式人工智能工具[2]。2023 年 9 月，一篇学术论文因出现“Regenerate response”字样而被撤稿，这是第一例因使用 ChatGPT 而被撤稿的 SCI 论文。2023 年 2 月，某校宣布禁止在学校中使用 ChatGPT 或其他生成式人工智能工具，但哈佛大学鼓励使用生成式人工智能技术进行负责人的研究[3]。学术机构在生成式人工智能应用问题上呈现的矛盾立场，折射出当前学界在面对技术冲击时普遍存在的被动性与应对滞后。这种现实迫切要求学术界展开系统性、前瞻性的研究，以形成科学有效的回应策略。必须清醒认识到，生成式人工智能融入学术实践已成为不可逆转的趋势，关键在于系统厘清其可能引发的学术伦理风险谱系，并以此为基础建立与之相适应的使用规范与治理框架。

2.2. 研究现状

生成式人工智能作为科学研究的辅助工具，在学术领域应具有很大的发展前景。主要包括：(1) 启发研究思路、设计研究方案；(2) 分析文献、数据，提升研究效率，如朱永新和杨帆以案例展示了 ChatGPT 可以协助研究者编程，极大提高了传统编程效率，从而加快科研创新进度；(3) 辅助学术写作改进语法、

自动翻译、生成语句等，如 Berg 指出研究者既可以通过被动方式从生成式人工智能快速获取内容、花费最少的精力来完善或发展论点，也可以通过苏格拉底式的方法，与它进行动态对话，以迭代地发展和完善他们的论点[4]。

已有研究多聚焦于使用者视角，通过实践观察与经验归纳，揭示了生成式人工智能对既有学术伦理体系的冲击，但普遍存在两方面局限：一是未能充分揭示该技术所引发伦理问题的特殊性与结构性，二是较少从学术机构作为治理主体的现实关切出发进行系统性回应。为弥补上述研究空白，本研究以生成式人工智能伦理规范多维度进行审视，系统归纳伦理失范重点关注的五类伦理议题，进而结合理论与文献，深入剖析其生成机制与内在特征，最终构建具有针对性与可操作性的治理路径。

3. 生成式人工智能伦理失范的具体呈现

这些数据可能包含用户的个人信息，如果数据收集和使用不当，就会导致用户隐私。这表明如何面对生成式人工智能带来的伦理失范成为一个非常迫切的话题。生成式人工智能在学术活动中的应用已成为不可逆转的时代潮流，因此我们首先要厘清生成式人工智能会带来哪些学术类型问题，才能进一步规范该技术的使用。生成式人工智能的伦理失范，已从潜在的治理议题，演变为一系列具体、广泛且相互关联的现实挑战。

(一) 学术诚信体系的冲击与知识生产失序。生成式人工智能首先对以诚信为核心的知识生产体系构成直接冲击。其能自动化生成论文、研究报告乃至编程代码，一方面极大降低了学术不端的实施门槛，催生了从观点剽窃、数据伪造到全文代写的多层次造假行为，使传统的查重与评审机制部分失效；另一方面，其生成的“无源内容”模糊了原创与模仿的界限，冲击了传统的知识产权与学术贡献认定框架，导致知识生产的源头性失序。这不仅是技术滥用问题，更动摇了学术共同体赖以生存的求真基石。

(二) 个人隐私边界的消融与数据权利侵蚀。在数据获取与模型训练层面，生成式人工智能构成了对个人隐私的系统性挑战。其训练依赖于海量的、多源异构的互联网数据，其中不可避免地包含未经脱敏处理的个人信息、敏感对话乃至生物特征数据。这不仅在收集阶段可能存在“告知-同意”原则的规避，更在于模型具备从看似非敏感的分散数据中重组、推断出个人敏感信息(如身份、健康状况、行踪)的“二次挖掘”能力，从根本上侵蚀了数据最小化与目的限定原则，导致个人隐私的边界在无形中被持续消解。

(三) 算法偏见的多重固化与社会公正威胁。需要注意的是，在这里，算法更多地被理解为“在一个特定的程序、软件或信息系统中一个或多个算法的实现和交互，而不是一种数学结构”[4]。生成式人工智能并非价值中立，其输出深刻反映了训练数据中蕴含的、甚至被放大的社会偏见与结构性不平等。这种算法歧视体现在多个层面：一是在内容生成中固化性别、种族、地域等刻板印象；二是在辅助决策场景(如招聘简历筛选、信贷评估、司法风险评估)中，系统性地区别对待特定群体，将历史性不公编码为看似客观的算法结论；三是因其输出具有“拟真权威性”，可能使歧视性结果更难以被识别和挑战。这不仅是技术偏差，更是对社会公平正义原则的自动化威胁。

(四) 信息生态的污染与公共信任的解构。生成式人工智能以极低成本、高逼真度批量生产文本、图像、音视频内容的能力，使其成为虚假与误导信息的“超级工厂”。这严重污染了信息生态：一是制造并传播难以甄别的虚假新闻与政治宣传，操纵公众舆论；二是生成用于诈骗的个性化欺诈内容；三是制造大量的“信息噪音”，挤压真实、优质信息的传播空间。其后果是侵蚀公众对媒体、机构乃至客观事实的基本信任，动摇理性公共讨论的基础，对民主社会与公共安全构成深层威胁。

(五) 知识产权的结构性困境与创造性异化。在知识产权领域，生成式人工智能引发了前所未有的纠纷与困境。一是权利归属模糊：AI生成内容的“作者”是开发者、使用者还是模型自身，法律上尚无定论；二是侵权认定困难：其输出是对海量受版权保护数据的融合性模仿，构成合理使用还是侵权难以界

定；三是创造性价值被异化：当 AI 能高效生成“合格”甚至“优秀”的标准化文化产品时，人类的原创性劳动与创造性价值可能被贬抑，长远看可能抑制文化创新的活力。

以上五个维度并非孤立存在，而是相互交织、彼此强化。学术失范削弱知识权威，信息污染破坏信任基础，隐私侵犯与算法歧视侵蚀个体权利与公平，知识产权困境则扰动创新生态。它们共同勾勒出生成式人工智能伦理失范的复杂图景，也指明了治理必须应对的系统性挑战。

4. 生成式人工智能引发的学术造假问题剖析

那么像这种由于人工智能引发的算法的伦理问题需要进一步剖析吗？答案可能是肯定的。就像多明戈斯所指出的，“只有工程师和机修工有必要知道汽车发动机如何运作，但每位司机都必须明白转动方向盘会改变汽车的方向、踩刹车会让车停下来”[5]对于生成式人工智能引发的学术造假问题的分析，并不是由某一方面单独引起的，是多方面作用的结果。接下来将从生成式人工智能本身、使用者以及管理者三个维度进行探析。

首先从生成式人工智能本身来讲。体现在学术造假现象的表现形式上。生成式人工智能在学术领域的应用为学术造假提供了新的手段。常见的学术造假现象包括论文抄袭、数据篡改、虚假研究报告生成等。一些学生和研究人员利用生成式人工智能快速生成论文内容，直接抄袭或稍作修改后提交，严重违反了学术诚信原则。在数据处理方面，生成式人工智能可以被用于伪造实验数据，使研究结果看起来更加“理想”，从而误导学术界和公众。此外，还有人利用生成式人工智能生成虚假的研究报告，用于申请科研项目、获取学术荣誉等。其次进一步探析学术造假的成因分析。技术层面：生成式人工智能的易用性是导致学术造假的一个重要原因。由于剽窃生成式人工智能输出的内容难以被检测技术工具精准识别。与简单地抄录他人内容的传统或基本剽窃相比，生成式人工智能工具有强大的转述能力，它们可以改变整个段落的句法和语法结构，但仍采用了转述来源中包含的观点，并向使用者呈现出省略了对参考来源的引用[6]。这就相当于生成式人工智能在一定程度上代替使用者对文本进行加工，又能规避论文查重带来的风险，这就使得这种新式剽窃很难被发现。

其次，从生成式人工智能的使用者来说。自 ChatGPT 的诞生到 Deepseek 的火爆出圈，被 AI 取代的焦虑，在全球高校研究者和学生之间愈演愈烈。人文社科界，正在经历一场静默的身份危机。当 AI 能够迅速完成史料爬梳、自动生成符合期刊规范的政治学模型，对哲学思想进行精准无误的权威解读时，研究者们遭遇着对自我价值的普遍怀疑。因此学术压力是促使部分人进行学术造假的内在动力。在当前的学术评价体系中，论文发表数量和质量往往是衡量学者学术水平和职业发展的重要指标。为了在激烈的竞争中脱颖而出，一些学者和学生不惜采用不正当手段，利用生成式人工智能进行学术造假。此外，部分使用者道德意识淡薄，缺乏对学术诚信的敬畏之心，没有认识到学术造假的严重后果。

最后，从对于生成式人工智能的管理层面来看。学术评价体系的不完善和监管不力为学术造假提供了可乘之机。一些学术机构过于注重论文的数量和影响因子，而忽视了研究的质量和创新能力，导致部分人为了追求论文数量而进行造假。同时，在学术评审过程中，存在评审不严格、人情关系等因素干扰的情况，使得一些造假论文能够顺利通过评审。此外，对于学术造假的处罚力度相对较轻，难以起到有效的威慑作用。

因此，不管是出于何种原因带来的学术造假，它的影响我们都不可轻视学术造假带来的影响。对学术研究质量的影响。学术造假严重损害了学术研究的质量。虚假的研究成果会误导后续的研究方向，浪费大量的科研资源和时间。同时，它也破坏了学术研究的公正性和严肃性，使得真正有价值的学术成果难以得到应有的认可和推广；对学术声誉的影响。学术造假事件的发生会严重损害学术机构和学者的声誉。一旦某个学术机构或学者被曝光存在学术造假行为，其在学术界的形象将一落千丈，不仅会影响其

未来的学术发展和合作机会，还会对整个学术界的声誉造成负面影响；对人才培养的影响。在学术造假的环境下，学生容易受到不良影响，形成错误的学术价值观和道德观念。他们可能会认为通过不正当手段获取学术成果是可行的，从而忽视了对自身学术能力和道德素质的培养，不利于培养具有创新精神和责任感的高素质人才。

5. 学术造假引发社会信任危机的路径分析

从信息传播维度来看。在信息时代，学术造假信息能够通过社交媒体、学术平台等渠道迅速传播。一旦学术造假事件被曝光，相关信息会在短时间内广泛传播，引起公众的关注和讨论。虚假信息的快速扩散会误导公众认知，使公众对学术研究的真实性和可靠性产生怀疑。此外，社交媒体上的信息传播往往具有情绪化和片面化的特点，一些不实的言论和猜测会进一步加剧公众的不信任情绪。

从公众认知维度来看。公众对学术造假的认知和态度变化是引发社会信任危机的重要因素。当学术造假事件频繁发生时，公众会对学术界的道德水平和学术规范产生质疑。他们可能会认为学术界已经失去了应有的公正性和严肃性，成为一些人谋取私利的工具。这种认知的转变会导致公众对学术成果和科学研究的信任度下降。

从社会心理维度来看。学术造假引发的社会心理反应会进一步加剧社会信任危机。公众在得知学术造假事件后，往往会产生焦虑、愤怒、失望等情绪。这些情绪会在社会范围内传播和扩散，形成一种负面的社会氛围。公众可能会对科技发展产生担忧，认为科技的发展并没有带来真正的进步，反而被一些人用于不正当的目的。同时，社会信任危机还会影响社会的稳定和发展。当公众对学术界和科学研究失去信任时，可能会减少对科研项目的支持和投入，从而阻碍科技的进步和创新。此外，社会信任危机还可能引发公众对政府和社会机构的不信任，影响社会的和谐与稳定。

6. 生成式人工智能伦理失范的多维透视

从技术维度来看，一是算法偏见与歧视。如今，机器算法能够自己给自己编写代码，“通过从数据中推断，它们自己会弄明白做事方法，掌握的数据越多，它们的工作就越顺利。”^[7]“在机器习得越来越像人的语言能力的过程中，它们也正在深度地吸取人类语言模式中隐含的种种偏见。”^[8]生成式人工智能模型的训练数据往往来自现实世界，而现实世界中存在着各种偏见和歧视。这些偏见和歧视会被模型学习并体现在生成的内容中，从而导致算法偏见和歧视问题。二是技术失控风险。随着生成式人工智能技术的不断发展，其自主性和智能性越来越高，存在技术失控的风险。一旦生成式人工智能系统出现故障或被恶意利用，可能会生成大量有害或虚假的内容，对社会造成严重的影响。

从法律维度来看，一是法律责任界定困难。生成式人工智能的应用涉及多个主体，包括开发者、使用者、数据提供者等。当生成式人工智能生成的内容引发法律问题时，很难确定各个主体的法律责任。二是知识产权保护挑战。生成式人工智能生成内容的版权归属问题是一个复杂的法律难题。生成式人工智能的运行原理可以分为两个过程：首先是通过机器学习和深度学习在接受大量人类语料的投喂后形成“利用语元关联度预测下一个语元”的能力，并且在从输入到输出再从输出到输入的双向互动中不断训练和优化模型；其次，在此基础上，生成式人工智能产品还会利用用户在实际参与时的语言表达习惯，不断地调整自身^[9]。目前，法律对于人工智能生成内容的版权认定尚未形成统一标准。从传统知识产权法角度看，版权通常赋予给具有创造性的自然人或法人。然而，生成式人工智能生成内容的过程并非完全由人类主导，它基于算法和大量数据自主学习生成内容，这使得其创造性难以简单归因于人类。若将版权赋予开发者，开发者可能并未直接参与内容的创作构思；若赋予使用者，使用者可能只是输入了简单提示语。这种版权归属的不确定性，不仅会影响创作者的积极性，还可能引发大量的知识产权纠纷，

阻碍生成式人工智能相关产业的健康发展。

从教育维度来看，一是教育公平问题。生成式人工智能的广泛应用可能加剧教育不公平现象。在经济发达地区，学校和学生能够更早、更广泛地接触到先进的生成式人工智能教育工具和资源，利用其提升学习效率和效果。而在经济欠发达地区，由于资金、技术等限制，学生可能无法享受到这些优势，导致教育差距进一步拉大。此外，不同家庭背景的学生在获取和使用生成式人工智能教育资源方面也存在差异，富裕家庭的学生可能拥有更多的学习辅助设备和资源，从而在学业竞争中占据更有利的位置；二是价值观引导缺失。生成式人工智能生成的内容具有多样性和不确定性，其中可能包含一些不符合社会主流价值观的信息。学生在使用生成式人工智能进行学习和研究时，如果缺乏正确的引导和判断，容易受到这些不良信息的影响，形成错误的价值观和世界观。

从文化维度来看，一是文化价值观冲突。不同文化背景下，人们对生成式人工智能的接受程度和使用方式存在差异，这可能导致文化价值观的冲突。在一些注重集体主义的文化中，生成式人工智能的应用可能被视为对人类集体智慧和协作的挑战；而在强调个人主义的文化中，人们可能更关注生成式人工智能对个人隐私和自主性的影响。此外，生成式人工智能生成的内容可能融合了多种文化元素，但这些元素的组合和呈现方式可能不符合某些文化的价值观和审美标准，从而引发文化冲突；二是文化多样性威胁。生成式人工智能的训练数据往往来自特定的文化群体或主流文化，这可能导致生成的内容偏向于这些文化，忽视其他小众文化或弱势文化的表达。长期来看，这可能会削弱文化的多样性，使得一些独特的文化元素逐渐被边缘化甚至消失。

7. 生成式人工智能伦理失范的路径优化

推进健全网络综合治理体系和优化公共安全治理机制，应重点完善生成式人工智能的统筹发展与管理机制，加快建立健全覆盖人工智能研发、部署与应用全流程的安全监管制度。我国需考虑，在人工智能未来风险不可完全认知、风险评估及成本效益分析等规制手段不完全有效的情况下，坚持人本主义立场，秉持人工智能系统与人类的协作与互补关系，保持人机共存社会中的人类主导地位，防控科技伦理风险，保障人的基本权利，探寻符合中国语境、面向中国问题的科技向善的实现之道^[10]。具体可以从以下几个方面进行。

第一，构建可信人工智能技术体系，夯实伦理风险的内生性屏障。治理的根本在于从技术源头嵌入伦理约束。首先，需推动可信人工智能技术范式的落地。这包括研发与应用可解释性人工智能技术，提升模型决策过程的透明度与可追溯性，破解“黑箱”困境。其次，必须在算法全生命周期贯彻公平性设计原则，通过偏见检测与缓解技术、采用多样化代表性数据集进行训练，并在输出端设置公平性约束，从源头抑制算法歧视的生成与放大。第三，强化隐私增强计算技术的整合应用，在模型训练与推理中广泛采用差分隐私、同态加密等技术，实现在数据“可用不可见”的前提下完成计算，从根本上捍卫个人隐私与数据主权。最后，开发并部署高效的深度伪造内容检测与溯源技术，通过数字水印、内容认证标准等技术手段，为生成内容提供可验证的“数字出身证明”，抵御虚假信息冲击。

第二，完善敏捷协同的法律规制体系，划定创新发展的刚性边界。法律规制需兼顾规范性与敏捷性。首要任务是推动立法更新，明确生成式人工智能相关主体的法律责任。需在法律上厘清开发者、部署者、使用者等各方在数据侵权、生成内容侵权、损害后果中的责任划分原则，特别是确立开发者在产品设计与投放市场前的强制性伦理风险评估义务。其次，应创新知识产权制度，探索设立适应人工智能生成内容的有限保护机制，明确其保护范围、期限与归属规则，同时改革合理使用制度，平衡训练数据获取与权利人利益保护。再者，建立分级分类的监管沙盒制度。根据应用场景的风险等级(如高风险的医疗、司法、教育，低风险的娱乐、创意辅助)，实施差异化的准入要求、合规标准和动态监测机制，在可控环境

中鼓励创新与快速迭代。最后，加强跨境数据流动与人工智能治理的国际规则协调，积极参与全球人工智能治理对话，推动形成具有广泛共识的伦理与规则框架，应对治理碎片化挑战。

第三，健全多层次的标准与行业规范，引导负责任的最佳实践。标准与规范是连接法律原则与具体实践的桥梁。国家层面应加快制定人工智能伦理与安全的国家标准体系，涵盖数据质量、算法公平、安全评估、可解释性等关键维度，为产业提供明确的技术合规指引。行业组织应牵头制定细化的行业自律公约与最佳实践指南，针对教育、金融、传媒等不同垂直领域，提出更具体的数据处理、模型审核、应用部署规范。企业层面，强制推行人工智能伦理影响评估制度，要求企业在开发、部署重大人工智能系统前，系统评估其对隐私、公平、安全、社会影响的潜在风险，并公开评估摘要。同时，在企业内部建立伦理官或伦理审查委员会机制，确保伦理考量深度融入产品管理全流程，形成内部制衡。

第四，培育多元主体参与的社会治理生态，凝聚价值对齐的广泛共识。有效的治理最终依赖于社会共识与多元共治。必须大力提升全社会的人工智能素养与伦理意识，通过教育体系与公共传播，普及人工智能基本原理、能力边界与潜在风险，培育公众的批判性使用能力。应鼓励建立包括技术专家、伦理学者、法律人士、社会科学家及公众代表在内的多元、透明的伦理审查与社会监督机制，对重大人工智能项目及应用进行独立的伦理审议与持续监督。此外，需支持跨学科的负责任人工智能创新研究，资助关于公平性、可解释性、价值对齐、长期社会影响等前沿课题的探索，为治理提供持续的知识供给。最终，应推动形成以人为本、向善发展的人工智能文化，促使技术创新主动服务于增进人类福祉、促进社会公平、保障全球安全的共同目标。

综上，应对生成式人工智能的伦理挑战，是一项涉及技术、法律、标准与社会的系统工程。唯有通过技术内置的伦理设计、敏捷前瞻的法律规制、细致入微的行业标准以及广泛深入的社会共治，形成多维度、多层次协同发力的治理格局，方能驾驭技术发展的巨浪，引导生成式人工智能在突破技术奇点的同时，守护好人类价值的基石，实现其推动社会进步的最大潜能。

8. 结论与展望

生成式人工智能在给社会带来巨大变革和创新机遇的同时，也引发了一系列复杂的伦理失范。从学术造假到社会信任危机，这些问题不仅影响了学术研究的公正性和严肃性，也破坏了社会的信任基础和稳定发展。通过从技术、法律、教育、文化等多维视角对生成式人工智能的伦理失范进行全面剖析，我们发现这些问题相互关联、相互影响，需要采取综合性的策略加以应对。

未来，随着生成式人工智能技术的不断发展和应用场景的不断拓展，其伦理问题将更加复杂和多样化。因此，需要政府、企业、学术界和社会公众共同努力，形成合力。政府应加强政策引导和监管，企业应积极履行社会责任，学术界应加强研究和探索，社会公众应提高伦理意识和参与度。只有这样，才能确保生成式人工智能在遵循伦理原则的基础上实现健康、可持续发展，为人类社会带来更多的福祉。同时，我们也需要持续关注生成式人工智能伦理问题的新动态和新挑战，不断调整和完善应对策略，以适应技术发展和社会变化的需求。

参考文献

- [1] Sleator, L. (2023) Almost Half of Cambridge Students Admit They Have Used ChatGPT. <https://www.thetimes.com/business/technology/article/cambridge-university-students-chatgpt-ai-degree-2023-msv7mw7z>
- [2] 李艳, 许洁, 贾程媛, 等. 大学生生成式人工智能应用现状与思考——基于浙江大学的调查[J]. 开放教育研究, 2024, 30(1): 89-98.
- [3] Rahman, M.M. and Watanobe, Y. (2023) ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied Sciences*, **13**, Article 5783. <https://doi.org/10.3390/app13095783>

-
- [4] Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016) The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*, **3**, 1-21. <https://doi.org/10.1177/2053951716679679>
- [5] 佩德罗·多明戈斯. 终极算法: 机器学习和人工智能如何重塑世界[M]. 黄芳萍, 译. 北京: 中信出版社, 2016.
- [6] Navarro-Dolmestch, R. (2023) Risks and Challenges Posed by Artificial Intelligence Generative Applications for Academic Integrity. *Derecho PUCP*, **91**, 231-270.
- [7] 佩德罗·多明戈斯, 黄芳萍. 终极算法: 机器学习和人工智能如何重塑世界[J]. 金融纵横, 2018(2): 102.
- [8] 温德尔·瓦拉赫, 科林·艾伦. 道德机器: 如何让机器人明辨是非[M]. 王小红, 主译. 北京: 北京大学出版社, 2017: 1.
- [9] 陈小平. 大模型: 人工智能思想及其社会实验[J]. 文化纵横, 2023(3): 70-77, 158.
- [10] 贾璐萌. 技术伦理实现的内在路径研究[D]: [博士学位论文]. 沈阳: 东北大学, 2018.