

中国中小学人工智能教材中伦理表征的认知网络分析

Mapping AI Ethics Representations in AI Textbooks of Chinese Primary and Secondary

Education

孟佶欣, 许柳桐, 何涛*

深圳大学 教育学部

* tao.he2016@gmail.com

【摘要】 随着全球 K-12 人工智能课程的扩展, 如何在教材中有效融入伦理内容仍面临重要挑战。本研究旨在通过描绘伦理表征的认知结构, 评估其教学适切性与发展性。研究采用大语言模型与认知网络分析 (ENA), 依据 UNESCO 十大人工智能伦理原则, 对中国 K-12 人工智能教材进行量化建模。结果发现教材存在“技术中心”倾向: 透明度 (46.3%) 显著高于可持续性 (5.4%)。同时, 教材呈现积极的三阶段发展路径: 从小学中低年级的直观感知, 到高年级的防御性规则, 再到初中阶段的公民治理。研究表明, 教材已具备一定引导作用, 未来设计重点应转向培养批判意识与生态责任。

【关键词】 K-12 人工智能教育; 人工智能伦理; 教材分析; 认知网络分析

Abstract: As K-12 AI curricula expand globally, integrating ethics into AI textbooks remains a significant challenge. This study maps the epistemic structure of ethical representations in textbooks to evaluate pedagogical appropriateness and structural progression. Using Large Language Models and Epistemic Network Analysis (ENA), we modeled ethical representations in Chinese K-12 AI textbooks based on UNESCO's ten ethical principles. Results reveal a techno-centric imbalance, with Transparency (46.3%) outweighing Sustainability (5.4%). At the same time, a three-stage progression emerges, from intuitive perception in middle primary grades, to Defensive Rules in upper primary, and Civic Governance in junior high. These findings suggest that while textbooks function as learning guides, future designs should move beyond technical demystification toward fostering critical awareness and ecological responsibility.

Keywords: AI education, AI ethics, textbooks analysis, epistemic network analysis

1. 引言

随着未来劳动力市场对人工智能基础素养的需求日益增长, 全球 K-12 教育系统正迅速将人工智能作为一门独立学科纳入课程体系 (Casal-Otero et al., 2023)。在此过程中, 专门的人工智能教材已成为课程实施的关键教学材料, 是界定学生应当学习何种人工智能知识的主要载体。与传统学科教材不同, 这些人工智能教材肩负着双重使命: 不仅需要帮助学生理解复杂的技术原理, 还必须培养其“人工智能素养”, 该素养涵盖技术能力与伦理意识 (Chiu & Chai, 2025)。因此, 经济合作与发展组织 (OECD) 和联合国教科文组织 (UNESCO) 等国际框架文件均明确提出, 此类教学材料应将公平性、环境可持续性伦理维度置于人工智能素养的核心位置 (OECD & European Commission, 2025; UNESCO, 2022)。在本研究中, 人

工智能伦理是指围绕人工智能设计、开发与应用过程中涉及的公平、公正、透明、责任与人类福祉等价值原则的规范性框架（UNESCO, 2022）。

然而，如何将这些宏观伦理原则转化为微观教学材料仍是一项关键挑战。教材不仅是知识的载体，更是塑造学生技术认知模式的“认知支架”（Gouseti & Brereton, 2024）。目前针对 K-12 阶段人工智能伦理教育的研究相对有限，而专门探究教材中伦理知识结构性组织方式的研究则更为匮乏（Adams et al., 2023）。伦理知识如何被表征与结构化（如隐私是被构建为技术安全问题，还是作为一项基本人权加以呈现），在人工智能教材研究中鲜有系统讨论。

为弥补上述研究空白，本研究采用认知网络分析方法（Shaffer et al., 2016），对教材中伦理表征的结构性共现关系进行建模分析。据此，本研究致力于回答以下三个核心研究问题：

RQ1：中国 K-12 人工智能教材中伦理内容呈现出何种结构模式？

RQ2：在小学中段（3-4 年级）与小学高段（5-6 年级）之间，人工智能教材的伦理关注点的认知结构存在何种差异？

RQ3：在小学与中学阶段的人工智能教材中，伦理表征在发展层面上存在何种差异？

2. 文献综述

2.1. K-12 教育中的人工智能伦理

尽管国际政策指南为人工智能伦理确立了普适性标准，K-12 教育需要一种契合青少年认知发展特点的独特路径。近年来，研究重点已从对成人伦理规范的“简化式移植”，转向强调“教学适切性”，即伦理教学内容与学生发展阶段相契合（Adams et al., 2023）。相关研究指出，K-12 阶段面临特有的人工智能伦理风险，如“监控式教育”，并指出课程应超越对技术规则的遵循，以培养对人工智能应用的批判性意识（Akgun & Greenhow, 2022）。Adams 等（2023）进一步批评了当前普遍的“合规导向”的教学方法，主张采用探究式模型，将抽象伦理与学生日常经验相联系。然而，如何在低龄学习者中实现这种“教学适切性”，仍是一个尚待解决的关键问题。此外，Gouseti 等（2024）也指出，现行教材中往往存在高层次价值观（如人的尊严）与学生现实生活体验之间的脱节。

因此，有必要深入考察教材能否发挥桥梁作用，使人工智能伦理教育超越简单的规则传授，将抽象的伦理价值有效地连接到学生的日常现实之中。

2.2. K-12 人工智能课程的教材设计

教材是将宏观教育政策转化为微观教学内容的主要载体。当前，K-12 人工智能课程的设计正经历一场从“编码技术中心”向“素养导向”的显著转型，并严格遵循价值驱动与能力本位的教育原则（Casal-Otero et al., 2023; Chiu & Chai, 2025）。在此背景下，伦理不再被视为孤立的模块，而是通过项目式学习等教学策略，与技术学习过程有机融合，旨在将“科技向善”的抽象理念转化为具体的行动框架（Bosarge, 2025）。然而，关于这些教材在教学层面是否具备充分的合理性，仍然存在一个显著的研究空白。尽管人工智能伦理的必要性已在学界形成共识，但针对不同年级人工智能教材中伦理表征的结构化呈现方式，相关研究依然匮乏。因此，评估教材中伦理表征的结构进阶性，对于评判当前人工智能教材设计是否真正达成教学适切性的目标至关重要。

2.3. 人工智能伦理的概念框架

随着人工智能时代的兴起，对教材中人工智能伦理标准进行分类变得日益重要，多种伦理框架已被提出。其中一项关键框架是经合组织的人工智能原则，其为各国政府提供了非约束性指导。另一项是联合国教科文组织的《人工智能伦理问题建议书》，其在促进创新的同

时，强调了问责、公平与透明（UNESCO, 2022）。此外，七国集团（G7）的广岛人工智能原则概述了关于可信人工智能、协作治理与负责任部署的国际指南（Group of Seven, 2023）。

在人工智能教材日益普及的 K-12 教育背景下（Nabi et al., 2025），联合国教科文组织的框架因其覆盖了从透明度、安全性到治理与问责等更广泛的维度（UNESCO, 2023），并得到实用实施工具的支持，而特别适用于教材的结构性分析。因此，本研究采用联合国教科文组织《建议书》中的十项伦理原则作为编码框架，用以绘制教材中人工智能伦理的内容图谱。

3. 方法

3.1. 数据收集

本研究从中国不同省份及主要城市使用的多套人工智能教材中获取原始数据。研究共分析了 50 本指定用于中小学教育的教材。鉴于目前适用于小学低年级（1-2 年级）的专门人工智能教学材料较为稀缺，该学段未纳入本研究的分析范围。

3.2. 分析流程

为满足后续分析的需要，研究首先对 50 本教材的原始文本进行了预处理，共计得到超过 108 万中文字符。为对预处理后的文本进行切分，本研究采用了基于语义的方法，结合“paraphrase-multilingual-MiniLM-L12-v2”模型与 ruptures 库进行变点检测，以替代传统的固定长度切分方式。该过程共得到了 3614 个可用于数据分析的文本片段。

为确保自动编码的信效度，本研究采用 3: 7 的验证规则对预处理数据进行划分。三位专家首先对 30% 的数据进行人工编码，并以此作为基准校验大语言模型的自动标注结果。结果显示，Gemini 3-pro-reviews 模型与专家编码的一致率为 82.6%，表明采用大语言模型进行内容自动标注的可行性。基于该结果，剩余 70% 的数据由 Gemini 3-pro-reviews 进行处理和评估。

研究依据联合国教科文组织《建议书》中的十项伦理原则进行编码，具体包括：人权与人类尊严（HR.HD），环境与生态系统繁荣（E.EF），多样性与包容性（D.I），和平、正义与强大机构（PJ.SI），相称性与不伤害原则（P.DNH）、安全与保障（S.S）、透明度与可解释性（T.E）、责任与问责（R.A）、数据治理与隐私（DG.P）、可持续发展与福祉（SD.W）。

之后，鉴于到认知网络分析在刻画知识要素之间共现关系与结构模式方面的优势，本研究采用 ENA 对教材中人工智能伦理原则的表征方式并相互关联关系进行定量分析。

4. 研究结果

4.1. 中小学阶段人工智能教材伦理表征总体分析



图 23 K-12 人工智能教材伦理表征总体分布的雷达图

总体而言，“透明度与可解释性”（T.E）在伦理表征中占比最高，共计 1673 条编码，占总样本量的 46.3%。相比之下，“环境与生态系统繁荣”（E.EF）的占比最低，仅有 195 条编码，仅占总语料的 5.4%。

表 11 各年级认知网络投影均值对比

年级	文本片段数量	X 均值	Y 均值	X 轴标准差 (SD)	Y 轴标准差 (SD)
G3-G4	558	-0.28	-0.12	1.83	1.65
G5-G6	481	0.31	0.42	2.12	1.94
G7-G8	828	0.01	-0.16	2.05	1.78
方差分析 (ANOVA)		F = 13.52	F = 15.86		
显著性水平 (p)		< .001	< .001		

方差分析结果显示，在两个投影维度上均存在显著的年级差异（X 轴：F=13.52，p<.001；Y 轴：F=15.86，p<.001）。在 X 轴上，G3-G4 年级组（均值-0.28）与 G5-G6 年级组（均值 0.31）呈现明显分离，这表明伦理认知从基础感知向以安全与规则为导向的思维方式转变。Y 轴则呈现出一个阶段性的跃升：G5-G6 年级组为正值（均值 0.42），强调安全、保障与隐私，而 G3-G4 和 G7-G8 年级组均为负值且数值相近。其中，G3-G4 年级组强调透明度与包容性；G7-G8 年级组则扩展到责任、权利与治理层面的伦理议题。

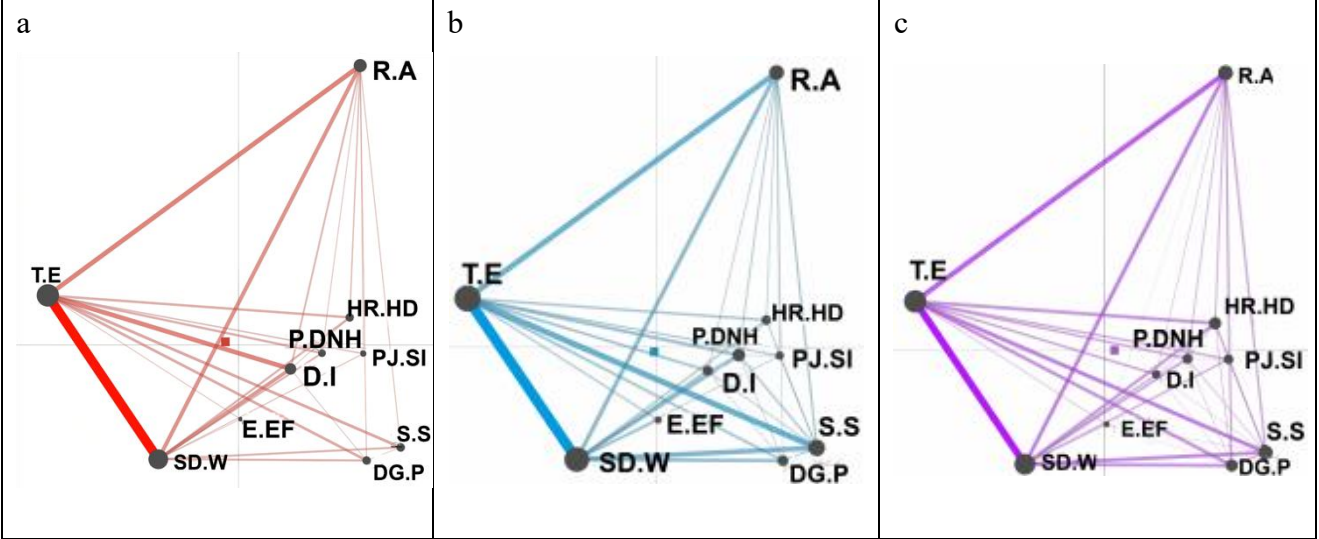


图 24 不同年级段的人工智能教材伦理表征的认知网络图

注：a 人工智能教材中 G3-G4 的网络图；b 人工智能教材中 G5-G6 的网络图；c 人工智能教材中 G7-G8 的网络图。最小边权重为 0.01。

图 24 展示了三个年级段人工智能伦理结构的认知网络。3-4 年級的伦理表征以“透明度与可解释性”、“可持续发展与福祉”及“多样性与包容性”构成稀疏簇，反映出“技术改善生活”的叙事取向。5-6 年級的网络则更为密集且更注重实践导向，其中“安全与保障”作为核心节点，与“透明度与可解释性”及“数据治理与隐私”紧密相连，强调安全使用和基于规则的推理。7-8 年級的网络复杂性最高，并转向社会与权利导向的治理观，表现为“人权与人类尊严”、“和平、正义与强有力的制度”及“责任与问责”与“数据治理与隐私”相互交织。总体而言，伦理理解在不同学段中呈现出阶段性重组特征，而非线性累积。

4.2. 人工智能教材中在小学中段与高段的伦理表征差异

表 12 小学中段（G3-G4）与高段（G5-G6）的稳健配对比较

维度	比较	均值差	标准误	P（校正后显著性）	结果
X 轴：基础感知 vs. 规则导向	G3-G4 vs G5-G6	-0.165	0.037	< .001	S.d.
Y 轴：个体取向 vs. 社会取向	G3-G4 vs G5-G6	0.013	0.025	1.000	n.s.

维度 1 显示，从 G3-G4 到 G5-G6 存在显著转变，表明伦理认知从基础感知向基于规则的伦理演进(p<.001)。维度 2 在两个组别间未发现显著差异。

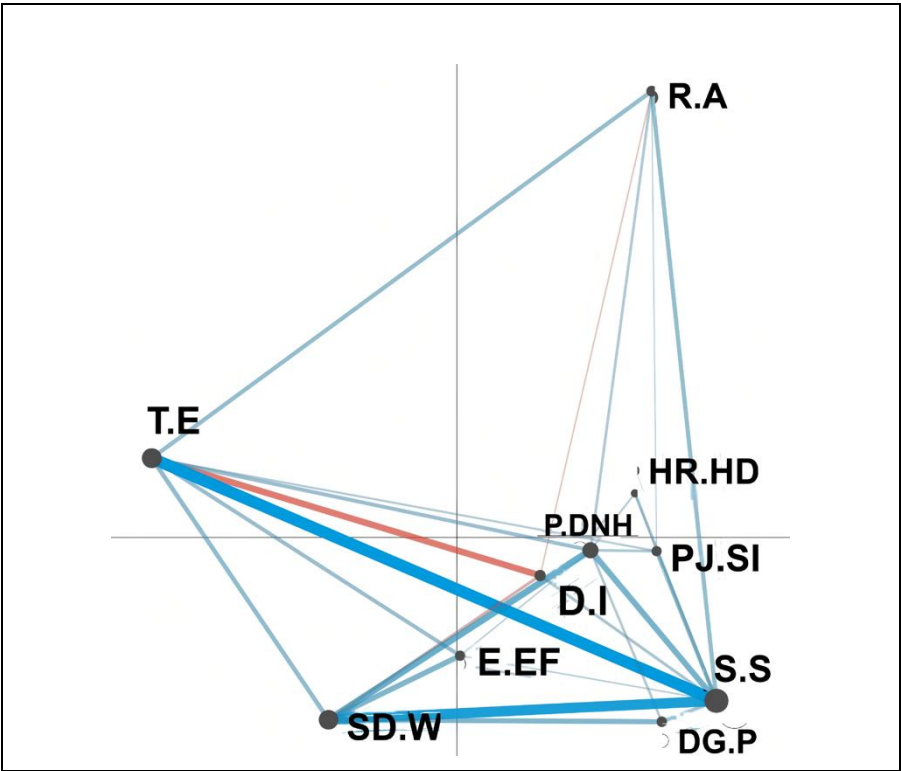


图 25 小学中段（G3G4）与高段（G5G6）人工智能教材的差值网络图
注：最小边权重为 0.01。

图 25 展示了小学中段与高段伦理差值网络。蓝线（G5-G6）显示其伦理表征与“安全与保障”有更强的关联，特别是“透明度与可解释性”及“数据治理与隐私”，这凸显了高年级教材中数据隐私意识的增强。相比之下，红线（G3-G4）则更多聚集于“多样性与包容性”，表明针对中年级阶段学生更侧重于直观与包容的伦理教育。

4.3. 人工智能教材在小学与中学阶段的伦理表征差异

表 13 初中阶段（G7-G8）与小学中段（G3-G4）、高段（G5-G6）的稳健配对比较

维度	比较	均值差	标准误	P（校正后显著性）	结果
X 轴：基础感知 vs. 规则导向	G3-G4 vs G7-G8	-0.108	0.029	< .001	S.d.
	G5-G6 vs G7-G8	0.058	0.037	.358	n.s.
Y 轴：个体取向 vs. 社会取向	G3-G4 vs G7-G8	-0.070	0.020	.001	S.d.
	G5-G6 vs G7-G8	-0.084	0.025	.002	S.d.

事后检验结果显示，在维度 1（认知复杂度）上，G3-G4 与 G7-G8 之间存在显著差异（ $p<.001$ ），但在 G5-G6 与 G7-G8 之间无显著差异。这表明，小学三、四年级可视为人工智能伦理理解的入门阶段，而小学五、六年级在认知复杂性层面已接近初中阶段。在维度 2（伦理社会化）上，则显示出明确的小学-初中分野：G7-G8 与 G3-G4（ $p=.001$ ）及 G5-G6（ $p=.002$ ）之间均存在显著差异，反映出伦理焦点从小学阶段的个人化、技术性伦理向初中阶段的社会化、制度性伦理的转变。

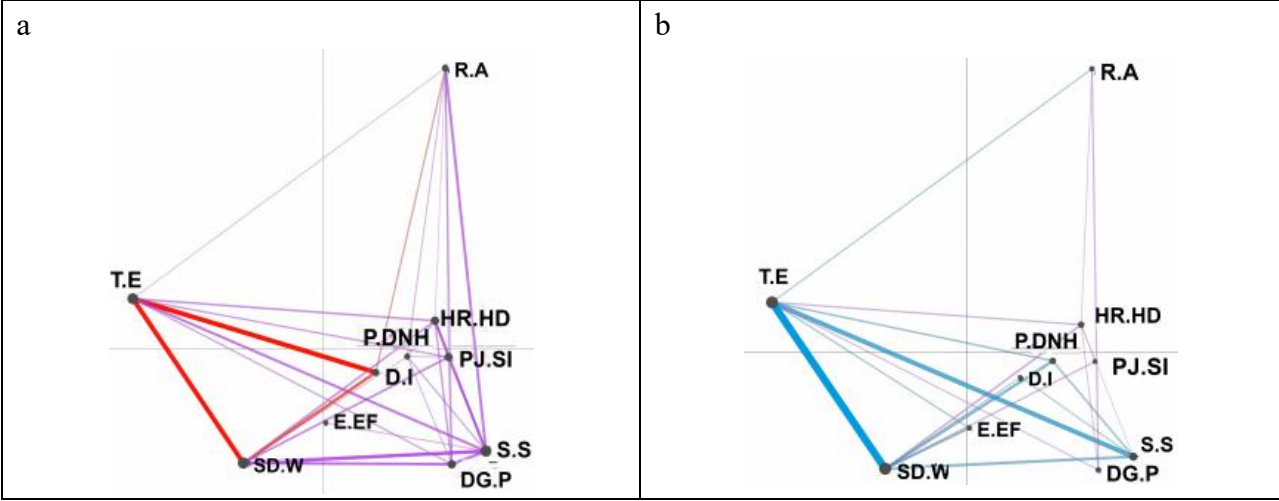


图 26 不同学段人工智能教材伦理表征的差值网络图

注：a 小学中段与初中阶段人工智能教材的差值网络图；b 小学高段与初中阶段人工智能教材的差值网络图。最小边权重为 0.01。

图 26 展示了小学中段与初中阶段之间，小学高段与初中阶段之间人工智能教材中伦理表征的差值网络结构。在小学中段与初中阶段的对比中，小学中段的网络将“透明度与可解释性”与“安全与保障”、“可持续发展与福祉”紧密连接，同时在“多样性与包容性”与“透明度”之间存在独特连接。相比之下，初中网络则紧密关联如“人权与人的尊严”、“和平、正义与强有力的制度”等伦理主题，体现出伦理理解由技术认知到伦理社会化的根本性跨越。

在小学高年级的对比中，其网络呈现出过渡性的桥梁特征：“透明度-安全”关联紧密，而“数据治理与隐私”的引入则凸显了对遵循保护性规则的关注。相反，初中网络将“数据治理与隐私”与“责任与问责”以及社会正义主题深度融合。由此可见，小学高年级的伦理焦点仍停留在技术合规层面，而初中阶段则转向在更广泛的系统性问责与社会治理框架内审视伦理问题。

5. 讨论与总结

5.1. 讨论

本研究运用认知网络分析揭示了中国 K-12 人工智能教材中伦理表征的深层结构。通过将实证发现与三个研究问题相结合，研究识别出课程中一种鲜明的“双重性”，即当前教材一方面在伦理内容分布上存在较明显的结构失衡，另一方面在表征组织上呈现出合理的阶段性进阶。

首先，关于伦理内容的整体结构模式（RQ1），课程反映出一种“技术中心论”。结果显示，“透明度与可解释性”在伦理图谱中占主导，占有所有编码片段的 46.3%。这一结果与 Ng 等（2024）的观察一致，即当前 K-12 课程倾向于优先培养技术性人工智能素养，而非社会性探究。许多教材将伦理问题简化为技术谜题，这可能限制了学生的批判能力（Holmes et al., 2022）。与此相比，“环境与生态系统繁荣”则处于边缘地位（5.4%）。这印证了 Biagini（2025）关于人工智能教育与可持续性之间存在张力的论点。由于忽略了诸如模型训练等实际成本，课程可能无意助长了对技术的盲目乐观，而这种态度将数字行为与生态后果割裂开来。

其次，关于小学阶段内部的差异（RQ2），认知网络比较揭示出从直观伦理感知向基于规则保护的认知转向。小学中段的网络主要围绕“多样性与包容性”聚集，强调对公平与非歧视的直觉理解；而在小学高段，结构焦点明显转向“安全与保障”及“数据治理”。这一转变印证了 Akgun 和 Greenhow（2022）的观点，即课程必须优先考虑特定的隐私素养，以帮助学生应对数字风险。因此，这些教材从培养广泛的伦理意识转变为通过基于规则的推理来建立必要的保护界限。

第三，在不同教育阶段的发展进阶（RQ3）上，人工智能教材呈现出了清晰的认知深化过程。从小学高年级以规则为导向的伦理焦点，到初中阶段以社会和制度伦理为核心的网络结构演化，这一过程符合 Lacerda Queiroz 等（2021）提出的“从具体到抽象”的设计逻辑。小学阶段主要通过识别偏见、保护个人数据等具体经验引入伦理议题，而初中阶段则进一步拓展至正义、人权与制度责任等抽象概念。这种清晰的分层表明，教材的结构序列具有内在逻辑性，引导学生逐步完成从个人保护到社会责任的认知发展。

5.2. 总结

本研究运用认知网络分析解构了 50 本中国 K-12 人工智能教材中的伦理图景。研究发现，当前课程呈现出“结构适切但内容失衡”的双重特征。在认知结构方面，教材实现了伦理表征的逐步进阶。从小学中段的“社会意识”，经由高段的“规则保护”，再到初中阶段的“公民治理”，整体与学生认知发展阶段高度契合，为伦理学习提供了有效支架。在伦理内容方面，课程暴露出技术中心论的倾向。过度强调“透明度”会掩盖对算法权力、环境可持续性 & 社会偏见等议题的批判性反思，影响学生对人工智能伦理的理解倾向。

需要指出的是，本研究以 UNESCO 十项原则作为编码框架，虽具国际规范性，但亦可能使教材中未被该框架涵盖的隐性伦理叙事较难被识别。综合而言，未来的课程建设可在保持现有逻辑的基础上，进一步扩展环境可持续性、社会正义与公共责任等议题的呈现，使人工智能伦理教育不仅关注技术解释，也更重视对社会影响的系统理解。未来的研究也可尝试结合多元伦理框架或开放式编码方法，以捕捉更为多样化的伦理叙事。同时，应超越文本分析，结合课堂实施与学习者数据，进一步检验教材伦理表征与学生伦理发展之间的关系。

参考文献

- Adams, C., Pente, P., Lernermeier, G., & Rockwell, G. (2023). The ethics of using AI in K–12 education: A systematic literature review. *International Journal of Information and Learning Technology*, 40(1), 1–18.
- Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(3), 431–440.
- Biagini, G. (2025). Towards an AI-literate future: A systematic literature review exploring education, ethics, and applications. *International Journal of Artificial Intelligence in Education*, 1–51.
- Bosarge, E. (2025, March). Project-Based Learning and the AI4K12 Framework in High School AI Curriculum: A Systematic Review. In *2025 ASEE Southeast Conference*.
- Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K–12: A systematic literature review. *International Journal of STEM Education*, 10(1), Article 29.
- Chiu, T. K. (2025). AI literacy and competency: definitions, frameworks, development and future research directions. *Interactive Learning Environments*, 33(5), 3225–3229.
- Group of Seven. (2023, October 30). *Hiroshima process international guiding principles for organizations developing advanced AI systems*. Ministry of Foreign Affairs of Japan. <https://www.mofa.go.jp/files/100573471.pdf>
- Gouseti, A., & Brereton, P. (2024). Unpacking the ethics of using AI in primary and secondary education: A systematic literature review. *Educational Technology & Society*, 27(1), 1–15.
- Lacerda Queiroz, R., Ferrentini Sampaio, F., Lima, C., & Machado Vieira Lima, P. (2021). AI from concrete to abstract: demystifying artificial intelligence to the general public. *AI & SOCIETY*, 36(3), 877–893.
- Lin, P., & Van Brummelen, J. (2021, May). Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. In *Proceedings of the 2021 CHI conference on human factors in computing systems* (pp. 1–12).
- Nabi, F., Vortia, W., & Shardey, E. (2025). Teacher readiness for AI and digital tools in K–12 classrooms: A review of professional development trends and gaps. *International Journal for Multidisciplinary Research*, 7(5), 17.
- Ng, D. T. K., Su, J., Leung, J. K. L., & Chu, S. K. W. (2024). Artificial intelligence (AI) literacy education in secondary schools: a review. *Interactive Learning Environments*, 32(10), 6204–6224.
- OECD & European Commission. (2025). *AI literacy framework for primary & secondary education*. OECD Publishing.
- Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of learning analytics*, 3(3), 9–45.
- UNESCO. (2022). *Recommendation on the ethics of artificial intelligence*. United Nations Educational, Scientific and Cultural Organization.

UNESCO. (2023). *Readiness assessment methodology: A tool of the recommendation on the ethics of artificial intelligence*. UNESCO.