

AI 替代人工护理的责任界定与老年人权益保护

吴莉婧*, 徐禹哲

北方工业大学文法学院, 北京, 中国

*通讯作者

【摘要】 AI 替代人工护理引发责任归属难题, 挑战传统法律伦理框架。本文以责任伦理学为基准, 运用哲学思辨与规范分析方法, 探讨 AI 护理的法律责任界定。研究发现现有立场冲突本质是保护性、建构性与制度性责任的优先级困境。文章据此提出层级化准则: 确立人身尊严与安全为绝对优先责任, 福祉促进为平衡责任, 角色能力差异化为分担责任。该层级化规范体系为 AI 照护责任归属提供了明确的价值排序与权衡原则, 为相关立法司法提供了理论基础。

【关键词】 人工智能; 老年照护; 法律责任; 责任伦理

1. 引言

在人口结构深刻转型与智能技术深度融合的时代背景下, 人工智能 (AI) 系统正逐步渗透至老年人照护领域, 从生活辅助到健康监测, 乃至情感陪伴, 呈现出替代传统人工护理的显著趋照势。这一技术跃迁在提升效率、缓解人力资源短缺的同时, 也从根本上动摇了以“人”为核心行动者和责任主体的传统法律与伦理框架。当护理行为的主体由具身的、富有道德直觉的“人”转向由算法驱动、具备一定自主性的“人工物”时, 一个尖锐的规范性问题便浮出水面: 当 AI 系统在替代性护理过程中造成老年人权益损害时, 应当依据何种伦理原则, 在法律上公平地界定与分配相关各方的责任? 这一问题并非简单的技术应用风险管控, 其内核是一场关于责任归属的价值冲突与伦理困境, 亟待通过深层的哲学思辨来确立可资遵循的规范性准则。

鉴于此, 本研究的核心目标在于: 以责任伦理学为价值基准, 通过哲学思辨, 推导出一套用于界定 AI 替代人工护理场景中法律责任的核心伦理规范。本研究将不致力于构建面面俱到的具体条文, 而是旨在为未来的法律解释与规范建构提供不可或缺的价值锚点与推理前提。为实现此目标, 思辨路径将遵循严谨的规范性论证逻辑: 首先, 确立以“关系性”、“非相互性”和“前瞻性”为核心的责任伦理学作为本研究的价值基准。这一选择源于其与照护伦理的内在亲和性, 它超越了个体主义与后果主义的局限, 强调在不对等的关系 (如老年人与护理系统) 中,

优势方 (技术的设计、提供、管理者) 对弱势方负有基于脆弱性保护的、面向未来的积极责任[1], 这为分析 AI 护理中的责任本质提供了恰切的理论透镜。其次, 运用此基准, 深入剖析前述价值冲突的伦理本质, 即辨明在 AI 护理关系网络中, 各主体 (如研发者、运营商、家庭、公共机构) 所负责任的伦理属性 (是补偿性、抑或是建构性) 及其优先层级。最终, 基于上述分析, 推导出具有普遍指导意义的规范性结论, 即关于责任归属应优先遵循的核心伦理准则。

本研究的学术价值在于, 直面法学在应对颠覆性科技时的理论滞后性, 通过引入并融合责任伦理学的哲学资源, 旨在为“人工智能法学”尤其是涉老科技伦理与法律规制研究, 贡献一个侧重于“责任原理”而非仅仅“归责技术”的规范性分析框架。它试图弥合单纯的技术风险管控论与抽象的权利宣言之间的沟壑, 推动责任法理由事后追究向事前建构、由个体本位向关系本位的范式演进。在实践层面, 所推导的伦理规范可为立法者设计 AI 护理产品的责任规则、为司法者裁断相关侵权纠纷、为行业构建自律标准, 提供深层的价值指引与论证理据, 从而在规范层面助力构建一个既保障老年人尊严与安全, 又促进技术向善发展的负责任创新生态。

2. 责任伦理学的价值基准: 为 AI 护理责任奠基

面对 AI 替代人工护理所产生的责任界定难题, 确立一个恰切且坚实的价值基准是进行规范性论证的逻辑起点。本研究选择责任伦理学 (Ethics of Responsibility) 作为核

心价值基准,其根本原因在于,传统的以个体自主、自由意志和过错归责为核心的规范伦理学(如义务论、功利主义)在应对具有高度不对称性、依赖性与技术中介特性的AI对老年人照护关系时,显现出解释力上的局限。责任伦理学,特别是自汉斯·乔纳斯(Hans Jonas)以降所发展的强调“前瞻性”、“非相互性”与“关系性”的理论脉络,为我们重构责任范式提供了更为适配的哲学透镜。

2.1 责任伦理学的内涵解析

责任伦理学的核心内涵可被拆解为三个相互关联的要素:

其一,关系性与非对称性责任。责任并非源于自由个体间的平等契约,而是诞生于一种存在论上的非对称关系之中,即一方(通常拥有更大的权力、知识或影响力)对另一方(处于脆弱、依赖状态)的存续与福祉负有不可推卸的义务[2]。在AI护理场景中,这直接对应着技术开发者、部署机构乃至监管者相对于认知与身体机能可能衰退的老年使用者所具有的显著能力优势。

其二,前瞻性与未来导向。责任伦理强调的并非仅仅是事后追溯性的归责与补偿,更是事前的预见、预防与积极建构。这意味着,对AI护理系统的责任考量,必须超越损害发生后的赔偿,延伸至技术设计之初的价值嵌入、风险评估以及整个生命周期中对老年人潜在影响的持续关怀。

其三,集体与制度性责任的凸显。在复杂的现代技术系统中,责任往往无法精准地锚定于某个特定个体,而是分散在组织、制度乃至整个社会网络之中。责任伦理学要求我们思考如何建构能够承载这种集体责任的制度与规范[3]。这与AI护理涉及多主体(硬件商、算法开发者、数据提供方、养老机构、家庭、政府)协作的特征高度契合。

2.2 责任伦理学的理论溯源

追溯其理论渊源,责任伦理学是对康德式形式主义义务论与结果主义功利论的一种批判性发展。它继承了现象学与存在主义对“他者”与“处境”的关切,将伦理重心从“我应当做什么”的普遍法则,转向“我对谁负责”的具体关系。在当代应用伦理领域,特别是在科技伦理、环境伦理与生命伦理中,责任伦理学已成为分析技术风险、代际正义与弱势群体保护的关键理论资源。将其应用于法学领域,尤其是侵权责任法与社会法研

究,旨在推动责任观念从“向后看的过错填补”向“向前看的风险分配与福祉促进”进行范式扩展。

2.3 责任伦理学的理论适配性分析

本研究主张责任伦理学与“AI替代人工护理的责任界定”问题具有内在的适配性。该基准精准覆盖了问题的核心矛盾:即如何在一种能力非对称、权力不平等且充满技术不确定性的照护关系中,公平地分配责任。传统的过错责任或严格产品责任框架,或纠缠于“算法黑箱”下的主观过错证明之难[4],或僵化地将AI视为普通物而忽视其作为“准照护者”的社会角色。责任伦理学则直接承认并立足于这种非对称关系,要求优势方(技术行动者)承担起保护脆弱方(老年人)的积极义务。因此,责任伦理学以其对关系、未来与制度的综合性强调,为剖析和化解AI护理中的责任界定困境,提供了最为周全且有力的价值标尺。

3. 多元价值诉求的冲突本质:基于责任伦理的审视

以责任伦理学为标尺,可以深入拆解导言中所述关于AI护理责任界定的价值立场分歧,追溯其诉求来源,并厘清其合理性与局限,从而凝练冲突的深层本质。

3.1 “用户中心主义下的责任强化论”的价值诉求分析

“用户中心主义下的责任强化论”的价值诉求根植于保护弱势群体与实现补偿正义的古老法律伦理原则。其合理性在于敏锐地洞察到老年人在面对复杂AI系统时可能存在的数字鸿沟、认知衰退与议价能力不足,从而处于极度脆弱(vulnerable)的地位[5]。从责任伦理学的“关系性与非对称性责任”视角审视,这种强调对技术提供方课以重责的诉求,正是对优势方(技术行动者)应承担保护义务的强烈呼唤。它要求法律规范必须向被照护者倾斜,以确保其基本尊严、安全与健康权益不受技术风险的侵蚀。然而,此立场的局限性在于,它可能过于侧重事后救济与风险静态分配,倾向于一种“后果主义”的严格责任模式。这种模式虽然强化了损害填补,但可能忽略了责任伦理同样强调的“前瞻性”。若责任配置过于严苛且缺乏弹性,可能抑制技术迭代与优化,从长远看反而不利于通过技术进步来提升整体照护质量与可及性,即可能损害对未来老年人福祉的“前瞻性责任”。

3.2 “产业促进导向下的责任限定论”的价值诉求分析

“产业促进导向下的责任限定论”的核心价值诉求是激励技术创新与促进社会福祉增长，其哲学基础接近于功利主义对效率与整体福利最大化的追求。它担忧过度的责任负担会形成“寒蝉效应”，阻碍资本投入与技术研发，最终延缓 AI 护理解决方案的普及，损害更广泛老年群体的长远利益[6]。从责任伦理学的“前瞻性”维度看，对技术发展的审慎支持本身具有合理性，因为技术的良性发展本身是履行对未来世代责任的一部分。然而，此立场的致命局限在于，它可能抽离了具体的“关系性”语境，将老年人抽象为技术产品的普通消费者，甚至将技术风险视为其应自行承担的市场选择后果。通过主张“技术中立”或过度依赖合同安排来限定责任，实际上可能架空优势方对脆弱方的具体保护义务，违背了责任伦理学中基于非对称关系的核心要求。这种立场在实践中可能导致责任虚化，将风险不当转嫁给老年人及其家庭。

3.3 “风险分担框架下的关系责任论”的价值诉求分析

“风险分担框架下的关系责任论”试图综合上述关切，其价值诉求指向分配正义与系统韧性。它认识到 AI 护理是一个多元主体参与的复杂生态系统，任何单一方的过错都难以完全解释损害，因此主张根据过错、控制力或收益比例来分配责任[7]。这一立场与责任伦理学强调的“集体与制度性责任”高度共鸣，它跳出了简单的二元对立，看到了责任承载者的多元性与网络化特征。其进步性在于将分析单元从“点对点”转向“网络对点”。然而，其当前的局限性在于规范层面的模糊性：它指出了分担的必要性，但未能提供清晰的价值排序与优先级准则。例如，当开发者、运营商与养老机构均存在不同程度过失时，何者的责任更具优先性？在保护老年用户权益与维系系统可持续运行之间，当出现资源分配冲突时，应依据何种价值进行裁定？这种“如何分担”的准则缺失，使得该框架在面临具体价值抉择时可能陷入困境。

综上所述，上述价值立场分歧的本质，并非简单的“保护 vs. 创新”或“个体 vs. 社会”的线性冲突，而是在一个由技术深度嵌入的、非对称的照护关系网络中，三种不同维度的责任伦理要求之间出现的张力与优先级困境：

1. 基于非对称关系的、对当下脆弱个体的保护性责任（用户中心主义的核心）；2. 面向未来的、通过技术发展促进整体福祉的建构性责任（产业促进导向的核心）；3. 在多元主体构成的系统中，公平分配义务与风险的制度性责任（风险分担框架的核心）。

真正的规范性难题在于：当这三种责任要求无法同时充分满足时，应当依据何种价值排序与权衡原则，来界定和分配法律责任？责任伦理学为我们提供了审视这一冲突本质的框架，但它本身并未给出自动的答案，必须进行进一步的规范建构，以确立在具体情境下履行责任优先级的准则。

4. 责任界定的规范性准则：层级化优先次序的建构

基于责任伦理学对关系性、前瞻性与制度性的综合强调，并针对上述冲突的本质——即不同维度责任要求的优先级排序难题，本研究推导出用于界定 AI 替代人工护理法律责任的三条核心规范性准则。这些准则构成一个具有内在逻辑的层级体系，旨在为立法、司法与行业自治提供价值指引。

4.1 人身尊严与安全保障的绝对优先责任

第一准则为优先性准则，即人身尊严与安全保障的绝对优先责任。在任何 AI 护理场景中，保障老年使用者的人身完整性、基本尊严与生命安全，必须被视为包括设计者、生产者、部署者、运营者等所有相关技术行动者所负有的、不可克减的绝对优先责任[8]。此准则直接源自责任伦理学中基于极端非对称关系的核心要求：当一方（老年人）在身体安全与生存层面完全依赖于一个技术系统及其控制者时，优势方的首要义务即是避免对其造成根本性伤害[9]。这构成了其他一切责任的底线和前提。在适用上包括以下内容：

1. 适用场景：涉及物理交互（如助力机器人、跌倒监测与干预）、生命支持（如智能健康监护与紧急报警）、直接影响生理与心理健康的场景（如用药管理、情感交互机器人）。

2. 规范内容：法律应在此领域设定强制的、高标准的技术安全与伦理规范，并推行与之对应的严格责任或过错推定责任。开发者与部署者必须证明其已穷尽当前技术条件下所有合理的预见与预防措施，即履行了“前瞻性责任”。任何以技术创新、成本效益

或用户协议为由，试图限制或免除对此类根本性损害赔偿责任的做法，均应被视为无效。

3. 优先级：此准则具有最高优先级，当与其他价值，如商业机密、开发效率等冲突时，人身安全与尊严保障必须胜出。

4.2 促进福祉与自主性的积极建构责任

第二准则为平衡性准则，即促进福祉与自主性的积极建构责任。在保障第一准则的前提下，技术行动者负有积极利用 AI 技术提升老年人生活质量、社会参与感与个人自主性的责任[10]。此准则体现了责任伦理学的前瞻性与建构性面向，旨在实现技术对美好生活的赋能。它要求责任界定不能仅限于损害防范，还应鼓励和规范那些能够实质性提升福祉的创新。

1. 适用场景：涉及生活便利辅助、认知训练、社交促进、个性化服务优化等提升生活质量的非安全性核心功能。

2. 规范内容：在此领域，责任范式应更倾向于过错责任原则，但需结合专业标准与合理注意义务进行判断。法律与行业规范应建立“负责任创新”的指引，明确技术开发者在设计时需进行“福祉影响评估”，考量其技术对老年人隐私、自主决策、社会连接等方面的长期影响。同时，应建立透明化机制，让老年人及其代理人能理解并适度控制这些辅助功能。

3. 优先级：此准则次于第一准则，但在资源分配与政策激励上应获得充分重视。当技术方案在提升福祉的同时可能引入可控的、非根本性的新风险时，应在充分告知和征得同意的基础上，允许基于获益-风险评估的审慎探索。

4.3 基于角色与能力的差异化制度性分担责任

第三准则为协同性准则，即基于角色与能力的差异化制度性分担责任。为实现上述两级责任，必须通过制度设计，在多元行动者之间建立清晰、公平的责任分担框架。此准则对应责任伦理学的集体与制度性责任维度，旨在解决“风险分担框架”中的模糊性问题。

1. 适用场景：所有 AI 护理产品与服务的全生命周期，尤其是当损害原因涉及多环节，如算法缺陷、数据质量、不当部署、运维失当时。

2. 规范内容：首先，应通过法律与标准明确不同主体的角色基准责任：开发者对算

法固有缺陷负责；生产者对硬件安全与系统集成负责；部署运营者对适用场景的合理性评估、人员培训、日常监控与应急响应负责。其次，建立按“风险控制能力与专业水平”分配责任的原则。对于难以追溯至单一源头的系统性风险或累积性损害，如长期使用某类交互算法导致的心理依赖，应建立由行业或政府主导的共同风险基金或社会保障补充机制，以体现集体责任。这既避免了用户中心主义下对单一企业的过度追责，也防止了产业促进导向下的责任蒸发。

3. 优先级：此准则是实现前两条准则的程序性与机制性保障。在责任认定中，应优先依据清晰的角色与能力标准进行划分；对于界限模糊的剩余风险，则启动制度化的分担机制。

本套规范性准则，并非对既有立场的简单折中，而是在责任伦理学统摄下的系统性重构。它吸收并升华了“用户中心主义”对弱势保护的绝对核心关切，将其确立为首要准则；批判性转化了“产业促进导向”对发展的追求，将其约束在“负责任创新”与福祉增进的轨道上，并置于安全保障之后；具体化并完善了“风险分担框架”，为之提供了“人身安全优先→福祉促进平衡→角色能力分担”的价值排序与操作优先级，从而将模糊的分担理念转化为有章可循的层级规范。这一规范体系，因其根植于对“人-机-老”照护关系伦理本质的深刻把握，并提供了清晰的价值权衡次序，因而在应对 AI 护理责任界定的复杂规范性挑战时，更具理论的彻底性与实践的指导力。

5. 结论

本研究围绕“AI 替代人工护理的责任应如何界定”这一核心规范性问题，通过确立以“关系性”、“前瞻性”与“制度性”为核心的责任伦理学为价值基准，究深入剖析了既有学术探讨中“用户中心主义”、“产业促进导向”与“风险分担框架”三者之间的价值立场分歧，揭示了其冲突本质乃是在技术中介的非对称照护关系中，对当下脆弱个体的保护性责任、面向未来福祉的建构性责任与多元系统中的制度性分配责任之间的优先级困境。基于此，本研究推导出一个层级化的规范性准则体系：以“人身尊严与生命安全保护的绝对优先责任”为不可逾越的底线；以“促进福祉与自主性的积极建构责任”为平衡发展的指引；并以“基于角色与能力的差异化制

度性分担责任”为实现前两者的程序性保障。这一规范体系，根植于责任伦理学对具体关系、未来导向与集体行动的深刻关切，为在法理与伦理层面廓清 AI 护理场景中错综复杂的责任归属，提供了明确的价值排序与权衡原则。本文研究有助于弥补当前人工智能法学研究中规范理论基础相对薄弱的环节，为老年法学、科技伦理与侵权责任法的对话与融合开辟了新的理论路径。

当然，本研究亦存在其思辨层面的局限：所提出的层级化规范准则，仍需在具体的部门法（如产品责任法、消费者权益保护法、养老服务监管条例）语境下进行更为细致的规则转化与调适；责任伦理学作为主要价值基准，虽具强大解释力，但其与本土法律文化及既有法教义学体系的衔接，仍需进一步的批判性对话与创造性转化。未来的研究可在此基础上，进一步探索这些规范性准则在不同法系与制度背景下的适用性边界，以持续丰富和完善应对技术时代责任伦理挑战的规范性理论工具。

参考文献

- [1] Kiršienė, J., Gruodytė, E., & Mosakas, K. (2023). AI and Moral and Legal Responsibility. In J.-S. Gordon (Ed.), *Future Law, Ethics, and Smart Technologies: The Future of Legal Education* (pp. 175–189) [M]. Brill.
- [2] Jonas, H. (1984). *The Imperative of Responsibility: In Search of an Ethics for the Technological Age* (pp.68-73) [M]. Chicago: University of Chicago Press.
- [3] Krkač, K., & Bračević, I. (2021). Artificial intelligence and social responsibility. In D. Crowther & S. Seifi (Eds.), *The Palgrave handbook of corporate social responsibility* (pp. 1153–1175) [M]. Springer International Publishing. <https://doi.org/10.1007/978-3-030-42465-7>
- [4] 邵红红. 破解算法侵权责任界定的中立性难题——以“算法推荐第一案”为切入点[J]. 新闻界, 2023, (09): 71-82. DOI:10.15897/j.cnki.cn51-1046/g2.20230810.002.
- [5] 陈红梅, 张耀英. 垄断监管、责任界定、利益平衡——数字平台治理的域外经验与启示[J]. 编辑之友, 2024, (08): 99-103+105. DOI:10.13786/j.cnki.cn14-1066/g2.2024.8.013.
- [6] 吕嘉鸿, 徐娟. 智慧养老模式运行中的潜在风险及其法律应对策略[J/OL]. 医学与法学, 1-11[2026-03-20]. <https://link.cnki.net/urlid/51.1721.R.20260202.1641.002>.
- [7] 杨渐雨, 王超男. 医疗人工智能赋能健康治理: 内在机理、伦理困境及其应对策略[J]. 中国卫生事业管理, 2026, 43(02): 141-145.
- [8] 顾加栋, 黄帅. 养老人工智能应用场景下老年人人格权保护研究[J/OL]. 医学与法学, 1-7[2026-03-25].
- [9] Pasquale III F. The price of autonomy: liability standards for complementary and substitutive medical robotics and artificial intelligence [J]. *Ius et Praxis*, 2022, 28(1):157-179.
- [10] 邢源恒. 生成式人工智能嵌入健康养老的伦理挑战及治理[J]. 卫生经济研究, 2026, 43(02): 72-76. DOI:10.14055/j.cnki.33-1056/f.2026.02.003.