

法庭中的机器人：人形机器人出庭作证对 缓和证人出庭难的启示

周是伦*

摘要：聊天机器人技术的实时交互性和高度自主性特征使得人形机器人能够出现于庭并回答法官询问，且该当庭陈述具有难以预测性，以及该陈述的可靠性具有不确定性，这引发了法官能否像采纳人类证人证言一样采纳人形机器人陈述的问题。如果人形机器人可以代替人类证人出庭作证，似乎可以缓和证人出庭难的问题。然而，基于美德司法理论，人形机器人不能与人类构建社会关系且无法承担责任，法官无法从人形机器人的陈述中获得正当信念，法官基于该陈述所作出的断言不具有正当性。由于对人形机器人负责的主体难以预测到法官询问以及人形机器人陈述的内容，对人形机器人应当出庭作证的要求会间接促使这些主体出庭作证，从而缓和证人出庭难的顽疾。

关键词：人形机器人 生成式人工智能 美德司法理论 证人证言 证人出庭作证制度

* 周是伦，爱丁堡大学法学院。

目录

一、问题的提出：人形机器人能否替代人类证人出庭作证	20
(一) 提供证词的 Ai-Da 机器人	20
(二) 中国证人出庭作证的困境	21
二、像人类证言一样的人形机器人陈述	22
(一) 聊天机器人技术的特征	23
1. 实时交互性	23
2. 高度自主性	23
(二) 似是而非的人形机器人陈述	24
三、像采纳证人证言一样采纳人形机器人陈述偏离美德司法的正当性要求	24
(一) 美德司法理论对法官作出断言的正当性要求	25
(二) 人形机器人与人类之间构建社会关系的困境	26
(三) 人工智能承担法律责任的困境	27
四、以人形机器人出庭作证促使人类证人出庭作证	28
(一) 以人类证言确证人形机器人的陈述	28
(二) 我国必要证人出庭作证的制度保障	29
五、小结	30

人形机器人产业作为一种新兴未来产业，与人工智能技术的发展密不可分。人工智能与机器学习的技术创新推动了人机交互技术的进步，生成式人工智能技术作为一种新型人工智能技术，可以从海量数据源中按照用户的提示输出新内容，包括文本、图像、音频和视频等。基于这种具有交互性的技术，对话式人工智能可以让人类用户参与和人工智能的互动对话。当人机交互系统基于文本时，该人工智能通常被称为聊天机器人。^[1]OpenAI 于 2022 年 11 月底向公众发布了名为 ChatGPT 的聊天机器人 (OpenAI 2022)，作为最知名的聊天机器人，该产品由大型语言模型 (LLM) 驱动，可以帮助用户进行编码、写歌、总结文本、理解文章内容、提出创意思法、撰写文本等无数任务，还可以通过在海量数据集上进行训练来生成类似人类的回复。^[2]除了基于文本的 ChatGPT，聊天机器人行业已经开始转向实时语音和逼真的数字角色，这些虚拟的人形机器人的外观、动作和表达方式都与真人无异，这些具有人类外观的机器人形成了人形机器人产业。近期，特斯拉发布了全新一代人形机器人 Tesla Optimus Gen 2，具有全新设计的执行器和传感器，增强了对环境的适应性，其手部运动更灵活。据国际投资银行高盛预测，预计到 2035 年，人形机器人产业的市场规模有望达到 1540 亿美元，从医疗保健和养老护理到制造业和娱乐，人形机器人有望成为人类生活的许多方面的重要组成部分。^[3]

为了应对人形机器人产业不可阻挡的发展趋势，中国工业和信息化部发布了《人形机器人创新发展指导意见》以更好的利用人形机器人产业的科技创新力量，支持并推动中国数字化倡议和数字中国的建设，从而抓住人工智能、高端制造和先进材料相互融合所带来的机遇。^[4]然而，随着生成式人工智能技术在司法领域重点应用日益广泛，人形机器人在刑事司法中存在的潜在问题也引发了学术界的关注，人形机器人技术也给法治带来了新的挑战，包括法官应当如何对待那些人形机器人所生成的材料，从而符合法治的要求。

一、问题的提出：人形机器人能否替代人类证人出庭作证

本文将基于上述人形机器人产业的发展背景以及实践中人形机器人 Ai-Da 提供证词的案例，将人形机器人给刑事证据领域带来的挑战进一步具体化，即存在人形机器人能否替代人类证人出庭作证的问题。针对这一问题，本文将结合中国现有证人出庭作证制度的固有困境，分析该问题值得探讨的法律意义。

(一) 提供证词的 Ai-Da 机器人

^[1] 参见周书环：《聊天机器人的法律人格争议与传播侵权责任研究：基于 ChatGPT 场景视角》，《南京社会科学》2023 年第 6 期。

^[2] 参见杨建武、罗飞燕：《类 ChatGPT 生成式人工智能的运行机制、法律风险与规制路径》，《行政与法》2024 年第 4 期。

^[3] 参见白静：《布局人形机器人赛道 推进智能产业创新发展——解读〈人形机器人创新发展指导意见〉》，载 <https://m.fx361.com/news/2024/0118/23446231.html>，2024 年 1 月 18 日。

^[4] 同前引。

人工智能技术已经发展到人工智能机器可以回应人类询问的程度，Ai-Da 机器人就是一个典型的例子。英国 Ai-Da 由艾丹-梅勒（Aidan Meller）委托、构思、设计并发明，由工程艺术（Engineered Arts）公司制造。^[5] Ai-Da 于 2019 年首次展出，售价远超 100 万英镑，其形象是一位身着连裤袜的白人女性，深色的头发剪成波波头短发。作为世界上第一位超逼真的仿人机器人艺术家，她可以使用她眼睛里的摄像头、她的人工智能算法和她的机械臂进行绘画，其艺术作品包括绘画、素描、雕塑、行为艺术和诗歌。^[6]

值得注意的是，Ai-Da 人形机器人于 2022 年出现在英国上议院委员会的会议中，就技术对创意产业的影响作证并回答问题。^[7] 当种能够回答问题的人形机器人步入司法领域时，此时法官可以在法庭上审问人形机器人，被告也可以对该人形机器人进行质询。本文用“人形机器人的陈述”的术语表示人形机器人在法庭上回答问题和主动作出陈述所形成的信息材料，并将其区别于“证人证言”，即证人就其感知的案件情况向法官所作的陈述。^[8] 有学者认为，如果人形机器人可以出庭回答法官质询，像人类证人一样接受律师的盘问，例如该机器人作为犯罪的目击证人对犯罪事实进行描述，人形机器人在法庭中的陈述可以被视为证人证言。^[9] 此外，有实证研究表明，人形机器人可以作为一种有效和高效地记录受害者和证人高质量证据陈述的方法，相比之下，人类的证人证言属于主观性较强的证据，由于证人自身感受、记忆和表达能力是有限的，再加上受到外界压力以及证人主观愿望的影响，可能会出现不诚实作证的情况，这似乎这意味着人形机器人的陈述相比人类证人的回忆更加具有准确性。^[10]

（二）中国证人出庭作证的困境

上述人形机器人提供证词的案例引发了学术界关于人形机器人的陈述能否替代人类证人证言的问题，当这一问题放置于中国证人出庭作证制度中，就会更加凸显其法律意义。证人出庭作证，是指证人出现于法庭并提供证据证明自己所言属实的法律行为。^[11] 证人出庭作证制度，是指知悉案件情节的自然人应法院传唤在案件审理时当庭作出陈述以查明当事人有罪与否的制度。^[12] 在司法实践中，证人不出庭作证是中国刑事庭审活动多年来一个顽疾，^[13] 长期存在且根深蒂固。我国证人不出庭的成因包括侦查中心主义的诉讼构造之下的证据审查模式、司法效率的追求以及对证人拒不出庭作证行为的惩罚太轻。首先，侦查中心主义

^[5] Alexandra Anikina, *Procedural animism: the trouble of imagining a (socialist) AI*, A Peer-Reviewed Journal About, Vol. 11:1 p. 134-151 (2022).

^[6] 同前引。

^[7] 同前引。

^[8] 参见卫凯博：《“以审判为中心”背景下证言可信性审查的困境与突破》，《河北法学》2023 年第 3 期。

^[9] Brian Sites, *The Future of the Confrontation Clause: Semiautonomous and Autonomous Machine Witnesses*, Vanderbilt Journal of Entertainment & Technology Law, Vol. 22: 3, p. 547-586 (2020).

^[10] Rashid Minhas, Camilla Elphick & Julia Shaw, *Protecting victim and witness statement: examining the effectiveness of a chatbot that uses artificial intelligence and a cognitive interview*, AI & SOCIETY, p.1-17 (2022).

^[11] 参见田国宝：《中国刑事证人出庭作证激励机制反思与重构》，《法学》2021 年第 2 期。

^[12] 参见李娅、郭立梅、冯艳响：《证人出庭难的法律原因及对策思考》，《南方论刊》2021 年第 3 期。

^[13] 参见陈卫东：《从关键证人回归必要证人：关键证人出庭作证逻辑反思》，《法学研究》2023 年第 6 期。

意味着侦查主导并决定着诉讼走向，侦查阶段之后的审判阶段是围绕侦查活动所收集的证据材料进行审查，其审查的材料包括侦查阶段所形成的卷宗记载的书面证人证言，该书面证词只要和其他证据能够相互印证则可以判断该书面证词的真实性可靠性，因此通常没必要再要求证人出庭作证接受交叉质询。^[14]其次，由于司法机关的工作考核强调结案率，且案多人少、司法资源紧张等现实客观上要求法官提高庭审效率。而证人出庭会延长司法审判时间，且带来更多的诉讼环节，这会增加法官办案的程序和时间。^[15]在证人出庭作证制度下，证人一旦出庭，其所带来的新的不确定信息就可能暴露案卷笔录中的漏洞与破绽，并给法官对该证据真实性可靠性的判断带来了难题和审判压力。^[16]此外，虽然我国《刑事诉讼法》规定了证人具有出庭作证的义务，但是当证人无正当理由拒不出庭作证时，证人面临的法律后果包括强制到庭或者训诫、处 10 日以下拘留，这些法律后果对拒不出庭作证的刑事证人而言惩罚力度和威慑力都不大。^[17]

在我国目前证人出庭率低的司法现状下，如果以人形机器人的出庭作证替代证人出庭，且如果人形机器人陈述能够替代人类证人证言，似乎可以缓解我国证人不出庭所带来的困境，且该困境的解决具有必要性。刑事证人出庭作证率低的问题与我国以审判为中心的司法改革要求和直接言词原则的立法趋势相背离。推进以审判为中心的刑事诉讼制度改革要求确立审判程序的核心地位，而非侦查中心。庭审实质化的核心内容是，控方和辩方当庭提出的有证明价值的证据材料，都能对庭审和裁判结果走向发生实质性影响。这就要求法官亲自听取控辩双方当庭提出的证据和主张，包括证人在法庭上所提供的证词，从而实质化地审查判断证据的相关性、可靠性和合法性。^[18]在此司法改革的进程之下，学术界呼吁在我国新刑事诉讼法的立法过程中确立直接言词原则，即所有证据都应在法庭上出示，所有定案的证据都要形成于法庭，在法庭通过调查和辩论，完成举证、质证、认证，使法官形成心证结论。^[19]

综上，生成式人工智能技术的发展使得人形机器人具有主动提供证词并回答问题的能力，这引发了法学界关于人形机器人的陈述能否替代人类证言的问题。该问题在我国证人出庭作证制度中可以凸显出重要的法律意义，我国证人出庭难现状具有解决的必要性，以人形机器人出庭替代人类证人出庭似乎可以缓解我国证人出庭难的顽疾，从而符合我国庭审实质化的司法改革和直接言词原则的立法趋势。

二、像人类证言一样的人形机器人陈述

[14] 同前引。

[15] 参见程金华：《中国法院“案多人少”的实证评估与应对策略》，《中国法学》2022年第6期。

[16] 参见元轶：《庭审实质化压力下的制度异化及裁判者认知偏差》，《政法论坛》2019年第4期。

[17] 同前引[12]。

[18] 参见熊晓彪：《中国刑事司法证明“唯客观化”批判》，《法学研究》2024年第2期。

[19] 参见陈卫东：《〈刑事诉讼法〉第四次修改前瞻》，《政法论坛》2024年第1期。

人形机器人 Ai-Da 提供证词的案例为人形机器人在证据领域的研究提供了场景，本章节将基于人形机器人所依托的聊天机器人技术的特征，分析人形机器人陈述替代人类证人证言的可能性，并进一步对人形机器人陈述的证据属性，即该陈述的可靠性进行评估。

（一）聊天机器人技术的特征

1. 实时交互性

人形机器人所依托的聊天机器人技术为人形机器人带来了实时交互性和高度自主性的特征，进而为人形机器人在法庭上像人类回答法官询问提供了可能性。聊天机器人是通过自然会话语言与人类用户互动的机器会话系统，该技术可以应用于社交机器人的开发中，例如社交机器人的技术设备，包括护理机器人、博物馆导游等。^[20]这些社交机器人的一个共同点是，它们都或多或少地依赖于机器人与人类之间的交流，因此也依赖于信息的交换。当社交机器人具有酷似人类外形时，用户与这种人形机器人的互动应表现出从与人类互动中相同的情况，即其信息的交换具有实时交互性。这意味着人形机器人能够及时的回复人类的对话，通过自动实时对话与人类用户进行互动，这种对话被认为是连贯的和自然的。^[21]详言之，人形机器人在作出独立判断的同时，能够对其最初的结论提出质疑，并通过根据不同的环境调整经验和数据，表现出与特定环境相一致的行动能力和情感反应倾向。

2. 高度自主性

此外，生成式人工智能技术以其高度自主性为特征，能够通过算法和机器学习生成新的信息，而这些信息并不是人类直接引导的，甚至不完全为人类所理解。其高度自主性表现为生成式人工智能可以通过输入较少的数据生成更多的信息，并且可以在不同的情形下，自主更新并根据不同情景调整其生成的结果。例如生成式人工智能技术的绘画功能可以根据目击者对嫌疑人的描述自动生成嫌疑人肖像，^[22]再者，美国哥伦比亚的法官使用 ChatGPT 来帮助论证其推理、回答法律问题并直接生成判决文书。^[23]聊天机器人技术作为一种新型生成式人工智能技术，也具有高度自主性。人形机器人作为一种自主或半自主装置，可以通过机器学习技术使其能够与周围环境互动，并能够在独立于人类干预的情况下自主做出决定。在聊天机器人技术的实时互动性与高度自主性的结合之下，如果把人形机器人掩盖住，我们可能无法分辨该证词是由人类提供还是由机器提供，这种技术特征人形机器人陈述替代人类证言提供了可能。以 2015 年《机械姬》(Ex Machina) 中艾娃这个角色为例，艾娃作为一个机器人，被精心设计成能够对一系列人类道德情感和行为做出恰当的反应，表现出与人类极

^[20] 参见李晟：《国家安全视角下社交机器人的法律规制》，《中外法学》2022 年第 2 期。

^[21] Marcel Finkel & Nicole C. Krämer, *Humanoid robots—artificial. Human-like. Credible? Empirical comparisons of source credibility attributions between humans, humanoid robots, and non-human-like devices*, *International Journal of Social Robotics*, Vol.14:6, p. 1397-1411 (2022).

^[22] Massimo Leone, *From fingers to faces: Visual semiotics and digital forensics*, *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique*, Vol.34: 2, p.579-599 (2021).

^[23] 参见於兴中、郑戈、丁晓东：《生成式人工智能与法律的六大议题：以 ChatGPT 为例》，《中国法律评论》2023 年第 2 期。

为相似的特征。艾娃通过了图灵测试，这表明“她”可以被视为人类，如果艾娃的存在被隐藏起来，只留下“她”的声音，人们可能真的会把“她”误认为人类。^[24]

（二）似是而非的人形机器人陈述

然而，聊天机器人的技术特征也会给评估人形机器人陈述是否真实可靠带来了挑战。人形机器人陈述具有似是而非的特征，即人形机器人陈述似乎是真实可靠的，然而也可能并非如此。这是由于高度自主性的聊天机器人技术导致人形机器人陈述的生成过程难以被控制，这意味着人们难以在该证据的生成过程中将其不可靠性不真实的信息进行控制和纠正，容易导致人形机器人生成错误的结果。在一个案例中，一位用户“玛丽亚”使用大型语言模型（LLM）对婴儿的病情症状进行描述，获得的对话结果是建议给孩子服用阿司匹林，并表示孩子早上应该会好起来。然而结果并非如此，由于没有及时接受适当的治疗，婴儿将有长期的认知障碍。该用户起诉了搜索引擎算法的创建者，认为搜索引擎应承担损害赔偿责任。搜索引擎公司认为，由于聊天机器人功能包含有关其准确性的警告和免责声明，该用户本应意识到该回复不具有权威性。^[25]据此，OpenAI 在 ChatGPT 的主页上发布了一条警告：聊天机器人偶尔会生成错误信息，用户不应就聊天机器人生成对话的真实性和可靠性抱有过高的期望，不应该完全相信聊天机器人。^[26]

同样，在司法实践领域中，这种可靠性的似是而非给法官的论证带来了挑战。似是而非的论证是指根据诉讼过程中某一时刻的证据，得出一个似乎为真的假设结论，但如果在诉讼过程的后期出现新的信息，该结论可能会被撤回。正如有学者指出，人工智能技术所带来的幻像，即往往给人们造成了认为该结果是真实可靠的错觉，在司法领域中，需要关注法官如何回应这种可靠性不确定的情况，法官对于人形机器人陈述是否具有可靠性的评估是在不确定条件下的进行的推论，这种不确定性以为着法官对证据属性的评估是一种暂时接受的结果。

综上，由于人形机器人所依托的聊天机器人技术具有实时交互性和高度自主性的特征，这为人形机器人作出陈述并回答问题提供了技术支持。然而，然而，由于聊天机器人技术具有高度自主性，导致人形机器人生成的错误信息在其生成过程中难以被识别和消除，人形机器人的陈述具有难以预测性以及其陈述是否具有可靠性是不确定的，这给法官应当如何在不确定性的情况下审查人形机器人陈述的可靠性带来了挑战和难题。

三、像采纳证人证言一样采纳人形机器人陈述偏离美德司法的正当性要求

^[24] Jess Hutto-Schultz, *Dicitur Ex Machina: Artificial Intelligence and the Hearsay Rule*, *Geo Mason L Rev*, Vol. 27:2, p.683-718 (2020).

^[25] Maura R. Grossman et al., *The GPTJudge: Justice in a Generative AI World*, *Duke Law & Technology Review*, Vol.23:1, p.1-26 (2023).

^[26] Chirag Arora, *Proxy Assertions and Agency: The Case of Machine-Assertions*, *Philosophy & Technology*, Vol. 37:1, p.1-19 (2024).

诚然，人类证人也会作出虚假的证言，关于回答法官能否像采纳人类证言一样采纳人形机器人陈述这一问题，需要将视角从拘泥于判断证言的真实性可靠性转变到法官作出断言（判决）的正当性视角，即法官形成对于人形机器人陈述具有或者不具有可靠性的断言应当基于正当信念。在美德司法理论视角下，人形机器人无法承担责任且无法与人类之间构建社会关系，因此人形机器人的陈述无法为法官提供正当信念，法官不应当相信人形机器人陈述。换言之，如果法官直接采纳人形机器人陈述，则意味着该断言不具有正当性，偏离美德司法理论的要求。

（一）美德司法理论对法官作出断言的正当性要求

在司法决策的过程中，法官不仅被要求做出判决（断言），还要求该判决具有正当性。^[27]美德法学理论关注如何使得法官作出的断言具有正当性，认为当且仅当一个美德的法律决策者在类似情况下做出断言时，该断言才是正当的。^[28]美德司法理论指出具有正当性的断言取决于法律决策者的个人美德。^[29]对该个人美德的评价并不完全取决于行为的结果是否正确，相反，它关注的是实施行为的个人的美德品质，这种方法将评估行为本身的道德与评估行为人的道德品质区分开来。例如，即使一个人出于错误的原因做出了正确的行为，虽然该行为是正确的，但在道德上却不值得称赞。换言之，一个人如果出于好意犯了错，他仍然可以被认为是具有美德的。美德司法理论将美德划分为道德美德和智力美德。道德美德评价的是一个人的道德品质，而不仅仅是他的行为，其清单包括智慧、勇气、仁慈、正义、诚实和忠诚等。^[30]智力美德包括良好的记忆力、感知力、健全的认知能力，以及先验直觉和准确推理的能力。^[31]智力美德鼓励人们在有充分理由相信自己的信念并为自己的信念辩护。实践智慧是将道德原则应用到现实生活中，在具体情况下做出符合道德规范的决定的能力。^[32]它不仅包括上述道德美德的要求，还包括人们如何在不同情况下落实这些道德美德，包括在实践中应当如何驾驭复杂的道德困境，个人如何根据理性思考做出合乎美德的行为等。^[33]

在证据法领域中，美德司法理论要求法官有正当理由相信断言为真或假，从而控制法官应该形成或不应该形成哪些信念。在该视角下，证据法应被理解和设计为一个结合道德和认识论原因的领域，美德司法理论可以为可采性的刚性约束提供正当性的证成，也可以对法官形成对证据可靠性的信念进行柔性约束。这种美德司法理论的责任观反映体现在上述证据分析方法的内在视角中。在法律决策过程中，证据分析方法的外在视角理论和内在视角理论都

^[27] Luiz Silveira, *Discovery and justification of judicial decisions: towards more precise distinctions in legal decision-making*, *Recht en Methode in onderzoek en onderwijs*, Vol.2014, p.9-27(2014).

^[28] Amalia Amaya & Hock Lai Ho, *Law, virtue and justice*, Hart Publishing, 2012, p.10.

^[29] Amalia Amaya, *Virtuous Adjudication; or the Relevance of Judicial Character to Legal Interpretation*, *Statute Law Review*, Vol.41:1, p.87-95 (2019).

^[30] Paul Bloomfield, *Virtue Epistemology and the Epistemology of Virtue*, *Philosophy and Phenomenological Research*, Vol.60:1, p.23-43 (2000).

^[31] 同前引。

^[32] Claudio Michelon, *Practical wisdom in legal decision-making*, University of Edinburgh School of Law Working Paper, Vol.13, p.1 (2010).

^[33] 同前引。

从认识论的角度判断证据的可采性，以追求事实认定的准确性，排除不公平偏见和认知风险。尽管如此，这两种证据分析方法体现了信念合理性和错误可消除性方面的差异。外在视角理论关注最终评估的状态，以追求结果的准确性为最高目标，它认为错判风险是不可避免的，而证据规则所能做的就是分配错误风险。^[34]内在视角关注的是追求真理（知识）过程的合法性，它主张证据法应该是规范和限制真理发现活动的法律。^[35]例如，它关注事实发现者认知过程中存在不公平偏见的风险，并搁置对结果准确性的假设。更重要的是，它假定事实发现者不仅受到认知上的约束，而且受到道德上的柔性约束。

证据法应当是针对决策者（法官）的，更具体的要求是关注法官的美德品格，并判断法官作出的断言是否具有正当性。证据的内在分析视角可以回答法官应当如何对待这些新证据从而形成具有正当性的断言，例如证人证言规则也关注到法官作出断言的过程中会不断面对新的证据，其中包括证人的举止以及被告席上的被告的举止。^[36]证据是能够证明其他信息的信息，当证人 W 就 H 事件作证时，其证词立即提供的信息是“W 说 H”，然而，这一信息本身并不能证明 H 为真或假。该视角还渗透了一种法律认识论的观点，即法官并不是科学家，法庭中的认知活动也区别于科学领域中探寻真理的认识论模式。在科学领域中，科学家们通过实验和不断的假设检验来获取知识作出结论，而法庭程序中的知识并不是通过实验作出的，而是通过对抗式辩论形成的说服性的社会认识论，不仅仅是一个能否作出断言的问题，还是一个该断言是否具有正当性的问题，对于这种正当性的判断美德司法理论的约束。

（二）人形机器人与人类之间构建社会关系的困境

在判断人类能否基于人形机器人陈述产生正当信念时，其一是要探索人形机器人与人类之间的是否存在能够产生这种正当信念的社会关系。美德，例如谦逊的价值在于它肯定所有个体的平等价值，并通过以平等为基础的社会和政治互动促进平等的社会关系。^[37]在人类社会中，模仿典范有助于鼓励社会成员传承美德，培养植根于情感、互助、关怀和爱护的社会关系。^[38]美德的社会属性代表了一种倾向，在这种倾向下，人们会自然而然地、下意识地遵守社会规范，并以这些社会规范的方式行事。^[39]基于这些社会关系，一方的言行可以激发人们的信念。例如，许多父母对子女与生俱来的关爱体现了一种美德倾向。这种本能向子女和其他社会成员传达了一种信息，即他们可以信任并期望父母也这样做。^[40]虽然事实可能并非

^[34] 参见熊晓彪：《迈向“最佳解释确信”的司法证明》，《法制与社会发展》2023年第5期。

^[35] 参见熊晓彪：《中国刑事司法证明“唯客观化”批判》，《法学研究》2024年第2期。

^[36] 参见熊晓彪：《证据评价的自由与规范法理》，《中国政法大学学报》2023年第4期。

^[37] Maartje MA de Graaf, *An ethical evaluation of human-robot relationships*, *International journal of social robotics*, Vol.8, p.589-598 (2016).

^[38] 同前引。

^[39] 参见张彦、胡俊：《品格教育中榜样示范的问题与回应——以亚里士多德美德论为考量视角》，《道德与文明》2020年第3期。

^[40] 参见韩燕丽：《道德典范示范效应再检视——一种基于美德伦理学的分析》，《华中科技大学学报(社会科学版)》2020年第1期。

如此，但这种对父母具有美德的信念向外界传达了一种期待，即如果他们不符合这种期待，例如虐待子女，他们将受到道德谴责甚至虐待罪的惩罚。

在社会关系理论的框架内，社会关系是在人与人之间形成的，人工智能无法被视为社会关系结构的一部分。^[41]换言之，人类与人形机器人之间无法存在或发展平等的社会关系。例如，人类可以购买人形机器人，但购买人类则是非法的。诚然，护理人员也可能从护理机器人那里获得重要的精神支持，护理机器人可以在护理人员睡觉时轻松监测生命体征、药物和疼痛程度，它们还能对药物相互作用或新症状发出警告，而这些往往是人类无法察觉的。护理机器人相当于一个人工智能代理人，与人类之间建立代理关系，人类作为被代理人可以指定代理人（人形机器人）追求特定的目标并实现特定的价值，代理人应当满足被代理人的需要并维护被代理人的利益，这种代理与被代理的关系区别于人类关爱实践中通过互惠来维系人与人之间的关系。由于人类与人形机器人之间不存在这种可以传递期望和合理信念的社会关系，人类也不会自动对人形机器人产生正当信念，而是对人工智能有条件的信任。例如，这些条件包括对人形机器人的使用应当遵守法律标准和道德原则。总之，一切美德都是人与人打交道时所要求的，而不是人与机器人之间的关系。正是因为人形机器人与人类之间缺少这种可以产生并传递正当信念的社会关系纽带，法官无法基于人形机器人的陈述产生正当信念。

（三）人工智能承担法律责任的困境

主体的责任性意味着该主体能够对自己行为的后果负责。根据美德司法理论，法官应当基于可承担责任的主体提供的知识中作出断言，这意味着能够承担责任的主体可以通过敢于承担犯错误的风险来换取法官的信任。^[42]这种美德责任主义与知识正义理论相一致，后者认为决策者应将其信念建立在证据的基础上，或在认识论上负责任，并形成正当的判断。^[43]在证据法领域中表现为对传闻证据的排除，禁止采纳传闻证据不是因为该传闻证据本身不可靠，而是因为法官对该证据是否可靠是不知道的，是无法确定证据的可靠性的。^[44]详言之，如果证人 S 就 p（陈述）向法官（H）作证，那么证人 S 必须有意让 H 认识到：（i）证人 S 相信其陈述的 p 为真；（ii）证人 S 有意让（H）基于（i）相信 p；如果没有证人 S 的存在，那么法官（H）就没有理由相信证人 S 关于 p 的证词。例如，卷宗中写有“S 说 S 听到 A 说他要杀人”这句话，只能证明“S 听到 A 说他要杀人”这句话是 S 说的，而不意味着“A 要杀人”这句话是真的，此时并不存在关于“A 要杀人”的知识（真理）。况且由于 S 没有出庭证明 S 相信“A 要杀人”是真的，所以无法知道 S 是否希望法官相信 A 要杀人，也不存在关于该主张的知识和正当信念像法官的传递。此外，在证人证言规则中，

^[41] Ahyeon Kim et al., *Effects of gender and relationship type on the response to artificial intelligence*, *Cyberpsychology, Behavior, and Social Networkin*, Vol. 22:4, p. 249-253 (2019).

^[42] Christopher Tollefsen, *Justified Belief*, *American Journal of Jurisprudence*, Vo 1.48, p.281-296 (2003).

^[43] 同前引。

^[44] 参见刘金松：《数字时代大数据辅助司法证明的构造及其风险防控》，《现代法学》2024年第1期。

如果机器人作证可以替代人类证人，可能会引发作证不再是人类不可推卸的责任，将出庭作证的责任推卸给机器人的危险。

然而，这种基于证人敢于承担错误风险来换取法官信任的信念传递，只存在于人与人之间，而不存在于人与人工智能之间，因为人工智能无法对其陈述负责。^[45]换言之，一个无法承担自主责任的实体无法向法官传达这种信念信息，法官必然无法据此形成正当信念。反之，如果直接采纳人形机器人的陈述则会陷入道德运气，这意味着在缺乏正当信念的情况下，尽管该证据具有准确性，也只不过是撞上了好运气，如果该证据不具有准确性，那也是不幸碰上了坏运气，而不是存在一个需要为该运气负责的主体。这种道德运气本身就不正当的，会使得人类将自己决定的命运交给人工智能，无异于听天由命。对这种道德运气的抵抗需要人类进行对风险进行背书，需要对该人形机器人证据负责的人类主体的证人证言将该证据转化为人类陈述，法官从而基于该人类陈述生成正当信念。综上，法官无法仅仅基于人形机器人的陈述来产生正当信念，因此，法官作出的断言不能直接来源于人形机器人的陈述。否则不仅将一个人的命运托付给了一个对结果漠不关心的实体，即人形机器人，还违背了美德司法对法官作出断言正当性的要求。

四、以人形机器人出庭作证促使人类证人出庭作证

面对生成式人工智能技术不可阻挡的发展趋势，需要思考的是如何更好的利用新兴技术以造福人类，而非一律禁止使用新兴技术。同样的，法官不应当像采纳人类证言一样采纳人形机器人的陈述并不意味着禁止人形机器人出庭作证，而是应当思考如何利用人形机器人出庭作证的现象从而更好的推动我国证人出庭作证制度的完善。由于人形机器人陈述具有不可预测性，对人形机器人陈述负责的人类主体会关注人形机器人在法庭上的回答情况，为促使这些人类主体出庭作证提供了可能。同时，近期我国学术界讨论的必要证人出庭作证制度为坚持并要求这些人类主体出庭作证提供了制度上的保障。

（一）以人类证言确证人形机器人的陈述

人形机器人与对其负责的主体（包括开发者和制造者）之间形成了代理与被代理的关系，这意味着人形机器人被拟制为代理人，这符合万物有灵论对于算法代理的论述，即将非人类赋予人格的地位，可以表现自己、发挥作用、说话并被倾听。有学者进一步认为人形机器人可以作为人类证言的代理人，^[46]人们在参与社会实践中与代理人打交道，相信代理人的话是真实可靠的是基于对被代理人的相信，这种信任关系正如美德司法理论所要求的，只存在人类与人类之间，而非人类与人形机器人之间。具体而言，人们相信人形机器人对话的信息是准确的意味着相信开发人形机器人的人能够对这种技术输出结果的准确性进行保障，包括承

^[45] 参见张琳琳：《人工智能司法应用的责任归结困境与解决路径》，《当代法学》2023年第5期。

^[46] 同前引[26]。

担输出结果不准确的后果、修复错误输出结果。且当人形机器人“说谎”时，人们不会将人形机器人的说谎归咎于人形机器人本身具备的不良品质，而是会指责人形机器人的设计缺陷，即应当将人形机器人设计成符合道德原则的机器。人工智能研究人员都有责任积极主动地确保他们正在开发和使用的技术不仅不会侵犯人权，而且还能促进人权。例如，开发护理机器人的程序员应当为护理机器人设立程序防止该机器人损害人类健康。

基于这种代理与被代理的关系，人形机器人陈述的内容会影响到人们对于人形机器人背后的开发者制造者的信赖情况。具体而言，当人形机器人陈述存在错误时，开发者会有动力作出指正，防止这种错误的信息影响到自身利益。当人形机器人在法庭中面对法官询问时，由于开发者无法预测法官会提问什么，且人形机器人陈述本身就具有不确定性，开发者为了确保自己利益不受损害会更加关注人形机器人在法庭中的表现，例如开发者通过参加庭审关注人形机器人在法庭上陈述什么，陈述的内容是否符合开发者的期待。因此，上述代理与被代理的关系会间接促使开发者出现于庭，此时法官应当通过对开发者进行确证，从而将人形机器人的陈述转化成证人证言，并形成对该证人证言可靠性的正当信念。这也符合上述美德司法理论对法官作出断言的正当性要求，即法官的断言应当基于知识和正当信念。具体的确证思路是指证明案件需要两个或两个以上的来源，如果一名证人的证言与第二名证人的证言在基本细节上一致，从而确证了第二名证人的证言。例如在苏格兰法律中，证人证言必须有两个或两个以上的来源才能被视为证据。这种确证的方式区别于我国相互印证模式，在相互印证证明方式之下，无需证人出庭即可达到判断证人证言真实性的目的，而确证的方式强调直接言词原则中要求人形机器人陈述和人类确证的证言都呈现于庭，这也符合前文中关于庭审实质化的司法改革要求和直接言词原则的立法趋势。

（二）我国必要证人出庭作证的制度保障

基于人形机器人与对其负责的主体之间的代理与被代理关系，对人形机器人陈述负责的人类主体应当属于必要证人，我国必要证人出庭作证制度为这些主体的出庭作证提供了保障，并有助于补足我国关键证人出庭作证的制度缺陷。关键证人出庭作证逻辑意味着证人并非原则上必须出庭作证，只有关键的少数证人才应当出庭作证，法官通过预先审查庭前证言的证明力的方式判断该证人是否属于关键证人。^[47]有学者指出，在我国推进庭审实质化的司法改革背景下，应当摒弃关键证人出庭作证逻辑，回归必要证人出庭作证的逻辑，从而合理界定我国《刑事诉讼法》第192条第1款的构造和内涵。^[48]

相比之下，必要证人出庭作证的逻辑要求证人能出庭的都应当出庭，这意味着在其主张只要符合证据关联性、作证必要性条件就有必要出庭，法官不必预先审查庭前证言的证明力，而是在法庭上通过法庭辩论、质询等环节的过程中对证人证言的证明力予以查明，这符合庭

^[47] 参见张迪：《关键证人出庭例外情形下的作证制度研究——以司法融合新型技术为视角》，《中国刑警学院学报》2021年第5期。

^[48] 同前引[14]。

审实质化的要求。^[49]从证人出庭作证制度的法律条款角度看，证人是否出庭取决于法院对“必要性”的裁量，该裁量等同于刑事定罪标准的审查，^[50]即证人是否有必要出庭作证，要看已有的证据是否已经确实成分，形成法官对于案件有罪或无罪的内心确信，可以作出断言。基于上述美德司法理论的正当性要求，法官不应当像采纳人类证言那样采纳人形机器人陈述，因此如果没有人形机器人背后的开发者的证言进行确证，则无法基于人形机器人陈述生成具有正当性的断言，对于其要求人形机器人背后的开发者出庭作证是具有必要性的。在此正当性的要求之下，对人形机器人负责的主体属于我国《刑事诉讼法》第192条第1款中的“人民法院认为证人有必要出庭作证”中的“有必要”的主体。综上，由于人形机器人和对人形机器人负责的主体之间可以视为为代理与被代理关系，法官可以通过要求人形机器人的出庭作证间接促使对人形机器人负责的主体出庭作证，从而可以对该人形机器人的陈述进行确证，并将其转换为人类证人证言，从而形成正当信念。从证人出庭作证的制度要求角度看，这种代理与被代理的关系赋予了对人形机器人负责的主体出庭作证的必要性，即属于该主体必要证人，应当作为证人出庭作证据此在诉讼程序中受到证人的待遇和保护。

五、小结

综上所述，人形机器人 Ai-Da 提供证词的案例引发了学术界关于法官能否像对待人类证人证言一样对待人形机器人的陈述以及能否相信人形机器人的疑问。证人出庭作证难一直以来被认为是我国刑事司法实践中的顽疾，严重阻碍着庭审实质化的实现，该顽疾具有解决的必要性。如果将人形机器人的陈述视为人类证人证言，似乎可以缓解我国证人出庭难的顽疾。然而，由于人形机器人所依托的聊天机器人技术具有实时互动性和高度自主性的特征，其人形机器人陈述的可靠性具有不确定性。诚然，人类证言也存在虚假的可能，如果仅仅拘泥于判断证据属性的视角，即判断人形机器人的陈述和人类证言的可靠性程度，则会陷入似是而非的论证，即不知道人形机器人的陈述是否真实可靠，而仅仅是暂时接受一个假象的可靠性。

美德司法理论看到了证据的可靠性与正当信念的关系，可以回答法官能否像对待人类证人证言一样对待人形机器人陈述的问题。该理论将其关注的重点从证据属性的判断转变到关注法官生成断言的过程，即要求当法官有正当理由认为断言为真或假时，该断言才具有正当性。由于人形机器人无法与人类构建社会关系且无法承担责任，法官无法从人形机器人的陈述中获得正当信念，因此无法作出具有正当性的断言。这有助于更好的控制人形机器人陈述的可靠性的不确定性所带来的错误判决风险，使法官避免直接采纳人形机器人的陈述，从而符合法治要求。

^[49] 参见肖杨钟：《论庭审实质化与证人出庭作证》，《铁道警察学院学报》2021年第3期。

^[50] 参见李奋飞、匡旭东：《证人出庭作证“必要性条款”解释适用研究》，《法学杂志》2023年第5期。

从证据制度的构建及完善角度看,虽然法官不应当像采纳人类证言一样采纳人形机器人的陈述,但是并不意味着人形机器人应当被禁止呈现于法庭,反而应当通过要求人形机器人出庭作证,间接促使对其负责的主体,包括人形机器人的开发者和制造者等出庭作证,从而使得这些主体可以更好的应对法庭询问人形机器人问题的不确定性和人形机器人陈述本身的不确定性所带来的损害自身利益的风险。基于人形机器人与对其负责的主体之间的代理关系,可以将这些主体其解释为必要证人,这些主体的出庭作证还受到我国必要证人出庭作证制度的保障。美德司法理论对证据规则的分析在英美证据法学界并不少见,期待在新兴技术的背景之下,未来能够有更多关于证据规则与美德司法理论的本土化研究。