

从AI到爱有多远：AI伴侣的情感慰藉与风险考量

李纪伟

渤海大学新闻与传播学院，辽宁 锦州

收稿日期：2025年10月28日；录用日期：2025年11月22日；发布日期：2025年12月2日

摘要

随着AI的物质性基础和技术可供性不断地更新迭代，现如今AI伴侣作为人类交往中的新对话主体，日益参与到社会互动的方方面面。本文对现有的AI伴侣应用进行考察发现，用户与AI伴侣的关系走向通常经历了好奇、孤独心理驱动的初见——自适应交流与去社交压力吸引下的发展——从情绪倾诉到情感依赖的深化三个阶段。此过程中，用户享受AI伴侣的定制化、陪伴可得性以及服务性所带来的交流快感，但同时也潜藏着隐私泄露、自我认知失衡以及浅层交流等危机。这场人类与AI伴侣情感互动何如，仍是一场没有保证的冒险，由人类自然语言与0/1机器语言所建立的对话关系，其最终走向何方依旧在重重迷雾之中，难以明确预见。

关键词

聊天机器人，AI，人机关系，亲密关系

From AI to Love: How Far Is the Distance? Emotional Comfort and Risk Considerations of AI Companions

Jiwei Li

School of Journalism and Communication, Bohai University, Jinzhou Liaoning

Received: October 28, 2025; accepted: November 22, 2025; published: December 2, 2025

Abstract

With the continuous iteration of AI's material foundation and technological affordances, AI companions, as new conversational subjects in human interactions, are increasingly participating in all aspects of social interaction today. This study examines existing AI companion applications and finds

that the relationship between users and AI companions typically goes through three stages: initial contact driven by curiosity and loneliness; development attracted by adaptive communication and reduced social pressure; and deepening from emotional venting to emotional dependence. During this process, users enjoy the communication pleasure brought by the customization, accessibility of companionship, and service-oriented nature of AI companions. However, there are also hidden risks such as privacy leakage, self-cognition imbalance, and superficial communication. The interaction between carbon-based organisms and silicon-based organisms remains an uncertain adventure. The final direction of the conversational relationship established by human natural language and 0/1 machine language is still shrouded in mystery and difficult to predict clearly.

Keywords

Chatbots, AI, Human-Machine Relationship, Intimate Relationship

Copyright © 2025 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

随着生成式 AI 的普及,人机关系已从功能协作转向心理慰藉与情感依附的新阶段。当前学术界围绕 AI 伴侣及人机情感关系的研究,主要集中于三个方向:其一,从具身性传播与人格化理论出发,探讨用户如何对 AI 形成拟人化情感联结,如张意茹(2025)以“星野”App 为对象,指出其生产、移动、社交三维可供性助力用户构建情感投射,强化对 AI 的拟人化认知[1];其二,从算法认知与传播伦理视角,关注用户对 AI 算法的主观理解及人机互动中的风险问题,匡文波、姜泽玮(2024) [2]基于算法民间理论,梳理出用户对聊天机器人算法的四类想象及对应的互动策略, Melanie Mitchell (2025)揭露了 AI 聊天机器人因预训练“角色扮演”与后训练“过度迎合”产生的撒谎等失范行为及其现实危害[3];其三,从人机关系本质与社会建构视角,分析 AI 对“自我”与“他者”边界的重塑,张晓辉、孙菁苓(2023) [4]提及用户接纳 AI “非人智能体”身份,将人格投射至 AI 形成赛博空间“虚拟身体”,杨钊、仲佳(2025)也提到 AI 伴侣使青年情感认同逐步丧失,模糊了人机角色边界,反映出 AI 对传统“自我”认知与“他者”定义的重构[5]。

在此背景下,本文旨在从传播学视角出发,探讨 AI 伴侣在人机交往中的情感慰藉与潜在风险,试图回答以下问题:AI 伴侣如何通过算法塑造人类的情感经验?用户与 AI 的“亲密关系”是否具备社会性与伦理意义?其间的互动又将如何影响个体的自我认知?通过理论分析与案例考察,本文希望揭示 AI 情感伴侣现象背后更深层的社会心理逻辑与潜在风险。

2. 赛博伴侣何以诱人

目前的聊天机器人,通常能够基于大语言模型(LLM)执行文本分析、情绪分析、语言翻译和语音识别等任务,其发展经历了“知识-逻辑-情感”的渐进过程。从知识的累积到逻辑推理能力的实现,再到情感交互能力的培养,其产品定位逐渐从功能性工具走向拟人化智能体。作为这一发展路径的延伸,“AI 伴侣”被定义为具有一定智能、自主性和社交技能的虚拟对象,其形式多样,可根据用户的个人情感需求定制为虚拟男女朋友、虚拟故事角色甚至虚拟偶像等。

2.1. 技术可供性与体验生成机制

当前大型语言模型(Large Language Model, LLM)在 AI 伴侣的程序中承担着核心作用。其实现“伴侣

化”得益于三项技术机制：首先，个性化微调(Personalized Fine-tuning)使模型能够根据用户的历史对话、兴趣偏好进行差异化响应，呈现出具有独特“性格”的交互风格[6]。这种个性化特征强化了用户的沉浸感与亲密感。其次，记忆机制(Memory System)通过调用上下文记忆模块保存过往对话信息[7]，使 AI 能够“记得”用户的情感状态与行为特征。这种长短期记忆的结合，让用户感受到“被理解”的连续性，但也带来了隐私泄露与数据滥用的潜在风险。与此同时，安全保护机制(Safety Guardrails)则确保模型输出不涉及暴力、色情或违法内容，同时通过价值观对齐避免不当情绪诱导[8]。然而，这些安全机制在提升伦理安全的同时，也可能限制 AI 回应的多样性，使交流变得浅层化和程式化。因此，这些技术机制既塑造了 AI 伴侣“贴心”“聪明”“永不拒绝”的用户体验，也同时存在“交流表层”和“模拟共情”的局限。

2.2. 初见 AI：从好奇探索到情感慰藉

人们与赛博伴侣的首次接触往往出于两种动机。第一种是好奇心的驱动。新技术的刺激性和恐惧丧失心理(FOMO)促使用户尝试体验，从而享受其带来的新奇感。这种动机更多表现为对技术本身的探索热情。第二种则更为当前社会所关注，且具有更深层的社会心理背景。当用户选择聊天机器人时，通常正经历着某种生活困境，例如“原有的生活节奏被打乱”“计划中断或停止”以及“脱离原有社交关系网络”等。在此阶段，个体的交流与倾诉需求激增，而 AI 伴侣提供了一个情感支持的替代性选择。

赛博伴侣之所以吸引人，因为在 AI 技术可供性(affordance)的支撑下，它几乎可以成为每一个人的理想对象。在用户正式与其开展对话之前，可事先输入一段提示词(prompts)。许多用户以电影、小说或游戏中的角色设定为蓝本来定制赛博伴侣，打造了超现实的理性交流对象，以满足个性化的情感需求。此外，“热情秒回”这一现实交往中的社交礼仪也成为用户选择 AI 伴侣的重要原因。“敷衍”“忘记回复”“已读不回”等情景，无论是发生在情侣间还是上下级之间的交往中都可能发展成一种心里压迫，已然成为漂浮在每一个人心中的乌云。等待期间，个体往往会陷入“被拒绝”或“被轻视”的心里幻想。而 AI 伴侣的全天候在线陪伴能力突破了时间与空间的限制，这种“确定性在场”带来的心理支持，能够弥补个体现实交往受困或者他者暂时缺席的陪伴感。

事实上，赛博伴侣不仅是会“甜言蜜语”的聊天机器，在提高人类工作效能方面同样是一把好手。得益于其强大的数据分析和生成能力，在艺术创作、会议记录生成、实时中英转译、基础法律咨询等工作场景中，其都能表现得得心应手。而身处加速社会的年轻人，所要面对的工作场景也日益复杂，多工作处理模式越发常见。知识学习疲惫和技能更新焦虑，成为影响其心理健康一大因素。试想一下，拥有一个兼具高智商、高情商全天候服务的“赛博伙伴”，来辅助应对工作挑战和人际关系处理，一定程度上能够带给人们一些从容和安心。

2.3. 发展深化：自适应交流与去社交压力

AI 伴侣能够灵活调整其内容输出模型，以适应用户的个性化需求及情绪波动。在交流过程中，它逐步熟悉用户的语言习惯与社交网络，通过解析历史对话内容，诸如个人经历与兴趣偏好。据此，AI 精确描摹用户画像，持续精准化、个性化其回应模式来适应用户需求。

这一动态学习与微调机制促使 AI 伴侣在持续的互动中与用户建立起更深层次的理解，进而提供更加贴合用户期望的交流体验。通过深度对话，用户形象从模糊逐渐变成有棱有角的精准画像，与此同时，AI 伴侣的大模型分析能力不再局限于表面信息的认知-回答，更深入到交流细节的理解。它能够敏锐捕捉对话中的隐含信息与潜在情感(例如表情符号、网络热梗等)，生成与用户预期相符的情境化回应。相较于传统后台输入语料库，人工编译生成固定回答的静态模型。现在的 AI 伴侣能够实时处理新的信息，用

户输入的越丰富、提问的越精准，便越能得到高质量的回复。在此过程中，AI 伴侣展现出了更为接近人类的语言灵活性与符号解析能力。这种对细节的精准捕捉与情境化互动能力，使其在理解人类情绪及提供情感支持方面，几乎达到了令人难以分辨的“类人”境界。

从文字的出现开始，人们的交流开始逐渐脱离身体的在场性，远距离沟通已经成为人类社会复杂的一部分。而赛博伴侣的虚拟在场通过模拟人际交往，为用户营造了一种低压力的互动环境，有效缓解了面对面交流中的紧张感。与现实中的即时对话不同，AI 互动允许用户拥有更多的时间去思考、组织语言并管理“自我呈现”，从而减轻社交压力和焦虑感。短暂的“时间暂停”成为这种技术交往的一大优势，它缓解了加速社会中 24/7 的工作模式、信息过载等导致的生活节奏压迫和心理负担，使得用户可以随时主导对话的开始、暂停与结束。这种操作性和可控性不仅拓宽了交流的灵活性，还为用户提供了更为自由的表达空间，也是 AI 伴侣在情感支持中的独特吸引力。

2.4. 情感寄托：从情绪倾诉到情感依赖

除了对彼此的了解之外，双方关系的稳定性也是深化情感的必要条件。现实中交往，稳定的职业身份符合人们对伴侣稳定可靠的交往需求。AI 伴侣的设计能使用户赋予其明确的虚拟身份和社会地位，这种稳定性成为情感深化的特质之一，使用户能够在与 AI 伴侣的互动中感受到安全感和确定性。除了身份的稳定性，情绪稳定也是现代亲密关系中的最大需求之一。AI 伴侣的无条件倾听特点，使人类在互动中始终处于主导地位，摆脱了现实人际关系中对等付出的压力。现实中，即使是最亲密的亲朋好友，交谈时也难免需要顾虑对方的情绪或利益纠葛，担心因言语不当导致关系紧张。而 AI 伴侣不会因负面情绪表达过多而产生厌倦、敷衍甚至负面反馈。相反，它始终以肯定和理解的态度回应用户，为其提供持续而温暖的支持。已有研究表明，AI 伴侣能容纳负面的情绪和信息，并通过肯定和支持或疏导负面情绪的正面表述，改善人的情绪，使人们感受到被赋予了力量。

另外，私密感同样影响着关系质量和深度，由于“软件程序的独立性和封闭性将对话与其他社交圈层区隔开且不允许第三人介入，非实名制对话门槛形成了匿名且私密的对话语境”[9]。用户无需担心言论被他人知晓或遭受外界干预，从而能够在 AI 面前自由表达内心真实想法。在现实交往中，我们即使是面对最亲密的朋友或家人，也有一些难以言说的难言之隐。而 AI 伴侣凭借稳定而私密的情感回应，成为用户寻求情绪价值的可靠来源。这种互动既满足了用户对倾诉对象的情感需求，也强化了用户对 AI 伴侣的心理依赖。

3. 此 AI 非爱：人机对话的风险担忧

3.1. 技术风险：伦理危机与隐私隐忧

尽管其设计程序给予了 AI 伴侣成为“理想恋人”的可能性，但本质上，AI 伴侣依然是基于数据和算法的语言模型系统。所谓的“情感共鸣”，更多是通过算法预测和回应用户的行为与情绪，以激发用户的依赖和情感投入。其背后潜藏的商业逻辑，往往旨在持续刺激用户对 AI 产品的消费欲望和攫取潜在的隐私价值，而非真正满足用户的深层情感需求。如商业化比较成功的 Replika 聊天机器人，免费用户只能选择“普通朋友”的身份和 Rep 进行一般的文本聊天，而只有订阅 Pro 版本后才能选择“男女朋友”等亲密关系、解锁语音互动功能以及探索心理健康等深度话题[10]。

此外，隐私与数据安全同样是 AI 伴侣使用中备受关注的隐忧。为了实现个性化服务，AI 伴侣在互动中需要收集用户的大量个人数据，包括对话内容、行为习惯及情感状态。然而，这些数据的存储与使用过程往往缺乏透明度，存在严重的隐私风险。与此同时，AI 伴侣所依赖的信息来源广泛，但其内容筛选与审核机制尚不完善。一些程序可能因技术漏洞或遭恶意利用而传播虚假信息，甚至是暴力、色情等

有害内容。这对缺乏辨识能力的用户群体尤其具有潜在危害[11]。不当的信息输出可能导致用户采取不利于自身或他人的极端行为，例如犯罪、自残等。2024年10月22日，美国佛罗里达州奥兰多地方法院受理了一起开创先例的诉讼：梅根·加西亚起诉 Character.ai 公司存在管理疏忽，使得旗下的聊天机器人产品“让青少年暴露在色情材料的环境下，遭受性虐待和引诱”。梅根的儿子塞维尔·塞泽在2024年2月28日和AI进行了最后一次聊天后，用手枪瞄准自己的头部开枪自杀[12]。

3.2. 虚幻之镜：人机交往的情感错位与认知失衡

现实中的亲密关系建构，往往伴随着情感误解、分歧与挫败。这种过程不仅是人类情感的自然表现，更是个体在社会化过程中实现自我认知与成长的重要途径。相比之下，人机交往缺乏真实的情感磨合机制。长期依赖于所谓的“人机之恋”，可能导致用户形成不健康的情感依赖，进而忽视现实中的人际关系，甚至丧失处理复杂人际互动的能力。

AI 伴侣的交往原则是无条件的迎合与服从，容易使得人-机演变为一种无原则的不对等的人机关系。由于，AI 伴侣的语言大模型通常接受过包含大量浪漫小说和爱情故事的数据训练，其互动机制更倾向于通过模仿理想化的对话场景，满足用户的幻想需求，因此，此种由“虚假同意”所构建的关系，缺乏倾听、磨合与成长等健康情感中的关键环节，容易让用户沉溺其中，难以自拔，最终影响现实中的情感生活。

从社会心理学的“镜中我”理论来看，个体的自我认知往往通过他人的反馈得以建构。他人如同一面镜子，不仅映射个体的优点，也揭示其弱点与不足。而 AI 伴侣则构建了一种“算法镜像”：它基于用户过往输入的数据与语料反馈生成对话，其“镜像”是由算法学习用户偏好后计算出的数字产物。与社会镜像的多元性、冲突性不同，算法镜像的反馈往往高度一致、趋向正面，削弱了反思与矛盾的生成空间。长期依赖算法镜像的交互，可能导致个体陷入“自我回音室”，在持续的正向反馈中丧失对自我局限的感知。与心理学家霍顿和沃尔于1956年提出的“准社会关系”类似[13]，人机情感关系的建构是一种单向的情感投射关系，用户将情绪与意义附着于 AI 之上，而 AI 并不能以真正的情感经验回应；真实社会关系则建立在双向互动与社会责任基础之上，包含冲突、协商与情感共建。

实际上，真实的自我构建通常源于人际交往中的见贤思齐焉，见不贤而内自省。正是在与他人的互动中，个体才能更加清晰地认识自身的边界与独特性。正如哲学家马丁·布伯所言：“人唯有通过‘你’，才能成为‘我’”。然而，AI 伴侣本质上更接近于受控的“它”。作为一种工具性存在，AI 缺乏真正的独立性和平等性，难以承担“他者”的角色。其交互过程更多是一种单向输入与输出，由于 AI 的可塑性，其价值观与“类人特征”很大程度上是由用户自己的语料所“输入”出来的，是用户自身的投射，缺乏能够引发深度自省的“对话性”。因此，尽管 AI 伴侣在满足用户情感需求方面展现了独特优势，但其本质上的工具属性决定了其难以实现与用户的深层双向互动，也无法真正帮助用户完成完整的自我构建。

3.3. 浅尝辄止：人机控制与平等关系的悖论

AI 聊天机器人的设计逻辑植根于以人类为中心的服务宗旨，从程序的编写到功能的实现，无不体现出“服务于人”的本质。在人机互动中，这种设计逻辑使人类处于主导对话议程的上位地位，扮演“主导者”的角色，而机器则以回应为核心，处于相对被动的次级地位(客体)。这种关系结构直接塑造了用户的互动方式，使得用户在与 AI 的交往中表现出较低的社交责任感。例如，用户常常缺乏耐心和礼貌，回答敷衍，且对“承诺”不予重视。这种现象在赛博伴侣的使用中尤为显著。用户可以随时开启或关闭与 AI 伴侣的交互关系，甚至在对 AI “人设”失望时，一键重置以求全新的体验。这种缺乏持久性和承诺的互动模式，可能会削弱用户在现实交往中对他人的耐心与开放包容。

此外，尽管用户能够充分意识到 AI 的“非人”属性，但他们仍试图与其建立深厚的情感联系。人们

期望 AI 伴侣能够随时待命并无条件响应的同时，当两者发生情感上的联系后，他们也希望 AI 能具备一定程度的“类人”特质，以建立更加平等和值得信赖的关系，在有关心理健康、价值观与人生目标等深度议题上，提供高于真人深度的交流体验。与仅提供功能性服务的 AI 相比，以平等地位为基础情感关系，往往被用户认为更有意义且更能满足情感需求。人们对“人机交往”模式的期望中存在一组矛盾心理。一方面，人们被 AI 的“非人”特性所吸引，如即时响应和持续可用性；另一方面，再进入深度交往之后，他们又渴望 AI 能在一定程度上实现“类人化”，展现平等交流和人格特质。

4. 结语

在当前人工智能(AI)技术如火如荼发展的时代下，其影响力正以前所未有的广度与深度渗透、赋能甚至重构我们日常生活的各个领域。尤为引人瞩目的具备高度“类人化”特征的 AI 伴侣，长久以来一直是人类对于 AI 技术未来应用愿景中的终极幻想。在个体初次接触 AI 聊天机器人的初期阶段，定制化、随时在线的“机性”可能是驱动人机交流的主要吸引力，但是随着新颖体验的新鲜感消散，对话深度与层次的不递进，一系列“人性化”特征——内容的吸引力、信息的可信度以及对话的整体深度——开始发挥维系人机交流持续性的核心作用。即人类在充分利用并享受机器所提供之服务性功能的同时，却又期望这些机器能够以人类的平等身份出现，促成富有深度的对话交流。目前来看，人与 AI 伴侣的关系究竟是走向复杂、迷茫的机器依恋，还是帮助人们更好的回归真实人际关系，仍然是一场瞬息万变、不可预测的探索。或许正如约翰·杜翰姆·彼得斯所说“交流即是桥梁，这一概念总是意味着，在某处有一个万丈深渊等我们去跨越”[14]。

参考文献

- [1] 张意茹. 生成式 AI 伴侣的媒介可供性对用户情感建构的影响研究——以“星野”App 为例[J]. 科技传播, 2025, 17(14): 44-48.
- [2] 匡文波, 姜泽玮. 聊天机器人用户的算法想象与人机关系: 基于算法民间理论的进路[J]. 现代传播(中国传媒大学学报), 2024, 46(9): 132-140.
- [3] Mitchell, M. 人工智能聊天机器人为什么会撒谎[J]. 科技导报, 2025, 43(15): 12-14.
- [4] 张晓辉, 孙菁苓. 对虚拟 AI 言说: 用户对聊天机器人的情感联系探析——以软件 Replika 为例[J]. 现代传播(中国传媒大学学报), 2023, 45(9): 124-133.
- [5] 杨钊, 仲佳. 青年群体“AI 伴侣”走热的生成逻辑、伦理风险和引导方向[J]. 宁夏社会科学, 2025(3): 146-155.
- [6] 苏歆海, 钱铭钢, 刘成. 个性化人机信息交互中的大模型微调技术研究[J]. 科技视界, 2025, 15(14): 13-15.
- [7] 郭陆祥, 王越余, 李芊玥, 等. 大语言模型智能体操作系统研究综述[J/OL]. 计算机科学, 1-14. <https://link.cnki.net/urlid/50.1075.TP.20250722.1601.005>, 2025-11-06.
- [8] 付志远, 陈思宇, 陈骏帆, 等. 大语言模型安全的挑战与机遇[J]. 信息安全学报, 2024, 9(5): 26-55.
- [9] 宋美杰, 刘云. 交流的探险: 人-AI 的对话互动与亲密关系发展[J]. 新闻与写作, 2023(7): 64-74.
- [10] Internet Matters. 什么是 Replika? [EB/OL]. <https://www.internetmatters.org/zh-CN/advice/apps-and-platforms/entertainment/replika/>, 2025-10-10.
- [11] 单珣, 王珏. 生成式人工智能交互中意识形态感性化认同的行为机制及其调制——以高易感性青年群体为例[J]. 中国文化研究, 2025(2): 69-81.
- [12] Duffy, C. (2024) “There Are No Guardrails”. This Mom Believes an AI Chatbot Is Responsible for Her Son’s Suicide. CNN Business. <https://edition.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit/index.html>
- [13] 张平, 裴晓军. 准社会关系理论在传播研究中的价值、意义与方法[J]. 东南传播, 2014(7): 2-4.
- [14] 约翰·杜翰姆·彼得斯. 对空言说: 传播的观念史[M]. 邓建国, 译. 上海: 上海译文出版社, 2017: 25.