

情感人工智能时代的用户体验塑造：人机共情的新路径¹

郭景

(北京邮电大学经济管理学院, 100876; mailguojing@bupt.edu.cn)

摘要：随着情感人工智能技术的迅猛发展，其在社会各个领域的应用不断加深，尤其在塑造用户体验方面展现出前所未有的潜力。本研究旨在呼吁整合多学科的丰富知识，从全面而深层次的角度探讨人机共情在服务接触中对用户体验的影响机制，以及如何通过情感层面的营销管理来提升用户体验。为此，本文回顾了人机共情的发展背景和理论基础，强调其在水机交互中的关键作用。其次详细论述了人工智能作为交互式代理在建立共情能力方面的相关技术准备。最后通过理论阐述指出人机共情是一种双向互动过程，包含人工共情能力与共情反馈行为两个主要维度。本文不仅为人机共情的理论和实践提供了新的研究视角，也为营销管理领域带来了新的用户体验塑造策略。

关键词：情感人工智能；人机共情；人机交互；非对称式社会互动；用户体验；营销管理

1 研究背景

随着技术的不断进步和社会需求的不断增长，人工智能已经从实验室走入社会生活中，与人类建立了各种形式的亲密关系。智能化应用正加速渗透社会各个领域，深刻变革用户体验的创造方式^[1]。当前，人工智能技术从简单模拟人类逻辑思维和任务执行，已发展至对人类情感的理解和模拟，使人机交互过程更具情感化^[2]。基于此，用户体验的塑造面临新的挑战：既需深刻理解人机交互对用户体验的作用机理，又需在情感层面调整营销管理模式。在实践中，具有共情特征的人工智能，包括智能助手（如 EVI、Siri、小爱同学）、智能助理（如 Google Assistant、Alexa）、智能数字人，以及生成式系统（如 chatGPT、SORA 等），开启了情感人工智能应用的新场景^[3]。一个重要议题是人工智能的共情能力，即机器人能否理解或感受他者的情感状态，并做出相应行为反应。

然而，鲜有研究关注人工智能在服务接触中共情能力对用户体验塑造的作用机制。众所周知，营销人员的共情能力是创造顾客体验的重要因素，企业品牌若要赢得消费者的心，需要与顾客建立情感联系^[4]。如何创造更有沉浸感和幸福感的顾客体验，始终是企业营销的难题之一。由于人工智能感知和处理能力不断增强，其处理社交和情感线索并与人进行自然地互动已经成为可能^[5]。目前，人工智能逐渐成为人类助手甚至伙伴，具备了更多的共情特征，承担了诸多一线营销人员的工作职能。然而，人机之间的社会互动面临情感、认知、道德等方面的能力不对称，构建人机共情关系是一项挑战性任务。人机社会互动不同于人与人之间的互动，也不同于以往计算机媒介中的人机互动，具有能力非对称、互动双向性和互动实时性的特点^[6]。

尽管人机共情在理论和实践中具有重要意义，但现有研究在概念上尚无统一且清晰的定义和分类标准来描述不同类型或程度的人工智能共情能力，在理论上缺乏系统框架来解释和预测人工智能共情能力对用户体验的促发机制和效果。为填补这些研究空白，本文旨在从人机共情的视角整合相关领域的研究成果，聚焦以下问题：如何认识人机共情的理论内涵，以及如何认识人机共情对人类用户体验的促发机制。为此，本文通过人机交互的现象描述与文献回顾，探讨人工智能共情能力的定义及其在实验研究中的构建过程。

¹ 基金资助：北京邮电大学博士生创新基金资助项目（CX2023103）

2 人机交互中的共情基础

21 世纪以来,人工智能技术迅猛发展,使得人机交互从传统的任务执行和功能性操作,逐渐拓展到情感层面。共情在这一过程中成为一个核心概念,用于描述和分析人类与智能系统之间的情感交流与理解。共情是一种人类在社会交往中常用的心理技巧,它可以帮助人类理解他人的想法和感受,从而建立信任关系和亲密关系^[7]。从人机交互的指向性来,应存在两种关系:一种是“人对机”,人类占据主动;一种是“机对人”,机器占据主动。

“人对机”关系可以追溯到人类对工具及物品的情感依附。这种情感关系在技术还不发达的早期阶段已经出现,例如人们对长期陪伴的物品产生情感依恋、对具有互动性的“智能物”产生探索欲。然而,受限于当时技术和历史条件,没有完备的理论来解释这种对“智能物”的情感关系。在“人对机”模式中,人类占据主动地位,主要通过输入指令和互动来控制机器。在此过程中,人类的情感依赖与互动需求成为情感共情的基础。尽管机器在这类关系中是被动的,但人类对机器的依恋和情感投射为共情基础奠定了扎实的情感依托。

另一方面,“机对人”关系则是随着人工智能技术的发展逐渐演变而来的,强调智能系统主动识别、理解和回应人类情感。此时,机器不仅仅是被动的工具,它们通过对人类情感的理解和模拟,识别和回应人类的情感需求并主动参与到人机交互之中。机器在交互过程中具有一定的主动性,能够基于情感线索主动调整行为策略,从而实现更深入的情感共情。从理论层面来看,随着人工智能技术的发展,人机交互中的主动性分配正在经历深刻变革。在“人对机”模式下,共情基础主要来源于人类对机器的情感依恋和互动需求,这种关系以人类为主体、机器为辅。而在“机对人”模式下,共情基础则依赖于机器对人类情感的理解和模拟,机器通过主动互动来影响人类的情绪和决策。

3 对共情的认知挑战

共情一词最早出现在 19 世纪末的心理学研究中,意为“理解他人思想”,是一种道德和审美的愉悦,存在情感交流的可能性^[8]。当人类经历共情的时刻,会把他人的情绪、态度等内化到自己身上,再从自己的角度、经历来推测他人的心理状态。这样就能够在一定程度上体验到他人的处境和心境,从而产生情感共鸣和情感反应。

当前对共情的认知挑战在于人工智能是否具有共情能力。这最早可以追溯到对机器人社交功能的探索。社交机器人是一个跨学科的研究领域,早在上个世纪 90 年代就已经引起了心理学领域学者们的关注。随着人工智能技术的不断进步,对社交机器人的情感属性和服务属性的研究范畴不断扩大,前者侧重情感表达与识别,后者侧重以机器人的身份提供各类的服务。社交机器人开始渗透到日常生活中,人与机器之间的亲密关系也开始受到了关注。这种亲密关系涉及到人类对于机器人是否具有共情能力的认知和评价,以及机器人如何模拟和展示共情能力^[9]。人与社交机器人的互动会形成深层次的情感联系,甚至可能发展为爱情或友谊。如果机器人要成为人类的伙伴,就需要有一种能够理解和感受人类情感的共情能力,即让机器人能够感知和回应人类的情感状态,并建立和维持良好的亲密关系。由于亲密关系对于人机互动的的话语机制和社会影响,使得机器人不再被视为单纯的物质对象,而是被赋予了社会属性和身份^[10]。而且这种亲密关系改变了人们对于机器人的认知和态度,使得机器人被视为有情感和意志的个体,而不是无感无意的工具。

高度老龄化的国家面临着严重的劳动力短缺和社会保障压力,因此也需要借助具有情感属性和社会属性的机器人来缓解服务人员劳动力不足的问题。日本的浅田捻教授基于机器人在社会中的作用和价值,提出机器人不仅可以帮助人类完成各种物理任务,还可以满足人类的心理需求。其研究的动机和目标之一,就是把人类的共情机制嵌入机器人系统,通过增加机器人的共情能力以改善人机交互过程,让机器人更加适应和融入人类社会^[11]。浅田捻等人将其定义为“人工共情”(artificial empathy),指那些社交能力建立在激发人类情感反应能力基础上的机器人的行为^[11, 12]。作为一种技术手段,其目的是让机器人更加适应人类的社会环境和情感需求,增强人机交互的效果和体验。改善人机交互过程的最终目标是实现人机协同,让机器人成为人类社会的一员,与人类共同创造和享有价值。浅田捻教授的研究不仅对于提高社交机器人的性能和适应性有着重要的意义,也对于探索人类情感和人工共情的本质和机制有着深远的影响。

4 人机共情的内涵

北京大学颜志强教授团队在《心理科学》期刊发表的论文中，将“artificial empathy”定义为“智能体共情”^[13]，这一定义强调了机器人系统在模拟人类情感交流过程中的能力。与此同时，何双百教授提出了“人工移情”的概念，并首次引入了基于互动产生的人机共情概念，进一步丰富了该领域的理论框架。此外，共情在语言翻译和文化语境中具有多重含义。“empathy”一词源自德语“Einfühlung”，意指能够深入理解和感受他人的情感。在英文中，“empathy”既可以描述机器对人类情感的响应，也可以描述人类对机器的情感投入。而在中文语境中，根据具体的应用情境，“empathy”可以被翻译为移情或共情，这反映了情感交流的双向性。鉴于何双百教授和陈多闻教授对共情概念的深入探讨，将“artificial empathy”译为“人工共情”，特指在机器人系统中嵌入共情机制，以实现机器对人类情感的模拟和响应^[10, 14]。

为了理解人机共情的内在机理过程，首先需要从人与人之间的互动寻找共情理论基础。传统意义上的共情通常发生在两人以上的竞争、协作、交流的社会互动过程中，而且要求共情主体和共情对象或者说社会互动的双方，都具有意识、情感等心理能力。然而，在人与机器的互动中，无论是高阶还是低阶社交能力的机器人，当前技术水平下它们普遍缺乏真正的意识、自主性和创造性等高级心理能力。这些能力是人类共情过程中的核心因素，能够促进人与人之间的深层次情感交流和理解。因此，人与机器人之间的互动在本质上存在心理能力的不对称。在这种心理能力不对称的互动中，机器虽然能够通过情感计算和编程模拟表现出一定的情感反应，但这种模拟情感往往缺乏真实性，无法与人类复杂、动态的情感体验相匹配。因为人类的情感含有主观体验和情境依赖性，这些特质目前的人工智能技术尚难以完全复制，机器情感响应不能与人类的情感体验同步。

从阶段性来看，人机共情是一个具有时间顺序、包含双向互动的过程。首先是机器对人类的共情关怀（机对人），该阶段涉及人工智能系统利用算法模拟人类共情行为和情感交流过程的能力，包括情感状态的识别、理解以及相应的情感响应，以主动识别和响应人类情感需求。其次是人类用户对共情机器人的情感体验与行为参与（人对机），该阶段关注人类用户如何感知和回应机器人的共情行为，包括人类用户对机器共情表现的情感体验及互动行为的反馈。人工共情能力体现了人工智能系统在情感层面的智能，共情反馈行为揭示了人类用户在情感交流中的主体性和参与度。

而从维度性来看，人机共情不是线性的时间过程，而是存在于情感、认知、行为维度，每个维度都可以独立塑造人机交互的共情体验。情感维度聚焦人机交互过程中情感状态的识别、共享和响应，指智能系统能够模拟人类的情感状态实现情感层面的互动。认知维度侧重于智能系统理解和处理人类情感背后的认知过程，以准确解读用户的情感表达和意图，并做出恰当的回。行为维度关注在人机共情过程中人类用户与智能系统之间的行为互动，强调用户在互动中对智能系统表现出的情感反应。情感维度强调情感识别与模拟，认知维度则注重情感理解与处理，而行为维度关注互动行为反馈，三者的协同构建了复杂的人机互动体系。

机器引发人类用户共情的机制可归纳为两大类：行为状态驱动型与情绪模仿驱动型。行为状态驱动型理论认为，用户对机器的共情感知源于对机器外部特征及其所处环境的观察^[15]。具体而言，机器人或代理的物理外观、面部表情、语言语调等特征，以及其执行的任务、目标与难度等环境因素，均可能对用户的情感态度产生影响。此观点强调了用户基于对机器外部特征与环境的观察来形成情感与态度。情绪模仿驱动型理论则认为，用户对机器的共情感知源于机器对用户情绪状态的共情性响应^[16]。换言之，机器人或代理能够识别并理解用户的情绪，并以共情的方式作出反馈。通常，这一过程涉及“情绪模仿”策略，即机器通过模仿用户的情绪状态来激发共情反应。此观点强调了机器对用户情绪状态的识别、理解与反馈能力，以及在此基础上建立的共情沟通与协作关系。

5 人工共情的构造与设计

学界正致力于对共情概念的扩展与模型构建，以阐释不同过程如何相互作用以产生共情。这些模型的目标是捕捉人类情感的复杂性，并将其应用于人工智能系统的情感模拟。人工共情的实现涉及神经网络、自然语言处理（NLP）、计算机视觉和生物传感器等先进技术的应用，这些技术被集成使用以捕捉和分析人类的

语言、声音、面部表情、身体姿势等，从而提升机器人识别和响应人类情感状态的准确性。在情感模型的构建领域，已有研究提出了多种理论框架。例如，德瓦尔的“俄罗斯套娃模型”、德赛第等人的“双加工模型”、以及明斯基的“情感机器”以及“批评家-选择器模型”。这些模型基于传统共情理论，为人工共情的模型构建提供了理论基础。

5.1 套娃模型

德瓦尔提出了“套娃模型”，该模型创新性地共情反应划分为三个层次：情绪感染、情绪共鸣和情绪/认知共情^[17]。情绪感染指的是个体对他人情绪状态的无意识模仿；情绪共鸣则体现为有意识的情绪同步；而情绪/认知共情则涉及对他人情绪状态的理性理解和评价。这三个层次在功能和神经生理层面上各有特点，为理解不同物种间共情反应的差异以及同一物种内部共情反应的变化趋势提供了理论基础。

基于德瓦尔的“套娃模型”，浅田捻等人进一步提出了智能体共情模型，旨在机器人中实现类人共情反应。该模型探讨了情绪共鸣和情绪/认知共情两个关键机制的实现。情绪共鸣与运动模仿紧密相关，依赖于镜像神经元系统。通过开发一个计算模型，浅田捻等人模拟了婴儿通过感觉-运动映射与照料者互动时的关联学习过程^[12]。该模型将机器人的学习过程分为两个阶段：建立运动命令与视觉影像之间的联系，以及在此基础上发展出区分“我”与“非我”的能力。这一模型对于研究机器人如何形成更自然的行动模式，提升其感知人类情绪的能力具有重要意义。

浅田捻等人的智能体共情模型与德瓦尔的“套娃模型”在理念上保持一致，都强调了个体拥有一套由核心共情系统和学习系统构成的共情反应机制，且学习系统能够通过与环境的互动不断优化核心共情系统。这些研究为理解共情反应的机制和实现人工共情提供了重要的理论基础和技术支持。

5.2 双加工模型

德赛第的研究指出，通过自下而上的自动化加工通路与自上而下的认知加工通路相结合，能为机器人表征情感提供理论支撑^[18]。自下而上的自动化加工通路主要负责初步的情绪识别和反馈，提供即时的情感反馈，可以使机器人迅速识别出人类用户情绪变化并作出基本反应。自上而下的认知加工通路涉及到更复杂的情感识别与响应策略，使机器人能够根据情境和用户特定需求，对情感信息进行深层次的理解和处理，确保情感反应的准确性和适时调整，从而提供更为人性化的互动。双加工模型的双通路设计使得智能体在交互过程中能够实现情感的多层次处理，更接近人类的情感交流模式，提升其共情能力。

罗德里格斯等人提出了一个基于双加工模型的智能体共情机制理论框架^[19]。该框架认为智能体共情涉及两个核心加工过程：共情反应特征的描述和通过智能体行为体现共情内在原理。共情评价系统使智能体能够根据事件线索进行情绪评估，共情反应系统则使智能体基于这些情绪执行适当的行为。智能体需利用场景知觉和情绪识别来筛选情绪，并计算情绪与场景的相关性，以选择适应当前情境的行为，展示合适的共情反应^[20]。

从双加工模型视角看，情绪共情和认知共情构成了人工共情的两个主要维度。人工共情设计需同时考虑情感反应的感性层面和认知评价的理性层面。感性层面上，智能体通过神经镜像等机制模拟他人心理状态，产生相似或同步的情绪反应；理性层面上，智能体利用元认知机制区分自我与他人，调节情绪和行为，并基于目标和环境选择最优策略。这两个层面的相互作用，共同构成智能体共情的完整过程。基于双加工模型，人工共情可细分为情绪识别、情绪共鸣、情绪理解和情绪回应四个连续阶段，这一理论框架也为智能体共情机制的研究提供了新的视角。

5.3 批评家-选择器模型

明斯基的理论框架将人类认知过程划分为六个层级：本能反应、后天反应、沉思、反思、自我反思及自我意识情感。本能反应作为基础层级，而道德和价值观则位于最高层级。中间层级承担着处理日常事务、冲突与目标的职能，涵盖常识性问题。明斯基进一步提出，在每个层级中，存在“批评家”与“选择器”的动态机

制：批评家负责识别问题与挑战，而选择器则基于批评家的评估，选择适当的资源或策略以应对挑战。这一机制在特定情境下触发，形成动态的“批评家-选择器”循环，策略的选择和执行导致情境变化，激活或关闭其他层级的批评家。人类情感状态的转变涉及大脑资源的激活与抑制，从而影响大脑运作模式。情感状态变化可能引发资源的“级联”效应，其广泛性决定了情感状态的多样性。情感状态可视为资源相互作用的产物，这些资源的运作方式是行为驱动的关键。资源的演化包括先天与后天学习获得的，它们代表了大脑中多样化的结构和过程，从认知到行动再到思维方式的转变，涉及特定或广泛的脑区功能^[21]。

明斯基阐释了心理批评家的概念，这些批评家通过经验学习和反馈调整形成资源，影响不同层次的认知与行为。心理批评家分为三种类型：纠正性警告、外显抑制和内隐束缚。纠正性警告是一种本能反应，用于在紧急或威胁情境下发出强烈信号，促使立即中断当前行为并采取应对措施。外显抑制与学习或规则相关，它在准备执行可能导致错误或失败的行为时发出警告，促使重新考虑或修改计划，体现了人工智能系统在提供建议时的信息检查功能。内隐束缚与价值观相关，它在思考过程中抑制可能导致错误的想法，促使人工智能系统在服务提供中寻求符合人类需求的解决方案。心理批评家在构建共情能力方面起着关键作用。它们通过分析外部刺激（如面部表情、语音语调）和内部状态（如用户的过往行为和偏好），识别用户的情绪状态。基于识别到的情感，选择器激活相应的情感反应机制，模拟人类的情感共鸣过程，使人工智能系统能够产生类似人类的情感信号。选择器根据批评家的评估，选择适当的行为策略，以适当的方式回应用户的情感状态，如通过语音、文字或行为表达关切或支持。心理批评家在人工共情的构造中扮演着至关重要的角色，不仅识别和响应用户的情绪状态，还通过分析和评估情感背后的原因和含义，为人工智能系统提供深层次的理解和适应性行为策略。

6 结束语

共情弥补了机器“物”的局限性，能够帮助人工智能走出“只是机器”的思维框架，转而走向“社会伙伴”定位^[10, 22]。本研究对情感人工智能的人机交互这一新兴领域进行深入探讨，旨在揭示机器人共情行为的内涵以及人机共情创造体验的过程。机器人的共情表现，无论是通过语言还是非语言的方式，均有可能对用户的体验造成直接影响，进而塑造顾客的互动行为。值得注意的是，机器人共情行为的潜在效应不容忽视。如何在人机交互的框架下，有效利用具备共情能力的机器人以促进人际交流，而非取而代之，已经成为一个既具有理论价值又涉及伦理考量的关键议题。

未来的研究亟需在以下关键领域取得突破：首先，明确人机共情的定义与分类，深化对共情概念及其边界的理解，并探索不同类型与程度的人机共情。其次，揭示人机共情的作用机制，构建心理与神经科学的理论模型，发展能够模拟人机共情的计算方法与交互技术。最后，探讨人机共情在营销管理领域的理论应用与实践价值，研究如何借助人机共情提升人机交互体验，促进人类福祉，并探索其在服务创新与产品开发中的潜力。

参考文献

- [1] 姜婷婷, 许艳闰, 傅诗婷, 等. 人机交互体验研究: 为人本人工智能发展注入新动力[J]. 图书情报知识, 2022,39(04): 43-55.
- [2] CHEN Q, LU Y, GONG Y, et al. Can AI chatbots help retain customers? Impact of AI service quality on customer loyalty[J]. Internet Research, 2023,33(6): 2205-2243.
- [3] LIU-THOMPSON Y, OKAZAKI S, LI H. Artificial empathy in marketing interactions: Bridging the human-AI gap in affective and social customer experience[J]. Journal of the Academy of Marketing Science, 2022,50(6): 1198-1218.
- [4] 沈鹏熠, 李金雄, 万德敏. “以情动人”还是“以理服人”? 人工智能聊天机器人角色对顾客情感依恋的影响研究[J]. 南开管理评论, 2023.
- [5] YALÇIN Ö N. Empathy framework for embodied conversational agents[J]. Cognitive Systems Research, 2020,59: 123-132.
- [6] PENTINA I, XIE T, HANCOCK T, et al. Consumer-machine relationships in the age of artificial intelligence: Systematic literature review and research directions[J]. Psychology & Marketing, 2023,40(8): 1593-1614.

- [7] MCBANE D A. Empathy and the salesperson: A multidimensional perspective[J]. *Psychology and Marketing*, 1995,12(4): 349-370.
- [8] YALÇIN Ö N, DIPAOLA S. Modeling empathy: building a link between affective and cognitive processes[J]. *Artificial Intelligence Review*, 2020,53(4): 2983-3006.
- [9] JECKER N S. You've got a friend in me: sociable robots for older adults in an age of global pandemics[J]. *Ethics and Information Technology*, 2021,23(S1): 35-43.
- [10] 何双百. “机器同伴”:新型亲密关系下的“人机共情”现象思考[J]. *西南民族大学学报(人文社会科学版)*, 2021,42(07): 185-192.
- [11] ASADA M. Towards Artificial Empathy: How Can Artificial Empathy Follow the Developmental Pathway of Natural Empathy?[J]. *International Journal of Social Robotics*, 2015,7(1): 19-33.
- [12] ASADA M. Development of artificial empathy[J]. *Neuroscience Research*, 2015,90: 41-50.
- [13] 颜志强, 苏金龙, 苏彦捷. 从人类共情走向智能体共情[J]. *心理科学*, 2019,42(2): 299-306.
- [14] 陈多闻, 汪姿君. 何为人工共情?——当代新兴人工共情技术本体的哲学凝思[J]. *自然辩证法研究*, 2023,39(09): 59-63.
- [15] EREL H, TRAYMAN D, LEVY C, et al. Enhancing Emotional Support: The Effect of a Robotic Object on Human-Human Support Quality[J]. *International Journal of Social Robotics*, 2022,14(1): 257-276.
- [16] CONCANNON S, TOMALIN M. Measuring perceived empathy in dialogue systems[J]. *AI & SOCIETY*, 2023.
- [17] de WAAL F B M. Putting the Altruism Back into Altruism: The Evolution of Empathy[J]. *Annual Review of Psychology*, 2008,59(1): 279-300.
- [18] DECETY J, JACKSON P L. The functional architecture of human empathy[J]. *Behavioral and cognitive neuroscience reviews*, 2004,3(2): 71-100.
- [19] RODRIGUES S H, MASCARENHAS S, DIAS J, et al. A Process Model of Empathy For Virtual Agents[J]. *Interacting with Computers*, 2015,27(4): 371-391.
- [20] LIM A, OKUNO H G. A Recipe for Empathy[J]. *International Journal of Social Robotics*, 2015,7(1): 35-49.
- [21] MINSKY M. A framework for representing emotional states[M]. *EMOTIONS*, 2008: 618.
- [22] 于雪, 翟文静, 侯茂鑫. 人工智能时代人机共生的模式及其演化特征探究[J]. *科学与社会*, 2022,12(04): 106-119.