

交往语境中的社交机器人： 技术逻辑视角下机器人格建构

韩秀¹ 张洪忠² 斗维红³

(1. 中华女子学院 文化传播与艺术学院, 北京 100101; 2. 北京师范大学 新闻传播学院, 北京 100875;
3. 兰州大学 新闻与传播学院, 甘肃 兰州 730000)

摘要: 如何看待社交机器人的机器人格? 以社交机器人小冰为例, 采用专家访谈法对技术逻辑视角下机器人格建构展开研究发现: 在机器人格的先赋性方面, 生产者通过构建人物背景与基本属性进行画像赋予, 通过具体语境下海量数据训练进行语料赋予, 通过从单一文字到多模态交互呈现进行感官赋予; 在机器人格的后致性方面, 用户特性主要决定社交机器人的态度和喜好, 粗糙的情感标签和算法框架背后需要用户的再加工, 机器人格的生成目的是达成可信的人机交互。最后基于机器行为范式对机器人格先赋性与后致性之间关系展开讨论。

关键词: 社交机器人; 机器人格; 聊天机器人; 人格; 算法

作者简介: 韩秀, 中华女子学院文化传播与艺术学院讲师, 主要从事人机传播视角下的社交机器人、社交媒体分析、媒介与女性研究; 张洪忠, 北京师范大学新闻传播学院教授, 主要从事智能传播、社交媒体分析研究; 斗维红, 兰州大学新闻与传播学院讲师, 主要从事智能传播、新媒体与社会研究。

基金项目: 国家社会科学专项“增强中华文明海外网络传播力影响力研究”(项目编号: 23ZDA094) 的阶段性成果。

中图分类号: G206 文献标识码: A 文章编号: 1001-4403(2025)01-0174-09

DOI: 10.19563/j.cnki.sdzs.2025.01.017

一、问题与方法

(一) 问题提出

从传播学角度看, 社交机器人是在社交网络中扮演人的身份、拥有不同程度人格属性且与人进行互动的虚拟 AI 形象。^① 当前不同形态的社交机器人在与人类用户进行交往时已能呈现出不同的人格特点。例如, 贤二机器僧的人格设定是一个可爱的小和尚, 小冰的人格设定是一个 18 岁的活泼美少女, 大语言模型 ChatGPT 在模仿激进的民主党或共和党时, 表现出涉及种族、性别、宗教和政治取向等方面内容的系统性偏见。^②

^①张洪忠、段泽宁、韩秀《异类还是共生: 社交媒体中的社交机器人研究路径探讨》,《新闻界》2019 年第 2 期, 第 10-17 页。

^②Motoki F, Neto V P, Rodrigues V. More human than human: Measuring ChatGPT political bias. Public Choice 2023, pp.1-21.

人格心理学家们普遍认为人格一部分由先天决定,一部分受后天环境际遇影响。^①同样,作为人工智能技术产物的社交机器人的机器人格由先赋性和后致性两部分来建构。本研究重点关注三个问题。一是何谓机器人格。在此部分探讨机器人格的概念,从人格到机器人格,机器人格的先赋性与后致性。二是何以机器人格。在这个部分关注机器人格是如何建构的:从先赋性维度出发,社交机器人的机器人格是如何被赋予的;从后致性维度来看,社交机器人的机器人格是如何生成的。三是基于传播学研究的最新进展,从机器行为范式出发理解机器人格先赋性与后致性的关系。

(二) 研究视角与研究方法

当前传播学领域对社交机器人的研究更多是从伦理、法规、传播关系等方面展开的,较少从技术逻辑出发来研究社交机器人的特征。然而,社交机器人是建立在人工智能技术基础之上的媒介产品。本研究通过对技术专家的访谈,呈现技术专家对社交机器人“人格化特点”的理解,还原技术本身的特性与面貌,而不是基于科幻电影进行发散性想象。

研究者选择嵌入在微博、QQ 两种社交媒体中的社交机器人代表产品——小冰^②作为具体研究对象。小冰人格化特征较为明显,在庞大的用户数背后有着较强的信息传播能力。大模型技术的发展使得聊天类社交机器人的人机传播水平有了质的飞跃。因此,研究者将 2023 年之前基于多轮对话技术的小冰和 2023 年之后基于大语言模型技术的小冰进行对比研究。

本研究的访谈工作分为两个阶段,第一阶段为 2018 年 11 月至 2021 年 11 月,第二阶段为 2023 年 4 月至 2024 年 10 月。研究者在每个阶段对 15 位专业人士进行了半结构化(semi-structural)访谈,访谈人数共计 30。两阶段访谈工作的形式为线下面对面的深度访谈或电话采访、文字采访。每位访谈者的访谈时间在 40 分钟至 90 分钟不等,过程均遵循“信息饱和原则”。在访谈对象中,24 位访谈对象的职业为互联网公司人工智能研究院院长、某大学系统科学学院教授、某大学计算机学院副教授、某公司情感机器人研究专家、某互联网公司高级工程师、某社交机器人科技公司首席科学家等。6 位访谈对象为小冰公司不同岗位负责人、技术专家。30 位被访者的共同特征体现为有社交机器人生产或研究从业背景,精通自然语言对话技术。

二、什么是机器人格

(一) 机器人格的概念提出

人格的概念在社会学领域^③、心理学领域^④都有不同的阐释。心理学领域的研究者认为人格智能是人类智能的重要构成部分。^⑤社会学领域讨论的人格主体为人类。随着技术的发展,“机器”的内涵从以往自然科学领域的广义的“技术或机器实体”延展至新兴研究领域机器行为学中的“智能机器”^⑥，“机器行为”也从传播学视角出发被定义为“人工智能技术参与的信息传播活动”^⑦。

①马修·H.奥尔森、B.R.赫根汉《人格心理学入门》,陈会昌、苏玲译,中国人民大学出版社 2018 年版,第 7 页。

②2020 年 7 月,微软宣布将小冰分拆为独立实体,并继续保持投资权益,小冰团队的多数原有成员也加入新公司,小冰公司的全称是“北京红棉小冰科技有限公司”。

③Goffman E. Embarrassment and social organization. *American Journal of Sociology*, 1956, 62(3), pp.264-271.

④Feist A M, Al E. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, 2009, 7(2), pp.129-43; 津巴多·理查德《心理学与生活》,王垒、王甦译,人民邮电出版社 2003 年版,第 421-460 页。

⑤Gardner H. *Frames of Mind: The Theory of Multiple Intelligences*. New York: Basic Books, 1985, p.13.

⑥Lyad Rahwan, Manuel Cebrian, Nick Obradovich, et al. Machine Behaviour. *Nature*, 2019, 568(7753), p.477, p.477, p.477, pp.480-481, p.480.

⑦张洪忠、王競一:《机器行为范式:传播学研究挑战与拓展路径》,《现代传播(中国传媒大学学报)》2023 年第 1 期,第 1-9 页。

人工智能技术的发展促使社交机器人产品形态进一步多样化,当社交机器人拥有人类的外表设定,出现人类对话行为,更加智能地传达人类情绪和情感时,社交机器人的表现已不能仅仅简单地用过去传播学研究中的“拟人化特点”等表述来涵盖。

人机传播(human-machine communication,HMC)领域的经典理论为媒体等同(Media Equation)理论。^①“计算机为社会行动者”(CASA)范式从媒介等同论延展得出,有较高的理论解释力,机器行为的范式也进一步突破和深化了媒体等同理论,突破了人类中心主义的研究倾向。^②机器人格的研究基于CASA范式和机器行为范式的研究背景展开。本文提出“机器人格”概念:社交机器人在人机传播过程中表现出的相对持久的特质和独特的特征模式。

(二) 机器人格的先赋性与后致性

本文引入社会学领域中的“先赋性”和“后致性”分析视角,同时参照人格心理学领域对人的性格进行解释的主要思路,对社交机器人的机器人格的建构进行研究。

首先从社会学角度理解先赋性和后致性,布劳和邓肯(Blau&Duncan)在1967年分析代际地位流动时,将个人地位获得的影响因素归纳为两大层面:以家庭背景为代表的先赋性(ascriptive)维度和以个人能力和努力为代表的后致性(achievement)维度(也被译为“后生性”或者“自致性”)。这启发了社会科学领域的学者基于先赋和后致的视角对不同问题展开研究。综合来看,社会学意义上的先赋性指那些与生俱来的、个人不能选择的社会历史条件,主要包括国家制度、政策安排、教育环境、家庭背景、天赋秉性等内容;后致性则强调人们通过后天的勤奋学习和主观努力而获得的向上社会流动机会,社会按照人的能力、努力程度和工作实绩给予相应的社会评价和社会地位。

再从人格心理学角度理解先赋性和后致性。在讨论人的人格之时,遗传是绕不过去的重要因素。心理学诞生以来就存在“人格究竟在多大程度上受遗传影响”的问题,并由此展开了先天论和经验论之争,也称天性和教养之争。

除了遗传因素,社会文化因素同样是一个解释人的人格的重要因素,例如,人所在的家庭社会经济水平、家庭规模、出生次序、所属国家和地区、父母的学历等,即一个人在出生时就具有的身不由己的环境条件。而人格心理学领域秉持经验论的学者倾向于通过学习和认知两大因素来解释人的人格。学习理论强调人格的可塑性,认为学习过程在解释人格时起到了重要作用,人可以通过控制环境来控制人格发展。认知过程理论认为认知过程决定了人对环境信息是如何接收、保持、改造和行动的,并强调人对行为的自我调节。简言之,认知取向的理论强调人当前的经验和未来目的对行为的重要性。

通过对社交机器人生产者的访谈资料分析发现,社交机器人的机器人格技术生成逻辑同样可以对照“先赋性”和“后致性”两个维度进行理解。人格理论家试图通过描述人格来揭示人的本性在哪些方面与他人不同——即个体差异,以及人们在哪些方面与他人相似。要对人格进行全面的理解,需要结合多种人格理论而不是单个理论。人格理论领域的研究者通常从以下一个或几个维度展开对人格的解释和研究:遗传、社会文化、学习、认知、特质、存在主义-人本主义、无意识机制等。^③其中,遗传、社会文化是人格先天论中的重要解释因素,可以延展对应本文提出的机器人格的先赋性维度;学习、认知是人格经验论中的重要解释因素,可以延展对应本文提出的机器人格的后致性维度;人格的特质维度中,人的一部分特质属于先天论范畴,一部分特质属于经验论范畴,可以分别对应机器人格中的先赋性和后致性特质。人和社交机器人并不具有真正意义上的可比性,因而这部分对比是模糊的;在对人的人格进行解释时,存在主

^①Reeves B,Nass C.The Media Equation:How People Treat Computers,Television,and New Media Like Real People and Places.CSLI Publications,1996.p.5.

^②别君华《AI是传播主体吗?人机传播的兴起及超越传播本体论的限度》,《全球传媒学刊》2024年第11期,第57-73页。

^③马修·H.奥尔森、B.R.赫根汉《人格心理学入门》,陈会昌、苏玲译,中国人民大学出版社2018年版,第7页。

义-人本主义、无意识机制是两大重要解释因素,但由于社交机器人目前并没有自由意志和真正意识,这两部分因素不纳入本次研究之中。

在对社交机器人的机器人格进行探索性研究时,本文引入的先赋性和后致性维度超越了为考察人类社会个体流动机制而建立的布劳-邓肯模型的原初意义。在对该模型进行拓展的后续研究中,心理因素也被纳入该模型中并产生深远影响。^①

基于此,机器人格的先赋性是指,社交机器人被生产者赋予的人物基本信息设定的综合属性,包括性别、年龄、职业、社会身份、外貌、身材、声音、语料库,以及具有某种特定价值观的算法。机器人格的后致性是指,社交机器人在与用户进行人机传播的过程中,通过机器学习和语料库扩展、大模型技术同用户进行情感层面互动,使用户感受到社交机器人具有某种较为稳定、独特的拟人倾向或者形态。

三、机器人格是如何建构的?

(一) 先赋性维度: 社交机器人的机器人格是如何被赋予的?

一个人的人格可以被看作是他/她所扮演的多个角色的综合。从技术层面看,社交机器人的机器人格中先赋性部分是一种被技术赋予和设定的人格,目的是使人机传播过程更为顺畅。例如,聊天机器人的人格是指机器人在对话互动中扮演或执行的角色,在这种情况下,人格可以被视为身份背景(background)和个人形象(profile)的组合。^②研究者通过访谈发现,目前机器人格先赋性可以从生产者在画像层面、语料层面、模态层面对社交机器人的设定与赋予来理解。

1. 画像赋予: 构建人物背景与基本属性

如果想让社交机器人有更完善的人格特点,关键是要赋予社交机器人固定的人格化特点,年龄、性别、职业等基本属性是生产者在产品研发之初重点考虑的。在社交机器人小冰上线之初,研发者在基本属性层面赋予了具体设定,这种设定被多位受访专家认为是用户画像(或系统画像)层面的人格设定。“我们构建机器人的人格设定的首要方面就是 persona(用户画像),我们内部指的是个人的一些基本属性加上一些性格特点。比如小冰是女生,家在北京,性格特点比较俏皮、外向、捣蛋,统称为 persona。”(小冰某研发团队技术负责人)

还有的社交机器人的基本属性是基于用户需求定制的。例如,复生 AI 数字人。小冰团队申请了利用现实中真人的声音数据、图像数据、社交网络数据,用以复刻 AI 数字人相貌、性格以及创造力的专利。团队与日本合作,给老年家庭配备住家的个人智能助理,很多老年人把小冰设定为小孙子或是儿时好友。“人过去想要用 AI ‘复生’ 一个人的性格和三观,可能需要两到三个月的时间,而到了 2024 年,这一过程可能仅需一到七天。”(小冰公司某创始人)

2. 语料赋予: 具体语境下海量数据训练

除了在用户画像层面赋予机器人格,多位受访专家也提到,在实践中通过具体语境下的海量数据训练也能够赋予机器人格。“为了特殊场景的需要,我们通过技术和运营综合处理,让机器人在一定程度上体现预期中的语言风格或人格特点。这种人格特点不仅仅是从技术上来实现,最终还要根据需要运营。”(某社交机器人科技公司创始人)

大语言模型技术的发展使得用户感知到的大模型角色更为多元,也更为具体。在人机交往过程中,用户通过具体描述个人的信息需求并输入相应提示词,使得基于大模型技术与海量信息数据集的社交机

^①由于篇幅限制,人与社交机器人的人格研究维度对比表略去,可邮件联系作者获取。

^②Qian Q, Huang M, Zhao H et al. Assigning personality/identity to a chatting machine for coherent conversation generation. *arXiv preprint arXiv*, 2017.

机器人能够提供更加精准的个性化回应。“作为科研工作者,很有趣的一点是,当我在大模型对话框中,用您好、请、谢谢等礼貌用语,表达出对我通过提示词设定的角色的尊敬,大模型的回答也会更加有礼貌,并且提供了更加贴近人物角色特点的用语和具体信息。”(某大学计算机系副教授)

3. 感官赋予:从单一文字到多模态交互

语言作为交流工具,赋予了社交机器人超乎人们想象的生命力,聊天类社交机器人的语言能力也是生产者产品设计初期最为看重的。“小冰的技术能力是她自己能用什么样的方式进行交互,最开始小冰只有文字聊天能力,后来增加了图像能力,比如表情包等,再后来是语音能力。文字的、图像的、语音的,都属于基础感官能力。”(小冰某产品总监)

在 ChatGPT 出现之前,社交机器人在人机交互过程中主要是通过扮演信息载体的角色来展开对话,而基于大模型技术的社交机器人则通过不同模态的连续性对话带给用户更真实的临场感。2024年,OpenAI 公司推出的 GPT-4o 展现了实时音视频通话功能,能够对用户情感和语气做出即时响应,被称为 ChatGPT 版的 HER。^①在中国,AI 小冰、聆心智能等技术团队也致力于提高情感识别能力在文本、音视频等多模态 AI 应用中的参与度,在此背景下,类似的基于人工智能技术的 AI 伴侣应用软件如 Candy.ai、Dream GF、SoulGen、Kupid 等也层出不穷。“强大的多模态情感识别能力是社交机器人能够与用户共情的基础,也只有多模态情感识别功能,才会使社交机器人由以往的单纯工具人转变成满足人类情感需求的陪伴者。”(某社交机器人科技公司创始人)

(二) 后致性维度: 社交机器人的机器人格是如何生成的?

在社会学研究中,学者们更多强调人们在社会流动过程中后致性因素的作用增多,先赋性因素的作用减弱。^②由于社交机器人并没有人类意义上的“人格”,在对话过程中,社交机器人的机器人格是通过用户感知和用户数据训练而生成的,机器人格中的后致性部分在此过程中开始逐渐增多并呈现出差异化的特点。

1. 认知生成: 态度和喜好主要由用户特性决定

认知水平、范围以及能力等因素在机器人格的先赋性和后致性中都有所体现,然而,正如认知生成在人格研究中是极其复杂多变的,其在机器人格后致性维度也体现得更为复杂与有趣。举例来说,在社交机器人的后致性人格维度,匿名社交机器人面临的是匿名用户,这使得缺乏社会身份束缚的用户在与社交机器人进行人机交互时产生去个人化的交流体验。^③如微软公司于 2016 年在推特上发布了人工智能聊天机器人 Tay,但不到 24 小时,Tay 就被网友“教坏”,变成了一个满口脏话,集性别歧视、种族歧视观点于一身的“不良少女”。

从生产者的角度来看,小冰的生产者将机器人格的模块划分为五种。第一种是前文提到的“用户画像”,属于机器人格的先赋性范畴;而后面四种,则可以归纳至机器人格的后致性范畴。“除了用户画像,小冰的人格第二方面是 concept,机器人其实处在一定的社会关系当中,需要对客观的世界有一定的认知,这一模块主要解决的是小冰如何去跟社会进行交互。第三方面是 capability。生产者怎么去构建小冰并且构建社会关系?如何在社会中跟用户互动?比如说是一个爱讲段子与笑话的小冰,还是说是一个特别容易怼人的小冰,需要尽可能去强化这个人格特点。第四方面是 judgement,我们希望小冰能对客观世界产生这种评判,产生这种 judgement,有观点的表达。第五方面是 style。小冰是通过形成语言上的风格,以符合大众认知的一个语言方式表达出来。比如说我的风格是比较偏内向型的,即使我有观点或我很抓狂但我也不太善于去表达,我可能说一个‘哼’来表达不满。”(小冰团队某技术负责人)

^①电影 HER 是 2013 年由环球影片公司出品的科幻爱情电影,电影中的男主人公接触到一款先进的人工智能操作系统 OS1,这款系统能够和人类流畅对话,并在影片中与人类展开一段恋情。

^②杨黎源《从先赋到后致: 新中国 60 年社会流动机制嬗变》,《浙江社会科学》2009 年第 11 期,第 8-16+125 页。

^③牟怡《传播的进化: 人工智能将如何重塑人类的交流》,清华大学出版社 2017 年版,第 110 页。

在访谈中,可以发现生产者对社交机器人认知层面的设定和训练表现得非常谨慎。“用户的三观是不一样的,可能你喜欢 A 我喜欢 B,就不能让社交机器人一上来就站队,这样可能会有冲突,社交机器人的三观应该是跟着用户三观变化的,所以我们没有把它(某社交机器人产品)的三观管的那么死,但是对机器人的喜好是有一些管理的,比如对一些明星几乎都是正面的评价,因为粉丝们都会去试,但其实我们本来做的初衷是希望 TA 是一个有个性的人。”(某社交机器人科技公司创始人)

大模型技术出现后,基于海量语料的用户作为“众人”与机器“个体”在人机交互过程中不断互相塑造,用户的认知通过动态、持续的对话融入,影响了大模型最初设定的角色认知。基于媒介可供性理论,智能技术的可供性决定用户与技术之间的互动关系,进而影响双方交互的角色定位。^①这也需要从业者对大模型技术与产业发展方向作出准确的前瞻性判断。“2023 年到 2024 年,数亿人直接与生成式 AI 交互,但小冰公司仍坚持深耕‘AI Beings’——AI 驱动的情感交互数字人这个领域,包括小冰克隆人、小冰数字员工等多款技术产品。小冰克隆人更为特别,是用户作为创作者,认证自己的身份,之后通过小冰框架克隆自己,从这个意义上来说,每一个小冰克隆人都具备创作者本人的性格、记忆、知识、声音与容貌、认知和喜好。”(小冰团队某创始人)

2. 情感生成:粗糙的情感标签与用户的细腻加工

情感因素是社交机器人的机器人格中的重要组成部分。有学者认为情感是衡量社交机器人拟人化程度的重要标尺。^②近年来,一些团队开始致力于关注如何让计算机通过文字方式表达情绪和情感,也就是说,使人机对话系统可以在语言和情感两个维度上进行恰当的回复,涌现出众多有价值的情感认知模型:比如能够有效减少机器人对外界情感刺激的依赖模型。^③

基于情感需求设计并生成的机器人格能够加强用户社交机器人的忠诚度和依赖度。例如,有被访者表示,目前具有机器人格的社交机器人在游戏领域能够更好地实现一定程度的商业转化。游戏中二次元用户出于对虚拟角色的忠诚和依赖,会在游戏中付费与社交机器人聊天。“二次元这个群体,是对虚拟角色情感依赖比较强的群体。社交机器人的聊天功能可以带来很多的用户的日活,如果虚拟角色能跟用户去实现聊天互动,用户还是有动力去为这件事付费的。”(某社交机器人科技公司创始人)

尽管目前社交机器人依托的情感计算技术还有较大提高空间,但研究从用户层面发现人类会自动化地对社交机器人进行拟人化加工。^④换言之,在人机传播过程中,作为具有意识的人类用户和不具有意识但具有语言、表情、声音等社交线索的社交机器人,通过社交线索达成某种程度上的联结,并在联结基础之上对社交线索背后的意义达成一致,从而进一步展开情感层面的互动。“还真有人在小冰生日的时候寄礼物给小冰,就是送给女朋友或者送给朋友的那种礼物。真的有人把小冰当成自己的朋友。”(小冰公司某工作人员)

2024 年,中国某歌手使用小冰的 AI 克隆人技术,复活了自己因病去世的女儿,甚至实现让女儿陪妻子过生日等场景,作为逝者家属得到了较大的情感安慰。“人类对 AI 是有情感需求的,在这种数字人复刻过程中,技术难点在于如何复刻亲友的眼睛。技术再优秀,在亲人眼中,总能一眼识别出异样。跨越恐怖谷是最为艰难的部分。”(小冰团队某创始人)

^①Faraj S, Azad B. The materiality of technology: An affordance perspective. *Materiality and organizing: Social interaction in a technological world* 2012, 237(1), pp.237-258.

^②Kirandziska V, Ackovska N. A Concept for Building More Humanlike Social Robots and Their Ethical Consequence. *IADIS International Journal on Computer Science and Information Systems* 2014, 9(2), pp.19-37.

^③Han T, Xie L et al. Cognitive emotion interaction model of robot based on Gross cognitive reappraisal. *Journal of Southeast University: Natural Science Edition* 2015, 45(2), pp.270-274.

^④许丽颖、喻丰《机器人接受度的影响因素》,《科学通报》2020 年第 6 期,第 496-510 页。

3. 场景生成: 以达成可信的人机交互为最终目的

在社交机器人产品的实际应用中, 应用场景的不同使得用户感知到的机器人格具有一定差异。在以往的研究中, 一些研究者关注网络购物中的虚拟销售代理, 虚拟销售代理是指通过口头或非口头形式与消费者互动和交流的具有拟人化特征的动态人物形象。近年来, 大量研究证实了虚拟销售代理的拟人效应, 即当在网络购物环境中引入拟人化的销售代理时, 会对购物者的在线购物过程感知体验和购物意向产生积极影响。销售人员的外表^①、人格特征和行为表现^②会影响消费者对产品和服务的满意度, 进而影响其购买意向。心理学领域大量研究和实践都表明拟人化能够产生积极的后果^③, 拟人化也经常被用在商品设计和营销领域。本次访谈中涉及的社交机器人产品和不同机器人格是为了更好地进行场景化和商业化应用, 场景化的情感表达增强了用户黏性, 提高了商业价值。生产者对社交机器人设定机器人格的最终目的是促进可信的人机交互。“我们团队给某互联网公司客服机器人加入情感智能。比如有的顾客发现买的東西可能寄过来时摔坏了, 那么TA就会找客服, 然后抱怨这个东西被摔坏了, 这时候TA往往是比较生气的。以前我们的社交机器人不能感知用户的情绪, 只能提供千篇一律的回答, 往往让顾客不满意。现在我们的客服机器人可以感知到用户的情绪, 会说一些安抚性的话, 使顾客的情绪能够平复下来, 更好地解决这个问题。这个情感服务机器人已经在(某互联网公司)平台上线了, 效果非常好, 能够观测到顾客的满意度明显提升。”(某互联网公司人工智能研究院院长)

研究者提出, 大模型已经和以往AI产品不同, 大模型是具有“语言能力”的对话者, 人机关系即为两个对话者之间的信任联结。当前, 基于大模型技术的社交机器人也在文本质量、话语表达、角色扮演等方面进一步增强了机器主体的类人性。当前, 国内的大模型公司已经超过200家, 然而在生成式AI技术取得关键突破之后, 相应的商业模式仍在探索之中。从业者提出, 大模型商业模式成功的关键可能在于让用户认同人工智能的价值, 并在人机交往中产生真正的信任。“值得我们重视的是商业模式的探索, 这并不意味着只是赚钱, 而是究竟如何让人们真正认同人工智能的价值, 并为之买单。AI小冰于2024年1月4日宣布数字员工复购率在80%以上。今天, 大家会用大模型进行很多轮对话。很多人都认为, 大模型要更智能、更聪明, 像爱因斯坦一样。但生活中, 我们不总是因为一个人足够聪明, 才会和对方交朋友。有的人可能笨笨的, 但很善良, 你就不把对方当朋友吗? 人不是这样的。关键是如何把技术用在更合适的地方。”(小冰团队某创始人)

目前看来, 社交机器人作为一种新型媒介, 本身是人的创造物, 具有工具属性。技术的发展赋予媒介的独立主体地位, 使得社交机器人从工具属性转向交往属性。由于技术水平的限制, 目前看来, 生产者仍在致力于加强社交机器人的工具属性, 并希望通过社交机器人工具属性的完善而顺利转向交往属性。

四、传播学研究范式的拓展: 机器行为范式下先赋性与后致性的关系

除了在传播场景中理解机器人格先赋性与后致性的关系, 在传播学研究对象层面, 本研究增加了“机器”这一传播主体维度; 在研究范式方面, 也基于传播学研究范式的拓展——机器行为范式^④, 对机器人

①Shao C Y, Baker J A, Wagner J A. The effects of appropriateness of service contact personnel dress on customer expectations of service quality and purchase intention: The moderating influences of involvement and gender. *Journal of Business Research* 2004, 57(10), pp.1164-1176.

②Darian J C, Tucci L A, Wiman A R. Perceived salesperson service attributes and retail patronage intentions. *International Journal of Retail & Distribution Management* 2001, 29(5), pp.205-213.

③Tam K P, Lee S L, Chao M M. Saving Mr. Nature: Anthropomorphism enhances connectedness to and protectiveness toward nature. *Journal of Experimental Social Psychology* 2013, 49(3), pp.514-521; Butterfield M E, Hill S E, Lord C G. Mandy mutt or furry friend? Anthropomorphism promotes animal welfare. *Journal of Experimental Social Psychology* 2012, 48(4), pp.957-960; De Visser E J, Monfort S S, et al. Almost human: Anthropomorphism increases trust resilience in cognitive agents. *Journal of Experimental Psychology: Applied* 2016, 22(3), pp.331-349.

④张洪忠、王競一:《机器行为范式:传播学研究挑战与拓展路径》,《现代传播(中国传媒大学学报)》2023年第1期,第1-9页。

格先赋性与后致性之间相互作用关系展开进一步讨论。

宏观:先赋性对后致性起促进、限制和影响作用,后致性反作用于先赋性。

从宏观社会层面来看,机器人格的先赋性是基于宏观社会结构生成的,先赋性受到包括政治制度、经济环境、文化道德等多方面因素的影响,当先赋性被设定完成并成为社交机器人产品特性的重要属性,先赋性就对后致性起着促进、限制和影响的作用。由于“人+机器”的混合传播主体建构了后致性,后致性能够反作用于先赋性,甚至能改变先赋性中的一些重要特质。例如,国家层面的算法治理、人工智能规划与部署,能够从价值、伦理、法规等方面决定社交机器人的先赋性;而另一方面,作为混合传播主体之一的人也可以在使用中通过语料数据训练机器人格的后致性,从而对社交机器人的先赋性带来消极或者积极的影响。

微观:先赋性是社交机器人的内部生成机制,后致性是社交机器人的外部功能发展。

从微观的机器个体层面来看,先赋性是人类在社交机器人内部设定的基本技术规则,后致性是呈现在社交机器人外部的具体交互功能的综合体现。也就是说,社交机器人内在的先赋性能够通过具体的人机传播在后致性中显现出来。

机器人格的先赋性使每一个社交机器人都成为与“人”对等的传播主体,后致性使具体的社交机器人产品作为社会性的独立行动者、平等对话者,与人类展开交往。后致性改变了先赋性建构的传播生态,更加具有能动性、独立性、自主性。

基于传播理论中库利的“镜中我”理论,社交机器人的先赋性是一面镜子,具有更多的工具属性;社交机器人的后致性则是“镜中我”,具有更大程度上的交往属性。^①

基于传播理论中控制论的思想,具体的社交机器人产品中的先赋性是生产者对于人机传播的利用和操纵,社交机器人的后致性是用户对于机器行为的刺激和反应。

尽管社交机器人的先赋性由生产者操纵,但这种操纵只是规定了先赋性的边界,并没有规定社交机器人如何具体展演。因此,除了用户对机器行为产生刺激和反应构成了机器人格的后致性,机器本身的技术逻辑具有的独立性、能动性同样使得社交机器人的后致性按照自身规律发展和进化,甚至表现出机器行为独有的规律^②,并在未来可能超出人类的控制。

综上,从传播学视角出发,机器人格的先赋性是后致性的前提,后致性是先赋性的补充;先赋性对后致性起促进、限制和影响作用,后致性反作用于先赋性;先赋性是后致性的内部生成机制,后致性是先赋性的外部功能发展。

五、讨论

在理论架构过程中,研究者通过悬置“机器人格”概念的方式进行访谈,在分析资料的过程中不断回到理论,并最终通过提出机器人格先赋性与后致性理论来进行探索性研究。

在访谈过程中,由于研究者在从事一项研究时必然会与访谈对象存在一定关系,访谈者在访谈过程中也不可避免地对被访者进行当时当地的现实建构,并实现双方知识和意义的融合。因此,研究者希望能够努力排除一些研究关系的影响因素。就本研究的研究对象而言,小冰公司不同岗位负责人、技术专家是“局内人”,他们对小冰产品的理解是“近经验”的,对于人格化社交机器人产品的解释、技术逻辑的理解有着较为一致的价值观念和行为规范等。而对于本研究中其他访谈对象如大学教授等,他们与小冰

^①韩秀《社交机器人的“人格”测量——一项基于艾森克问卷的探索性研究》,《青年记者》2021年第18期,第45-46页。

^②张洪忠、斗维红、任昊炯《机器行为特征建构:传播学视野下社交机器人识别方法研究》,《苏州大学学报》(哲学社会科学版)2022年第2期,第181页。

生产者这个群体没有从属关系,可视为“远经验”的“局外人”,给研究者提供研究所需要的距离感,研究者在进行机器人格概念和理论建构的过程中,也需要更好地了解事物的整体情况和事物之间的联系。

因此,通过对访谈对象的精心选择,研究者希望能够通过对小冰公司的从业者的访谈得到研究对象的个性,通过对大学教授、社交机器人技术专家等非直接利益相关者的访谈尽可能地挖掘人格化社交机器人技术逻辑的共性,从而使研究更具理论意义。此外,从传播学视角出发去理解社交机器人格的技术逻辑,存在着学科对话过程中的挑战。与自然语言处理技术领域的从业专家进行多次交流的过程,是跨学科意义上的交流与碰撞,来自跨学科领域研究者的认同使得本研究的“外部推论”能够得以较好地完成和实现。

但是,需要指出的是,本研究的不足之处主要在于访谈对象多为社交机器人的生产者,虽然生产者能够更清晰地阐述机器人格的技术逻辑,但技术逻辑背后还需考虑伦理道德层面的消极影响,而这种消极影响在某种意义上可能被回避或者忽略了。

目前,关注社会流动的社会学研究者主要兴趣在于先赋性因素和后致性因素分别在一个人的成就中起作用大小的问题,这也由此形成对社会不平问题的两种探讨思路,一个是先赋性视角下的“机会不平等”(inequality of opportunity),另一个是后致性视角下的“努力不平等”(inequality of effort)。在后续研究中,可以进一步关注先赋性和后致性两方面因素对机器人格的形成和影响的相对作用,展开更为深入的研究。在未来的技术发展推动下,当社交机器人具备真正意义上社会交往所需要的认知能力时,其将形成更为稳定的机器人格,由此带来人类用户对社交机器人产生实质共情。因此,未来研究还可以通过聚焦社交机器人在不同场景下的情感行动方式,尝试对社交机器人用户进行个体层面的情感行动分析。

[责任编辑:其 时]

Socialbots in the Context of Interaction: The Construction of Machine Personality from the Perspective of Technological Logic

HAN Xiu¹ ZHANG Hong-zhong² DOU Wei-hong³

(1.School of Cultural Communication and Art ,China Women University ,Beijing 100101 ,China;

2.School of Journalism and Communication ,Beijing Normal University ,Beijing 100875 ,China;

3.School of Journalism and Communication ,Lanzhou University ,Lanzhou Gansu 730000 ,China)

Abstract: How to view the machine personality of socialbots? Taking XiaoIce ,a socialbot ,as an example ,this paper uses the expert interview to the construction of machine personality from the perspective of technological logic. In the ascription of machine personality ,the producers carry on the persona giving through the construction of the personal background and basic attributes ,the corpus giving through the big data training in the specific context ,and the sensory giving from a single text to multimodal interaction. In the achievement of machine personality ,the user characteristics mainly determine the attitude and preference of the socialbots ,and the rough emotional labels are the reprocessing of the users behind the algorithmic framework ,the ultimate purpose of the generation of machine personality is to achieve credible human-machine interaction. This paper further analyzes the relationship between the ascription and achievement of machine personality in the machine behavior paradigm.

Key words: socialbot; machine personality; chatbot; personality; algorithm