

悲伤机器人的隐私困境与伦理出路

The Privacy Dilemma and Ethical Solution of Griefbots

张旭 /ZHANG Xu¹ 王金柱 /WANG Jinzhu²

(1. 东南大学人文学院, 江苏南京, 211189; 2. 安徽大学哲学学院, 安徽合肥, 230039)

(1. School of Humanities, Southeast University, Nanjing, Jiangsu, 211189;

2. School of Philosophy, Anhui University, Hefei, Anhui, 230039)

摘要:“悲伤机器人”作为人工智能应用于哀悼场景的发明之一,通过模拟共情使哀悼者能够与逝者的数字形象进行互动,从而使其完成悲伤情感的疗愈。然而,这种技术介入使隐私问题的边界跨越生死,对逝者和生者的隐私提出挑战,既威胁逝者的记忆完整性与人格尊严,又破坏哀悼者的情感自主与悲伤真实性。困境深化引发主体的存在危机:对逝者的数据降格和叙事分解使其面临人格消散的“二次死亡”风险,对哀悼者的数据化监控和算法情绪引导使其失去通过“向死而生”确证自己是有意存在的机会。面向危机,以对技术工具理性批判为基,唤醒隐私的自我价值,将“隐私敏感性”作为技术设计的原则,使技术回归守护人之本性的使命。

关键词: 悲伤机器人 人工智能 隐私困境 数字哀悼 隐私敏感性

Abstract: As one of the AI innovations applied to the context of mourning, “griefbots” employ simulated empathy to enable mourners to interact with digital avatars of the deceased, thereby facilitating their emotional healing. However, this technological intervention presents a challenge to the privacy of the deceased and the living, threatening the integrity of the deceased’s memory and their personal dignity, while undermining the mourner’s emotional autonomy and the authenticity of their grief. The deepening of this predicament could trigger the existential crisis of the subject: the data-driven reduction and narrative fragmentation of the deceased expose them to the risk of a “second death”—the dissolution of their personhood—while the data surveillance and algorithmic emotional manipulation directed at the mourner deprive them of the opportunity to affirm their own meaningful existence through “being-towards-death”. In the face of this crisis, the application of technologies like griefbots must be based on the critical examination of technological instrumental rationality, grounded in an awakening of the intrinsic value of privacy, and rooted in “privacy sensitivity” as a core principle of design. Only in this way could technology be re-anchored to its fundamental mission of safeguarding human nature.

Key Words: Griefbots; Artificial Intelligence; Privacy problem; Digital mourning; Privacy sensitivity

中图分类号: B82; TP242.3 DOI: 10.15994/j.1000-0763.2026.11.010 CSTR: 32281.14.jdn.2026.11.010

基金项目: 国家社会科学基金一般项目“数字时代隐私的哲学研究”(项目编号: 24BZX081)。

收稿日期: 2026年1月15日

作者简介: 张旭(1996-)女, 内蒙古乌海人, 东南大学人文学院博士研究生, 研究方向为隐私哲学、技术哲学。Email:

230219063@seu.edu.cn

王金柱(1975-)男, 安徽淮南人, 安徽大学哲学学院教授, 研究方向为科技哲学、设计哲学、隐私哲学。Email:

anhwjz@163.com

一、“悲伤机器人”的兴起与内涵

亲人离世产生的“悲伤情绪是痛苦的,甚至是创伤性的,严重的结果将导致抑郁症,甚至出现自杀行为”。^[1]为帮助哀悼者疗愈悲伤情绪,人类曾尝试各种办法来建立“与逝者的想象对话以使他们保持在场”。^[2]从传统社会利用墓碑、铭文、悼词和棺木等寄托对逝者的追思与敬意,到之后摄影、电话和广播等技术的兴起将抽象的情感追忆具象化为可视的影像与可触摸的相片,再到数字时代,技术设计虚拟墓园、数字纪念空间等使逝者的存在得以在云端持续延展,哀悼者亦可突破地理与时间的限制,在线开展缅怀与互动。可见,技术在哀悼活动中扮演着重要角色,不仅承载记忆、调节情绪,更在哀悼者与逝者之间建构起情感联结。但是,“悲伤并非对死亡的一种自动反应,而是一种复杂的情感回应”,^[3]这些与逝者的单向沟通无法满足哀悼者“被理解”和“被回应”的更深层次的情感需求。于是,一种以模拟逝者的语言特征和行为模式来与哀悼者进行互动交流的人工智能应用形式最近出现。这是一种基于逝者的数字足迹所创造的聊天机器人,它被赋予多个名称,包括“幽灵机器人”(ghostbots)、“死亡机器人”(deadbots)和“悲伤机器人”(griefbots)。各类名称虽指向相同的人工智能技术,但概念的差异反映的是在实际应用中对相关技术核心问题思考侧重的不同。这里,我们选取“悲伤机器人”作为主要表述,原因在于,“悲伤”本质上是一种存在论的过程,“尽管悲伤是一种令人痛苦的情感体验,但它对人类自我认知具有积极意义,它为我们提供了一个以面对死亡的方式认识自我的机会,当我们所依恋的人离世时,悲伤不仅促使我们直面死亡这一根本性存在问题,更引发对自我同一性危机的担忧,进而促使我们在反思中实现对自我的重新理解和身份的重构”。^[4]因此,“悲伤机器人”的命名选择,既不同于“死亡”所传递的生物学意义上那种对人的瞬时的、不可逆的状态描述,亦不同于“幽灵”

所表达的对人工智能复活技术所呈现状态的描述,而是指向哀悼者和逝者共同在存在意义上的建构过程。作为情感人工智能具身化的重要实践,“悲伤机器人”进入哀悼活动领域,旨在通过塑造一个具备自主学习能力、高度拟人化的数字逝者形象,与哀悼者开展持续性的互动,从而使哀悼者在互动过程中获得更真实的情感体验,为个体提供心理疏导和情绪安抚,助力完成悲伤的疗愈过程。现有的技术原型在功能设定方面,充分体现了这一命名选择的合理性。无论是基于大语言模型建立的“聊天机器人”,例如金尼娅·库伊达(Eugenia Kuyda)精心设计的Replika或由OpenAI推出的“Project December”,还是用于哀悼的数字永生项目,例如HereAfter、Enternime,或是用于悲伤干预治疗的陪护机器人,例如PARO,其核心目标并非“处理死亡”或“制造幽灵”,而是通过介入人类的情感世界,影响逝者与哀悼者之间关系的构建以及个体对自我存在意义的理解。

从技术特征上看,“悲伤机器人”的正常运作依赖对逝者和哀悼者情感数据的持续、深度和多模态的采集。然而,对此类内心世界高敏感数据的广泛获取和处理,已成为引发潜在隐私侵害的技术性根源。在各种类型的机器人中,即便是同样具备情感感知能力的社交陪伴型机器人,其所涉及或可能引发的隐私问题始终局限于“生者”范畴。然而,“悲伤机器人”第一次将隐私侵犯的触角明确伸向生的另一头——逝者。借助人工智能构建的与逝者相似的机器人,意味着技术的信息收集将不再局限于哀悼者的情感数据,而是包括了逝者生前的数字遗骸——任何可以转化为数据的逝者历史信息(如文字记录、语音影像、社交痕迹、行为日志等)都能成为“悲伤机器人”模型训练的数据原料。因此,虽然“悲伤机器人”的出现为哀悼者和逝者之间的沟通和联系提供了新的方式,却也引发更为深刻的隐私问题:逝者是否享有隐私?哀悼者是否有权代表逝者行使隐私权利?对逝者隐私的侵犯是否构成对其人格尊严的终极侵犯?哀悼者与机器人互动的

过程中是否会暴露更多的心理隐私?针对“悲伤机器人”在研发与应用过程中可能引发的隐私问题,敏锐的理论家们已经意识到,该技术与哀悼者之间的双向互动,不仅可能对个体的心理隐私构成潜在风险,而且对哀悼者的自我身份认同和社会关系建立也带来更为深远的影响。有学者表示,“悲伤机器人”通过“邀请哀悼者与其进行互动,可能干扰个体建立新的情感联结。在本应通过社会活动与他人建立联系的阶段,诱使其选择向虚拟逝者寻求心理支持,从而阻碍个体在现实世界中继续前进”。^[5]

目前,关于这类机器人应用所引发的隐私问题虽然已引起各界广泛关注,但是,由于该技术的不同命名使得对这一技术现象的讨论分散于不同问题领域,致使有关隐私的讨论也大多混杂于其他问题之下,或被一笔带过,或浅尝辄止,对隐私困境的审视尚未触及问题之根本,故难以形成有效的应对方案。理论的滞后不仅使这类机器人的开发计划无法轻易推进,亦令公众在面对此类新兴技术时隐忧丛生。鉴于此,选择“悲伤机器人”而非其他命名,意味着对其分析将置于存在论情感分析框架下,考察“悲伤机器人”在应用中如何介入人类情感过程,如何通过侵犯隐私削弱个体的自主性和自我意义的生成能力,并最终引发存在论层面的主体危机。

二、“悲伤机器人”引发隐私问题的生成逻辑

“悲伤机器人”所承载的技术理念,为哀悼者和逝者之间的情感沟通提供了一种更为直接的互动方式,但其依赖数字足迹的对话模式也引发新的隐私问题。

首先,“悲伤机器人”的研发和应用需要大规模采集逝者的个人数据(包括年龄、性别、职业在内的基本信息、网络上遗留的照片、语音和文字记录,以及他人对逝者所存留的记忆碎片)。然而,当这些数字足迹被整合至“悲伤机器人”的数据库时,其属性将经历本质性转变——它们不再是逝者身份的象征,而是算

法模型的训练参数和资源。此种转化不仅加剧了对逝者隐私的侵犯,亦对其记忆的完整性、人格的独立性以及尊严构成威胁。

记忆作为个体心理隐私的关键组成部分,在“悲伤机器人”构建已故逝者的数字形象过程中,可能遭遇隐私侵犯与结构破坏的风险。由于该类技术旨在生成高度拟人化的虚拟形象,其数据采集的范围将超越传统的年龄、性别以及外貌等表层信息,而是涉及那些承载逝者心理活动与情感状态的内在数据。机器人进入逝者的精神世界,对其自主掌控和决定何人能够进入其最私密的心理领域的的能力构成威胁。然而,“悲伤机器人”的情感计算和语言学习能力不仅会分析已有数据,还会从看似无关的数据中推断出逝者生前可能尚未同意公开的秘密,并将“不同对话中提取的数据不成比例地混合在一起”^[6]形成新内容。如此行为不仅未能尊重逝者的记忆完整性,而且由于损害隐私生成和理解所依赖的情境基础,使“逝者一生在不同数字论坛的对话中所展示的多重公共自我面临暴露的风险,而其中可能有家庭成员不知道的方面,且逝者很可能不愿意透露这些方面”。^[6]不仅如此,悲伤机器人“还会自主发起对话,要求哀悼者给出更多信息”,^[7]这种诱导行为不仅是对逝者个人隐私的侵犯,更是对哀悼者和逝者之间共享的情感记忆完整性的损害。

通过对逝者的数据化处理,“悲伤机器人”将逝者简化为可处理的数据集合,这种将逝者的数字遗骸视为一种工具或资源的态度与行为“是对个人死后生活隐私的最大威胁”。^[8]我们无法确定逝者是否同意在死后以这种方式被数字化“复活”,因为他本人在离世后便失去对自己信息支配和管理的权利。不仅如此,“悲伤机器人”的商业运作模式倾向于将逝者的数字遗骸视为一种可以满足情感需求的数字消费品,具有一定商业价值。这也是为何“虽然逝后隐私已经是一种公认的现象,但大多只在分析数字资产和死亡的一般法律方面或在数据保护的背景下简要地讨论逝后隐私”。^[6]由此导致的结果是企业在制定机器人的相关法律合同

时,往往表现出随意性和模糊性,“Mozilla基金会的一项研究发现,现有的大部分人工智能伴侣应用程序都没有通过隐私认证”。^[9]因此,即便逝者在生前可能签署过相关的隐私知情同意协议,但仍然难以保证逝者数据在其死后的使用能够符合最初同意的范围。在此情境下,“悲伤机器人”通过对逝者的数据降格和商品化,不仅可能损害逝者应当享有的隐私权益,还可能对其死后的人格尊严造成贬损,面临沦为可被随意调取、分析与使用的透明化客体的风险。

其次,“悲伤机器人”的设计初衷是为哀悼者提供一种与逝者互动的新方式,以此缓解并帮助其走出悲伤情绪。然而,这种双向情感交流方式在使哀悼者获得情感慰藉的同时,也会在不经意间使其暴露更多关于自身身体和心理状况的深层次的敏感信息,这一过程不仅会增加隐私泄露的风险,还可能导致哀悼者面临情感自主性丧失和人际关系失真的问题。

“悲伤机器人”通过模仿逝者的说话语气、表达风格甚至特定记忆点,建立了一个使哀悼者会产生“似曾相识”的在场感和熟悉感的场景,以此唤起他对逝者深刻的情感回忆。这里,机器人看似承担起“安全的倾诉容器”的角色,但实际是使哀悼者进入一种全息监控的环境,所有的身体活动(提及特定事件时的面部表情、肢体动作和心率变化等)和心理活动(内心感受、自我回忆等)都可能被技术同步记录作为悲伤生理维度的量化指标。数据生成后,处于情感脆弱状态的哀悼者不仅难以做出理性判断,而且极易做出非理性授权,将个人数据的管理与使用随意交由技术系统处理。这种状况不仅会增加隐私泄露的可能性,更会导致个体的情感自主性遭遇隐性丧失的风险,因为它不是在强迫中发生,而是在看似“被理解”的过程中悄然发生,其后果令人堪忧,“如果不作伦理方面的考虑,就很有可能使悲伤商品化,将深刻的个人经历变成逐利的冒险”。^[10]

“悲伤机器人”对哀悼者提供的“无评判的安全倾听”使他们更容易卸下心理防备,分享那些曾经并未与旁人言说的内心感受。在这种

更沉浸、更真实、更即时的情感互动中,哀悼者极有可能“产生一种持续性的期望,希望从他们所爱的人的机器人那里得到回应,但这可能会培养对这项技术的某种情感依赖”,^[11]结果“将很容易受到机器人开发公司的剥削和操纵”。^[12]并且,这种操纵往往以“关怀”的面目出现,极具隐秘性和渗透性,哀悼者以为自己在被理解、被陪伴,实则是在无意识中接受技术的操纵和规训。加之“悲伤机器人”的“功能的视觉超现实主义倾向于通过强化用户的情感依赖来创造更强的连续性错觉”,^[13]这种连续性不仅制造了虚假的关系预期,使哀悼者陷入与机器人难以分割的联系之中,而且弱化了哀悼者对现实与虚拟世界的区分,使哀悼者极易“陷入不真实、自我欺骗甚至妄想”^[14]境地。当哀悼和悲伤的自然进程被破坏,哀悼者也将失去在真实的社会交往中经历哀悼、学习如何面对死亡并处理悲伤情绪的机会。

三、“悲伤机器人”引发的人类生存危机

“悲伤机器人”引发的隐私问题,不仅是对逝者和哀悼者隐私权利的侵害,更是通过消解其主体性,动摇人类存在的本体论根基。隐私作为守护人之为本体论条件,通过划定主体内在性的神圣疆域,使个体得以实现对自我主体性的确认,并在与他者的张力中建构自我意义。具体而言,一方面,隐私为自我提供了一个免受外部世界持续冲击的缓冲地带,在这个领域内得以反思其存在的意义,直面自身的畏与死,并能通过自身经验感知进行价值判断、情感体验和道德抉择;另一方面,在隐私守护下具有自我内核的主体,能够以稳定的身份去面对他者,同时他者也能避免在此过程中被自我完全同化,从而确保个体间的相遇是主体间性的碰撞,而非单方面的淹没或吞噬。然而,“悲伤机器人”穿透自我边界,对个体最私密的情感领域——逝者的记忆、生前未公开的秘密、哀悼者对逝者的哀思、内心回忆以及生理反应等极度私密的数据的无差别收集,使得自我内心世界被迫打开,致使人们的内在体验被彻底

“外部化”和“客体化”，成为被外部系统审视和分析的数据。此时，个体所遭遇的挑战已超越单纯的隐私泄露风险，更涉及其内在自我发展与存在空间被逐步侵蚀的深层问题。这种侵犯既是对个体人格完整性的破坏，更是对其自主性的消解。

对逝者而言，“悲伤机器人”构建的“数字逝者”将其降低为一种数字化资源，从而削弱其作为生命主体的完整性与独特性。本质上，“逝者作为独特的他者，呈现着绝对的超越性和不可占有性”，^[15]是不可替代的存在。然而，“悲伤机器人”将逝者的生活经历和生物特征分解为孤立的数据点，消解其作为“不可知整体”的神圣性，使不可言说的“逝者”降格为可操控的数据客体。这种对逝者隐私侵犯所导致的数据化解码，实则用算法消解其死亡的超越性，取消逝者作为独特生命存在的主体性，导致逝者无法以“绝对他者”的身份在世存在。结果是使逝者的数字遗骸没有被视为“构成一个人人格的东西”^[8]而获得尊重。因此，逝者主体性的侵蚀，核心是其自我表达权利的剥夺。起初，“悲伤机器人”对逝者数据的占有和监管，使逝者失去对自我“编辑”的控制力，进而影响其在不同人际关系中自我形象的建构与维护。随后，在处理不同情境下的哀悼情绪时，“悲伤机器人”会对逝者的数据进行交叉使用，从而削弱并消解逝者生前在多重语境下所呈现的复杂而丰富的自我形象。这种对逝者数据的混合和强行并置，本质上是将逝者的多重自我简化为可编辑的数据模型，导致逝者作为独特他者叙述自我的功能退化。结果是，“悲伤机器人”试图通过采集逝者的隐私数据塑造一个“数字形象”以实现其在现实世界中的继续存在，但在这个过程中，通过数据降格与叙事分解的双重操作，使逝者在失去隐私的同时遭遇主体性瓦解的危机，面临存在论层面上人格消散的“二次死亡”的风险。

对哀悼者而言，“悲伤机器人”的加入改变传统的哀悼方式，哀悼过程的数据化监控和悲伤情绪的算法引导，破坏其作为本真性实践所需要的私密空间，继而使哀悼者失去按自然

节奏和方式经历、学习和处理“失去与死亡”意义的的能力，逐渐丧失定义其自身存在的可能性。哀悼者作为“此在”，在面对死亡时，能够重新审视自身与逝者存在的意义，从中获得对生命与死亡的新认识，进而深入思考“生活应如何延续以及我们应当成为怎样的人”（[4]，p.103）这一重要问题。作为一种本真性的实践形式，哀悼活动使哀悼者可以正视死亡的不可逾越性，清晰地认识到生与死之间的界限，从而在日常沉沦中抽离，领会自身作为“向死而生”的有限存在者。传统哀悼的私密性为这种本真性实践提供条件，通过空间隔离、亲属关系筛选和仪式加密等方式，使哀悼者能够在相对私密的空间内释放自己的悲伤情绪，在目睹逝者肉身消逝的过程中意识到生与死的界限，在独处中体验“畏”并领会自身的“向死存在”，最后通过记忆叙事的方式整合生死关系，以此重建对生命意义的理解。然而，“悲伤机器人”通过塑造出一个“数字逝者”，营造了一种逝者永久在场的感觉，从而破坏“死亡作为不再能在此”^[16]的时间断裂点，使哀悼者难以清晰地区分生与死、真实与虚拟的界限，更无法在“空无”中深刻体悟死亡与存在的本质。“悲伤机器人”对哀悼者情感数据的收集和操控，更是使其从情感主体降格为情感数据的提供者，从意义的创造者降格为算法输出的接受者，导致其真实记忆被算法生成的记忆取代，从而逐渐丧失对自我记忆整合为连贯生活叙事的自主能力。这一过程从根本上削弱的是哀悼者构建内在自我的可能性，使其从“凝视死亡的主体”沦为“被数据凝视的客体”，失去通过直面死亡进行自我反思和确认自身存在的机会，陷入“存在的彻底遗忘”的危险境地。

“悲伤机器人”的隐私困境所引发的双重危机，对逝者而言，是使其失去作为“绝对他者”的不可还原性；对哀悼者而言，是使其失去通过“向死而生”确认自己作为有意义的“本真存在”的最后机会。对双重主体存在危机的揭示将我们引入一个更为根本性问题的思考之中：当“悲伤机器人”的发展以如此深刻的方式介入哀悼，我们该如何保有作为有意义的存

在者的主体地位?又该如何保有对逝者作为不可侵犯的“他者”的尊重?这些问题没有现成的答案,但对危机的思考已为后续的伦理路径探寻指出方向。

四、规避“悲伤机器人”隐私问题的伦理设计路径

“悲伤机器人”引发的隐私困境,实为一场在技术座架下针对人类存在根基的“去主体化”工程。于逝者,它通过数据占有将其压缩为可操控的数字客体,消解其“不可言说”的伦理重量;于哀悼者,算法标准化殖民情感自主,剥夺其“向死而生”的本真觉醒。当技术主导悲伤情绪,死亡被技术驯服,隐私侵犯最终演变为对主体性的剥夺。应对危机,需从“悲伤机器人”发展的工具理性霸权批判出发,通过对以隐私为核心的价值理性的人文复归,将那些正在被侵蚀的存在价值重新嵌回技术设计之中,使技术的进步能够真正服务于人的发展,而非削弱人的主体性。

“悲伤机器人”作为一种新兴的数字哀悼技术,其设计初衷根植于价值理性的关怀。从设计理念上,该技术旨在回应人类自古以来的情感需求:在不可逆转的死亡面前寻求一种连续性的精神联系。这一目标具有浓厚的人文色彩,是技术赋能情感关怀的重要进步。然而,在实际的研发与应用中,“悲伤机器人”却因陷入工具理性的逻辑,将哀悼和悲伤简化为一种可通过数据模拟和算法优化解决的“问题”。当哀悼过程被视为数据处理,控制就取代了理解,那些在逝者和哀悼者之间的隐私、尊严、自主性以及人类真实的情感交往需求便被忽视。此刻,“悲伤机器人”的隐私困境深刻暴露该技术的一个尖锐悖论:本欲通过增强人文关怀而设计的技术却在实际研发和应用中破坏实现关怀所需要的基础——隐私和自主性。“悲伤机器人”引发的隐私危机揭示了一个更为深层的技术伦理问题:技术如若缺乏价值理性的引导,则可能演变为一种新型的“情感殖民”工具。截至目前,“悲伤机器人”的设计与开发依旧

遵循着“自由放任的精神,当涉及他们的产品对用户的影响时,开发人员通常会采用一种‘我们会去解决’的态度”,^[17]这种“先行后思”的工作方式将伦理考量降级为事后的装饰品,致使用户的隐私同意往往流于形式,更多时候是基于不平等信息条件的被迫选择,那些伦理价值所重视的隐私、尊严、自主等原则在这一技术应用面前被边缘化。

对“悲伤机器人”可能带来的风险及其根源的揭示说明,“无论是阻碍接受亲人的离世,还是提供安慰和缓解悲伤,悲伤机器人的应用效果不仅取决于个人的反应,还取决于这些技术的设计方式”。^[18]技术可以增强人文关怀,但前提是须以人的价值为导向,而非使人的价值臣服于技术的逻辑。因此,解决“悲伤机器人”带来的隐私问题,就需要将“守护主体自主性”为核心的隐私道德观念嵌入其设计结构之中,从而实现技术的人文转向。“悲伤机器人”在实际应用中所引发的隐私争议,根源在于对隐私概念的误解。在“数据控制”的概念框架下,隐私被简化为对数据的管理与控制,相关的保护机制集中于知情同意原则或对用户提供访问权和删除权。但是,此类观点存在局限,它使逝者与哀悼者被视为数据对象,忽略了隐私在构成人的存在、关系和意义方面的关键地位。并且,“现行法律规范具有现世性……有关个人信息处理的相关规范也主要围绕自然人展开”,^[19]导致逝者隐私通常不在考虑范围内,在“死后隐私的合法地位方面,法律尚未提供连贯的支持论据”。^[6]这表明,将隐私视为一种数据并局限于在世之人层面的考量,实质上忽视了其内在所承载的人格、尊严、情感联结以及文化层面的神圣意义。因此,若要将隐私作为伦理观念嵌入“悲伤机器人”的技术设计之中,应首先实现隐私从“权利”向“价值”层面的升维。

隐私是人的内在本质属性,其“本性在于对自我的捍卫,它使人们在不同的社会关系中有所保留地表达自己,以维护自己的尊严和自主性”。^[20]在哀悼语境中,隐私为哀悼者划出一个免受外界干扰的空间,使其能够以本真性

的方式面对死亡议题并进行深刻的自我对话和情感调节。通过这样的内在体验,他们才能将“死亡作为最本己的可能性承担起来,才能获得个体化,成为本真的自己”。^[21]即使逝者已故,他依然具备不可被随意践踏的道德地位。作为具有独立人格的道德主体,逝者应以真实、完整的人格形象被铭记与尊重,任何对他们的自我形象或人格相关数据的挖掘和不当使用,都是对其作为独立个体在道德层面的亵渎。因此,维护逝者隐私不仅是对其人格权的维护,更是对其作为“绝对他者”地位的伦理确认。

对工具理性的超越,不仅体现在对隐私价值在理论层面的重新审视,更反映在技术实践中,通过技术设计将伦理原则有效融入技术构建。这就要求将关于隐私的价值“物化”为技术设计的特征,确保“悲伤机器人”从设计到应用的全过程都坚持隐私价值取向的统一。以“隐私”为核心的价值敏感性设计,主要包含三个层次:第一,“悲伤机器人”的设计需要遵循隐私的情境完整性。“隐私的意义不是固定和绝对的,而是存在于一定的情境之中”。^[22]在哀悼活动中,逝者的生前数字足迹和哀悼者的悲伤情感都有其独特的交往和社会情境,一旦脱离原始的倾诉语境和当时的情感状态,便会引发隐私泄露的风险。隐私敏感性设计不是要完全放弃对数据的采集,而是要执行精准地退让,以防止数据在解构与再整合的过程中出现隐私侵犯问题;第二,“悲伤机器人”的技术设计需要实现对哀悼者和逝者的关系性守护。“悲伤机器人”在哀悼过程中起着中介调节作用,它承担着连接哀悼者与逝者之间情感纽带的功能,这就要求其设计不能仅追求创造一个安抚情绪的产品,而是致力于构建一个具有互动意义的关系场域,“悲伤机器人”在此承担保护功能而非窥视功能,目标是强化而非削弱哀悼者和逝者之间的伦理纽带;第三,“悲伤机器人”的技术设计需要对知情同意的选择机制做存在论层面的重构。传统知情同意的失败在于数据授权是一种静态的不可逆行为,根据“被遗忘权”的理论原则,个人与数据的关系会随时发生变化。同样,悲伤是一个持续变化的情

感过程,“悲伤机器人”的设计若要真正实现人类自主,就需要从设计结构层面重新进行权力配置,从“静态授权”转向“动态调整”,从而确保用户随时拥有数据被遗忘的自由,以此维护人之主权。

对“悲伤机器人”技术本质的深入探讨可见,对该技术的伦理设计并非对其发展的过度干预,而是一种旨在捍卫人类核心价值、防范技术主导风险的主动调控策略。通过构建以“隐私价值敏感性”为导向的技术设计路径,可促使“悲伤机器人”的技术创新朝着更加重视人类隐私与尊严的方向发展,实现对人类存在价值的维护。这一设计理念,不仅是“悲伤机器人”技术发展的未来,更是我们作为伦理主体必须为自己设计的未来。

[参考文献]

- [1] Jellesma, F. C., Vingerhoets, Ad J. J. M. 'Crying in Middle Childhood: A Report on Gender Differences'[J]. *Sex Roles*, 2012, 67(7): 412-421.
- [2] Walter, T. 'Communication Media and the Dead: From the Stone Age to Facebook'[J]. *Mortality*, 2015, 20(3): 215-232.
- [3] Brinkmann, S. 'Learning to Grieve: A Preliminary Analysis'[J]. *Culture and Psychology*, 2020, 26(3): 469-483.
- [4] Cholbi, M. *Grief: A Philosophical Guide*[M]. Princeton: Princeton University Press, 2022, 65-102.
- [5] Elder, A. 'Conversation from Beyond the Grave? A Neo-Confucian Ethics of Chatbots of the Dead'[J]. *Journal of Applied Philosophy*, 2020, 37(1): 73-88.
- [6] Harbinja, E. 'Post-Mortem Privacy 2.0: Theory, Law and Technology'[J]. *International Review of Law, Computers and Technology*, 2017, 31(3): 26-42.
- [7] Jiménez-Alonso, B., Brescó de Luna, I. 'Griefbots. A New Way of Communicating with the Dead?'[J]. *Integrative Psychological and Behavioral Science*, 2023, 57: 466-481.
- [8] Öhman, C., Floridi, L. 'An Ethical Framework for the Digital Afterlife Industry'[J]. *Nature Human Behaviour*, 2018, 2(5): 318-320.
- [9] Caltrider, J., Rykov, M., MacDonald, Z. 'Romantic AI Chatbots: Don't Have Your Privacy at Heart'[EB/OL]. <https://www.mozillafoundation.org/en/privacynotincluded/articles/happy-valentines-day-romantic-AI-chatbots-dont->

- have-your-privacy-at-heart/. 2024-02-14.
- [10] Lemma, A. 'Mourning, Melancholia and Machines: An Applied Psychoanalytic Investigation of Mourning in the Age of Griefbots'[J]. *The International Journal of Psychoanalysis*, 2024, 105(4): 542-563.
- [11] Lindemann, F. N. 'The Ethical Permissibility of Chatting with the Dead: Towards a Normative Framework for "Deathbots"'[D]. Osnabrück: Osnabrück University, 2022.
- [12] Dewitte, P. 'Better Alone than in Bad Company: Addressing the Risks of Companion Chatbots Through Data Protection by Design'[J]. *Computer Law and Security Review*, 2024, 54: 106019.
- [13] Renfree, I., Harrison, D., et al. 'Don't Kick the Habit: The Role of Dependency in Habit Formation Apps'[A], Kaye, J., Druin, A., et al. (Eds.) *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*[C], New York: Association for Computing Machinery, 2016, 2932-2939.
- [14] Morris, R. M., Brubaker, R. J. 'Generative Ghosts: Anticipating Benefits and Risks of AI Afterlives'[J]. arXiv preprint, arXiv: 2402.01662, 2024.
- [15] Levinas, E. *Totality and Infinity: An Essay on Exteriority*[M]. Pittsburgh: Duquesne University Press, 1969, 194.
- [16] Heidegger, M. *Being and Time*[M]. Stambaugh, J. (Trans.) New York: State University of New York Press, 1996, 39-40.
- [17] Nowaczyk-Basińska, K. 'Digital Afterlife Leaders: Professionalisation as a Social Innovation in the Digital Afterlife Industry'[J]. *Mortality*, 2025, 30(2): 333-354.
- [18] Voinea, C. 'On Grief and Griefbots'[J]. *Think*, 2024, 23(67): 47-51.
- [19] 陈曦宜. "数字逝者" 技术的媒介性与关系型规制 [J]. 法学家, 2025, (3): 14-27; 191.
- [20] 王金柱、张旭. 关于“隐私”本质的活动论探析 [J]. 哲学分析, 2021, (6): 113-124; 192-193.
- [21] 马丁·海德格尔. 存在与时间 [M]. 陈嘉映、王庆节 译, 北京: 商务印书馆, 2019, 305-306.
- [22] Nissenbaum, H. 'Privacy as Contextual Integrity'[J]. *Washington Law Review*, 2004, 79(1): 119-157.

[责任编辑 李斌]