大模型时代人机传播中的隐私风险及其应对

郭巧敏　张琳颖　董博越　郭明军　杨伯溆 *

【摘　要】人工智能大模型的应用让机器人变得愈发智能，与此同时伴随个体隐私的让渡。本文以人机传播中的隐私话题为切入点，在人机交互过程中，机器人融入生活对个体隐私带来的风险分为三个方面，一是收集信息的直接监视，如性别、年龄、个人居住地等；二是伴随数据集的丰富，加之“同意—使用”原则中个体和平台的不对等，情绪和精神状态等多维度敏感信息被收集；三是互动风险导致人机边界模糊，导致伴随性风险如安全风险、立法风险、社会风险、信任风险等产生，最终不利于网络空间的有序发展。基于此，本文试图从政策法律、技术规制及伦理设计等角度，提出人机交互中隐私保护问题的破解路径，以平衡技术创新、社会福祉和个人隐私保护之间的关系。

【关键词】隐私；人机传播；赋权；祛权

人工智能技术正在为机器人赋能，并改变机器人的纯工具与客体地位，不同于传统的人机关系即控制与被控制，人机关系正在走向分享控制的情境，不同类型的机器人将会承担起劳动替代、任务合作、情感沟通等角色，以上技术的进步为机器人步入日常生活奠定了技术基础，随之带来的是数据安全和隐私泄露，这将是大模型时代推动人机交互的重点问题①。从当前研究来看，目前大模型时代人机传播的隐私主要分为三个角度：一是特定类型机器人如医疗机器人、智能家居、新闻写作机器人使用中的隐私风险②；二是从法学、传播学、情报学等视角探求用户在使用大模型过程中，机器人的法律人格、传播侵权风险和感知行为研究③；三是从社会影响角度如社会公平、责任归属、机器对人的奴役和异化等角度探讨如何应对风险④。本文认为，人机传播的隐私话题是技术和社会互动的结果，基于此，本文从社会治理的角度出发，对人机传播的发展进行系统梳理，结合主要面临的隐私风险展开分析，并提出应对举措。整体脉络围绕以下问题展开，社会机器人给人机传播带来哪些变化？当前人机传播在隐私视角下存在哪些风险和问题？从隐私规制的角度，如何推动人机传播向良性发展？

*　郭巧敏，国家信息中心大数据发展部规划与应用处助理研究员；张琳颖，国家信息中心大数据发展部规划与应用处实习研究员；董伯越，新华社新华网编委、主任编辑；郭明军，国家信息中心大数据发展部规划与应用处处长；杨伯溆（通讯作者），北京大学新媒体研究院教授、博士生导师。

① Zhao J, Zhang A, Rau P L P, et al. *Trends in Human-Computer Interaction in the 5G Era: Emerging Life Scenarios with 5G Networks*, Interacción, 2020.

② 赵玲玲、郭遥:《智能医疗机器人应用的伦理风险及其治理路径——基于利益相关者视角》,《科技管理研究》2023年第11期。

③ 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。

④ 刘壮、胡景谱:《社会化机器人引发的“社会问题”探析》,《科学·经济·社会》2023年第3期。

一、人机传播：诞生、影响和担忧

（一）社会机器人的崛起

机器人（robot）概念最早由捷克作家恰佩克提出，就该词的原始意义来看，被解释为完成人的工作的机械①。伴随技术的进步，众多学者将机器人定义为“具有传感器、执行器和移动性的网络物理系统”②，突出了机器人的机动性，并将其置于移动传播领域，与智能手机和移动媒体相比，机器人可以增强自治功能。在此基础上，Bekey等人将社会机器人的概念定义为能够感知、思考和行动的机器③。这是因为机器人不仅在工业生产领域广泛应用，也正在进入生活领域，下一代社会机器人将渗透至健康管理、家庭看护、娱乐、教育等领域，并可以和人类展开互动，如人形机器人Nao、Pepper和类似海豹的护理机器人Paro都是典型的研究对象。过去有关机器人的研究多集中于工业情境中如何实现人机的安全协作和信任以及改善机器人的设计等问题，但对隐私主题的研究仍不足。当越来越多的机器人融入生活，隐私问题就必须要面对。伴随大模型发展，社会机器人通过捕捉人类行为数据变得更加智能化，与此同时，网络社会的成熟正在让人们的隐私保护意识越来越强，如何处理二者之间的隐私让渡问题，是人机传播得以实现的重要影响因素。因此，下文从人机传播的发展脉络出发，指出其交互不仅是技术进步的体现，更是一种社会过程，其中伴随技术的成熟，社会机器人的主体意识越来越强，带来了对身份与权利的讨论。

（二）人机传播带来身份与权利的探讨

社会机器人可以和人类进行互动，区别于计算机介导的交流（CMC），CMC是通过电子中介进行的人与人之间的沟通交流，计算机是一个通用术语，指的是那些物理上分离的人能够跨距离即时相互通信的电子设备，电话是典型代表。社会机器人则不同，他们一定程度上可以说是人类的替代品，研究将其称之为后人类的电子化（Post-human cyborgization）④。也就是说，人工智能和深度合成技术正在让人机互动不仅可以展开语言层面的交流，还可以让社会机器人通过深度学习技术更加符合社会道德规范，推动人机关系发生新变化。如今，技术与人之间的传播越来越逼近于类人化的人际传播，具有拟人化、人性化特征的社会机器人不仅会说话，还会和我们进行对话，它们知道个体的名字，能够分辨个体声音和学习个体偏好，它们作为积极的参与者进入了人们日常的生活世界。这些机器逐渐成为了传播的主体，标志着技术正迎来身份的转向，这些转向都推动了人机传播研究的发展。人机传播（HMC）的概念由萨奇曼（Suchman）提出，用来指代真实人类与智能体之间的传播交流活动⑤。随着人机传播研究的不断深入，传播领域学者们如古兹曼（Guzman）、甘克尔（Gunkel）等开始用HMC（Human Machine Communication）一词来搭建人机传播学术研究框架，伴随着拟人化、智能化的机器人的崛起，这一领域逐步改为Human Robot Interaction（HRI）。这充分说明，在智能化社会，技术不仅仅是具有工具属性，更拥有主体“意识”，而人机传播是一个在传播中正在形塑的问题，不仅是因为机器以虚拟的算法形式进入人们的日常生活，而且也作为社会行动者进入大众生活，如Jibo、Kubi、Pepper等具身化的社会机器人，Alexa、

① 郁乐：《智能机器人能够拥有权利吗？》，《华中科技大学学报（社会科学版）》2020年第5期。

② Tim, Smithers. Autonomy in Robots and Other Agents, *Brain and Cognition*, 1997, 34(1).

③ Lin P, Abney K, Bekey G. *Robot Ethics: The Ethical and Social Implications of Robotics*, MIT Press, 2012, p.33.

④ Zhao S. Humanoid Social Robots as a Medium of Communication, *New Media & Society*, 2006, 8(3).

⑤ Suchman L. *Human-machine Reconfigurations: Plans and Situated Actions*, Cambridge University Press, 2007, p.33.

Google Home等对话型机器人的商业化使用。技术越来越被设计成为传播者，也被更加广泛地使用。在智能物联的技术逻辑下，从冰箱到汽车、手表甚至家庭都将能够直接与人交换信息。在传播学研究中，CMC—HMC—HRI的进化路径显示人类和机器交互获得意义的过程，理应对社会机器人的主体意义做出回应。

一方面，需要我们重新思考社交关系。虚拟共存是后人类社会的一部分①，特征是高度连接，人类与类人生物的融合并进行信息交换，因此我们需要将社会关系进行新的概念化，传统的社会关系是强调人与人之间的互动，但社会机器人的崛起意味着人际关系不再局限于传统意义上的人，还包含与类人个体的互动。

另一方面，在人机传播环境下，社会关系概念化的核心离不开对机器人身份与权利的探讨，这意味着需要对机器人的人格权问题展开思考。人格权作为一个法律概念，强调的是民事主体所固有而由法律直接赋予民事主体所享有的各种人身权利。目前业界和学术界针对机器人主体地位探讨存在三种声音：第一种声音认为主体机器人可以作为行为主体，主要是媒体报道的案例，比如沙特的女机器人索菲亚（Sophia）拥有了永久居民权，日本机器人Shibuya Mirai拥有永久居留权。第二种声音认为机器人不能作为行动主体，机器人要满足道德行动者，就必须满足三个问题，一是机器人是否自主；二是机器人是否有意识，一般来讲意识是人脑特有的，机器只能通过预定程序来进行，且目前的脑科学研究表明必须通过神经系统才能实现意识阶段；三是机器人是否处于负责任的地位，这是确定机器人身份与权利的前提②。机器人目前与人最大的差别是创造力③，而且也没有自我意识、道德意识以及肉体感受性，其生命是可以修复的，而人类生命则不是。此外，机器的自主性是因为在“感知—思考—行动”基础上的智能机器人具有技术自主性，即可以被界定为在外部世界作出独立于外在控制或影响的决定并予以实施的能力，这一自主性具有纯技术本质，而这种自主性程度取决于机器人被设计成可与环境进行交互的复杂程度，在可见的未来都只是高度自动化的机器对人类亲自掌控的逐步摆脱，因此这种自主性与规范意义上的自主性并不相同，而且其自主性本身是基于感知技术和决策控制技术产生的④。第三种是综合两种声音，认为机器人具备数字人格特征，作为身边的助手，帮助处理日常琐碎工作，能够让自己作出一些判断⑤，比如在具体机器人的设计中，可以根据机器人工作的场景确定自主性程度，比如地面、水下、空中场景对人机交互提出了含蓄（implicit）与直接（explicit）的不同要求，自主性上也分完全自主性（full autonomy）、大多自主性（mostly autonomy）和部分自主性（partial autonomy）⑥，在政府官方文件中，2017年2月6日欧洲议会通过的《就机器人民事法律规则向欧盟委员会的立法建议〔2015/2103（INL）〕》认为，现今具备深度学习能力与自主性的机器人能够从经验中学习并作出“准独立”的决策。该能力使其越来越接近于能够与环境发生交互并且作出重大改变的智能体（agents），从而考虑赋予具有深度

① Pepperell R. *The Posthuman Condition Consciousness Beyond the Brain*, The Posthuman, 2003, p.56.

② 陈俊秀、李立丰：《“机器意识”何以可能——人工智能时代“机器不能被骗”立场之坚守》，《大连理工大学学报（社会科学版）》2020年第6期。

③ 韩东屏：《未来的机器人将取代人类吗？》，《华中科技大学学报（社会科学版）》2020年第5期。

④ 王莹：《法律如何可能？——自动驾驶技术风险场景之法律透视》，《法制与社会发展》2019年第6期。

⑤ Anderson M, Anderson S L. When Is a Robot a Moral Agent? *International Review of Information Ethics*, 2006, (6).

⑥ Islam M J, Hong J, Sattar J. Person Following by Autonomous Robots: A Categorical Overview. *The International Journal of Robotics Research*, 2018(12).

学习能力和自主性的最复杂的智能机器人以“电子人格”（Electronic Person）的法律地位和责任主体地位①。机器人身份与权利关系讨论的背后是赋权祛权问题，关于这一问题，二者是此消彼长的，机器人被赋权过多，则意味着人类被祛权越多，这就像国家和社会、商业机构与消费者、传统职业与新职业之间的关系问题②。

（三）赋权祛权问题折射出隐私担忧

赋权祛权问题是技术发展带来的新问题。对于社会机器人而言，人身权利的讨论核心则是人与机器人的隐私让渡问题，在网络空间环境下，这种隐私担忧与无处不在的数据收集密切相关。早在20世纪90年代中期，当时互联网还处于萌芽阶段，David Lyon就预见到了监视社会的出现，在大公司和政府部门的庞大计算机数据库里，个人生活的精确细节（个人信息、行为信息、位置信息等）被不断收集、存储、检索和处理。社会机器人的智能化离不开大量的数据训练，但也助长了人们对大数据侵犯个体信息的恐惧③。这是因为在大数据时代，即使是碎片化的个人信息，经过整合后也可以刻画出信息主体的数据形象④，使商家的精准营销成为可能，个体的信息裸奔成为现状。不得不说，这些信息是网络社会得以流动的重要因素，但从治理的角度来看，这些被收集的信息可以是财富，可以是权力，对生活在数字时代的个体产生结构性规制，当个体的太多信息被让渡出去且失去保护之后，社会性死亡就开始了。

可以说，信息技术的发展让隐私问题成为全球话题，世界各国都出台了相应的法律规范保护隐私，如经济合作与发展组织（OECD）发布的《隐私权指南》，美国的《电子通信隐私法》，欧盟的《通用数据保护条例》等，都从不同角度保护个体隐私和信息安全。就我国而言，2020年5月28日，十三届全国人大三次会议表决通过了《中华人民共和国民法典》，规定人格权是民事主体享有的生命权、身体权、健康权、姓名权、名称权、肖像权、名誉权、荣誉权、隐私权等权利。除前款规定的人格权外，自然人享有基于人身自由、人格尊严产生的其他人格权益，并对隐私权和个人信息保护做了相应的法条阐释。从法律的角度讲，隐私是一项基本权利，对于民主社会的建立、个人的自由、心理健康、个性和创造力都具有重要意义⑤。因此，就人机传播而言，隐私到底是什么？为什么值得保护？它有多大的价值？破解人机传播中的隐私担忧，必须正视隐私的意义以及当下人机传播中隐私保护存在的风险和利益博弈等问题。

二、隐私视角：意义、风险和问题

（一）隐私的概念与意义

在全球范围内，隐私都是极为重要的话题，学界最早对隐私问题的探讨起源于1890年，美国法学家塞缪尔·沃伦（Samuel Warren）和路易斯·布兰代斯（Louis Brandeis）撰写了《The Right to Privacy》文章，他们主张用“right to be alone”作为隐私权的定义。信息技术的崛起，

① Solène Gérardin. European Parliament Resolution with Recommendations to the Commission on Civil Law Rules on Robotics, 2017(2).

② 张卫：《数据的赋权与祛权：基于微观权力的数据伦理分析》，《伦理学研究》2019年第2期。

③ König R, Uphues S, Vogt V, Kolany-Raiser B. The Tracked Society: Interdisciplinary Approaches on Online Tracking. *New Media & Society*, 2020, 22(11).

④ 郝思洋：《大数据时代个人信息保护的路径探索》，《北京邮电大学学报（社会科学版）》2016年第5期。

⑤ Solove D J. Understanding Privacy. *Social Science Electronic Publishing*, 2008, 59(7).

隐私话题成为世界一线问题，1960年起，隐私问题的研究稳步上升。

就概念而言，“隐私”的阐释与客观环境紧密相关，并具有深刻的文化内涵。按照上文塞缪尔·沃伦和路易斯·布兰代斯的解释，隐私就是“不受打扰”的权利，存在于“私人和家庭的神圣领域”。因此隐私一般指的是个人不愿他人干涉与侵入的私人领域，与人的私密方面相关。这体现出隐私的个人信息控制特征，也就是说隐私的核心不在于是否有人知道，而在于在什么时间和场合，让谁知道。明确隐私概念之后，可以发现导致隐私问题的技术原因是海量数据的共享与挖掘，其社会性后果是主体身份的数据化，直接表现是信息的泄露和信息的破坏。伴随数据商用的发展和个人主体意识的崛起，隐私关注就成为技术使用和接受的重要因素，根据Lanier和Saini的分析，影响个体的隐私关注分为三类：（1）收集，消费者希望被公司告知其个人信息的收集和使用情况。（2）控制，消费者希望感觉自己对个人信息的收集和公司之间的信息共享具有一定的控制权。（3）安全，大多数消费者希望确保他们提供给公司的个人信息，尤其是在线信息的存储是安全的、不被滥用的①。因此，具有更高隐私关注度的消费者在提供其个人身份时认为会面临更高的风险，这也被认为是影响新技术普及的重要因素。而国内目前的技术形态使用与隐私的研究中侧重社交媒体与隐私保护，比如揭示微信用户隐私保护的现状及其制约因素，将自我效能感、风险感知、信任、隐私担忧及隐私保护作为变量进行测量，探究社交媒体使用中隐私担忧的现状②和社交媒体不持续使用的模型建构，还有基于社交媒体中的隐私悖论、隐私侵权与隐私保护展开研究，均属于网络隐私权的研究范畴。在关于隐私的定性研究中，通常是将网站隐私声明作为文本分析重要的切入点，以此提出对应规范网络平台的隐私入侵的建议。

维基百科上针对隐私的讨论围绕有限的访问权限、个人对自己信息控制的能力、隐私被保护的状态（孤独、亲密、匿名和保留）、人格权和自主权、删除权和拒知权、被遗忘权、隐私与自我认同的建立、技术的隐私侵犯与保护、隐私监控、不同国家隐私的法律保护、移动互联网时代的隐私、隐私与新技术、隐私悖论、隐私素养、青少年隐私保护、隐私的经济估值、自拍文化与隐私等，显示了隐私的跨学科属性，法学、信息管理学、新闻传播学和心理学都有对应的研究。

可以看出，国内外都对隐私展开了对应的研究，侧重点有所不同，但共识是隐私被视为一项基本权利，这具有重要的意义。在个人层面上，无监督行动自由被认为是个人自我发展和产生创造力的重要基础，在社会发展层面，隐私被认为是对自由主义人格的保护，它可以促进思想自由、政治民主和社会进步。在此基础上，隐私被概念化为“与个人的思想、观念和行为有关的状态”，不同学者对隐私的边界展开讨论，区分了隐私的四个方面，即信息隐私（informational privacy）、心理隐私（psychological privacy）、社交隐私（social privacy）和物理隐私（physical privacy）③，具体讨论内容如表1所示。

① Lanier C D, Saini A. Understanding Consumer Privacy: a Review and Future Directions, *Academy of Marketing Science Review*, 2008,12, (2).

② 徐敬宏、侯伟鹏：《“隐私担忧”的中介效应：基于对大学生微信使用的结构方程模型分析》，《传播与社会学刊》2020年第5期。

③ Lutz C, Maren Schöttler, Hoffmann C P. The Privacy Implications of Social Robots: Scoping Review and Expert Interviews. *Mobile Media & Communication*, 2019, 7(3).

表1 四种隐私类型的关注焦点

隐私类型	包含内容
信息隐私	数据、敏感数据、安全、第三方、云计算、透明度、信息收集的知情权
心理隐私	心理依赖性、自主性、隐私担忧、寒蝉效应(Chilling Effect)
社交隐私	人和机器人的连接、信任和影响
物理隐私	私人空间的可接近性与不可接近性、不舒服的交往距离

不得不说，正如韦伯的理想类型，这种隐私类型学的界定为研究隐私划定了界限。在人机传播的研究中，四种隐私穿插其中，信息隐私是产生担忧的基础，也是目前个体最为关注的，与个人信息紧密相连，如性别、年龄、身高、收入、职业、住址、个人网络活动信息、个人网络存储信息等，集中体现为互联网使用中的“同意原则”失灵、个体面临数据收集具有极大的被动性最终带来的信息泄露。伴随智能设备深入社会生活领域，与之相伴的心理隐私、社交隐私和物理隐私成为个体必须要面临的问题。

（二）人机传播中面临的隐私风险

信息数据的收集是网络社会运转的要素，不同于传统社会，新媒体使社会空间、个人空间、私人空间和公共空间的边界变得模糊，人机传播中的隐私问题显得更加重要。在人机交互过程中，机器人融入生活对个体隐私的影响分为三个方面，一是收集信息的直接监视，信息社会里个体信息隐私边界的模糊化、个人信息数据的易得性（电子监控系统的记录；社会管理机构、公共服务部门或商业消费场所等保留的个人信息和上网记录等）、个人信息数据的难隐匿性、个人对信息数据掌控力下降，当然这也与政府、媒体、商业机构对信息隐私的监视、展示与利用密不可分①，这些为机器人监视个体和家庭提供便利；二是访问权限的增加触及敏感信息，个人不同意就没办法享受机器人在居家生活中带来的便利，社会机器人可以随时随地收集有关用户日常生活的信息，并且通过众多传感器收集有关敏感特征（如情绪和精神状态）。例如，用户在与机器人互动时随便评论自己的感受，从而可以记录以前难以掌握的情感信息。敏感信息还包括用户图像，如在卧室和浴室中的用户、房屋内的布局、疾病和健康状况，这些都是特别敏感的信息，会让个体产生不舒服的感觉；三是社交风险，机器人在设计上越来越像人类，并且在社交方面具有交互性，从而使其对用户和更大的社区更具吸引力。许多研究表明，人们很难对诸如机器人之类的拟人化技术做出反应，因此与机器人互动中带来的伴随性风险就产生了，如安全风险、立法风险、社会风险和信任风险。

就安全而言，机器人让黑客获取全面信息更加便利。机器人是信息时代的产物，因此其特征具备了迭代中物质特性的融合，也就是说机器人既有工业时代作为机器的物理性风险，也有信息时代的信息泄露风险，还有在未来物联网时代下的叠加风险，而且这些风险伴随机器人逐步融入社会生活，同时具有场景的开放性特征，因此人机传播更是一个公共安全问题。这是因为社会机器人一方面带来新的安全风险，另一方面也导致已有的安全风险被放大，主要原因是社会机器人无需随身携带，但具有独立且增强的移动性，这增加了它们监视和进入私人空间的潜力，可能会对用户的身体隐私产生负面影响。除了物理意义之外，我们还看到了对其他形式隐私的影响，机器人可能会影响用户的心理和社会隐私②，这是因为社会机器人可以和人类建立纽带，对社交和信息隐私具有潜在

① 张霆：《大数据时代信息挖掘、利用中的公民隐私保护》，《河北师范大学学报（哲学社会科学版）》2016年第5期。

② Calo, Ryan. *Robot Ethics: The Ethical and Social Implications of Robotics*, Patrick Lin, George Bekey, and Keith Abney, eds., Cambridge: MIT Press, 2010, p.27.

影响。另外由于算法技术的不透明性，机器人越来越依赖基于云的数据处理以及大量的传感器和执行器，这使它们成为高度复杂和不透明的系统，其数据收集常常为用户所未知或误解①，最终使得它们成为隐私信息的载体，从而增加了日常生活的安全风险。

人格权是重要的法律概念，在人机传播环境中，人类行为的界定出现了模糊，一方面社会机器人具备了一定的主体“意识”，另一方面，人类在与社会机器人互动的情况下产生的社会行为对他人带来影响。就像自动驾驶技术融合也带来了参与主体与行为主体（汽车整车生产商、自动驾驶系统不同模块与功能的制造商与提供商、驾驶员、智能网联服务提供者等）的多元化、因果链条的延长与责任界限的模糊等问题。人机传播过程中，不仅有机器人开发者、提供者与生产商，还有参与其互动的使用者。因此，在立法层面，机器人与人的民事责任主体地位、刑事责任主体地位以及物理信息风险的叠加，都给立法带来挑战。

就社会风险而言，一是不可否认代码和算法在网络空间秩序崇高的重要性，而其背后是一种技术规制的考量，但技术规制与传统的法律规制、道德规制还有很多需要弥合的地方，因此也带来了设计机器人伦理机制的结果不确定性；二是非公正性更加突出，大众传播时代强调媒介素养，在信息社会、数据时代，信息素养、数据素养相伴而生，这些都与网络使用能力或者说ICT效能密不可分，这些受到政治、经济、文化、社会等多层因素影响，如果鸿沟问题尚未解决，那么势必带来技术普及的不公平性，最终的结果是少数人通过机器人实现便利；三是社会机器人的伦理建构规则复杂，当前人工智能的反馈机制和人类智力的反思平衡之间仍然存在功能鸿沟②。此外，和游戏、药品、赌博、网络、手机成瘾一样，需要考虑机器人的设计是否会导致人类上瘾与滥用。

在信任层面，以往人机传播研究指出，推动技术普及的核心要素之一是信任，这其中涉及人、机器人和人机传播中的特定环境，具体如表2所示。

表2　影响人机传播的信任因素③

人类因素	机器因素	环境因素
基于能力	基于绩效	团队合作
注意能力、参与程度、专业性、工作强度、竞争性、先前经验、环境感知。	行为独立性、可靠性、可预测性、自动化程度、失败比例、透明度、错误警报。	团队文化、团队的信息传播和共享方式。
个体特征	基于属性	任务特征
人口统计学、个人特质、对机器人的态度、自信心、与机器人相处的舒适度、信任倾向。	机器人的个性、适应性、类型、拟人性、接近性。	任务类型、任务复杂度、物理环境、任务需求。

毫无疑问，隐私和机器人技术之间的风险以及其他相关性将在未来智能社会中详细展现出来。随着技术的发展，互联网平台作为信息基础设施的作用逐步凸显，这意味着机器人背后的平台技术公司对用户信息的垄断需要引起重视，但当前个体和平台技术公司利益处于失衡状态。

（三）提供者使用者利益失衡

在人机传播中，存在两个主体，一是作为技术使用者的人，二是作为技术研发者的平台公司。

① Lee M K, Tang K P, Forlizzi J, et al. *Understanding Users! Perception of Privacy in Human-robot Interaction*, ACM International Conference on Human-robot Interaction, 2011.

② 陈凡、徐旭：《当代人工智能伦理设计的困境和超越》，《华中科技大学学报（社会科学版）》2020年第5期。

③ Hancock P A, Kessler T T, Kaplan A D, et al. Evolving Trust in Robots: Specification Through Sequential and Comparative Meta-Analyses. *Human Factors*, 2021, 63(7).

在个体层面，隐私保护面临两个典型问题，一是隐私悖论，这一概念描述了隐私态度与行为之间的差异①，即个体虽然有隐私保护意识，却不自觉在社交媒体上披露更多个人信息，对于这一行为的解释有两种，通常是用户缺少风险意识或者用户隐私素养不够；二是用户的隐私疲劳和隐私冷漠（privacy apathy and privacy fatigue）②，这描述了用户在隐私数据保护时的态度，我们可以理解为隐私麻木，即用户会感到无能为力，所以就放任平台、机构和其他的用户可以前所未有访问数据，这种行为也可以称之为数字辞职③。这与平台公司的垄断密不可分，为预防和制止平台经济领域垄断行为，引导平台经济领域经营者依法合规经营，促进线上经济持续健康发展，虽然我国目前出台了《关于平台经济领域的反垄断指南》，但在网络社会中，整体技术风险都存在一定的滞后性，这是因为平台除了传统的规模效应大者越大之外，新的平台型公司还有另外两个能力：一是网络效应，二是数据智能。在享受平台便利之时，个体成为互联网平台争夺用户注意的工具，不仅知情同意权利被掠夺，而且匿名化技术在算法的加持下可以整合出完整的用户画像。因此，我们不得不思考在与机器人相处中，平台的数据所有权和个人的隐私权保护问题，并解决用户和平台的权利不对等问题。平衡个体隐私保护和平台技术公司垄断是今后网络社会治理的主要问题。

三、治理归途：法律、技术和伦理

只要未来人类依然希望与其他智能体的无碍交流，那么，构建出一个人类与机器人共同认可的规范将是网络社会治理的必经之路④。如上文所述，人机传播中关键问题的破解与人机交互中的隐私保护、隐私担忧、隐私风险、隐私让渡等隐私问题密不可分，只有解决这些问题，人机传播才得以可能。从网络社会治理的角度上看，本研究试图从影响人机交互的三个层面原因即机器人、人、人机交互中的政策环境出发，围绕政策法律、技术规制和伦理设计角度展开基础性思考，同时认为应该引入可允许的风险理论，以平衡技术创新、社会福祉和个人隐私保护之间的关系。

（一）建构适应网络社会的政策法律

如上文所述，个体和平台技术公司在处理个人信息层面是不对等的，因此从网络社会治理的角度看，如果说PC互联网和移动互联网时代，人机传播的发展离不开技术和市场，在物联网时代，政策法规的保护是人机传播得以顺利发展的基石⑤。从《民法典》中将隐私权作为人格权开始，紧接着国家出台了《个人信息保护法》，到国家市场监督管理总局发布《关于平台经济领域的反垄断指南》，都说明政策法律对规制技术发展具有重要作用。结合人机传播中的隐私问题，可以围绕以下角度展开。

一是区分隐私边界。特别是在信息隐私领域，社会和机构隐私问题之间的区别可以用于衡量不同利益相关者（例如制造商、黑客、其他用户、第三方软件提供商）的隐私风险感知，有关网络隐私边界的研究，可追溯到美国学者S.Petronio的传播隐私管理理论，该理论用边界来限定私人领域

① Bib S. Barnes, Susan. A privacy paradox: Social networking in the United States, *First Monday*, 2006(11).

② Lutz C, Hoffmann C P, Ranzini G. Data Capitalism and the User: An Exploration of Privacy Cynicism in Germany. *New Media & Society*, 2020, 22(7).

③ Draper N A, Turow J. The Corporate Cultivation of Digital Resignation. *New Media & Society*, 2019, 21(8).

④ 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67页。

⑤ 胡泳、年欣：《中国数字化生存的加速与升级》，《新闻与写作》2020年第12期。

与公共领域界限，边界的一边不披露私人信息，而另一边披露私人信息。在传统社会中，私人领域与公共领域的边界较为固定，而网络社会的连接性和界限模糊引发边界连接和边界渗透等过程，尤其在社会化媒体中隐私存在多重边界，信息传播的过程中个人与外部连接完成初步的信息传播，再进一步的信息转发、分享后达到边界渗透的效果。边界的模糊化是智能社会的典型特征，比如5G环境下的VR、AR空间环境。Helen Nissenbaum建议在具体的社会、技术背景下分析隐私，一方面是要基于一定的社会情景理解隐私，因此不仅需要宣称“私人”与“公共”之间的二分法，更重要的是需要认识到在一个领域被认为是公共的东西，在另一个领域可能是私人的东西。我们愿意把财务生活的细节呈现给我们的会计师，但通常不会向他们展示我们赤裸的身体。相反，尽管医生对我们的健康了如指掌，但他们并不会掌握我们的收入情况。简而言之，隐私的保护只能通过语境来理解。当然，法学家Solove也提醒我们单从语境看待隐私往往无法为决策或法律决定提供充分的方向，因为决策或法律决定依赖于归纳总结。因此隐私的评估必须在一般性与特殊性之间、抽象与具体之间的张力中航行。

二是基于机器人位阶赋予对应权利。针对人机传播中隐私保护的法律制定，还可以基于人工智能技术的发展程度，围绕实力等级和发展阶段赋予机器人相应的人格权利①，具体如表3所示。

表3　机器人权利位阶表

名称	技术基础	社会阶段	人格定位
初级人工智能下的机器人	算法模拟	工业社会	无人格
弱人工智能下的机器人	大数据	信息社会	有限人格
强人工智能下的机器人	云计算	智能社会	完整人格
超人工智能下的机器人	技术叠加	物联社会	同等人格

三是基于位阶开展个人隐私保护。社会机器人技术的进步为人机交互提供了技术基础，因此机器人的技术等级是人机交互中隐私分配的基础。在无人格的机器人技术条件下，个体可以完全控制机器人，可以不用释放隐私数据。在机器人具备有限人格的条件下，个体可以让渡一部分个人信息隐私数据到机器人，但是这其中，人依然占据主导地位。在完整人格的条件下，需要释放敏感数据，但是要有更高的标准，即可以每一次交互中都做到具体的同意，逐次地授权。在同等人格条件下，人与机器人之间的信息可以实时传输，权利义务可以实现转移，但此时就需要对机器人的生产使用进行实名认证。可以看出，伴随机器人技术的成熟，人类让渡的隐私数据越来越多，甚至到没有隐私的程度。

（二）技术负效应呼吁更先进的技术

媒介决定论的研究者麦克卢汉、伊尼斯一定程度上认为，新媒介技术对社会文化、社会心理都产生深远的影响，一种新媒介的出现也会带来一种新文明。但新技术的发展也会带来一定的弊端，如James Katz所说，技术发展带来的传播问题，只能用更先进的技术解决。针对影响人机交互中的隐私问题，可以从机器人类型、隐私保护的风险识别、隐私计算和评估三个角度出发寻求出路。

一是明确机器人的分类。如前文所述，社会机器人不是一个现代概念，但是人们普遍同意当代社会机器人项目始于1956年著名的AI达特茅斯夏季会议，当时确定了构建具有普通人类一般智慧

① 杨学科：《论人工智能机器人权利及其权利界限》，《天府新论》2019年第1期。

的机器的议程。1960年后期，第一台自主式人形机器人“Shakey”在斯坦福研究所建造。一段时间以来，许多研究人员认为仿人般的人工智能是可以实现的。然而，在反复失败以克服所谓的“世界知识”问题之后，人工智能项目在20世纪70年代和80年代陷入停滞。互联网在20世纪90年代的普及促成了众多在线会话代理的兴起，例如Julia和Alice，他们可以与人类讨论多种话题。根据他们所扮演的角色，研究者将社交机器人分为两种主要类型：功能型社会机器人和情感型社会机器人。功能型社会机器人是旨在与人类进行互动以实现工具性目的的人形社交机器人。当前，这样的机器人在商业领域中被广泛使用，以代替人类来服务人类客户，比如销售接线员、服务器导游、商场接待员，这些机器人与人类互动，要么是无形的实体，要么是能够表达语言（基于文本或语音）以及非语言表达的拟人化人物。情感型社会机器人是指在情感上与人类互动的社会机器人，这些机器人主要用于两种环境：在线聊天和私人住宅。在线聊天的匿名设置已成为聊天机器人的自然栖息地，聊天机器人被设计伪装成人类，可以帮助人类减轻孤独感，建立友谊关系，在私人家庭中，情感型社会机器人充当宠物和玩偶，它们与人居住在一起，并且能够执行简单的家务劳动，有娱乐和陪伴功能①。人机互动中，两类机器人都是常见的类型，需要基于机器人功能的不同，展开隐私边界的确定。

二是隐私决策的风险感知。隐私决策在很大程度上与风险感知密不可分②。因此在人机传播中，进行隐私风险识别、风险分析和风险减轻是隐私问题的解决策略③。具体措施如表4所示。可以看出，风险的感知是基于语境的，因此有研究者通过使用熵作为一种手段来衡量上下文感知服务中的隐私保护问题，该熵用于衡量定位用户的下落和识别个人选择的能力，用于计算对上下文感知服务器的查询中位置和个人偏好报告的抽象级别，这种方法在服务提供期间的用户数据报告中应用。

表4 隐私保护的风险判定阶段

过程阶段	解决措施
风险识别（risk identification）	七种隐私类型：个人隐私、行为隐私、交流隐私、数据和图像的隐私、思想和感情的隐私、位置和空间隐私、联合隐私。
风险分析（risk analysis）	隐私保护目标：机密性、完整性、可用性、不可链接性+数据最小化、介入性、透明性。
风险减轻（risk mitigation）	监控得以缓解，维护个体和机器人的独立空间。

三是隐私保护操作化。在隐私风险感知的基础上，有研究者基于人员流动数据的观察发现粗糙的数据集也几乎没有匿名性，这从另外一个角度说明个人隐私的基本限制对设计致力于保护个人隐私的框架和机构具有重要意义。可以预见未来，机器人的智能化必然离不开数据的收集，而数据一旦得不到保护，其普及就会成为难题，因此机器人平台公司必须重视隐私技术评估的研究。可以从隐私设计原则④、隐私计算（Privacy Computing）技术和隐私影响评估⑤展开，具体见表5。

① Zhao S. Humanoid Social Robots as a Medium of Communication. *New Media & Society*, 2006, 8(3).

② Frik A, Gaudeul A. A measure of the Implicit Value of Privacy Under Risk. *Journal of Consumer Marketing*, 2020, 37(4).

③ Heuer T, Schiering I, Gerndt R. Privacy-centered Design for Social Robots. *Interaction Studies*, 2019, 20(3).

④ 郑志峰：《人工智能时代的隐私保护》，《法律科学：（西北政法大学学报）》2019年第2期。

⑤ Information Commissioner' s Office. Conducting Privacy Impact Assessments Code of Practice, 2014.

表5　隐私保护的评估标准

隐私设计原则	隐私计算技术	隐私影响评估
第一，积极预防，强调从最开始的设计阶段，就考虑隐私保护问题。 第二，隐私默认保护，让隐私保护成为企业商业实践和系统运行的默认规则。 第三，将隐私嵌入设计之中，成为系统组成部分。 第四，功能完整，正和而非零和，主张实现用户、企业等多方共赢。 第五，全生命周期保护。 第六，可见性和透明性，并接受独立核查。 第七，尊重用户隐私，确保以用户为中心。	在保护数据私密性的前提下，完成对数据的计算分析任务，从而构建了一种可信任的执行环境。从技术角度来看，当下“隐私计算”指的是利用可信执行环境、安全多方计算、同态加密、零知识证明、差分隐私和联邦学习等系统安全技术与密码学技术，在保证原始数据安全隐私性的同时，实现对数据的计算和分析。美国人口普查局、苹果公司和Facebook作为差分隐私（隐私计算的一种类型）的研究者，认为这种技术可以帮助建立人机传播时代的信任机制。	确定隐私评估的需求、描述数据流、识别隐私和相关风险、界定和评估隐私解决方案、签署并记录隐私影响评估的结果、将评估结果反馈。

（三）伦理道德设计回应数据正义

2023年国家网信办联合7部门印发《生成式人工智能服务管理暂行办法》，提出尊重他人合法权益，不得危害他人身心健康，不得侵害他人肖像权、名誉权、荣誉权、隐私权和个人信息权益，确保人工智能规范、有序、可持续发展。而在伦理设计层面，则要考虑主体的道德后果、网络安全和隐私素养的普适性。

《2001：太空漫游》（2001：A Space Odyssey）经常被用作机器人伦理设计的讨论，但核心问题是在多大程度上为智能自动化机器赋予道德责任。伦理规范的讨论对那些拥有道德地位的实体负有义务或责任。一方面，目前学术界主要讨论是立足将机器人作为道德承受体的角度，比如享受终极关怀的依据、公共规范层面的内容，最终要尊重机器人身上的人性，人是机器人的道德监护人，要持续扩展道德关怀与伦理共同体的范围①。另一方面，道德承受体需要基于人机交互关系展开讨论，最终核心变量还是拟人化。这呼吁社会机器人的系统掌控者即机器人专家开发基于隐私保护的道德规范，并能够普遍监控机器人发展带来的道德后果。比如可以在编程级别上根据社会规范、价值和义务来调整个人数据输入和输出的阈值，而不仅仅是基于经济价值和效用。从这个意义上说，这是一种比技术上更道德的方法，不仅仅是工具上的观点，意味着算法可以被视为制度，但其作用必须服从公共利益。同样，数字平台中的隐私保护应被视为一种制度性产品，由此与用户的互动承担了监管范围内的公共义务。基于此，政策制定者、学者和倡导者可以针对基于实用程序的机器学习和大数据智能证明监管监督和干预措施的合理性②，消费者可以通过道德设计规范对机器人进行编程，使其不仅可以在特定时间操作或完成特定任务，也保护个人空间。

在网络安全层面，必须警示数据资本主义和监视资本主义，二者都是通过对用户信息的监视获取经济利益，这些个人数据既是用户访问服务所必需的，又是用于收集、分析和出售给第三方平台可无限获利的③。伴随技术的进步，这些监视将不仅局限于物理观察、监听或窃听，还包括心理监视（使用性格测试和测谎仪作为人员选择的手段）和数据监视（集中收集计算机银行中个人的信息），这给网络安全带来巨大挑战。

在隐私素养层面，与数字鸿沟密不可分。美国哈佛大学教授Pippa Norris从3个方面定义“数

① 杨通进：《论机器人的道德承受体地位及其规范意涵》，《哲学分析》2019年第6期。

② Jin P Y, Eun C J, Hee S D. The Structuration of Digital Ecosystem, Privacy, and Big Data Intelligence. *American Behavioral Scientist*, 2018, 62(10).

③ Shoshana Zuboff. The Age of Surveillance Capitalism: the Fight for a Human Future at the New Frontier of Power, *Journal of Social Political Philosophy*, 2019, 1(2).

字鸿沟”：一是全球鸿沟，指经济不平等造成的网络接入差距；二是社会鸿沟，指国家内部信息富足者和信息贫困者之间的差距；三是民主鸿沟，指在社会动员或参与公共生活的人们电子资源使用能力的差别。这一定义比较全面地概括了“数字鸿沟”的本质，受年龄、经济、技术素养、地域、教育水平和种族、社会制度等多重因素影响①，人机互动中隐私素养保护的差异更是另外一种新型“鸿沟”，需要基于人口统计学特征分层次分梯队地展开隐私的教育。

综上所述，本文试图从法律、技术、伦理视角回应当下人机交互中的隐私保护问题。现代风险社会中，由于自然科学的未知因素，每个社会主体不得不作出面对风险的决定，这意味着个人在行为时作出完全没有风险的决定和行为往往是不可能的。因此，应基于产业特征、创新成本和生产效率等多个维度综合权衡利弊，坚持包容审慎的原则，在新事物探索过程中允许适当的风险。人机传播带来的社会问题和产生的社会影响是多层面的，不仅涉及多学科交叉，也涉及多领域、多主体的互动融合。随着人际传播方式的快速扩散及广泛普及，隐私成为其绕不开的基础性、关键性议题，这背后是人与机器人权利的让渡②。未来的研究中，需要在遵循各项隐私政策基础上，综合平衡各项治理措施的利弊，探寻人机交互中隐私保护的最佳平衡点。

Privacy Risks and Countermeasures in Human Robot Interaction: in the Era of Large Models

GUO Qiaomin　ZHANG Linying　DONG Boyue　GUO Mingjun　YANG Boxu

[Abstract] The application of large model of artificial intelligence makes robots become more and more intelligent, and at the same time, with the transfer of individual privacy, this paper takes man-machine communication as the entry point. In the process of man-machine interaction, the risk of integrating robots into life to individual privacy is divided into three aspects. First, direct surveillance that collects information, Such as gender, age, and where an individual lives ETC. Second, with the abundance of data sets and the asymmetry between individuals and platforms in the “consent - use” principle, multi-dimensional sensitive information such as emotions and mental states is collected; Third, interactive risks lead to the blurring of human-machine boundaries, resulting in accompanying risks such as security risks, legislative risks, social risks, trust risks, etc., which is ultimately not conducive to the orderly development of cyberspace. Based on this, the author tries to propose a solution to the problem of privacy protection in human-computer interaction from the perspectives of policies, laws, technical regulations and ethical design, so as to balance the relationship between technological innovation, social well-being and personal privacy protection.

[Key words] Privacy; Human-Robot Interaction; Empowerment; Dispel Power.

（责任编辑：朱瑞　责任校对：蒙柯键）

① 薛伟贤、王涛峰：《“数字鸿沟”研究述评》，《科技进步与对策》2007年第1期。

② Jia L, Ruan L. Going Global: Comparing Chinese Mobile Applications’ Data and User Privacy Governance at Home and Abroad. *Internet Policy Review*, 2020, 9(3).