社会科学进展
2025 年 4 月第 7 卷第 4 期

生成式人工智能对个人信息的法律保护

李 瑾 马 飞

内蒙古工业大学，呼和浩特

摘 要 | 以ChatGPT为代表的生成式人工智能横空出世，已广泛应用于多领域和行业，推动了技术创新与突破。然而，生成式人工智能的运行机制依赖海量数据，这使得个人信息面临严重威胁。在生成式人工智能背景下，个人信息保护存在法律与监管滞后、知情同意规则流于形式、目的限制原则与最小必要原则形同虚设等挑战，亟需完善专业化立法、秉持包容审慎的监管理念、优化知情同意规则、贯彻目的限制原则和最小必要原则。

关键词 | 生成式人工智能；个人信息；个人信息保护

1 背景

生成式人工智能的历史最早可以追溯到20世纪50年代，“人工智能”一词是美国学者约翰·麦卡锡（John McCar-thy）在1956年提出，至今差不多发展70年。[1] 2022年年底，Open AI公司推出一款的智能聊天机器人模型：ChatGPT，这款生成式人工智能一经问世，迅速风靡全球，在发布仅一周用户就已超百万，一月后用户数量突破了1亿，成为人工智能界的“顶流”，是历史上增长最快的消费应用。2023年Open AI公司又发布的GPT-4，使得以ChatGPT为代表的生成式人工智能成为当年的全球科技热点，推动了教育、医疗、金融等多个领域的发展，对各行各业产生了深远影响。在2023年，我国以百度为首的互联网企业也发行了属于自己的生成式人工智能大模型，比如文小言（原文心一言）、讯飞星火、通义、豆包等生成式人工智能，推进了我国生成式人工智能的发展。

以ChatGPT为代表的生成式人工智能掀起了全球第四次科技革命浪潮，不仅推动数字经济快速增长，促进技术交叉创新与协同发展，更为全球经济注入新活力，成为带动全球经济增长的重要引擎。[2] 虽然生成式人工智能拓宽应用场景、提升审查效率，但同时也带来了许多挑战与风险，个人信息泄露和保护无疑是重要话题之一。例如，2023年3月20日，ChatGPT发生了一起重大信息安全事故，部分用户对话数据、个人敏感信息遭到泄露，Open AI公司在事故发生后公开承认了是开源库出现错误，表示已经修复了错误并完成验证。[3] 同年6月和9月Open AI公司因ChatGPT未经授权收集并泄露用户个人信息，两次遭到集体诉讼。出于对数据安全与个人信息保护等多方面的考量，意大利个人数据保护局（DPA）宣布从2023年3月31日起暂时禁止使用ChatGPT，同时展开隐私安全立案调查，这是全球第一道针对ChatGPT的政府禁令；美国、德国、法国、加拿大、西班牙、爱尔兰等多个国家纷纷开始加强对ChatGPT的监管，出台相关监管措施；2024年7月12日，欧盟正式公布了《人工智能法案》，旨在确保人工智能技术的安全性、可靠性、

通讯作者：马飞，内蒙古工业大学硕士研究生，研究方向：民商法。
文章引用：李瑾，马飞. 生成式人工智能对个人信息的法律保护 [J]. 社会科学进展，2025，7（4）：307-311.
https://doi.org/10.35534/pss.0704054

可追溯性，同时保护数据安全和个人信息安全，这是全球首部对人工智能领域进行监管的法律，填补了全球在人工智能监管领域空白，为其他国家和地区提供了重要参考。

2 生成式人工智能下个人信息的概述

2.1 个人信息的内涵

个人信息是生成式人工智能领域中个人信息法律保护的核心概念。个人信息的内涵在我国不同的法律中有不同的规定。2016年颁布的《网络安全法》第七十六条第五项规定个人信息的内涵，其主体主要是自然人个人身份的各种信息，采取识别与列举的方定义式。2020年通过的《民法典》第一千零三十四条规定了个人信息的内涵，主体能够单独或者与其他信息结合识别特定自然人的各种信息，也是采用了识别与列举的方定义式，但相较于《网络安全法》，对个人信息内涵进行了扩张，增加了“电子邮箱、健康信息、行踪信息”。2021出台的《个人信息保护法》第四条第一款规定，个人信息是指以电子或者其他方式记录的与已识别或者可识别的自然人有关的各种信息，不包括匿名化处理后的信息。与《网络安全法》《民法典》规定的内涵相比较，《个人信息保护法》未采取列举的方式规定，只是定义了何为个人信息，并新增了“不包括匿名化处理后的信息”。[4] 三部法律虽对个人信息内涵的界定各有侧重，但共同构建了个人信息保护的法律基础，体现了对个人信息保护的高度重视，尤其是在当今信息化时代，个人信息作为一种重要的“战略资源”，具有极高的商业和社会价值，需要采取多种措施加强个人信息保护，确保个人信息安全和合规使用。

2.2 生成式人工智能时代下个人信息的特征

2.2.1 个人信息收集的多样化和场景化

在生成式人工智能技术驱动下，个人信息收集呈现出显著的多样化和场景化的特点。生成式人工智能系统主要是通过大规模数据学习和训练来生成文本、视频、图片等内容的技术。为提高模型的准确性，针对用户需求，个人信息提供者和信息收集者会从多个渠道收集信息，包括网络爬虫、社交媒体平台、公开数据集等多元渠道，收集的信息包括但不限于个人一般信息，甚至由于算法黑箱的特性，个人敏感信息也会被收集，致使个人信息泄露风险加大。与此同时，生成式人工智能系统涉及各个领域，根据不同应用场景对个人信息需求的不同，制定个性化推荐。例如，在某个社交媒体上，生成式人工智能可以根据用户的浏览历史、购买记录等信息，推断用户的兴趣偏好，为用户推荐个性化的商品或服务。

2.2.2 个人信息的敏感性与隐私性增强

生成式人工智能系统开发的工作原理是模拟人脑神经网络，由大量互相连接的人工神经元组成，按不同层级结构排列构成。生成式人工智能模型越要发挥其功效越需要海量的数据，经过“投喂”海量数据进行模型训练，锻炼其自动生成用户需求的文本能力，因此，数据成为预训练模型的关键要素。[5] 生成式人工智能系统能够自主学习，深度挖掘出潜在的信息，原本这些关于个人的信息可能不会涉及敏感信息，但经过生成式人工智能系统的算法处理后，可能就会揭露出个人的身份、偏好、情感等敏感内容。随着生成式人工智能系统不断迭代升级，数字技术不断优化，个人敏感信息就会不断被收集，而个人却不知敏感信息已被收集，导致个人信息泄露风险加大，个人隐私被侵犯，甚至对个人造成严重的财产损失和精神损害。因此，保护个人信息的重要性日益凸显，为应对生成式人工智能带来的隐私挑战，需要不断出台保护个人信息的法律法规，完善隐私政策，提升隐私保护技术，加强个人信息保护意识。

2.2.3 个人信息的价值性和利用性提升

在生成式人工智能时代背景下，个人信息的价值性和利用性尤为凸显。个人信息具有人身专属性，且信息主体对自己的个人信息享有专属权，尤其是个人信息中的敏感信息，根据《个人信息保护法》第二十八条第一款规定，敏感信息包括但不限于生物识别、宗教信仰、行踪轨迹等信息，一旦被非法利用，会致使个人的人格尊严受到侵害。个人信息具有财产性，比如个人的交易记录、通讯记录、健康生理信息等信息具有巨大的潜在商业价值，会给相关企业带来巨大经济利益。与此同时，根据相关资料显示，2018年推出的GPT模型参数量高达1.17亿，2020年推出的GPT-3模型参数量已经达到了1750亿，据报道2023年推出的GPT-4参数量达到了约5000亿，这种规模的增长说明了生成式人工智能需要海量的数据，导致个人信息的收集和利用呈指数上升，生成式人工智能经过算法加工，进行分析和预测，制定更加精准的营销策略和产品服务，提供个性化服务和体验。

2.3 生成式人工智能时代下个人信息法律保护的必要性

2.3.1 保障个人信息安全

生成式人工智能在预训练和正式运行时，会持续不断地收集、分析、处理个人信息，这会使得个人信息随时处于非法使用或者泄露的风险之中，尤其是具有财产属性和敏感内容的信息更容易被泄露。因此，保障个人信息安全是生成式人工智能应用过程中不可忽视的问题，必须加强保障个人信息安全的措施，建立健全个人信息安全防护机制，有效防止个人信息被非法使用或者防范网络犯罪。

2.3.2 实现信息主体与处理主体的利益平衡

生成式人工智能产业的健康发展高度依赖海量数据的“投喂”。而个人信息是数据形成的基础，国家在推

动生成式人工智能产业发展时，必须平衡个人信息保护与信息处理主体之间的利益。一方面，为了保护个人信息，限制信息处理主体收集、分析数据，阻碍生成式人工智能产业发展；另一方面，大力扶持生成式人工智能产业发展，会造成个人信息严重泄露，甚至威胁到人身安全或者财产安全。因此，国家层面必须一面保障个人信息安全，一面推动生成式人工智能产业发展，两手都要抓、两手都要硬，在实现信息主体安全的基础上促使生成式人工智能产业发展。

2.3.3 维护社会秩序和稳定

生成式人工智能技术的普及加剧了个人信息泄露或者非法利用，会给社会秩序和稳定带来极大挑战，影响国家长治久安、和谐安定。例如，大规模信息泄露将削弱公众对生成式人工智能服务的不信任，阻碍数字经济发展；而非法利用个人信息从事网络犯罪活动则直接影响社会稳定。因此，强化个人信息保护既是防范违法犯罪、维护社会安全的必然要求，也是推动技术健康发展的重要保障。

3 生成式人工智能下个人信息面临的法律挑战

3.1 法律与监管滞后

面对以ChatGPT为代表的生成式人工智能带来的新挑战，使得现有的个人信息保护机制显得力有未逮。一方面，生成式人工智能领域个人信息保护立法在专业化有所欠缺。[6] 我国对于个人信息保护的法律有《民法典》《数据安全法》《网络安全法》《个人信息保护法》，但对于生成式人工智能中个人信息保护的法律法规，目前效力层级最高的是2023年7月13日公布的《生成式人工智能服务管理暂行办法》，属于部门规章，可见，对于生成式人工智能领域个人信息保护立法在专业化还不够完善。另一方面，我国对生成式人工智能的监管涉及多个部门相互配合，并非由一个部门牵头负责，这就会导致生成式人工智能发生个人信息侵权时，会产生监管主体责任竞合，加上监管职责界限不清晰，就会出现各部门相互推诿或者踢皮球的现象。

3.2 知情同意规则流于形式

知情同意规则是个人信息保护的基本准则。《民法典》第一千零三十五条确立同意原则，《个人信息保护法》第十三条第一款专门规定了个人信息处理者必须取得个人同意才能处理信息；第十四条确立“告知–知情–同意”的个人信息处理规则，处理个人信息应当在个人被告知且知情的情况下，经过个人同意，才能允许处理个人信息，并在处理目的、方式发生变化，要重新取得个人同意。生成式人工智能模型的语料库来源复杂，需要庞大的体量支撑，这就导致知情同意规则难以有效落实。首先，生成式人工智能大模型预训练时需要海量数据“投喂”，才能使得服务更加准确，再加上生成式人工智能自主性和黑箱性的特点，生成式人工智能在收集个人信息时，未能提前充分告知个人，并取得个人同意。其次，信息处理者在收集和处理个人信息时，如需取得每位用户的同意，需要付出巨大的经济成本，为规避责任和追逐利益，使得现有的知情同意规则日益僵化。最后，用户注册平台或者App时面对冗长复杂的用户协议与隐私政策，若不同意该协议，用户就无法获得所需的服务或者产品，被迫接受条款或者条件严重削弱了知情同意规则的法律效力。

3.3 目的限制原则与最小必要原则形同虚设

目的限制原则是和最小必要原则属于个人信息保护的基本原则。《个人信息保护法》第6条进行了规定，明确了目的限制原则与最小必要原则，要求处理个人信息时，必须先明确处理目的，其处理行为必须与处理目的直接相关，不得超出处理目的的范围，并要求采取影响最小、侵害最小的方式，采取的处理方式必须与处理目的适当，不应导致严重的后果。一方面，很难基于目的限制原则明确划分处理个人信息的范围及边界。当在生成式人工智能开展人机对话时，生成式人工智能自动就会储存所有对话内容，作为后续预训练的数据库，生成式人工智能无法识别超出其预定目的的范围达到信息内容，可能会造成个人信息过度收集、泄露，监管挑战等一系列问题。另一方面，对于收集个人信息没有统一标准，如何界定最小范围难度较大。因生成式人工智能收集信息越多、越全面，输出的信息等越准确，信息收集者就会通过多渠道、多途径收集个人信息，使得最小必要原则难以有效落实。因此，在生成式人工智能收集、利用、生成信息时，目的限制原则与最小必要原则形同虚设，必须要完善目的限制原则与最小必要原则，以避免个人信息脱离主体控制。

4 生成式人工智能下个人信息法律治理措施

4.1 完善专业化立法和包容审慎监管理念

完善生成式人工智能下个人信息保护立法专业化、多维度法律体系。尽管我国多部法律涉及生成式人工智能个人信息保护领域，但很明显这些立法重心不在生成式人工智能个人信息保护领域。当下亟需以《个人信息保护法》为基础性法律，加快制定相关生成式人工智能个人信息保护领域的行政法规、规章制度、司法解释，制定配套政策与措施，尤其是要进一步加强对敏感个人信息的法律保护，维护个人信息安全，减少个人信息被非法收集、利用、泄露的风险，为生成式人工智能的健康发展保驾护航。

深化包容审慎监管理念。包容审慎监管包含“包

https://doi.org/10.35534/pss.0704054

容性监管”和“审慎性监管”两大核心理念。一方面，实施包容性监管。给予生成式人工智能必要的发展时间和试错空间，建立个人信息分级监管模式，依据风险大小适时适度干预，不能监管过度，限制生成式人工智能的健康发展。另一方面，强化审慎性监管。形成以事前审查为主，事中监管为辅，事后评估和跟踪的监管模式，设置统一的监管标准，要求各部门分工协作、相互配合，形成监管合力，避免因监管不力相互推诿，做好“数字守门人”，最大程度消除监管死角和缩小监管盲区，优化监管体系。

4.2 完善知情同意规则

完善知情同意规则需从增强规则的有效性与可操作性入手。该规则分为“知情”和“同意”两项子规则，一方面，在处理个人信息时，除了根据《个人信息保护法》第十三条第二款到第七款情形除外，生成式人工智能服务提供者要明确告知信息主体处理目的、方式和范围，还应当告知生成式人工智能算法过程、算法风险等事项，采用简明扼要的方式告知信息主体，涉及个人敏感信息的，根据《个人信息保护法》第三十条规定，处理个人敏感信息必须告知处理的必要性和对个人权益的影响。另一方面，构建分层知情统一规则。对于已经去识别化处理的个人信息，能够减少间接识别风险，可以采取默示同意，不必取得信息主体的同意，对于其他的一般个人信息原则，应当取得个人的同意，但涉及个人敏感信息的，根据《个人信息保护法》第二十九条规定，必须取得个人的单独同意；《个人信息保护法》第三十一条规定，对于不满14周岁未成年人的个人信息，需经过监护人的同意。

4.3 贯彻目的限制原则和最小必要原则

第一，灵活适用目的限制原则。作为个人信息保护的核心原则，旨在确保个人信息处理活动的合法性、正当性和必要性。生成式人工智能服务提供者在进行数据发分析、算法预测、模型训练等活动时，服务提供者应当清晰、准确及完整地向个人说明其处理个人信息的目的，且处理方式与处理目的之间具有“直接关联性”，不得将个人信息用于无关用途。

第二，践行最小必要原则。遵循最小必要原则有助于保护个人信息权益，降低个人信息泄露和滥用的风险，也助于促进生成式人工智能技术的健康发展，信息收集者在收集个人信息时应当采取对个人影响最小的方式，处理个人信息时应当限于实现处理目的的最小限度内和最小范围内。国家要出台收集个人信息的最低限度标准和最小范围标准，避免过度收集个人信息，减少对信息主体隐私的干扰和侵犯。[7]

5 结语

随着生成式人工智能技术的不断优化升级，其在推动数字社会与数字经济发展的同时给个人信息保护带来严峻挑战。因此，为了实现网络技术创新与个人信息保护的动态平衡，需加快推进专门立法进程，加强包容审慎监管的落实，完善相关法律规则，切实履行“数字守门人”的职责。

参考文献

[1] 朱荣荣. 生成式人工智能应用中间接识别个人信息的法律保护[J]. 科技与法律（中英文），2024（4）：104-114.
[2] 麦肯锡. 生成式人工智能的经济潜力:下一波生产力浪潮[EB/OL]. （2023-06-14）[2024-01-10]. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-eco-nomic-potential-of-generative-ai-the-next-productivity-frontier #intro-duction.
[3] 斜晓东. 风险与控制：论生成式人工智能应用的个人信息保护[J]. 政法论丛，2023（4）：59-68.
[4] 张涛. 生成式人工智能中个人信息保护风险的类型化与合作规制[J]. 行政法学研究，2024（6）：47-59.
[5] 王东方. 生成式人工智能对个人信息权益的侵害风险及其法律规制[J]. 征信，2024，42（2）：31-37.
[6] 王大志，张挺. 风险、困境与对策：生成式人工智能带来的个人信息安全挑战与法律规制[J]. 昆明理工大学学报（社会科学版），2023，23（5）：8-17.
[7] 周莉. 学术期刊用户画像个人信息保护：风险与规制——以《个人信息保护法》为视角[J]. 武汉科技大学学报（社会科学版），2023，25（1）：95-99.

Generate Artificial Intelligence's Legal Protection of Personal Information

Li Jin Ma Fei

Inner Mongolia University of Technology, Hohhot

Abstract: The generated artificial intelligence represented by Chat GPT was born, and it was widely used in multi -field and industry to promote technological innovation and breakthroughs. In the context of generating artificial intelligence, the protection of personal information protection has the lagging law and regulation, the rules of informed consent flow in the form, the principles of purpose restrictions, and the minimum necessary principles. The principles of rules, the principle of implementing the purpose and the principle of minimum necessary necessary.

Key words: Generate artificial intelligence; Personal information; Personal information protection

https://doi.org/10.35534/pss.0704054