机器学习中已公开个人数据的合法利用路径

王婉清
（华东政法大学　中国法治战略研究院，上海　200042）

摘　要：已公开个人数据作为机器学习的重要训练语料，应对其秉持开放利用的目标取向。但采取宽松获取策略，却由于爬取范围不清晰、用于生成式 AI 存在侵权风险、个人数据主体难以行使信息自决权而面临合法利用的实践困境。检视困境成因，应围绕机器学习应用全周期，构建已公开个人数据的合法利用路径：在数据获取阶段，评估爬取行为的正当性与潜在影响，若涉及竞争性权益，应转向 API 授权等合法路径，确保数据来源合法；在机器学习智力成果投入应用阶段，应依据个人信息种类设置分类安全机制，并实时监督以防范隐私泄露与滥用风险；在应用投放市场后，应构建训练数据披露机制，以透明度支持用户干预，保障个人信息自决权的实现。

关键词：已公开个人数据；机器学习；竞争性权益；个人信息保护；信息披露
中图分类号：D922.17；TP181　　**文献标志码**：A　　**DOI**：10.19358/j.issn.2097－1788.2026.02.010

中文引用格式：王婉清. 机器学习中已公开个人数据的合法利用路径［J］. 网络安全与数据治理，2026，45(2)：73－80.
英文引用格式：Wang Wanqing. Legal use of publicly available personal data in machine learning［J］. Cyber Security and Data Governance，2026，45(2)：73－80.

Legal use of publicly available personal data in machine learning

Wang Wanqing
(China Institute for Rule of Law Strategy, East China University of Political Science and Law, Shanghai 200042, China)

Abstract: Publicly available personal data, as a crucial corpus for machine learning training, should in principle be governed by an orientation toward open utilization and more permissive acquisition strategies. However, practical challenges arise in lawful use due to ambiguities in the scope of web scraping, potential infringement risks in generative AI applications, and the difficulty for data subjects to exercise informational self-determination. To address the dilemma of lawful use, it is necessary to construct a legal utilization pathway for such data throughout the full machine learning cycle. During the data collection stage, the legitimacy and potential impact of web scraping should be assessed. If competitive interests are involved, access should shift to lawful channels such as API authorization to ensure data sources are legal. In the application stage of machine learning outputs, a classified security mechanism should be established based on the type of personal information, with real-time supervision to prevent privacy breaches and misuse. After deployment in the market, a data disclosure mechanism should be implemented to support user intervention through transparency and safeguard the right to personal information autonomy.

Key words: publicly available personal data; machine learning; competitive interests; personal data protection; information disclosure

0　引言

目前，我国的人工智能（Artificial Intelligence，AI）已经进入统筹安全与创新发展的新阶段[1]。人工智能系统多以机器学习（Machine Learning）为基础技术路径。例如，生成式 AI 的工作原理是基于海量数据学习总结规律，不断优化模型，依据操作者指令生成新的内容。而总结规律的过程便是机器学习环节[2]。机器学习利用数据和算法，通过模型训练学习、参数调优来逐步提高决策准确性[3]，最终形成预测、判断等信息智能，实现特定目标[4]。

在以数据为核心驱动的人工智能技术体系中，机器学习对训练数据的依赖性愈发显著。与传统软件开发的预设固定规则不同，机器学习通过对海量数据的自主学习来完成能力迁移与性能优化。因此，高质量语料成为影响模型效果的关键变量。而网络空间中的已公开个人数据因获取便利、信息密度高等特征，符

合生成式人工智能研发对训练语料的需求，因而被广泛采集并成为训练集的重要组成部分，用于支撑机器学习模型构建和优化，应用于用户个性化推荐、自然语言处理、人脸识别训练、金融风控与信用评估等场景。因此，在机器学习中如何高效规范地利用已公开个人数据，已成为人工智能发展和个人信息权益保护的重要课题。

1 问题的提出

在实践中，已公开个人数据用于机器学习场景面临合法性困境。个人数据之上承载着数据主体的人格权益，同时因大规模关联分析与智能计算而承载财产性权益。即使个人数据已经处于公开状态，其能否被自由获取以及获取后的进一步利用是否面临限制，仍然具有诸多争议。另外，在数据训练与系统开发的过程中，开发者往往认为用户对于其自行公开的个人数据放弃控制权，或者将匿名化或者去标识化处理后的已公开个人数据视为公共资源，而对于后续数据处理、训练与输出等阶段缺乏必要的限制，因此存在侵害个人数据主体人格权的风险。

由此可知，立足于促进人工智能创新研发的战略定位，需要解决如何实现已公开个人数据在机器学习场景下的合法利用问题，有效规制其中的法律风险。为此，有必要审视当下的制度需求与实践困境，厘清机器学习场景下已公开个人数据的合法利用逻辑与规则，构建机器学习全周期的合法利用路径，促进包括个人信息在内的数据要素有序流动，助力人工智能产业发展。

2 机器学习中已公开个人数据的利用需求与实践困境

2.1 机器学习中已公开个人数据的利用需求

在加快培育数据要素市场与推动人工智能创新发展的战略背景下，网络空间中广泛存在的大规模已公开个人数据成为机器学习技术实现的重要保障，可以在获取端适当采取宽松获取策略，原则上允许将其作为训练数据集，以满足人工智能研发对个人信息的利用需求。

2.1.1 制度需求：在输入端对已公开个人数据适当放宽限制

在数据已成为第五大生产要素的数字经济时代，国家发展战略高度重视数据资源的高效流通与利用，对个人数据的规范利用也作出部署。我国国家数据局等部门印发的《“数据要素×”三年行动计划（2024—2026年）》指出，提升数据供给水平，在保护个人隐私前提下促进个人信息合理利用。国家发展改革委等部门于2024年12月印发的《关于促进数据产业高质量发展的指导意见》也指出，支持在保护个人信息权益的前提下，加强个人数据开发利用。在2025年1月印发的《关于完善数据流通安全治理 更好促进数据要素市场化价值化的实施方案》中，提出强化个人数据流通保障，推动数据高质量发展和高水平安全良性互动。

个人数据的有效利用也为人工智能产业发展提供保障。当前人工智能系统已经由以模型为中心转向以数据为中心[5]，进一步推进数据的单纯工具属性转向关键生产要素属性[6]。已公开的个人数据是生成式人工智能的主要语料来源[7]，个人数据的法律保护和合法利用路径将从源头上影响人工智能产业的发展[8]。在此背景下，在输入端对已公开个人数据适当放宽限制，避免制约个人数据的开发利用，成为推动人工智能创新发展与数据要素市场培育的重要环节。

2.1.2 理论基础：将已公开个人数据用于机器学习落入合理处理范围

在当前规范体系下，已公开个人数据用于机器学习可视为落入合理范围，不会直接影响数据主体的权益。数据允许公众访问并不意味着可以被随意收集或使用[9]，《中华人民共和国个人信息保护法》（以下简称《个人信息保护法》）第27条规定了已公开个人信息处理的合法性基础，但由于表述模糊，仍存在较大解释空间。一般以公开方式为分类依据，如果来源于个人自愿公开，则以推定同意为处理的合法性基础；如果来源于依法强制公开，则以目的一致为合法性基础[10]。

一方面，若个人自愿公开其个人信息，代表同意他人在可预期的风险范围之内作出处理，自愿承担可能存在的侵权风险，因此无需再次获取同意，将其个人信息用于机器学习可视为合理范围内的处理行为。公开行为将会使得个人信息处于难以控制、便于获取的状态，主动将其暴露于较高风险之中，作为理性的自然人，应当意识到此类后果，并且谨慎地实施其个人信息的公开行为。因此，在风险可控范围内，可以直接处理个人自愿公开的个人信息[11]。而机器学习的目的在于挖掘海量数据中的统计学规律，而非识别个人信息的潜在联系[12]，因而已公开个人数据用于机器学习的主要风险在于泄露，但公开数据的泄露并不必然造成危害后果；同时，若开发者严格落实个人信息安全保护措施，可进一步降低个人数据泄露的潜在危

害。在个人数据被采集后，通常会对其采取脱敏、截断、归一化等技术手段，模型只提取数据的语义、向量、统计结构，极大降低了将数据与特定个人关联的可能性[13]。

另一方面，强制公开行为通常出于实现公共利益的目的，体现了个人信息权益与公共利益之间的平衡。如果后续的处理目的与公开目的一致，则可以延续合法性[14]。而将已公开的个人数据用于机器学习中是否契合初始的公开目的，如新闻报道、舆论监督等，似乎并不能确定，甚至可能有所偏离。但是从促进产业发展的视角出发，此处应当进行宽松解释。机器学习的智力成果为人工智能大模型打下基础，通用大模型可服务于多元化的公共目标，间接契合了个人数据的公开目的，而垂直大模型也通常是在通用大模型的基础上，进一步利用大量特定领域的数据集、针对特定任务进行优化，因此亦满足目的一致性。

2.2 机器学习中已公开个人数据的利用困境及反思

2.2.1 已公开个人数据的获取界限不明确

企业间的数据流通往往涉及竞争法效果，已公开个人数据的流通利用亦如此[15]。已经拥有一定数据资源的企业往往抗拒其他企业随意获取其控制的公开数据以攫取其竞争优势，而其他企业则希望能够无限制获取公开数据以满足利用需求。反不正当竞争法逐渐成为解决企业数据权益纠纷的主流司法路径，竞争性权益也逐渐成为裁判关键要点。如果企业对于其控制的个人数据投入了治理成本，产生了潜在的竞争优势和交易可能，使其具备的市场竞争性利益价值应当获得法律保护，过度爬取将侵犯已有的数据竞争性权益[16]。

以腾讯与北京联云天下科技有限公司等不正当竞争纠纷案为例[17]，法院认为，海量的聊天记录在微信平台集成之后衍生出新的价值，即使对单个数据不享有权利，但基于平台的个人信息保护义务，以及对微信平台的投入和运营，原告对其合法持有的上述数据集合，享有数据资源基础上的竞争利益。被告收集、使用包括其用户与其聊天相对方的微信相关数据的方式有悖商业道德，且损害了相关微信用户的信息安全、原告关于微信用户数据的安全保障利益等竞争利益，扰乱了相关市场秩序，构成不正当竞争。

2.2.2 机器学习智力成果用于生成式 AI 存在侵权风险

在机器学习场景下，大规模已公开个人数据经训练后，可能会通过语义关联等重新指向具体个体，输出个人信息，进而构成对隐私权、名誉权等人格权的侵害。以 OpenAI 的 ChatGPT 模型争议为例[18]，有用户反馈模型在特定提示下能生成或“回忆”与其真实经历高度相似的文本内容，或复现过往公开发布的博客与社交媒体发言，导致隐私泄露。因此，需要遵循个体的合理预期原则[15]，即考虑数据主体对于保障隐私权等权利的合理预期，以平衡好数据利用与权益保护[19]。在机器学习的过程中，对大规模个人数据的收集和分析将形成对个人画像，产生累计效应，引发隐私权侵权风险[20]。同时，一旦机器学习训练数据集中纳入人脸信息等生物特征，即使该类信息已经向公众开放，仍然应当避免直接爬取并用于训练，否则可能引发对肖像权及个人信息权益的侵害风险。

2.2.3 个人数据主体难以行使信息自决权

目前，《个人信息保护法》第 13 条与第 27 条初步构成了已公开个人信息处理的规范框架。其中，第 13 条确立了已公开个人信息的合法处理基础，第 27 条则细化了处理要求，延续《民法典》第 1036 条要义，进一步明确，即便是已公开的个人信息，只要信息主体明确拒绝，处理者仍需事先取得其明确同意，不能随意将其用作机器学习训练语料。

现实中，对于将已公开个人数据作为训练语料的处理行为，数据主体往往难以明确拒绝，这便暴露出现行规则难以有效保障数据主体个人信息权益的困境。AI 高速发展伴随着海量的数据训练，对于机器学习处理已公开个人数据的行为，若其落入“合理范围”内，且处在对个人权益不产生重大影响的前提下，那么个人明确拒绝将会成为排除其合法性的原因。而现实中，个人信息的公开方式多种多样，主要分为个人自行公开与其他方式公开，包括来源于政府公开信息、合法新闻报道等[21]。这就使得数据主体很难掌握自身公开过的个人信息范围，包括其具体处理情况，并不了解其是否被用于机器学习中，缺乏行使个人信息拒绝权的有效条件，导致数据主体即便自始不希望其个人信息作为训练语料，或者在 AI 应用中存在担忧，但是因不明晰其个人信息流动链条，而无法明确拒绝。

例如，2025 年 7 月，有网友称被微信“AI 搜索”功能生成“个人简历”，并呈现涉及该姓名的相关公众号推文，而后投诉未果[22]。尽管腾讯方面回应称，AI 搜索仅利用了公众号及其他互联网渠道的公开信息[23]，然而，对于涉及用户姓名的公众号推文等处理行为依然可能引发用户担忧——用户不希望其个人信息作为训练语料，或者在后续 AI 应用中希望能够明确拒绝对其个人信息的处理行为。由此可见，即使对于

已公开个人信息，也并不能无限制地用于机器学习训练，需要考量主体个人信息自决权的保障实现。

因此，面对规模化、多元化应用的机器学习技术，已公开个人数据处理的基础规范与实践产生了冲突，不能有效契合人工智能价值链上数据来源者的治理利益诉求，也对实现“保护个人信息权益”的立法目的带来了阻碍。为形成可信赖的人工智能规范体系，需要完善数据来源者权益的保障路径，主要为个人信息权益。在此基础上，还应配套个人信息匿名化技术等处理规则，以对冲训练数据的不可控性[24]。

3 机器学习中已公开个人数据的合法利用路径构建

机器学习中已公开个人数据的合法利用路径可以纵向展开：在获取阶段，应关注是否存在竞争性权益，确保数据来源合法；在智力成果投入应用阶段，应在输出端进行安全过滤，并根据个人信息类型设置分类分级安全机制，实行实时监管，防范隐私泄露与滥用风险；在应用投放市场后，为保障数据主体个人信息权益，应以训练披露制度与用户干预实现技术创新与权利保障的动态平衡，促进人工智能健康有序发展。

3.1 获取边界：保障已公开个人数据的竞争性权益

3.1.1 已公开个人数据竞争性权益的理论证成

对企业控制的已公开个人数据认定竞争性权益具备可行性。目前，数据保护类司法案例中，普遍以《中华人民共和国反不正当竞争法》（以下简称《反不正当竞争法》）以及《网络反不正当竞争暂行规定》等作为法律依据。2024 年 9 月 1 日起施行的《网络反不正当竞争暂行规定》第 19 条明确提出了数据不正当竞争条款，从反不正当竞争角度对数据获取提出要求。2025 年 1 月 1 日起施行的《网络数据安全管理条例》第 18 条也规定，爬取数据时应评估对网络服务造成的影响。在市场竞争的规范视角下，评估已公开个人数据的利用行为是否正当，需要结合竞争者、消费者、其他市场参与者的多重利益考量，因此反不正当竞争法是契合数据开发利用链上多主体利益诉求的较优路径[25]。

对企业控制的已公开个人数据认定竞争性权益具备合理性与必要性。有学者提出，爬取数据用于机器学习对被爬取经营者的影响较小，不会抢夺其他经营者的市场利益，而人工智能技术只不过使得分析数据的规模更大、效率更高而已，本质上依旧相当于人类分析数据的行为[26]。

然而，企业对其在网络空间中所采集的个人数据进行改进和生产，付出了足够的成本后，应享受合理的竞争性权益。从数据权益归属的角度来看，尽管个人数据主体拥有其数据所有权，如果企业对于个人数据的开发利用付出了足够的劳动以及其他成本，例如在对已公开的原始数据进行汇集编排、处理、加工并最终形成衍生数据的过程中，付出了时间、劳动力、金钱等成本，企业针对衍生数据也应享有相应权益。这导致数据权益群中的部分权益归属到企业端[27]。

另外，企业对个人数据享有竞争性权益，意味着在数据遭受不正当获取或者滥用时，企业有权请求合理赔偿。企业个人数据竞争性权益的客体应当限于包含用户或者其他主体的个人信息的数据，该权利本质上源于用户对个人数据的让渡。因此，此种权益保护路径能够激励企业对收集的已公开个人数据进行开发利用和加工处理，并对于企业付出的成本给予保障。如果仅仅出于促进流通的考量而不设任何限制、允许随意获取他人控制的衍生公开数据集作为训练语料，反而将打击企业对数据开发利用的积极性，在一定程度上阻碍个人数据的流通利用。因此，在当前数据产权制度尚未完善的背景下，应允许企业在达到合理条件时获得个人数据的竞争性权益，但是认定过程需要始终秉持审慎态度，避免数据垄断等恶意行为的出现。

3.1.2 已公开个人数据竞争性权益的认定路径

从司法层面上认定企业对于个人数据享有竞争性权益，一般从以下几个方面出发：第一，企业处理和利用个人数据的行为是否合法。为了保护个人信息权益，需要防止在个人数据利用过程中识别个人身份，即直接将侵权后果联系到数据主体，在保障数据流通的基础上应将数据识别分析的风险最小化[13]。第二，在满足合法性基础上，企业是否事实上控制了数据。第三，关注企业对数据的资本积累与成本投入。第四，数据能否给数据权益人带来流量和竞争优势。经营者之间即使经营模式有差异，但面向的目标用户群体相同，也可初步认定经营者之间存在竞争关系。第五，对于具体的数据权益归属，网络服务协议中的约定可以提供一定的参考意义，例如用户是否自愿将数据开发利用权能让渡给经营者。

在爬取已公开个人数据用于机器学习的过程中，关注企业在针对其控制的个人数据的经营过程中能否获得竞争性权益，一方面能够避免侵犯其合法权益，保障获取合法；另一方面聚焦于自身已控制的数据资源，如若能够认定获得竞争性权益，则可基于此在损害发生时请求赔偿，从而有效规范市场竞争秩序，促

进数据流通行业健康有序发展。

3.1.3 已公开个人数据合法获取的方式

为规避可能侵犯竞争性权益的风险，企业在采集已公开的个人数据进行模型训练时，应遵循合规要求，综合考虑数据抓取方和被抓取方间是否具有竞争关系、被抓取方是否对抓取的数据享有竞争性权益、抓取方的行为是否具有正当性、抓取方对抓取数据的使用是否具有正当性、是否给被抓取方带来相应的危害结果等因素。在爬取过程中采用频率控制、调度分布等方式防止对原本信息服务造成干扰，保留采集日志以备审查，同时在数据预处理阶段应执行去标识化等处理方式，确保模型参数无法映射至具体平台结构或用户行为。具体而言，对于对特定平台数据依赖度高的模型开发任务，可以优先考虑通过 API 授权、商业数据合作等方式合法获取数据资源，构建共赢式数据生态。

3.2 内容安全：构建生成式人工智能的个人信息保护机制

个人数据处于公开状态，仍涉及数据主体人格权益保障问题。机器学习产生的智力成果多应用于生成式人工智能，提供模型训练和优化支持，应采取必要措施防止生成侵犯他人个人信息权益的信息。除却有必要，尽量避免将敏感个人数据用作训练语料。

3.2.1 模型部署前：已公开个人数据的分类安全策略

生成式人工智能的生成模型，其训练和优化需要海量的数据做支撑，这就意味着收集的个人数据可能被用于模型训练和优化，将有可能在与用户交互过程中被输出，存在着个人信息或者隐私泄露的风险。泄露原因主要分为两种类型，一是生成式人工智能本身的系统设计所导致的个人信息泄露，例如抓取的个人信息被泄漏至生成内容中；二是在交互过程中，用户不慎输入个人信息或隐私信息，生成式人工智能收集后可能在其他交互过程中将其泄漏，或是因安全防护措施不到位而遭受恶意攻击导致泄露。因此，企业应在数据预处理阶段采用充分的脱敏技术，并且履行内容安全控制义务，避免在推荐系统、生成式人工智能等任务中输出可感知身份特征的结果，禁止关联输出个人信息。

已公开个人数据中不乏生物特征等敏感个人信息，对此不宜适用宽松获取策略，应当优先进行保护。依据《个人信息保护法》第 6 条、第 28 条，敏感个人信息的处理需要遵循三方面的实质性要求：特定目的性、充分必要性、落实严格保护措施[28]。作为高度关系到自然人的人格尊严、人身和财产安全的个人信息类型，若无充分的必要性，应避免将敏感个人信息数据用于机器学习。

因此，机器学习系统开发者和后续部署者应当对收集的已公开个人数据构建分类安全策略，并在内容输出前进行事先审查，避免模型输出用户的个人信息甚至是敏感个人信息。第一，对于公众人物或者社会知名人士的个人数据，可适用更为宽松的利用规则，赋予开发者较高的处理自由。第二，对于一般的已公开个人数据，在内容输出前进行事先审查，避免模型过度关联并输出用户的个人信息。第三，对于生物特征等敏感个人数据，即便处于公开状态，也应尽量避免用作机器学习训练语料，确有必要用于医学、金融等特定领域的垂直大模型训练时，应依法告知数据主体并取得其单独同意，并评估相关风险与可能造成的影响。例如，在未取得被编辑个人同意的情况下，不得克隆其公开于社交平台中的声纹信息并用于语音合成的数据训练中。其中，对于未成年人个人信息，即使处于公开状态，也应当履行特殊保护要求，并征得监护人的单独同意[26]。并且需要防范深度合成等技术对其进行篡改、拼接甚至是伪造。

3.2.2 模型部署后：内容安全的监督处置义务

在模型产品化、商业化的后续阶段，服务提供者应当对产品中是否存在违法违规内容进行实时监督和处置。根据《生成式人工智能服务管理暂行办法》第 14 条，生成式人工智能服务研发人员在选择训练数据集、设计算法、优化模型的过程中，应采取必要措施防止生成违法内容。模型部署上线后，服务提供者一旦发现生成内容存在侵权行为，应立即采取更正、屏蔽、删除等措施，后续对生成式人工智能模型进行内容过滤策略调整或模型优化训练。如果用户企图利用生成式人工智能模型获取他人的个人信息，应及时进行提示与更正，对达到恶意用户应限制其使用服务[10]。例如，对用户输入信息进行分词处理并检测关键词，若用户短时间内连续输入违法不良信息或明显诱导生成违法不良信息的，应采取暂停提供服务等处置措施，具体策略需视实际情况适当调整。

同时，也应对个人信息异常流动或者外挂现象保持实时监测预警，若发现异常现象，需立即启动预警程序并采取必要措施防止信息大规模泄露。如发生个人信息泄露，应当立即通知信息主体，以减少对信息主体的不利影响，同时应当及时向监管机构报告信息泄露的原因、范围、所造成的影响以及已采取的具体处置措施，如漏洞修补方案、补丁更新日程等。最后，

信息泄露的预警与处置机制需持续更新，例如根据不断变化的攻击手段储备应对能力，根据违规案例更新安全过滤词库、调整预警和处置策略。此外，还可以结合其他技术落实监督机制，例如利用区块链技术来建立不能篡改的数据记录系统，便于追溯来源和核查成因[29]。

3.3 权益保障：以信息披露与用户干预实现个人信息自决权

放宽对已公开个人数据的获取将会给个人信息带来一定的风险，除了输出端的安全控制之外，还需要在机器学习成果投入应用之后采取措施，保障个人信息高效安全利用。《个人信息保护法》第27条规定了“个人明确拒绝的除外”，保障了个人信息主体对于已公开个人信息的自决权[30]，也体现了与撤回同意所一致的个人信息权益保障理念[31]。服务提供者在规范处理已公开个人数据的同时，应当提供拒绝路径并依法处理拒绝请求，但是个人信息保护请求权的实现不可避免地受到技术现实的限制。因此，有必要针对机器学习训练语料构建合理的披露机制，保障数据主体的知情权，并以清晰便捷的退出机制降低个人行使权利的门槛。

3.3.1 人工智能训练数据披露制度的比较法分析

聚焦生成式人工智能的数据风险问题，各国都开始意识到对其进行监管的必要性，并加大审查与监管力度，力求在控制生成式人工智能风险的同时助力生成式人工智能产业健康发展，重视用户数据主体权利保障。

欧盟以《人工智能法案》（AI Act）确立AI模型训练数据的披露规则，包括披露范围、豁免情形等，而后又以《通用人工智能行为准则》（General-Purpose AI Code of Practice）向AI企业提供了实践指引。2025年7月24日，欧盟委员会发布《通用人工智能模型训练内容公共摘要说明与模板》（Explanatory Notice and Template for the Public Summary of Training Content for General-Purpose AI Models），作为通用人工智能模型训练内容披露的官方实施指引，模板中包括模型名称等基本信息、公开数据集等详细的训练数据来源类型、安全过滤与自决权保障等相关数据处理机制。

披露机制的完善体现了欧盟在技术进步与可信治理之间的平衡[32]。一方面，披露AI模型训练集中公开数据的来源类型，能够消除个人信息保护请求权的实现所受到的现实限制；另一方面，对于安全过滤与自决权保障等相关数据处理机制的披露，也进一步反向促进服务提供者自觉落实安全义务，完善个人信息权益保障。目前，欧盟AI办公室负责监督落实情况，违规者将被处以最高为全球年营业额3%或1 500万欧元的行政罚款，这亦彰显了信息披露在人工智能治理效能上的优化作用，披露机制也将进一步成为监管合规的关键要件。

美国国会研究处于2023年5月发布报告《生成式人工智能和数据隐私：初探》（Generative Artificial Intelligence and Data Privacy：A Primer），分析了生成式人工智能相关的数据利用问题，并提出了建立通知、披露机制与删除机制等对应解决方案，可以要求开发者在收集或使用个人数据之前获取数据主体的同意，并且告知其个人数据的使用目的，数据主体有权要求对其数据进行删除或者选择存储最短期限[33]。州级立法也在不断推进。加州《生成式人工智能训练数据透明度法案》（Generative Artificial Intelligence：Training Data Transparency）将于2026年1月1日起生效，其中以列举方式详细规定了开发者需要在其网站中披露的训练数据文档内容，包括：数据来源、训练目的、采集方式、是否包含特定数据（公共数据、个人信息、版权作品与合成数据等）、数据集的首次使用时间以及是否对数据集进行处理或修改等。此外，该法案提出了面向国家安全领域或者AI安全领域的披露豁免条款，以增强该披露机制应对实际应用场景的能力，进一步体现了立法灵活性。

与欧盟不同，美国的立法体系中尚未确立严格的AI训练数据披露义务规范，目前正处于立法推进过程中，体现出了对AI训练数据透明度的积极治理态度以及鼓励创新的立法价值取向。其中，面向公众，在网站中披露训练数据文档，旨在说明其训练数据利用情况，促进透明度义务的落实；面向用户，通过通知、披露、删除等一系列机制构造出完整的个人数据披露流程，有效保障数据主体的信息隐私权益。

3.3.2 我国已公开个人数据披露规则的具体构建

在机器学习中利用已公开个人数据的过程中，建立有效的信息披露制度，是保障数据主体控制权、尊重其个人信息权益、预防潜在侵权风险的关键环节[34]。虽然数据已处于公开状态，但在用于机器学习时，其处理目的、范围与方式已经发生转变，初次披露的场景可能已不再适用。通过适度的信息披露，可以回应公众对算法透明性与个人信息权益保护的关切，为个人数据主体行使拒绝权提供条件。企业在构建或部署依赖已公开数据的机器学习模型时，应披露与数

据处理相关的基本信息，包括数据采集来源、利用目的、处理方式、是否进行脱敏、是否存在对外共享或输出风险等内容。在不泄露商业秘密或核心算法逻辑的前提下，保障公众对数据使用的基本知情权。

信息披露制度应体现差异化原则，以面向对象与应用场景为区分标准，采取不同披露方式。其一，向公众公开个人数据利用信息。对于使用范围广、社会关注度高的模型，应在官网、产品介绍页或用户协议中设立专门章节，概述企业在模型训练中使用已公开个人数据的基本原则与措施，回应公众对技术伦理与数据利用合法性的普遍关注。其二，向用户告知个人数据处理方式。当使用的数据与具体用户之间存在潜在关联性时，应通过用户协议、隐私政策、弹窗提示等方式，清晰告知用户其数据可能被用于训练、测试等场景，包括是否采取脱敏、能否识别个人身份、是否设置退出机制等信息。其三，向监管部门提供个人数据来源记录。对于训练数据来源于第三方平台、接口或网络爬虫的，应保留完整的数据来源记录、获取方式说明与授权凭证，接受监管部门的抽查与合规性审查，体现可溯源、可核查的治理逻辑。

3.3.3 数据主体拒绝权保障机制的落实方案

为保障个人数据主体拒绝权的有效实现，企业应建立便捷有效的投诉反馈机制。例如，在机器学习智力成果应用于生成式人工智能时，应当在服务使用界面提供明确的投诉反馈入口，面对生成隐私侵权信息或者存在侵害个人信息权益的其他行为，或者发现存在滥用公开数据或误用敏感信息的迹象，用户可以及时反馈给服务提供者或开发者，并且可基于正当理由行使其拒绝权，要求终止对其个人数据的处理[35]。对此，服务提供者或开发者应设置专人受理，采取必要措施避免侵权范围进一步扩大，并及时进行反馈，借助平台的信息处理能力提升权益保障效率。在收到数据主体关于终止处理其个人数据的通知后，如果具备采取必要应对措施的条件却未实施，应视为侵犯其个人信息权益[36]。

4 结束语

已公开个人数据已成为机器学习领域技术研发和业务创新的重要燃料，随着通用大模型的不断发展及广泛应用，如何实现个人信息保护和新技术新业态创新之间的平衡，显得尤为迫切。促进开发利用并不意味着无限制利用，模型开发者不明确已公开个人数据的爬取界限，后续投入生成式人工智能将面临内容安全问题，个人数据主体担忧自身权益受到侵害，造成当前的合法利用困境。因此，需要结合机器学习的技术特征和产业需求，构建符合生成式人工智能客观发展规律的合法利用路径。在采集过程中，需要明确对于个人数据竞争性权益的认定逻辑，避免涉及侵害数据权益的不正当竞争行为，以 API 授权、商业合作等方式激励企业对于数据的开发创新。在后续利用过程中，需要关注输出端的隐私泄露等内容安全问题，采取有效措施防范语义关联等技术特性所引发的个人信息泄露等风险，尽量避免将敏感个人信息用作训练语料。同时建立适当的训练数据披露制度，并且为个人行使拒绝权提供便利条件和及时处置，回应公众对人工智能透明度与个人信息权益保护的需求，推动包含已公开个人信息在内的数据要素的流通利用，促进新技术新业务发展。

参考文献

[1] 刘宪娟，薛念文. 习近平关于人工智能重要论述的生成逻辑、核心要义及实践指向 [J]. 学术探索，2025（7）：32－38.

[2] 詹爱岚，田一农. 生成式人工智能机器学习中的著作权风险及其化解路径 [J]. 电子知识产权，2023（11）：4－14.

[3] 汪荣贵，杨娟，薛丽霞. 机器学习及其应用 [M]. 北京：机械工业出版社，2019.

[4] 付文博，孙涛，梁藉，等. 深度学习原理及应用综述 [J]. 计算机科学，2018，45（S1）：11－15，40.

[5] WHANG S E，ROH Y，SONG H，et al. Data collection and quality challenges in deep learning：a data-centric AI perspective [J]. The VLDB Journal，2023，32（4）：791－813.

[6] 刘艳红. 数据要素全生命周期安全风险的刑事保障制度研究——以数字经济安全法益观为视角 [J]. 法学论坛，2024，39（1）：39－50.

[7] FABRE B. Generative AI is scraping your data. So，now what? [EB/OL]. （2023－08－21）[2025－07－15]. https：//www. darkreading. com/vulnerabilities-threats/generative-ai-is-scraping-your-data-so-now-what.

[8] 汪庆华. 人工智能的法律规制路径：一个框架性讨论 [J]. 现代法学，2019，41（2）：54－63.

[9] Office of the Privacy Commissioner of Canada. Principles for responsible，trustworthy and privacy-protective generative AI technologies [EB/OL]. （2023－12－07）[2025－07－15]. https：//www. priv. gc. ca/en/privacy-topics/technology/artificial-intelligence/gd_principles_ai/.

[10] 张新宝. 生成式人工智能训练语料的个人信息保护研究 [J]. 中国法学，2024（6）：86－107.

[11] 程啸. 论公开的个人信息处理的法律规制 [J]. 中国法学，2022（3）：82－101.

[12] MANHEIM K，KAPLAN L. Artificial Intelligence：risks to pri-

vacy and democracy［J］. The Yale Journal of Law & Technology，2019，21：120－121.
［13］高富平. 个人信息流通利用的制度基础——以信息识别性为视角［J］. 环球法律评论，2022，44（1）：84－99.
［14］张新宝，昌雨莎. 已公开裁判文书中个人信息的保护与合理利用［J］. 华东政法大学学报，2022，25（3）：6－21.
［15］丁道勤. 已公开个人信息处理规则及其重构［J］. 政法论坛，2025，43（2）：47－57.
［16］张冰. 企业数据权益保护的司法路径分析——以反不正当竞争法适用现状为切入点［C］//《智慧法治》集刊2023年第2卷——城市数字化转型的法治保障研究文集. 上海市嘉定区人民法院，2024：176－186.
［17］北京政法网. 北京高院发布2023年度知识产权司法保护十大案例和商标授权确权司法保护十大案例［EB/OL］.（2024－04－26）［2025－07－26］. https：//www. bj148. org/sytjlb/202404/t20240426_1664446. html.
［18］王俊，冯恋阁. 窃取个人数据？OpenAI遭集体诉讼！［EB/OL］.（2023－06－30）［2025－07－15］. https：//www. stcn. com/article/detail/905689. html.
［19］谢琳. 大数据时代个人信息使用的合法利益豁免［J］. 政法论坛，2019，37（1）：74－84.
［20］纪庆全. “合理隐私期待”标准及其对中国的借鉴意义［J］. 西部法学评论，2021（5）：118－132.
［21］赵祖斌. 已公开个人信息在合理范围内处理的规范逻辑与实践因应［J］. 行政法学研究，2025（1）：146－159.
［22］PChome电脑之家. 网友称被微信AI搜索“开盒”：可根据名字查到个人资料［EB/OL］.（2025－07－02）［2025－07－26］. https：//news. qq. com/rain/a/20250702A049PP00.
［23］新浪科技. 腾讯回应微信AI搜索争议：仅整合公开信息，不触碰用户隐私［EB/OL］.（2025－07－03）［2025－07－26］. https：//finance. sina. com. cn/tech/roll/2025－07－03/doc-infecyit3718962. shtml.
［24］姚佳. 人工智能的训练数据制度——以“智能涌现”为观察视角［J］. 贵州社会科学，2024（2）：51－57.
［25］周樨平. 大数据时代企业数据权益保护论［J］. 法学，2022（5）：159－175.
［26］黄武双，谭宇航. 机器学习所涉数据保护合理边界的厘定［J］. 南昌大学学报（人文社会科学版），2019，50（2）：44－52.
［27］李园园. 数据抓取行为不正当竞争司法认定规则探究［J］. 黑龙江社会科学，2024（3）：59－67.
［28］杨尚东. 论大数据时代政府信息公开中的敏感个人信息保护［J］. 行政法学研究，2025（4）：101－114.
［29］JEREMIAS P A . Regulating algorithms at work：lessons for a ‘European approach to artificial intelligence’［J］. European Labour Law Journal，2022，13（1）：30－50.
［30］萧鑫. 个人信息拒绝权的界定与适用［J］. 社会科学研究，2023（2）：74－85.
［31］解正山. 论已公开个人信息的“合理处理”［J］. 学习与探索，2022（9）：69－78.
［32］KUSCHE I. Possible harms of artificial intelligence and the EU AI act：fundamental rights and risk［J］. Journal of Risk Research，2024：1－14.
［33］HARRIS L，ZHU L. Generative artificial intelligence and data privacy：a primer［EB/OL］.（2023－08－21）［2025－07－25］. https：//www. congress. gov/crs-product/R47569.
［34］高富平，张启航. 可信AI：人工智能法律治理的内在逻辑与实现路径［J/OL］. 学术探索，1－11［2025－08－01］. https：//link. cnki. net/urlid/53. 1148. C. 20250709. 1742. 004.
［35］刘辉，雷崎山. 生成式人工智能的数据风险及其法律规制［J］. 重庆邮电大学学报（社会科学版），2024，36（4）：40－51.
［36］梅傲，张优涵. 已公开个人信息合理处理的实践困境及因应之策［J］. 西华大学学报（哲学社会科学版），2024，43（6）：51－61.

（收稿日期：2025－08－01）

作者简介：

王婉清（2001－），女，硕士研究生，主要研究方向：数据法。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部