Frontiers of Humanities and Social Sciences

Axon Connecting Ideas to Impact.

生成式人工智能训练信息的侵权认定不足与完善路径

黄琛

福建师范大学 福建福州 350108

摘　要： 当前生成式人工智能技术快速发展迭代，生成式人工智能的逻辑在于通过自身固有程序进行数据抓取并训练学习从而根据使用者指令生成相应的内容，当前对于生成式人工智能训练信息规范存在边界相对模糊、对不同性质数据主体保护未作区分、算法黑箱导致因果关系与过错难以证明而面临适用障碍等问题，由此影响生成式人工智能训练信息的侵权认定。为此应构建以“合理使用例外”为核心的数据抓取边界规则，实施针对个人与非个人数据的分级同意与管理体系，并通过算法影响评估与透明度报告制度破解黑箱难题，以期在促进技术创新与保护各方权益间实现平衡。

关键字： 生成式人工智能；训练信息；侵权认定；合理使用；个人信息保护

Abstract: Currently, generative artificial intelligence (AI) technology is undergoing rapid development and iteration. The underlying logic of generative AI lies in using its inherent programs to capture data, conduct training and learning, and thereby generate corresponding content based on user instructions. The current regulatory framework for generative AI training information faces several issues, including relatively blurred boundaries, lack of differentiation in protecting data subjects of varying natures, and application obstacles due to the difficulty in proving causation and fault resulting from the algorithmic "black box." These issues consequently impact the infringement determination of generative AI training information. Therefore, it is essential to establish boundary rules for data capture centered on "fair use exceptions," implement a tiered consent and management system for personal and non-personal data, and address the black box challenge through algorithmic impact assessments and transparency reporting mechanisms, aiming to achieve a balance between fostering technological innovation and protecting the rights and interests of all parties involved.

Keywords: Generative Artificial Intelligence; Training Information; Infringement Determination; Fair Use; Personal Information Protection

1. 问题的提出

生成式人工智能，即利用自身程序，当使用者发出指令后其通过运行内部程序自行搜集需要的数据进而得出相应的结论的人工智能程序，这也被理解为一种自动化决策。生成式人工智能是一种通过学习大规模数据进而生成新的原创内容的新型人工智能。生成式人工智能需要依靠算法设计者设计的程序语言以及大量基于算法程序的数据训练来实现运行，其中就涉及到需要选取特定的用于训练的数据信息，包括公开信息与未公开信息，自然人相关信息与非自然人相关信息等，现行生成式人工智能相关规范未有对相关训练信息内容进行合适的区分、程序自身透明度也相对缺乏，进而影响了相关信息主体的权益保护，而通过探索合适的生成式人工智能训练信息的侵权认定进而规范生成式人工智能训练信息搜集就显得尤为迫切和必要。

2. 生成式人工智能训练信息的规制现状

生成式人工智能发展目前仍然处在早期阶段，存在安全方面的不足和个人信息泄露风险。相较于过去的人工智能技术，生成式人工智能可以输出特定内容，输出端的个人信息风险是其特殊性之所在[1]。对于生成式人工智能这一类新事物，立法规制上始终存在着鼓励和限制两种声音，客观而言规范过多确实会限制新事物的发展，规范过少又有引起新事物发展异化风险，对于相关领域的特别立法有必要持特别审慎的态度。

我国高度关注生成式人工智能领域的立法，2023 年 8 月国家网信办联合国家发展改革委、教育部、科技部、工业和信息化部、公安部、国家广电总局公布实施《生成式人工智能服务管理暂行办法》（以下简称《暂行办法》），这是中国首次对生成式 AI 研发及服务作出明确规定。该管理办法从技术发展与治理、服务规范、监督检查与法律责任层面对当下生成式人工智能的提供者和使用者的权利和义务提出了具体要求，同时规定了如果不履行义务可能产生的责任。《暂行办法》明确了人工智能的使用原则及规定其适用范围为向我国境内公众提供服务或利用境外技术向境内提供服务的人工智能，针对使用生成式人工智能可能存在的风险对各主体提出了具体的责任要求，如需要进行安全评估、履行备案变更注销等手续、配合有关部门提供必要的数

据支持与协助等。2022 年 11 月国家互联网信息办公室、中华人民共和国工业和信息化部、中华人民共和国公安部联合发布第 12 号令《互联网信息服务深度合成管理规定》，明确了组织和个人进行互联网信息深度合成服务的基本规则，包括有关机构的登记、技术支持者数据和技术管理规范、有关部门的监督和管理等。

同时我们也会注意到，当前人工智能侵权的立法规制并不完善，具体表现为当前我国针对人工智能领域的立法特别是生成式人工智能领域立法相对薄弱，位阶较高的规范性法律文件仅是由国家网信办联合多部委公布的《暂行办法》，该办法对当下较为常见的人工智能服务可能出现的问题进行规制，但是其并没有解决人工智能侵权的整体性问题，仅是就部分频发的领域或问题做出规定[2]。同时人工智能的高度电子化特性也让对其的规制相较于一般侵权而言专业化更强，复杂程度更高。其次是人工智能侵权的认定与救济等配套机制不完善。如前文所述，由于人工智能侵权的立法规制等缺乏，导致出现人工智能侵权情形时其认定和处理需要类比相关规则进行。如实务中出现人工智能的生成物侵权采《民法典》侵权责任编关于产品侵权责任的认定和处理方式，但是民法典产品侵权中的产品一般指代是指经过加工、制作，用于销售的产品，直接套用进人工智能领域未免牵强。救济上存在困难使得受害人往往难以有效维护自身的合法权益，缺乏人工智能专门法导致的对于人工智能主体性质和生成物性质的认定争议使得受害者救济不畅。

3. 生成式人工智能训练信息的侵权认定

侵权之债作为一种法定之债认定时应当有法律的明文规定。鉴于目前生成式人工智能侵权没有特殊的侵权法律规范，故而生成式人工智能的训练信息侵权原则上仍然应遵循一般侵权的归责原则即过错责任原则[3]。法律依据则相应的是民法典侵权责任编的规定，认定侵权的要件适用四要件，包括侵权行为、损害后果、因果关系、过错等，四个要件相结合综合判定生成式人工智能的训练信息是否侵权。侵权行为即事实上的对他人造成伤害的某些行为，可以表现为作为，也可以表现为不作为。在生成式人工智能训练信息领域的侵权行为的表现形式可以是作为也可以是不作为。作为的形式如设计者在设计人工智能程序时有意植入秘密搜集涉及其他主体性别、民族、地域等个人信息的程序或是使用者恶意利用人工智能创造侵害他人知识产权利益的物。不作为的形式如作为程序生成式人工智能不可避免存在系统漏洞，使用人在使用程序时已知晓可能存在的问题，故意使用或是放任人工智能程序继续实施一定的行为致使侵害他人法益的情况。损害后果即由侵害行为导致

的一定损害。一般的侵权损害后果主要以显性为主，如故意伤害致受害人损伤，而在人工智能领域更多以抽象形式呈现，具体的表现形式如人工智能生成物（AIGC）即基于生成对抗网络、大型预训练模型等人工智能的技术方法，通过已有数据的学习和识别，以适当的泛化能力生成相关内容的技术。对相关利益人知识产权或者是平等权、人格权造成侵害。因果关系即侵权行为和损害后果之间的某种必然的联系。在人工智能领域侵权中侵权行为与损害后果之间的关系把握起来并不容易。由于对侵权主体的性质认定存在争议，即是讨论人工智能本身行为与损害结果还是设计者或提供者或使用者与损害结果之间的关联责任主体理论及实务界意见不一，现实中主要观点有人工智能主体说、法定主体说和折中说等。同时因果关系适用理论也存在争议，如袁文权认为可以以大数据相关关系取代传统因果关系[4]，王莹则认为算法侵权未脱离传统自然法侵权因果关系理论[5]。故而生成式人工智能侵权的因果关系认定并未有一个相对权威的路径。主观态度（过错）：即侵权主体实施侵权行为时的主观心理状态，在侵权行为的过错责任归责原则之下需要考虑行为人是否具有过错，生成式人工智能侵权认定时侵权主体判定存有不同观点，而以不同侵权主体进行过错认定时的标准并不相同。

生成式人工智能作为一种新技术，其与传统意义上理解的工具又有所不同，这也影响了与之相关的训练信息侵权认定。生成式人工智能能够自行运用自身程序处理信息数据，与之相关的概念包括算力、算法、数据等。算力是设备根据内部状态的改变，每秒可处理的信息数据量，通俗来说就是计算能力，也是对数据的处理能力。算力包含了软件、硬件系统的开发，通过计算机、芯片等载体提供基本运算能力，算力的大小代表着对数字化信息处理能力的强弱。目前提供算力的类型有多种形式，常用的有云计算算力、高性能计算算力、智能计算算力、混合计算算力、算力网络等。算法可比喻为人工智能发展的大脑，是通过一系列人工智能算法，比如机器学习从海量数据中获得规律，并利用规律对位置数据某些特性进行预测与判断，是处理数据信息的规则与方式。人工智能算法众多，常见的有朴素贝叶斯、决策树、深度学习、强化学习、逻辑回归、支持向量机、遗传算法、蚁群算法、元学习等。人工智能算法的类型可依据人工智能算法理论基础知识，如概率统计、集合论、空间几何、图论、矩阵论等加以分类。例如，基于集合论分类则可分为 K-means（K 均值算法）、k-NN（k 近邻算法）以及 Apriori 算法等。数据是对客观事实的描述，或是人们通过观察、实验或计算得出的结果，是信息的表现形式和载体。数据的类型有多种，其中最简单的就是数字，也可以是文字、符号、图像、语音、视频等。数据可以是连续的值，比如声音、图像，称为模拟数据。也可以是离散的，如符号、文字，称为数字数据。在计

算机系统底层，数据通常以二进制信息单元 0、1 的形式表示[6]。在具体实践中，生成式人工智能的运行在抽象层面上，它通常采用神经网络，即由可执行计算的互联节点组成，节点之间的相互影响或连接即模型参数或权重，其是在数据训练过程中获得的。大模型是一个参数神经网络，即对数据中的实体、模式和关系的编码表示，可以根据输入而自主输出内容。大模型输出的过程实质上是将这些具有统计相关性的特征予以随机性重组。相对于训练数据而言，其输出结果是全新的，但其构成特征均符合训练数据集中各样本的统计相关性[7]。生成式人工智能算法训练是抓取作品与输出内容的关键环节：抓取作品是算法训练的物质基础；输出内容决定了算法训练模型设定的目标与价值取向，直接影响算法训练作品利用的具体方式[8]。这些在传统侵权领域未曾出现的全新要素深刻影响了生成式人工智能的训练信息侵权的认定，侵权方式出现了较大变化。简而言之对生成式人工智能训练信息的侵权认定需要转化为算法语言来加以判定而非传统意义上的直接判定，要结合生成式人工智能程序的实际运行流程实际处理信息来判断其是否构成侵权。

4. 生成式人工智能训练信息侵权认定之不足

如前文所述，目前生成式人工智能领域尚无特别侵权的相关规定，故而其适用一般侵权规则，但由于其作为新技术的特殊性，故而在生成式人工智能训练信息的侵权认定上存在一定程度的空白。具体包括生成式人工智能的训练信息收集边界相对模糊的不足、对象未有区分的不足以及自身程序透明度不足等。

4.1 边界相对模糊的不足

即生成式人工智能的训练信息搜集的边界模糊。具体而言就是什么数据可以搜集，什么数据不能搜集，这里面临的是一个数据信息使用合法化的问题。与之相对应的侵权构成要件即侵权行为，不合理的信息搜集不仅可能严重侵犯相关主体的合法民事权益给当事人或者当事企业造成不良影响，情节严重的甚至可能构成刑事犯罪。民事权益方面如“何某诉上海自某人工智能科技有限公司网络侵权责任纠纷案”[注 1]，该案中某智能科技公司设计出的算法规则可以根据使用者提供的相关信息拟制出一个虚拟人物，使用者可以与上述虚拟人物对话。何某在未授权上述科技公司搜集信息的情况之下，人工智能软件违法搜集何某的个人信息，构成对何某人格权益的侵犯。又如黄某诉邵某人格权纠纷案[注 2]，该案中黄某与邵某系邻居，邵某在家门口安装了人工智能程序的摄像装置，该装置可自动识别及拍照，由于两户之间间隔

较近且通过该人工智能摄像系统可以拍摄到黄某家的阳台等内部环境一定程度影响黄某隐私权利，最终该装置的使用被人民法院依法认定构成侵权并要求限期拆除。由此可见信息搜集边界的模糊影响了生成式人工智能的运用及发展。刑事犯罪方面如唐某侵犯公民个人信息案[注3]，该案中被告人唐某开发所谓的“星空互语人工智能电话机器人”软件对外销售，并向客户收集利用该软件拨打电话号码，后通过技术手段对上述电话号码对应的机主姓名等信息进行匹配，形成可以识别特定自然人身份的信息并加以记录。唐某这一利用人工智能系统搜集公民个人信息的行为最终被法院认定犯侵犯公民个人信息罪。

4.2 对象未有区分的不足

即生成式人工智能的训练信息搜集的对象没有区别开来。一般而言侵权行为的认定与侵权对象没有必然联系。但生成式人工智能对不同民事主体的训练信息搜集及使用侵权认定应该有所区分，因为不同对象的权利保护客观而言存在差异，特别是人格权益的部分。具体而言，对自然人来说，训练信息的搜集使用与人格权益息息相关，除非法律明确规定或者有当事人的明确授权，否则不应擅自使用自然人相关个人信息作为训练信息。生成式人工智能不合理的信息搜集与训练可能侵犯自然人的相关人格权利如隐私权等。而对于非自然人主体如公司法人等，因非自然人主体涉及到的人格权利有限，类型上生成式人工智能的不合理数据搜集更多可能侵犯其商业秘密权利、知识产权等财产性权利。如某某（北京）科技有限公司、上海妙克某某科技有限公司等著作权权属、侵权纠纷案[注4]，该案中上海妙克某某科技有限公司利用某人工智能软件生成相关乐谱等与某科技公司创作的乐谱高度雷同，其被指控相似的乐谱系使用相关生成式人工智能软件利用自身程序搜集的训练信息而训练生成，该生成的乐谱被认定侵犯原告相关财产权利。

4.3 透明度程度欠佳的不足

生成式人工智能程序往往与相关的算法企业的商业秘密或著作权等权益息息相关，加之由于生成式人工智能算法的专业性以及一定程度的秘密性即所谓算法黑箱，一般人很难了解到生成式人工智能信息搜集的方式以及限度，有时就算知道了算法程序的所谓源代码，在没有专业人士介绍的情况下一般人也很难理解其中的奥义，如某电脑贸易（上海）有限公司诉国家知识产权局、上海某网络科技股份有限公司发明专利权无效行政纠纷案[注5]，该案中所提及的包括“过滤器将用户语句区分为格式化语句或自然语句、查询模块或者对话模块可以和游戏服务器相交互、UTF-8 编码语句具体为格式化

语句抑或自然语句、网络学习扩充对话数据库的具体技术手段、网络学习扩充对话数据库、人工智能服务器检索对话数据库并选择最合适的应答语句的具体方式”等专业术语难以为一般人所理解。而如果相关算法设计方不公开相关程序语言，实际运用算法技术的生成式人工智能使用者更是无从知晓人工智能程序实际运用时搜集或使用了多少数据以及这些数据来源于何处。

5. 生成式人工智能的训练信息侵权认定之完善

针对生成式人工智能训练信息的侵权认定之不足，可以从明确合理边界、区分不同对象、完善算法透明度等方面加以完善。

5.1 明确合理边界

即明确生成式人工智能什么信息可以搜集，什么信息禁止搜集。笔者认为，生成式人工智能信息的搜集边界应当有法定与意定之区别，具体而言即法律禁止搜集的信息无论如何生成式人工智能不得搜集，哪怕是当事人同意。而法律没有规定禁止搜集的信息则可以经当事人同意后进行信息搜集与训练。以公开方式呈现的信息应当视为允许进行信息搜集，如上市公司财务报表、公开的裁判文书等，需要明确的是，信息的公开并不意味着当事人许可了完全的使用权[9]，也应当注意训练信息在处理和使用时不能违反其他相关法律规定与公序良俗等。如使用个人信息时要给予足够尊重，不能进行对个人人格进行污名丑化，在利用非公民个人信息时要注意训练信息使用的保密。训练信息搜集与处理边界的明确，一方面能够促进生成式人工智能产业发展，另一方面也有助于保护相关主体的合法权益，这也是一种利益平衡与产业激励的方式[10]。

5.2 区分不同对象

应当针对不同的对象训练信息抓取侵权进行区别认定。对自然人对象采取严格的保护主义，没有法律明确规定或者当事人没有明确授权，一般而言人工智能算法程序不得擅自搜集与使用自然人个人相关信息，这以要求也延展到了公开领域搜集的个人信息数据。如欧盟就要求对于并非从数据主体处获得的个人数据，必须履行告知义务；英国要求从公开可及的资源获取个人数据，需要具有合法依据，并通知个人，对于超出个体期待的数据处理，需要告知和评估[11]。而对非自然人对象采取宽容主义，一方面非自然人主体涉及的人格权益内容更少，另一方面笔者认为非自然人主体进行对外活动特别是商事活动本身就需要以自身部分信息的公开透明为前提，如预备上市企业应按照法律规定公开自身近期真实的财务状况、人民法院依法公开相关主体

的涉案状况等。在这样的前提之下生成式人工智能对这些公开信息进行搜集并作为训练信息使用并无不妥，此时在人工智能训练信息搜集使用的侵权认定上就同自然人主体的认定存在一定差异。

5.3 完善算法透明度

由于生成式人工智能算法程序存在作为商业秘密或著作权的可能，同时算法程序语言高度的专业性与更新变化始终存在，故而实际上不存在生成式人工智能算法能够让所有人完全理解的可能，更多的应是对算法程序透明化不足的完善。完善算法透明度不足主要分为内向和外向两个维度。一方面是内向维度，即生成式人工智能算法企业自身对程序设计的优化。作为生成式人工智能的开发设计者，其最了解该人工智能的程序设计，所谓解铃还需系铃人，从设计者角度出发解决问题相较于其他参与主体而言效率更高，效果也更好。另一方面是外向维度，即有关部门加强对生成式人工智能的训练信息的监管，《暂行办法》第十六条明确了网信、发展改革、教育、科技、工业和信息化、公安、广播电视、新闻出版等部门，应当依据各自职责依法加强对生成式人工智能服务的管理。有关主管部门针对生成式人工智能技术特点及其在有关行业和领域的服务应用，完善与创新发展相适应的科学监管方式，制定相应的分类分级监管规则或者指引[注6]。这也对相关规范的标准提出了高的要求，即在管理的同时也要保护行业的良性发展，特别是对于生成式人工智能在信息抓取行为上进行规范，不能仅有纲领性的文件，相关部门也应出台更加细化的部门规章和行业规范标准等，如此才能将监管和规范落到实处。

6. 结语

生成式人工智能本质上作为一种算法程序，包括其训练信息侵权认定在内的法律规范必然应当与算法程序侵权法律规范相联结，当前我国没有关于算法侵权的特别法律规范而是以一般侵权规范对算法侵权相关行为予以规制，考虑到算法侵权与传统侵权的差异，客观而言有进一步区别完善之必要，包括但不限于侵权的认定、承担责任的主体及相应的救济措施等，这种规范应当是统一的范式以对不同形式的人工智能模式进行相关的侵权认定。

参考文献

[1] 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,(06):86-107.

[2] 中国政府网：生成式人工智能服务管理暂行办法[EB/OL].(2023.7.10)[2025.12.10].https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm

[3] 王利明.生成式人工智能侵权的归责原则与过错认定[J/OL].中国法律评论,1-15[2025-07-03].http://kns.cnki.net/kcms/detail/10.1210.D.20250702.1616.002.html.

[4] 袁文全.算法歧视的侵权责任治理[J].兰州大学学报(社会科学版),2023,51(02):89-99.

[5] 王莹.算法侵害责任框架刍议[J].中国法学,2022,(03):165-184.

[6] 人工智能发展的三大基础要素，你知道多少？深圳市人工智能研究会[EB/OL].(2023.3.17)[2025.12.7]https://www.saiia.org.cn/index.php/2023/03/17/rgzndsdjcys/

[7] 梁志文.版权法上生成式人工智能输出的定性及其责任规则[J].法学,2025,(06):129-146.

[8] 倪朱亮.生成式人工智能训练使用作品的许可复合机制研究[J/OL].法律科学(西北政法大学学报),2025,(04):1-14[2025-07-04].https://doi.org/10.16290/j.cnki.1674-5205.2025.04.004.

[9] 程啸.论公开的个人信息处理的法律规制[J].中国法学,2022,(03):82-101.

[10] 杨利华.生成式人工智能服务提供者注意义务研究[J].比较法研究,2025,(03):54-68.

[11] 丁道勤.生成式人工智能训练阶段的数据法律问题及其立法建议[J].行政法学研究,2024,(06):16-28.

注释

[注 1] 参见北京互联网法院（2020）京 0491 民初 9526 号民事判决书

[注 2] 参见上海市青浦区人民法院（2020）沪 0118 民初 15600 号民事判决书

[注 3] 参见上海市浦东新区人民法院（2019）沪 0115 刑初 5251 号刑事判决书

[注 4] 参见（2024）最高法知民终 704 号民事判决书

[注 5] 参见最高人民法院（2017）最高法行再 34 号行政判决

[注 6] 《生成式人工智能服务管理暂行办法》第十六条之规定