

论ChatGPT侵犯原始数据所有者隐私权的责任归属

李卓玲

江西理工大学, 中国·江西 赣州 341099

【摘要】 ChatGPT是美国OpenAI公司开发的一款生成式人工智能产品, 其对人类社会的生产生活都能带来巨大的便利。但是, 在ChatGPT的原始数据的收取阶段和运用阶段, 其原始数据的所有者都暗藏着许多被侵犯隐私权的风险, 关于原始数据的所有者被侵犯隐私权的责任归属, 各学者也各执一词。有学者认为应由ChatGPT作为独立的民事主体担责; 有学者认为ChatGPT不具有独立的人格, 应当由其开发商OpenAI担责; 还有学者认为除OpenAI外仍应存在其他的责任主体。

【关键词】 ChatGPT; 原始数据; 隐私权; 侵权责任主体

引言

2022年, 美国OpenAI团队开发出了一款名为ChatGPT的生成式人工智能产品, 引起了社会广泛关注, 对人类社会生活产生深刻的影响。受ChatGPT的影响, 我国的百度和阿里等公司也相继推出了文心一言、道义千问等生成式人工智能产品, 国内外都兴起了一股AIGC产品的热潮。

有的学者认为ChatGPT具有独立的人格可以自行承担侵权责任, 有的学者认为ChatGPT不具有独立的人格, 而应当由其开发商OpenAI担责, 而也有学者认为归责主体不应局限于开发商, 还应包括算法的设计者、网络侵权者等等, 应当具体情况具体分析, 明确不同ChatGPT运作阶段下的不同责任主体。

一、ChatGPT原始数据相关理论

(一) ChatGPT概念及内涵 ChatGPT (ChatGenerative Pre-trained Transformer) 是美国OpenAI公司在2022年11月30日发布的人工智能聊天机器人程序, 属于生成式人工智能的一种重要类型。生成式人工智能是指基于数据学习训练后能够输出类似人类思维的复杂内容, 并可执行通用问答系统或自动创建艺术图像 (如StableDiffusion) 等任务的技术。此外, 其生成内容的能力基于其开发商大规模地收集数据以及Transformer模型都才用, ChatGPT将开发商收集的原始数据作为其训练的养料, 用Transformer模式使得ChatGPT迅速获得生成内容的能力, 再运用其原始数据按照用户的指令生成内容。

(二) ChatGPT原始数据特点

依据《中华人民共和国数据安全法》的条款八, 执行

数据操作时应遵循法律法规, 尊崇公众道德和社会伦理, 恪守商务操守与行业规范, 保持诚信并尽到数据保全的责任, 对社会的贡献不可忽视, 同时不能有损国家和个人的权益。由此可以看出, 数据本身就依法被保护。而原始数据本质也是数据, 也应受到保护。

原始数据是OpenAI公司及开发商为了训练ChatGPT的生成能力而收集的, 同时也被运用于生成内容的过程之中。与一般的数据不同, ChatGPT的原始数据所有人具有人格属性和财产属性。

(三) 对于利用聊天机器人(ChatGPT)可能引发的数据拥有者隐私泄露问题, 责任方应承担起相应的法律职责, 确保用户提供的资料仅用于合法用途, 避免采集过多的私人信息或存储可识别用户身份的信息, 同时禁止未经授权地分享这些信息给第三方。此外, 他们还需遵守相关法规, 积极回应并妥善处理用户提出的查询、拷贝、修改、增补或者清除他们的个人信息的申请。而法条之所以会做出此规定, 便也体现出现实中存在着此种原始数据所有者隐私权被侵犯的现象。

1、原始数据收取阶段中侵犯隐私权的风险

在2023年的6月29日, 16位匿名者在美国加州旧金山联邦法庭上对微软和openai提起诉讼, 控告chatgpt未给用户充分通知或获得其同意, 非法搜集并泄露他们的个人资料。原告寻求高达30亿美元的赔偿。这个情形显示出在ChatGPT都原始数据的收取阶段, 开发商便会对此非法收集, 从而侵犯原始数据所有人所有者的隐私权。

在这个时期, OpenAI公司致力于提升ChatGPT的语言技

能，他们通过大量的原始资料来实现这一目标。而其中的核心技术就是使用了Transformer模型。这个模型的使用使ChatGPT能迅速理解数据间的关联，借助人工智能教练员对其输入和输出的频繁调整，让它初次领悟到自然的语言结构规律。为了ChatGPT能够运用此模型快速掌握自然语言的准则，OpenAI公司往往会收集大量数据，但是关于其原始数据的来源是否合乎法律规定的问题，ChatGPT没有给出明确答复，其中往往暗藏着原始数据所有者的隐私权被侵犯的风险。

2、原始数据运用阶段中侵犯隐私权的风险

在此阶段，用户往往会输入一个指令，ChatGPT会根据该指令，在原始数据中寻找与指令具有最密切联系的数据，通过算法整合分析数据，然后生成与用户指令所相对应的内容。此外，用户输入的数据也可能被ChatGPT转化为原始数据，被用以提高ChatGPT的数据处理能力。在该ChatGPT运用原始数据的阶段，可能会发生因算法黑箱而导致的原始数据的泄露，因网络入侵者的干扰而导致的原始数据的侵权，以及因用户诱导而导致的原始数据的披露的情形，从而使原始数据所有者的隐私权具有被侵犯的风险。

二、ChatGPT自身能否作为侵权责任主体的探析

对于生成式人工智能能否拥有法律人格即民事主体资格，总和国内外学者观点主要有以下三种学说。一是否定说，即生成式人工智能不应拥有法律人格，当然也不能被视为法人，将生成式人工智能侵权看作产品侵权，责任承担则是开发者、销售者等责任主体承担不真正连带责任；二是折中说，也称有限人格说，即赋予生成式人工智能有限的法律人格，类似于法人；三是肯定说，即生成式人工智能应当拥有法律人格，使生成式人工智能能够像自然人一样享有民事权利和承担民事义务和民事责任，能够更好从事民事活动。但是我国目前对以上三种观点并没有作出明确的选择，对ChatGPT自身能否作为侵权责任主体的问题没有给出明确的解答，各方学者也各执一词。

(一) ChatGPT背后的责任主体承担不真正连带责任有学者提议，赋予人工智能民事主体资格，由人工智能作为侵权责任主体对其造成的损害独立承担侵权责任，并建立相应基金账户，当人工智能给他人造成损害时，优先以账户

资金进行赔付。然而，倘若从实质意义视角考量，ChatGPT不能以其独立的意志从事民事活动，不能独立对外承担民事责任，即使赋予其民事主体资格，实际承担赔偿责任的主体仍是自然人，此项制度没有任何实际意义。

当原始数据的所有者的权利遭受侵害时，应当由其背后的开发商、算法设计者等责任主体承担不真正的连带责任，而非由ChatGPT自身担责。

三、ChatGPT在不同阶段下侵犯隐私权的不同责任归属

2023年7月13日发布的《办法》的条文较少，对于隐私权保护的条文更是屈指可数。根据《办法》的规定，第9条指出，任何向公众提供的网络信息内容的创建方都应遵守法律规定的责任和网络信息安全的职责。如果这些内容包含个人数据，则需要遵循相应的法律法规来负责并保证对该类信息的隐私权保障。此外，为确保用户权益，所有此类服务的提供商必须与其所使用的生成型人工智能服务的使用者签署一份详细的服务合同，以明晰各自的责任范围和权力关系。该条文也即仅仅明确了服务提供者及开发商和算法设计者的责任以及个人的责任，但也没有进行进一步细化和规范。因此，对原始数据所有者隐私权被侵犯的问题，进一步细化其责任主体是具有实质意义的，而对其责任主体的归属问题，应当根据ChatGPT运作不同阶段进行分别讨论。

(一) 原始数据收取阶段中侵犯隐私权的责任主体

初始化ChatGPT的语用技能通常需要大量输入的数据来实现。在此过程中，开发者会在网站上公开他们的隐私策略，其中明确指出他们收集了“通过您使用我们的服务而获得的个人资料”，这包含了如日志记录、使用情况、设备特征及cookie等信息。虽然这类信息主要为技术性的内容，但是它们仍然可能泄露出用户的私人信息。比如，日志记录里的“网络IP地址、网页浏览器的种类及其设定”以及使用情况里“用户所在的时间区域和国别”等都对保护用户的安全构成潜在的风险。除此之外，从其隐私政策中我们也可得知如果用户创建账户来使用服务或与其沟通，ChatGPT会收集个人信息，具体如下包括账户信息、用户内容、通讯信息、社交媒体信息，而用户对于其信息将被运用于何处以及将以何种方式被应用往往是不知情的，从而极有可能导致原始数据所有者的隐私权被侵犯，而导

致此种现象发生的主体是开发商OpenAI公司及数据的收集主体也及其侵权行为的实施主体。

(二) 原始数据运用阶段中侵犯隐私权的责任主体

1、网络入侵者作为侵权责任主体

首先,在原始数据运用的阶段,可能会有网络入侵者入侵ChatGPT系统盗取原始数据,侵犯原始数据所有人的隐私权。其次,网络入侵者也可能在原始数据被运用阶段扰乱ChatGPT系统的秩序,入侵者对系统进行入侵可能会导致算法的错误,使其系统不能正常运转,从而使系统丧失对原始数据的保护能力,从而导致原始数据的泄露。

2、用户作为侵权责任主体

在原始数据的运用阶段,用户可能会利用ChatGPT的训练机制对ChatGPT进行诱导,故意诱导其一步步泄露ChatGPT的原始数据,侵害原始数据所有人的隐私权。用户可能会故意不按照ChatGPT的操作步骤要求下达指令,ChatGPT也对其行为进行了筛选,选择性地回答用户下达的指令,但若用户仍不停止对ChatGPT的诱导,导致ChatGPT的原始数据保护系统崩盘,从而也会导致原始数据的泄露。

(三) 侵犯原始数据所有者隐私权的兜底责任主体

OpenAI作为开发商,且ChatGPT又无独立从事民事活动的的能力,故OpenAI公司对ChatGPT的行为应当具有监督和管理义务。其次,OpenAI公司作为ChatGPT的开发商,其为实际上的受益者。根据《民法典》的公平原则和民事损害填平原则,受益者应在不存在其他侵权主体的情况下对被侵犯隐私权的原始数据所有者的承担责任,及作为兜底条款的侵权责任主体。同时,将OpenAI作为原始数据所有者隐私权被侵犯的责任主体能使被侵权人的损害赔偿请求权获得最大限度的实现,使被侵权人的损害能够被完全地填平,故将OpenAI作为ChatGPT原始数据侵权的兜底条款的责任主体是合理的正当的。

结语

ChatGPT作为一项新兴的生成式人工智能,拉开了强人工智能的序幕,对人类生产生活带来了巨大的便利。微软创始人比尔·盖茨宣告,生成型的人工智能将会成为继互联网和智能手机之后的第三个具有颠覆性的科技进

步。在这个过程中,ChatGPT作为一个新颖的生成型人工智能工具,它的发展离不开原始数据这个关键因素,而在它从采集到应用的过程中都存在可能遭受侵犯的风险。虽然ChatGPT具有先进性和实用性,但是我们也应当看见ChatGPT具有的原始数据所有者的隐私权被侵犯的实际风险,其也可能冲突现有的法律秩序。

要而论之,我们应当辩证地看待ChatGPT的实用性以及风险性,在看到ChatGPT给我们生活带来便利,给我们社会带来进步的同时,也应当看见ChatGPT所暗含的法律风险,看见该风险下所蕴含的不完善的法律规制,通过对法律不断的完善,并且灵活地运用现有的法律,不断推动ChatGPT在我们生活中的运用,促进社会不断发展。

参考文献:

- [1] 刘霜,张潇月.生成式人工智能数据风险的法律保护与规制研究——以ChatGPT潜在数据风险为例[J].贵州大学学报(社会科学版),2023,41(5)
- [2] 秦俊龙.生成式人工智能创作物的侵权责任界定探析——以ChatGPT为例[J].法学(汉斯),2023,11(6)
- [3] 陈思敏.ChatGPT侵权责任问题研究[J].中阿科技论坛(中英文),2023(12)
- [4] 张高源.试论生成式人工智能侵权的责任主体——以ChatGPT为例[J].产品可靠性告,2023(12)
- [5] 邓建鹏,朱恽成.ChatGPT模型的法律风险及应对之策[J].新疆师范大学学报(哲学社会科学版),2023,44(5)
- [6] 郭玉洁,丁国峰.ChatGPT生成式AI的数据安全风险及应对[J].大数据时代,2023,(07)
- [7] 孙露,刘健男.ChatGPT场景下数据安全风险与法律应对[J].中国广播电视学刊,2023,(09):10-12+29.
- [8] 谭佐财.ChatGPT的法律风险与治理路径[J].湖南科技大学学报(社会科学版),2023,26(03):
- [9] 朱孟垚,李兴华.ChatGPT安全威胁研究[J].信息安全研究,2023,9(06)
- [10] 程乐.生成式人工智能的法律规制——以ChatGPT为视角[J].政法论丛,2023,(04)

通讯邮箱:706704356@qq.com