

解读《生成式人工智能服务管理暂行办法》第9条第1款

闫子晗

上海大学法学院，上海

收稿日期：2026年1月9日；录用日期：2026年2月2日；发布日期：2026年2月12日

摘要

《生成式人工智能服务暂行办法》第九条第1款明确规定：“提供者应当依法承担网络信息内容生产者责任，履行网络信息安全义务。涉及个人信息的，依法承担个人信息处理者责任，履行个人信息保护义务。”通过对该条款及相关规定进行对比，可以发现义务主体与责任主体不一致，由此对生成式人工智能服务提供者承担网络信息内容生产者责任的合理性问题进行反思。要求某一主体承担某一责任，前提需要该主体有相应义务。从服务提供者的义务来源即其与生成式人工智能技术的密切关系出发，分析生成式人工智能技术原理与特点，阐明内容的生成实质上受到数据、算法、人机交互等众多因素的共同影响，从技术层面指出立法者将承担内容侵权责任的主体归为生成式人工智能服务提供者是基于其对数据收集与算法设计具有较强控制力。然而，仅就这两方面的理由没有满足由服务提供者承担内容生产者责任的充分必要条件，并没有全面考虑在生成内容的过程中还涉及到服务使用者的影响以及服务提供者对最终的内容生成起到的控制力有限的情况。对该条款的解读还应当回到法律文义解释上，这要求必须透彻理解“生成式人工智能服务提供者”与“网络信息内容生产者责任”两个关键词。相关规定对“服务提供者”的定性不同，定性决定主体的法律地位，进而决定相关主体承担何种法律责任。因此，生成式人工智能服务提供者的法律地位的认定是重中之重，《生成式人工智能服务暂行办法》将其定性为“内容生产者”，但是由于生成内容的控制力方面二者具有显著不同，使得将二者等同缺乏合理依据，进而要求生成式人工智能服务提供者承担内容生产者责任缺乏正当性，立法者在制度设计上忽视其他主体尤其是生成式人工智能服务使用者通过“指令”对内容生成发挥的作用，进而只要求服务提供者承担全部内容侵权责任缺乏合理依据。

关键词

生成式人工智能服务提供者，网络信息内容生产者责任，ChatGPT

Interpret Article 9, Paragraph 1 of the Interim Measures for the Administration of Generative Artificial Intelligence Services

Zihan Yan

Abstract

Article 9, Paragraph 1 of the Interim Measures for Generative Artificial Intelligence Services clearly stipulates: “Providers shall assume the responsibility of producers of network information content in accordance with the law and fulfill the obligation of network information security. Where personal information is involved, it shall assume the responsibility of a personal information processor in accordance with the law and fulfill the obligation to protect personal information.” By comparing this provision with the relevant regulations, it can be found that the subjects of obligation and responsibility are inconsistent, thereby reflecting on the rationality of generative artificial intelligence service providers bearing the responsibility of producers of online information content. To require a certain subject to undertake a certain responsibility, it is necessary that the subject has corresponding obligations. Starting from the source of service providers' obligations, that is, their close relationship with generative AI technology, analyze the principles and characteristics of generative AI technology, and clarify that the generation of content is essentially influenced by many factors such as data, algorithms, human-computer interaction, etc. From a technical perspective, it is pointed out that legislators attribute the subject of content infringement liability to generative AI service providers based on their strong control over data collection and algorithm design. However, the reasons for these two aspects alone do not meet the necessary and sufficient conditions for a service provider to assume the responsibility of a content producer, and do not fully consider the influence of the service user involved in the process of generating content and the limited control that the service provider has over the final content generation. The interpretation of this provision should also return to the literal interpretation of the law, which requires a thorough understanding of the two key terms “generative AI service provider” and “network information content producer responsibility”. The relevant provisions have different characterizations of “service providers”, which determine the legal status of the subject and thereby determine what kind of legal liability the relevant subject bears. Therefore, the determination of the legal status of generative AI service providers is of Paramount importance. The Interim Measures for Generative AI Services categorizes it as a “content producer”, but due to the significant differences in the control of generated content, there is no reasonable basis for equating the two, and thus it is not justified to require generative AI service providers to assume the responsibility of content producers. Legislators' institutional design neglects the role of other subjects, especially users of generative AI services, in generating content through “instructions”, and thus only requires service providers to bear full liability for content infringement.

Keywords

Generative AI Service Provider, Network Information Content Producer Responsibility, ChatGPT

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

将国家互联网信息办公室(以下简称“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、广电总局公布的《生成式人工智能服务管理暂行办法》(以下简称《办法》)与此前公布的《生成式人工智能服务管理办法(征求意见稿)》(以下简称《征求意见稿》)进行对比,针对人工智能生成内容风险的义务与责任主体设定以及具体内容方面可以发现前者做出了部分更改,但部分保持不变。

首先,对义务主体进行修改。《办法》第四条规定了生成式人工智能服务的提供者与使用者应当共同遵守的五项义务,改变了《征求意见稿》第四条仅对提供者设置相应义务,并未对使用者做出任何限制性规定。其次,在具体义务的规定上也做出调整。将《征求意见稿》第四条第四项义务要求即“利用生成式人工智能生成的内容应当真实准确,采取措施防止生成虚假信息”修改为“提高生成内容的准确性和可靠性”。《办法》考虑到服务提供者仅能确保用于生成式人工智能“预训练”的数据的真实性与准确性^[1],而由于算法技术与大型语言模型技术所呈现出的自主性和不可预测性等特征,提供者无法控制生成内容的真实性,且人工智能本身也不能验证生成内容的真实性与准确性,故将服务提供者保证生成内容之真实准确性的严苛要求下降到仅要求提高生成内容在准确性和可靠性方面的程度即可。再次,对责任主体的设定保持一致。据《征求意见稿》第五条以及《办法》第九条可知,都将承担生成式人工智能生成内容风险责任的主体规定为生成式人工智能服务提供者。

这意味着由双主体——生成式人工智能服务提供者与生成式人工智能服务使用者对生成式人工智能生成内容履行保证义务,但当生成的内容风险涉及危害国家与公共安全、个人名誉、企业名誉与经济利益以及知识产权侵权等问题,仅由生成式人工智能服务提供者承担相应责任,排除了使用者对内容生成承担责任。不仅如此,《办法》已经在具体义务方面考虑到了服务提供者对生成内容控制力的欠缺问题,但是在责任承担方面又不予考虑,直接规定服务提供者承担“内容生产者责任”,即要求其承担全部责任。上述问题的核心争议点在于要求生成式人工智能服务提供者承担网络信息内容生产者责任是否具有正当性。

本文将从以下几个方面进行说明,首先从生成式人工智能服务提供者与相应技术之间的关系来探讨《办法》与《征求意见稿》将责任主体聚焦于生成式人工智能服务提供者的原因;其次通过界定生成式人工智能服务提供者的法律地位来解读《办法》第九条第1款的含义;最后得出本文结论,即仅由生成式人工智能产品提供者承担相应责任不具有正当性。

2. 服务提供者承担内容生产者责任的原因

责任的产生源自对义务的违反,分析立法者将承担内容涉及侵权风险的责任主体规定为生成式人工智能服务提供者的原因,必须先厘清服务提供者对网络信息安全等方面负担义务的原因。显然地,这来源于这一认定——服务提供者与生成式人工智能技术、生成式人工智能技术与生成内容之间具有紧密联系,即由于内容直观来看是人工智能生成的,因此由提供生成式人工智能这类技术与服务的主体来承担内容责任是最合适不过的。然而,若要将具体原因分析透彻,则有必要深入研究生成式人工智能的技术原理与特点,在此基础上试图从技术层面探讨要求服务提供者承担内容责任是否具有一定合理性。

2.1. 生成式人工智能技术特性

《办法》给出了生成式人工智能技术的相关定义,是指具有文本、图片、音频、视频等内容生成能力的模型及相关技术。以 ChatGPT 为代表的新一代生成式技术的应用模型,其算法框架是基于“生成式预训练转化器模型(GPT)”而实现生成式的内容输出,本质在于利用预训练语言模型和人类反馈的强化学习(RLHF),对给定指令输出最适当的内容^[2]。以 ChatGPT 为例,其采用的是大规模语言模型预训练的技术方式来实现文字内容的自动生成。

具体来看,生成式人工智能技术特点主要体现在,其一,生成式人工智能的技术原理发生改变。生

成式人工智能之所以能自动生成文字、图像、音视频等内容是建立在数据收集和算法设计的前期工作基础之上,人工智能对采集到的庞大数据进行分析,自主识别、分类和处理任务,得以实现内容生成,且生成的内容因人而异、因时而异,甚至同一使用者输入相同“引导词”,ChatGPT 输出的内容都有所不同。人工智能具有自主、无法预测和不确定性等特点。在这一过程中不需要人工参与,与传统人工智能技术区别开来,正因如此也加大了对人工智能生成内容进行控制的难度。

其二,影响人工智能生成内容的因素增多。在传统网络环境下,网络信息内容完全由内容生产者制作、复制、发布,直接影响内容的生成,服务提供者(这里指网络平台)仅为其提供帮助,居于辅助地位。然而,运用生成式人工智能技术,影响内容生成的因素包括大数据、算法、使用者等。前两个因素可以由生成式人工智能服务提供者来控制,但是也不能完全忽略使用者在生成过程中起到的地位。况且,根据第一个特性,由于技术原理已然发生改变,即使服务提供者可以控制数据收集与算法设计两个因素,也可能出现生成内容风险问题。

其三,人工智能生成侵权内容的原因认定困难。综合前两个特性,由于算法复杂性带来的算法黑箱问题以及内容的生成是诸多因素共同作用的结果,导致了人工智能生成涉及侵权问题的内容究竟源于何种因素,在现有技术层面下难以准确判断。不仅如此,内容侵权问题不可能只归结为上述提到的任何单一因素,尤其是在不同情形下,也许是算法的设计问题(包括算法偏见)发挥了关键性影响,也有可能某些用户规避了算法预防机制,有意引导 ChatGPT 生成一些歧视性内容。

2.2. 服务提供者承担内容生产者责任的直接依据

前文已经指出,在生成式人工智能生成内容的过程中,大数据是内容生成的基础资源,算法是分析、加工以及处理前者的技术手段,共同影响生成式人工智能最终输出的内容。而生成式人工智能服务提供者可以在数据收集与算法设计两个前端环节(也包含对算法进行修改和更新)发挥控制作用,这也是立法者将内容责任归于生成式人工智能提供者的直接依据。

2.2.1. 数据收集方面

服务提供者通过收集和筛选庞大复杂的数据,来对生成式人工智能进行预训练和优化训练。要知道,数据的真实与准确性对于设计人工智能和构建语料库模型的成功至关重要,乃至影响最终生成内容的呈现。而在纳入初始语料库时,只有服务提供者才能对其进行把控与选择。服务提供者知道哪些数据可以被合法合规使用,哪些数据不可以被使用,否则可能产生侵权风险,相应的甄别责任就需要由服务提供者承担,很显然,生成式人工智能其他主体没有相关权限对于初始数据的合法性进行把控和监管。

2.2.2. 算法设计方面

据前述,生成式人工智能主要基于神经网络,通过预训练、优化训练大规模的数据集,学习抽象出数据的本质规律和概率分布,并利用生成模型生成新的数据,根据使用者的指令,最终输出最优化内容[3]。

针对算法设计方面,由于这一模式——“模型设计 + 数学规则”构建的算法具有高度技术性和复杂性[4],导致非算法专业人士尤其是普通公众,无法掌握或理解算法的运行和决策原理。更不必说,算法“黑箱”的技术壁垒使得非算法专业人士无法判断服务提供者在此方面是否存在过错,这足以影响对服务提供者能否依据过错责任归责[5]。据此,生成式人工智能服务提供者是决定算法的关键,则其在一定程度上决定了生成内容的合法合规性与价值取向,由此产生了提供者在算法设计方面的注意义务,进而需要承担内容侵权责任。

然而,仅就这两方面的理由没有满足由服务提供者承担内容生产者责任的充分必要条件,并没有全

面考虑在生成内容的过程中还涉及到服务使用者的影响以及服务提供者对最终的内容生成起到的控制力有限的情况。

3. 生成式人工智能服务提供者的法律地位

《办法》第九条第1款明确指出,提供者应当依法承担网络信息内容生产者责任。对该条款的理解应当从“服务提供者”和“内容生产者”两个概念入手,实质在于界定生成式人工智能服务提供者的法律地位。

3.1. “服务提供者”定位模糊

“服务提供者”这一概念在《互联网信息服务深度合成管理规定》(以下简称《深度合成管理规定》)与《办法》中均有明确定义。前者将深度合成服务提供者定义为提供深度合成服务的组织、个人。后者将生成式人工智能服务提供者定义为利用生成式人工智能技术提供生成式人工智能服务的组织、个人。但从上述定义中无法准确得出二者对“服务提供者”的定性是否一致的结论,还需要结合服务提供者责任与义务的相关规定。

《深度合成管理规定》第七条要求深度合成服务提供者承担信息安全主体责任,第八条、第十条、第十二条针对深度合成的内容对服务提供者规定了事前审核、事中标识、事后投诉的各项义务。该规定实际上是将服务提供者定性为网络技术服务提供者,通俗理解为内嵌了类似 ChatGPT 等深度合成技术的网络平台。而在《办法》中已经明确规定生成式人工智能服务提供者承担网络信息内容生产者责任,即将其定性为网络信息内容生产者,属于《网络信息内容生态治理规定》中有“网络信息内容生产者”的概念,即“制作、复制、发布网络信息内容的组织或者个人”。

显然地,网络技术服务提供者与网络信息内容生产者对于生成内容侵权所承担的责任并不相同。前者自身并不生产内容,而仅为内容生产者或服务使用者生产、发布或传播深度合成内容提供服务,因此其不是承担内容责任的直接责任人,但是与《网络信息内容生态治理规定》规定的“网络信息内容服务平台”略有不同,前者不仅具有“通知-删除”义务即在收到合格通知以及明知或应知侵权时应当采取必要措施制,还应当履行前述《深度合成管理规定》规定的各项义务,否则承担侵权责任。而内容提供者自身直接制作、复制或发布了侵权内容,是内容责任的直接责任人,应当承担直接侵权责任。

由此,两项法规基于对“服务提供者”的定位不同,要求其承担不同的侵权责任。但是,《征求意见稿》第十五条与《办法》第十四条均对服务提供者规定了诸如“发现违法内容,应当及时采取各种阻止措施”的义务,而“发现-阻止”的义务模式要想发挥实质作用,只能适用于网络信息平台这类主体,当其主动发现或被用户举报违法内容时,只要履行了法定阻止义务,就不再承担法律责任。而生成式人工智能服务提供者承担的是内容生产者责任,一旦发生生成内容造成侵权损害,其就需要承担相应责任,无论是否采取了及时停止或阻止措施,“发现-阻止”机制就失去了其激励性的意义[6]。

因此,《办法》对生成式人工智能服务提供者即要求其承担内容生产者责任,又规定了类似网络信息平台的相应义务与责任,究竟是将服务提供者认定为直接责任人还是间接责任人仍尚不明确。

3.2. “服务提供者”不能被认定为“内容生产者”

上述内容实际上是以将生成式人工智能服务提供者认定为网络信息内容生产者为前提进行讨论的,笔者将对该前提进行反思,即服务提供者与内容生成者能否等同?

前文已经指出,网络信息内容生产者是指制作、复制、发布网络信息内容的组织或者个人。而《办法》中将生成式人工智能服务提供者定义为利用生成式人工智能技术提供生成式人工智能服务的组织、

个人。生成式人工智能服务提供者向用户提供的是其自动生成的内容，例如，在 ChatGPT 场景中，用户通过输入特定指令，由 ChatGPT 自动生成相应的内容，实质上就是由服务提供者 OpenAI 公司提供，而非用户上传。在形式上，人工智能自主生成的内容已经与人类创作的内容无异，因此，服务提供者所做的工作与网络信息内容生产者相当，由此，要求生成式人工智能服务提供者依法承担网络信息内容生产者责任具有“合理依据”。

然而，生成式人工智能服务提供者与网络信息内容生产者在实质上具有显著不同。核心在于生成式人工智能服务提供者并不是直接的内容生成者，而是通过复杂的算法设计间接地影响内容的生成。这一间接性影响体现出二者对于生成内容的控制力具有显著区别，从而导致对其划定侵权责任时应当适用不同规则。

具体来看，应当分析生成式人工智能服务提供者在人工智能生成不同内容的过程中发挥的作用。在服务提供者内部具体到技术开发者，由他们设计算法和训练人工智能自主生成的能力，不同于传统人工智能技术——直接从机器内部具有的大数据库直接提取出来以回答使用者的问题，在此阶段，生成式人工智能与早期人工智能的区别就是前者具备了“机器学习”的能力。例如世界知识产权组织发布的《经修订的关于知识产权政策和人工智能问题的议题文件》指出，“人工智能生成的”与“人工智能自主生成的”均是指在没有人干预的情况下由人工智能生成产出¹。而“没有人干预”与“直接决定”具有互斥性，因此，一旦人工智能开始运行，不断生成海量内容并且因人而异，事实上就脱离了服务提供者能够直接决定的范围，服务提供者甚至无法预测人工智能需要生成何种内容，其与生成式人工智能的联系仅停留在设计与更新算法、修改和弥补漏洞等技术层面之间的关系，在二者中间通过“自主生成的算法”隔断开来，无法达到网络信息内容生产者由于自身直接制作和上传网络信息内容以至于能完全控制何种内容出现的控制力标准。

因此，生成式人工智能服务提供者并不当然是网络信息内容生产者，则要求生成式人工智能网络信息承担内容生产者责任，只能是一种法律拟制。但该法律拟制在笔者看来不具有正当性，其在生成内容的侵权责任主体划定上剔除了人工智能服务使用者。

在事实层面，对生成式人工智能最后输出的具体内容施加影响的并非只有服务提供者，服务使用者可以通过对生成式人工智能下达各种“指令”从而影响生成内容，甚至直接影响内容的生成与修改。再次以 ChatGPT 生成物举例，使用者下达一个“制作著作权法的教课提纲”指令，ChatGPT 随即生成一个提纲；使用者继续下达“将提纲缩短至 10 课时”指令，ChatGPT 生成一个按照要求修改好的提纲。我们可以看出，人对 ChatGPT 下达的指令其实更类似于一种修改意见，根据《著作权法实施条例》第三条第二款，为他人创作进行组织工作，提供咨询意见、物质条件，或者进行其他辅助工作，均不视为创作。ChatGPT 面对人类的一次又一次要求，按照“文字接龙”的方式进行修改。当然，具体修改哪些内容以及如何修改等都是由 ChatGPT 自主把握。

综上所述，无论是生成式人工智能服务提供者还是服务使用者与人工智能生成内容之间都不具有紧密联系，仅产生间接性的影响，承担内容生产者之直接责任不具有正当性，而且仅将内容侵权归责给单一主体——服务提供者也不具有合理性。

4. 结论

本文首先从技术层面指出立法者将承担内容侵权责任的主体归为生成式人工智能服务提供者是基于其对数据收集与算法设计具有较强控制力。然而，仅就这两方面的理由没有满足由服务提供者承担内容生产者责任的充分必要条件，并没有全面考虑在生成内容的过程中还涉及到服务使用者的影响以及服务

¹WIPO/IP/AI/2/GE/20/1 REV. Revised Issues Paper on Intellectual Property Policy and Artificial Intelligence.

提供者对最终的内容生成起到的控制力有限。

其次,从生成式人工智能服务提供者的法律地位方面,分析得出《生成式人工智能服务暂行办法》将其定性为“内容生产者”,但是由于生成内容的控制力方面二者具有显著不同,使得将二者等同缺乏合理依据,进而要求生成式人工智能服务提供者承担内容生产者责任缺乏正当性,立法者在制度设计上忽视其他主体尤其是生成式人工智能服务使用者通过“指令”对内容生成发挥的作用,进而只要求服务提供者承担全部内容侵权责任缺乏合理依据。

参考文献

- [1] 徐伟. 论生成式人工智能服务提供者的法律地位及其责任——以 ChatGPT 为例[J]. 法律科学(西北政法大学学报), 2023, 41(4): 69-80.
- [2] 庄嘉. 涉罪前瞻: AI 解锁暗黑应用技能包[J]. 检察风云, 2019(21): 11-13.
- [3] 蒲清平, 向往. 生成式人工智能——ChatGPT 的变革影响、风险挑战及应对策略[J]. 重庆大学学报(社会科学版), 2023, 29(3): 102-114.
- [4] 钟晓雯. 算法推荐网络服务提供者的权力异化及法律规制[J]. 中国海商法研究, 2022, 33(4): 63-72.
- [5] 任安麒. 网络服务平台算法推荐的著作权侵权认定规则[J]. 北京航空航天大学学报(社会科学版), 2023, 36(3): 190-198.
- [6] 姚志伟, 李卓霖. 生成式人工智能内容风险的法律规制[J]. 西安交通大学学报(社会科学版), 2023, 43(5): 167-170.