

# 生成式人工智能生成内容标识义务的反思

文怡筱

上海大学法学院, 上海

收稿日期: 2026年4月26日; 录用日期: 2026年5月19日; 发布日期: 2026年5月29日

## 摘要

生成式人工智能生成内容具有高仿真性、易传播性与识别困难性, 可能引发人格权益侵害、虚假信息扩散及社会秩序风险。标识义务是区分人工智能生成内容与一般内容的重要规制工具。我国现行制度已初步形成以服务提供者为核心的标识义务体系, 但仍存在“显著标识”概念模糊、义务范围不清、主体责任分配不均的问题。对此, 应细化标识标准, 适度扩展服务使用者的义务, 并完善行业自律机制。

## 关键词

生成式人工智能, 标识义务

# Reflections on the Labeling Obligations for Generative AI-Generated Content

Yixiao Wen

Law School, Shanghai University, Shanghai

Received: April 26, 2026; accepted: May 19, 2026; published: May 29, 2026

## Abstract

Content generated by generative artificial intelligence is highly realistic, easily disseminated, and difficult to identify. It may give rise to risks such as infringement of personality rights and interests, the spread of false information, and threats to social order. Labeling obligations constitute an important regulatory tool for distinguishing AI-generated content from ordinary content. China's current legal framework has preliminarily established a labeling obligation system centered on service providers. However, problems remain, including the ambiguity of the concept of "conspicuous labeling", the unclear scope of obligations, and the uneven allocation of responsibilities among relevant subjects. Accordingly, labeling standards should be further refined, the obligations of service users

should be moderately expanded, and industry self-regulation mechanisms should be improved.

## Keywords

Generative Artificial Intelligence, Labeling Obligations

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. 生成式人工智能生成物的风险

### (一) 技术与现实背景

所谓生成式人工智能(Generative Artificial Intelligence, 以下简称 GenAI), 是指基于深度学习技术所开发的能够根据其训练数据自主生成高质量文本、图片以及其他生成内容的人工智能技术。从 2020 年开始, 基于 Transformer 模型的深度神经网络的进步使许多生成式人工智能系统能够接受自然语言提示作为输入。其中包括大型语言模型聊天机器人, 如 ChatGPT、Bing Chat、Bard 和 LLaMA, 以及能够将文本转换为图片的人工智能艺术系统, 如 Stable Diffusion、Midjourney 和 DALL-E。与传统人工智能相比, 生成式人工智能在学术研究和技术应用上都已经取得了突破性的进展。在自然语言处理领域, 生成式人工智能已经可以生成质量较高的自然语言文本。在计算机视觉领域, 生成式人工智能系统也已经能够生成高质量的图像和视频, 在图像编辑、海报设计、艺术创作等领域都已经得到了广泛的应用。

基于生成式人工智能技术开发的人工智能产品不断曝光在大众视野并引发使用热潮, 伴生技术的风险也同样引人担忧。使用法律对新型技术进行规制是各国政府与立法者目前重点关注的领域, 我国也不例外。为促进生成式人工智能技术健康发展和规范应用, 国家互联网信息办公室起草了《生成式人工智能服务管理办法(征求意见稿)》<sup>1</sup>, 于 2023 年 4 月 11 日发布向社会各界公开征求意见, 并最终出台《生成式人工智能服务管理暂行办法》(以下简称《管理暂行办法》)<sup>2</sup>, 专门针对生成式人工智能技术进行规制。

对生成式人工智能技术予以法律规制, 一是为了更好地引领行业发展, 二则是出于风险控制的考虑。本文的主要研究对象生成式人工智能生成内容的标识义务, 正是一项被设计用于控制生成式人工智能内容风险的义务。具体而言, 生成式人工智能内容存在着如下风险。

### (二) 生成式人工智能生成内容的风险

#### 1. 侵犯公民权益之风险

生成式人工智能技术在其应用过程中, 不可避免地会使用到包含公民个人生物识别信息的数据集。所谓个人生物识别信息, 指人脸、指纹、虹膜、声纹等能够用以识别自然人身份的经数据化处理后的信息, 具有不可变更性与高度敏感性。作为一种权益的个人生物识别信息, 在此应用过程中, 面临着被侵害的风险, 其中最为显著的就是信息泄露与信息失真两大风险[1]。所谓信息失真是指由生成式人工智能技术制作的具有混淆性或虚构性的内容一经公布、传播, 可能使公众对相关事实产生误认, 进而削弱真实信息的可识别性与可信度的现象。而信息泄露之风险, 则是在应用生成式人工智能技术的过程中, 包

<sup>1</sup>[http://www.whwx.gov.cn/wxd/t202304/t20230411\\_2184451.shtml](http://www.whwx.gov.cn/wxd/t202304/t20230411_2184451.shtml).

<sup>2</sup>[https://www.gov.cn/zhengce/zhengceku/202307/content\\_6891752.htm](https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm).

含于个人隐私的个人敏感信息、生物识别信息等不可避免地作为数据集被调用，在此过程中所涉及的信息泄露风险。

目前，基于现有的生成式人工智能技术，已经可以快速生成大量具有仿真性的图片、视频、音频、文本文件，或完全“无中生有”捏造虚假的信息数据，这完全有可能构成诽谤。这些以公民的基本信息为学习材料合成的，或真或假的或掺真掺假的文本的内容的存在本身就可能对公民的人格权、个人信息自决权、名誉权等构成威胁，遑论经广泛传播以后所可能造成的更为严重的后果。同时，这些合成内容在检测与识别方面也存在着相当程度的困难。比如音画型媒体文件在传播过程中，经常会受到压缩或缩小尺寸<sup>[2]</sup>。这种缩减实际上对于内容的检测造成阻碍，这更是为内容侵权提供了一定程度上的便利。

基于上述高仿真性、难检测性与易于传播的特征，内容对于相关公民权益的侵犯早已不再是耸人听闻，而是确已发生。在杭州互联网法院审理的萧山区人民检察院诉虞某龙个人信息保护民事公益诉讼案中，被告人将从互联网等渠道收集而来的人脸信息，未经许可地利用换脸软件，将之与部分淫秽视频中主体的人脸信息进行替换，伪造生成淫秽视频、图片。并通过网络社交群组出售其合成的违法视频，为他人提供换脸视频的定制服务。法院最终判决被告人立即停止侵权，删除在社交软件群组中留存的所有涉案个人信息，在《法治日报》向社会公众刊发赔礼道歉生命以消除影响进行警示，并处以6万元人民币的罚款。<sup>3</sup>该案表明，生成式人工智能技术一旦与人脸信息、淫秽内容和网络传播相结合，其侵害后果便不再停留于个体权益层面，而会进一步扩散为公共利益风险。随着生成式人工智能技术的不断迭代与更新，社会公众也将面临更为严峻的侵权风险。

## 2. 威胁社会稳定之风险

信息是我们与世界接触的媒介与基础，我们在对事物进行判断时，不可避免地依赖于我们所接收到的信息。相较于对公民个人权益的侵害可能性，深度伪造技术的快速发展，将可能引致严重的社会后果——事实真相的消解。

由假情报(Disinformation)引起的公共社会事件层出不穷。所谓假情报，指公布于众的或透露的有意迷惑人的情报，意在影响他国的大众舆论或政府的信息和情报。生成式人工智能技术在新闻媒体领域，大量地被运用于伪造政客、公众人物的公开演讲。

### (三) 标识义务的重要性与履行困境

标识义务概言之是要求相关主体对人工智能生成内容施加可见或不可见的标识，以达到将人工智能生成内容与一般内容区分开来的目的的义务。将标识义务作为本文的研究对象，是因为标识义务是规制生成式人工智能生成内容风险的最有力措施之一。学者姚志伟认为履行标识义务是一种披露行为。所谓披露，通常是指通过一定的标识方法，向外界特别是向公众揭示特定的内容是由人工智能生成而非人类创造。披露可以使外界区分人工智能生成内容和人类创造的内容，只有先将人工智能生成内容区分出来，才能对内容风险进行规制。从这个意义上而言，披露是人工智能生成内容风险规制的基础性方法。而作为基础性规制方法的标识义务却在构建上存在着一定的问题，这不仅不利于实现立法目的，同时也可能放纵生成式人工智能生成内容的风险滋生，造成不可估量的后果。

## 2. 标识义务体系的构建现状

生成式人工智能的标识义务，是指《管理暂行办法》中所规定的，由生成式人工智能服务提供者所承担的对视频、图片等内容进行标识的义务。该规定实际上是一条转置性规定，《管理暂行办法》规定在履行标识义务时，应当以《互联网信息服务深度合成管理规定》(以下简称《管理规定》)中的规定为准。

<sup>3</sup>参见杭州互联网法院(2023)浙0192民初4563号判决。

《管理规定》具体地规定了标识义务的相关内容<sup>4</sup>，服务提供者是标识义务主要的义务主体，负有隐形标识义务与显著标识义务两大不同种类的标识义务。<sup>5</sup>其中，隐形标识义务要求服务提供者对使用其服务生成或者编辑的信息内容采取技术措施添加不影响用户使用的标识。显著标识义务则可以被分为两个部分，一是主动显著标识义务，要求在提供具有生成或者显著改变信息内容功能的服务并可能导致公众“混淆或者误认”时，应当“在生成或者编辑的信息内容的合理位置、区域进行显著标识”。二是提示显著标识义务，服务提供者在所提供的服务并非具有生成或显著改变信息内容的功能时，应当要提供显著标识的功能，同时提示服务使用者可以进行显著标识。针对服务提供者之外的主体，《管理规定》在第18条规定了任何主体、组织都不得采用技术手段删除、篡改本规定标识的禁止性义务。

除此之外，有关标识义务的要求，在如下两项法律文件中也有体现。《互联网信息服务算法推荐管理规定》要求算法推荐服务提供者在传输算法合成内容之前，应当对其进行显著标识<sup>6</sup>。《网络音视频信息服务管理规定》则要求网络音视频服务提供者和网络音视频服务使用者在利用基于深度学习、虚拟现实等新技术新应用制作、发布、传播非真实音频信息等，应当以显著方式予以标识。不符合标识要求时应当立即停止传输信息，并且补充显著标识。<sup>7</sup>

有关标识义务的规定在我国的人工智能技术相关立法体系中占据了一定的篇幅，总览我国人工智能生成内容标识义务法律的规定，不难发现如下两个现象：第一，强制性的义务大多由服务提供者承担，人工智能服务的使用者仅承担非强制性的标识义务或不承担任何义务，在责任主体的分配上，呈现出以人工智能服务提供者为重心的倾向；第二，所有有关人工智能生成内容标识义务的规定都散见于部门规章之中，法律文件的层级普遍不高，并且规定大多十分模糊，仅概括地描述了各个标识义务，也存在有直接采取引致条款进行规范的倾向。这两大现象则对应到后文中即将进行讨论的有关现行生成式人工智能标识义务在构建上存在的两大问题，即义务内容设置模糊问题、主体责任分配不均问题，在下一节中对这两个问题展开具体的分析。

### 3. 标识义务的构建不足

生成式人工智能生成内容的标识义务见于《管理暂行办法》第12条，如前文所述作为引致条款，履行标识义务的标准应当参照《管理规定》中的有关条文。因此在讨论生成式人工智能生成内容的标识义务时，也将主要围绕《管理规定》中的有关条款进行讨论。

#### (一) 义务内容设置模糊

##### 1. 显著标识义务中“显著”尚无明确法律定义

<sup>4</sup>隐形标识义务参见《互联网信息服务深度合成管理规定》第16条：深度合成服务提供者对使用其服务生成或者编辑的信息内容，应当采取技术措施添加不影响用户使用的标识，并依照法律、行政法规和国家有关规定保存日志信息；显示标识义务参见第17条：深度合成服务提供者提供以下深度合成服务，可能导致公众混淆或者误认的，应当在生成或者编辑的信息内容的合理位置、区域进行显著标识，向公众提示深度合成情况：（一）智能对话、智能写作等模拟自然人进行文本的生成或者编辑服务；（二）合成人声、仿声等语音生成或者显著改变个人身份特征的编辑服务；（三）人脸生成、人脸替换、人脸操控、姿态操控等人物图像、视频生成或者显著改变个人身份特征的编辑服务；（四）沉浸式拟真场景等生成或者编辑服务；（五）其他具有生成或者显著改变信息内容功能的服务。深度合成服务提供者提供前款规定之外的深度合成服务的，应当提供显著标识功能，并提示深度合成服务使用者可以进行显著标识。（[https://www.cac.gov.cn/2022-12/11/c\\_1672221949354811.htm](https://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm)）

<sup>5</sup>《管理规定》的第16条和第17条可以被视为是分别规定了服务提供者应当进行隐形披露和显性披露，这是以服务提供者按条例规定所应当做的标识的种类为依据进行界分的。考虑到条文的具体表述，没有使用“披露”一词，尽管披露一词在涵义上更能够概括义务内容甚至于设立义务的目的。另外在第十七条特别强调了应当进行“显著标识”的情况下，将这两种不同的标识义务命名为隐性标识义务和显著标识义务更为妥当。

<sup>6</sup>参见《互联网信息服务算法推荐管理规定》第9条：算法推荐服务提供者应当加强信息安全管理，建立健全用于识别违法和不良信息的特征库，完善入库标准、规则和程序。发现未作显著标识的算法生成合成信息的，应当作出显著标识后，方可继续传输。（[https://www.cac.gov.cn/2022-01/04/c\\_1642894606364259.htm](https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm)）

<sup>7</sup>参见《网络音视频信息服务管理规定》第11条：网络音视频信息服务提供者和网络音视频信息服务使用者利用基于深度学习、虚拟现实等新技术新应用制作、发布、传播非真实音视频信息的，应当以显著方式予以标识。（[https://zwgk.mct.gov.cn/zfxxgkml/zcfg/gfxwj/202012/20201204\\_906347.html](https://zwgk.mct.gov.cn/zfxxgkml/zcfg/gfxwj/202012/20201204_906347.html)）

在上一节中已经明确了《管理规定》中规定服务的提供者负有显著标识义务。在《管理规定》即《管理暂行办法》颁布以前，涉及人工智能生成内容显著标识的有关立法，散见《网络音视频信息服务管理规定》和《互联网信息服务算法推荐管理规定》之中。这两部文件也都提到了应当对相关的算法合成内容进行显著标识，但前述四部文件，并没有任何一部法律文件对显著标识中的“显著”作出过明确的定义，这是标识义务在设置时的一大缺陷。

“显著”一词，以《新华词典》作解，意为“非常明显”。因此从字面意义上理解，在合成内容之上进行显著标识的目的，在于以非常明显的方式提示服务的使用者或一般信息的接受者该内容系生成式人工智能技术生成，以达提示目的。进而将合成内容与普通内容进行区分，在信息传播的源头就削弱虚假信息传播的风险。但缺少对于“显著”的法律界定，将直接导致服务提供者履行标识义务时缺乏一个可行的参照。即设置的提示标识需要满足何种具体条件才算是具备了应有的显著性？缺乏针对显著性的定义，将导致服务提供者对显著标识义务的履行与履行程度的认定都面临相当的困难。

## 2. 义务的履行范围缺少明确界定

如前节所述，目前标识义务制度的规定过于笼统，尚未对“显著标识”作确切定义。除此之外，针对“可能导致公众混淆或误认的”信息的判别标准，在现行法律制度之下也处于缺位状态。从文义可以得知，服务提供者仅需对可能导致公众混淆或者误认的信息内容履行显著标识义务，那么在可能导致混淆、误认的情况的判断标准缺位的情况下，服务提供者的义务履行范围则面临着无法确定或无限扩大的困境。在这种情况下，服务提供者或可能出于避免合规风险的考虑而无差别地对所有的生成内容都进行标识，这将极大程度上地提高运营成本并也缩减用户的服务体验。

大众对于信息的辨别能力各有不同，虚假信息、合成信息给不同个体所带来的影响也千变万化。对接触合成内容的一方来说，对“显著标识”“混淆”或“误认”的判断具有相当的主观色彩。这种情况下如果不事先就对需要履行标识义务的具体范围进行确定，将导致在实践中，很难判断义务主体是否完全履行了其所负有的标识义务。笔者认为如此构建标识义务制度是否过分流于形式了。并且当涉及侵权举证的情景时，则不论是相关服务的提供者还是其使用者，都会面临相当的举证困境。

## (二) 主体责任分配不均

主体责任最初兴起于企业安全生产领域，之后逐渐扩展至多个经济领域。近年来由于互联网经济的纵深发展，将主体责任理论引入该领域乃是必然趋势。责任，一般是指当事人违反相关行为规范而应当承担的不利后果。但主体责任中的责任主要是指主体应当承担的积极作为和不作为的义务，学者刘权认为主体责任也可以被称之为主体义务<sup>[3]</sup>。本节所讨论的主体责任也是更多指向主体义务的方面，而不是违反规范而应当承担的不利后果方面。

### 1. 服务使用者义务承担畸轻

《管理暂行办法》中，要求提供者承担网络信息内容安全生产者责任，对生成内容按照《管理规范》中的相关规定进行标识。但对于服务使用者，却并没有做出强制性的规定。使用者作为服务、生成式人工智能技术应用过程中的重要参与主体之一，将不可避免地参与到各个环节之中。使用者除使用服务进行内容制作之外，在社交媒体平台持续蓬勃发展的背景之下，同时也承担着信息的传播者的任务。信息传播过程中，包括传播者、受传者、讯息、媒介和反馈五大要素<sup>[4]</sup>。其中的传播者即承担着搜集、加工、制作传递信息的职能。传播者逐渐成为传播活动的重要主体，极大程度上地推动着舆情的产生。据此，在规范约束服务及其衍生物的过程中，对服务使用者的规范应是立法过程中的应用之义。不同主体，所需承担的主体责任因其在现实过程中所发挥的作用的不同而不同。哈特认为，责任是一个复杂的概念，具体而言又主要具有五种责任，即角色责任、因果责任、法律责任、道德责任和能力责任<sup>[5]</sup>。上述五种责任虽各有各的不同内涵，但不论主体究竟扮演何种角色，都应承担其应有的责任。并且也只有在进行行为

主体都对相应责任予以充分履行的基础上，各个环节才能协调一致，秩序才能得以维护。

但在现有法律制度之下，针对服务使用者有关义务的规定，仅存在于《中华人民共和国民法典》第1194条中有关网络用户侵权责任的条款及《互联网信息服务管理规定》第18条：“要求任何组织和个人不得采取技术手段删除、篡改隐匿标识”之中。相较于服务应用过程中的其他责任主体，如服务提供者的有关义务设置，对服务使用者的义务设置明显过于宽松，并不利于有效地调整待规范客体。

## 2. 服务提供者义务承担畸重

相比之下，服务提供者不仅要承担网络信息内容安全主体责任，承担显著标识义务与隐形标识义务，还要承担输入信息和使用记录的保护义务、建立投诉、举报机制、承担监督检查等平台安全义务。

由生成式人工智能服务提供者负担这部分义务，是弥补数字时代法律缺陷的迫切需求。赋予其相应的主体责任，可以缓解成文法所固有的僵化、滞后的缺陷，规范“无法可依”的平台行为，并且能促进平台发挥主观能动性实现预防式的治理。赋予其主体责任，即上述诸多义务，有助于平台从源头开始合理约束自身需求，进而使管控措施在最大程度上地发挥效用。这是因为单纯强调通过追究不利后果责任对平台进行事后惩戒，对于促进平台积极履行义务保障各方主体权益的助力不大。而主体责任及其相关义务，是一种积极责任。积极责任要求主体积极作为，积极履行法定义务、道德义务与约定义务，在一定程度上能起到治本的作用。除上述两点外，赋予服务提供者以主体责任，也是推进监管的必然要求，有利于促进平台自治。强调落实平台自身的主体责任与相关义务，实际上是要求平台自身加强自我规制，而非全然以政府介入监管的模式，从而可以消除过度监管、监管不到位等“反法治现象”。

然而，过度强调平台的主体责任，也会导致平台等运营成本过大。平台作为市场主体，不应被要求不计成本地维护公共利益，有悖公平原则。例如《互联网信息服务管理规定》第7条所谓“落实安全主体责任”就呈现出对平台等安全保障义务扩大化适用的趋势。但技术发展固有的弊端，即破解技术的发展速度往往慢于技术发展的速度这一特征，将导致平台在技术层面难以对致害性进行完全的管控。此时如赋予其过分的法律义务负担，将增大网络服务的运营成本，则不利于互联网技术的进一步发展。

## 4. 对生成式人工智能标识义务制度的完善建议

学理上将主体责任分为法律责任、契约责任与道德责任。其中，仅法律责任具有国家强制力，因为其在规定了义务的同时也规定了违反义务的不利后果。而道德责任属于道义上的要求，法律往往只会规定最低限度的道德责任。三者之间并不存在泾渭分明的界限，但是我国在构建标识义务制度时，并没有在规定义务的同时规定违反义务的不利后果

这也就导致了生成式人工智能标识义务在构造上是非刚性的，甚至呈现出一种道德责任的特征。施加这项义务并不会直接提高服务提供者的运营成本，因为服务提供者可以不履行义务。但是规范性法律文件当然起到引领性的作用，这是法律的主要作用之一。

为了让标识义务更好地发挥其作用，对标识义务制度的完善提出如下建议。

### (一) 针对显著标识义务内容的完善建议

#### 1. 明确“显著方式”的定义

以《中华人民共和国商标法》与《商标审查审理指南》中围绕商标显著性概念所做的一系列阐述为参照，在未来对内容标识义务做进一步细化规定时，要想最大程度上发挥标识义务对阻断依托于技术所生产的虚假信息传播的作用，有必要对显著方式标识作进一步的定义。这是因为只有以显著方式进行了标识的内容，才能最大程度上地与其他普通信息区别开来，从而起到实质上的提示作用，可知显著标识义务的构建基础即在于显著的定义。

进一步明确显著方式的定义，除了对显著进行法律上的详细描述之外，笔者认为还可以采取从反面

明确的方式，即明确何种标识方式是不属于“显著方式”的，也被称为是负面清单的模式。例如专列条款列举未在合理位置进行标识，或采取的标识类型非嵌入性水印，易被消除等情况均不符合显著方式标识的规定，以此从反面对标识义务者、审查人员进行指引，进而使抽象笼统的规定更便于适用。

明确上述显著方式相关概念，是为了使审查人员能够更好地掌握显著方式标识的内涵与外延，进而能够更准确地判断其是否符合标准地履行了内容的标识义务，从而保证标识义务能够最大程度上地发挥其作用，而不是沦为形式与表面；另一方面也提醒服务提供商、服务的使用者在设计、使用产品时，能有更有力、更规范的操作方案参考，有助于实现行业自律。通过行业自律来实现新型技术的治理是我国目前的基本立场，也是成本最小、最富有经济效率的规制措施之一。

## 2. 制定具体的标识义务实施细则

目前针对标识义务的规定，仅笼统地规定为应当以显著方式对内容进行标识。但针对具体应当如何履行标识义务，即标识义务的实施细则，未能有更为详细的规定。制定具体的标识义务实施细则，有助于显著标识义务人更规范充分地履行义务，以及部分地消解具体侵权纠纷发生时需承担有关举证责任的权利人的举证困难。

首先，在现有条例已经对不同类型的内容作出明确的界分基础上，应当此为基础，分别规定不同种类的内容的标识义务实施标准。其次，在具体的规定上，尽量言简意赅地规定具有可操作性和实践性的标识方法，具体到采取何种技术标准的水印、应当做出包含合成、修改具体情况的声明等。如针对“人脸生成、人脸替换、人脸操控、姿态操控等人物图像、视频生成或者显著改变个人身份特征的编辑服务”此类图像、视频类内容，应当以嵌入式可移动水印的方式，对其进行显著标识。同时，要求义务人作出明确的书面声明，简要描述对有关图片、视频中的相关元素所作出的具体合成、修改内容。

设置标识义务是从源头上杜绝虚假信息传播风险的最有力措施之一，我国应当尽快完善有关的实施细则及配套法规，促进落实标识义务的履行，从而营造更好的互联网环境。

### (二) 针对标识义务主体分配不均的完善建议

有失公平的主体责任分配，不利于技术行业发展的同时，也助长其他责任主体的惰性。不论各责任主体究竟扮演着何种“角色”，都应公平地承担其应负的责任<sup>[6]</sup>。在制定法律法规时，所呈现出的有偏颇的责任分配倾向，应当通过重新分配相应的主体责任来达到的平衡的目的。

#### 1. 对使用者增设标识义务及征求多方同意义务

以法律规范服务使用者、内容制作者的行为，能从源头上有效治理现有的内容乱象。笔者认为，应当在赋予服务提供者以强制义务的同时，也规定服务使用者履行相同的显著标识义务。通过法律规定，促使服务使用者在利用服务提供者所提供的服务制作内容时，就主动地对其予以符合要求的标识，作有关其所制作的内容的主动披露。

仅赋予服务提供者以标识义务，在加重服务商运营成本的同时，也存在着漏管、监管不到位的可能。技术作为流行的新型技术，具有前文所述的诸多显性风险，但尚未能引起大众的关注。而现有的《互联网信息服务管理规定》仅规定服务提供者应当提示使用者对生产的内容进行标识，但并未直接赋予其一致的标识义务。因此，使用服务进行合成的行为人对内容进行标识、披露的意识是相当薄弱的。但法律具有激励行为人的功能，通过立法赋予服务使用者以标识义务，充分发挥法律的权威性，令使用者意识到对内容进行标识的重要性，能够有效地促进标识义务的履行。从而更好地发挥平台自治的作用，以达规制技术的最主要目的之一——防止虚假信息的传播。

除将显著标识义务主体范围扩大至服务使用者外，还可对服务的使用者增设征求多方同意义务。要求使用者在利用技术制作合成内容时，应当尽可能取得其所利用的用以合成的内容的有关权利人的授权。设置此种义务有助于保护权利人的知情权，改善信息不对称的情况，也有利于在出现侵权情况时被侵权

权利人的举证。

## 2. 建立行业自律机制缓解服务提供者义务负担

通过建立合理的行业自律机制,将有效地减轻服务提供者的负担,从而达到使得主体责任分配变得相对均衡的目的。互联网行业自律,是指从事互联网各领域如研发、生产、维护等活动的相互独立的主体,通过制定互联网行业规约、规章,在法律缺位或缺乏详细实施细则的情况下,自发地对行业进行规制约束的一种治理模式[7]。

目前,关于互联网的自律机制统共分为三大类,分别是政府主导模式、政府监督模式与政府与行业共同调控模式。在针对技术领域的规制上,我国显然采取的是政府主导模式的行业自律,即不重视实质上的行业自律,而采取一定的立法措施与技术手段,赋予部分法律主体以强制性义务,来对相关内容进行审查与过滤。但法律本身的滞后性与民众对政府干预中所可能出现的争议性执法活动的排斥心理,是技术乃至互联网技术在治理过程中的一大阻力。除此之外,现有立法体制也过分地将针对技术的有关监管义务倾斜地施加在服务提供者之上,这样的分配结果有失公平。

在治理互联网服务领域,将政府主导模式向政府与行业共同调控模式逐渐过渡,可以有效缓解上述问题。首先,采取这种治理模式政府与行业自律组织合作监管某一具体领域的方式,能够充分发挥行业内部的专业技术、资源优势,当各主体都有序地参与到治理过程中时,个别义务主体的责任负担也就当然地减轻了,从而缓解畸重地将监管义务倾斜到某一类特定主体之上的现状。采取政府与行业共同调控模式,能够使得监管措施能够更灵活地应对技术的不断更迭发展。考虑到技术的新颖性与复杂性,在运作过程中涉及多方主体的技术的治理并非简单的技术问题,而是复杂的社会系统工程,应当在立法明确主体功能范围的同时,给予各主体以法律范围内的相当的自治权,令其在具体治理过程中能有灵活调整、回旋的余地,而非沿用传统的“大包大揽”的思路。

## 参考文献

- [1] 万志前,陈晨.深度合成技术应用的法律风险与协同规制[J].科技与法律(中英文),2021(5):85-92.
- [2] Zhang, H., Zhu, Z., Sun, L., Mao, X. and Wang, Y. (2020) A Detection Method for Deepfake Hard Compressed Videos Based on Super-Resolution Reconstruction Using CNN. *Proceedings of the 2020 4th High Performance Computing and Cluster Technologies Conference & 2020 3rd International Conference on Big Data and Artificial Intelligence*, Qingdao, 3-6 July 2020, 98-103. <https://doi.org/10.1145/3409501.3409542>
- [3] 刘权.论互联网平台的主体责任[J].华东政法大学学报,2022,25(5):79-93.
- [4] 李菲,王旭昇,张婉秋.舆情场中传播者的信息行为及监管对策[J].长春师范大学学报,2021,40(6):191-194+197.
- [5] H.L.A.哈特.法律、自由与道德[M].钱一栋,译.北京:商务印书馆2021:1-128.
- [6] 张凌寒.深度合成治理的逻辑更新与体系迭代——ChatGPT等生成型人工智能治理的中国路径[J].法律科学(西北政法大学学报),2023,41(3):38-51.
- [7] 许玉镇.网络治理中的行业自律机制嵌入价值与推进路径[J].吉林大学社会科学学报,2018,58(3):117-125+206.