

知网个人查重服务报告单 (全文标明引文)

报告编码:BC20260429064942370356941

检测时间:2026-04-29 06:08:24

篇名: 标识义务视角下AIGC的法律规制路径探析

作者: 相昀含

检测类型: 毕业设计

比对截止日期: 2026-04-29

检测结果

去除本人文献复制比: 4.1% 去除引用文献复制比: 4.1% 总文字复制比: 4.1%
单篇最大文字复制比: 1.0% (查重试一下)

重复字符数: [364] 单篇最大重复字符数: [92] 总字符数: [8852]

4.1% (364) 4.1% (364) 标识义务视角下AIGC的法律规制路径探析_第1部分 (总8852字)



(注释: 无问题部分 文字复制部分 引用部分)

1. 标识义务视角下AIGC的法律规制路径探析_第1部分 总字符数 8852

相似文献列表

去除本人文献复制比: 4.1% (364) 去除引用文献复制比: 4.1% (364) 总文字复制比: 4.1% (364)

1	查重试一下 佚名 - 大学生论文联合比对库 - 2024	1.0% (92) 是否引证: 否
2	3.0查重 佚名 - 大学生论文联合比对库 - 2024	1.0% (92) 是否引证: 否
3	人工智能生成内容可版权性研究 佚名 - 大学生论文联合比对库 - 2024	1.0% (92) 是否引证: 否
4	论人工智能生成物的著作权保护——以“2019年腾讯诉盈讯侵犯著作权及不正当纠纷案”为例 4 佚名 - 大学生论文联合比对库 - 2024	1.0% (89) 是否引证: 否
5	本科毕业论文注释版 佚名 - 大学生论文联合比对库 - 2024	1.0% (89) 是否引证: 否
6	《AIGC的数据权保护模式研究——以ChatGPT为例》(定稿1950修4月7日) 佚名 - 大学生论文联合比对库 - 2024	1.0% (88) 是否引证: 否
7	1查重-普通版 佚名 - 大学生论文联合比对库 - 2024	0.9% (76) 是否引证: 否
8	人工智能生成内容著作权保护问题研究 于立芳 - 兰州大学硕士学位论文 - 2020	0.7% (66) 是否引证: 否
9	AIGC技术对新闻媒体创作的影响探究以“百度AI开放平台”的媒体技术应用为例 佚名 - 大学生论文联合比对库 - 2024	0.5% (47) 是否引证: 否
10	AIGC技术在动画创作中的应用与影响 佚名 - 大学生论文联合比对库 - 2023	0.5% (47) 是否引证: 否
11	AIGC在数字摄影中的应用研究 佚名 - 大学生论文联合比对库 - 2024	0.5% (47) 是否引证: 否
12	AI绘画、视频、音乐等作品著作权归属如何界定?_财经头条_新浪财经 佚名 - 互联网文档资源 - 未知	0.5% (41) 是否引证: 否
13	MOBA游戏整体画面的作品属性研究	0.5% (41)

	杨琴 - 海南大学硕士学位论文 - 2021	是否引证: 否
14	人工智能生成物刑法保护的基础和限度 刘宪权; - 华东政法大学学报 - 2019	0.5%(41) 是否引证: 否
15	结果视角下人工智能生成物的保护路径检讨 郑远民;贺栩溪; - 科技与法律 - 2020	0.5%(41) 是否引证: 否
16	生成式人工智能作品性认定与法律保护 佚名 - 大学生论文联合比对库 - 2024	0.5%(41) 是否引证: 否
17	人工智能生成物著作权问题研究 刘美媛 - 吉林大学硕士学位论文 - 2021	0.5%(41) 是否引证: 否
18	中国保护知识产权网 佚名 - 互联网文档资源 - 未知	0.4%(35) 是否引证: 否
19	当AI像人一样互动,著作权法如何鼓励创作? 人工智能_新浪财经_新浪网 佚名 - 互联网文档资源 - 未知	0.3%(28) 是否引证: 否
20	最后一次查重 佚名 - 大学生论文联合比对库 - 2023	0.2%(19) 是否引证: 否
21	论汇编作品的独创性 朱翔 - 华东政法大学硕士学位论文 - 2017	0.2%(19) 是否引证: 否
22	风险社会语境下刑法功能的反思 赵静 - 甘肃政法学院硕士学位论文 - 2016	0.2%(14) 是否引证: 否
23	刑法谦抑理论研究 李磊明 - 吉林大学博士学位论文 - 2020	0.2%(14) 是否引证: 否
24	民事与刑事责任区分问题研究 安之楠 - 西南政法大学博士学位论文 - 2018	0.2%(14) 是否引证: 否
25	论积极刑法观在我国的确立及展开 严秀琴 - 甘肃政法大学硕士学位论文 - 2022	0.2%(14) 是否引证: 否
26	风险刑法的自我抑制 刘东行 - 山东大学硕士学位论文 - 2013	0.2%(14) 是否引证: 否
27	期待可能性理论再讨论 佚名 - 大学生论文联合比对库 - 2024	0.2%(14) 是否引证: 否
28	《合同诈骗中的刑民交叉问题研究》(二次降重版) 佚名 - 大学生论文联合比对库 - 2024	0.2%(14) 是否引证: 否
29	刑法作为风险规制措施的局限性 张伦伦 - 天津商业大学硕士学位论文 - 2017	0.2%(14) 是否引证: 否

原文内容

标识义务视角下AIGC的法律规制路径探析

摘要：生成式人工智能的快速发展，使AIGC深度嵌入信息生产与传播，其生成内容的“不可见性”对传统法律体系形成系统性挑战。本文认为，现行困境的根源在于人类创作与AI生成之间缺乏可识别的制度边界。标识义务作为一种信息披露型规制手段，以知情权保障与风险预防为法理基础，具有成本低、覆盖广、操作性强的优势，可在AIGC规制体系中发挥“前置性锚定”功能。借鉴欧盟风险分级立法与美国版权登记披露等域外经验，本文主张构建以风险分级为基础的差异化标识规则，明确基础模型开发者、内容生成者与传播平台的分层义务体系，建立显性与隐性标识相结合的技术标准，并完善民事、行政、刑事梯次递进的责任机制。以标识义务为核心支点，有助于在促进技术创新与保障公共利益之间实现动态平衡。

关键词：AIGC；标识义务；法律规制；知情权；风险分级

引言

2022年以来，以ChatGPT、Midjourney、Stable Diffusion为代表的生成式人工智能技术实现了跨越式发展，**AIGC（Artificial Intelligence Generated Content, 下文中称AIGC）**，也即人工智能生成内容，迅速从实验室走向商业应用，深刻改变着内容生产的底层逻辑。根据中国信息通信研究院发布的《2023年全球人工智能产业研究报告》，截至2023年底，全球AIGC市场规模已超过150亿美元，其中中国市场的占比约为25%，显示出AIGC在全球范围内的快速渗透与应用趋势。基于深度学习和自然语言处理技术的AI模型不断取得突破，使得AI不仅能够模仿人类的语言表达，还能生成高质量的文本、图像、音频乃至视频内容，在文学创作、艺术设计、影视制作等领域得到了广泛应用。AIGC的快速崛起对传统法律体系构成了全方位的挑战。传统法律体系在面对相关问题时表现出明显的滞后性与适应性不足，亟需构建适应AI时代特征的法律规制路径。

本文以AIGC引发的法律属性认定、版权归属模糊、侵权责任不明及信息真实性缺位等规制困境为背景，围绕生成式人工智能的技术特征与社会影响，论证标识义务作为核心规制手段的正当性。通过厘清AIGC的法律属性与责任认定逻辑，明确标识义务的前置锚定地位及其与版权保护、数据透明、责任追溯的协同关系，借鉴域外经验，结合我国产业与社会治理需求，分类分级

构建AIGC标识义务的制度框架与实施机制。旨在为完善我国AIGC法律规制体系提供理论支持和实践参考，推动人工智能技术在合法、合规的前提下健康发展，在促进技术创新的同时，有效防范潜在的法律风险和社会伦理问题，实现技术进步与权利保护、产业发展与社会治理的动态平衡。

一、AIGC的法律属性认定与规制困境

（一）AIGC的法律定性争议

1. 创作主体认定困境

AIGC法律定性问题的核心是创作主体的认定。在传统著作权法的理论框架下，“作者”始终是以自然人为中心构建的，作品被视为创作主体人格的外化表现，著作权的授予以人类的智力劳动为前提。然而，生成式人工智能技术的迅猛发展，深刻颠覆了这一传统假设，使得创作主体的法律认定陷入前所未有的困境。

AIGC的产生至少涉及三类主体：算法模型的设计者、AI工具的提供者，以及具体生成指令的使用者。三类主体对最终生成物均有不同程度的贡献，却无法被任一单一主体完整涵盖。

杨立新、黄露的研究指出，AI的生成过程受到三重人类决定力的约束：其一，训练语料的约束性，即用以训练的数据种类和内容完全由人为控制，生成物不可避免地受到训练语料的范围限制；其二，算法机制的基础性，设计者通过编程决定了AI的底层运行逻辑，即便深度学习表面呈现出自主性，也无法脱离人类赋予的算法框架；其三，指令的决定性作用，利用者通过输入提示词、选择参数和风格，直接决定了生成物的主要构成与表现形式[12]。由此可知，AI本质上是人类意志的延伸，主体认定困境源于多元主体的责任分散。

2. 作品性认定的理论争议

AIGC是否构成《著作权法》意义上的“作品”，是理论界与司法实践长期争论的核心命题。围绕这一问题，学界目前主要形成了否定说与肯定说两种基本立场。

否定说的核心论据在于：《著作权法》第3条规定，作品须为“文学、艺术和科学领域内具有独创性并能以一定形式表现的智力成果”，而“智力成果”的认定以创作主体具有自然人身份为前提。王迁教授是否定说的代表性论者，其认为AIGC只是根据算法进行运算，不同于著作权法意义上的“创作”，正如猕猴自拍照因缺乏人类作者而不受版权保护，AI的“运算”行为同样不能等同于具有独创性智力投入的“创作行为”。该观点进一步指出，若将人工智能生成的内容视为用户的独立创作作品，则以此为由提交学位论文的学生也应被视为论文作者，这将明显违背学术规范，由此形成“悖论”，暴露了肯定说在理论逻辑上的内在矛盾[9]。

肯定说则主张应从生成内容的客体属性出发进行判断，而非拘泥于创作主体的身份认定。吴汉东教授从法理反思层面提出，生成式人工智能的出现要求著作权法重构“独创性”认定的理论框架，应当在“人机合成创作”中坚持“人本主义原则”，承认机器作者的事实创作地位，同时建立“创作主体与权利主体相分离的二元结构”，即承认AI具有拟制作者地位，但著作权人仍须是具有意思能力的自然人或法人[11]。蒋舸则从“用户独创性表达”视角切入，指出AIGC既包含来自机器的选择，也包含来自人类用户的选择，当后者满足独创性表达要求时，AIGC便足以被认定为作品；判断独创性应“关心人贡献了什么，而不是工具贡献了什么”，“关心人贡献的实质，而非纠结于其形式”[4]。

3. 司法实践认定的演变与分歧

我国法院对AIGC的作品性认定问题的裁判立场经历了从否定到有条件肯定的演变。

2018年，北京互联网法院在“菲律宾律所诉百度公司案”中，认定AI自动生成的分析报告不构成作品，理由是该报告为软件使用者“一键生成”，未能体现自然人的独创性智力投入。

2019年，深圳市南山区人民法院在“腾讯公司诉盈讯公司案”中则作出了截然不同的认定，认为涉案财经报道文章系由包含编辑团队、技术团队在内的主创团队利用AI辅助完成，其特定表现形式体现了创作者的个性化选择与安排，构成受保护的文学作品，并采用“两步判断法”作为裁判思路。

2023年，北京互联网法院在“李某诉刘某AI绘画侵权案”中首次认定，用户通过输入提示词、调整参数等方式利用AI生成的图片构成美术作品，用户为著作权人，该判决被认为是国内首例认可AI生成图片著作权的生效判决。

但进入2025年，苏州市中级人民法院在“丰某某诉朱某某案”中再度否定了简单“提示词+生成”方式的作品属性，认定该案系我国首例“AI文生图不构成作品案”，由此表明司法标准仍在动态调适之中。

（二）现行法律规制的主要困境

1. 版权归属模糊

现行《著作权法》以“作者”为权利归属的起点，但面对AIGC中的多元主体结构，既有的归属框架无法给出清晰回答。实践中，AI工具的服务协议则进一步加剧了归属的混乱，契约主导的归属方式缺乏明确法律依据，且服务协议效力边界模糊，难以形成稳定的权利预期。

2. 侵权责任认定复杂

一是训练数据阶段的侵权认定困难。生成式AI模型在训练阶段大量使用受版权保护的作品，大模型公司通常不完整披露训练数据集来源。而传统著作权侵权认定采用的“接触+实质性相似”标准，正面临双重障碍：原告难以证明被告“接触”其作品，且生成内容经过算法深度学习转化，实质性相似的判断标准受到严峻挑战。姚叶的研究进一步指出，由于立法刚性、司法适用受限及技术发展超前，现行“著作权限制与例外”制度在应对AIGC侵权问题时同样呈现失灵状态[13]。

二是生成内容阶段的责任主体模糊。当AIGC侵犯他人权利时，责任链条中的各主体往往相互推诿：算法设计者以“仅提供技术工具”为由主张免责；平台服务商援引“避风港原则”主张不承担主动审查义务；内容使用者则以“不知晓生成内容侵权”为由抗辩。这种责任分散结构使权利人的救济途径受阻，违法成本过低，难以形成有效威慑。

3. 信息真实性与用户知情权缺失

在信息真实性方面，AI模型通过学习训练数据的概率分布生成新内容，其底层逻辑是拟合而非验证事实，当训练数据存在偏见或信息缺口时，模型可能生成看似语义连贯却与客观事实存在偏差的内容。王岭等学者指出，“在缺乏有效内容审核机制的情况下，人工智能生成内容面临真实性与可靠性挑战，可能引发公众认知混乱、社会争议乃至群体对立等问题”，并可能催生虚假信息、智能伪造和伦理失序等新型风险[8]。

更为严峻的是AIGC的“不可见性”问题：随着生成技术的持续进步，AIGC与人类创作内容在表现形式上日益趋同，难以可靠区分。这一特征在高度敏感的场景中可能引发更为严重的后果。而在用户知情权保护方面，用户在接触AIGC时，通常无法获知内容的生成方式、数据来源、算法逻辑等关键信息，这直接导致其丧失自主判断内容可信度的事实前提。这不仅构成对用户知

情权的侵害，也使得事后的权利救济缺乏事实基础，当用户因信赖AI生成的虚假信息而遭受损失时，往往无从举证其决策是在受欺骗情况下作出。

二、标识义务的正当性基础与制度功能优势

从制度设计的角度审视，对AIGC设定标识义务，正是应对上述多重困境的关键前置性手段：通过强制要求AIGC明确标注来源，一方面能保障用户获取关于内容性质的基础信息，使其能够据此作出理性决策；另一方面，标识记录也为事后的侵权追溯、责任认定与监管执法提供了可操作的制度基础。这也弥补了现行规制框架中最明显的制度空缺之一——在缺乏有效标识义务的条件下，AI内容的“隐形传播”将使所有规制手段都失去发力支点。正因如此，我们有必要从法理层面系统阐释标识义务的正当性基础，并进一步剖析其在整体规制框架中所展现的独特制度与功能优势。

标识义务的理论基础

1. 用户知情权保障与信息透明原则

知情权是现代法律体系中具有基础地位的公民权利，其核心在于保障权利主体对涉及自身利益的信息享有获取、了解与判断的权利。在AIGC广泛渗透信息生态的背景下，用户是否知悉其所接收内容来源，直接关系到其信息判断的准确性与意思自治的完整性。

信息透明原则与知情权保障相互印证，共同构成标识义务的首要理论依据。该原则要求信息的生产者与传播者在内容发布环节充分披露内容的生成来源与生成方式，以确保信息接收方能够在知情基础上作出理性判断。易耀森对我国160种医学期刊的调研结果显示，超过60%的医学期刊在防范AIGC不当使用方面存在责任主体模糊问题，相当比例的学术期刊尚未建立针对AIGC的披露声明制度，由此暴露出信息透明原则在制度落地层面的严重缺位[14]。这一现象绝非学术出版领域独有，而是整个AIGC应用生态系统中信息透明缺失的缩影与预警。因此，将标识义务制度化，是保障用户知情权、贯彻信息透明原则的逻辑必然。

2. 社会信任维护与风险防范原则

当大量AIGC以“人类原创”面目流通，社会公众对信息来源的判断能力与信赖基础将遭到根本动摇，进而引发认知信任的结构性危机。邓建鹏的研究中揭示，AI绘画的大量侵权现象已导致创作者群体对生成式人工智能技术产生强烈的信任危机，其核心症结正在于缺乏有效的来源标注机制，使得人类原创作品与AIGC之间的边界难以识别与区分[2]。这一现象表明，信任危机的修复依赖于识别秩序的重建，而标识义务恰恰是重建工作的起点与前提。

风险防范原则作为现代法律规制的重要方法论基础，主张在重大技术风险尚未充分具现之前即采取前置性规制措施，以避免事后补救的高昂社会代价。在此基础上，标识义务通过强制要求内容来源的透明披露，使受众得以事先启动必要的甄别与批判性阅读，从而在风险前端形成认知过滤屏障，有效压缩虚假信息的传播空间与社会危害。这一前置性的制度安排与风险防范原则的立法逻辑高度契合，构成了标识义务正当性的重要规范来源。

（二）标识义务的制度优势

在AIGC法律规制体系中，标识义务之所以居于前置的关键位置，源于其独有的制度优势：实施成本较低、适用范围广泛、实操性强，且无需直接介入技术研发与内容生成环节，即可对信息传播秩序实现系统性规范与矫正。此外，一旦AIGC未经标识而引发侵权或损害，标识义务的缺失本身即可作为过错认定的重要依据，从而在既有侵权责任归责框架内为受害方提供可预期的救济路径。陈兵明确指出，AI服务提供者未履行标识义务，可被直接推定存在主观过错，并由此承担相应的民事侵权责任[1]。这种“标识即义务、缺标即过错”的责任逻辑，赋予了标识义务在规制体系中无可替代的地位。

（三）标识义务的多方协同功能

1. 厘清版权保护的边界

在缺乏来源标注的情况下，人类创作成分与AI生成成分之间的边界难以厘清，权属认定由此陷入困境。强制标识义务的确立，将为区分人类创作与AI生成提供制度性前提，使版权归属问题的个案判断具有更为清晰的事实基础。

2. 配合版权注意义务

版权注意义务是指在版权登记、使用、保护及网络服务中，权利人或平台应采取合理措施避免侵权并维护合法权益的行为规范。一方面，已标识的AIGC便于权利人识别潜在侵权；另一方面，服务提供者主动履行标识义务的行为，亦可作为其尽职尽责、无主观过错的抗辩依据，有助于在责任链条的认定中厘清各方义务边界。

3. 保障数据透明性

强制标识义务保障了内容层面的透明，而其规制溢出效应则体现于对训练数据与算法过程的间接透明推动。许茂恒的研究揭示，要求AIGC提供者在法律上承担内容标识与溯源责任，将倒逼其在技术上建立更为完整的数据来源记录与算法透明机制，以应对潜在的合规核查与举证需要[10]。这一机制传导路径，使得标识义务超越了简单的信息公示功能，成为推动AI全链条数据合规的制度杠杆，进而与数据治理体系形成正向反馈循环。

三、标识义务的域外经验考察与启示

（一）欧盟《人工智能法案》的风险分级与全链条责任模式

欧盟于2024年正式通过《人工智能法案》，构建了迄今为止全球最系统的AI综合性立法框架，其核心是基于风险等级的分类分级监管机制，将AI应用划分为四个风险层级，并施以差异化的合规义务。

在标识与透明性义务方面，《人工智能法案》明确要求：生成式AI系统所生成的文本、图像、视频等内容必须以机器可读的方式标注；在互动型AI场景下，必须在交互开始时向用户明确披露其正在与AI交互，不得误导用户将AI误认为真实人类。

与此同时，欧盟《通用数据保护条例》从数据保护维度对AI训练数据的合法性、用户个人数据的处理透明度提出了严格要求，形成了与《人工智能法案》相互呼应的双重约束机制。

欧盟模式的优点在于体系完整、层级分明、强制性强，但其严格的合规要求也引发了对创新阻碍效应的担忧。

（二）美国与国际技术标准的行业实践

与欧盟的系统性立法路径不同，美国在联邦层面尚无针对AIGC的综合性立法，主要依赖版权局行政执法和联邦法院判例法形成裁判规则。

美国版权局通过2023年发布的《版权登记指南：包含人工智能生成材料的作品》，确立了“披露义务”：版权登记申请人有义务披露其提交注册的作品包含AIGC，并说明人类作者对作品的具体贡献。这一做法实质上将标识义务嵌入了版权注册程序，以“信息披露换权利保护”的机制引导申请人主动标注AI生成来源。

在州立法层面，加利福尼亚州、德克萨斯州等多个州已相继出台针对AI合成内容的专项立法，要求在选举相关内容中强制标

注AI生成来源，并对违规行为设置了罚款与刑事责任。

美国模式的优点在于灵活性强、创新空间大，但联邦层面立法缺失导致各州标准不一，跨州执法困难，整体规制效果较为有限。

（三）域外经验对我国标识制度构建的启示

其一，提升立法位阶，推动体系化规制立法。就我国而言，现行规制体系以部门规章与规范性文件为主，立法层级较低，规制效能有限。在条件成熟时，应积极推动《人工智能法》的专项立法进程，并在此基础上相应修订《著作权法》相关条款，从制度层面明确AIGC的法律定性与权利归属规则，以填补现行立法体系中的制度空白。

其二，推动标识义务与权利保护机制的有机衔接。我国可参考美国版权局将披露义务嵌入版权登记程序的制度设计，探索建立“标识—确权—保护”的联动机制：以完成规范性标识为取得著作权保护的程序性前提条件，将标识合规义务的履行与版权权利的生成相挂钩，从而在强化义务约束的同时，为权利人提供清晰、可预期的确权路径，实现规制目标与权利激励的双重整合。

四、我国AIGC标识义务的制度构建与实施路径

（一）构建分层设定的全链条标识义务体系

基础模型开发者与服务提供者作为AIGC的技术源头，拥有最为强大的技术控制力，理应承担底层技术标识与前端提示的核心义务。欧盟《人工智能法案》以及我国《生成式人工智能服务管理暂行办法》均对其标识责任作出明确规定，该类主体不仅需要在生成工具底层嵌入数字水印、元数据标签等技术模块，同时还要为用户提供显性标识选项。

内容生成者与使用者直接决定AIGC的用途与传播范围，是面向公众的直接披露主体，故而需要承担披露与声明的直接义务。传播平台作为网络空间的管理者，虽不直接参与内容生成，但掌控着内容传播的关键渠道，因此需要承担核验与协管的义务。

我国《网络信息内容生态治理规定》也明确平台需对可疑信息进行风险提示。在实践操作中，平台应当开发AIGC检测接口，要求上传者勾选AI生成属性，针对未标识的高逼真AIGC主动开展二次标注或风险提示。

（二）确立分类分级的差异化标识规则

分类分级标识应始终遵循“风险越高、义务越重”的核心逻辑，将AIGC应用场景划分为三个类别，并针对设定梯度化的标识要求。借鉴欧盟AI法案风险分级监管的思路，结合我国算法分类管理的基本原则，实施差异化的标识规则，以此兼顾监管实效与产业发展需求。

高风险场景主要涵盖新闻报道、医疗诊断、学术出版、政治选举、公共决策等关乎公共利益与社会伦理的领域，此类场景需要实施强制性深度标识，不仅要强制披露AI参与的具体环节，还需同时采用强鲁棒性数字水印与显性声明，确保内容可追溯、可核验，从而有效防范虚假信息引发的重大社会风险。

中风险场景主要包括商业营销、AI绘画、艺术创作等商业应用领域，该类场景侧重实施防混淆商业标识，核心目标是防范消费者误认与不正当竞争行为，要求在商业宣传中明确标注AI生成属性，以此保护人类创作者的市场权益。

低风险场景主要为个人娱乐、内部测试、日常聊天等非公开、非商业用途，此类场景适用自愿性轻量标识，依靠模型服务商的底层隐性水印标识，豁免个人用户的显性标识义务并采取自愿标识原则。

（三）健全技术、监管与责任并重的实施保障机制

1. 技术保障与标准规范

现有水印技术普遍面临对抗性扰动与去水印等技术挑战，因此一方面需要强制推行防篡改、强鲁棒性的水印技术，保障标识信息难以被轻易去除；另一方面应当建立国家级AIGC内容溯源与检测核验平台，形成正向水印验证与反向AI特征检测的双重技术屏障，实现AIGC全流程可追溯。

标识内容标准层面，至少应当涵盖所使用AI工具名称、版本、生成时间的基础信息，包含内容生成来源、训练数据合法来源说明的来源信息，针对存在“事实幻觉”的内容所作出的风险提示，以及涉及人脸、地点、事件等真实要素的深度伪造内容所标注的“非真实”属性。

技术标准层面，我国应当建立显性标识与隐性标识相结合的技术规范体系，与C2PA、Google SynthID等国际主流标准相契合。

显性标识是在文本标题下方、图像边角、视频固定帧等位置添加肉眼可见的声明或图标；隐性标识则采用深度学习数字水印技术，将标识信息嵌入内容冗余空间即频域、空间域之中，同时统一元数据标准，确保AIGC在脱离原始平台后依旧能够被机器读取溯源信息。

2. 监管执法协同机制

应建立网信办牵头，工信、公安、市场监管及版权局多部门联动的机制。在执法中，版权部门负责处理因未标识引发的著作权纠纷；市场监管部门依据《反不正当竞争法》对利用未标识AIGC进行虚假宣传、扰乱市场秩序的经营者进行查处；网信部门则督促平台履行协管责任，形成“技术对抗技术、机制管好平台”的共治格局[8]。同时鼓励行业组织出台自律公约，建立黑名单共享机制。

3. 法律责任梯次体系

为强化标识义务的威慑力，需要构建民事、行政、刑事梯次递进的完整责任体系。

第一，以填补损害与行为矫正为导向，完善民事责任规则。在涉AIGC纠纷中，若AI服务提供者或使用者未按规定嵌入数字水印或进行显性披露，应当在诉讼中适用“过错推定”原则，由其承担举证不能的不利后果。对于恶意去除隐性标识、利用未标识内容进行深度伪造以攫取非法商业利益的主体，司法机关不仅应判令其承担停止侵害、赔偿实际损失的责任，还应大幅提高技术滥用者的违法成本。

第二，以维护行业生态与公共利益为核心，强化行政监管责任。针对基础模型开发者未从底层嵌入标识代码、内容分发平台怠于履行AIGC核验义务，以及生成者恶意规避强制披露等违法违规行，网信、工信及市场监管等部门应依据《生成式人工智能服务管理暂行办法》及相关专门法规，实施穿透式监管，建立阶梯式的行政处罚机制。

第三，以捍卫国家安全与社会秩序为底线，激活刑事责任。刑法作为社会治理的最后一道防线，当规避标识义务的行为突破民刑交叉边界，对社会整体信任体系造成严重破坏时，必须予以严厉打击。

（四）加强国际合作与跨境监管协调

AIGC技术的无国界性决定了单边治理的效力受限，跨境侵权与虚假信息的大规模传播对国家安全提出了全球性挑战。

一方面，我国应积极争取AIGC国际标准制定的主导权。依托庞大的国内市场与应用场景优势，主动在世界知识产权组织等国际框架下提出生成式AI数字水印技术规范与通用标识符号的“中国方案”，推动技术标准的国际互认，从而降低国内AI企业

的 出海合规成本。

另一方面，则应探索建立跨境监管协调机制。在应对境外开源大模型或跨境AI服务时，确立适度的管辖原则。对于向境内用户提供AIGC服务或其生成内容在境内大规模传播的境外主体，必须强制其符合我国的隐式与显式标识要求。同时，加强与国际 社会在版权认定、跨境数据追踪与虚假信息阻断方面的多边执法合作，共同构建安全、互信的全球人工智能治理新秩序。

五、结语

标识义务并不试图回答“AI能否成为作者”这一终极命题，而是以信息披露为支点，在人类创作与AI生成之间划出一道制度性边界，使法律体系得以在前端重建识别与区分的能力。这道边界向前保障公众知情权，在风险前端形成认知过滤屏障；向 后 为版权确权、侵权归责与数据合规提供事实基础，使事后救济不再因内容的“不可见”而落空。

本文所主张的风险分级差异化标识规则，正是技术创新与权利保障之间动态平衡的制度表达——在低风险场景保持克制，在高风险场景施加刚性约束，使规制力度与技术风险相匹配。以标识义务的制度化为起点，推动形成“标识—确权—追溯—问责”的完整治理闭环。

展望未来，在这场人与机器的秩序重构中，标识义务或许是一个看似谦抑的起点，但正是这一道看似微小的边界，守护着人类表达的可信度、法律责任的确定性，与信息社会赖以运行的认知秩序。

参考文献

陈兵. 人工智能生成内容权益保护的实践检视及优化[J]. 社会科学战线, 2025(6): 191-200.

邓建鹏, 苏星月. 人工智能绘画的版权风险与应对思考[J]. 学术论坛, 2024(3): 124-136.

郭钊均, 李美玲, 周杨铭, 等. 人工智能生成内容模型的数字水印技术研究进展[J]. 网络空间安全科学学报, 2024, 2(01): 13-39.

蒋舸. 论人工智能生成内容的可版权性：以用户的独创性表达为视角[J]. 知识产权, 2024(1): 36-66.

李永志. 人工智能生成内容的保护模式研究[J]. 北方经贸, 2025(10): 1-5.

刘桂新, 冯世烽. 传播平台人工智能生成合成内容标识的制度逻辑与机制完善[J]. 重庆邮电大学学报(社会科学版), 2026(1): 51-60.

刘晓春. 生成式人工智能数据训练中的“非作品性使用”及其合法性证成[J]. 法学论坛, 2024, 39(3): 67-78.

王岭, 胡丹妮, 熊艳. 生成式人工智能的内容风险挑战与监管政策[J]. 东北财经大学学报, 2025(5): 15-24.

王迁. 人工智能生成的内容是作品吗？——以学术规范和著作权法的关系为视角[J]. 中国法律评论, 2025(5): 37-51.

许茂恒. AIGC大模型中内源性幻觉难题溯源与版权规制路径[J]. 北京理工大学学报(社会科学版), 2024, 26(06): 98-108.

吴汉东. 论人工智能生成内容的可版权性：实务、法理与制度[J]. 中国法律评论, 2024(3): 113-129.

杨立新, 黄露. 利用人工智能绘画工具生成物的性质认定及其保护[J]. 求是学刊, 2025(4): 102-114.

姚叶. 机器窃书为雅罪？——人工智能生成内容著作权侵权认定及风险纾解[J]. 暨南学报(哲学社会科学版), 2025(8): 71-89.

易耀森. 人工智能生成内容应用出版伦理政策现状分析与完善建议——以我国160种医学期刊为例[J/OL]. 科技与出版, 2025, 44(10): 82-93.

于嘉诚. 人工智能服务提供者对生成内容的版权注意义务[J]. 中国政法大学研究生学报, 2025(1): 158-171.

张平. 人工智能生成内容著作权合法性的制度难题及其解决路径[J]. 法律科学(西北政法大学学报), 2024, 42(3): 18-31.

朱阁. “AI文生图”的法律属性与权利归属研究[J]. 知识产权, 2024(1): 24-35.

Höltvervenhoff S., Ricker J, Raphael M M, et al. Security Benefits and Side Effects of Labeling AI-Generated Images[R]. arXiv preprint, 2025.

Letrone W, Hayashi M. Current Developments on Labelling of AI-Generated Content: Comments on the AI Convention and the EU AI Act[J]. 國際協力論集, 2025, 33: 59-67.

Li Y., Shao S., He Y., et al. Rethinking Data Protection in the (Generative) Artificial Intelligence Era[R]. arXiv preprint, 2025.

Wang T., Zhang Y., Qi S., et al. Security and Privacy on Generative Data in AIGC: A Survey[J]. ACM Computing Surveys, 57(4), Article 82, 2024.

说明：1. 总文字复制比：被检测文献总重复字符数在总字符数中所占的比例

2. 去除引用文献复制比：去除系统识别为引用的文献后，计算出来的重合字符数在总字符数中所占的比例
3. 去除本人文献复制比：去除系统识别为作者本人其他文献后，计算出来的重合字符数在总字符数中所占的比例
4. 单篇最大文字复制比：被检测文献与所有相似文献比对后，重合字符数占总字符数比例最大的那一篇文献的文字复制比
5. 复制比按照“四舍五入”规则，保留1位小数；若您的文献经查重检测，复制比结果为0，表示未发现重复内容，或可能存在的个别重复内容较少不足以作为判断依据
6. **红色文字**表示文字复制部分；**绿色文字**表示引用部分（包括系统自动识别为引用的部分）；**棕灰色文字**表示系统依据作者姓名识别的本人其他文献部分
7. 系统依据您选择的检测类型（或检测方式）、比对截止日期（或发表日期）等生成本报告
8. 知网个人查重唯一官方网站：<https://cx.cnki.net>