

人工智能时代声音权益保护的困境与体系回应

贾姝萱

宁波大学马克思主义学院, 浙江 宁波

收稿日期: 2026年8月6日; 录用日期: 2026年8月17日; 发布日期: 2026年9月8日

摘要

伴随着人工智能合成与深度伪造技术的快速发展, 对自然人声音权益的保护面临前所未有的侵害风险。声音与肖像虽同属标表个人特征的人格标识, 但在可识别性的判断标准、侵权的具体表现形式、权利主体的界定等方面却存在着实质性差异。因此, “参照适用”的统一标准已难以充分回应人工智能时代下声音权益保护的差异化需求。如何建构差异化的声音可识别性判断标准并合理界定AI生成声音场景中的责任主体, 实现人格权保护与技术创新的平衡成为了一项重要议题。

关键词

声音权益, 人格权, 人工智能, 可识别性

Voice Rights Protection in the AI Era: Dilemmas and Systemic Responses

Shuxuan Jia

School of Marxism Studies, Ningbo University, Ningbo Zhejiang

Received: August 6, 2026; accepted: August 17, 2026; published: September 8, 2026

Abstract

With the rapid advancement of AI synthesis and deepfake technologies, the protection of natural persons' voice rights faces unprecedented risks of infringement. Although voice and portrait both serve as personal identifiers that signify individual characteristics, they differ substantially in terms of identifiability criteria, specific forms of infringement, and the definition of rights holders. Therefore, the uniform standard of “application by reference” can no longer adequately address the differentiated needs of voice rights protection in the AI era. How to establish differentiated identifiability criteria for voice and reasonably define the responsible parties in AI-generated voice scenarios, while striking a balance between personality rights protection and technological innovation,

文章引用: 贾姝萱. 人工智能时代声音权益保护的困境与体系回应[J]. 法学, 2026, 14(9): 108-114.

DOI: 10.12677/ojls.2026.149262

has become a critical issue.

Keywords

Voice Rights, Personality Rights, Artificial Intelligence, Identifiability

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. 引言

人工智能声音合成技术的迅速发展正深刻地改变着声音的利用方式与保护路径。人工智能声音克隆、深度伪造等技术的广泛应用,使得自然人的声音可以被低成本、高效率地采集、合成、模仿甚至篡改。从配音演员的声音被 AI 化后用于商业产品,到名人声音被深度伪造并用于直播带货,再至网络主播的声音被“偷走”后用于推销,可见人工智能技术对声音权益的侵害已经从理论推演转变为现实常态。在此背景下,《中华人民共和国民法典》(后文简称《民法典》)第 1023 条第 2 款的规定:“对自然人声音的保护,参照适用肖像权保护的有关规定”构成了声音权益保护的核心规范依据。但是声音与肖像在物理属性、识别机制、侵权形态等方面的实质性差异,却造成了对声音权益保护在现实实践中面临了重重困难。简单地“参照适用”肖像权保护规则显然已经难以回应 AI 时代声音权益保护的需求。从既有研究来看,目前学界对声音权益的关注主要集中在声音权益的法律性质与“参照适用”肖像权规则的具体操作方案,对人工智能技术给声音权益带来的侵权标准模糊、主体归责不明确等问题的研究则相对较少。那么在人工智能语音合成与深度伪造技术的广泛应用下,如何立足实践中的具体问题,明确人工智能对自然人声音权益的侵权边界并寻找出一条切实保护声音权益的现实路径呢?

2. 声音权益的规范基础与法律性质

2.1. 《民法典》第 1023 条第 2 款的意涵

《民法典》第 1023 条第 2 款规定:“对自然人声音的保护,参照适用肖像权保护的有关规定。”这一规定具有三层意义。其一,确认了声音具有独立人格的利益属性。在《民法典》颁布之前,我国法律并未对声音保护作出明确规定,声音利益长期处于法律保护的模糊地带。第 1023 条第 2 款的出台,标志着立法者明确将声音确立为独立于肖像、姓名等人格标识之外的特殊人格利益。声音作为身体的一部分,是由人类发声器官活动所产生的结果,对人的生命、身体、健康、自由和尊严等具有重要意义,承载着与姓名、肖像和名誉等其他人格要素同等重要的人格利益[1]。其二,采用了“参照适用”的准用性立法技术。所谓“参照适用”,是指法律明确规定某一事项可以适用其他法律规范,但并非完全的“适用”,而是需要在考量规范对象之间的相似性与差异性的基础上,有选择地、变通地适用。就声音保护而言,《民法典》第 1018 条规定“依法制作、使用、公开或者许可他人使用”的肖像权权能、第 1019 条有关“不得以丑化、污损,或者利用信息技术手段伪造等方式”侵害他人肖像权行为的规定等,均可参照适用于对自然人声音权益的保护。其三,明确了声音保护的基本框架。人格权利是民事主体享有的、为实现人格尊严和人格自由而排斥他人侵害的权利[2]。而民法典第 110 条规定肖像权由自然人享有¹,因此

¹ 《民法典》第 110 条第 1 款规定:自然人享有生命权、身体权、健康权、姓名权、肖像权、名誉权、荣誉权、隐私权、婚姻自主权等权利。

参照肖像权,声音保护的主体为自然人。究其客体而言,声音具有识别特定自然人身份的功能,法律保护的对象只能是声音所体现的利益,而不是声音自身[3]。此外,对“参照适用”的法律技术内涵还需进一步厘清。参照适用不同于准用或直接适用,其赋予了裁判者在具体个案中根据声音特性的差异进行裁量调整的空间。立法者之所以选择参照适用的技术路径,一方面是基于对声音与肖像同质性的肯定,另一方面也是出于对声音权利边界尚不清晰的审慎考量。在人工智能技术尚未充分展现其全部风险时,采用更具弹性的保护模式,可以为未来法律回应技术变革预留必要的制度空间。然而,这种弹性在当下却逐渐演变为不确定性,亟需通过教义学上的精细作业加以填充。

2.2. 声音权益的法律性质

《民法典》第1023条第2款使用了“保护”而非“权利”的表述,与肖像权条文中“肖像权”的表述形成对照。这表明立法者有意将声音定义为一种受法律保护的人格利益,而非一项独立的具体人格权。声音与肖像同属表型人格标识,二者皆具有基于人本身产生的唯一可识别性的特征。除非通过科技手段,声音一旦定型就难以改变,因此与肖像一样可以成为表型自然人的人格标志。从保护需求来看,声音权益已经超越了纯粹的精神利益范畴,兼具显著的财产属性。现代社会中,声音成为商业运作的对象,其商品化权益逐渐受到重视,自然人可通过使用、制作或许可他人使用自身声音等方式获得经济利益[4]。映射声音权益的商业化利用,声音权益包含人格权益、财产权益,且财产权益之产生乃基于声音之人格标识利用[5]。尤其是知名人士的声音,在导航、广告、有声读物等商业场景中具有可观的经济价值。因此,无论理论上将声音界定为“权利”还是“利益”,都不应否定其作为独立人格标识的法律地位以及精神利益与财产利益并存的二元属性。这种二元属性在人工智能时代被进一步放大。一方面,人工智能技术使得声音的财产价值可以被规模化利用,权利人对其声音经济利益的支配需求空前迫切;另一方面,深度伪造技术对声音的篡改和嫁接,又可能对权利人的人格尊严造成严重侵害。财产属性的扩张与人格属性的脆弱,共同构成了当下声音权益保护的张力。如果过度强调财产属性而忽视声音作为人格标识的尊严内核,可能导致声音的商品化失控;反之,如果固守传统人格权的非财产性立场,则可能阻碍声音作为数据要素的合法流通与利用。

2.3. 声音与肖像的同质性及其限度

声音与肖像之所以可以“参照适用”,根源在于二者同属表型人格标识。声音与肖像均具有独特性、唯一性、稳定性的特点,能够展示出个人的行为和身份。在实践中,声音与肖像经常同时使用,如在广告中使用他人肖像的同时往往会配上其声音。从功能的角度来看,声音与肖像都发挥着识别特定自然人身份的作用,都承载着人格尊严与人格自由发展的价值。然而,声音与肖像在物理属性和识别机制上却存在明显差异,这些差异构成了“参照适用”的限度。其一,声音的识别依赖于听觉感知,而肖像的识别依赖于视觉感知,二者所依赖的感官通道不同,导致识别主体和标准的设定上存在差异。其二,声音具有时间性,即声音的呈现需要一定的时间维度;而肖像具有空间性。这一差异使得声音的固定、复制和传播比肖像更加隐蔽、更加难以追踪。其三,声音的模仿门槛显著低于肖像的模仿门槛。自然人可以通过训练模仿他人的声音,但难以通过非科技方式模仿他人的面部特征。人工智能技术则进一步降低了声音模仿的门槛。这意味着,在将肖像权保护规则参照适用于声音权益保护时,必须根据声音的特殊属性作出必要的调适。

3. 人工智能时代声音权益保护困境

3.1. 人工智能合成技术造成声音识别困难

人工智能语音合成技术,包括文本转语音、语音克隆、声音深度伪造等的迅速发展造成对声音识别

的困难不断加剧。人工智能技术下声音的采集与使用变得更为便捷。在传统场景中,声音的使用通常伴随着公开的录制或播放行为。而在人工智能时代,一段数秒的公开讲话录音即可作为训练素材,通过算法生成高度仿真的合成声音。这种对声音的侵权往往采取极为隐蔽的手段并增加识别的难度。传统的声音识别主要依赖于通过音色、语调、发音风格等声学特征对自然人原始声音的比对。而在人工智能合成声音的场景中,争议声音并非对原始声音的直接复制,而是经过算法处理后形成的合成物,如果算法仅保留了原始声音的部分特征,对声音可识别性的判断将更为困难。在我国首例 AI 生成声音人格权侵权案中,五被告均否认侵权,主张人工智能合成后的声音产品,与自然人声音在人身权属性上有所区别,目前的技术都会对人工智能合成声音进行水印标记,这切断了人工智能合成声音与自然人声音之间的联系,不会产生对应自然人的人格属性,不具有对原告人格的可识别性²。

3.2. 侵权主体多样

相较于传统对声音权益侵害的案件中主体的单一性,人工智能技术下的声音权益侵权行为的链条更为复杂。在“殷某桢案”³中,一条完整的侵权链条涉及配音演员、文化传媒公司(录音制品的著作权人)、软件公司(AI 化处理者)、云服务平台(销售者)、智能科技公司(使用者)等多个主体。如何在不同主体之间分配侵权责任,成为司法实践中的难题。可见,人工智能的多方参与者加剧了责任主体的模糊性。生成式人工智能作为算法、算力、数据和用户的集合体,是一种更为复杂的技术生态,涉及的主体范围更广。从技术维度考量,生成式人工智能涉及模型设计者、数据供应者、数据所有者、算力供应商、服务提供者和用户这六大主体。上述六大主体全部参与了人工智能的内容生成。一旦生成式人工智能生成侵权内容,那么必然陷入辨识侵权责任主体的困境[6]。

3.3. “参照适用”模式张力不足

“参照适用”模式在人工智能时代逐渐显露出弊病,其面临的根本困境在于:声音与肖像的差异在技术背景下被放大,而“参照适用”的统一路径已经难以充分回应这些差异,甚至忽视声音的特殊性以及二者在侵权行为认定方式上的区别。声音作为听觉类标志,其标表自然人身份的识别性相较于肖像这一视觉类标志明显较弱,与此同时,对声音的使用往往与肖像、声音内容相结合,因此,不能简单挪用肖像的可识别性判断标准。[7]而传统简单地可识别性标准以及参照肖像权的“制作、使用、公开”框架,显然无法涵盖人工智能合成技术声音侵权的全部形态。其次,“参照适用”加剧了法律适用的不确定性。《民法典》第 1023 条第 2 款虽然指明了参照方向,但并未明确具体应该参照哪些条款以及如何参照。这一规定导致了对声音权益的界定和侵权形态均不够明晰,对声音的许可使用和法律保护解读不一。在人工智能技术快速迭代的背景下,这种不确定性会阻碍对声音权益的有效保护。

4. 人工智能时代声音权益保护的现实路径

可见,单纯参照肖像权已无法完备的保护权利人的声音权益,基于现实问题寻求一条能够完善声音权保护的路径成为了亟待解决的问题。

4.1. 建构差异化的声音可识别性判断标准

可识别性是声音权益保护的核心要件。无论是使用客观的鉴定技术还是参考人耳的主观判断,其目的都在于证据标准上保证“确是使用该自然人的声音”的高度盖然性[8]。在人工智能时代,应当建构差异化的可识别性判断标准。在保护人格尊严的场景中,例如利用人工智能技术丑化声音、深度伪造语音

²参见北京互联网法院(2023)京 0491 民初 12142 号民事判决书。

³同 2。

用于诽谤他人等，可识别性的门槛应适当放宽。只要相关公众或特定群体能够将争议声音与权利人联系起来，即应认定其具有可识别性。这是因为人格尊严的保护具有优先性，不应因技术手段的隐蔽性而导致保护真空。正如“殷某桢案”裁判要旨所揭示的，经人工智能技术处理后的声音是否落入自然人声音权益的保护范围，关键在于通过该声音是否仍能识别出该特定自然人⁴。在保护人格同一性利益的场景中，例如利用人工智能技术生成声音以冒用权利人身份，应当以一定范围内的熟人能否确信争议声音属于权利人为准。同一性利益的核心在于维护自然人人格标识与其人格之间的稳定对应关系。如果权利人生活、工作范围内的熟人能够确信该人工智能合成声音指向权利人，即应认定具有可识别性。在保护财产性利益的场景中，例如人工智能合成声音并用于商业广告、有声读物中等，有必要将声音的商业利用排除出人格权规制范畴，以声音经济价值来源为标准，区分保护知名人士和普通自然人的声音[9]。如果采用过于宽松的可识别性标准，将不当限制人工智能语音技术的合法应用，抑制产业创新。因此，应当要求在隔离对比的情境中，相关领域的普通公众能够将人工智能声音与特定自然人的原始声音明确对应。在判断方法上，声音可识别性的认定应当综合考虑行为人的使用情况。具体而言，技术维度上，如人工智能合成声音与原始声音在音色、语调、发音风格等方面的相似程度；场景维度上，如声音的使用场景是否足以引导公众将声音与特定自然人关联；主体维度上，根据自然人的社会知名度，采取一般社会公众或一定范围内公众的主观标准。

4.2. 合理界定人工智能声音侵权中的责任主体

人工智能技术下的声音侵权往往涉及多个主体，传统模式中谁侵权谁负责的直接责任原则在多元传播场景下面临适用困境。因此，应当根据各主体在侵权链条中的角色和作用，合理界定责任。对于声音素材提供者的责任的认定，在“殷某桢案”⁵中，文化传媒公司作为录音制品的著作权人，将含有殷某桢声音的录音制品提供给软件公司用于AI化处理。法院认定，文化传媒公司依照约定对录音制品享有著作权等权利，但不享有授权他人处理并使用殷某桢声音的权利。因此，声音素材提供者在未获得权利人明确授权的情况下将声音素材提供给第三方用于AI化处理，应当承担侵权责任。对于人工智能技术提供者的责任的认定，软件公司利用AI技术对声音素材进行处理并生成文本转语音产品。法院认定，未经许可对他人声音进行AI化处理与使用构成侵权。AI技术提供者不能以“技术中立”为由免责，当其将AI技术应用于特定自然人的声音并生成可识别的合成声音时，即已构成对声音权益的积极侵害。对于平台运营者的责任认定，法院认定本案中智能科技公司通过应用程序接口直接调取并生成文本转语音产品在其平台中使用构成侵权。在另一涉AI“深度伪造”名人声音带货案中⁶，法院进一步明确，平台商家对其网络主播发布的推广内容负有合理审查义务，未尽到审核注意义务的，需与网络主播承担连带责任。这表明，平台运营者不能以“仅为技术服务提供者”为由规避责任，而应当根据其在侵权内容生成和传播中的实际作用承担相应责任。在责任形态的具体认定上，应当区分直接侵权责任与间接侵权责任，并根据不同主体的风险控制能力和注意义务水平确定责任边界。模型设计者与数据供应者位于产业链上游，其行为属于事先的数据收集与模型构建，距离最终侵权内容的生成较远。对于此类主体，不宜施加过重的审查义务，而应侧重于要求其在数据来源合法性方面履行合规审查职责。服务提供者与平台运营者处于产业链中下游，其不仅具有更强的风险识别和管控能力，也因侵权内容的传播而直接获利，理应对其平台上的AI合成声音内容承担更高的注意义务。具体而言，当平台知道或应当知道其用户利用AI技术合成侵权声音时，应当及时采取删除、屏蔽、断开链接等必要措施，否则应当就损害的扩大部分承担连

⁴同2。

⁵同2。

⁶参见北京互联网法院(2024)京0491民初10085号民事判决书。

带责任。此外，在强化声音权益保护的同时，亦须审慎划定保护边界，避免过度保护抑制人工智能技术创新与文化创意自由。在 AI 声音模仿情境下，若合成声音系用于评论、新闻报道、学术研究、讽刺模仿等非商业性目的，且未对权利人的人格尊严或财产利益造成实质损害，应承认行为人享有一定的抗辩空间。这种“分层递进”的责任认定逻辑，既能有效惩戒和预防侵权行为，又不至于过度抑制 AI 产业链各环节的创新发展。

4.3 强化技术治理

人工智能生成声音技术的隐蔽性使得侵权行为的发现和证明异常困难。在“殷某桢案”中，被告甚至主张 AI 合成声音经过水印标记“切断了与自然人之间的联系”⁷。因此，利用声音水印、来源披露等技术治理措施，正是将技术手段由侵权的遮羞布转化为治理关键口的重要工具。

我国已初步建立起以内容标识为核心的技术治理规范体系。《人工智能生成合成内容标识办法》(以下简称《标识办法》)⁸以部门规章形式明确了 AI 生成内容的标识义务，确立了“显式 + 隐式”双重标识制度。此外，技术措施在侵权责任认定中也发挥着重要作用。其一，标识的缺失或篡改可作为主观过错的推定依据。《标识办法》第 10 条明确规定：“任何组织和个人不得恶意删除、篡改、伪造、隐匿本办法规定的生成合成内容标识。”若 AI 合成声音服务提供者未按规定添加标识，或者使用者故意删除标识后传播侵权内容，法院可将此作为认定其具有主观故意的有力证据。其二，隐式标识可为溯源追责提供技术支撑。隐式标识记录着服务提供者、生成时间、内容编号等关键信息。即便侵权内容经过多次传播或简单编辑，其 AI 生成属性依然可追溯。除却以上《标识办法》的强制性措施外，还应增加激励措施以确保技术治理措施的落地。在激励层面，应当建立技术合规与责任减免的联动机制。服务提供者如能证明其已按规定添加隐式标识且侵权声音的生成记录可追溯，在涉及第三方恶意篡改或删除标识的案件中，可主张减轻或免除赔偿责任。这种制度设计既能避免对技术开发者施加过重的审查负担，又能通过利益驱动促进行业自律。

5. 结语

人工智能时代的声音权益保护，既不能脱离《民法典》第 1023 条第 2 款所确立的规范框架，也不能简单地将肖像权保护规则“照搬”适用于声音场景。声音与肖像虽有同质性，但在物理属性、识别机制、侵权形态等方面的差异不容忽视。“参照适用”的精义在于“参照”而非“适用”。因此，在借鉴肖像权保护成熟经验的同时，必须根据声音的特殊属性作出必要的调适。在立法层面，条件成熟时可以考虑通过司法解释或专门立法，明确声音权益的具体内容和保护规则，填补“参照适用”模式留下的规范空白；在司法层面，通过指导性案例和典型案例逐步形成类型化的裁判规则，为声音可识别性的判断、侵权责任的认定、损害赔偿的计算提供更具操作性的指引；在技术治理层面，探索声音水印、合成标识等技术手段在声音权益保护中的应用，实现法律规制与技术治理的协同。

伴随着人工智能技术的发展，声音的商业价值和社会价值将持续提升。唯有在“参照”与“独立”之间寻求动态平衡，在借鉴肖像权保护成熟经验的同时充分回应声音的特殊性，才能为数字时代的声音权益提供坚实的法律保障。

参考文献

- [1] 张新平. 数字时代自然人声音权的性质与构造[J]. 法商研究, 2026, 43(3): 50-64.

⁷同 2。

⁸https://www.gov.cn/zhengce/zhengceku/202503/content_7014286.htm

- [2] 王利明, 程啸. 中国民法典释评·人格权编[M]. 北京: 中国人民大学出版社, 2020: 7.
- [3] 王绍喜. 《民法典》时代声音保护的解释与适用[J]. 法律适用, 2023(6): 35-44.
- [4] 陈笑春, 李泽林. 谁的声音: 生成式人工智能语境下声音的主体与权益[J]. 西南民族大学学报(人文社会科学版), 2025, 46(3): 141-150.
- [5] 李国强, 孙一理. 声音权益商业化利用的法律困境及消解路径[J]. 东南大学学报(哲学社会科学版), 2025, 27(6): 22-31, 155.
- [6] 陆小华, 陆赛赛. 论生成式人工智能侵权的责任主体[J]. 南昌大学学报(人文社会科学版), 2024, 55(1): 119-131.
- [7] 北京互联网法院课题组. AI生成声音侵害声音权益的法律认定——以殷某某诉北京某智能科技有限公司等人格权侵权案为例[J]. 法律适用, 2024(9): 123-133.
- [8] 苏艺. 声音商品化权可识别性认定的二阶构造——兼评“人工智能声音权第一案”[J]. 河北法学, 2025, 43(11): 102-121.
- [9] 孙雅琪. 智媒时代声音权益的法律保护: 困境成因与纾解之策[J]. 出版发行研究, 2026(2): 70-78.