

虚拟数字人系列之六 — 深度合成技术合规篇

作者：段志超 | 徐源璟 | 王媛娇¹

“深度合成”是虚拟数字人制作的关键技术之一，其主要依赖自动编码器（Autoencoders）和生成对抗网络（Generative Adversarial Networks）²等人工智能技术提取原始数据、进行模型训练，并在训练后对数据重建合成。为了确保该等服务提供者在技术开发和使用过程中的健康有序发展，监管部门一直在摸索对深度合成等人工智能技术合规监管的方法和路径。从我国已出台的部门规章和规范性文件来看，《网络信息内容生态治理规定》、《互联网信息服务算法推荐管理规定》、《网络音视频信息服务管理规定》均已包含概括性地对生成合成类算法和利用深度学习、虚拟现实等新技术新应用制作音视频内容等进行监管的规定。而国家互联网信息办公室于2022年1月28日发布的《互联网信息服务深度合成管理规定（征求意见稿）》（“《深度合成规定》”），则进一步厘清和细化了深度合成技术的应用场景，明确了深度合成服务提供者和使用者的信息安全义务，并规定国家网信部门负责统筹协调全国深度合成服务治理和相关监督管理工作。

一、深度合成的应用场景

根据《深度合成规定》的定义，深度合成是指利用以深度学习、虚拟现实为代表的生成合成类算法制作文本、图像、音频、视频、虚拟场景等信息的技术，并以包括但不限于的形式列举了六类技术。结合虚拟数字人的制造和商业化运营中可能出现的元素，我们将《深度合成规定》所明确列举的几类技术分类总结如下：

环节	元素内容	具体场景	对应《深度合成规定》的技术分类
创设环节	虚拟数字人的形象外观	以真人视觉形象为基础创造虚拟数字人的视觉形象（人脸替换）、无真人形象为基础创造虚拟数字人视觉形象（人脸生成）	人脸生成、人脸替换、人物属性编辑、人脸操控、姿态操控等对图像、视频内容中人脸等生物特征进行生成或者编辑的技术（“生物特征类”）
	虚拟数字人的声音	以真人声音制作音频样本、以真人声音为基础的声音创设（语音转换）、仅基于文本生成的声音创设（文本转语音）	文本转语音、语音转换、语音属性编辑等对语音内容进行生成或者编辑的技术（“语音内容类”）
	虚拟数字人的动作	真人驱动类虚拟数字人的动作创设（姿态操控）	生物特征类

¹ 实习生邹奕对本文的写作亦有贡献。

² 参见曹建峰：《AI生成内容发展报告2020—“深度合成”（Deep synthesis）商业化元年》。

环节	元素内容	具体场景	对应《深度合成规定》的技术分类
商业运营环节	虚拟数字人的商业化作品	虚拟数字人出演话剧、MV 等视听作品或其直播（语音转换、人脸操控、姿态操控）	语音内容类
			生物特征类
		虚拟偶像举办演唱会	音乐生成、场景声编辑等对非语音内容进行生成或者编辑的技术（“非语音内容类”）
	已故人物虚拟分身重现表演	三维重建等对虚拟场景进行生成或者编辑的技术（“三维重建”）	
	虚拟数字人通过 AI 自动生成的内容	AI 智能写作、智能作画	篇章生成、文本风格转换、问答对话等对文本内容进行生成或者编辑的技术（“文本内容类”）
		虚拟主播语音播报（文本转语音）	语音内容类
虚拟人问答聊天（如虚拟男友、虚拟女友等）		文本内容类	

二、深度合成技术的合规要点

深度合成服务提供者，是指提供深度合成服务以及为深度合成服务提供技术支持的组织³。《深度合成规定》明确要求深度合成服务提供者应落实主体责任，从技术、内容等多角度提出了对服务者的制度管理要求。结合《深度合成规定》，我们建议虚拟数字人的技术方应重点关注以下要求：

（一）标识深度合成信息内容

《深度合成规定》对深度合成信息内容进行了分类并要求进行标识管理。若所提供的服务为具有生成或者显著改变信息内容功能的，应当使用显著方式在合理位置或区域对深度合成信息内容进行显著标识。《深度合成规定》第十四条列举的典型情境如下：

#	类型	典型情境	标识方式
1	提供模拟自然人进行文本生成或者编辑服务	智能对话、智能写作等	在文本信息内容的稿源说明处等位置
2	提供语音生成或者显著改变个人身份特征的编辑服务	合成人声、仿声等	在音频信息内容的合理区域以语音说明等方式
3	虚拟人物图像、视频生成或者显著改变个人身份特征的编辑服务	人脸生成、人脸替换、人脸操控、姿态操控等	在图像、视频信息内容的明显位置
4	生成或者编辑服务	沉浸式拟真场景等	在虚拟场景信息内容的明显位置
5	提供其他具有生成或者显著改变信息内容功能的的服务的	/	在文本、图像、音频或者视频、虚拟场景等的合理位置或者区域进行显著标识

³ 《深度合成规定》第二条：深度合成服务提供者，是指提供深度合成服务以及为深度合成服务提供技术支持的组织。

若所提供的服务为其他深度合成服务的，仅需向深度合成服务使用者提供对深度合成信息内容进行显著标识的功能并提示使用者可进行该标识即可。考虑到虚拟数字人依赖于多种合成技术的组合，相关运营方仍有必要高度关注显著标识义务。

若运营方发现按《深度合成规定》理应进行显著标识的深度合成信息内容未被恰当标识，应立即停止传输该等信息、待整改完成后再继续进行。同时运营方需关注，前述标识的添加不应影响用户使用，并应依法保存可识别、可追溯的网络日志信息。

（二）加强技术管理：建立健全算法机制机理审核，定期审核、评估、验证算法机制机理

在《深度合成规定》发布之前，我国已于 2021 年 12 月对包含生成合成类技术在内的“应用算法推荐技术”出台了《互联网信息服务算法推荐管理规定》（“《算法规定》”），从信息内容规范、算法技术审核、用户权益保护及服务者备案等多方面对算法推荐服务提供者提出要求。虚拟数字人运营方使用生成合成类算法技术，应遵守《算法规定》的基础内容，包括但不限于（1）如果构成“具有舆论属性或者社会动员能力的算法推荐服务提供者”，需在提供服务之日起十个工作日内通过互联网信息服务算法备案系统进行服务者备案；（2）建立健全科技伦理审查、反电信网络诈骗及安全事件应急处置等管理制度及规则；（3）为管理制度配备相关专业人员及技术支撑等。

而在《深度合成规定》中，亦有细化规定要求深度合成服务提供者加强深度合成技术管理，定期审核、评估、验证算法机制机理；提供具有对人脸、人声等生物识别信息或者可能涉及国家安全、社会公共利益的特殊物体、场景等非生物识别信息编辑功能的模型、模板等工具的，应当自行开展安全评估，预防信息安全风险。

除对特定算法工具需要履行安全评估要求外，《深度合成规定》第二十条对“深度合成服务提供者开发上线具有舆论属性或者社会动员能力的新产品、新应用、新功能的”也提出了开展安全评估的要求，该规定和 2019 年出台的《网络音视频信息服务管理规定》中对“网络音视频信息服务提供者”，以及 2021 年发布的《网络直播营销管理办法（试行）》对“直播营销平台”使用虚拟形象从事网络直播营销的要求实质一致。

其中，深度合成服务提供者何时构成具有舆论属性或者社会动员能力的算法推荐服务提供者尚不明确。2018 年国家网信办的《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》一文将“论坛、博客、微博客、聊天室、通讯群组、公众账号、短视频、网络直播、信息分享、小程序等信息服务或者附设相应功能”均定义为“具有舆论属性或社会动员能力的互联网信息服务”。结合已生效的监管规定，考虑到目前监管部门对于“舆论属性”的宽泛定义，以及算法技术在虚拟人中的深度应用，虚拟数字人运营方应当密切关注立法和监管动态，在法规要求的情况下及时做好算法备案和安全评估。

（三）使用者需实名认证，同时深度合成服务提供者应当制定并公开管理规则和平台公约，完善服务协议

要求用户进行实名认证是所有网络运营者的义务，《深度合成规定》亦进行重申，不进行真实身份信息认证的，深度合成服务提供者不得为其提供信息发布服务。除此之外，深度合成服务提供者应当制定并公开管理规则和平台公约，完善服务协议，以显著方式提示深度合成服务使用者信息安全义务，并依法依约履行相应管理责任。

（四）加强对深度合成信息内容管理

《深度合成规定》要求深度合成服务提供者强化内容审核，具体措施包括：采取技术或者人工方式对深度合成服务使用者的输入数据和合成结果进行审核；建立健全用于识别违法和不良深度合成信息内容的特征库，完善入库标准、规则和程序；对违法和不良信息依法采取相应处置措施，并对相关深度合成服务使用者依法依规采取警示、限制功能、暂停服务、关闭账号等处置措施。

根据深度合成的具体应用场景，深度合成服务提供者还应遵守该领域相对应的法规。例如，虚拟数字人商业化经营中涉及提供互联网新闻信息服务的，应按照《互联网信息服务算法推荐管理规定》，“规范开展互联网新闻信息采编发布服务、转载服务和传播平台服务，不得生成合成虚假新闻信息⁴”。

（五）遵守个人信息保护的相关规定，使用个人信息需要取得事先同意

如我们在之前系列文章 [《虚拟数字人系列之二 人格权篇》](#) 所提及，已有多款国内外软件已经开发出“捏脸”技术，在使用过程中需要全方位收集面部特征信息等私密信息和生物识别信息。与《个人信息保护法》的规定相一致，《深度合成规定》中再次明确：“深度合成服务提供者提供人脸、人声等生物识别信息的显著编辑功能的，应当提示深度合成服务使用者依法告知并取得被编辑的个人信息主体的单独同意。法律、行政法规另有规定的除外。”据此，提请虚拟数字人运营方特别注意，若所使用的深度合成技术或其他技术需要收集或以任何形式使用人声、人脸等生物识别特征时，运营方需履行全面告知义务并提前取得个人信息主体的“单独同意”，且按照《个人信息保护法》第五十五条规定，需事前进行个人信息保护影响评估。

（六）遵守数据安全管理的规定，加强训练数据管理

虚拟数字人的创造及商业化过程中多要素、多场景均依赖深度合成技术，而深度合成技术的模型训练需要提取大量原始数据。针对网络数据处理活动，我国已有《中华人民共和国数据安全法》、《网络数据安全管理条例（征求意见稿）》等法律法规进行专门规定。虚拟数字人的运营方若符合“为用户提供信息发布、社交、交易、支付、视听等互联网平台服务”⁵的特征，则将因其“互联网平台运营者”的主体身份而受到《网络数据安全管理条例（征求意见稿）》等规范的特别规制，如恰当披露平台规则、隐私政策和算法策略，若利用人工智能、虚拟现实、深度合成等新技术开展数据处理活动的需进行安全评估等。

除此之外，《深度合成规定》中还首次明确了“深度合成服务提供者应当加强训练数据管理，确保数据处理合法、正当，采取必要措施保障数据安全。”虽然该条及《深度合成规定》中并未就深度合成服务提供者进行训练数据管理的细则要求进行明确，但作为虚拟数字人运营方，我们理解其涉及大量数据的收集、存储、使用、施工、传输及对外提供等数据生命周期的多个环节，运营方需要从我国数据安全监管的整体框架着眼履行数据安全监管领域对数据处理者设定的各类义务和要求。

同时提请注意，若虚拟数字人运营方在商业化过程中涉及跨境业务或虚拟数字人商业布局于公共通信和信息服务、金融、电子政务、医疗、汽车、工业和信息化领域等重要行业或特别监管行业，运营方还需注意该等行业的特别监管制度并需关注相关数据是否属于重要数据、运营方是否会被定义为关

⁴ 《互联网信息服务算法推荐管理规定》第十三条。

⁵ 《网络数据安全管理条例（征求意见稿）》第七十三条“本条例下列用语的含义：（九）互联网平台运营者是指为用户提供信息发布、社交、交易、支付、视听等互联网平台服务的数据处理者。”

键信息基础设施运营者等。

（七）建立健全辟谣机制并提供投诉举报入口，及时受理、处理并反馈处理结果

《深度合成规定》要求深度合成服务提供者应当建立健全辟谣机制，发现深度合成信息服务使用者利用深度合成技术制作、复制、发布、传播虚假信息的，应当及时采取相应的辟谣措施，并将相关信息报网信等部门备案。此规定承袭了此前《网络音视频信息服务管理规定》第十三条的要求，并与虚拟数字人运营方对“虚假信息”的内容管理义务相衔接。

同时为便利用户申诉，深度合成服务提供者应当设置便捷有效的用户申诉和公众投诉、举报入口，公布处理流程和反馈时限，及时受理、处理并反馈处理结果。

三、结语

在本虚拟数字人系列文章中，我们梳理了虚拟数字人的发展历程，以人格权益为引剖析虚拟数字人的“人”之属性，从商业化实例入手分享知识产权布局路径，在广告代言和直播带货场景下讨论责任承担方，就虚拟偶像演艺活动探讨周全之道，并在本篇文章中总结了深度合成技术合规的相关问题。无论全民元宇宙时代是否会到来，伴随着新技术的发展和新场景的出现，总会有新的观点碰撞，亦会有新的共识交融。届时法律法规又会给我们怎样的回应，我们翘首而望！

特别声明

汉坤律师事务所编写《汉坤法律评述》的目的仅为帮助客户及时了解中国或其他相关司法管辖区法律及实务的最新动态和发展，仅供参考，不应被视为任何意义上的法律意见或法律依据。

如您对本期《汉坤法律评述》内容有任何问题或建议，请与汉坤律师事务所以下人员联系：

段志超

电话： +86 10 8516 4123

Email: kevin.duan@hankunlaw.com