Journal of Economic Governance and Business Studies
ISSN 3134-5603

AIGC 时代新闻真实性的危机与应对策略研究

作者简介：符伊娜

工作单位：浙江越秀外国语学院，新闻学专业。

摘要

随着生成式人工智能技术快速迭代，AIGC 广泛应用于新闻采集、内容写作、素材剪辑与传播分发全流程，重塑新闻行业生产模式。AIGC 凭借高效、低成本的内容生产优势，为新闻业赋能的同时，也带来新闻真实性层面的全新挑战。虚假图文、深度伪造音视频、AI 自动生成不实信息大量涌现，模糊事实与虚构的边界，冲击传统新闻真实性准则。本文采用文献研究法、案例分析法，梳理 AIGC 技术在新闻领域的应用现状，剖析新闻真实性危机的多重表现，挖掘危机背后技术、行业监管、媒体伦理、受众媒介素养等深层成因，探索适配 AIGC 时代的新闻真实性保障路径，以期为新闻行业规范使用生成式人工智能、守住新闻真实底线提供理论参考。

关键词

AIGC；新闻真实性；深度伪造；媒介伦理；新闻生产

引言

真实性是新闻的生命，是新闻行业存续发展的核心根基。自近代新闻行业诞生以来，无论媒介形态如何演变，报纸、广播、电视再到互联网新媒体，“新闻必须忠于客观事实”始终是新闻从业者恪守的基本准则。进入数字时代，人工智能技术持续突破，以 ChatGPT、AI 绘图、AI 视频生成工具为代表的 AIGC，能够依托大语言模型快速生成文字、图片、音频、视频等

多模态内容，大幅降低内容创作门槛。传统新闻生产流程中，采访、撰稿、剪辑、配图等环节往往耗费大量人力与时间，而 AIGC 工具可以在数分钟内产出完整新闻稿件、新闻配图甚至新闻短片，被众多自媒体、新闻机构尝试投入日常新闻生产。

技术本身具备中立属性，但 AIGC 存在模型幻觉缺陷，生成内容可能凭空编造人物、事件、数据，在缺乏人工严格核验的前提下直接对外发布，极易催生虚假新闻。深度伪造技术可以复刻人物语音、面部形象，制作以假乱真的伪造新闻视频，普通受众仅凭感官很难分辨真伪。大量虚假 AI 新闻借助社交平台快速裂变传播，不仅误导公众认知，扰乱公共舆论环境，还会损害媒体公信力，消解大众对于新闻内容的信任。

当前国内学界已经围绕 AIGC 与新闻业开展大量研究，学者们从技术伦理、新闻生产变革、版权保护等多个维度展开论述，但针对新闻真实性危机的系统性梳理，以及落地可行的应对策略还有进一步探讨空间。基于此，本文结合近年 AIGC 虚假新闻典型案例，分析 AIGC 给新闻真实性带来的各类风险，挖掘危机产生的多层原因，并从媒体机构、新闻从业者、监管部门、受众四个层面提出对策，探索新技术环境下维护新闻真实性的可行方案。

一、AIGC 在新闻领域的应用现状

1.1 辅助新闻稿件撰写，简化文字生产

AIGC 大语言模型最基础的应用场景便是文字生成。在新闻行业，很多媒体利用 AI 完成快讯、数据新闻、体育赛事快讯、财经简报等标准化稿件写作。例如股票行情、赛事比分、气象快讯等信息固定、格式模板化的新闻，AIGC 可以抓取公开数据源，自动整理信息、生成稿件，减少记者重复性文字工作，让记者把更多精力投入深度采访、调查报道。部分媒体也利用 AI 完成资料整理、文献梳理、采访纪要转写等辅助工作，提升新闻采编整体效率。

但是该模式存在明显短板，大模型的训练数据来源于互联网海量文本，当数据源本身混杂不实信息，模型输出内容就可能出现事实错误。模型幻觉会凭空捏造事件细节、引用不存在的采访对象、编造数据，若编辑未逐条核实，直接发布 AI 稿件，虚假新闻就此产生。

1.2 多模态内容生成，丰富新闻素材

除文字之外，AIGC 可以生成新闻配图、新闻插画、短视频、虚拟主播播报视频。传统新闻配图需要实地拍摄或者图片采购，AI 绘图工具可以快速根据文字描述生成图像素材；虚拟数字主播技术，能够实现 24 小时不间断新闻播报，降低演播室运营成本。多模态 AIGC 丰富新闻表达形式，适配短视频平台的传播特点，契合新媒体时代受众碎片化阅读习惯。

多模态内容的伪造风险远高于纯文字内容。AI 图片、深度伪造视频具备极强迷惑性。伪造图片可以虚构灾难现场、集会场景；深度伪造视频可以伪造公众人物发言片段。图片和视频一直被大众视作“眼见为实”的证据，当视听素材能够被 AI 随意篡改，传统依靠影像素材判断新闻真伪的方式失效，新闻真实性的判断难度显著提升。

1.3 用于新闻分发与个性化推送

AIGC 技术还可以辅助内容标签化，根据受众兴趣，自动拆分、改写新闻内容，适配不同平台传播风格，完成个性化推送。算法结合 AIGC，会优先推送受众感兴趣内容，容易形成信息茧房。在信息茧房环境中，受众只接收符合自身固有认知的信息，虚假新闻更容易获得传播土壤，事实核查的声音难以触达受众。

二、AIGC 时代新闻真实性危机的具体表现

2.1 模型幻觉催生事实失实新闻

模型幻觉是大语言模型与生俱来的缺陷。AIGC 生成内容并不是基于对客观事实的理解，而是根据海量文本数据，预测文字之间的关联概率。当提问信息缺少足够参考资料，模型就会编造信息，虚构人名、时间、地点、数据、事件过程。部分自媒体为追求流量，直接使用 AI 生成社会热点类稿件，不经事实核查就发布，大量编造的细节混入新闻，造成新闻事实失真。这类虚假新闻往往逻辑通顺，文本阅读起来没有明显漏洞，普通人很难识别内容虚构部分。

2.2 深度伪造技术制造虚假视听新闻

深度伪造（Deepfake）是 AIGC 多模态技术带来最突出的风险。借助 AI 换脸、语音合成技术，能够制作伪造视频，模拟人物说出从未讲过的话语。在新闻传播场景中，伪造公众人物讲话视频、伪造突发事件影像，极易引爆网络舆论。此类虚假新闻传播速度极快，往往谣言大范围扩散之后，权威媒体才发布辟谣信息，而谣言已经对当事人名誉、社会舆论造成负面影响。

传统新闻行业，影像素材作为新闻事实的重要凭证，视频和照片被认为可信度更高。深度伪造打破“影像记录真实”的固有认知，一旦伪造内容大规模泛滥，受众会对所有影像新闻产生怀疑，真实新闻素材也会被质疑，整体媒体公信力被消耗。

2.3 新闻来源模糊，溯源难度提升

传统新闻拥有清晰来源，新闻会标注采访对象、信息出处、影像拍摄者，受众和媒体从业者可以追溯消息源头，核验真伪。而 AIGC 生成内容很难溯源。AI 生成的图文视频，无法直接定位原始素材，内容经过多次改写、二次 AI 加工之后，信息链条断裂。很多自媒体转载 AI 新闻时，不会标注原始来源，多层转发之后，很难判断信息最初出自哪里，事实核查工作难度大幅增加。

2.4 虚假新闻量产，放大谣言传播规模

AIGC 极大降低虚假新闻制作成本。过去制作一则虚假新闻，需要人为撰写、修图剪辑，耗费较多时间精力；现在只需要输入简单提示词，短时间内一次性产出大量不同版本虚假稿件、配图。低成本、批量化生产，让虚假新闻数量爆发式增长。海量虚假信息充斥社交平台，媒体事实核查人员人力有限，很难做到逐条甄别，谣言管控压力显著加大。

三、AIGC 时代新闻真实性危机产生的深层原因

3.1 技术层面：人工智能本身存在局限性

当前 AIGC 大模型不具备认知能力，无法区分客观真实与虚构创作。模型学习网络数据，互联网本身包含大量不实自媒体内容、谣言、错误信息，训练数据集自带错误，模型输出自然会继承偏差。同时，模型无法自主进行事实核验，不会主动辨别内容是否符合客观现实，只能按照语言规律生成文本。多模态生成模型在图像、视频细节处理上，可以以假乱真，进一步放大伪造风险。

3.2 行业层面：流量导向弱化媒体把关机制

新媒体环境下，部分自媒体、小型内容机构以流量收益为核心目标。流量变现的商业模式，促使创作者追逐热点，优先追求内容产出速度，压缩事实核查环节。传统媒体内部拥有多层把关人，记者撰稿、编辑审核、值班总编终审，层层校验事实。而自媒体生产链条短，很多账号没有专职事实核查人员，直接将 AI 生成内容一键发布，把关环节缺失，虚假新闻直接流入公共传播空间。

同时，部分媒体机构盲目追求生产效率，过度依赖 AI 工具，把 AIGC 当成直接产出成品新闻的工具，忽视人工核验，把关流程简化，事实出错概率上升。

3.3 监管层面：相关规范滞后，识别技术不足

AIGC 技术发展速度远超法规完善速度。近几年国内陆续出台生成式人工智能服务管理暂行办法，对 AIGC 内容提出规范要求，但网络内容体量巨大，平台人工审核压力大。识别 AI 生成内容的检测工具存在短板，部分 AI 内容检测软件识别准确率有限，面对经过改写、二次加工的 AIGC 内容，识别效果下降。跨平台传播的虚假新闻，跨平台溯源、处置存在难度，谣言传播之后追责存在困难。

3.4 受众层面：媒介素养参差不齐，辨别能力有限

普通受众缺少专业新闻辨别训练，很难分辨 AI 生成虚假内容。多数受众习惯依靠直觉判断新闻真假，看到配图、视频就默认事件属实。面对社会热点事件，受众容易被情绪带动，在没有核实信息真伪情况下转发内容，助推虚假新闻传播。还有部分受众信息茧房效应明显，只愿意相信符合自身情绪和立场的内容，就算出现辟谣信息，也不愿意接受真相，虚假新闻

持续发酵。

四、AIGC 时代维护新闻真实性的应对策略

4.1 媒体机构：完善内部 AI 使用规范，重建把关流程

媒体机构需要制定明确的 AIGC 使用准则，界定 AI 在新闻生产中的使用边界。AI 只能作为辅助工具，用于资料整理、初稿草拟、素材检索，禁止 AI 直接生成成品新闻发布，所有 AI 产出内容，必须由持证新闻从业者逐条核验事实、核对来源。媒体应当建立专门事实核查岗位，针对 AI 生成内容做专项校验，核查时间、地点、人物、数据、引文，确认全部信息真实无误后，才能发布。

同时坚持标注 AI 生成内容。凡是新闻稿件、配图、视频由 AIGC 参与创作，需要在内容显著位置标注 AI 参与创作说明，保障受众知情权，区分 AI 生成素材与实拍采访素材。媒体内部完善追责机制，因 AI 内容未经核查发布虚假新闻，追究相关编辑、责任人责任。

4.2 新闻从业者：提升 AI 媒介素养，坚守新闻伦理

新闻专业学生与在岗记者，需要主动学习 AIGC 相关知识，了解大模型幻觉、深度伪造的风险，掌握 AI 内容识别方法。不能过度依赖 AI 工具，牢记新闻采访、实地调查、多方求证的专业基本功。在使用 AI 辅助写作时，要对 AI 输出全部信息溯源，交叉验证多个信息源，不能直接引用 AI 编造的资料。

新闻从业者坚守新闻伦理，恪守客观、真实、公正的职业准则，拒绝为流量刻意制造博眼球的虚假内容。在发现 AI 虚假新闻时，积极发布辟谣报道，澄清事实，引导舆论。作为新闻学专业人才，也要持续关注技术变革带来行业变化，持续更新专业知识体系。

4.3 监管部门与平台：健全法规，升级 AI 检测技术

监管层面持续细化 AIGC 内容管理规则，压实平台主体责任。各大社交平台、资讯平台，升级 AI 内容识别系统，建设深度伪造图片、视频检测能力，对明显伪造的虚假新闻进行拦截。建立虚假新闻快速辟谣通道，当重大事件出现 AI 谣言，快速推送权威辟谣信息。完善溯源机制，对刻意制作传播深度伪造虚假新闻、恶意造谣的主体依法追责。

平台完善内容发布规则，要求创作者标注 AIGC 生成内容；对于批量发布 AI 虚假新闻的账号，采取限流、封禁等处罚手段，减少谣言传播源头。

4.4 面向公众开展媒介素养教育，提升受众辨别能力

高校、媒体、教育机构共同开展全民媒介素养科普，普及 AIGC 虚假新闻识别技巧。教会受众辨别 AI 生成图文、深度伪造视频，提醒大众不能单凭图片视频判断新闻真伪，遇到热点消息优先查看权威媒体发布信息，不要随意转发未经核实网络消息。高校新闻专业，例如新闻学专业课程，可以增设人工智能与新闻传播相关课程，培养未来新闻人才，同时面向社会开展科普，提升全社会辨别 AI 谣言的能力。

结论

AIGC 是新闻行业变革的重要技术力量，它提升新闻生产效率，拓展新闻表达形式，给新闻行业带来全新机遇，但同时冲击新闻真实性底线。模型幻觉、深度伪造、批量生产等问题，制造大量虚假新闻，挑战传统新闻把关模式，损害媒体公信力。

新闻真实性是不可动摇的行业底线，技术发展不能牺牲事实真实。面对 AIGC 带来的危机，不能一味抵制技术，而要理性看待，善用技术优势，管控技术风险。媒体机构完善 AI 使用规范，守住内容把关关口；新闻从业者夯实专业能力，恪守新闻伦理；监管平台完善法规与检测技术；提升全民媒介素养，多方协同发力。

技术一直在迭代，新闻行业需要持续调整规范，在拥抱人工智能技术的同时，牢牢守住新闻真实性这一核心准则，推动新闻行业在智能时代健康可持续发展。

参考文献

[1] 彭兰.生成式人工智能对新闻业的影响、风险与伦理调适[J].新闻与写作,2023(04):5-14.

[2] 喻国明.AIGC 对新闻生产的重构：逻辑、风险与治理[J].新闻爱好者,2023(06):11-18.

[3] 张洪忠,王廷威.深度伪造技术对新闻真实性的挑战及治理路径[J].国际新闻界,2023,45(07):23-42.

[4] 陈昌凤.人工智能时代新闻伦理的新议题[J].新闻与传播研究,2022,29(12):5-22+126.

[5] 方师师.算法虚假新闻：生成逻辑、传播机制与治理对策[J].现代传播(中国传媒大学学报),2023,45(02):138-145.

[6] 胡泳.大语言模型与新闻业的未来[J].新闻大学,2023(03):1-12+117.

[7] 李良荣.新闻学概论（第八版）[M].上海：复旦大学出版社,2021.

[8] 周勇.智能媒体时代新闻真实性的内涵重构[J].电视研究,2023(05):28-31.

[9] 何苏六.多模态 AIGC 的传播风险与内容规制[J].新闻记者,2023(08):46-56.

[10] 林爱珺.人工智能新闻的版权、伦理与真实性问题[J].新闻界,2023(01):34-43.

[11] 史安斌.全球视野下深度伪造虚假信息的治理研究[J].新闻与传播评论,2023,76(04):5-18.

[12] 刘建明.新闻真实理论再探：数字智能时代的新挑战[J].当代传播,2023(02):22-27.