从失范到规范：生成式人工智能的监管框架革新

刘学荣

（吉林大学　法学院，吉林　长春　130000）

摘　要：生成式人工智能在技术变革下引发的失范性风险，对既有人工智能监管框架提出了挑战。从底层技术机理出发，可知当前生成式人工智能呈现出“基础模型－专业模型－服务应用”的分层业态，分别面临算法监管工具失灵、训练数据侵权风险加剧、各层级间法律定位不明、责任界限划分不清等监管挑战。为此需以分层监管为逻辑内核，对我国既有人工智能监管框架进行革新。在监管方式上应善用提示工程、机器遗忘等科技监管工具；在责任划定上应进行主体拆解与分层回溯，从而规范“基础模型－专业模型－服务应用”的分层监管框架，以期实现有效监管，促进生成式人工智能的高质量发展。

关键词：生成式人工智能；算法黑箱；技术监管；法律责任

中图分类号：D922；TP399　　**文献标识码：**A　　**DOI：**10. 19358/j. issn. 2097－1788. 2024. 06. 009

引用格式：刘学荣. 从失范到规范：生成式人工智能的监管框架革新［J］. 网络安全与数据治理，2024，43(6)：58－63，71.

From illegal to legal: evolving regulatory frameworks for generative artificial intelligence

Liu Xuerong

(School of Law, Jilin University, Changchun 130000, China)

Abstract: The risk of aberration caused by generative artificial intelligence under technological change challenges the existing artificial intelligence regulatory system. Starting from the underlying technical mechanism, it can be seen that the current generative artificial intelligence presents a hierarchical format of "basic model-professional model-service application", and faces regulatory challenges such as the failure of algorithm supervision tools, the intensified risk of training data infringement, the unclear legal positioning between different levels, and the unclear division of responsibility boundaries. Therefore, it is necessary to take layered regulation as the logical core and reform the existing artificial intelligence regulatory framework in China. In the way of supervision, we should make good use of technology supervision tools such as prompt engineering and machine forgetting. In the delineation of responsibilities, the main body should be disassembled and hierarchical backtracking should be carried out, so as to standardize the hierarchical regulatory framework of "basic model-professional model-service application", in order to achieve effective supervision and promote the healthy and high-quality development of generated artificial intelligence.

Key words: generative artificial intelligence; algorithm black box; technical supervision; legal responsibility

0　引言

随着人工智能的迭代升级，对其进行的深层监管不仅关系到法律治理实效，也直接影响到技术发展与应用安全。生成式人工智能作为当前新质生产力发展的主要驱动力，需加以重点关注。相较于传统的人工智能，生成式人工智能因其深度学习属性而使技术原理变得更加复杂且难以理解，并由此导致算法黑箱、算法歧视、算法异化、算法权力失范等过去人工智能算法模型中常见的技术伴生风险问题更为严峻。与此同时，算法解释、算法审计、算法评估等过去对人工智能进行法律监管的传统工具在生成式人工智能面前也面临着失灵风险，法律监管体系的稳定性与安全性都受到了极大的冲击。

虽然我国人工智能法律监管始终走在世界前沿，并形成了具有中国特色的算法模型监管体系[1]，但就目前

针对生成式人工智能以及深度合成算法推出的监管规定，仍主要停留在人工智能模型治理衍生出的信息安全层面，偏重服务应用监管而轻视底层技术监管[2]，无法克服因人工智能模型的技术升级而产生的监管困境。

在技术失控风险日益严重，现有方案又无法实现有效监管的双重困境下，生成式人工智能的监管难度急剧增长。面对生成式人工智能蓄势待发的落地应用，需要针对性的法律监管方案对风险进行治理。因此，本文将从生成式人工智能的底层技术出发，首先对其采用的算法模型进行技术穿透，在解析技术原理后清晰定位生成式人工智能的监管困境，而后在底层技术特征的基础之上挖掘生成式人工智能技术监管的可行路径，弥补当前生成式人工智能法律监管工具的失灵，并结合生成式人工智能的底层运行机理与相应的运行主体进行精准分层责任落实，避免因“技术中立”滥用而引发法律责任逃避问题，以实现底层技术与分层主体有机协调的法律监管模式。

1 生成式人工智能的技术原理与监管困境

技术是所有“实现目的的手段”的总体[3]，要想实现对生成式人工智能的底层技术监管，避免陷入过去传统算法规制的思维惯性，应从科技现实主义角度出发[4]，首先实现技术穿透，弄清生成式人工智能所采用的算法模型到底是什么，与传统人工智能使用的经典算法模型有何区别。在掌握生成式人工智能的算法模型运行机理之后，明确生成式人工智能的风险来源，直达生成式人工智能的法律监管痛点所在。

1.1 生成式人工智能的技术原理

当前，生成式人工智能主要采用的算法模型是生成式对抗网络（Generative Adversarial Networks，GANs）和生成式预训练转化器（Generative Pre-trained Transformer，GPT）[5]。以生成式对抗网络为例，其有两个组成部分：一个生成器，即生成式模型，学习生成虚假数据；一个鉴别器，即判别式模型，学习区分虚假信息，并且采用神经网络进行训练，使深度合成功能逐步增强。其训练是将模型的数学公式设计好之后，寻找合适的参数，使模型对指定数据集的评估或分类结果与真实情况的差距达到最小化的过程。这一过程导致算法背后的专业技术人员都无法解释算法模型运行的具体原理与方式，沦落到只能依靠控制算法中的某些参数来尝试对算法模型进行调整。由此可见，生成式人工智能监管的复杂性提升至一个新的高度，这无疑加重了非专业人士对生成式人工智能进行监管的难度与挑战。

1.2 生成式人工智能的监管困境

我国对人工智能的监管规范主要集中在确保人工智能的安全性、使用的透明性、算法的可解释性以及技术的伦理性等方面，常见的技术监管工具为算法审计、算法评估和算法解释等。但这些监管工具在生成式人工智能的算法黑箱面前均面临着技术失灵风险，同时也无法满足生成式人工智能所产生的训练数据与隐私侵权等问题的监管需求，使生成式人工智能深陷技术监管工具失灵及法律监管缺位的桎梏。

以算法解释为例，过去对算法模型监管的探讨中，“打开黑箱”是算法监管所追求的目标，大多学者主张算法模型提供者应履行算法解释义务，提升算法透明度，以加强算法模型的可信任性，认为算法解释是治理算法模型的得力工具之一，甚至有学者曾言“算法解释是人工智能治理研究的圣杯”[6]。但实质上，过去试图“打开黑箱”的算法监管方式在生成式人工智能面前是一条行不通的路径。如今生成式人工智能所采用的算法模型是经典的黑箱算法，因其技术底层蕴含着流形分布定律和学习概率分布，其所依托的技术理论本身尚未发展成熟，算法解释与法律监管的难度极高，经常需要放弃理论上的最优性，多数情况下严重依赖于调参以维持理论的稳定性，算法解释技术的抓手便更无法触及技术理论底层，算法解释的动力学机制逐渐失灵。且算法解释作为一种技术手段，并不是为算法模型监管而专门设计的，其本身受被解释对象的技术条件发展限制，算法解释的技术能力划定了可解释的最大边界，当解释技术的发展追赶不上被解释对象的技术发展时，算法解释便无法满足算法模型的监管需求[7]。诚然，没有合适的解释技术，规范层面的多数制度设计都将无从谈起，但如果继续按照我国大多数学者所讨论的原有算法监管思路，期待仅依赖算法解释工具实现对生成式人工智能算法模型的监管[8]，无疑加重了算法开发者所承担的解释义务，在一定程度上或将阻碍生成式人工智能的落地应用，并使算法模型监管深陷解释的泥沼无法自拔。

除此之外，生成式人工智能呈现出“基础模型-专业模型-服务应用”的分层业态（如图1所示），无法在我国过去基于平台治理而形成的“技术支持者-服务提供者-内容生产者”的监管框架中找到适配的法律定位[9]。并且生成式人工智能的基础模型层、专业模型层产生了同为技术支持者的法律责任复合状态，亟须对其进行分层责任抽离，以期实现对生成式人工智能的精准法律监管，减轻技术端的注意义务，促进产业发展，并且鼓励企业根据不同的业态层次承担不同水位的风险防范义务与法律责任。

因此面对即将发生的生成式人工智能数据侵权与隐私安全法律监管困境，尚无有效的技术规制手段，这就需要从底层技术监管角度出发，将现有人工智能的监管

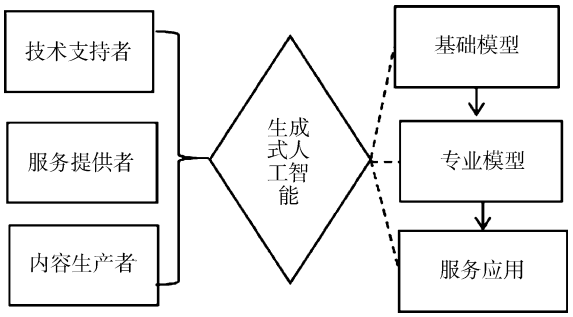

图 1　生成式人工智能的分层业态

工具进行技术融贯，以适应因技术发展而带来的监管工具缺陷，同时明确生成式人工智能在不同业态层次下的责任主体、归责原则以及问责限度。

2　生成式人工智能的底层监管路径

技术带来的问题在技术发展中也会被逐步解决。倘若一项新兴技术引发的法律问题能够依靠技术手段予以内部优化，那么监管路径应该以技术为先，外部政策与法律规制无需过早介入[10]。生成式人工智能伴随着算法模型的复杂升级而出现了显著性能飞跃，传统的法律监管工具难以实现有效监管。但若以技术控制先行的理念进行监管优化，可最大限度地转化生成式人工智能所带来的技术发展优势，挖掘底层技术对生成式人工智能进行监管的可能性，充分利用生成式人工智能的技术与工具属性，发挥人类的主观能动性，优化技术变革与法律监管之间的平衡发展。

2.1　算法黑箱监管工具：提示工程（Prompt Engineering）

因生成式人工智能算法模型是典型的黑箱算法模型，过去常见的算法解释监管工具在此刻逐渐失灵，须寻求其他监管工具来解决现有算法模型监管中因黑箱问题所带来的输出失控风险与输出责任问题。在此背景下，应优化提示工程的使用，在满足高质量输出的基础上保障急用先行，建立生成式人工智能高质量发展的法律监管体系。

提示工程在生成式人工智能领域属于一种编程方式，是算法黑箱问题最为可行的优化工具，其不仅能作为输出内容质量的监管手段，也可作为输出责任问题的法律保障依据。首先在输出内容质量监管上，算法模型提供者可根据问题类型与其使用的底层算法模型架构不同的提示技巧，使输出内容能够根据用户需要而精准生成，直接影响生成式人工智能算法模型的生成结果和质量，进而实现对算法黑箱问题的优化。在我国“AI 服务平台”著作权纠纷案中，法院便指出被告应进一步采取关键词过滤等措施，防范其服务继续生成与涉案作品实质相似的图片，使算法黑箱的防范程度达到用户正常使用涉案提示词，而不能生成与涉案作品实质相似的图片。而其中所要求的关键词过滤、提示词选取等监管措施均在提示工程的范畴之内，可见提示工程作为输出内容的监管工具，已得到了司法层面的认可，并逐步在司法监管实践中展开应用。且随着生成式人工智能产品的场景化应用，提示工程逐渐发展成为了体系[11]，出现了 PromptPerfect、PromptHero、ClickPrompt 等多个专业网站，用以辅助提示词设计和提示方式优化，实现了对生成式人工智能输出内容质量监管的进一步优化。

其次在输出内容责任问题上，因提示工程包括提示词的选取、设计与对话修正等内容，属于人类智慧主导下的创作媒介，故而成为了我国法律层面中人类“独创性”属性的司法确认依据。在我国“AI 文生图”著作权纠纷案中，法院便首次从提示工程的角度出发，在提示词选取和提示方式设计中认定了生成式人工智能应用过程中的人类“独创性”，做出了与美国法院截然相反的判决，确立了我国人工智能生成物的可版权性法律判断标准，进而为后续生成式人工智能输出内容的法律责任划分奠定基础。

目前，提示工程的主要表现形式为语言文本提示，主要运行模式为正向提示词和反向提示词的输入。但在技术层面上，提示工程并非简单地停留在输入语言文本提示。算法模型提供者也会深入到算法模型内部，配置一些参数来调整输出不同的结果，如 Temperature 调整和 Top-p 调整，甚至在一些高阶提示中，自动提示工程（Automatic Prompt Engineer）可以将指令生成问题定义为自然语言合成，作为算法模型黑箱优化问题的候选解决方案。为此，生成式人工智能必须为提示工程留足空间，弥补算法解释工具在生成式人工智能时代的失灵，优化算法黑箱监管难题，防止生成式人工智能的放任输出以致毁灭其应用生态。并且，有效的提示工程还可以减少算法偏见和算法歧视等常见的算法模型输出失控风险，确保生成式人工智能生成内容符合伦理性要求。

2.2　数据与隐私保障工具:机器遗忘（Machine Unlearning）

根据《生成式人工智能服务管理暂行办法》（以下简称《暂行办法》）第 14 条规定：提供者发现违法内容的，应当及时采取停止生成、停止传输、消除等处置措施，采取模型优化训练等措施进行整改，并向有关主管部门报告。可见，《暂行办法》中也对生成式人工智能的发展发出了底层技术治理的呼唤，提出了对违法内容采用算法模型优化训练等措施进行整改，这一点在司法实践中也得到了印证。在我国“AI 服务平台”著作权纠纷

案中，原告向法院主张将涉案数据从生成式人工智能的训练数据集中删除，而囿于被告未实际进行模型训练行为而未得到法院采纳。这从侧面证明实际进行模型训练的开发者有义务对模型进行停止生成和消除等优化措施。

与机器学习相对，机器遗忘是指出于隐私、可用性和被遗忘权利的需要，将有些特定样本信息从模型中移除，满足机器学习因隐私、知识产权、可用性或其他权利要求而带来的从模型中删除训练样本的请求[12]。机器遗忘相较于传统的训练数据集删除方式，可以凭借技术路径最大限度地降低模型使用成本。因训练数据集删除后模型开发者往往面临的是模型重新训练的巨额成本，而机器遗忘则仅意味着某些任务性能变化的微调成本，且二者最终都能取得数据删除的效果，使得开发者以较低的成本实现模型监管过程中的数据与隐私安全保障。

目前，机器遗忘技术主要分为数据重组和模型操纵两种方式，并且在简单的机器学习模型上，比如线性模型或者调整深度学习的线性决策层上已取得了一定的成就，并有望在更为严格和复杂的生成式人工智能算法模型上取得技术突破，从而实现生成式人工智能的数据安全与隐私安全保护的技术治理路径，对底层监管需求进行技术上的回应。

机器遗忘作为最有潜力成为生成式人工智能的技术监管手段之一，倘若生成式人工智能的产业部门加以引进和吸收，则由生成式人工智能所带来的数据和隐私安全恐慌便能在一定程度上得到缓解，并且也能更好回应我国对生成式人工智能的监管要求，如《暂行办法》第11条第二款规定：提供者应当依法及时受理和处理个人关于查阅、复制、更正、补充、删除其个人信息等的请求。机器遗忘能使生成式人工智能的底层技术治理成为可能。

由此可见，技术的发展是一把双刃剑，带来风险的同时也带来无限机遇。深度合成算法模型将生成式人工智能带到人工智能的权力之巅，同时与其相应的针对性技术发展也为生成式人工智能带来了一些技术维度的监管视角，用AI监管AI在生成式人工智能时代或将不再是一个遥不可及的愿景。

3 生成式人工智能的精准分层问责

为促进技术的规范发展，应完善配套的法律监管方案。通过前文分析可知当前生成式人工智能的底层技术极具复杂性，若采纳“技术中立”的观点，则复杂的算法黑箱便成了技术服务主体的免死金牌，一旦面临算法运行产生的风险便主张技术中立来逃避法律制裁，甚至不同主体之间互相推诿，使法律责任的认定更加困难，也使技术发展面临严峻的问责挑战。因此，为实现生成式人工智能的高效监管，应采取精准的分层问责监管模式。

3.1 生成式人工智能的责任主体

关于生成式人工智能是否具有法律上的主体资格从而成为侵权责任的承担者，学界的讨论颇多。目前的主流观点认为，生成式人工智能本身仍属于技术力量，是具有使用价值的工具和手段，是算法模型开发者价值观的技术性体现，因而不宜成为法律上的适格责任主体，提供和使用工具的人才应是法律上具有主体资格的侵权责任承担者[13]。况且，如果生成式人工智能技术本身可以成为承担责任的主体，将对我国既有的法律制度框架造成颠覆性冲击，会打破我国目前已基于主体治理范式而建立的问责机制[14]。

但需注意的是，我国过去基于平台治理而形成的“技术支持者-服务提供者-内容生产者”监管框架在生成式人工智能的分层业态下逐渐消解，生成式人工智能所涉及的主体更加多元，因而学界呼唤生成式人工智能精准分层治理[15]，依照技术、行业、应用的上下游分为“基础模型-专业模型-服务应用”三层，既包括以算法技术为核心的模型基座方，也包括以相关行业数据为训练内容的专业模型方，还包括以应用场景为导向的服务应用方，甚至服务终端的使用用户，都有可能成为侵权责任的承担主体，不能简单认定其中的一方承担侵权责任，而应当进行全面分析，但具体判断标准并不明确。质言之，生成式人工智能在不同主体层次上的归责原则和问责限度都引发了新的问题。

3.2 生成式人工智能的归责原则

对于生成式人工智能归责原则的讨论，采用过错责任、过错推定和无过错责任原则的学者均有之。袁曾从责任后果产生的因果关系出发，以过错责任为核心对生成式人工智能的归责原则进行展开[16]；陈全真则从生成式人工智能的平台集权工具属性出发，追崇以过错推定原则为核心来重构侵权规则[10]；李彤则从侵权免责视角出发，认为无过错责任原则是生成式人工智能服务提供者免责事由的归责架构基础[17]。可见，当前学界对生成式人工智能的归责原则并未达成一致，形成了从过错责任出发的因果型、从过错推定出发的场景型以及从无过错责任出发的底线型三种判断标准。

但仔细剖析可知，学界之所以在归责原则上产生分歧是由于其出发的主体视角不同。有学者侧重于生成式人工智能的技术主体视角，立足于生成式人工智能服务

提供者进行归责，也有学者侧重于生成式人工智能的应用主体视角，立足于生成式人工智能服务应用者进行归责，并未对生成式人工智能背后的运行主体进行分层拆解。为此，需将生成式人工智能的归责原则进行“基础模型-专业模型-服务应用”的主体分层有效讨论，如图2所示。

首先在基础模型层，以高度复杂的算法模型基座为运行基础，足以阻碍过错责任的认定方式，用户及服务应用者很难拿出确凿的证据证明模型基座算法的侵权行为，因此在模型基座层应采用过错推定或公平责任原则进行归责[18]。面对数据侵权情形，若模型基座方所提供的算法模型不包括原始训练数据，有单独的专业数据提供方，则此时可采用公平责任原则的归责方式，与他方分担侵权损失结果；若模型基座方提供的算法模型包含原始的训练数据，此时面对数据侵权情形则须按照无过错责任原则进行追责，以确保数据源的安全。面对非数据侵权情形则应采用过错推定的归责原则，来最大限度地承担侵权所造成的损害后果，使服务应用者和使用用户得到生成式人工智能模型基座的可追责保障。

其次在专业模型层，以相关行业的数据为模型训练的主要内容具有鲜明的行业特征。面对专业模型层行业数据的侵权行为应贯彻最严厉的归责原则，即无过错责任原则。对专业模型层的训练数据提供者来说，无过错责任原则的落实有其合理的现实依据[19]。数据作为生成式人工智能的专业模型训练来源，训练数据提供者自然清楚其所使用的数据来源是否合法，是否会对他人隐私和国家安全造成威胁。尽管数据可能是从他人处进行采集，但是训练数据提供方也应进行严格的审核，对自己训练所使用的数据负有高度安全责任义务。

最后在服务应用层，服务应用者对依据算法模型运行的服务终端大多扮演监管者的角色，即监管生成式人工智能的应用对象和应用场景，若因自我监管不足而导致的侵权风险，则应采取过错责任原则的责任认定方式[20]，要求终端使用者在使用生成式人工智能产品时，具有自我权益保护意识，必要时政府部门可要求对服务应用终端进行全流程记录监管，将使用记录交由具备公信力的独立第三方政府部门留存，为终端使用者提供强有力的支持。

3.3 生成式人工智能的问责限度

在责任落实上，生成式人工智能也应根据“模型基座－专业模型－服务应用”三层进行精准分层问责限度划分，避免“一刀切”式的统一问责尺度，在符合比例原则的基础上坚持民事责任为主、刑事责任为辅的问责限度，为生成式人工智能的落地应用提供良好的发育土壤。

一是基础模型层，主要侧重基础支撑平台的搭建，包括传感器、AI芯片、算法服务和计算平台。目前生成式人工智能算法模型内部运行极其复杂，导致责任认定中的因果关系盘综错杂，很难清楚认定其是否属于侵权责任主体，因此建议技术层涉及的主体，大多扮演补充责任的角色，在合理限度范围内承担义务。并且随着我国自主研发水平与生产能力的不断提高，芯片、传感器和计算平台等技术迎来了光明的发展前景，应以鼓励支持为主，避免法律规制过度阻碍技术发展。但自由并不代表放任，模型基座中一旦涉及数据内容侵权则需承担天然的问责重担。数据的来源是否合法、是否安全、是否侵犯国家秘密或者个人隐私，都需要进行细致分析。

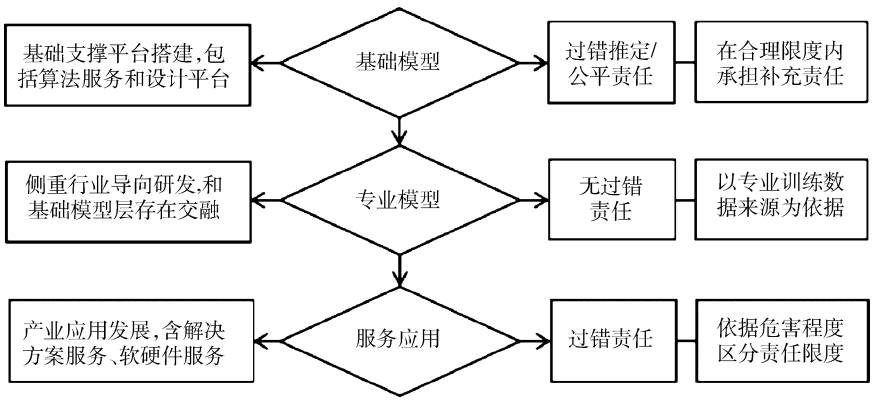

图2 生成式人工智能责任监管图谱

二是专业模型层，主要侧重行业导向的研发。虽然目前推出的 ChatGPT 是融合算法与数据于一体的生成式人工智能代表产品，但依据生成式人工智能的专业化应用趋势来看，专业模型层和模型基座层未必能保持一体化的状态进行产业发展。以法律行业的生成式人工智能专业模型为例，其在模型基座的选择上可能采取的是主流算法交互方式，但在专业模型的训练过程中，法律法规、裁判文书等数据将成为主要的数据训练对象，因不同行业数据的特殊性与保密性要求，模型基座方无法直接与专业模型方进行模型融合设计，否则法律风险和法律责任都将呈现出混杂的局面。因此应以专业训练数据来源为依据进行责任限度划分，若数据来源主要取决于专业模型方，则应由专业模型方承担主要法律责任。同时考虑到目前生成式人工智能模型中也存在不需要单独提供专业行业数据的类型，即模型基座自身含有基础训练数据，此时模型基座方对数据来源的侵权责任便难逃其咎，应和专业模型方采用不真正连带责任的方式来弥补复杂情形下侵权责任落实的缺失。

三是服务应用层，主要注重产业应用发展，包含解决方案服务、硬件产品和软件产品。作为服务应用提供者，将生成式人工智能模型应用至不同场景，如医疗、金融、教育等主流领域中，所引发的侵权后果危害程度不尽相同，因此可参照美国的《算法问责法案》逻辑，要求对“高风险”的应用系统进行影响评估，对应用对象、应用场景进行具体规制，依据应用结果危害程度的不同执行差异化的责任承担限度。

4 结束语

生成式人工智能的全面落地应用正蓄势待发，即将为新质生产力的发展带来全新动力。但需注意的是，生成式人工智能只是高度凝聚人类智慧的技术产品，难逃其工具属性，因此面对生成式人工智能的技术失控风险，可寻求提示工程和机器遗忘等有效的技术工具进行治理与监管，从底层技术出发刺破生成式人工智能的失控面纱，从而实现高效规制。同时面对生成式人工智能的法律责任监管，首先应从问责逻辑出发，确定何种主体应为生成式人工智能的侵权结果负责；其次，因生成式人工智能所带来的侵权因果关系复杂，应对生成式人工智能的归责原则进行讨论，为实现问责的正当化与层次化，需厘清“模型基座-专业模型-服务应用”这三层不同的责任认定原则与限度，进行主体精准拆解与责任分层回溯，以完善生成式人工智能的法律监管体系；最后，在对生成式人工智能进行问责时，应以民事责任、行政责任为主，非必要不采用刑事手段进行规制，避免刑法越位，扼杀技术创新的积极性，同时兼顾比例原则，避免对生成式人工智能技术生态产生严重的内部干预，重蹈欧盟“围追堵截”式监管的覆辙。

参考文献

[1] 曾雄，梁正，张辉. 欧美算法治理实践的新发展与我国算法综合治理框架的构建［J］. 电子政务，2022（7）：67－75.

[2] 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.

[3] 张恩典. 数字时代版权的算法实施：类型、困境及法律规制［J］. 暨南学报（哲学社会科学版），2023，45（5）：35－49.

[4] 於兴中，郑戈，丁晓东. 生成式人工智能与法律的六大议题：以 ChatGPT 为例［J］. 中国法律评论，2023（2）：1－20.

[5] 毕文轩. 生成式人工智能的风险规制困境及其化解：以 ChatGPT 的规制为视角［J］. 比较法研究，2023（3）：155－172.

[6] 中国法律评论. 全国首届智能科技法治青学者论坛综述［EB/OL］.（2019－12－16）［2024－01－15］. https://www.ilawpress.com/material/detail?id=4391485218414924 80&t=material.

[7] 周翔. 算法可解释性：一个技术概念的规范研究价值［J］. 比较法研究，2023（3）：188－200.

[8] 袁曾. 生成式人工智能责任规制的法律问题研究［J］. 法学杂志，2023，44（4）：119－130.

[9] 张凌寒，于琳. 从传统治理到敏捷治理：生成式人工智能的治理范式革新［J］. 电子政务，2023（9）：2－13.

[10] 陈全真. 生成式人工智能与平台权力的再中心化［J］. 东方法学，2023（3）：61－71.

[11] 苏宇. 大型语言模型的法律风险与治理路径［J/OL］. 法律科学（西北政法大学学报），2024（1）：1－13.［2024－03－01］. http://doi.org/10.16290/j.cnki.1674-5205.2024.01.010.

[12] 郁建兴，刘宇轩，吴超. 人工智能大模型的变革与治理［J］. 中国行政管理，2023，39（4）：6－13.

[13] 郭春镇，勇琪. 算法的程序正义［J］. 中国政法大学学报，2023（1）：164－180.

[14] 张欣. 生成式人工智能的算法治理挑战与治理型监管［J］. 现代法学，2023，45（3）：108－123.

[15] 张凌寒. 生成式人工智能的法律定位与分层治理［J］. 现代法学，2023，45（4）：126－141.

[16] 袁曾. 生成式人工智能的责任能力研究［J］. 东方法学，2023（3）：18－33.

（下转第 71 页）

侦查机关面临的全新挑战。对此，《网络犯罪程序意见》等相关司法解释规定的抽样取证规则，为缓解网络犯罪侦办中司法资源紧张的局面提供了又一思路。然而抽样取证规则在实践中也暴露了诸如抽样取证启动条件不明确、程序不规范、被追诉人权利保障不健全等问题，对于以上问题、本文进行了初步探讨，从明确抽样取证适用标准、完善网络犯罪抽样取证程序、限缩抽样取证范围、赋予被追诉人抽样取证救济权几个角度提出初步建议。但囿于笔者理论水平有限以及相关司法实务经验的匮乏，未来如何进一步完善网络犯罪抽样取证规则，仍需要在理论和实践层面做深层次的探究。

参考文献

[1] 江溯. 中国网络犯罪综合报告［M］. 北京：北京大学出版社，2021.
[2] 左卫民. 反思过度客观化的重罪案件证据裁判［J］. 法律科学（西北政法大学学报），2019，37（1）：112－122.
[3] 熊晓彪. 概率推理：实现审判智能决策的结构化进路［J］. 中外法学，2022，34（5）：1278－1298.
[4] 万毅，纵博. 论刑事诉讼中的抽样取证［J］. 江苏行政学院学报，2014（4）：120－128.
[5] 林喜芬. 大数据证据在刑事司法中的运用初探［J］. 法学论坛，2021，36（3）：27－36.
[6] 杨帆. 海量证据背景下刑事抽样取证的法治应对［J］. 法学评论，2019，37（5）：105－112.
[7] 许昊. 从证明标准角度看刑事缺席审判制度的适用——以刑事诉讼法关于贪污贿赂犯罪缺席审判程序的规定为视角［J］. 人民司法，2019（28）：4－8.
[8] 刘品新. 网络犯罪证明简化论［J］. 中国刑事法杂志，2017（6）：24－37.
[9] 高童非. 刑事抽样证明的类型化重释［J］. 中国刑事法杂志，2022（3）：106－121.
[10] 邹列军. 电信网络诈骗犯罪多发高发原因之探讨［J］. 新型犯罪研究，2021（1）：83－85.
[11] 张曙，孔佳玉. 海量证据的抽样取证规则探析——以信息网络犯罪为视角［J］. 浙江工业大学学报（社会科学版），2023，22（2）：194－199.
[12] 王铼，雍晓明. 对利用网络进行诈骗犯罪的侦查取证问题研究［J］. 政法学刊，2010，27（1）：79－83.
[13] 于志刚. 网络思维的演变与网络犯罪的制裁思路［J］. 中外法学，2014，26（4）：1045－1058.
[14] 程龙. 论大数据证据质证的形式化及其实质化路径［J］. 政治与法律，2022（5）：96－114.
[15] 吴芳，罗斌飞. 抽样取证在网络犯罪中的应用研究［J］. 江西警察学院学报，2022（5）：82－88.
[16] 谢小剑，肖宪涛. 信息网络犯罪案件抽样取证制度的完善［J］. 法治论坛，2023（1）：253－268.
[17] 潘金贵，吴庆棒. 证据与技术：刑事抽样证明的科学面向［J］. 中国人民公安大学学报（社会科学版），2023，39（4）：45－59.
[18] 王志刚，刘思卓. 论网络犯罪证明中的数额认定方法［J］. 重庆邮电大学学报（社会科学版），2020，32（2）：35－43.
[19] 陈东升，刘向. 网络新型犯罪取证定性难题如何破解［N］. 法制日报，2017－01－19（005）.
[20] 程雷，曲育铮. 电信网络诈骗案件中两项证据审查难题及破解［J］. 人民检察，2023（17）：21－26.
[21] 马忠红. 论网络犯罪案件中的抽样取证——以电信诈骗犯罪为切入点［J］. 中国人民公安大学学报（社会科学版），2018，34（6）：69－78.
[22] 郑飞. 证据属性层次论——基于证据规则结构体系的理论反思［J］. 法学研究，2021，43（2）：123－137.
[23] 刘方权，庄嘉伟. 刑事诉讼中的见证人问题研究［J］. 中国人民公安大学学报（社会科学版），2020，36（6）：108－120.

（收稿日期：2024－01－24）

作者简介：

刘宇浩（2000－），男，硕士研究生，主要研究方向：诉讼法学、行政法与地方法制。

（上接第63页）

[17] 李彤. 生成式人工智能技术提供者侵权免责事由的识别重整［J］. 南京社会科学，2024（2）：86－97.
[18] 陈英达，王伟. 由“急用先行”走向“逐步完善”：生成式人工智能治理体系的构建［J］. 电子政务，2024（4）：113－124.
[19] 朱荣荣. 类 ChatGPT 生成式人工智能对个人信息保护的挑战及应对［J/OL］. 重庆大学学报（社会科学版）：1－14.［2024－03－01］. http://kns.cnki.net/kcms/detail/50.1023.C.20230921.1151.002.html.
[20] 刘文杰. 何以透明，以何透明：人工智能法透明度规则之构建［J］. 比较法研究，2024（2）：120－134.

（收稿日期：2024－03－14）

作者简介：

刘学荣（1999－），女，硕士研究生，主要研究方向：知识产权法、数据法。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部