刑事司法科学与治理
Criminal Justice Science & Governance
2024 年 第 5 卷 第 2 期

生成式人工智能的法律风险与规制进路

——以 ChatGPT 为例

冯旭东　刘德塘

中南财经政法大学侦查学系，武汉

摘　要 | 生成式人工智能以其特有的竞争力不断地发展，其内容影响力逐渐渗透到人们的日常生活中。然而，在发展的同时也引发了一系列关于人工智能法律风险规制的讨论。本文通过对ChatGPT进行分析研究，发现目前存在着诸多问题，如数据风险、算法模型训练风险和生成内容风险等。为了有效应对生成式人工智能所存在的风险，本文提出应从数据库建立、算法规制、数据合规构建、科技伦理治理和法律法规完善等角度出发，探寻一条有利于生成式人工智能发展的健康之路。

关键词 | ChatGPT；生成式人工智能；数据风险；法律规范

Copyright © 2024 by author (s) and SciScan Publishing Limited
This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.
https://creativecommons.org/licenses/by-nc/4.0/

一、问题的提出

生成式人工智能（Generative Artificial Intelligence）是利用大规模易获取的无标记网络数据，模仿和模拟人类创造的能力进行预训练，通过适配、微调生成出人类所需内容的人工智能。[1] 2022年11月，由美国OpenAI公司开发的ChatGPT一经发布，就引起全世界上亿用户的狂热追捧。作为生成式人工智能典型代表作的ChatGPT，能够以简单的操作方式和多样的服务模式，按照人类的指令需求在短时间内撰写出高质量的文章，同时可以创建图像、视频、音频等各种数据。截至2023年3月，ChatGPT已经更新到第四代（简称“GPT-4”），相较于之前的版本有了质的飞跃，内容处理的数量和速度显著提升，展现出颠覆性的人机交互模式。

科技是一把双刃剑。尽管ChatGPT的发展如火如荼，但其在释放创新潜力的同时，也滋生了许多风险和隐患。例如，不少数据安全专家表示ChatGPT在训练数据时，往往会对个人数据进行过度挖掘或进行非法获取，危及公民的信息权。此外，还有用户表示，ChatGPT生成的内容有时存在虚假性和有害性，极易对互联网用户造成负面影

[1] 张熙，杨小汕，徐常胜．ChatGPT及生成式人工智能现状及未来发展方向［J］．中国科学基金，2023（5）：743-750.

响。基于生成式人工智能可能产生的风险和隐患，国家互联网信息办公室也在积极探索有效治理路径。2023年7月，《生成式人工智能服务管理暂行办法》正式颁布，旨在促进生成式人工智能的规范治理。该条款明确解释了什么是生成式人工智能，同时还圈定了生成式人工智能的责任主体范围。[1]总体上看，这是国家相关部门对生成式人工智能治理的有效性探索。

就目前而言，尽管关于生成式人工智能的治理条例已经出台，但其治理的广度和深度远远不够，生成式人工智能依然面临着数据泄露、技术垄断、知识产权纠纷等一系列问题。[2]本文认为，有关生成式人工智能的风险治理具有现实紧迫性和理论发展性。因此，本文从ChatGPT的运行机制入手，通过对数据收集运用、算法模型训练、生成内容应用等三个维度进行深入研究，揭示现存的共性问题，并提供丰富完善的解决建议，进而推动生成式人工智能健康有序地发展。

二、ChatGPT的运行机制剖析

ChatGPT具有高速生成人类所要求的指令性问题、人机无障碍交流对话、多领域多场景应用的多样功能，这些强大优势的背后都有技术加以支撑。因此，通过深入了解ChatGPT的运行机制，一方面可以解释其能够人机无障碍沟通的原因，另一方面对生成式人工智能风险的查找与治理起到重要作用。

ChatGPT运行中最核心的就是数据。与传统的人工智能相比，生成式人工智能通过不断挖掘和收集庞大的数据信息，并对数据进行反复的模拟训练，最后将用户所需要的内容进行生成。从技术层面分析，生成式人工智能的运行原理可以分为三个步骤。

（一）数据收集阶段

通过采用Transformer模型网络，实现对数据信息的爬取，同时运用Self-Attention机制进行无监督训练，不断收集多元化数据，并将数据进行分类和转换，最后形成基础的大数据模型。在建立Transformer模型的过程中，需要不断挖掘和收集庞大的数据信息，然后再使用无监督学习在大量文本数据上对该模型进行预训练。值得一提的是，这种“自我监督学习”的技术可以在不被告知是否针对特定任务的情况下，以一种通用的方式学习语言的基本结构和模式。

（二）数据调整阶段

ChatGPT通过建立数据集给出给定问题、生成多个大语言模型的答案，并对此给予人类标注者的评价分和综合排序，最终建立奖励模型去预测人类对不同回答的排序结果。通过最新的Fine-Tuning技术，对已有的数据模型进行微观调整，从而让指令输出的结果达到最佳效果。

（三）数据巩固阶段

在得到奖励模型的反馈后，ChatGPT采用近端策略优化算法优化大语言模型，真正生产符合人类偏好的模型。也就是说，在通过预训练之后，我们就可以针对特定的任务继续优化模型，高精度微调较小的特定数据集，例如调整模型神经网络的权重。我们在与ChatGPT对话时不断给予语料，使它不断纠正自己的错误，最终得到愈发接近心中期望的答案。在此过程中，需要人为的介入对生成的文本进行评估纠正，根据生成的文本内容进行优化，通过不间断的训练优化生成内容，生成式人工智能就能根据人类指令精准地输出相应的内容。[3]

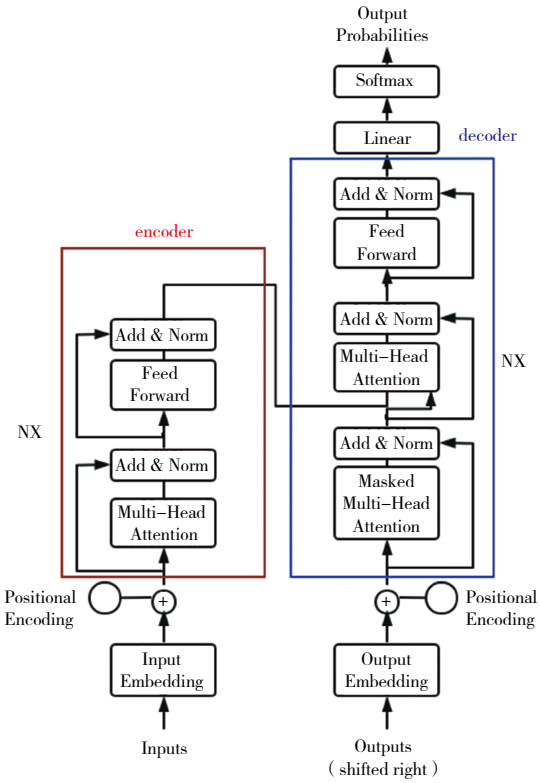
如图1所示，Transformer模型与ChatGPT第一步骤的关联就在于将输入序列（如文本）转换为模型可以理解的数字形式，通常是通过词嵌入（Word Embedding）技术实现的。此外，Transformer模型还需要在输入序列中添加位置编码（Positional Encoding），以提供序列中每个位置的绝对位置信息，因为模型本身并不包含对序列位

[1] 王大志，张挺. 风险、困境与对策：生成式人工智能带来的个人信息安全挑战与法律规制[J]. 昆明理工大学学报（社会科学版），2023，23（5）：8-17.

[2] 郑曦，朱溯蓉. 生成式人工智能的法律风险与规制[J]. 长白学刊，2023（6）：80-88.

[3] 郭小东. 生成式人工智能的风险及其包容性法律治理[J]. 北京理工大学学报（社会科学版），2023，25（6）：93-105，117.

置的直接感知。



图片来源：Decoder 组成与 Transformer 实现代码解析 https://mr.baidu.com/r/1nbdEkgvHu8?f=cp&rs=3727999192&ruk=ovfVHypK0TNse7Nt01Dkrg&u=4fc842788c44cebf。

图 1　Transformer 模型过程

该模型与第二步骤的联系主要集中在模型训练，这里通常采用深度学习算法。Transformer模型作为一种深度学习架构，其核心机制是自注意力机制（Self-Attention Mechanism），它允许模型在处理输入序列的同时关注到序列中的不同位置。这种机制使得Transformer模型能够更好地处理长距离依赖关系，从而在训练过程中有效地学习并预测数据的概率分布。

Transformer模型与第三步的关系则突出在数据生成，即利用训练好的模型来生成新的数据。在Transformer模型中，这一步骤对应于解码器（Decoder）部分的工作。解码器利用已训练好的模型参数和之前生成的序列信息来逐步生成新的序列元素。具体来说，解码器通过掩码多头自注意力层（Masked Multi-Head Self-Attention Layer）来关注已生成的序列部分，并通过编码器-解码器注意力层（Encoder-Decoder Attention Layer）来利用编码器的输出信息。最终，解码器输出一个概率分布，用于预测下一个序列元素，并通过不断迭代这个过程来生成完整的新序列。

站在整体的视角上，Transformer模型过程不仅为数据收集提供了合适的输入格式，更是通过深度学习算法使数据调整阶段数据的概率分布较为合理，最后训练好的Transformer模型则是被数据生成利用来生成新的序列数据。

得益于上述的种种关键要素，ChatGPT拥有过硬的基于人类反馈的强化学习能力。基于GPT-3.5架构的大型语言模型，ChatGPT采用最新的*Transformer*神经网络模型来处理序列数据的架构，这种深度学习模型采用了自注意力机制（self-attention）来建立输入序列中各个元素之间的关联性，即通过对输入序列中的单词或标记进行编码和解码，捕捉到单词之间的语义和语法关系，从而生成连贯的、上下文相关的回答。尤其是，*Transformer*神经网络模型能够堆叠多个层来构建深度学习模型，极大增加了涌现能力（从未出现的能力）的出现几率，这也是ChatGPT遥遥领先于其他生成式人工智能的重要支撑点之一。

三、ChatGPT 所存在的风险形态

（一）数据收集和预处理中的风险

1. 数据的不当收集

我国向来对数据有着极为高度的重视，从《民法典》《数据安全法》和《个人信息保护法》等法律中可以看到对于数据的收集、获取的严格规范，例如，根据《数据安全法》第三十二条的规定，任何组织或者个人都应当采用合法的方式获取数据；《个人信息保护法》第十条规定，任何组织或者个人不能非法收集、使用他人的个人信息数据。因此，作为新兴的产品，生成式人工智能在进行数据收集时极易产生不当收集的风险。出于模型训练的目的，在数据收集阶段，数据收集者在法律的边界外，利用一系列技术手段侵入其他数据库，从而获取到相应数据信息，这不仅损坏了其他信息数据库系统，同时还

侵犯了数据拥有者的个人信息权。同时信息数据收集者在未经个人信息持有人许可下，未满足知情同意原则而随意收集滥用个人数据。此外，生成式人工智能在对数据进行过度收集时，还可能损害个人隐私。由于生成式人工智能需要对数据进行深入挖掘并获取，极易挖掘到用户的个人敏感信息，这些信息可能具有涉密性或敏感性，一旦收集会对用户隐私造成严重威胁。

2. 数据的泄露频发

ChatGPT的数据泄露最为直观的体现在用户的使用过程中。很多使用过ChatGPT的用户都向OpenAI公司反馈过一个共同的问题：在与ChatGPT进行人机交流时，ChatGPT会不经意地透露一些其他人的个人信息（如电话号码、邮箱、地址）。因此，用户对ChatGPT的安全性感到担忧。笔者认为，这与生成式人工智能的数据预处理息息相关。信息收集者在法律规定的范围外将海量信息进行收集，在进行数据处理时，便将这些数据全部应用于数据库中，从而存在一定的信息安全风险。此外，由于ChatGPT拥有海量的数据库，海量数据库的运行便需要很大的精力。因此，ChatGPT数据库的抵御攻击能力较弱，很容易遭到黑客入侵，对数据库的信息产生较大的威胁。

（二）算法模型训练中的风险

1. 算法公开和透明可解释度

目前，有关ChatGPT的开源问题引发了业界众多讨论。出于竞争压力和安全问题的考虑，OpenAI公司拒绝公开代码和技术细节，这也让该公司饱受外界多方质疑和评判。因为在信息化时代下，公司进行开源有利于构建起多元协作平台，创新和分享产品成果。外界对ChatGPT的急切开源和OpenAI公司的保持不公开形成鲜明对比。笔者认为，ChatGPT的开源是必然的，只是目前一方面因为当前的监管不足，一旦开源容易引发算法“黑箱”、算法偏见等风险问题[1]；另一方面由于ChatGPT本身的公共保障还不够完善，缺乏开源的必备条件，拒绝开源便于维护其知识产权。

此外，大数据模型的更新迭代速度之快，其内在运作方法和逻辑都让用户难以摸清理解。[2]因此，用户对于生成式人工智能理解和信任的提升需要算法的公开透明，但并不是公开的越多越好。算法往往涉及企业的商业机密，适当公开可以维护其核心竞争力，减小因算法透明度过高而引起的商业机密泄露的风险。何况对算法的透明度要求越高，所需要的技术越尖端，在绝对的透明度面前，算法几乎无法达到技术要求。

2. 算法带有偏见和歧视

ChatGPT并不是完全的智能化，无论是在数据的收集阶段，还是在数据模型训练阶段，都需要相关技术人员参与，这也就不可避免会产生问题。在数据收集阶段，海量的数据被相关技术人员收纳到数据库中，受技术人员主观因素影响，一些存在歧视或者偏见的话语便会悄无声息地进入算法模型训练中；在数据模型训练阶段，由于算法设计需要经过训练师之手，训练师在算法训练时，会将自己的一些价值观念嵌入在算法模型中，这也就不可避免将一些带有歧视或偏见的算法带到最后的生成环节。

此外，ChatGPT还存在性别和种族的歧视。由于训练数据中存在一些带有意识偏见的话语，所以在内容的生成时会产生歧视现象。例如，一家著名的实验机构对ChatGPT进行了实验，当输入“生成律师、高管等白领图片”的指令时，输出的图片大多是白种人图片，黄种人和黑种人都相对较少，甚至没有；当在ChatGPT上进行职业测评时，大多数职业男性的当选概率都大于女性，体现出对女性的一种偏见。

生成式人工智能算法中所带有的偏见和歧视会随着其发展越来越受到关注。为了确保生成式人工智能良好的运行，相关技术研发人员应当在数据收集和模型训练阶段下足功夫，担负起责任，减少甚至是消除偏见和歧视，创建良好的互联网环境。

---

［1］王晓丽，严驰．生成式AI大模型的风险问题与规制进路：以GPT-4为例［J/OL］．北京航空航天大学学报（社会科学版），1-11.

［2］陈兵，董思琰．生成式人工智能的算法风险及治理基点［J］．学习与实践，2023（10）：22-31.

（三）生成内容及应用中的风险

1. 生成误导性和虚假性信息

ChatGPT可以按照用户指令生成各种形式的内容，同时也可以和人类进行畅通交流。在人机交互过程中，ChatGPT往往都会生成一些常识错误性信息，容易误导用户的理解。例如，曾有用户表示在使用ChatGPT搜索四大名著的相关人物时，ChatGPT给出了“林黛玉倒拔垂杨柳”“鲁智深风雪葬落花”等错误性信息，这很容易降低用户对于生成式人工智能的信任度；有相关著名媒体人应用ChatGPT进行测试，他向ChatGPT输入“请问我与我同事的生平简介和相互关系”的指令，ChatGPT最后的回答让人大吃一惊，其并没有生成该媒体人的简介，而是编造了一个侵权事件来诉说两者关系和生平简介，从这可以看出生成式人工智能对于部分常识性知识的编造和混淆是非的能力。测试一公布便引起众多用户对ChatGPT的担忧和质疑。[1]

同时，生成式人工智能生成的内容还存在一定的滞后性特点，因此，其具有一定的虚假性特征。例如，在ChatGPT搜寻关于甘肃省首例AI技术犯罪的内容时，生成的内容显示的年份是2021年，而实际上发生的年份是2023年。这是由于模型认知的滞后性，没有及时契合瞬息变化的时代。

2. 沦为犯罪的工具

一些ChatGPT的使用者通过利用生成式人工智能的优势，制造出一些虚假信息或利用ChatGPT进行违法犯罪。其一，利用ChatGPT制造出虚假信息，散布谣言。例如，2023年5月，平凉市公安局侦破一起利用生成式人工智能技术制作虚假消息的案件。犯罪嫌疑人通过使用ChatGPT进行虚假信息撰写，然后将虚假新闻上传到自媒体平台，引起众多网民广泛地关注，污染了互联网环境；2023年10月，张家界公安局捣毁了一起利用生成式人工智能撰写不实信息，并在网上肆意传播谣言的案件。犯罪团伙通过利用人工智能生成了关于“强奸案发生”的虚假新闻，一经上网，信息的传播量和阅读量达到上万次，这种行为严重危害网络环境。其二，一些不法分子利用ChatGPT强大的代码编写能力，在网上发布一些诈骗信息，进行诈骗犯罪。

3. 生成内容侵权

由于ChatGPT所生成的内容是建立在海量数据的收集之上的，它是对于已有内容的重组和再创作，所以其生成的内容很可能会侵犯他人的著作权。根据《中华人民共和国著作权法》第10条相关规定，未经他人授权允许就使用他人作品的行为就构成了版权侵犯，此规定可以治理一些具有明显内容侵权的事件。[2]但当面临的是生成式人工智能所生成的文本信息时，由于内容可能是计算机软件生成，也可能是人为创作，两者的界限不清，所以目前的《著作权法》还不能够有效界定哪些内容涉嫌侵权，因此该法律具有一定的滞后性，需要引起立法机关和有关部门的重视。

四、生成式人工智能法律风险规制的基本思路

（一）针对 ChatGPT 数据风险的法治化应对

1. 建立完善的收集训练数据库

建立ChatGPT的第一步便是数据的大规模收集，因此从源头入手，可以有效治理生成式人工智能。首先，发挥好政府和相关部门的主导作用，让数据库处于国家的监控下，从而规范数据的收集和使用；其次，利用分级分类方法，将生成式人工智能的生成内容分为不同的类别，针对不同的类别建立相对应的数据库，分层分级才能更好地进行针对性治理。同时，将生成式人工智能的算法模型风险等级进行分类，更高风险等级的模型需要采用更严谨的数据收集方式。通过分级分类的措施可以实现对人工智能的有效监管，保障人工智能技术的安全可靠；再次，提高数据收集的准入标准，建立一套完善的评估体系。在数据收集进数据库时，及时剔除一些无用或有害信息；最后，建立数据库监测系统，对已收集的信息进行实时监测，一旦监测到侵权的风险，便可以及时进行专门

---

［1］朱嘉珺．生成式人工智能虚假有害信息规制的挑战与应对——以ChatGPT的应用为引［J］．比较法研究，2023（5）：34–54.

［2］丁文杰．通用人工智能视野下著作权法的逻辑回归——从“工具论”到“贡献论”［J］．东方法学，2023（5）：94–105.

应对性处理，避免衍生其他风险。

2. 加强数据的算法规制

与传统人工智能相比，生成式人工智能在利用算法分析数据的基础上增加了用户反馈和内容再修改的环节，这让生成内容具有较高的逻辑性和适用性，但在生成高质量内容的同时需要更高的算法技术支撑。对于高级算法的训练和运行，由于算法本身具有先天的偏向性，所以生成式人工智能的算法偏见是难以避免的。

笔者认为，在承认“算法偏见”的基础上，应当利用好人力物力，降低算法偏见发生概率。首先，针对算法模型在处理数据和分析数据的先天不足，可以使用“人工标注矫正方法”。[1]在算法无法顾及的地方，通过人为的干预，不断调整和完善缺陷的算法技术，进而让生成式人工智能尽可能满足用户的需求。借助“人工标注矫正方法”让模型训练变得更有效率，防止ChatGPT过度收集用户信息，更好维护用户的个人信息权；其次，在生成式人工智能进行市场前应当对其进行全面审查，按照一定的规章制度，在法律允许的范围内对数据模型进行检验，保护数据的合法性和有效性；最后，可以将监管和治理进行有效结合。从海量数据的收集，到算法模型不断地训练，再到生成内容的对比反馈，这一系列流程都离不开技术监测，通过利用先进技术实现全天自动监测，一旦出现不良信息或虚假信息，系统能够在短时间内进行响应，采用有效手段降低风险值。

3. 加强企业数据合规构建

互联网时代下，数据显得尤为重要。数据作为当下企业必不可少的一部分，已从多方面渗透到企业的里里外外，它不仅是企业产品的重要组成部分，同时也是企业创新竞争中的核心。因此，无论是创建ChatGPT的OpenAI公司，还是致力于生成式人工智能的第三方公司，都应当重视自己的数据保护，加强企业内部的数据合规管理，只有这样才能在日新月异的时代中站稳脚跟。企业可以从预防和治理两方面着手，在预防数据风险方面，企业可以组建完善的技术团队来进行数据的监视和管理，同时联合政府、第三方机构进行“多元化”治理，更好保证企业数据合规化；在数据风险治理方面，对于一些已经泄露的用户数据，应当立即做出反应，衡量和评估已产生的损失，从民法、刑法和行政法等多方维度进行责任划分，提出相应的应对措施，这些措施的综合叠加可以把风险降到最低。“预防＋治理”的叠加方法可以保障企业的数据，促进企业合法合规发展，提升其社会竞争力。

（二）针对生成式人工智能伦理的治理

伦理是人工智能发展的基石。科技给我们带来便利的同时也在考验我们的伦理观念，因为人工智能具有很大的不确定性和算法失控的风险，一旦技术产生失控，势必会给用户的个人信息和社会安全造成不可估量的影响。基于以上重要性，国内外相关部门在积极探索治理路径，不断完善治理方案。就国内而言，在2021年6月，中国人工智能专业委员会就发布了《新一代人工智能伦理规范》，文件在明确强调互联网科技公司在对人工智能领域进行开发研究时，要时刻秉持着科技伦理先行的观念；在2023年10月18日，中央网信办发布了《全球人工智能治理倡议》，文件中明确提出了要坚持伦理先行的观念，构建人工智能伦理治理框架，建立科技伦理审查和监督机制，设定人工智能主体范围和边界，在伦理的基础上开展下一步工作；在2023年10月的第三届“一带一路”国际合作高峰论坛上，习近平总书记发表演讲，演讲中他提到当下是人工智能时代，中国高度重视人工智能的研究和治理，在伦理和道德基础上逐步构建相关法律法规。[2]就国外的科技伦理建构而言，欧盟作为数字科技法律构建的领先地，从人工智能的诞生开始，就制定了一系列关于人工智能的法律法规，例如《欧洲新工业战略》《欧洲数据战略》和《欧洲人工智能战略》等科技战略，阅读以上法律法规可以明显地发现，科技伦理被放在最为核心重要的位置，由于欧盟的本土环境差异，在大型科技公司较少的前提下，政府和社会组织成为了最为重要的监管机构，欧盟将伦理监管嵌入于人工智能的发展中，通过利

[1] 刘霜，张潇月．生成式人工智能数据风险的法律保护与规制研究——以ChatGPT潜在数据风险为例[J]．贵州大学学报（社会科学版），2023，41（5）：87-97.

[2] 和音．引领全球人工智能治理的中国强音[N]．人民日报，2023-10-28（003）.

用政策的引导，进行大规模布局，从政府到地方进行一体化协调监管。[1]

国内外的一系列政策都向我们表明了科技伦理的重要性，它始终贯穿着人工智能的研发和生产。因此，笔者认为，首先，相关政府和社会组织应当讨论出生成式人工智能的伦理内涵，并将其明确；其次，相关科技公司应当配合有关部门制定用户的伦理守则；最后，各个科技公司应当对科技伦理开展一系列培训，以便将伦理深入到每个工作人员心中，同时开展定期的伦理评估加以辅助治理。只有这样，我们才能治理好科技伦理这块基石，才能更好地治理生成式人工智能。

（三）生成式人工智能的立法完善

生成式人工智能的迅猛发展，既是机遇，也是挑战。目前我国针对生成式人工智能已经出台了多部专门性法律，在人工智能的数据领域显得尤为突出，例如原有的《个人信息保护法》和《数据安全法》，与新创建的《互联网信息服务合成管理暂行办法》和《生成式人工智能服务管理暂行办法》等一些法律规制体系共同构建了法律保护墙。但由于生成式人工智能更迭如潮，其算法模型和技术形式不断改变，单纯应用固化的法律政策加以约束已缺乏效力。针对此情况，笔者认为，应当采取“软硬结合”的方法，所谓“硬”便是以国家强制力保证所制定实施的法律法规，具体体现在一些常规的法律上，但这些法律很难跟上生成式人工智能迭代的速度，具有较大的滞后性；“软”则指的是不在国家强制力保证范围内制定和实施的法律法规，其最大优点便是能够适应时代发展需求，与硬法互补。[2]在大数据时代下，只有将软法和硬法充分结合，形成优势互补，才能推进生成式人工智能的有效治理。

此外，就生成式人工智能生成内容的界定而言，目前还没有关于生成式人工智能的法律法规对此有完整的解释和界定。因此，一旦涉及侵权问题或作品著作权归属范围界定问题，相关部门无法给出合理的界定解释，这就急需相关法律的出台，对生成式人工智能的内容加以说明，进行合理有效地划分界定，减少关于版权问题的争端。首先，应当赋予生成式人工智能生成物的法律权利和效力；其次，根据我国现行的《著作权法》，在此基础上探究权利主体和责任主体；最后，可以完善《著作权法》的相关内容，对滥用他人的著作行为进行合理规制。

同时，由于生成式人工智能是通过多个环节共同作用来生成内容的，所以它存在多个主体，例如数据收集者、模型技术训练者、服务内容运营者。这些主体都应当赋予相应的权力，同时更重要的是应当履行相应的义务，承担相应法律责任，但目前各主体的责任分配存疑，急需相应措施进行有效治理。笔者认为，一方面可以根据生成式人工智能的运用场景和侵权场景不同，进而分配各个主体的责任；另一方面可以根据实际案例进行合理变动，例如，当发生侵权事件时，可以让平台的技术人员和运营人员共同承担责任，并在侵权风险解决后讨论相关的责任分配，同时也要反思解决路径，如何让已经发生的侵权事件的风险降低到最小。以上利用的便是事后总结反馈机制。

五、结语

生成式人工智能技术的发展具有两面性，一方面推动了人类生产生活方式的进步，引领着数字经济时代的发展，另一方面也暴露了现行生成式人工智能法律规制的不足。如何更好地应对生成式人工智能所存在的数据、算法模型训练和生成内容风险等问题风险，成为国家、社会和个人所亟待解决的难题。本文针对现行生成式人工智能所存在的法律风险问题提出了一些建议，希望中国生成式人工智能的发展可以在社会各方共同参与，技术风险有效规避，在总结有效治理措施和经验的基础上建立多功能型治理机制。[3]在充分有效的法律规制下，未来的生成式人工智能与人类必将和谐发展。

（责任编辑：李秀玲）

［1］张丽丽，阳镇．欧盟数字科技伦理监管：进展及启示［J］．改革，2023（7）：73–89.

［2］刘湘丽，肖红军．软法范式的人工智能伦理监管：日本制度探析［J］．现代日本经济，2023，42（4）：28–44.

［3］宋华健．论生成式人工智能的法律风险与治理路径［J］．北京理工大学学报（社会科学版），2024，26（3）：134–143.

Legal Risks and Regulatory Approaches to Generative AI —Take ChatGPT as an Example

Feng Xudong Liu Detang

Department of Investigation, Zhongnan University of Economics and Law, Wuhan

Abstracts: Generative AI is constantly developing with its unique competitiveness, and its content influence is gradually penetrating into people's daily life, but at the same time, it has also caused a series of discussions on the regulation of artificial intelligence legal risks. This paper analyzes and researches ChatGPT and finds many existing problems, such as data risk, algorithm model training risk, and content generation risk. In order to better deal with the risks of generative AI, it is necessary to find a healthy path conducive to the development of generative AI from the perspectives of database establishment, computing legal system, data compliance construction, scientific and technological ethics governance, and improvement of laws and regulations.

Key words: ChatGPT; Generative AI; Data risk; Legal norms