法学年鉴
Annals of Law
ISSN Print: 3079-5389
ISSN Online: 3079-5397
https://doi.org/10.63313/Law.8013

生成式人工智能数据风险及其法律规制

邓峰林

北京联合大学法律（法学）硕士，北京市

发布日期：2026年1月15日

摘　要

生成式人工智能以语言大模型为核心迅猛发展，深刻改变人们生活，但数据安全风险凸显。数据输入阶段，被动与主动获取数据模式均存非法获取风险，个人数据易被过度分析利用，侵犯知识产权，甚至威胁国家数据主权。数据处理阶段，算法黑箱致输出结果难解释，算法偏见易误导用户思想行为，引发社会问题。数据输出阶段，易产生虚假信息，被用于网络犯罪，扰乱社会秩序。数据存储阶段，数据泄露风险上升，危害多元主体利益。需从多方面规制：输入阶段强化市场准入与企业合规经营；处理阶段落实算法透明原则，引入事后规制措施，建立辟谣举报机制；输出阶段强化科研伦理法治化，构建多元监管机制；存储阶段开发机构承担数据安全保障与应急义务。我国需坚持创新与底线相结合，完善立法，强化科研伦理，协同推进算法规制与隐私保护，细化企业数据安全合规管理，完善用户数据主体权利保障，推动其在法治轨道稳健发展。

关键词

生成式人工智能；数据风险；法律规制

Generative AI Data Risks and Legal Regulation

FengLin Deng

Beijing Union University, Bejing 100000, China

Abstract

Generative AI poses significant data risks across its lifecycle, from potential rights infringement during data collection and opaque algorithmic processing to harmful outputs and storage vulnerabilities. Effective regulation requires robust oversight, including compliance enforcement, algorithmic transparency, and strong data security mandates. China must balance innovation with governance by enhancing laws, ethics, and user protection to ensure responsible development within a legal framework.

Keywords

文章引用：邓峰林. (2026). 生成式人工智能数据风险及其法律规制. 法学年鉴, 卷 1 期 2, 14-21. ISSN Print: 3079-5389; ISSN Online: 3079-5397.
DOI: 10.63313/Law.8013

Generative AI; Data Risks; Legal Governance

1. 问题的提出

以 ChatGPT 为典型代表的生成式人工智能，正以一种前所未有的态势深度融入并逐步重塑人们的生活模式。生成式人工智能之所以显著区别于过往的人工智能形态，关键在于其构建起了以语言大模型为核心支撑的技术架构。在数据获取层面，它借助强大的爬虫技术，大规模地采集网络空间中已公开的海量多元数据，并且在持续的“人机互动”进程中，敏锐地捕捉并积累数据信息，以此实现对语言大模型数据库的不断丰富与优化。在人工智能领域，业界普遍达成共识，数据、算法以及算力作为驱动人工智能发展的“三驾马车”，各自发挥着不可或缺的作用，而高质量的数据更是生成式人工智能得以稳健发展的基石。

然而，伴随生成式人工智能的迅猛发展，各类数据问题在全球范围内呈爆发式涌现，涵盖数据泄露、知识产权侵权、不正当商业竞争等多个领域。这些问题不仅对个人隐私、企业权益造成了直接损害，还对正常的市场竞争秩序与社会公共利益构成了潜在威胁。以数据泄露为例，大量个人信息的非法外流，可能导致用户遭受诈骗、骚扰等侵害；在知识产权侵权方面，生成式人工智能所生成的内容极有可能侵犯他人的版权、商标权等知识产权；而在商业竞争领域，企业间围绕数据资源的争夺引发了一系列不正当竞争行为，破坏了公平竞争的市场环境。

鉴于此，数据法律保护对于生成式人工智能的健康、有序发展的重要性不言而喻。我国国家网信办等七部门联合发布的《生成式人工智能服务管理暂行办法》，这一举措无疑在提升我国针对生成式人工智能法律规制力度方面迈出了重要一步。但不容忽视的是，该规定多为倡导性、原则性的规定有关数据风险规制的具体应用场景细则还有待完善。本文旨在通过系统梳理生成式人工智能数据风险的具体类型，在全面剖析的基础上，积极探索并构建更为完善的数据风险治理路径。

2. 生成式人工智能的数据安全风险

2.1. 数据输入阶段：数据源合规风险

生成式人工智能在语言大模型数据库的构建与发展进程中，主要依托被动与主动两种模式。被动模式下，语言大模型数据库的构建与更新是大量用户通过对话框自主输入信息，系统自动将这些信息进行保存并整合纳入数据库之中。这一过程看似自然流畅，但背后涉及到用户信息的自主性与隐私保护等复杂问题。用户在输入信息时，往往难以充分意识到这些信息会被如何使用以及在多大程度上影响自身权益，而系统在收集和存储这些信息时，也缺乏足够透明的告知机制。主动模式以数据爬虫技术为典型代表。数据爬虫技术凭借其强大的程序设定，能够在互联网的海量信息中自动收集大量数据。这种数据收集方式的显著优势在于其高效性，能在短时间内获取规模庞大的数据资源，为模型的训练提供丰富素材，从而使模型在学习和应用过程中表现得更为准确、可靠。然而，无论是被动还是主动的数据获取方式，都潜藏着导致语言大模型数据库非法获取数据的风险。依据我国《个人信息保护法》的明确规定，个人信息处理者只有在取得个人同意的前提下，才具备处理个人信息的合法权限。若生成式人工智能所获取

和利用的数据集中包含公民个人信息，那么获得用户同意便成为其合法处理这些信息的必要前置条件。但在实际的应用场景中，逐一征求用户同意面临着巨大的现实阻碍，几乎难以实现。这一困境不仅源于技术层面的复杂性，还涉及用户体验、成本效益等多方面因素。大规模地向用户征求同意，需要投入大量的人力、物力和时间成本，同时可能会严重影响用户的使用体验，导致用户对相关应用产生抵触情绪。此外，个人数据还面临着过度分析利用的法律风险。由于生成式人工智能收集个人数据的边界存在模糊性，它常常借助算法技术对数据进行深度剖析，以追求答案的更高准确性。这种深度分析可能会超越合理的界限，侵犯用户的个人隐私和信息权益。例如，通过对用户的浏览历史、消费习惯等数据进行过度挖掘，可能会泄露用户的敏感信息，对用户的生活造成不必要的干扰。生成式人工智能在获取和使用这些文本语料时，如果未经权利人许可而对文本进行复制、改编或者传播，极有可能涉嫌侵犯他人知识产权，将可能引发知识产权纠纷。非法爬取的数据往往具有机密性、高密度性和保护性等特征，非法获取和利用这些数据的行为不仅对个人的信息权益构成直接侵害，还可能涉及国家的数据安全和数据主权问题。当生成式人工智能开发机构运用非法数据爬取技术成功获取我国境内数据并将其传输出境时，我国便丧失了对这部分数据出境的有效管理以及对境外数据的自主控制。这一状况严重挑战了我国的数据主权，可能导致我国在数据领域的战略利益受损，影响国家的安全和发展。

2.2. 数据处理阶段：算法黑箱与偏见

对于生成式人工智能而言，数据作为其技术应用的根基，为模型的训练和运行提供了原始素材，而算法则是运用技术手段对数据进行重构并使之产生价值的关键途径。在生成式人工智能高度智能化的大语言模型实际应用场景中，能否恰当地运用算法技术对数据进行分析处理，成为衡量其安全性的重要考量因素之一。大语言模型的算法内部机制和决策过程往往呈现出不可解释或难以理解的特性，这导致算法的输出结果存在难以捉摸的“黑洞”，即产生了备受关注的“算法黑箱”风险。这种风险的存在，使得人们在使用生成式人工智能时，无法清晰地了解模型为何会给出特定的输出，难以对结果的合理性和可靠性进行有效评估。此外，算法开发者出于商业秘密保护、竞争优势维护等多种原因，往往倾向于隐藏算法决策的规则。这使得算法对于被决策主体缺乏必要的透明性，导致用户在与生成式人工智能交互过程中，难以理解模型的决策过程以及预测结果是如何产生的。由于无法洞悉模型的决策逻辑，用户也就无法准确评估模型的可靠性和稳定性。在实际应用中，这种信息不对称可能导致用户在依赖生成式人工智能的决策时面临诸多潜在风险，例如在金融领域的风险评估、司法领域的辅助裁判等场景中，若用户无法判断模型决策的可靠性，可能会做出不合理的决策，进而造成严重的后果。

生成式人工智能的文本生成能力主要由其自然语言模型所决定，而自然语言模型的性能和特点本质上取决于算法的选择以及用于模型训练的庞大数据库。这一特性使得模型开发者具备了影响甚至操控模型输出文本价值观的能力，他们可以通过裁剪数据库或操控算法的方式，将自身的偏好植入训练数据之中。当带有开发者特定偏好的数据被用于模型训练时，生成式人工智能输出的文本便会呈现出相应的价值观倾向。以 OpenAI 开发的 ChatGPT 为例，为了确保其输出信息具有高度准确性，OpenAI 实施了“人类反馈强化学习”的训练方法。具体而言，开发主体从 GPT 中选取样本，并对这些样本进行人工标注，随后利用标注得到的评分结果来训练反馈模型。通过将反馈模型与原模型进行对抗强化训练，不断对原模型的输出结果进行优化，最终获得一个符合人类语言习惯、偏好和价值观的语言生成模型。然而，这种看似科学有效的模型生成机制背后，却隐藏着不容忽视的隐患。生成式人工智能开发机构有可能出于特定的政治目标、商业利益或其他特定利益考量，在语言模型训练过程中使用带有偏见的数据样本。当这些带有偏见的数据被用于训练模型时，生成的回复文本将完全符合开发机构的意识形态标准，从而对用户产生“潜移默化”的影响。用户在长期与这类生成式人工智能交互的过程中，其思想或行为偏好可能

会不自觉地朝着有利于开发机构的方向发展。这种数据偏见风险一旦形成，可能会带来极为严重的安全后果。它不仅可能影响个体的认知和判断，导致个体在信息接收和决策过程中产生偏差，还可能在社会层面引发舆论导向偏差、价值观冲突等问题，对社会的稳定和发展造成潜在威胁。

2.3. 数据输出阶段：数据滥用

生成式人工智能凭借其强大的数据处理与运算能力，在为用户提供信息服务时，呈现出的最终信息往往是在对海量数据进行筛选、整合后，所输出的较为单一化且标准化的内容。这种生成模式，从本质上决定了生成式人工智能难以对生成内容自身的真实性与准确性进行精准判断。其原因在于，生成式人工智能的运行机制主要依赖于既定的算法模型和大规模数据训练，在数据处理过程中，它缺乏对信息所涉事实的直接感知与理性思辨能力，无法像人类一样基于常识、经验以及深入的调查来甄别信息的真伪。这一特性使得生成式人工智能在实际应用中极易产生大量虚假信息。由于算法在筛选和整合数据时，可能受到数据偏差、算法缺陷以及训练数据质量等多种因素的影响，生成的内容可能偏离客观事实。更为严重的是，这种技术漏洞可能被恶意利用，导致生成式人工智能成为恶意内容的制造源头。虚假信息的大量传播，无疑会对人们的思维方式和行为模式产生误导与负面影响。在信息传播迅速且广泛的当下社会，公众在缺乏有效辨别手段的情况下，容易盲目接受这些虚假信息，进而在决策、认知等方面产生偏差，扰乱正常的社会认知和行为秩序。近年来，随着人工智能技术的迅猛发展，AI 诈骗事件层出不穷，引发了社会各界的广泛关注。AI技术的进步使其具备了强大的伪造能力，不仅能够逼真地伪造他人面孔，还能高度还原和合成他人声音，这为不法分子实施诈骗行为提供了更为隐蔽和高效的手段。借助深度合成技术，不法分子能够以更低的成本、更高的效率伪造图片、视频等多媒体内容，生成极具欺骗性的虚假信息。这些虚假信息被广泛应用于侮辱、诽谤、诈骗等网络犯罪活动中。在侮辱与诽谤方面，不法分子通过生成虚假的图文或视频内容，恶意诋毁他人名誉，损害他人的人格尊严，严重侵犯公民的名誉权，对受害者的精神和生活造成极大的伤害。在诈骗领域，借助生成式人工智能伪造的虚假信息，不法分子能够更加逼真地模拟真实场景，诱骗受害者上当受骗，导致其财产遭受损失。此类犯罪行为的频发，不仅严重威胁网络安全，破坏了网络空间的清朗秩序，还对现实社会的稳定与和谐构成了巨大挑战，严重扰乱了正常的社会经济秩序和社会公共安全。

2.4. 数据存储阶段：数据泄漏

生成式人工智能作为依托海量数据进行学习与模型生成的前沿技术，在运行过程中需对规模庞大的数据进行处理。而这一数据处理过程，不可避免地会扩大其潜在的攻击面。随着攻击面的拓展，数据泄露风险发生的概率显著上升，且一旦发生泄露，其影响范围也将更为广泛，波及众多层面。从实践和理论层面来看，这类数据泄露风险通常呈现出以下三种典型类型。其一，涉及用户个人数据泄露引发的隐私权侵害问题。在生成式人工智能广泛应用的场景下，大量用户的生物识别信息、健康信息等高度敏感的个人数据被纳入数据处理范畴。这些数据与用户的个人隐私紧密相连，一旦泄露，将对用户的隐私权造成严重侵害。其二，企业内部在使用生成式人工智能产品时，由于操作不当或模型本身存在固有缺陷，可能导致商业机密的泄露，进而引发不正当竞争问题。其三，因生成式人工智能导致国家秘密泄露，进而威胁国家安全的问题。国家秘密涉及国家安全和利益的各个方面，如国防、外交、经济等重要领域的关键信息。生成式人工智能在处理海量数据的过程中，若因技术漏洞或恶意攻击等原因，导致国家秘密被不当获取或泄露，将对国家的安全和稳定构成严重威胁。这种威胁不仅可能影响国家的政治、经济和军事利益，还可能损害国家在国际社会中的地位和形象。尽管生成式人工智能服务提供者通常会采取一定的措施来保障数据安全，如在用户协议中声明将采用匿名化、加密等技术手段对数据进行安全防护，

但大量公开报道表明，数据泄露风险绝非虚构。以 OpenAI 为例，2023 年 3 月 24 日，OpenAI 官方发布声明，指出有 1.2%的 ChatGPT Plus 用户数据存在泄露风险，涉及姓名、聊天记录片段、电子邮箱和付款地址等敏感信息。这一事件清晰地表明，生成式人工智能模型具备感知敏感信息的能力。即便训练数据集中未包含隐私信息，且使用者在操作过程中不存在疏忽，仍有可能由于算法技术漏洞等原因导致数据泄露。

3. 生成式人工智能数据风险规制的路径

3.1. 数据输入阶段：数据源合规风险规制

生成式人工智能数据风险是机遇与挑战并存的复杂问题。鉴于目前无法完全消除此类风险，适度接受残留风险并最大限度地进行风险控制，成为维持安全与发展平衡、实现成本最小化的合理策略。在输入阶段的风险控制上，可从两个关键思路着手。一方面，强化生成式人工智能市场准入规则。首要之举是构建生成式人工智能安全评估规制。欧盟最新通过的《人工智能法案》，将人工智能划分为最低风险、低风险、高风险、不可接受风险四个等级。依据该法案，高风险等级的生成式人工智能系统，其开发机构必须提交合规评估报告，并在投放市场前接受相关机构审查与批准。这种基于风险分级的准入机制，强调不同风险层级的差异化监管要求，突出了风险管理的重要性，为我国完善生成式人工智能应用准入规范提供了有益参考。其次，需明确生成式人工智能语料库数据获取规则。以 ChatGPT 为例，其既未公开数据获取方式，也未标明语料库数据来源，导致中文数据库来源的合法性及数据爬取的合理性难以判别。因此，披露语料库数据获取方式与来源，是保障个人隐私和国家安全的关键环节。再者，开发机构应主动运用数据清洗和去标识化等技术，对采集数据进行筛选、去噪、去重和标注等处理，保障数据的完整性与真实性，防止数据被篡改、删除或损毁，确保模型训练数据的准确可靠。此外，开发机构还应定期对产品进行合规审查，以识别并解决潜在的法律与道德风险，保障数据来源合法合规。另一方面，在数据分类分级的基础上，强化企业合规经营，确保数据来源合法。我国《数据安全法》虽已将数据分为核心数据、重要数据和普通数据，但未明确各数据类型的基本内容，也缺乏相应量化衡量标准。应针对不同数据类型设定差异化保护标准，并通过量化计算判断情节严重程度。企业作为数据业务的直接参与者，更易察觉数据安全风险。因此，研发企业有必要事先制定数据合规计划，明确合法获取数据的方式以及可能面临的刑事风险，并制定具体风险规制措施。当企业采用“爬虫”“自动识别算法”等自动化工具收集数据时，需自行评估其行为对目标网络系统的潜在影响，确保不干扰系统正常功能。若企业通过获取开源数据集间接收集数据用于模型训练，应严格审查数据提供方的授权及合法性来源，避免侵犯他人知识产权，并要求数据提供方配合定期开展合规审计。在涉及个人信息时，企业内部需设立专门审查机构，对数据采集内容进行合规审查，确保非法数据及未经用户明确同意的数据不被纳入训练数据集。

3.2. 数据处理阶段：算法风险规制

算法所具有的难以理解和非直觉性特质，如同为生成式人工智能输出文本背后的价值判断披上了一层厚重的面纱，由此衍生出意识形态安全风险。而算法透明原则，旨在揭开这层面纱，使算法的真实面目得以展现。该原则要求通过公开和披露算法的设计原理、数据输入输出等关键要素，提升算法的可解释性与可问责性，进而保障算法的公正性和可信性。在《生成式人工智能服务管理暂行办法》生效前，我国在算法解释权方面缺乏明确规定，仅在部分人工智能相关规范和标准中提及“算法应当具有可解释性”这一模糊要求。直至《生成式人工智能服务管理暂行办法》第 19 条出台，首次规定“有关主管部门

依据职责对生成式人工智能服务开展监督检查，提供者应当依法予以配合，按要求对训练数据来源、规模、类型、标注规则、算法机制机理等予以说明，并提供必要的技术、数据等支持和协助”。这一规定在应对人工智能算法安全风险上，无疑是重大的进步。然而，该规定存在明显的局限性。其一，有权要求人工智能服务提供者就算法进行解释说明的主体仅限定为监管机构，却未将其他重要的利益相关方，特别是个人信息主体纳入其中。这在实际执行过程中，极有可能致使个人信息主体的合法权益难以得到充分保障。其二，该规定对算法机制机理的描述过于宽泛，缺乏明确且精准的指向，这无疑会增加执行的难度，甚至可能导致该规定在实践中难以真正落地。欧盟和美国在算法治理领域起步较早，已将算法透明原则付诸实践，并将算法解释权法定化。其经验表明，生成式人工智能中算法透明原则的有效落实，不能单纯依赖可解释权，还需借助算法影响性评估等事后规制措施作为补充。此外，除了通过落实算法透明原则来应对恶意内容生成风险外，还需做好事后的应对与惩处工作。尤其要促使平台建立辟谣和举报机制，并对违法传播虚假有害信息的主体采取停止传输等限制措施，以此构建全方位的风险防控体系。

3.3. 数据输出阶段：数据滥用风险规制

生成式人工智能的蓬勃发展，为众多领域带来了前所未有的创新契机，在推动经济发展、社会进步等方面展现出巨大潜力。然而，如同硬币的两面，其发展过程中也滋生了一系列严峻问题，其中数据滥用风险尤为突出，如虚假信息的广泛传播、AI 诈骗的猖獗泛滥等，这些问题对社会的稳定构成了严重威胁。为有效应对可能存在的数据滥用风险，我们亟需强化科研伦理的法治化建设，推动科技向善治理，并构建多元监管机制，实现对生成式人工智能全链条的合法性监管。科技伦理作为在科学研究与技术创新等科技活动中必须遵循的行为准则与价值理念，不仅涵盖科研行为应恪守的学术规范，还包含现实社会基本道德对科研成果所划定的规范边界。在生成式人工智能产业高歌猛进的当下，我们务必时刻铭记“以人为本，伦理先行”以及“科技为人服务”的核心理念，绝不能为盲目追求利润最大化而逾越研发的道德底线与法律红线。在此背景下，建立独立的人工智能伦理审查机构，如科技伦理委员会，显得尤为必要。该机构应当组织制定生成式 AI 的伦理指南、自律公约等行业规范，以此提高行业准入门槛，在生成式 AI 进入市场前实施严格的道德审查。此举旨在从源头上把控生成式人工智能的发展方向，确保其在符合伦理道德的轨道上运行，避免因追求技术进步而忽视对社会可能造成的负面影响。

同时，构建多元监管机制以实现全链条合法性监管也势在必行。一方面，应考虑设立覆盖研发运行全程的独立监管机构，制定统一的生成式人工智能责任框架，在数据安全与科技创新之间寻求合理平衡，精准划定二者的保护边界。监管者在现有的数据安全监管体系基础上，可以积极引入影响评估、监管沙盒等先进制度。影响评估制度能够在生成式人工智能项目实施前，对其可能产生的社会、经济、伦理等多方面影响进行全面评估，为决策提供科学依据；监管沙盒制度则允许在相对可控的环境中对创新产品或服务进行测试，既鼓励创新，又能有效防范风险。此外，要求服务提供者如实依法报送算法、技术等必要信息，以便监管机构全面掌握行业动态，及时发现并解决潜在问题。另一方面，研发企业自身也肩负着不可推卸的责任，需通过多种机制和技术手段强化对数据滥用风险的监管。从 OpenAI 等企业受到广泛监管和社会关注的案例中可以清晰地认识到，一家致力于可持续发展的 AI 企业，必须具备卓越的风险治理水平以及持续的合规风险治理更新和改善能力。企业应建立健全内部监管机制，加强对数据收集、存储、使用和共享等各个环节的管控，运用加密、去标识化等技术手段保护数据安全，防止数据被滥用。只有通过政府监管与企业自律的协同发力，才能形成全方位、多层次的监管体系，有效应对生成式人工智能带来的数据滥用风险，推动该领域的健康、可持续发展。

3.4. 数据存储阶段：数据泄漏风险规制

在生成式人工智能迅猛发展的当下，其数据存储端所面临的重要数据泄露风险已成为制约行业稳健前行、威胁多元主体权益的关键隐患。为有效应对这一挑战，需从前端的数据防护以及后端的应急处置两个关键维度出发，明确生成式人工智能系统开发机构所应承担的数据处理环境安全保障义务与数据安全突发事件应急义务，从而筑牢生成式人工智能的数据管理保障机制，为行业的可持续发展保驾护航。

从事前的数据保护视角来看，生成式人工智能开发机构需构建起一个全方位、多层次的数据处理环境安全防护体系，确保语料库数据时刻处于安全无忧的存储环境之中。该安全防护义务主要涵盖以下三个核心层面：其一是数据分级保护义务。开发机构需基于一套科学且严谨的数据评估标准，包括数据的来源途径、内在价值高低、敏感程度深浅以及泄露后可能产生的影响大小等关键要素，对语料库中的数据展开精准的分类分级工作。在此基础上，制定具备针对性与差异化的数据保护规范制度，针对语料库中的重要数据与核心数据，必须实施更为严格、高级别的防护策略。不仅如此，开发机构还需进一步细化内部规则，对语料库数据从分类流程、标记方法到评估标准等各个环节进行全面、细致的规制，以保障数据分级保护工作的科学性、规范性与有效性，进而切实降低数据泄露的潜在风险。其二是数据风险监测和评估义务。依据《数据安全法》第 29、30 条的明确规定，生成式人工智能开发机构在我国境内开展数据处理活动期间，必须将数据风险监测作为一项常态化、持续性的工作。一旦发现诸如数据安全漏洞、异常访问行为等风险信号，应即刻启动应急响应机制，迅速采取有效的补救措施，以防止风险的进一步扩散与恶化。与此同时，开发机构还需定期对自身的数据处理活动进行全面、深入的风险评估，评估内容应涵盖数据处理流程的安全性、系统架构的稳定性、外部攻击的可能性等多个维度。完成风险评估后，应及时向有关主管部门报送详细、准确的风险评估报告，为主管部门的监管决策提供有力的数据支持，助力主管部门全面把握行业的数据安全态势，实现宏观层面的有效监管与精准指导。其三是设置数据安全负责人和管理机构的义务。欧盟的《通用数据保护条例》第 37 条明确要求所有数据处理者设立“数据保护官员”，专门负责数据保护相关工作。我国《数据安全法》亦有类似的规定，依据该法第 27 条第 2 款，生成式人工智能开发机构在处理重要数据时，必须明确指定数据安全负责人与相应的管理机构，并切实将数据安全保护责任落实到位。关于数据安全负责人和管理机构的具体职责，《网络数据安全管理条例（征求意见稿）》第 28 条作出了较为详尽的规定，可作为开发机构的重要参考依据。这些职责涉及数据安全重大决策提议、数据安全风险监测工作的组织开展、数据安全宣传教育培训活动的策划实施等多个关键领域。通过明确专人负责与专门机构管理，确保数据安全工作能够得到高效的组织、协调与执行。

在事后应急处理方面，生成式人工智能开发机构需要建立一套健全、完善的应急预案。该预案应详细规定数据安全突发事件发生后的响应流程、处置措施以及报告机制等关键内容，确保在面对突发事件时，能够迅速做出响应，采取高效的处置措施，最大程度降低数据安全事件对用户数据安全造成的损害，并及时向上级主管部门和相关用户报告事件进展情况，维护用户的合法权益，保障社会的稳定运行。

结语

2024 年底，OpenAI 发布的 03 模型表现卓越，借助思维力链（COT）技术模拟人类慢思考，处理博士水平科学问题时超越普通博士生，标志 AI 推理能力大幅提升，推动人工智能迈向新阶段。生成式人工智能加速发展并广泛应用，引发诸多复杂法律问题，且未来法律规制难题会不断增多。数据安全风险是法律规制核心，如何在鼓励创新、推动产业发展的同时防范负面影响，成为全球亟待解决的课题，我国也在积极探寻平衡创新与安全的治理路径。我国出台的《暂行办法》，初步回应了生成式人工智能带来的挑战，为产业规范发展搭建了法律框架。但科技与产业不断变化，法律规制也要与时俱进。未来，要坚持鼓励创新与严守底线的监管原则。立法上，持续完善法律法规，适应科技和产业动态需求。同时，

强化科研伦理观念，让科研人员和企业将伦理融入技术研发应用。在技术体系构建上，倡导“以人为本”，使人工智能服务人类福祉。算法规制与隐私保护需协同，规范算法各环节，保护用户隐私。企业要细化数据安全合规管理体系，保障数据合法安全使用。还要完善用户数据权利保障机制，赋予用户更多控制权与知情权。通过这些多维度举措，为智能生态健康发展提供法律保障，推动生成式人工智能在法治轨道稳健前行，实现技术创新与社会稳定的良性互动。

参考文献

[1] 刘辉,雷崎山.生成式人工智能的数据风险及其法律规制[J].重庆邮电大学学报(社会科学版),2024,36(04):40-51.

[2] 郑晓华.算法时代网络意识形态风险防范与实践逻辑[J].重庆邮电大学学报(社会科学版),2023,35(01):163-170+189.

[3] 斜晓东.论生成式人工智能的数据安全风险及回应型治理[J].东方法学,2023,(05):106-116.

[4] 何哲,曾润喜,秦维,等.ChatGPT 等新一代人工智能技术的社会影响及其治理[J].电子政务,2023,(04):2-24.

[5] 邓建鹏,朱怿成.ChatGPT 模型的法律风险及应对之策[J].新疆师范大学学报(哲学社会科学版),2023,44(05):91-101+2.

[6] 汪庆华.人工智能的法律规制路径:一个框架性讨论[J].现代法学,2019,41(02):54-63.

[7] 樊春良.科技伦理治理的理论与实践[J].科学与社会,2021,11(04):33-50.

[8] 王玎.论数据处理者的数据安全保护义务[J].当代法学,2023,37(02):40-49.