Dispute Settlement 争议解决, 2026, 12(4), 155-162
Published Online April 2026 in Hans. https://www.hanspub.org/journal/ds
https://doi.org/10.12677/ds.2026.124112

Hans汉斯

生成式人工智能数据安全风险及刑法应对

张文潇，马春茶

上海政法学院刑事司法学院，上海

收稿日期：2026年3月2日；录用日期：2026年3月26日；发布日期：2026年4月7日

摘　要

生成式人工智能的广泛应用，在带来技术红利的同时，也使得数据安全风险从传统的信息系统安全层面，延伸至更加复杂关键的数据内容安全与数据利用安全层面，这些风险贯穿于其数据采集、算法处理、内容输出三个环节。我国现行刑法在应对数据安全风险时存在一定问题，突出表现为刑法“数据控制”治理理念与生成式人工智能“数据利用”发展需求的矛盾，以及涉生成式人工智能犯罪刑事责任归责困难的问题，这便需要刑法对数据安全治理理念从“数据控制”向“数据利用”进行转变，同时对涉生成式人工智能犯罪中涉及的多主体归责机制进一步完善，明确生成式人工智能能否成为具备刑事责任能力的独立刑事主体。

关键词

生成式人工智能，数据安全风险，刑法应对，数据利用，刑事责任

Data Security Risks and Criminal Law Response of Generative Artificial Intelligence

Wenxiao Zhang, Chuncha Ma

School of Criminal Justice, Shanghai University of Political Science and Law, Shanghai

Received: March 2, 2026; accepted: March 26, 2026; published: April 7, 2026

Abstract

The widespread application of generative artificial intelligence has brought technological dividends while extending data security risks from traditional information system protection to more complex and critical aspects of data content security and data utilization security. These risks permeate three key stages: data collection, algorithmic processing, and content output. China's current

文章引用: 张文潇, 马春茶. 生成式人工智能数据安全风险及刑法应对[J]. 争议解决, 2026, 12(4): 155-162.
DOI: 10.12677/ds.2026.124112

criminal law exhibits certain limitations in addressing data security risks, notably manifesting as conflicts between the "data control" governance philosophy of criminal law and the developmental needs of generative AI's "data utilization", as well as challenges in determining criminal liability for AI-related offenses. This necessitates a shift in criminal law's data security governance paradigm from "data control" to "data utilization", alongside further refinement of multi-stakeholder accountability mechanisms for AI-related crimes. It is imperative to clarify whether generative AI can be recognized as an independent criminal entity with full criminal liability capacity.

Keywords

Generative AI, Data Security Risks, Criminal Law Response, Data Utilization, Criminal Liability

1. 引言

生成式人工智能，凭借其大语言模型与深度学习技术，可依据使用者[1]指令自动生成文本、图像、代码等多种原创内容，已在营销、客服、互联网运维和数据分析等多个领域中创造出巨大价值[1]。其核心生产力在于对大规模数据的高效“利用”，即通过数据训练实现内容生成，通过数据反馈完成模型升级。在这一技术原理下，数据的价值不再仅仅体现为静态的“占有”或“控制”，更是体现为动态的“流通”与“利用”。数据作为核心生产要素的地位空前凸显，数据安全日益成为关乎国家安全、社会公共利益与公民个人权益的重要因素。随着生成式人工智能的广泛应用，数据安全风险从传统的信息系统安全层面，延伸至更加复杂关键的数据内容安全与数据利用安全层面。

2. 生成式人工智能及其数据安全风险

从运行原理来看，生成式人工智能的运行过程可以大致分为三个阶段，首先是前置性学习训练及算法辅助升级的数据输入阶段，其次是算法处理数据及形成结果的算法训练阶段，最后是内容生成投入应用并产生影响的内容输出阶段[2]。而这几个阶段都可能产生与数据有关的安全风险，具体可以归纳成数据输入阶段的非法获取数据风险、算法训练阶段的数据偏见风险与内容输出阶段的数据泄露风险等。

(一) 非法获取数据风险

根据数据库采集数据的过程和结果，在数据输入阶段，生成式人工智能可能产生的非法获取数据风险主要有以下几种：

第一，恶意手段收集的风险。生成式人工智能可能会通过隐私拦截、窃取攻击、灰色交易等非法手段恶意获取更多信息数据。这些行为严重威胁数据安全与隐私保护：隐私拦截会侵入用户设备，窃取敏感信息；窃取攻击会利用系统漏洞或社交工程手段，非法获取数据；灰色交易会涉及数据的非法买卖，使数据被用于诈骗、勒索等违法犯罪活动，扰乱社会秩序，破坏网络环境的信任基础。

第二，过度收集使用者隐私的风险。生成式人工智能在为使用者提供服务时，需要通过与使用者进行深度的信息互动来满足其需求。生成式人工智能倾向于收集使用者过往的数据和个人习惯，例如浏览记录、社交信息、地理位置以及个人偏好等网络信息，通过对这些数据的分析，来预测使用者的个人特

[1]《生成式人工智能服务管理暂行办法》第二十二条将“生成式人工智能服务使用者”定义为使用生成式人工智能服务生成内容的组织、个人。

性，从而生成更符合使用者期望的结果。然而，这种对使用者个人信息的过度开发和收集，可能侵犯使用者的隐私权等合法个人权益。

第三，未经或超出授权收集的风险。依据《生成式人工智能服务管理暂行办法》第 7 条 [2]的相关规定，生成式人工智能进行涉及个人信息处理的相关操作时，必须事先获得个人的明确同意。如果没有取得用户授权或者超越其授权范围，擅自将敏感个人信息等特定数据纳入数据库中，将会直接侵害用户个人的知情权、同意权等合法权益[3]。

(二) 数据偏见风险

生成式人工智能依靠自然语言和算法模型生成结果，这会产生潜在的偏见，包括数据偏见和算法偏见。生成式人工智能在运行过程中，受其研发者的政治立场、个人喜恶，以及训练数据的来源、数量和质量等因素的影响，生成的结果存在一定的偏向性；同时，研发者也有可能将生成式人工智能的某道算法设计为会产出偏见结果的程序。如果生成式人工智能前期训练的数据存在偏见，算法运作也存在偏见，其生成的结果必然带有偏见，甚至会在使用过程中进一步延续和放大偏见，使得最终生成的结果无法达到客观与全面的要求[4]。

(三) 数据泄露风险

生成式人工智能的数据库数据呈指数级增长，并采用应用软件、网页、插件等多模态提供服务[5]。数据库中的大量数据不可避免地面临着直接或间接数据泄露的风险。直接泄露通常由网络攻击或系统漏洞引起。例如，生成式人工智能系统若遭受网络攻击，或者本身存在安全漏洞，可能被黑客或恶意攻击者利用，导致系统内的数据被窃取并泄露。而间接性的数据泄露则与人工智能的迭代训练过程有关。生成式人工智能在进行迭代训练时，需要持续使用大量数据来学习新知识、提升理解能力，并以此实现自身的升级和优化[4]。在使用生成式人工智能时，用户为获取服务而提供的个人信息、商业秘密等敏感内容，若因数据量庞大而未得到充分有效的脱敏处理，便可能在后续的人机交互中被意外重现，从而造成信息泄露[6]。

3. 我国刑法对数据安全的保护现状

在人工智能蓬勃发展的新时代，数据安全犯罪是指以经过数字化处理的一切数据为犯罪对象或工具，利用人工智能技术实施的违法犯罪行为，其危害后果不仅会破坏计算机信息系统的功能，还会侵犯个人的合法权益，甚至可能对经济秩序、国家安全造成严重威胁。

为保护数据安全，规范数据使用行为，我国相继出台了《中华人民共和国网络安全法》《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法》等多部专门性法律规范。但我国刑法中尚不存在以数据法益为核心的数据犯罪体系，相关罪名分散于分则不同章节，通过不同行为类型予以呈现。

我国刑法主要通过区分数据在犯罪行为中所起的作用，采用两种模式来保护数据安全：一种是在犯罪构成要件中直接将数据规定为犯罪对象，如非法获取计算机信息系统数据罪、破坏计算机信息系统罪等；另一种是将数据作为信息传递媒介实施其他犯罪的条件或工具，如侵犯公民个人信息罪、侵犯商业秘密罪、泄露国家秘密罪等。

从犯罪行为的类型来看，我国刑法将侵害数据安全的犯罪行为主要划分为以下五种情形：一是通过非法手段如窃取、购买或者其他方式获取数据的行为，例如非法获取计算机信息系统数据罪；二是篡改数据的行为，包括修改、删除、增加数据等，例如破坏计算机信息系统罪；三是编造并传播虚假数据的行为，例如编造并传播证券、期货交易虚假信息罪；四是非法提供或者泄露数据的行为，例如泄露内幕

[2]生成式人工智能服务提供者应当依法开展预训练、优化训练等训练数据处理活动，遵守以下规定：涉及个人信息的，应当取得个人同意或者符合法律、行政法规规定的其他情形。

信息罪；五是利用数据进行犯罪的行为，例如利用未公开信息交易罪[7]。

4. 现行刑法规制数据安全犯罪的困境

在生成式人工智能快速发展的背景下，刑法对于数据安全的“数据控制”治理理念与生成式人工智能技术的发展需求之间存在一定的冲突。同时，关于生成式人工智能能否成为犯罪主体的问题仍存在较大争议，这使得当生成式人工智能超越其既定算法和程序，实施危害数据安全的违法犯罪行为时，研发者、服务提供者、使用者和人工智能体之间刑事责任的归属和分配难以明确界定。据此，在生成式人工智能快速发展的背景下，刑法规制数据安全犯罪的困境主要表现为以下几点：

(一) 刑法数据控制治理理念与生成式人工智能发展需求的矛盾

依据《中华人民共和国数据安全法》第 3 条第 3 款 [3]的定义，数据安全的核心在于通过采取必要措施，使数据既能得到有效保护，又能实现合法利用，并且具备维持这种安全状态的能力。从这一定义出发，数据安全可以分为两方面：其一是“数据控制安全”，其核心目标是维护数据主体对自身数据的控制权，确保数据主体能够自主决定数据的使用方式和范围；另一是“数据利用安全”，其重点在于保障数据在收集、存储、处理、传输以及使用的各个环节中的安全状态。

当前，刑法的数据安全治理理念总体上倾向于“数据控制”。这一模式的核心理念是，通过严格规制数据获取行为，强化数据主体对数据的排他性支配，从而实现对数据滥用风险的前置防范[8]。该理念认为，数据主体对其个人数据享有绝对的控制权，任何未经授权的获取或使用数据的行为都应被视为对数据主体权益的侵犯，应受到法律的严格限制和严厉制裁。其目的在于，通过严格规制数据获取和使用行为，提前防范数据滥用的风险，从而保护数据主体的隐私和安全[9]。然而，这种控制模式的深层逻辑建立在两个前提之上：一是数据主体能够有效控制所有与其相关的数据；二是通过禁止非法获取就能有效防范数据滥用。在生成式人工智能时代，这两个前提都已发生动摇[10]。一方面，海量数据的流动使得数据主体事实上难以实现对个人数据的全程控制，用户知情同意规则在实践中往往流于形式；另一方面，非法获取行为只是数据滥用的前端环节，真正造成法益侵害的往往是后续的数据分析和利用行为[11]。

生成式人工智能技术的核心在于对数据的学习和分析，因而需要海量、优质、可传播可共享的信息。这种对数据的特殊需求与数据控制理念所强调的数据主体排他性控制权形成了直接的冲突。

首先，海量数据需求与数据共享限制的矛盾是冲突的核心表现。生成式人工智能的训练过程依赖于海量的数据，但刑法数据控制理念会限制数据的传播和共享，要求数据的获取和使用必须经过数据主体的明确授权，否则将被视为非法行为。这种严格控制理念在一定程度上阻碍了数据的广泛传播和共享，从而对生成式人工智能的发展形成了制约。如果刑法仅停留于前端控制，不仅难以有效保护数据安全，反而可能因过度干预数据获取而阻碍技术的创新发展[9]。

其次，数据获取行为的合法性问题进一步加剧了这一冲突。生成式人工智能依赖于数据爬虫等自动化技术获取大量数据，为模型训练提供丰富的素材。然而，从刑法数据控制理念的角度来看，数据爬虫技术的合法性边界并不清晰。一方面，数据爬虫可能在未经授权的情况下获取数据，这无疑是对数据主体合法权益的侵犯；另一方面，如果完全禁止数据爬虫技术的使用，可能会导致生成式人工智能无法获取足够的数据来实现其功能。对于仅通过公开渠道获取、未突破技术防护的数据行为，即使未经数据主体明确同意，也不应轻易认定为刑事不法[12]。

综上所述，刑法数据控制治理理念与生成式人工智能的发展需求之间存在着难以调和的矛盾。数据控制理念强调数据主体对数据的排他性控制权，而生成式人工智能则需要获取和共享大规模的数据来实现其功能和价值。这种矛盾不仅影响了生成式人工智能的发展，也对传统的数据保护理念提出了新的挑战。

[3]数据安全，是指通过采取必要措施，确保数据处于有效保护和合法利用的状态，以及具备保障持续安全状态的能力。

(二) 涉生成式人工智能犯罪归责困难

生成式人工智能的自主性与智能性对传统犯罪构成理论的主体范围提出了根本性挑战。传统犯罪构成理论中的主体主要围绕自然人和单位展开，强调主体应当具备相应的辨认能力和控制能力，即刑事责任能力。然而，生成式人工智能的出现打破了这一传统框架。生成式人工智能具有高度的自主性和智能性，能够通过深度学习和算法优化进行复杂的逻辑推理、创造性生成内容并进行决策，这使得生成式人工智能在某些情况下能够超越其算法与程序设计，独立实施行为。如果其行为造成了危害社会的结果，是否应当将生成式人工智能视为犯罪主体？

传统犯罪构成理论中，辨认能力和控制能力是认定犯罪主体的关键要素。学界对于生成式人工智能是否具备辨认能力和控制能力存在着较大争议。支持者认为，生成式人工智能具备通过深度学习和算法优化分析和处理输入数据并生成结果的能力，这种能力在一定程度上类似于人类的辨认能力和控制能力。而反对者认为，生成式人工智能的行为基于算法和数据，未体现人类意识和意志，尚不具备“脱离编程，在自主意识和意志下独立实施行为”的能力[13]。因此，生成式人工智能不具备真正的辨认和控制能力，即使其行为造成了危害社会的结果，也不能将其视为犯罪主体。

如果生成式人工智能具备一定的辨认能力和控制能力，能够被视为犯罪主体，那么随之而来的问题就是刑事责任的划分与承担。在传统犯罪构成理论中，犯罪行为的实施者是自然人或单位，刑事责任划分标准比较明确。然而，生成式人工智能的出现使得这一问题变得更加复杂。生成式人工智能如果作为犯罪主体实施了犯罪行为，应当由谁承担刑事责任？是研发者、提供者、使用者还是生成式人工智能本身？

从研发者角度来说，研发者在涉及生成式人工智能时，通过数据训练和算法设计赋予其一定的自主智能，因而研发者应当对生成式人工智能的行为所造成的危害结果承担刑事责任[4]。然而，研发者在设计时可能会采取一定防范措施来防止生成式人工智能被用于违法犯罪活动。这种情况下，生成式人工智能实施犯罪行为，研发者是否应当承担刑事责任则需要视具体情况分析。从提供者角度来说，提供者作为服务的部署者和运营方，在技术能力和风险管控上具有优势地位，若其未履行法律法规设定的安全管理义务，明知或应知存在风险而放任不管，则同样应承担相应责任[5]。从使用者角度来说，使用者利用生成式人工智能实施犯罪行为应当承担刑事责任。但是，如果使用者的行为是受生成式人工智能自主决策产生的结果的影响，应当如何确定使用者的责任范围和程度？[6]从生成式人工智能体自身来说，如果其本身具有独立的意识和意志，是否应当被视为犯罪主体？如果其行为造成了危害结果，是否应当承担相应的刑事责任？

综上所述，生成式人工智能的出现对传统犯罪主体认定标准带来了前所未有的挑战。生成式人工智能的自主性和智能性使得其在某些情况下能够独立作出决策并实施行为，从而引发了是否可以将其视为犯罪主体的讨论，进一步涉及到研发者、提供者、使用者与生成式人工智能体之间的刑事责任划分，这都是亟待解决的问题。

[4]根据《生成式人工智能服务管理暂行办法》第七条，服务提供者(通常包含研发主体)应当依法开展预训练、优化训练等训练数据处理活动，包括使用合法数据、不侵害知识产权、涉及个人信息需取得同意等。研发者作为技术创造者，其注意义务贯穿于数据训练和算法设计的全过程。

[5]《生成式人工智能服务管理暂行办法》第三条明确了国家坚持发展和安全并重、促进创新和依法治理相结合的原则，对生成式人工智能服务实行包容审慎和分类分级监管。第九条规定，提供者应当依法承担网络信息内容生产者责任，履行网络信息安全义务；涉及个人信息的，依法承担个人信息处理者责任，履行个人信息保护义务。第十四条规定，提供者发现使用者利用生成式人工智能服务从事违法活动的，应当依法依约采取警示、限制功能、暂停或者终止提供服务等处置措施，保存有关记录，并向有关主管部门报告。此即提供者安全管理义务的规范依据。

[6]《生成式人工智能服务管理暂行办法》第十条要求提供者指导使用者科学理性认识和依法使用技术，采取有效措施防范未成年人用户过度依赖或沉迷。使用者作为生成行为的发起者，对其利用服务实施的违法犯罪活动应承担直接责任。

5. 完善刑法规制数据安全犯罪的探索

(一) 转变刑法数据安全治理理念

面对生成式人工智能带来的种种风险，刑法对数据安全的治理理念急需从“数据控制”模式向“数据利用”模式转变[9]。这不仅是技术发展的必然要求，也是实现数据价值最大化和社会进步的必要条件。这一转变，并非完全放弃对数据安全的保护，而是要将规制重心从数据的“静态控制”转向数据处理的“动态安全”，在保障数据主体合法权益的同时，为数据流通与技术发展留出必要空间。

数据利用模式的核心在于平衡多重利益。数据利用模式并非忽视对数据主体的保护，而是将保护方式从“排他性控制”转向“规范性利用”。这一模式的理论根基在于承认数据的公共产品属性[11]。数据不同于传统财产，其价值恰恰在于流动与共享。对数据的过度控制不仅会抑制数据价值的释放，也难以兼顾其他主体的正当利益诉求。数据利用模式的本质，是通过将规制重心从数据获取转向数据利用，引导数据控制者合理使用数据，从而实现数据主体、数据利用者和社会公共利益之间的动态平衡。在这一模式下，单纯的数据获取行为不再成为刑法关注的重点，真正进入刑法视野的，是利用数据过程中对他人合法权益造成实质侵害的行为。

向数据利用模式转变，必须在制度层面设定清晰的边界条件，防止因理念转型而滑向数据滥用的另一端。首先，应当建立数据获取的合法性标准。对于仅通过公开渠道获取、未突破技术防护的数据行为，即使未经数据主体明确同意，也不应轻易认定为刑事不法[12]。其次，应当明确数据利用的正当性要求。数据处理者在利用数据时，必须遵循目的限定原则和比例原则，不得超出合理范围对数据进行加工、分析或传播。再次，应当区分不同类型数据的保护强度。对于涉及国家安全、商业秘密、个人隐私等特殊数据，仍需保留必要的控制性保护；而对于一般性数据，则应更多着眼于利用行为的规范性。

为平衡刑法数据控制理念与生成式人工智能发展需求，应构建非法获取行为的豁免机制，明确获取数据的合法方式。一方面，需明确数据获取行为的合法方式，如在特定条件下允许数据爬虫技术的合理应用，以满足生成式人工智能对数据的需求。另一方面，可引入“合法利益豁免”原则。该机制允许数据使用者在为自身或第三方合法利益所必需，且经权衡证明数据利用利益高于数据主体利益时，无需征得同意即可处理个人信息。这种做法不仅能支持生成式人工智能的数据获取需求，也不违背数据主体对数据的严格控制。

综上，数据安全治理理念从“数据控制”向“数据利用”的转变，是适应数字时代发展的必然选择。通过建立非法获取行为的豁免机制，明确获取数据的合法方式，可以在保护数据主体合法权益的同时，促进数据的传播和共享，推动生成式人工智能技术的发展。这种平衡的保护模式不仅有助于实现数据价值的最大化，还能为社会的数字化转型提供坚实的法律保障。

(二) 探索生成式人工智能刑事责任的替代路径

生成式人工智能的犯罪主体地位问题是一个极具前瞻性和复杂性的议题。随着技术的飞速发展，人工智能体的智能水平和自主性不断提升，其是否能够成为犯罪主体以及如何明确其刑事责任边界，成为刑法学界急需探讨的问题。然而，在当前弱人工智能时代乃至可预见的强人工智能时代初期[7]，一步到位地承认人工智能独立的刑事主体地位，不仅在理论上存在难以逾越的障碍，也缺乏现实的必要性[14]。因此，更务实的研究方向应当是探索在现行刑法框架内，通过引入或改造既有责任制度，来有效规制涉生

[7]弱人工智能时代是指人工智能系统只能在特定领域内执行人类预设的特定任务，其“智能”表现为对人类某些脑力或体力劳动的局部替代与辅助，系统本身不具备独立的意识和意志，也无法超越其设计和编制的程序范围自主决策。当前，以 DeepSeek、ChatGPT 为代表的生成式人工智能虽已展现出强大的内容生成与推理能力，但仍属于弱人工智能的范畴。强人工智能时代则是指人工智能系统能够具备与人类相当的、通用的、自主的意识和意志，可以在未经特定编程的领域内独立进行学习、推理、决策并实施行为。届时，人工智能将不再仅仅是人类的工具，而可能成为具有独立辨认能力和控制能力的行为主体。

成式人工智能的刑事风险。

首先，应当明确否认生成式人工智能的刑事主体资格。判断人工智能体能够成为犯罪主体可以从其独立意识和意志及其辨认和控制能力两方面综合。一方面，人工智能体是否具有独立意识和意志是判断其能否成为犯罪主体的关键。这需要进一步明确标准，可以从人工智能体的行为模式和决策过程入手，分析其是否能够独立于人类指令进行自主决策；也可以考察人工智能体是否能够对自身行为的后果进行预判和评估。另一方面，人工智能体的辨认和控制能力的判断也是确定其犯罪主体地位的重要依据。辨认能力是指人工智能体对其行为性质和社会危害性的认知能力，控制能力是指人工智能体能够自主决定是否实施某种行为的能力。在判断时可以借鉴人类行为能力的规定，结合人工智能体的技术特点进行综合考量。例如，通过模拟不同的场景和情境，观察人工智能体在面对复杂问题时的决策过程和行为表现，从而评估其辨认和控制能力的强弱。此外，还可以考虑引入专家评估机制，由专业的技术团队和法律专家共同对人工智能体的能力进行评估，以确保评估结果的科学性和准确性。

尽管生成式人工智能展现出强大的内容生成能力，但这种能力本质上仍是基于对人类既有数据的统计、学习和模仿，并未产生真正意义上的“自由意志”和“规范意识”[15]。其缺乏对法律规范的内在认同和情感体验，不具备刑法所要求的道义非难可能性。因此，其始终是人类的创造物和工具，难以成为独立承担刑事责任的主体[14]。

在明确人工智能体的犯罪主体地位的基础上，进一步需要厘清研发者、提供者、使用者和人工智能体之间的刑事责任。研发者作为人工智能体的创造者，对人工智能体负有首要责任。如果研发者在研发过程中存在故意或过失，导致人工智能体具有潜在的犯罪风险，那么应当承担相应的刑事责任。提供者作为服务的部署和运营方，在技术能力和风险管控上具有优势地位，若其未履行法律法规设定的安全管理义务，明知或应知存在风险而放任不管，则同样应承担刑事责任。使用者是人工智能体的实际操作者，对人工智能体的使用行为和目的负有直接责任。如果使用者利用人工智能体实施犯罪行为，或者明知人工智能体具有犯罪风险而使用，那么应当承担主要的刑事责任。而人工智能体本身，如果具备了独立意识和意志，在实施犯罪行为时具有辨认和控制能力，那么在理论上也应当承担相应的刑事责任。然而，由于现行刑法体系中尚未将人工智能体纳入犯罪主体的范畴，因此在实践中对其刑事责任的追究还存在一定的困难和争议。

综上所述，随着生成式人工智能技术的不断发展，关于其刑事责任的问题不应陷入“是否为人”的形而上学争论。明确其独立意识和意志的认定标准以及辨认和控制能力的判断方法，对于确定其犯罪主体地位至关重要。同时，合理划分研发者、提供者、使用者和人工智能体之间的刑事责任，也是应对人工智能犯罪问题的关键所在。

6. 结论

生成式人工智能的广泛应用为社会带来了巨大的便利，但同时也产生了诸多数据安全风险，如非法获取、算法偏见和信息泄露等，这些风险对个人隐私、商业机密和国家安全构成了严重威胁。然而，现行刑法在应对这些新型数据安全挑战时存在诸多困境，主要体现在刑法数据控制理念与生成式人工智能需求之间的矛盾，以及犯罪主体认定标准与生成式人工智能地位之间的冲突。传统的刑法数据控制理念强调数据主体对数据的排他性控制权，而生成式人工智能则需要大规模的数据获取与共享来实现其功能和价值，这种矛盾在一定程度上阻碍了生成式人工智能的发展。此外，生成式人工智能的自主性和智能性使其在某些情况下能够独立实施行为，引发了是否可以将其视为犯罪主体的讨论，进一步涉及到研发者、提供者、使用者与生成式人工智能体之间的刑事责任划分问题。

为应对上述问题，本文共提出两点了完善刑法规制数据安全犯罪的建议。一方面，应转变刑法数据

安全治理理念，从传统的“数据控制”向“数据利用”转变，通过建立非法获取行为的豁免机制、明确获取数据的合法方式，力求在保护数据权益与促进数据流通、支持技术创新之间达成平衡。其次，应明确生成式人工智能的犯罪主体地位，合理划分研发者、提供者、使用者和人工智能体之间的刑事责任，判断人工智能体是否具有独立意识和意志及其辨认和控制能力，并在此基础上合理划分研发者、提供者、使用者和人工智能体之间的刑事责任，以确保责任的合理公平承担。

综上所述，面对生成式人工智能技术的持续升级发展，刑法需要主动进行调整以提供更为有效的保障。通过转变数据安全治理理念和明确生成式人工智能的犯罪主体地位，不仅能为技术发展明确更为清晰的法制边界，也能够推动刑法理论的完善与发展。未来，随着生成式人工智能技术的进一步发展，刑法在数据安全保护领域的研究需要不断深入，重点关注如何在技术发展与法律保护之间寻求平衡，为生成式人工智能产业的健康发展提供坚实的法律保障。

参考文献

[1] IBM (2023) What Is Generative AI? https://www.ibm.com/topics/generative-ai

[2] 刘艳红. 生成式人工智能的三大安全风险及法律规制——以 ChatGPT 为例[J]. 东方法学, 2023(4): 29-43.

[3] 邹开亮, 刘祖兵. 生成式人工智能个人信息安全挑战及敏捷治理[J]. 征信, 2024, 42(1): 41-50, 57.

[4] 赵梓羽. 生成式人工智能数据安全风险及其应对[J]. 情报资料工作, 2024, 45(2): 30-37.

[5] 张欣. 生成式人工智能的数据风险与治理路径[J]. 法律科学(西北政法大学学报), 2023, 41(5): 42-54.

[6] 徐伟, 何野. 生成式人工智能数据安全风险的治理体系及优化路径——基于 38 份政策文本的扎根分析[J]. 电子政务, 2024(10): 42-58.

[7] 李振林, 潘鑫媛. 生成式人工智能背景下数据安全的刑法保护困境与应对——以 ChatGPT 为视角的展开[J]. 犯罪研究, 2023(2): 25-33.

[8] 刘宪权. 非法获取公开数据行为的刑法规制[J]. 法律适用, 2025(8): 97-113.

[9] 于改之. 从控制到利用: 刑法数据治理的模式转换[J]. 中国社会科学, 2022(7): 56-74, 205.

[10] 杨志琼. 数字经济时代我国数据犯罪刑法规制的挑战与应对[J]. 中国法学, 2023(1): 124-141.

[11] 梅传强, 盛浩. 数据安全刑法保护的模式转换: 从管理安全到利用安全[J]. 重庆大学学报(社会科学版), 2025, 31(1): 272-288.

[12] 聂立泽, 王祯. 生成式人工智能对数据安全保护的挑战及刑法应对[J]. 河南社会科学, 2025, 33(1): 56-63.

[13] 刘宪权. 生成式人工智能的发展与刑事责任能力的生成[J]. 法学论坛, 2024, 39(2): 18-28.

[14] 张成东. 强人工智能体刑事主体地位之否定[J]. 时代法学, 2019, 17(5): 54-62.

[15] 刘艳红. 人工智能法学研究的反智化批判[J]. 东方法学, 2019(5): 119-126.