计算机科学
COMPUTER SCIENCE

人工智能训练中合成数据的融贯性法律治理

张涛

引用本文

张涛. 人工智能训练中合成数据的融贯性法律治理[J]. 计算机科学, 2025, 52(2): 20-32.

ZHANG Tao. Coherent Legal Governance of Synthetic Data in AI Training[J]. Computer Science, 2025, 52(2): 20-32.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.240900163

人工智能训练中合成数据的融贯性法律治理

张 涛
教育部哲学社会科学实验室——中国政法大学数据法治实验室 北京 100088
中国政法大学数据法治研究院 北京 100088
中国政法大学数字社会治理研究院 北京 100088

摘 要 人工智能需要大规模、多样化和高质量的数据来训练机器学习模型,而收集这些真实世界的数据可能成本高昂,并可能威胁个人隐私、引发偏见或歧视以及侵犯版权。在实践中,合成数据作为一种替代性解决方案,受到广泛关注,被越来越多地用于训练机器学习模型。从数据法学的角度,借助数据科学以及计算机科学领域的研究成果,对人工智能训练中合成数据的治理框架进行了探索。首先,从规范层面分析了在人工智能训练中合成数据之所以受到重视的逻辑前提,即个人信息保护法所追求的"小隐私"保护与人工智能训练的"大数据"需求之间存在明显的不兼容性,使训练数据的开发面临挑战,而现有的法律和技术解决方案均存在治理效能不彰的问题。在此基础上,探讨了人工智能训练中合成数据的应用场景与风险类型。最后,提出以"法律3.0理论"和"数据治理理论"作为指引,从3个方面构建人工智能训练中合成数据的融贯性法律治理框架:制定合成数据的处理规则,强化合成数据的过程治理,开发合成数据的评估工具。

关键词:人工智能;合成数据;法律3.0;融贯性治理;数据法学
中图分类号 P181;DF37;DF0-059

Coherent Legal Governance of Synthetic Data in AI Training

ZHANG Tao
Ministry of Education Laboratory of Philosophy and Social Sciences—Data Law Laboratory of China University of Political Science and Law, Beijing 100088, China
Institute for Data Law, China University of Political Science and Law, Beijing 100088, China
Institute of Digital Society Governance, China University of Political Science and Law, Beijing 100088, China

Abstract Artificial intelligence requires large, diverse, and high-quality data to train machine learning models, and collecting this real-world data can be very difficult and can threaten individual privacy, trigger bias or discrimination, and violate copyright. In practice, synthetic data, as an alternative solution has received widespread attention and is increasingly being used to train machine learning models. This paper explores the governance framework of synthetic data in AI training from the perspective of data jurisprudence, drawing on research from both data science and computer science. It first analyzes the logical premise of the importance of synthetic data in AI training from the normative level, i.e., there is an obvious incompatibility between the protection of "small privacy" pursued by the personal information protection law and the demand for "big data" in AI training, which makes the development of training data challenging, and the development of synthetic data for machine learning models challenging. The development of training data faces challenges, while existing legal and technological solutions suffer from ineffective governance. On this basis, the application scenarios and risk types of synthetic data in AI training are discussed. Finally, it is proposed to build a coherent legal governance framework for synthetic data in AI training from three aspects, guided by the "law 3.0 theory" and "data governance theory": formulating rules for handling synthetic data, strengthening process governance of synthetic data, and developing assessment tools for synthetic data.

Keywords Artificial intelligence, Synthetic data, Law 3.0, Coherent governance, Data law

1 引言

数据是现代人工智能技术的重要驱动力,它们塑造了人工智能的认知边界,决定了人工智能"看"世界的界限[1]。欧盟《人工智能白皮书》(White Paper on Artificial Intelligence)指出,"若没有数据,就不会有人工智能的存在。许多人工智能系统的功能以及它们可能产生的动作和决策在很大程度上都取决于用于模型训练的数据集"[2]。大规模、高质量和多种类数据集的可用性,再加上算力的大幅提升,可以改善人工智能的性能并促进技术迭代[3]。然而,真实世界数据(Real-

到稿日期:2024-09-27 返修日期:2024-12-09
通信作者:张涛(zt0856@163.com)

world Data）通常具有高度敏感性，可能包含大量个人信息，收集、存储、共享和使用这些数据集可能构成个人信息保护法意义上的“个人信息处理”，因而需要遵循个人信息保护的一系列原则与规则[4]。研究也表明，人工智能训练与个人信息保护法之间存在明显的“不兼容性”，既有原则及规则可能限制收集、共享和使用大型训练数据集的能力，从而对机器学习和数据科学方法的研发与部署造成瓶颈[5]。

当训练数据收集、存储、共享和使用可能不可行或不具有成本效益时，合成数据（Synthetic Data）便有望成为平衡个人信息权益保护与数据效用提升的一种替代性方案[6]。用于机器学习的合成数据，在《麻省理工科技评论》中被认为是2022年“十大突破性技术”之一[7]。欧盟委员会联合研究中心（European Commission’s Joint Research Centre）在其研究报告中对合成数据给予高度认可，认为合成数据将成为人工智能的关键推动因素[8]。英国信息专员办公室（ICO）在《隐私增强技术指南》（PETs Guidance）中专门对合成数据的定义、作用及风险进行了规定，并指出“在无法访问大型数据集的环境中，合成数据可作为训练人工智能模型的有用工具”[9]。已有研究表明，合成数据在一定程度上可以保留原始真实数据的模式或统计特征，同时最大限度地减少共享的个人信息[10]。此外，合成数据还可以用于扩展小型数据集，从而增加数据集的多样性并减少偏见[11]。据美国权威机构高德纳（Gartner）的估计，到2030年，合成数据“将完全盖过人工智能模型中的真实数据”，成为人工智能可持续发展的又一推动因素[12]。

在实践中，合成数据在许多领域均有应用，既包括金融、保险和医疗卫生等受规制领域，也包括在数据稀缺或数据收集成本较高且不安全的能源、交通等领域。以医疗卫生领域为例，合成数据的用途主要包括以下几个方面：1）模拟和预测研究；2）假设、方法和算法测试；3）流行病学及公共卫生研究；4）健康信息技术开发；5）教育和培训；6）数据集的公开发布；7）数据关联[13]。就医疗人工智能的训练而言，Ive等利用基于心理健康疾病患者的出院报告创建的合成数据集来训练自然语言处理模型，不仅能够有效地针对心理健康疾病进行分类和诊断，而且降低了泄露患者敏感信息的风险[14]。美国国立卫生研究院也曾与合成数据公司Syntegra合作，利用超过260万COVID患者的健康信息创建了一个全面的合成数据库，以解决在疾病理解、药物与设备开发等关键领域面临的数据访问问题[15]。

然而，尽管高质量的合成数据可以缓解人工智能训练的诸多迫切问题，但是部署不当或低质量的合成数据也会带来一些潜在的危害[11]。一方面，与现实过于相似的劣质合成数据会加剧真实世界数据的局限性。研究表明，合成数据越接近真实数据，就越有可能泄露个人信息，这可能导致原始真实世界中个人信息被重新识别，进而违反信息隐私法的规定[16]。另一方面，与现实差异过大的劣质合成数据还可能造成更大的危害，这不仅可能会导致人工智能难以执行关键功能，而且可能引发系统性风险。这种系统性风险被学者们称为“模型崩溃”（Model Collapse），即模型性能会随着每次模型迭代而下降，从而导致较新的模型趋于无用。因为在受污染的数据上进行训练，模型可能会错误地感知现实，这对那些支持众多下游应用的基础模型尤为有害[17]。

有鉴于此，有必要将合成数据与人工智能治理体系结合起来，创建一个系统性的治理框架，以便在人工智能训练中能够负责任地使用合成数据。在理论研究中，已有研究成果主要从技术视角探讨合成数据的生成方法、技术缺陷、有效性验证等问题[18-20]。虽然人工智能已经在法学界引起了极大的关注，但是用于机器学习的合成数据之相关问题却尚未得到充分研究。基于此，本文以“法律3.0理论”和“数据治理理论”为指引，借助数据科学和计算机科学领域的研究成果，尝试从数据法学的角度对合成数据用于人工智能训练的逻辑机理、潜在风险与融贯性法律治理路径进行初步探讨，以引发学界对此问题的关注和讨论。

本文第2章探讨了在海量真实世界数据上训练的人工智能系统是如何与个人信息保护法产生冲突的；第3章介绍了合成数据，它作为一种替代性解决方案，可以突破现有法律路径和技术路径的局限性；第4章探讨了人工智能训练中合成数据的应用场景与主要风险；第5章探讨了如何通过融贯性法律治理来促进合成数据在人工智能训练中的负责任使用；最后总结全文。

2　人工智能训练与个人信息保护法之间的张力

人工智能的先决条件与广泛应用以及个人信息保护法规定的规制要求，引发了一系列复杂的、多维度的张力。人工智能的发展高度依赖于大量数据的访问，但这种访问受到个人信息保护法律制度施加的重大限制。这些限制主要适用于涉及个人信息的场景，并且主要源于个人信息保护法的预防性禁止原则（即需要授权）以及与个人信息处理相关的一般规则。

2.1　训练数据作为人工智能的基础

训练数据作为人工智能的基础，承担着至关重要的角色。欧盟《人工智能法》将“训练数据”界定为：“通过拟合可学习的参数来训练人工智能系统的数据。”在构建人工智能模型的过程中，训练数据不仅仅是输入信息的集合，更是模型学习和适应的基石[1]。一般认为，人工智能的训练过程通常包括以下阶段：首先，必须收集与特定问题或领域相关的原始数据；其次，必须根据问题设置对数据进行标注；再次，模型在生成的标注数据集上进行训练；最后，利用预留的测试数据集验证和微调模型的性能[18]。

从理论的角度来看，训练数据的质量是人工智能模型性能的决定性因素。训练数据的多样性保证了算法在各种情况下都能得到有效的训练；代表性确保了训练数据能够真实反映未来数据的特征；准确性则是指数据标签或输入的正确性，这直接影响到学习的结果；而完整性涉及到数据集中是否存在缺失值或不完整的记录，这可能导致算法无法有效学习数据的真实分布。数据的这些属性共同作用于模型的泛化能力，即模型对未见过数据的预测能力。一个具有良好泛化能力的模型能够在新的、未知的数据上做出准确的预测。然而，如果训练数据存在偏差，比如样本不足以覆盖所有可能的

情况，或者数据集中的信息与实际应用场景不匹配，那么模型的泛化能力就会受到影响。在这种情况下，模型在现实世界的应用就会受到限制，从而可能导致错误的判断或决策。例如，在自动驾驶汽车的开发中，如果训练数据没有包含所有可能的道路和天气条件，那么模型可能无法正确识别某些罕见的情况，从而导致事故。此外，数据质量不仅会影响模型的性能，还关系到算法的公平性和偏见问题。如果训练数据中存在系统性偏见，比如某一群体的数据被过度代表或者被忽略，那么学习算法可能会无意中学习并放大这些偏见，导致不公平的结果。因此，确保训练数据的多样性和代表性，对于开发公平、无偏见的人工智能模型至关重要。

此外，训练数据的收集、共享和使用也涉及众多的伦理和法律问题，尤其是在个人信息和隐私保护的背景下。在这个数据驱动的时代，如何平衡技术创新与个人权利的保护，成为了一个亟待解决的问题。因此，我们不仅需要关注训练数据的技术层面，更应该深入探讨其社会、法律和伦理层面的影响。这要求我们在设计和部署人工智能系统时，采取全面的视角，确保数据处理的透明度，以及对个人隐私的尊重和保护。在比较法中，欧盟《人工智能白皮书》在设计未来人工智能的监管框架时，首先对训练数据的监管提出了要求：一是要确保人工智能系统接受足够广泛的数据训练，并且需要涵盖避免危险情况所需的所有相关场景；二是要采取合理措施，确保人工智能系统后续的使用不会产生歧视的结果；三是确保在使用人工智能产品和服务期间，个人数据和隐私都应当获得充分的保护[2]。

2.2 个人信息保护法的不兼容性

高质量的训练数据并非自行生成，需要投入大量的资源进行开发，通常涉及不同的阶段。第一阶段是收集和整理数据。这个阶段的目标是获取足够数量和多样性的数据，以便模型能够从中学习到有用的信息。这可能需要从各种来源获取数据，包括公开可用的数据源、私人数据库，甚至是通过实验或调查收集的新数据。然后，这些数据需要被整理和清洗，以确保它们的质量和一致性。第二阶段是对训练数据进行结构化和分类。在这个阶段，数据被组织成一种可以由模型理解的格式。这可能涉及到将数据转换为特定的数据类型，或者将其标记为特定的类别。这个阶段的目标是创建一个结构化的训练集，其中包含了模型学习所需的所有信息。鉴于个人信息保护法对“处理”的宽泛定义，一般可以认为，在训练数据开发的每个阶段均可能涉及对相关个人信息的处理[21]。因此，个人信息保护法在训练数据的开发过程中也具有适用性。

然而，个人信息保护法所追求的“小隐私”保护与人工智能训练数据的“大数据”需求之间却存在明显的不兼容性[22]。

一方面，人工智能的训练过程对数据的需求是巨大且多样的，这往往涉及大量的个人信息处理。然而，根据《个人信息保护法》第 13 条的规定，个人信息的处理一般受所谓的“禁止原则”和“保留授权”的约束，即除非根据《个人信息保护法》第 13 条第 1 款的规定有合法性依据例外允许处理，否则一般禁止处理。其中，“个人同意”机制是个人信息处理最为常见且最为重要的合法性要求之一，意味着任何个人信息的处理都应获得信息主体的明确同意。在实际操作中，由于人工智能训练所需的数据量庞大，涉及的信息主体众多，要获得每一名信息主体的明确同意几乎是不可行的，这使得训练数据的开发难以满足个人信息处理的合法性要求。

另一方面，人工智能训练数据的开发与个人信息处理的一般原则存在冲突。例如，数据最小化原则要求仅收集和处理实现特定目的所必需的最少量的个人信息；目的限制原则强调个人信息的处理应有明确、合法的目的，并且后续的信息处理不得违背初衷；透明度原则要求信息处理的过程对信息主体应当是公开透明的。然而，人工智能的训练往往需要大量、多维度的数据集以提高其准确性和鲁棒性，这与数据最小化原则相悖。同时，人工智能技术的特点使得其数据处理的目的可能随着算法的调整和应用的变化而变化，这与目的限制原则相冲突。此外，由于人工智能算法的复杂性，在数据处理过程中难以对普通信息主体做到完全透明，这与透明度原则不符。

个人信息保护法与人工智能训练的不兼容性，受到社会背景与技术发展的双重影响。在《个人信息保护法》制定之时，正值社会迈向数字化转型的关键时期，个人权益面临巨大的风险，因此强调个人权利保护便成为《个人信息保护法》的核心价值底色[23]，其制度的设计也主要以工业社会甚至是传统农业社会为模板[5]。然而，随着数字技术的广泛使用，人类变得越来越数据化，人们与数字技术的接触和互动产生了大量有关其身体、习惯、偏好和社会关系的数字化信息。正是由于大数据时代的来临，人工智能技术得以快速迭代，这要求对当前的法律环境进行理论和实践上的变革。遗憾的是，个人信息保护法未能正确因应人工智能发展与应用的需求，大量规定与人工智能所需的大规模数据聚合环境不兼容，其严格执行可能会大幅改变人工智能开发和部署的方式，使其转为次优和低效的方式。这种不兼容性不仅会使个人信息保护法的诸多条款失去意义，不一定会给个人提供更多的信息保护，还可能限制人工智能的技术创新[24]。

3 合成数据：人工智能训练中个人信息的替代性选择

从短期来看，彻底解决个人信息保护法与人工智能训练之间的不兼容性存在困难，因此学者们开始通过理论与实践来探寻能够最大限度实现“小隐私”保护与“大数据”需求之间平衡的替代性方案。一种基本的理论假设是：既然难以改变个人信息保护法规定的诸多监管要求，那就尽可能减少个人信息保护法的适用性。在此背景下，学者们通过理论与实践提出了两种方案：一种是纯法学的路径，即重新界定个人信息的内涵，改变“可识别性”和“个人关联”的认定标准[25]；另一种是纯技术的路径，即依靠“匿名化”技术，剥离可能泄露隐私的标识符[26]。然而，随着数字技术的不断进步，上述两种方案均存在局限性，难以适应人工智能训练提出的新挑战。有鉴于此，“合成数据”便成为一种新的、更具吸引力的替代性方案。它可以模拟真实数据中的某些关键信息，为广泛访问数据进行分析提供支撑，同时又能减轻对隐私与保密性的担忧[27]。

3.1　法律路径:重新界定个人信息的困境

个人信息的定义在个人信息保护法律体系中占据核心地位。我国《个人信息保护法》第4条第1款将“个人信息”界定为“以电子或其他方式记录的与已识别或可识别的自然人有关的各种信息”;欧盟《通用数据保护条例》(GDPR)第4条第1项将“个人数据”界定为“与已识别或可识别的自然人有关的任何信息”。由此可知,将“信息”定性为“个人信息”是确立个人信息保护法实体范围的关键。据此,一般认为,可以将“可识别性”和“个人关联”作为个人信息概念的核心要素。前者是从信息到个人,即由信息本身的特殊性直接回溯到特定个人;后者是从个人到信息,即已知既定个人,知晓“关于”该个人的进一步信息[28]。

尽管从理论上看个人信息的概念似乎很明确,但在实践中其界限却极为模糊,学说和判例大多倾向于对个人信息的概念进行扩张性解释[29]。首先,确定“可识别性”的门槛通常较低。一般认为,为确定“可识别性”,通常可以考虑所有直接或间接用于识别信息主体的合理可能之手段。其次,判断“个人关联”的标准通常较宽泛。“与……有关”这一措辞界定了信息与个人关联的广泛方式,信息不仅可以因内容与自然人相关,而且也可以因处理的目的或结果与自然人相关。对此,有观点认为,在不久的将来,所有东西将是或将包含个人信息,个人信息保护法可能成为“无所不包的法律”:技术正在迅速向信息的完美可识别性发展;数据化和数据分析的进步使所有东西都包含信息;在越来越“智能”的环境中,任何信息都可能在目的或效果上与自然人有关[30]。由此引发的问题是,在所有信息均可能是个人信息并触发个人信息保护合规的情况下,个人信息保护法所产生的高度密集且不可扩展的权利义务制度将无法得到有效维持。

为了突破“扩张主义”立场下个人信息界定面临的困境,一些学者提出了“还原主义”立场,主张限缩个人信息的界定标准。例如,Schwartz等认为,可识别信息是一个广泛的连续体,识别信息需要付出不同程度的努力,而且信息可能被识别的风险也各不相同,将“可识别信息”与“已识别信息”等同,将导致信息隐私法的硬性触发。基于此,他们提出了“PII 2.0”模式,将信息置于一个连续统一体中,一端是无识别风险,另一端是已识别个人,信息可以是关于已识别的人、可识别的人和不可识别的人,并为每个类别提供不同的监管制度,最终实现激励相容的个人信息保护法律制度[25]。又如,Corte也认为,应当限缩个人信息的范围,一方面,需要通过综合评估“可识别性”和“个人关联”来抑制单独适用的扩张效应;另一方面,需要提高确定“可识别性”的门槛,既要考虑信息处理者的主观相对能力,又要考虑信息处理者所处的环境或场景[31]。

从理论逻辑的角度看,“还原主义”立场下的个人信息界定方案似乎可以缓解个人信息保护法与人工智能训练之间的张力。通过变革个人信息的内涵标准,限缩个人信息的范围,从而提高个人信息保护法的触发门槛,最终在实现个人信息保护的同时减轻信息处理者的合规压力。然而,此种解决方案仍然存在不足之处,难以契合人工智能训练的发展与规范要求。

其一,“还原主义”立场下的个人信息界定方案并未解决个人信息界定的主观性。如前所述,“还原主义”立场下的个人信息界定方案并未否定“可识别性”和“个人关联”在个人信息界定中的地位,而是对其进行改良和调整。事实上,从“数据感知”的角度看,由于数字数据与数字设备和软件等非人类实体的联系,数字数据在话语中经常被非人性化和非物质化,但是它们通常也是个性化的,这意味着人们如何收集和理解自己的数据可以被简化为认知或行为心理学模型,通常具有较强的主观性(一个人的垃圾可能是另一个人的宝贝)[32]。主观性导致“还原主义”立场下的个人信息界定方案难以为人工智能训练提供可预期性,潜在的不确定性风险可能影响技术创新中的“试错”心态,反而会增加合规成本。

其二,“还原主义”立场下的个人信息界定方案忽视了个人信息界定的场景性。个人信息的界定并非一成不变,它是一个动态的、多维的构建,需要根据具体的处理场景、环境和特点来确定。任何试图将其简化为单一维度的“还原主义”立场都难以捕捉到个人信息的全貌,而人工智能训练数据通常又是跨场景的,这就导致个人信息保护法与人工智能训练之间的冲突变得更为复杂。例如,一个人脸识别系统可能需要在不同的光照、角度和表情下收集人脸图像。这种跨场景的数据收集为人工智能系统提供了丰富的学习材料,但同时也带来了个人信息保护法律与技术需求之间的潜在冲突。换言之,基于单一场景的个人信息保护合规要求难以适应跨场景的训练数据处理实践。

3.2　技术路径:匿名化解决方案的失效

鉴于“还原主义”立场下的个人信息界定方案存在不足,技术专家们便从技术角度去探寻替代性解决方案,其中最具代表性的便是匿名化技术的应用。国家市场监督管理总局和国家标准化管理委员会于2020年3月发布的技术标准《信息安全技术　个人信息安全规范》(GB/T 35273-2020)将“匿名化”界定为“通过对个人信息的技术处理,使得个人信息主体无法被识别或关联,且处理后的信息不能被复原的过程”。在此基础上,我国《个人信息保护法》第73条第4项则将“匿名化”界定为“个人信息经过处理无法识别特定自然人且不能复原的过程”。从技术的角度看,匿名化技术有两大支柱:掩蔽(Masking)和去标识化(De-identification)。二者处理的是数据集中不同的字段,这意味着有些字段将被掩蔽,而有些字段将被去标识化。具体又可以分为泛化、压缩、分解、置换以及干扰等技术解决方案[33]。总体而言,匿名化技术的一个基本假设是:若删除或干扰信息能够将个人信息转化为非个人信息,那么信息本身就超出了个人信息保护法的范围,从而能够减少对信息使用的限制[26]。

基于上述技术逻辑和理论假设,主要国家(地区)的个人信息保护立法均将匿名信息排除在个人信息之外。这意味着个人信息保护原则及规则将不再适用于经过匿名化处理的个人信息。匿名化处理便成为保存信息效用和缓解隐私风险的良策。为了确保信息处理者能够负责任地使用匿名化技术,一些国家(地区)的个人信息保护机构还专门制定了各种指南。例如,欧盟第29条数据保护工作小组于2014年发布的

《第 05/2014 号意见:匿名化技术》对代表性匿名化技术的原理、优缺点以及使用技术时常见的错误和失效情况进行了阐述,帮助数据处理者设计有效的匿名化处理流程。

在实践中,已有研究表明,尽管匿名化技术在不断创新与发展,但与此同时"去匿名化"(De-anonymization)技术也在不断进步,这导致许多已经匿名化的数据通过"攻击"也可以被还原[34-35]。在理论上,学者们亦对匿名化解决方案提出批判和质疑。一是匿名信息存在再识别风险。例如,Ohm 认为,技术专家们通常依靠匿名化来证明不加区分地共享数据和永久存储数据是合理的,同时向用户承诺其隐私正在受到保护。但是再识别技术的进步揭示了这些承诺通常都是虚幻的,它破坏了人们对匿名化的信任,从而扰乱了隐私政策的格局,"数据可以是有用的,也可以是完全匿名的,但是绝不可能两者兼而有之"[36]。二是匿名化可能降低数据效用。无论使用怎样的技术手段,匿名化过程都将在一定程度上减少数据集的原始信息,因此匿名化程度越高,数据集的可用性(如可靠性、准确性、多样性等)就越低,即使是适度的隐私保护,也会极大地降低数据挖掘的效用[37]。

3.3 迈向新的替代性方案:合成数据的引入

重新界定个人信息的不确定性,再加上匿名化技术遭遇的批判与质疑,促使理论界和产业界将目光转向一种新的替代性方案,即合成数据。前文提及的英国信息专员办公室(ICO)便认为,合成数据作为隐私增强技术的一种表现形式,有望成为促进数据共享和再利用的有效手段,能够解决人工智能训练中的诸多问题,进而实现负责任的创新[9]。

3.3.1 合成数据的定义

理论界和产业界尽管对合成数据抱有极大的兴趣,但是目前对合成数据的定义仍未形成共识[38],学者们从不同角度进行了界定。例如,有观点认为,"合成数据是计算机生成的数据,用于模拟和替代经验观察,与现实世界中的现象没有直接对应关系"[39];有观点则认为,"合成数据是由算法生成并用于算法的数据"[40];也有观点认为,"合成数据不是真实数据,而是从真实数据中生成的且具有与真实数据相同的统计属性的数据"[41];还有观点认为,"合成数据是使用专门建立的数学模型或算法生成的数据,目的是解决一系列数据科学任务"[38]。尽管学者们对合成数据的概念有不同的描述,但是仔细分析上述定义可以发现,合成数据一般具有如下几个特征:1)不是通过直接测量获得;2)通过算法生成;3)与数学或统计模型相关;4)模仿真实数据[42]。此外,对合成数据意义下的"数据"也需要进行广义理解,既可以是结构化数据(如常见的关系数据库中的数据),也可以是非结构化文本(如聊天记录、病历、新闻报道等),还可以是图像、视频、音频和虚拟环境等[41]。

根据是否由真实数据集生成这一标准,一般可以将合成数据分为 3 种类型。第一种是基于真实数据集生成的合成数据。数据分析人员将根据现存的真实数据集建构一个模型,以捕捉真实数据的分布和结构。第二种是非基于真实数据集生成的合成数据。在没有真实数据集的情况下,数据分析人员可以根据其自身的知识背景或既有的模型来生成合成数据。第三种是基于真实数据和非真实数据的混合体所生成的合成数据。数据分析人员可以根据不同的使用目的,通过合适的方式并采用适当的比例将真实数据与非真实数据进行混合,以生成新的合成数据[41]。

3.3.2 合成数据的内在逻辑

随着数据生成技术的快速进展,如生成式对抗网络,使得生成与真实世界数据几乎无差别的合成数据成为可能[20]。在数据科学和人工智能领域,合成数据被认为蕴含全新的技术逻辑,可以用于训练人工智能模型,在真实世界数据稀缺、昂贵或难以获得的情况下尤其有益。

1)合成数据的补充逻辑。在人工智能训练中,合成数据越来越受到研究界和产业界的推崇。从诸多文献对合成数据的"益处"或"优势"的阐述来看,其首先体现为一种补充(Supplement)逻辑:由真实数据训练的算法模型所生成的合成数据,有能力增强机器学习算法及其训练数据集,使其更加完整,更具代表性和多样性[43]。早期有关算法偏见的研究中,便初步体现了所谓的"补充逻辑"。尽管解决方法不尽相同,但基本假设是一致的:可以对算法及其训练数据集进行各种补充,从而解决机器学习的法律及伦理问题。例如,Buolamwini 等认为,黑人面孔在训练数据集中的代表性不足,导致了人脸识别系统出现歧视性结果。对此,可以通过补充数据来缓和算法的偏见风险,使其更具代表性和包容性[44]。合成数据的出现进一步强化了补充逻辑,若无法从社会中提取补充数据,则可以通过算法生成这些数据,这将使人工智能系统能够从更大规模、更多样化的数据集中学习,从而提升其在真实世界数据上的性能[42]。例如,研究表明,在医学和医疗保健领域,准确的合成数据可用于增加数据集的多样性,并提高人工智能模型的鲁棒性和适应性[45]。

2)合成数据的去风险化逻辑。研究表明,机器学习模型的法律及伦理风险主要源自其训练数据集[46],而通过合成数据进行完善控制则成为将这些风险视为可消除的一种手段,达致"没有真实数据,就没有真正的风险"之目标。换言之,对人工智能开发者而言,合成数据能够使算法去风险化(De-risking),可以从根源上为算法系统的运行创造一个无风险区[40]。一方面,合成数据的生成过程本身就是一种风险控制的体现。通过精心设计的算法,合成数据生成器可以确保生成的数据在保持原始数据集的关键统计特征的同时,不包含任何个人可识别信息。另一方面,合成数据在机器学习模型的训练过程中提供了一种风险缓解的机制。合成数据由于不再被视为可识别的个人信息,因此也就不适用个人信息保护法律规范,二次利用也就不需要额外的合法性基础。

3)合成数据的非归属感逻辑。对信息主体而言,数据到底是如何变得重要的呢?这与所谓的"归属感"(Belongingness)密切相关。研究表明,人类的基本需求是在心理和情感上有一种属于他们所珍视的人和事物的感觉;归属感是基于人类维持人际关系和积极社会纽带的动机需求,因此它对人类的整体发展和福祉具有重要意义[47]。在此种意义上,"个人信息"实际上是一种"人类-数据集合体"(Human-Data Assemblages)概念,旨在突出社会物质观点所强调的主体性和

具身性的分布式和动态性,在数据与人类之间构建一种互相学习、共同进化的归属感[32]。然而,合成数据的核心魅力就在于唤起人造性或虚构性,即经过算法模型合成的数据不属于任何特定个人,它们不仅可以在诸多涉及敏感个人信息的领域(如人脸识别、医疗卫生、金融服务等)证成人工智能训练的合理性,同时反过来又会对真实的个人及其生活产生不同的影响。因此,合成数据便成为"非归属感"(Non-belonging-ness)的一种表现形式,可以向信息主体传递一种信号:合成数据能够消除与个人身体的可识别性、可视性和可管理性相关的风险,进而增强社会信任[40]。

4　人工智能训练中合成数据的风险检视

与真实数据相比,合成数据有许多优势,如合成数据更可控,可以生成所需的数量及类型[41]。这些因素使合成数据成为许多应用场景的宝贵工具。然而,合成数据也带来了显著的挑战和潜在的风险,有必要对其应用进行审慎考虑。

4.1　合成数据的应用场景

在人工智能领域,合成数据可以通过多种方式来改进机器学习模型并保护个人隐私。首先,合成数据在缺乏真实数据的情况下,为训练机器学习模型提供了一种有效的替代方案。在许多实际场景中,获取足够的真实数据可能是一项昂贵且耗时的任务。此时,合成数据可以填补数据空白,使得模型能够更好地进行训练。例如,在医疗图像识别领域,由于隐私和安全问题,真实的病患数据可能受到限制。因此,通过合成数据生成类似的图像,可以帮助训练出更准确的模型。其次,即使存在真实数据,合成数据也具有重要作用。在遵循数据保护法规的前提下,合成数据可以通过模仿真实数据的统计属性来增强模型的泛化能力。例如,可以使用生成对抗网络生成与真实数据具有相似分布的合成数据,从而避免直接使用真实数据,减少潜在的隐私风险。这种方法在隐私敏感领域,如金融、医疗和个人身份识别中,具有广泛的应用前景。最后,在机器学习模型正式应用于真实数据之前,合成数据可以用于测试和验证。通过合成数据集,可以评估模型的性能、鲁棒性和泛化能力,而无需涉及真实数据。这有助于减少对真实数据的依赖,从而降低隐私泄露的风险[48]。

从具体场景的角度看,目前合成数据已经开始应用于如下领域。1)金融服务领域。合成数据可以帮助训练机器学习模型,以执行欺诈检测和信用评分等任务。这可以使金融机构的流程自动化,并提高效率[49]。2)医疗卫生领域。医疗保健行业面临的最大挑战之一是数据的可靠共享。患者数据对于医生快速诊断和治疗患者非常重要。因此,许多医院和卫生机构都非常重视患者数据。通过合成数据,医疗保健组织可以创建基于真实数据但不包含任何实际患者信息的模型和模拟[45]。3)自动驾驶领域。训练自动驾驶系统需要大量的数据,这些数据可用于训练机器学习模型,并可用于预测自动驾驶系统在不同情况下的行为。然而,真实世界的数据往往是稀缺的、昂贵的和难以获得的。合成数据可用于为需要测试的所有不同场景生成数据,可以对系统进行更彻底的测试,有助于确保车辆的安全[41]。

4.2　合成数据本身的缺陷

尽管合成数据具有诸多好处且应用广泛,但是其仍然面临潜在的挑战和限制。

4.2.1　数据分布偏差问题

数据分布偏差通常源于合成数据生成算法的内在限制,以及在生成过程中未能充分捕捉到原始数据的复杂性和多样性。例如,合成数据可能在某些特征的边缘分布上表现出过度简化,或者在类别之间的关联性上缺乏真实数据所具有的细微差别。这种分布偏差不仅会影响数据的统计代表性,还可能导致基于这些数据训练的机器学习模型在实际应用中产生误导性预测。例如,如果一个模型仅在合成数据上进行训练,它可能会对那些在合成数据中过度表示的特征赋予过多权重,而忽略了在真实数据中更为重要的特征。这会导致模型在面对真实世界数据时出现预测偏差。

4.2.2　数据的时间性和动态性问题

在合成数据生成中,时间和动态方面的准确捕捉是构建高效模型的关键。然而,某些生成方法可能会忽略或无法充分捕捉到这些维度的细微差别。在真实数据集中,时间序列和动态变化往往承载着至关重要的信息,它们揭示了数据随时间的演变规律和内在联系。例如,在金融市场分析、气象预测或医疗健康监测中,时间序列数据的微小波动可能预示着重大事件的发生。如果合成数据生成方法不能忠实地模拟这些错综复杂的时间关系,它们生成的数据将无法反映现实世界中数据的真实动态。这种缺陷不仅会削弱模型在训练阶段的学习效果,而且在模型被部署到现实世界应用时,也可能因为无法适应真实世界数据的动态变化而失效。

4.2.3　数据的过度平滑问题

过度平滑(Over-smoothing)的问题在深度学习模型中尤为突出,特别是在使用图神经网络和其他基于邻域聚合的模型时。这些模型通过迭代地聚合邻近节点的信息来更新节点的表示,但如果聚合过程过于强烈,就可能导致不同节点的表示趋于一致,从而丧失了区分度。例如,在图分类任务中,模型需要能够识别和利用图中的关键结构特征来进行准确分类。如果出现过度平滑现象,即使是结构上相差甚远的节点也可能拥有相似的表示,这将直接影响模型对图结构的理解和分类之准确性[50]。

4.3　合成数据应用的风险

在数据科学和人工智能领域,合成数据日益成为研究的重点。因其独特的优势,如成本效益、数据多样性、细粒度控制和可定制性等,合成数据被广泛应用于人工智能模型的训练过程中。然而,尽管合成数据的使用为人工智能模型的训练提供了便利,但其应用也伴随着不可忽视的法律及伦理风险[51]。

4.3.1　合成数据的应用存在隐私属性泄露风险

在理论上,合成数据提供了一种避免直接处理敏感个人信息的方法,这在很大程度上降低了数据处理过程中的隐私泄露风险。然而,一系列研究也表明,如果对合成数据进行逆向工程,就有可能泄露底层真实数据或特定隐私属性。例如,Jordon 等的研究表明,即使是合成数据,也存在从中推断出

个人敏感信息的可能性，尤其是当合成数据与其他辅助数据集结合使用时。这种推断的可能性挑战了合成数据作为隐私保护工具的有效性。因此，合成数据的使用者必须认识到，隐私保护并非绝对的，而是一个需要不断评估和调整的动态平衡[38]。Stadler 等进一步指出，由于合成数据保留了某些真实数据的统计特性，因此隐私和数据保真度之间的权衡变得尤为复杂。如果过度强调隐私保护，可能会导致合成数据的实用价值降低，因为数据的实用性在很大程度上取决于其保留的真实数据特性。相反，如果保留过多的真实数据特性，则可能无法提供足够的隐私保护[16]。

4.3.2 合成数据的应用可能加剧算法决策的黑箱化

算法决策的黑箱化指的是算法内部运作机制的不可见性和不可解释性，这在使用机器学习和人工智能技术时尤为突出。当决策过程缺乏透明度时，即使是算法的开发者和用户，也可能无法完全理解其决策依据。在合成数据的背景下，黑箱化问题可能因几个原因而加剧：首先，合成数据通过复杂的统计模型生成，这些模型旨在捕捉和模仿真实数据的分布。然而，这些模型本身可能就是高度复杂和不透明的，增加了理解和解释算法决策的难度[51]。其次，由于合成数据不是真实数据的直接反映，它可能包含与原始数据集不同的噪声和偏差，这可能会在算法训练过程中引入额外的不确定性。此外，合成数据在保护隐私的同时，可能会隐藏或扭曲数据中的重要关联和模式，这些关联和模式对于算法决策可能至关重要。如果这些关联被不恰当地处理或丢失，算法可能会基于不完整或误导性的信息做出决策，从而影响决策结果的质量和可靠性。

4.3.3 合成数据的应用可能引发新的信息“污染”问题

自生成式人工智能的兴起与广泛使用以来，人工智能生成内容对网络的“污染”(Contamination)便成为学术讨论中经常表达的担忧[52]。如前所述，合成数据包括文本、图像、音频、视频等多种形式，若模型设计不当或原始数据本身存在质量问题，那么生成的合成数据就可能包含误导性或虚假性的信息，进而引发新的信息“污染”问题[39]。一方面，合成数据驱动的虚假信息的传播可能会削弱人们对合法信息来源的信任。在数字化世界中，信息的真实性是建立公众信任的基石。合成数据由于其高度仿真的特性，可能被用于制造和传播虚假信息，这不仅会误导公众，还可能对社会稳定和民主过程产生负面影响。另一方面，合成数据的使用可能会加剧所谓的“模型崩溃”现象。Shumailoy 等指出，当模型使用由自身或其他模型生成的内容进行训练时，会出现不可逆的缺陷，原始内容分布的尾部信息会消失。这意味着模型开始忘记训练数据中不太可能发生的事件，导致模型趋向于产生有限的输出，而这些误差在多次迭代后可能会累积，导致最终结果与现实世界的差距越来越大[53]。

4.3.4 合成数据的应用可能导致价值对齐的异化

人工智能的发展和壮大为技术专家乃至整个人类提出了新的重要问题，其中最重要的问题便是，人工智能系统应当遵循何种价值观，或者说遵循谁的价值观[54]。在此背景下，人工智能价值对齐(AI Value Alignment)的目标便是确保人工智能系统的行为符合人类的意图和价值观[55]。然而，研究表明，由于合成数据是通过算法生成的，而非直接从现实世界中收集，因此它可能无法准确捕捉到人类价值观和偏好的细微差别和复杂性[56]。这一点对于依赖数据驱动决策的人工智能系统尤为重要，因为这些系统的输出往往直接影响到人类用户的生活和决策。合成数据的这种局限性可能导致人工智能模型学习到有偏见、无依据或错误的信息。这些问题的存在，意味着根据合成数据训练的人工智能系统可能会表现出与人类期望不一致的行为，导致意想不到的后果甚至有害行为[57]。

4.3.5 合成数据的应用可能存在法律规制盲点

如前所述，合成数据的共享与使用涉及复杂的伦理和法律问题，需要法律规范予以规制。然而，现行的数据保护法规，如欧盟的《通用数据保护条例》(GDPR)和我国的《个人信息保护法》，主要以“可识别性”作为个人信息识别与认定的核心标准，经过“匿名化”处理的个人信息通常不再受到数据保护法的保护。从技术原理的角度看，合成数据在某种程度上确实符合法律所规定的“匿名化”含义，因为它通常无法与原始信息主体直接联系起来。从这个角度看，数据保护规则恐难以适用于合成数据，这可能导致法律规制的盲点，使合成数据的应用逃避现有的法律约束[58]。

此外，合成数据的特性也可能导致现有信息隐私保护机制的不足。传统的个人同意和匿名化机制，旨在通过限制信息流动来保护个体信息隐私，却未能充分考虑间接数据关系。在合成数据流动过程中，即便表面上看似未涉及真实的个人信息，但由于合成数据与原始数据集之间的潜在联系，数据使用者仍然可能利用其他信息重建个体的敏感信息。人工智能模型创建者很容易混淆用于创建模型的数据来源，并且从根本上架空或者绕过个人同意机制，这可能会增加对人工智能模型进行有效法律规制的难度[59]。

5 人工智能训练中合成数据的融贯性治理路径

在理论上，为了解决法律的滞后性与新技术快速迭代发展之间的紧张关系，Brownsword 提出了所谓的“法律 3.0 理论”。这是一种双管齐下的思维和方法，既关注规则的修改，也关注技术方案。一方面，法律规则需要被更新或修改，从而符合法律的目的或政策，支持这些规则的机构和资源(如监督和执行机构)需要被维持和升级，从而让这些规则不仅在纸面也在实践中符合立法目的；另一方面，除了利用规则以外，还需要积极寻求技术性解决方案，利用补充或代替规则的措施，通过关注“架构”(Architecture)和“设计”，使技术本身成为解决方案的一部分[60]。

在数据治理理论与实践中，“关系理论”(Relational Theory)的重要性逐渐受到重视[61]，在合成数据的治理中尤为重要，它强调重点关注数据生成的环境、使用目的以及数据作为证据的背景[39]。传统的数据保护法主要强调通过个人主义的信息主体权利，来赋能信息主体对数据化条件的个人控制权。然而，合成数据可能通过以下两种方式从根本上挑战以“个人主义”为中心的数据治理模式：一是通过基于他人数据

的推论对个人产生影响,即所谓的“溢出效应”,有学者据此主张个人在算法时代应当享有“合理推论的权利”(Right to Reasonable Inferences)[62];二是对非个人数据的广泛使用虽然可能不会危及个体权益,但是“群体特征”的识别与分析却可能导致集体层面的数据危害[63]。因此,在合成数据治理中,应当增加以“群体”为基础的关系模式,强调更具公共性、社会性和集体性的数据治理形式[61]。

如前所述,合成数据的治理不仅是技术问题,更是一个涉及技术、伦理和法律的复杂议题。只有通过多方面的协同努力,才能在促进人工智能技术创新的同时,确保个人或群体层面的信息隐私权益得到有效保护。有鉴于此,本文认为,“法律 3.0 理论”下的“融贯性治理方案”以及数据治理中的“关系理论”可以为合成数据的有效治理提供一个相对完整和系统的框架,如图 1 所示。1)在治理主体方面,应当坚持协同治理原则,既要将与合成数据处理有关的所有利益相关者均纳入治理框架,也要根据不同的角色定位,赋予不同的治理责任,同时强调对治理主体的监督与问责;2)在治理动因方面,要充分考虑合成数据可能引发的技术、伦理及法律风险,重点强调隐私保护、版权合规、数据安全等目标的实现;3)在治理内容方面,要根据治理动因,通过影响评估机制对伦理风险、业务一致性、合规性、数据质量等进行评估;4)在治理措施方面,要合理设定相应的治理规则与标准,将技术工具和法律措施与数据生命周期、人工智能价值链结合起来。本章将对重要的治理举措展开论述,以推动实现合成数据的融贯性法律治理。

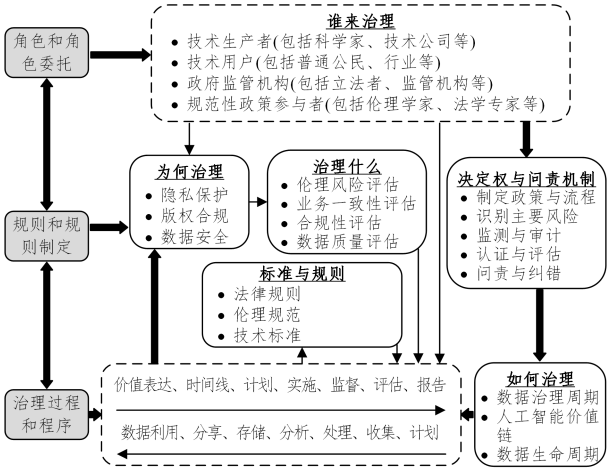

图 1　合成数据的融贯性法律治理框架

Fig. 1　Coherent legal governance framework of synthetic data

5.1　制定合成数据的处理规则

尽管人们对合成数据的兴趣日益浓厚,但尚无法律或伦理框架来规范其使用。为了确保合成数据的合规性、安全性和道德性,有必要从融贯性治理的角度出发,从以下 3 个方面完善合成数据的处理规则:1)明确合成数据的法律规则;2)制定合成数据的技术标准;3)健全合成数据的伦理规范。

5.1.1　明确合成数据的法律规则

如前所述,现行的数据保护法律,如欧盟的《通用数据保护条例》(GDPR)和我国的《个人信息保护法》,虽然为个人数据的收集、加工、存储、共享和使用设定了严格的规范,但对于合成数据的处理尚缺乏明确指导。合成数据的效用与风险要求我们必须重新审视和调整现行的法律框架,规则设计应当充分体现透明、问责和公平的原则,以最大限度地减少合成数据的生成和使用所产生的潜在社会危害[64]。

首先,应当明确合成数据的定义和分类。合成数据与匿名化数据、假名化数据的区别在于其生成过程和目的。合成数据通常是通过复杂的算法模型生成,旨在创建一个无法追溯到任何个人的数据集,同时保留原始数据集的统计特性。因此,法律应当明确合成数据的概念,区分其与其他形式的数据处理方式的差异,对不同类型的合成数据进行分类,如基于隐私保护的合成数据、用于科研的合成数据,以及用于商业分析的合成数据等,同时可以考虑引入基于风险的评估框架,并设置相应的数据处理规则,这也契合我国数据治理中的“分类分级”理念。

其次,应当强化合成数据处理者的透明度义务。透明度要求数据处理者在生成合成数据的过程中履行如下义务:一是提供明确且详细的信息,包括合成数据的来源、生成算法的构成、使用目的以及潜在风险等,以保障外部监督的有效性。二是使用合成数据进行模型训练时,应当清楚标明这些数据的合成性质,以避免用户在无意识的情况下对其准确性与可靠性产生误解[64]。2024 年 9 月,国家互联网信息办公室发布《人工智能生成合成内容标识办法(征求意见稿)》,明确以人工智能生成合成内容为治理客体,就标识义务主体、适用范围、标识类型、标识部署场景与部署方式等作出规定,也体现了透明度义务对于合成数据治理的重要性。

最后,应当建构处理合成数据的问责机制。随着合成数据的广泛应用,相关决策者和数据提供者的责任也愈加突出。在法律规则的制定中,应当明确制定合成数据生成和处理的责任主体及追责程序。这不仅包括算法设计者、合成数据生成者,还应涵盖使用这些数据的主体。建立明确的问责机制能够有效防止合成数据在生成和使用过程中可能发生的失误和滥用行为[63]。例如,如果合成数据的使用导致了对个人隐私及名誉的侵害,应当追究相关责任方的法律责任,并采取纠正措施。只有在法律框架内明确问责,才能确保各方对自身行为的负责,进而维护社会的公正与信任。

5.1.2　制定合成数据的技术标准

在现代规制体系中,技术标准不仅是一种重要的规制工具,也是预防型法治的一种重要实践形态[65]。在人工智能治理中,技术标准也成为一种非常重要的治理手段,被许多国家(地区)及国际组织寄予厚望。一方面,技术标准有助于为人工智能系统制定基于风险的关键性测试,以确定特定的人工智能系统是否会危及个人基本权利或民主价值;另一方面,技术标准可以在可靠性、鲁棒性、性能和功能安全性方面为人工智能系统制定质量标准和测试程序,从而为人工智能系统的统一合格评估和认证程序铺平道路[66]。在此背景下,制定合成数据的技术标准对于确保数据的质量和可靠性至关重要。

其一,应当明确基于数据生命周期的质量控制机制。一般而言,数据生命周期包括数据的生成、存储、处理、传输、

使用和销毁等各个阶段。具体到合成数据，数据生命周期可以分为以下5个阶段。1)数据创建。在这一阶段，通过直接输入或从外部来源获取信息，首次生成合成数据。2)数据输入和存储。这一阶段包括将合成数据输入计算机系统并存储到数据库。3)数据处理。这一阶段包括在计算机系统内对合成数据进行处理，将其转换成对用户更有用的格式。4)数据输出和传播。这一阶段是从计算机系统中生成合成数据并提供给用户的过程。5)数据处置。这一阶段包括处置不再需要的合成数据，可能涉及从数据库中删除数据或物理性地销毁存储媒体(Storage Media)[42]。技术标准制定者应当充分利用技术专家、企业界、法律界的智慧，评估每个阶段可能存在的风险，并且提供最低限度的质量控制标准和程序。

其二，应当阐明代表性合成数据方法的优势与不足。在计算机科学领域，使用基于计算机生成的合成数据来解决特定任务的概念并不新颖，其根基可以追溯到20世纪40年代。Ulam等提出的蒙特卡罗模拟方法(Monte Carlo Simulation Methods)就是早期合成数据应用的典型例子，其不仅开创了使用随机抽样方法解决复杂物理问题的先河，也为后来合成数据技术的发展奠定了理论基础[67]。进入21世纪，随着机器学习和人工智能的飞速发展，合成数据生成方法也经历了革命性的变革。生成对抗网络和变异自动编码器等深度学习结构，以其强大的数据生成能力，成为现代合成数据技术的代表[68]。然而，每种方法都有其优势和局限性。例如，生成对抗网络在生成高质量图像方面表现出色，但可能需要大量的训练数据和计算资源。因此，技术标准应当提供对这些方法的评估框架，明确每种方法的适用场景、所需资源、预期效果以及潜在的风险。此外，技术标准还应当指导用户如何根据特定任务的需求，选择最合适的合成数据生成方法。

其三，应当为具体场景中合成数据的处理提供指引。不同的应用场景(如金融服务、医疗卫生、自动驾驶等)对合成数据的要求各不相同，因此技术标准应当根据合成数据的用途、敏感性和所处的环境，提供详细的操作流程和使用约束。例如，在医疗卫生领域，合成数据的生成和使用则更加注重隐私保护和数据的代表性。医疗数据往往涉及患者的健康信息，这些信息的保护对于维护患者的隐私权至关重要。技术标准应当指导如何在不泄露任何患者身份的前提下，生成能够反映真实疾病模式和治疗效果的合成数据。同时，医疗合成数据还应保证足够的多样性，以避免在临床研究中引入偏差。在比较法中，英国金融行为监管局(FCA)于2023年3月成立了合成数据专家组(SDEG)，其汇集了来自各行业的21位专家，并于2024年3月发布了专家报告《金融服务业中合成数据的使用》，旨在帮助行业和监管从业人员全面了解与合成数据相关的技术、工具、实际挑战和机遇，以促进合成数据的有效和安全部署[49]。

5.1.3 完善合成数据的伦理规范

在人工智能领域，科技伦理治理不仅是技术发展的必要条件，更是确保技术进步服务于人类共同价值的基石[69]。合成数据作为人工智能研究中的关键资源，其伦理规范的健全性直接关系到技术应用的正当性和社会接受度。正如学者所言："要对合成数据的各种问题进行总结，必须重点讨论其伦理影响，包括有关诚实、公平和他人福祉的问题。"[39]因此，构建一套全面的合成数据伦理规范体系，对于平衡各方利益冲突及实现技术与社会价值对齐具有重要意义。

首先，伦理规范应当确立合成数据处理的伦理原则。与所谓的"硬法"(即立法机构通过的具有法律约束力的法规来界定允许或禁止的行为)不同，伦理准则不具有法律约束力，但具有说服力。Jobin等对全球84份人工智能伦理规范进行了实证分析，发现透明度、正义与公平、非恶意、负责任和保护隐私这5项伦理原则具有全球趋同性[70]。以此为参考借鉴，本文认为，应当在合成数据的伦理规范中明确下列原则。1)尊重隐私。在生成和使用合成数据的过程中，必须确保个人信息的匿名性，防止任何形式的隐私泄露。2)公开透明。合成数据的生成算法和应用场景应当向公众公开，以便于社会各界对其进行监督和评估。3)公正无偏。在合成数据的生成和应用中，应当避免引入或强化任何形式的偏见，确保数据的公平性和代表性。

其次，建立伦理审查机制是确保合成数据伦理规范得以执行的关键。在人工智能伦理治理体系中，伦理原则固然重要，但其可操作性不强，存在过度关注抽象伦理原则的确定，利益相关者之间缺乏有效合作，伦理原则与现实世界应用缺乏沟通等问题，因此需要建立相应的审查机制来确保伦理规范从"原则到实践"[71]。这一机制应包括如下要素：1)定期评估，对合成数据的生成和应用进行定期的伦理评估，以监控其是否符合既定的伦理原则；2)多元参与，伦理审查应当包括多方利益相关者，如技术开发者、数据科学家、法律专家、伦理学者以及公众代表；3)动态调整，伦理规范和审查机制应当能够根据技术发展和社会变化进行动态调整，以适应新的挑战和需求。

5.2 强化合成数据的过程治理

近年来，数据治理的重要性日益增加，受到不同组织的高度重视。对于数据处理者或控制者而言，数据治理的预期效益包括：1)通过与组织战略保持一致，优化数据的组织价值；2)优化与数据获取、使用和开发有关的风险，确保符合规制标准；3)优化所需的人力和技术资源，为涉及数据的各种业务提供更高效的支持[72]。就合成数据的治理而言，随着全球数据流动的推进，数据供应链可能跨越多个地域、多个组织、多个领域，庞大且冗长的供应链增加了数据治理的复杂性，再加上合成数据与真实数据的交织共存，目前尚未建立有效的可追溯框架或透明度机制来确保数据在整个生命周期中的完整性、安全性、隐私性和可问责性。

从已有的数据治理经验来看，若要实现有效的数据治理，数据治理机制必须解决与业务流程中数据的获取、持有、共享、使用和开发有关的纵向问题；同时解决与数据管理有关的横向问题，如质量、伦理和隐私问题、互操作性、知识管理以及部署组织结构[73]。在当前的数据治理实践中，尤其是数据共享中，区块链技术已经成为建构可追溯性机制的一种重要手段。究其原因，是由区块链网络的如下特征所决定的：1)去中心化，区块链中没有负责验证和批准账本记录的中央机构；

2)不可篡改性，区块链中存储的记录是永久性的，任何网络节点都无法更改、编辑或删除；3)透明度，区块链网络中的所有节点都保存着交易账本的完整且可审计的副本；4)可追溯性，可以跟踪所有交易，从而能够检索任何给定记录的完整历史记录[74]。例如，在医疗数据治理中，Masood 等设计了一种基于区块链的访问控制模型(Blockchain-Based Access Control Model)，该模型专门用于管理访问患者数据的授权权限，能够提高患者个人信息的精细访问控制、安全性、隐私性、可扩展性和可用性[75]。又如，在深度伪造视频治理中，Hasan 等设计了一种基于区块链的去中心化真实性验证系统(Proof of Authenticity)，其不仅能够提供公开、可信且可靠的数据来源，而且能够跟踪和追溯已发布的在线视频的历史记录，用于验证数字资产的真实性，其中包括视频、音频、图像等[76]。

结合已有的理论与实践经验，本文认为，有必要将区块链技术引入合成数据的治理实践中，建构数字监管链(Digital Chain-of-custody)[77]，实现对合成数据的过程性治理。如前所述，区块链技术，以其不可篡改性、去中心化和透明度高的特点，为合成数据的治理提供了新的可能性。通过建立数字监管链，可以在区块链上记录合成数据的每一个生成和使用步骤，从数据源头到最终应用，每一次的访问和修改都将被永久记录和验证。这种机制能够确保合成数据的来源可追溯，使用过程可审计，责任归属清晰。

在具体实施过程中，数字监管链应当与合成数据的生命周期进行有机结合[58]。首先，在数据生成阶段，数字监管链必须详细记录创建合成数据所用的方法和技术，包括但不限于差分隐私等隐私保护措施。这些信息应被精确地记录在案，并在合成数据的整个生命周期中持续跟踪，确保数据的生成过程透明可追溯。其次，在数据共享阶段，数字监管链应构建一个安全且可控的流程，以便在数据提供者、研究人员和决策者等利益相关方之间安全传输数据。该流程应采用加密技术、安全认证和访问控制机制，以防止数据遭到未经授权的访问或篡改。同时，每一次数据共享的交易都应被详细记录，包括发送方、接收方、时间戳和共享目的。再次，在数据存储阶段，数字监管链必须确保合成数据通过加密和访问控制机制得到安全存储，以防止未经授权的访问、修改或泄露。数据存储的位置和采取的安全措施应被记录在案，并接受定期审计，以确保持续符合既定的数据保护标准和法规。最后，在数据处置阶段，数字监管链应制定明确的协议来安全地删除或销毁不再需要的合成数据，并保留详细的处置记录，阐明数据销毁的时间、方法和具体原因。

5.3　开发合成数据的评估工具

在“法律 3.0 理论”看来，技术本身就应当成为解决方案的一部分，应当持续关注一系列技术工具的潜在用途。Lessig 在其著名的“代码即法律”(Code as Law)理论中，提出了一个关于互联网治理和规制的深刻见解。他认为，计算机代码(即软件和硬件的架构)在很大程度上决定了网络空间的体验方式，就像法律一样对人们的行为进行规制[78]。结合已有的理论与实践，从“技术规制技术”的角度来看，本文认为，可以采取如下措施对合成数据进行规制。

1)积极开发新技术以提升合成数据的质量。尽管现有技术，如生成对抗网络和扩散模型(Diffusion Models)，已经在合成数据的生成上取得了显著进展，但要生成既高质量又具有丰富属性的合成样本，仍然面临挑战。因此，未来的技术研究应当致力于开发新的先进技术，或在现有技术的基础上进行创新，以实现对生成数据特定属性的精确控制，从而创造出更加多样化和可定制的合成数据集。事实上，长期以来，理论界一直在尝试探讨技术规制(Techno-regulation)的可能性，将法律与其他约束性因素相结合，从而实现法律的持久规制潜力[79]。上文提及的“代码即法律”便是典型例证。还有一些学者从广泛的意义上讨论“设计”在法律中的作用，提出了诸如“通过设计进行法律保护”(Legal Protection by Design)或“通过设计进行规制”(Regulation by Design)等理论[80-81]。这些理论背后的科学假设便是，在技术发展过程中，不断发展的技术已经吸收了目标、目的和价值观，弱化了“工具”和“目的”之间的区别，因此必须在设计阶段进行法律和伦理分析，以确定其合理用途[82]。就合成数据的生成模型而言，在设计阶段就应当考虑如下要求：(1)句法准确性，即生成的数据应当是合理的；(2)隐私性，即应该能够精确量化通过发布合成样本泄露的原始数据信息；(3)统计准确性，即应该能够精确量化合成数据与原始数据之间的统计相似性；(4)效率性，即算法应当能够很好地适应数据空间(即特征空间)的维度[38]。

2)善用各类技术手段评估合成数据的效用。一般认为，合成数据若要有意义(Meaningful)，必须具备 3 个基本属性：一是实用性(Utility)，通常取决于合成数据对于特定任务是否有用；二是真实性(Fidelity)，主要是指合成数据与真实数据在“统计学”上的匹配程度；三是隐私性，通常取决于合成数据揭示的用于生成该数据的真实数据的信息量[38]。上述 3 个属性便成为评估合成数据效用的重要标准。在未来的技术研究中，不仅要进一步细化上述 3 个属性的具体测量指标，而且还要开发能够用于评估上述属性的技术工具。例如，对于合成数据的“实用性”之评估，技术研究应当进一步细化具体的测量指标。这些指标不仅包括合成数据在特定任务中的准确度和精确度，还应考虑数据在不同群体中应用时的公平性和无偏性。此外，实用性的评估还应包括合成数据在现实世界场景中的可转换性和可扩展性，即合成数据能否被有效地应用于不同的环境和规模[41]。

结束语　合成数据已成为解决人工智能发展中面临的数据稀缺、隐私保护以及成本高昂等问题的有效途径。通过生成既真实又多样化的数据集，合成数据能够支持在多个领域内对人工智能模型进行大规模训练。然而，合成数据亦并非“灵丹妙药”，其本身存在诸多不足，其广泛应用也带来了一系列法律及伦理风险，有必要将合成数据的规制放置在人工智能治理的整体框架中，以便我们能够为未来几年合成数据使用的增加做好更充分准备。囿于其社会技术特性，合成数据的规制是一个多维度、多层次的系统性工程，不仅涉及法律规则、伦理规范及技术标准的制定及修改，还涉及开发和使用作为规制手段的技术工具。本文提出在“法律 3.0 理论”和“数据治理理论”的指引下，既要通过规则之治，为负责任且高效

地利用合成数据划定规范框架，也应当通过技术之治，聚焦于提升合成数据的实用性、真实性和隐私性；既要关注个体权益的保护，也要通过设置相应的治理机制来防范群体层面的信息危害。合成数据治理不仅是一个技术问题，同时也是一个法学命题，其治理体系及机制的设计需要兼顾技术发展与法律规制的动态平衡，这也有赖于“数据法学”“人工智能法学”“科技法学”等新兴交叉学科的持续深入研究。

参 考 文 献

[1] CRAWFORD K. Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence[M]. New Haven: Yale University Press, 2021: 97-98.

[2] EUROPEAN COMMISSION. White Paper on Artificial Intelligence: a European approach to excellence and trust[EB/OL]. [2024-11-06]. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0065.

[3] KAPLAN J. Generative Artificial Intelligence: What Everyone Needs to Know[M]. New York: Oxford University Press, 2024: 31-32.

[4] KHAN M, HANNA A. The Subjects and Stages of AI Dataset Development: A Framework for Dataset Accountability[J]. Ohio State Technology Law Journal, 2023, 19(2): 171-256.

[5] DING X D. On Data Institution that Promotes Artificial Intelligence[J]. China Law Review, 2023(6): 175-191.

[6] KURAPATI S, GILLI L. Synthetic Data: A Convergence between Innovation and GDPR[J]. Journal of Open Access to Law, 2023(11): 1-12.

[7] HEAVEN W D. Synthetic data for AI[EB/OL]. [2024-11-06]. https://www.technologyreview.com/2022/02/23/1044965/ai-synthetic-data-2/.

[8] HRADEC J, CRAGLIA M, DI L M, et al. Multipurpose synthetic population for policy applications[M]. Luxembourg: Publications Office of the European Union, 2022: 15.

[9] ICO. Privacy-enhancing technologies (PETs) [EB/OL]. [2024-11-06]. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-sharing/privacy-enhancing-technologies/.

[10] BELLOVIN S M, DUTTA P K, REITINGER N. Privacy and Synthetic Datasets[J]. Stanford Technology Law Review, 2019, 22(1): 1-52.

[11] LEE P. Synthetic Data and the Future of AI [EB/OL]. [2024-11-06]. https://ssrn.com/abstract=4722162.

[12] ALEXANDER L. Is Synthetic Data the Future of AI? [EB/OL]. [2024-11-06]. https://www.gartner.com/en/newsroom/press-releases/2022-06-22-is-synthetic-data-the-future-of-ai.

[13] GONZALES A, GURUSWAMY G, SMITH S R. Synthetic data in health care: A narrative review[J]. PLOS Digit Health, 2023, 2(1): e0000082.

[14] IVE J, VIANI N, KAM J, et al. Generation and evaluation of artificial mental health records for Natural Language Processing [J]. NPJ Digital Medicine, 2020, 69(3): 1-9.

[15] CAIRO M. Synthetic Data and GDPR Compliance: How Artificial Intelligence Might Resolve the Privacy-Utility Tradeoff[J]. Journal of Technology Law & Policy, 2023(28): 71-113.

[16] STADLER T, OPRISANU B, TRONCOSO C. Synthetic Data -- Anonymisation Groundhog Day[EB/OL]. [2024-11-06]. https://doi.org/10.48550/arXiv.2011.07018.

[17] TAORI R, HASHIMOTO T B. Data feedback loops: model-driven amplification of dataset biases[C]// Proceedings of the 40th International Conference on Machine Learning (ICML'23), New York: JMLR.org, 2023: 33883-33920.

[18] NIKOLENKO S I. Synthetic Data for Deep Learning [M]. Cham: Springer, 2021: 2.

[19] NASSIF J, TEKLI J, KAMRADT M. Synthetic Data: Revolutionizing the Industrial Metaverse[M]. Cham: Springer, 2024: 10-11.

[20] FIGUERIRA A, VAZ B. Survey on Synthetic Data Generation, Evaluation Methods and GANs[J]. Mathematics, 2022, 10(15): 2733.

[21] HACKER P. A legal framework for AI training data—from first principles to the Artificial Intelligence Act[J]. Law, Innovation and Technology, 2021, 13(2): 257-301.

[22] ZARSKY T Z. Incompatible: The GDPR in the Age of Big Data [J]. Seton Hall Law Review, 2017, 47(4): 995-1020.

[23] ZHOU H H. Legal position of personal information protection [J]. Studies in Law and Business, 2020, 37(3): 44-56.

[24] ZARSKY T Z. The Privacy-Innovation Conundrum[J]. Lewis & Clark Law Review, 2015, 19(1): 115-168.

[25] SCHWARTZ P M, SOLOVE D J. The PII Problem: Privacy and a New Concept of Personally Identifiable Information[J]. New York University Law Review, 2011, 86(6): 1814-1894.

[26] ELLIOT M, HARA K, RAAB C, et al. Functional anonymisation: Personal data and the data environment[J]. Computer Law & Security Review, 2018, 34(2): 204-221.

[27] RAGHUNATHAN T E. Synthetic Data[J]. Annual Review of Statistics and Its Application, 2021(8): 129-140.

[28] ZHANG X B. Interpretation of the Personal Information Protection Law of the People's Republic of China[M]. Beijing: People's Publishing House, 2021: 41.

[29] QI A M, ZHANG Z. Identification and reidentification: The definition of personal information and the legislative choice[J]. Journal of Chongqing University (Social Science Edition), 2018 (2): 119-131.

[30] PURTOVA N. The law of everything: Broad concept of personal data and future of EU data protection law[J]. Law, Innovation and Technology, 2018, 10(1): 40-81.

[31] CORTE L D. Scoping personal data: Towards a nuanced interpretation of the material scope of EU data protection law[J]. European Journal of Law and Technology, 2019, 10(1): 1-26.

[32] LUPTON D. How do data come to matter? Living and becoming with personal data[J]. Big Data & Society, 2018, 5(2): 1-11.

[33] EMAM K E, ARBUCKLE L. Anonymizing Health Data[M]. Sebastopol: O'Reilly Media, 2014: 4-5.

[34] JI S L, MITTAL P, BEYAH R. Graph Data Anonymization, De-Anonymization Attacks, and De-Anonymizability Quantification: A Survey[J]. IEEE Communications Surveys & Tutorials,

2017,19(2):1305-1326.

[35] RUBINSTEIN I S,HARTZOG W. Anonymization and Risk[J]. Washington Law Review,2016,91(2):703-760.

[36] OHM P. Broken Promises of Privacy:Responding to the Surprising Failure of Anonymization[J]. UCLA Law Review,2010,57(6):1701-1778.

[37] BRASHER E A. Addressing the Failure of Anonymization:Guidance from the European Union's General Data Protection Regulation[J]. Columbia Business Law Review,2018(1):209-253.

[38] JORDON J,SZPRUCH L,HOUSSIAU F,et al. Synthetic Data-what,why and how? [EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2205.03257.

[39] OFFENHUBER D. Shapes and frictions of synthetic data[J]. Big Data & Society,2024,11(2):1-16.

[40] JACOBSEN B N. Machine learning and the politics of synthetic data[J]. Big Data & Society,2023,10(1):1-12.

[41] EMAM K E,MOSQUERA L,HOPTROFF R. Practical Synthetic Data Generation[M]. Sebastopol:O'Reilly Media,2020:1,2-3,4-6,19-20,69.

[42] GURSAKAL N,ÇELIK S,BIRISÇI E. Synthetic Data for Deep Learning[M]. New York:Apress Media,2022:1,3,5-6.

[43] JACOBSEN B N. The Logic of the Synthetic Supplement in Algorithmic Societies[J]. Theory,Culture & Society,2024,41(4):41-56.

[44] BUOLAMWINI J,GEBRU T. Gender Shades:Intersectional Accuracy Disparities in Commercial Gender Classification[J]. Proceedings of Machine Learning Research,2018(81):1-15.

[45] CHEN R J,LU M Y,CHEN T Y,et al. Synthetic data in machine learning for medi-cine and healthcare[J]. Nature Biomedical Engineering,2021(5):493-497.

[46] MAYSON S G. Bias in,Bias out[J]. Yale Law Journal,2019,128(8):2218-2301.

[47] PARDEDE S,KOVA V B. Distinguishing the Need to Belong and Sense of Belongingness:The Relation between Need to Belong and Personal Appraisals under Two Different Belongingness-Conditions [J]. European Journal of Investigation in Health,Psychology and Education,2023,13(2):331-344.

[48] CRISTOFARO E D. Synthetic Data:Methods,Use Cases,and Risks[EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2303.01230.

[49] FCA. Using Synthetic Data in Financial Services [EB/OL]. [2024-11-07]. https://www.fca.org.uk/publication/corporate/report-using-synthetic-data-in-financial-services.pdf.

[50] RUSCH T K,BRONSTEIN M M,MISHRA S. A Survey on Oversmoothing in Graph Neural Networks[EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2303.10993.

[51] LIU R B,WEI J,LIU F Y,et al. Best Practices and Lessons Learned on Synthetic Data [EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2404.07503.

[52] ZHI Z F. Information Content Governance of Large Model of Generative Artificial Intelligence[J]. Tribune of Political Science and Law,2023,41(4):34-48.

[53] SHUMAILOY I,SHUMAYLOY Z,ZHAO Y R,et al. The Curse of Recursion:Training on Generated Data Makes Models Forget[EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2305.17493.

[54] GABRIEL I. Artificial Intelligence,Values,and Alignment[J]. Minds and Machines,2020(30):411-437.

[55] RUSSELL S. Human Compatible:Artificial Intelligence and the Problem of Control[M]. New York:Viking Press,2019:137.

[56] ZHOU X H,SU Z,EISAPE T,et al. Is this the real life? Is this just fantasy? The Misleading Success of Simulating Social Interactions With LLMs[EB/OL]. [2024-11-07]. https://doi.org/10.48550/arXiv.2403.05020.

[57] ZOU A,WANG Z F,CARLINI N,et al. Universal and Transferable Adversarial Attacks on Aligned Language Models[EB/OL]. [2024-11-08]. https://doi.org/10.48550/arXiv.2307.15043.

[58] GIUFFRE M,SHUNG D L. Harnessing the power of synthetic data in healthcare:innovation,application,and privacy[J]. NPJ Digital Medicine,2023,6(1):1-8.

[59] WHITNEY C D,NORMAN J. Real Risks of Fake Data:Synthetic Data,Diversity-Washing and Consent Circumvention[EB/OL]. [2024-11-08]. https://doi.org/10.1145/3630106.3659002.

[60] BROWNSWORD R. Law 3.0:Rules,Regulation,and Technology[M]. New York:Routledge,2021:32-33.

[61] VILJOEN S. A Relational Theory of Data Governance[J]. Yale Law Journal,2021,131(2):573-654.

[62] WACHTER S,MITTELSTADT B. A Right to Reasonable Inferences:Re-Thinking Data Protection Law in the Age of Big Data and AI[J]. Columbia Business Law Review,2019(2):494-620.

[63] GAL M S,LYNSKEY O. Synthetic Data:Legal Implications of the Data-Generation Revolution[J]. Iowa Law Review,2024,109(3):1087-1156.

[64] BEDUSCHI A. Synthetic data protection:Towards a paradigm change in data regulation? [J]. Big Data & Society,2024,11(1):1-5.

[65] HUANG W Y. On Precautionary Rule of Law[J]. Chinese Journal of Law,2024,46(2):20-38.

[66] EBERS M. Standardizing AI:The Case of the European Commission's Proposal for an 'Artificial Intelligence Act'[M]// The Cambridge Handbook of Artificial Intelligence:Global Perspectives on Law and Ethics. Cambridge:Cambridge University Press,2022:331.

[67] METROPOLIS N,ULAM S. The Monte Carlo method [J]. Journal of the American Statistical Association,1949,44(247):335-341.

[68] ABUFADDA M,MANSOUR K. A Survey of Synthetic Data Generation for Machine Learning[C]// 2021 22nd International Arab Conference on Information Technology. Muscat,Oman,2021:1-7.

[69] ZHAO P. The Legal Implications of 'Ethical' Governance of

Technology[J]. Peking University Law Journal, 2022, 34(5): 1201-1220.

[70] JOBIN A, LENCA M, VAYENA E. The global landscape of AI ethics guidelines[J]. Nature Machine Intelligence, 2019, 1(9): 389-399.

[71] ZHOU J L, CHEN F. AI ethics: from principles to practice[J]. AI & SOCIETY, 2023, 38(6): 2693-2703.

[72] CABALLERO I, GUALO F, RODRIGUEZ M, et al. Maturity Models for Data Governance[M]// Data Governance. Cham: Springer, 2023: 139.

[73] ABRAHAM R, SCHNEIDER J, BROCKE J. Data governance: A conceptual framework, structured review, and research agenda [J]. International Journal of Information Management, 2019 (49): 424-438.

[74] ALMASLUKH A, ALAMEER A, ALSALEH H, et al. Data Mesh Meets Blockchain[J]. International Journal of Computational Intelligence Systems, 2024(17): 1-15.

[75] MASOOD I, DAUD A, WANG Y L, et al. A blockchain-based system for patient data privacy and security[J]. Multimedia Tools and Applications, 2024(83): 60443-60467.

[76] HASAN H R, SALAH K. Combating Deepfake Videos Using Blockchain and Smart Contracts[J]. IEEE Access, 2019(7): 41596-41606.

[77] PESTANA G, ANTUNES W, CARVALHO J. Digital Chain of Custody Operational Framework[C]// 2023 IEEE International Workshop on Technologies for Defense and Security. Rome, Italy, 2023: 417-422.

[78] LESSIG L. Code: Version 2.0[M]. Cambridge: Basic Books, 2006: 6-7.

[79] ZACCAGNINO R, CAPO C, GUARINO A, et al. Techno-regulation and intelligent safeguards[J]. Multimedia Tools and Applications, 2021(80): 15803-15824.

[80] HILDEBRANDT M. Legal Protection by Design: Objections and Refutations[J]. Legisprudence, 2011, 5(2): 223-248.

[81] ALMADA M. Regulation by Design and the Governance of Technological Futures[J]. European Journal of Risk Regulation, 2023, 14(4): 697-709.

[82] VANNA F D. The Construction of a Normative Framework for Technology-Driven Innovations: A Legal Theory Perspective [M]// Use and Misuse of New Technologies. Cham: Springer, 2019: 193-194.

ZHANG Tao, born in 1991, Ph.D, associate professor, is a member of CCF (No. Y1355M). His main research interests include data law, administrative law and artificial intelligence law.

(责任编辑:柯颖)