面向人工智能的数据治理框架

李继峰[1]，张成龙[2]，刘鑫[3]，陈劲宇[4]，张津铭[5]，毕超[3]
1. 国务院发展研究中心资源与环境政策研究所，北京 100010；
2. 国网能源研究院，北京 102209；
3. 中国农业发展银行，北京 100045；
4. 国网福建省电力有限公司经济技术研究院，福建 福州 350013；
5. 中国信息通信研究院人工智能研究所，北京 100191

摘要
数据对人工智能的开发应用具有至关重要的作用，这已成为工业界和学术界的共识。基于人工智能与数据的互动关系，以及以数据为中心的开发实践，提出面向人工智能的数据治理框架，包含源数据治理、预训练数据治理、评测数据治理、微调数据治理、推理数据治理和运维数据治理6个方面，每个方面都有其重点任务和技术。同时，深入分析ChatGPT、Ziya2和能源领域部分人工智能模型的数据治理案例和成功经验，以验证该框架的有效性。结果表明，该框架在提高人工智能模型性能、优化数据管理流程等方面具有积极作用，对面向人工智能的数据治理的理论和技术创新具有参考价值。

关键词
人工智能；源数据治理；预训练数据治理；评测数据治理；微调数据治理；推理数据治理；运维数据治理

中图分类号：TP18 文献标志码：A doi：10.11959/j.issn.2096-0271.2025004

Data governance framework for artificial intelligence

LI Jifeng[1], ZHANG Chenglong[2], LIU Xin[3], CHEN Jinyu[4], ZHANG Jinming[5], BI Chao[3]
1. Research Institute of Resources and Environmental Policy, Development Research Center of the State Council, Beijing 100010, China
2. State Grid Energy Research Institute, Beijing 102209, China
3. Agricultural Development Bank of China, Beijing 100045, China
4. Economic and Technical Research Institute of State Grid Fujian Electric Power Co. , Ltd. , Fuzhou 350013, China
5. Artificial Intelligence Institute, China Academy of Information and Communications Technology, Beijing 100191, China

Abstract
Data plays a crucial role in the development and application of artificial intelligence, which has become a consensus in industry and academia. Based on the interactive relationship between artificial intelligence and data, as well as the data-centric AI development practice, this paper proposes a data governance framework for artificial intelligence, which consists of six stages: source data governance, pre-trained data governance, evaluation data governance, fine-tuning data governance, inference data governance, and operation and maintenance data governance. Each stage has its key

tasks and technologies. At the same time, this paper deeply analyzes the data governance cases and successful experiences of ChatGPT, Ziya2 and artificial intelligence models in the energy field to verify the effectiveness of the framework. The result shows that the framework plays an important role in improving the performance of artificial intelligence models and optimizing data governance processes. The framework provides reference for theoretical and practical innovation of data governance oriented to artificial intelligence.

Key words
artificial intelligence, source data governance, pre-trained data governance, evaluation data governance, fine-tuning data governance, inference data governance, operation and maintenance data governance

0 引言

在1956年达特茅斯会议上，人工智能（artificial intelligence，AI）作为独立研究领域被提出，其后经历了多个发展阶段，包括早期的探索期、研究的低谷期以及近年来的快速发展期。AI是一个多维度、跨学科的研究领域，旨在通过计算机科学和工程学的方法模拟、延伸和扩展人类的智能。从广义上讲，AI使用计算机系统来执行部分需要人类智能的任务，如视觉感知、语言理解、决策制定和翻译，其应用范围非常广泛，包括但不限于专家系统、自然语言处理、机器学习、模式识别、人工神经网络等，这些技术被应用于工业、医疗、金融、安防等多个领域，极大地提高了人们的工作效率和生活质量。

AI的发展依赖于多种技术和要素，包括基础理论和算法、存储、计算、通信、数据以及相关软硬件的协同优化等。其中，大规模、高质量的数据是AI系统的基础，有效的数据收集、处理、分析等治理技术对于AI系统至关重要，直接影响着AI系统的性能。AI系统的开发应用离不开数据治理，从理论和实践的角度，面向AI的数据治理技术框架值得深入研究和探讨。

1 面向AI数据治理的理论分析

1.1 数据治理与AI发展的关系

数据治理是一套组织策略和实践[1]，通过制定和实施一系列政策、标准、责任和流程来管理、使用和保护数据，以保证数据的准确性、完整性、可用性、安全性、合规性，核心目标是提高数据质量[2]。

AI的发展与数据治理密切相关[3]，良好的数据治理是AI发展的前提条件，而AI技术的应用进一步提高了数据治理的能力和效率。数据治理通过数据清洗、去重和标准化等手段提高数据质量，高质量的数据可以显著提升AI模型的性能和可靠性。此外，数据治理需确保数据处理过程符合法律法规，保护个人隐私数据，这对于建立大众对AI系统的信任至关重要。AI技术在数据治理中发挥着越来越重要的作用，它不仅提升了数据处理的效率和质量，还增强了数据的安全性和合规性，推动数据治理向智能化和自动化方向发展。例如，AI系统能够自动化地进行数据收集、清洗、转换和集成等步骤，显著提高数据治理的效率和准确性；AI可用于数据加密存储和传输、访问控制等，提升了数据的安全性，同时AI还能自动识别和处理潜在的

安全威胁，保障数据免受攻击。

1.2 面向AI进行数据治理的必要性

数据是AI发展的核心要素和关键基础[4]。随着AI技术的不断进步，对数据的需求量和质量要求也在不断提高。因此，建立有效的数据治理框架，确保数据的准确性、完整性和安全性，是推动AI发展的基础。合理的数据治理可实现数据的共享和复用，充分挖掘数据价值和提高数据利用效率，从而降低AI系统的开发和运营成本，推动AI技术和产业健康发展。

数据治理有助于提高AI系统的可信度、可靠性和安全性[5]。规范数据的采集、存储和使用过程，可以减少数据偏差和数据错误，从而提高AI系统的决策准确性和稳定性。随着AI技术的广泛应用，个人数据的采集和使用越来越普遍，建立完善的数据治理体系、加强数据安全和隐私保护，有助于建立大众对AI的信任。

1.3 面向AI进行数据治理的理论逻辑

AI系统的全生命周期主要包括系统规划与需求分析、预训练、评测、微调、部署与推理、维护与退役6个阶段。数据治理贯穿AI系统的整个生命周期[6]，从系统规划到退役，每个阶段都有特定的数据治理任务，以确保数据的质量、安全和合规性，为AI系统的开发应用提供基础支撑（AI系统与数据治理全生命周期各阶段的对应关系见**表1**）。

在AI系统的规划与需求分析阶段，确定AI系统的目标、范围和需求（包括数据需求、功能需求和性能需求），需对源数据进行治理。数据治理包括确定数据需求、数据质量标准和数据安全要求。该阶段需要对数据源进行评估，确保其可靠性、多样性和相关性。

表1 AI系统与数据治理的生命周期对应关系

AI系统的生命周期	数据治理的生命周期
规划与需求分析阶段	源数据的治理
预训练阶段	预训练数据的治理
评测阶段	评测数据的治理
微调阶段	微调数据的治理
部署与推理阶段	推理数据的治理
维护与退役阶段	运维数据的治理

在AI系统的预训练阶段，准备和处理数据，选择合适的算法和模型，进行模型的训练和优化，以达到预期的性能指标，需对预训练数据进行治理，治理的重点是对数据进行清洗、转换和增强，以满足模型训练的要求。这包括处理缺失值、异常值和噪声数据，以及进行特征选择等。

在AI系统的评测阶段，对训练好的模型进行评估（包括性能测试、泛化能力和鲁棒性测试），以及进行必要的调整和优化。为持续提升AI系统性能、避免出现预训练数据被污染（即评测数据被包含在预训练数据集中，从而影响模型评估结果）等情况，AI系统的评测数据不能一成不变，需要适应AI系统性能的提升、工程应用场景的拓展、预训练数据的变化等情况。需对评测数据的多样性与代表性、质量与准确性、规模与结构、复杂性与逻辑性、公平性与无偏性等进行全面动态的治理，以客观反映AI系统的实际性能和任务表现。

在AI系统的微调阶段，为确保AI系统能够在具体应用场景中具有良好的适应性和泛化能力，需要对微调指令数据集进行治理，以便AI系统从微调指令数据集中充分学习到具体场景涉及的专业知识和能力。

在AI系统的部署与推理阶段，应用AI系统进行推理、预测或生成，并监控系统的性能和可靠性，确保满足用户的需求，需从推理数据的输入出发，系统化治理推理数据，关注实时数据的质量监控和异常检测，以确保模型在生产环境中的推理准确性和稳定性。同时，需要管理数据的访问权限，保护用户隐私。

在AI系统的维护与退役阶段，定期更新和维护系统，处理数据和模型漂移，以及在系统不再满足需求或无法继续使用时，进行退役和数据的归档或销毁，需对运维数据进行治理，包括对系统日志、性能指标等数据的收集和分析，用于系统的故障诊断和性能优化。在系统退役时，还要对数据进行归档或销毁，以遵循合规要求。

2 面向AI的数据治理实践范式

2.1 以数据为中心的AI

在早期阶段，AI研究的重点是在给定数据集的前提下，优化模型架构算法设计。然而，局限于给定数据集，把过多注意力聚焦于模型的参数、结构或算法，并不能确保AI模型在现实应用中表现优秀。因为实际任务的数据对于解决实际问题非常重要，通常模型难以从一个领域泛化到另一个领域。更进一步，忽视数据质量与多样性可能引发的数据级联效应，导致准确性下降和持续存在偏差等负面后果，这些问题在高风险领域的AI应用中尤为严重。

鉴于此，学术界和产业界的关注点逐渐转向以数据为中心的AI[7]，致力于实现数据的高质量和多样性。以数据为中心的AI强调在模型架构算法相对稳定的情况下，提升数据的质与量。尽管这一转变仍在进行中，但已有许多成功案例证明了这种范式的优势。

2.2 以数据为中心与以模型为中心的关系

以数据为中心的AI方法并不是要取代以模型为中心的方法，而是二者相互补充，共同推动AI系统的发展[8]。一方面，以模型为中心的技术可以支持以数据为中心的技术的目标实现。例如，可以利用生成模型（如生成对抗网络和扩散模型）进行数据增强，从而生成更多高质量的数据样本。另一方面，以数据为中心的方法也能够促进以模型为中心的技术的进步[9]。例如，数据可用性的提高可能会推动模型设计的进一步创新和改进。因此，在现实生产环境中，数据和模型往往是相互影响、交替演进的，以适应不断变化的环境需求。

数据和模型之间的界限逐渐模糊[10]。传统上，数据和模型被视为两个独立的概念。然而，随着模型能力的增强，算法、架构、参数等模型本身要素已转变为一种特殊形式的数据，可视为数据的载体。通过精心设计的提示，人们利用大语言模型（large language model，LLM）生成所需的数据，而这些数据又可以被用来进一步训练模型。这种方法的潜力已在GPT-4模型上得到了初步验证。

2.3 面向AI数据治理的重点任务

按照以数据为中心的AI实践范式[11]，基于数据治理的全生命周期以及大语言模型开发应用的全过程，面向人工智能的数据治理重点任务及相关技术[12]如下。

一是源数据的治理。根据AI系统规划设计目标，主要从源头和供给侧解决大语言模型训练耗费数据量大、耗费速度快，可能引起“数据短缺”的问题，同时也解

决数据质量不高的问题[13]，以推动训练数据有较为稳定的“源头活水”。

二是预训练数据的治理。预训练数据的治理旨在构建丰富多样且高质量的数据集，以支持机器学习模型的训练，包括数据收集、数据准备、数据浓缩和数据增强。

三是评测数据的治理。这些评测数据集能够对模型的性能进行全面客观的评价，为模型优化升级提供动力，包括同分布评测、异分布评测和评测数据集构建与治理。

四是微调数据的治理。微调数据的治理涉及数据收集、清洗、标注、验证和持续监控，以确保模型系统能够进一步学习专业领域数据的知识，在特定场景任务上具备应有的性能和可靠性。

五是推理数据的治理。其重点是在大语言模型运行推理过程中，通过一些特定的数据设定和输入，或者利用工程化的数据输入来激发模型的特定功能，提高模型的推理性能。

六是运维数据的治理。人工智能持续发展需要不断地维护更新相关基础数据，运维数据治理的目标是在不断变化的环境中确保数据的质量和可靠性，包括数据理解、数据质量保证、数据存储与检索、数据安全治理及合规处置数据与知识数据迁移。

3 面向AI数据治理的技术框架

3.1 源数据的治理

数据是大语言模型的基础，为了提升大语言模型的性能，加强数据源头治理是关键。数据源主要分为通用数据和专业数据两大类。由于规模大、多样性高和易于获取，通用数据（如网页、图书、新闻和对话文本）对于大语言模型的建模能力和泛化能力至关重要。专业数据（如多语言数据、科学数据、代码和特定领域资料）在提升通用大语言模型的性能方面占比较低，但能够有效提升模型在特定任务上的解决能力。

在通用数据方面，网页数据的数量最大，其内容的多样性有助于大语言模型获取丰富的语言知识。然而，网络数据的处理和筛选是复杂的，需要去除低质量内容（如垃圾邮件），以确保数据质量。对话数据（如社交媒体评论和聊天记录）对于提升模型的对话能力和问答任务表现有显著效果，但其收集和处理相对困难。书籍数据作为人类知识的重要载体，能够丰富模型的词汇量和理解能力，尤其是在理解长文本结构和语义连贯性方面可发挥重要作用。

在专业数据方面，多语言数据在提升模型的多语言理解和生成能力方面发挥着关键作用[14]。科学文本数据（如教材、论文和百科）对于提升模型在理解科学知识方面的能力具有重要意义。代码作为一种格式化语言，具有长程依赖和准确的执行逻辑，其语法结构、关键词和编程范式对生成式人工智能的生成功能起着重要作用。编程问答社区和公共软件仓库是代码数据的主要来源，提供了丰富的语境和真实世界中的代码使用场景。

随着模型的复杂度的提高和规模的扩大，其对数据的需求也在不断增加。例如，OpenAI的GPT-3模型接受的数据训练量达到了3 000亿token，而2023年谷歌推出的新一代语言模型PaLM 2的token数量已经突破了3.6万亿。对数据的持续需求可能导致训练数据枯竭。为了解决数据枯竭的问题，必须加强源头数据治理，从供给侧拓宽数据来源，从源头上增加数据规模，提高数据质量。例如：在宏观政策

层面，加快数字化转型，推动产业数字化、治理数字化进程；在中观层面，推动行业、区域数字化转型和数据治理；在微观层面，鼓励引导企业等主体参与产业数字化，将更多的实体关系、经营活动、知识积累转化为高质量的数据资源、数据资产。

3.2 预训练数据的治理

预训练数据为AI模型构建基石，模型的性能在很大程度上取决于数据的质量和数量。预训练数据治理旨在收集并生成丰富且高质量的训练数据，以支持AI模型的训练。

3.2.1 数据收集

传统上，数据集构建从零开始，通过人工收集相关信息来完成，这一过程极为耗时。随着技术的进步，数据集发现、数据集成、数据合成等一系列更加高效的方法被提出和应用，较好地提高了数据收集的效果。

数据集发现是训练数据收集的第一步，旨在识别和选择与目标任务相关的高质量数据集。选择合适的数据集能够确保模型在训练过程中接触到多样化且具有代表性的信息，从而提高模型的泛化能力。数据集发现主要包括以下任务：一是数据源识别，确定潜在的数据源，包括公开数据集、学术资源、互联网内容等；二是数据集评估， 评估数据集的质量、规模、多样性和相关性，以确保其适用于预训练；三是数据许可与合规管理，确保数据集的使用符合相关法律法规和伦理标准，包括数据隐私和版权问题[15]。

数据集成是将不同数据源的数据进行整合和统一的过程，以创建一个大规模、多样化且一致的训练数据集。数据集成可以消除“数据孤岛”，提高数据的可用性和一致性，从而提升模型的训练效果。数据集成主要包括以下步骤：一是数据清洗与预处理，去除数据中的噪声、错误和冗余信息，进行格式转换和标准化，以确保数据的一致性和质量；二是数据融合与匹配，将来自不同数据源的数据进行融合，采用实体匹配和数据对齐技术解决数据冲突和不一致问题；三是数据增强，采用数据扩充、数据变换等增强技术，增加数据的多样性和规模，以提高模型的鲁棒性和泛化能力。

数据合成是通过生成新的数据样本来补充现有数据集的过程，以解决数据稀缺或不平衡的问题。数据合成可以增加数据的多样性和扩大数据的覆盖范围，从而提升模型的训练效果和鲁棒性。数据合成主要包括以下内容：一是构建数据生成模型，基于已有真实数据或数据规律构造生成模型；二是使用生成模型（如生成对抗网络、变分自编码器等）生成新的数据样本，以模拟真实数据的分布和特征；三是数据增强与混合，采用数据插值、数据融合等数据增强和混合技术，将生成的数据与现有数据相结合，以增加数据的多样性和规模；四是数据评估与验证，评估合成数据的质量和有效性，确保其与真实数据具有相似的分布和特征，以避免对模型训练产生负面影响。

3.2.2 数据准备

数据准备是将原始数据转换为适合AI模型训练的格式的过程。数据准备是非常重要的一步，因为原始数据通常存在噪声、不一致性和无关信息，如果不进行适当的清洗和转换，会导致模型过拟合、泛化能力不足等问题。

数据准备包括以下步骤：一是数据清

洗，识别并修正数据中存在的错误、不一致和不准确等问题，如填补缺失值、去除重复数据等；二是特征提取，从原始数据中提取相关的特征，如图像的颜色、纹理特征，时间序列数据的统计和频谱特征等；三是特征转换，将原始特征转换为新的特征，以提高模型性能，如归一化、标准化、对数变换等。

3.2.3 数据浓缩

数据浓缩通过减少数据的特征数量或样本数量来降低数据复杂度，同时尽可能保留数据的关键信息。它有助于减少对内存和计算资源的需求，提高模型训练和部署的效率；缓解过拟合的情况，提高模型的泛化能力；提高模型的可解释性，使模型更容易理解。

数据浓缩的主要方法有：一是特征规模压降，选择最相关的特征子集，具体包括过滤法、包裹法和嵌入法；二是维度压降，将高维特征映射到低维空间，如主成分分析（PCA）和线性判别分析（LDA）等线性方法，以及自编码器等非线性方法；三是实例选择，选择最具代表性的样本子集，包括基于模型性能的包裹法和基于统计特性的过滤法。

3.2.4 数据增强

数据增强是一种通过人工创造新的训练样本来增加数据集大小和多样性的技术，其主要目的如下。一是提高模型的准确性、泛化能力和鲁棒性。现代机器学习算法通常需要在大量数据上学习，但获取大规模数据困难且耗时，数据增强通过自动化生成相似的新样本来解决数据不足的问题。二是缓解数据类别不平衡的情况。数据增强可以通过增加对少数类别的数据样本的采样来平衡数据分布。

数据增强的主要方法如下：一是基本简易操作方法，如图像的缩放、旋转、翻转、模糊化处理等，这类方法直接对原始数据进行简单的变换；二是数据合成方法，利用生成模型学习数据的分布，并生成新的合成样本，这类方法从全局角度学习数据模式，生成更具有代表性的新样本；三是针对数据类别不平衡的方法，如合成少数类过采样技术（synthetic minority over-sampling technique，SMOTE）、自适应合成（adaptive synthetic，ADASYN）采样方法等在少数类别样本附近插值生成新样本。SMOTE是一种针对数据类别不平衡问题的数据增强方法，通过在少数类别样本与其最近邻样本之间进行线性插值来生成新的合成样本，这可以有效增加少数类别的样本数量，缓解类别不平衡的情况。ADASYN是SMOTE的一种扩展方法，根据每个少数类别样本的学习难度（由其最近邻样本中的多数类别样本比例决定）来动态调整生成新样本的数量。对于那些更难学习的少数类别样本，该方法会生成更多的合成样本，这可以进一步提高模型对少数类别的学习能力。

3.3 评测数据的治理

评测数据治理的目标在于顺应AI大语言模型技术发展的趋势和实际应用场景的需要，构造合理的评测数据集，并适时更新或升级此数据集，以对大语言模型的综合性能和单项能力进行评定。

3.3.1 同分布评测

同分布评估是指生成符合训练数据分布的样本作为评测数据集，以评估模型在

特定子群体上的性能，并验证检查模型的性能边界。同分布评估旨在更细粒度地评估模型的性能，以发现其在特定子群体上的不足，并检查模型的伦理合规性，这对于构建可靠和安全的AI系统至关重要。这种评测方式有以下作用。一是发现模型在哪些训练数据集的子集上性能欠佳。模型在整体上表现良好，但可能会在某些特定训练数据子集上表现不佳，需要识别这些代表性的子集并进行调整，以避免出现偏差和错误，特别是在高风险应用中。二是分析验证模型的能力边界。理解模型的决策边界并在部署前检查其伦理合规性是至关重要的，尤其是在涉及政策制定等的高风险应用中。

同分布评测主要方法包括：一是数据切片方法，将数据集划分为相关的子群体，并分别评估模型在每个子群体上的性能，这可以使用预定义的标准（如年龄、性别、种族等）或自动化的切片方法；二是算法可解释性方法，生成一组假设性样本，这些样本可以改变模型的决策结果，帮助识别导致模型预测错误的最小输入变化，以检查模型的决策边界。

3.3.2 异分布评测

异分布评测使用与训练数据分布不同的样本作为评测数据集，以全面评估模型的性能，为模型部署前的安全性和可靠性提供保障。其主要作用如下：一是评估模型在意外场景下的泛化能力，训练数据和实际部署环境的数据分布可能存在差异，异分布评测可以揭示模型在这种差异情况下的表现；二是检测模型的鲁棒性，将生成对抗样本作为评测数据集以发现模型存在的弱点，从而采取措施提高模型的安全性[16]。

异分布评测的方法主要有：一是将生成对抗样本作为评测数据，通过对输入数据施加人为扰动，制造能够误导模型的样本，评估模型的鲁棒性；二是将生成分布偏移样本作为评测数据，通过偏斜采样或学习生成模型的方式，构造与训练数据分布不同的样本数据作为评测数据，评估模型在分布差异下的表现。

3.3.3 评测数据集的治理

评测数据集是评估和比较不同模型性能的关键工具。评测数据集的治理需要关注以下几个方面。一是评测数据集的数量。大语言模型开发应用进入快速发展阶段，单模态、多模态、通用型、垂直型等各类大语言模型不断涌现，需要更多类型、更多数量的评测数据集对各类大语言模型进行评测，但目前评测数据集的类型和数量都相对较少。二是评测数据集的质量。其对于提高模型评估的准确性至关重要，直接影响评测结果的准确性和可靠性，直接或间接影响大语言模型开发应用各环节的数据治理效果。三是评测数据集的设计和选择。设计选择评测数据集，还应考虑信度、效度和难度等因素，以确保数据集能够有效地反映模型的真实性能。即使是小型或合成的数据集也能够驱动模型创新，在选择评测数据集时，不仅要考虑数据的规模，还要考虑其能否全面覆盖模型应用场景中可能出现的各种情况。四是评测数据集的多样性和代表性。高阶多数据集建模的研究表明，利用多模态、多类型的数据集可以更有效地解决传统数据处理和分析方法失效的问题，在设计评测数据集时，应尽可能地考虑数据的多样性和代表性，以确保模型能够在多种不同的场景下被有效评估。五是评测数据集的隐私保护和用户参与问题。在设计和使用评测数据集时，必须平衡数据质量与数据

隐私、用户权益保护。

大语言模型评测数据集治理面临的问题包括但不限于评测数据集的数量、质量、设计和选择、多样性和代表性以及隐私保护等方面[14]。需要综合考虑数据集的设计原则、应用场景以及技术手段等，以确保评测数据集能够有效地支持大语言模型的性能评估和优化。

3.4 微调数据的治理

大语言模型经过预训练具备了通用知识能力，要将其应用于具体的行业实际，还需具备行业的专业知识和能力，这需要借助指令微调来实现。指令微调的基础是构建指令微调数据集，让大语言模型在指令微调数据集上进行学习，要使大语言模型取得预期的微调效果，需对微调数据集进行科学有效的治理。

3.4.1 数据标注

数据标注是为数据集中的元素分配描述性标签的过程，对于大语言模型微调至关重要，因为大语言模型微调使用的数据最好是标注过的高质量数据。传统上，因极其耗时且资源密集，尤其在处理大规模数据集时，数据标注面临巨大挑战。近年来，研究焦点逐渐转向通过减少人工干预同时保持标签准确性的方式来提升标注效率。具体策略包括利用未标注数据的半监督学习和主动学习方法，减少对显式标签的需求，以及通过众包技术加速标注过程，尽管这带来了数据一致性和质量控制的新难题。此外，先使用预训练模型进行初步标注、再由专家审核的半自动标注工具的开发应用，已成为有效降低数据标注劳动强度的途径。

3.4.2 指令微调数据集的治理

虽然经过大规模预训练，模型能够捕获语言的普遍规律和潜在知识，模型最初的设计目标是预测文本序列中的下一个词，这限制了模型直接理解和执行详细指令的能力。指令微调使大语言模型学习有标注的特定任务数据，熟悉如何解读和响应具体的指令性文本，从而实现从通用语言理解向任务导向型智能的转变。有效构建、治理指令微调数据集是进行指令微调、确保模型性能的关键步骤，具体策略和方法如下。

一是注重指令数据的来源和收集。从公开数据集、人类标注数据、自动生成数据等多渠道收集高质量数据，挖掘合适的指令模板，或使用种子指令进行改写，形成指令和对应输出的数据配对，从而提高指令数据的多样性，确保数据集覆盖多种指令类型和领域，以提高模型的泛化能力。

二是注重数据标注和管理。统筹自动标注和人工标注，为指令数据添加高质量的标签。借助自动标注平台或工具提高标注效率，如利用预训练模型生成初步的指令和输出，然后进行人工审核和修正。利用人工标注提高准确性，专业人员进行高质量的数据标注，可确保数据的准确性和伦理合规性。加强数据版本控制，使用版本控制系统管理数据集的不同版本，确保数据的可追溯性。

三是注重数据预处理。进行数据清洗，去除噪声数据和不一致的指令－输出对。对数据进行标准化处理，统一指令和输出格式，以便模型更好地理解和处理指令微调数据。进行数据增强，通过同义词替换、随机插入或者删除等操作增加数据的多样性。

四是注重数据集评估和验证。对数据集开展质量评估，建立包括准确性、及时

性、一致性等在内的客观指标，以及专家评估的主观指标，定期评估指令数据的质量。结合运用自动评估和人工评估方法：自动评估可使用BLEU、ROUGE等指标评估微调数据集的质量；人工评估即通过人工审核数据集，确保指令和输出的准确性和一致性。根据评估结果，采取数据清洗、特征工程等措施来提高指令数据的质量。

五是注重数据集组合。多任务微调可提升大语言模型的泛化性能，增加微调任务数量的好处在不同规模模型上得到了验证，因此，有必要组合多个不同任务构成具有多样性的指令微调数据集。不同任务数据的混合比例很关键，通常由实验和经验决定。为了让大语言模型解决特定任务，可依据表示相似性和梯度相似性选择相关多任务子集。但是需注意，不同任务间可能存在冲突，组合数据量过大可能因数据格式和分布的相似性削弱模型能力。

六是注重数据集的持续改进。建立反馈机制，收集模型在实际应用中的表现，持续改进数据集。定期更新数据集，确保数据集的时效性和相关性。同时，加强数据集的文档和元数据管理。详细记录数据集的来源、构建方法、标注过程和使用说明，确保数据的透明性和可追溯性。管理数据集的元数据，包括数据格式、标注信息、使用场景等，方便检索和使用数据。

3.5 推理数据的治理

推理数据治理是指在应用大语言模型进行推理的过程中，根据具体应用场景或执行推理任务的特点，有针对性地设计数据输入或者指令提示，嵌入必要的检索增强数据知识库，引入思维链，激发模型的特定能力，提高推理决策的准确性。

3.5.1 提示工程的数据治理

提示工程是一种通过设计和构建高质量的模型输入提示来实现特定任务的方法。它通过设计构造输入数据而不是调整模型本身来达到预期目标，可以指导大语言模型完成复杂的任务，相比于微调模型更加灵活高效，可以快速探索模型的知识能力。

为更好地提升模型推理性能，有必要从数据生成、质量控制、存储检索等多个角度对提示工程数据集进行系统性的设计、优化和治理，以确保提示数据的高质量和可用性。一是手动设计提示模板并自动生成提示数据集。可以从外部语料库中挖掘模板，或使用种子提示进行改写，以丰富提示的多样性。二是采用梯度搜索或生成模型等学习方法自动生成提示。该方法可更有效地探索模型的知识，发现最优的提示。三是建立提示工程数据质量评估机制。定期检查提示数据集的质量，并采取措施进行改进。使用机器学习模型自动检测数据质量问题，并通过人工参与等方式持续优化数据质量。四是设计高效的提示数据存储和检索系统。为确保在模型部署、推理时能够快速获取所需的提示数据，可以采用资源分配优化、查询加速等方法来提高提示数据的获取效率。

3.5.2 检索增强生成的数据治理

检索增强生成（retrieval-augmented generation, RAG）技术是在模型推理阶段引入外部数据知识进行辅助增强的技术，可以显著提高大语言模型的推理性能和准确性，预防出现幻觉。RAG的框架主要由索引、检索器、增强器和生成器4个核心组件构成。在索引阶段，对外部数据知识进行向量化索引；在检索阶段，利用向量相似性技术快速检索与用户

查询相关的文档；在增强阶段，将用户查询与检索到的上下文结合，形成较精炼准确的组合查询提示；在生成阶段，将组合后的查询提示传递给模型，生成最终响应和输出。

大语言模型检索增强技术实现的基础在于构建和治理大语言模型外挂的数据知识库。检索增强知识库的治理步骤如下。一是进行数据收集与预处理，根据大语言模型推理应用的专业领域需求，收集大量的基础专业数据，包括书籍、文章、网页内容等，并进行清洗、格式化和标准化处理。二是进行知识表示与抽取，将知识以结构化或半结构化的形式进行表示，并使用自然语言处理技术从文本中抽取知识。三是进行知识融合与推理，将抽取的知识融合到知识库中，解决知识冲突和冗余问题，并利用知识库进行推理，发现新的知识或关系。四是进行知识更新与维护，定期更新知识库，以反映最新的信息和知识，并保持知识库的动态性和准确性。

3.5.3 思维链的数据治理

思维链可提升大语言模型的推理能力。思维链是类似于人类思维的逐步推理过程，通过构建一个包含这些思维链的数据库，模型可以参考它们来改进自身的推理过程。为了确保这些思维链的质量，需要专家进行审核和标注，以保证其正确性和逻辑性。此外，问题的多样性也至关重要，数据库应包含来自不同领域的例子，以提高模型的泛化能力。

在模型推理过程中，如何有效地访问和利用这个数据库是一个值得考虑的问题。可能需要采取混合方法，一部分思维链用于训练，另一部分在实时推理时进行检索。随着问题和思维链数量的增加，数据库的可扩展性和检索效率成为一个挑战，可采用图数据库或索引系统来优化管理。安全性和隐私性也是不可忽视的问题[15]，特别是在数据库包含敏感信息或被应用于重要系统时，必须采取措施防止未经授权的访问和潜在的篡改。此外，评估该数据库对模型性能的影响是必要的，需要开发相应的指标来检验思维链的引入是否提升了模型的推理能力和预测生成的准确性。

3.6 运维数据的治理

大语言模型运维阶段在全生命周期中占据较大的时间比例，这一阶段的数据治理范围不仅覆盖大语言模型运维数据的监控管理运用，还包括前4个阶段数据的维护优化更新。运维数据的治理是一个多层次、持续进行的过程，致力于提高数据在动态环境中的质量和可靠性。

3.6.1 数据理解

为了进行有效的维护，首要任务是深入理解数据。数据理解不仅要识别数据类型和结构，还要求深入探究数据的内涵，包括但不限于数据的来源、演变历程、内在关系和潜在偏见。数据理解可借助高级可视化、数据估值等技术。高级可视化技术可以揭示数据的分布模式和异常，数据估值技术则评估数据对特定目的的价值，确保维护的数据是相关的、有价值的，并且适合于预定的应用场景。

3.6.2 数据质量保证

实际应用中，数据基础设施频繁、持续更新，影响了数据质量。因此，数据治理不仅需要构建高质量的训练或推理数据，更要在不断变化的环境中维持其卓越性。在动态环境中确保数据质量有两个核心方

面：一是持续监控数据质量，实际应用中的数据复杂多变，可能包含与预期目标不符的异常数据点，因此建立定量的评估标准来衡量数据质量至关重要；二是质量改进，如果模型受低质量数据的影响，实施质量改进策略以提升数据质量变得至关重要，这直接关联到模型性能的提升。

3.6.3 数据存储与检索

存储与检索为AI系统快速准确地提供数据，目前已有多种加速数据获取的策略。数据存储不仅要确保数据的安全性和完整性，还要优化数据的访问速度。查询加速技术，如索引优化、数据缓存策略，以及利用分布式存储和并行处理技术，大幅缩短了数据检索的时间，提升了系统的响应效率。然而，这些策略的实施也带来了存储空间管理的复杂性、数据一致性和分布式系统中的同步等问题。因此，设计灵活且高效的存储架构，平衡存储效率与检索速度，成为AI系统数据管理的重要内容。

3.6.4 数据安全治理

数据安全治理始终是数据治理不可忽视的重要内容[17]，需综合采取以下治理策略：遵循数据最小化原则；实施加密传输与存储；严格进行访问控制及身份验证；实时监控并检测异常；定期开展安全审计与渗透测试；应用隐私保护技术确保合规；构建分层防御体系，建立应急响应计划，全方位保护数据免受内外威胁，保障服务稳定与用户信息安全。

3.6.5 数据合规处置与迁移

处置数据是数据治理的最后一步。一是对数据进行归档与备份。对大语言模型训练和运行过程中产生的大量数据进行分类和评估，将其划分为核心数据或辅助数据。针对核心数据（如高质量的训练样本、模型参数等），应进行长期归档备份，以备未来研究、审计或复用；针对辅助数据，应依据其价值决定保留或销毁。二是保护隐私。在数据处置过程中注重隐私保护与合规处理，严格遵守数据保护法规，对涉及用户个人信息的数据进行匿名化处理或彻底删除，确保不违反隐私保护政策。三是数据迁移与整合。为仍有价值的数据规划合理的迁移路径，将其整合至新的数据管理系统中，以便后续利用。当大语言模型退役时，可考虑利用迁移学习技术将大语言模型在特定任务上的学习成果转移到新模型或新任务上，实现模型知识数据的迁移和复用。

4 面向AI数据治理的案例与经验

4.1 ChatGPT的数据治理实践

在探讨大语言模型的发展历程中，特别是GPT系列大语言模型，模型性能的提升不仅与参数量的增加相关，还与数据质量优化紧密相关。GPT系列模型的相关研究揭示了大语言模型数据治理方面的细致工作，其策略涵盖了上述数据治理框架的多个重要方面。

4.1.1 训练数据的治理演进

GPT模型的成功依赖于多个因素，模型参数的数量增加只是其中之一。对比研究GPT-1、GPT-2、GPT-3、InstructGPT和ChatGPT/GPT-4的相关论文发现，GPT模型通过改进的数据收集、标记和准备策略，显著提升训练数据的数量和质量。

训练数据的治理是大语言模型性能提升的关键性因素。

GPT-1：在BooksCorpus数据集上进行训练，该数据集包含4 629 MB原始文本，涵盖各种书籍类型，对训练数据的治理不够重视。

GPT-2：通过爬取Reddit链接创建WebText数据集，并将其用于模型的预训练。研发团队开始重视训练数据的治理，具体策略如下：一是对Reddit链接进行过滤，爬取高质量的文本数据；二是使用Dragnet和Newspaper工具对文本数据进行提纯；三是基于启发式策略进行数据去重和数据清理（数据准备）。经过数据治理，得到40 GB文本（约为GPT-1使用数据量的8.6倍），GPT-2无须微调即表现出良好的性能。

GPT-3：主要在Common Crawl数据集上训练，这是一个庞大但质量较差的数据集。采用的数据治理策略如下：一是训练分类器，过滤低质量文档；二是使用WebText判断文档质量；三是使用Spark的MinHashLSH进行数据去重；四是扩展WebText训练数据集，添加较高质量的书籍语料库和Wikipedia数据。对45 TB纯文本数据进行治理后，获得570 GB文本（进行了严格的数据质量控制，选用率仅为1.27%），在此更高质量更大规模训练数据集上训练得到的GPT-3模型，其性能超过GPT-2。

InstructGPT：在人类反馈的基础上进行强化学习微调，以符合人类期望。采用的数据治理策略如下：一是使用数据标注技术，用人类反馈答案的数据进行监督学习微调；二是通过考试和问卷的严格过程选择标注者，确保数据标注质量；三是构建比较数据集（按质量排序的人类评估答案）以训练奖励模型，然后使用人类反馈的强化学习（reinforcement learning from human feedback，RLHF）进行微调。通过前述数据治理，InstructGPT生成了更真实、无偏见、更符合人类期望的答案。

ChatGPT/GPT-4：随着产品商业化进程推进，数据治理等相关训练信息不再披露。ChatGPT/GPT-4很大程度上遵循了Transformer的架构设计，并在更高质量、更大规模的强化学习数据集上使用RLHF方法对模型进行微调，大幅提升模型性能。

从GPT-1到ChatGPT/GPT-4的训练数据治理经历了如下变化：较低质量、较小规模的数据集→更高质量、更大规模的数据集→更高质量、更大规模、引入人类反馈的标注数据集。与此同时，除了增加参数以适应更多的训练数据，模型算法结构设计没有重大调整，这表明了数据治理的重要性。

4.1.2 推理数据的治理演进

针对大语言模型的推理数据开发与治理研究仍处于初期阶段。在不久的将来，基于特定任务的推理数据开发方法将逐渐适应大语言模型，如构建对抗性攻击数据以测试模型鲁棒性。

当前的ChatGPT/GPT-4模型已达到高度复杂的水平，可以通过仅调整提示（推理数据输入）来实现各种目标。未来，许多AI从业者可能不再需要训练或微调模型，而是专注于提示工程。然而，提示工程是一个依赖经验的、具有挑战性的任务，即使是语义上相似的提示也可能产生显著不同的输出。在这种情况下，需要采用更加多样化的推理数据治理技术或策略，以提高模型的推理效果。

4.1.3 运维数据的治理演进

ChatGPT/GPT-4在数据维护方面花费了大量精力。作为商业产品，ChatGPT/GPT-4不可能只训练一次就停滞，其运维数据需要不断被更新和维护。一是持续进行数据收集，通过用户输入的提示和提供的反馈进一步改进模型。在这个过程中，模型开发者需要设计指标来监控数据质量以及维护数据质量的策略，以收集更高质量的数据。二是加强数据理解，开发各种工具来可视化和理解用户数据，以更好地理解用户需求并指导未来的模型改进。三是采用高效的数据处理技术，随着ChatGPT/GPT-4用户的快速增长，要开发高效的数据管理系统，以便快速检索用于训练和测试的相关数据。

4.2 Ziya2大语言模型的数据治理实践

Ziya2研究团队致力于持续预训练技术的开发[18]，在保持模型的大小和结构基本不变的前提下，深入分析高质量的预训练数据如何显著提升大语言模型的性能。为此，研究团队以Meta AI公司130亿参数的Llama2模型为基础，在高质量训练数据集（约7 000亿个中英文token）上进行了持续预训练，最终推出了Ziya2模型。预训练过程分为3个阶段，具体采取了以下数据治理策略。

在第一阶段，对接近LLaMA2原始分布的英文数据进行采样，并对中文数据进行了清洗，对代码数据进行了格式化，对这些数据进行混合，形成了高质量的无监督数据集，并进行预训练。在此阶段，训练数据被完全随机化，不同的数据片段被拼接成4 096个token的样本，并利用注意力掩码避免不同数据片段之间相互干扰，从而最大限度地提高训练效率。

在第二阶段，引入中文和英文标注数据，如Wanjuan-Idea数据集，增强Ziya2在下游任务上的性能。与第一阶段随机组合数据的方式不同，这一阶段将相同类型的标注数据拼接成一个样本，并确保每个样本中拼接的数据都是完整的。

在第三阶段，增加了与数学相关的标注数据，如MetaMath数据集，数据的拼接方式与第二阶段保持一致。经过这一阶段的预训练，Ziya2显著提升了数学推理能力和编程能力。这一结果表明，数学推理数据对于编程这类逻辑性较强的任务至关重要。为了防止Ziya2在预训练中出现灾难性遗忘，第二阶段和第三阶段额外采样了与标注数据同比例的中英文无标注数据构建训练数据集，以进行持续的预训练。

经过这一系列的训练，Ziya2团队成功打造了130亿参数的Ziya2模型。对比基准模型，Ziya2模型在各项评估指标上均展现了显著的性能提升。具体而言，以LLaMA2为标准进行LLM评估，Ziya2在MMLU上提高了10%，在CMMLU上提高了61%，在C-Eval上提高了68%，在GSM8K数学问题解答任务上提升了138%，在MATH数学问题解答任务上提升了120%，在HumanEval代码生成任务上提升了89%。相较于其他开源的、规模相当的大语言模型，Ziya2在中文及英文通用任务上取得了领先地位，在数学和编程领域任务上的表现显著优于同类模型。这表明，采用高质量的数据集和恰当的持续预训练策略，可以在不大幅度增加模型参数规模的情况下，有效提升大语言模型的性能表现。

4.3 能源领域AI大语言模型的数据治理实践

在能源领域，AI大语言模型的应用已

经取得了显著进展，以数据为中心的人工智能开发应用范式发挥着重要作用。

中国南方电网有限责任公司的“大瓦特”大语言模型主要应用于智能客服、输电巡检、负荷预测等任务。该模型整合了电力行业的专业知识和海量数据，构建了一个跨自然语言和计算机视觉模态的大语言模型，能够处理复杂的电力系统任务，如巡检报告自动生成和故障预测等。在“大瓦特”大语言模型的构建过程中，数据治理发挥了基础性作用，重点在数据的收集、清洗和标注，并通过不断优化数据质量和丰富数据样本，提升了模型的准确性和泛化能力。

国家能源集团的能源通道大语言模型主要用于煤炭、电力、铁路、港口、航运、化工等多领域的智能查询、智能平衡、智能预警和智慧分析。该模型利用生产运营过程中的设备、货物、物流、销售、气象等数据，对通用大语言模型进行强化训练，形成了具备能源专业知识的行业大语言模型。该模型数据治理融合了产业特定数据与通用数据，注重提高数据的质量和多样性，从而提升模型在特定能源场景中的应用效果。

上海全应科技有限公司的热电云平台模型的应用场景主要是热电生产的智能调控，以提升发电效率和减少碳排放。该模型通过AI技术对热电生产过程进行全自动智能调控，优化发电过程中的各项参数。该公司在数据收集和处理上投入大量资源，确保数据的准确性和实时性，从而使AI模型能够进行精准的预测和调控。

国网山东电力公司的AI中台代表性应用场景包括智能巡检、智能营销与客服等。该公司与百度智能云合作，搭建了AI中台，利用大语言模型技术提升电力系统的智能化水平。其数据治理的重点是数据的标准化和统一化管理，通过构建高质量的数据集提升了AI模型的训练效果和应用性能。

上述案例充分体现了以数据为中心发展人工智能的核心思想，即通过高质量的数据治理来驱动AI模型的性能提升，主要治理策略如下：一是注重数据收集与清洗，提升源数据治理效果，确保数据的全面性和准确性；二是注重数据集成与增强，提升预训练数据治理效果，将不同来源的数据进行集成融合，提升数据的多样性和覆盖面；三是突出数据治理的中心地位，注重数据标注与管理，通过专业的数据标注和管理工具，提升数据的可用性和训练效果；四是注重数据持续优化与模型升级迭代，加强运维数据的治理，通过不断的数据治理和模型迭代，提升AI模型的性能和适应性。

5 结束语

在人工智能研究及开发应用领域，以数据为中心的方法逐渐占据核心地位。经过学术界和产业界多年的不懈努力，人工智能相关模型架构设计日趋完善，特别是自Transformer架构问世以来，其潜力被持续挖掘中。目前，提升数据集的规模和质量已经成为增强AI系统性能的关键途径。源数据治理、预训练数据治理、评测数据治理、微调数据治理推理数据治理和运维数据治理将更紧密地融合在AI系统开发应用全过程中，成为推动人工智能发展的关键支撑力量。目前，大语言模型技术未被应用于双碳目标、节能减排、应对气候变化等细分领域，下一步相关人员可结合能源环境和应对气候变化专业领域的特点，对该专业领域的大语言模型开发应用进行尝试，将面向人工智能的数据治理框

架和技术应用于能源－环境－经济复杂系统和应对气候变化建模，以对省间多区域协同减排关键技术进行智能化组合生成、发掘评价，进而在具体应用中进一步丰富和完善面向人工智能的数据治理理论框架和技术实践。

参考文献：

[1] 李汶龙，袁媛，安筱鹏．刍议大数据治理的三大基础思维[J]. 大数据，2022，8(4)：34-45.
LI W L，YUAN Y，AN X P. Modus operandi of big data governance：some preliminary observations[J]. Big Data Research，2022，8(4)：34-45.

[2] 印鉴，朱怀杰，余建兴，等．大数据治理的全景式框架[J]. 大数据，2020，6(2)：19-26.
YIN J，ZHU H J，YU J X，et al. A panoramic framework of big data governance[J]. Big Data Research，2020，6(2)：19-26.

[3] POLYZOTIS N，ZAHARIA M. What can data-centric AI learn from data and ML engineering?[EB]. arXiv preprint，2021，arXiv：2112.06439.

[4] HAJIJ M，ZAMZMI G，RAMAMURTHY K N，et al. Data-centric AI requires rethinking data notion[EB]. arXiv Preprint，2021，arXiv：2110.02491.

[5] 夏正勋，唐剑飞，罗圣美，等．可信AI治理框架探索与实践[J]. 大数据，2022，8(4)：145-164.
XIA Z X，TANG J F，LUO S M，et al. Exploration and practice of trusted AI governance framework[J]. Big Data Research，2022，8(4)：145-164.

[6] 代红，张群，尹卓．大数据治理标准体系研究[J]. 大数据，2019，5(3)：47-54.
DAI H，ZHANG Q，YIN Z. Study on big data governance standard system[J]. Big Data Research，2019，5(3)：47-54.

[7] MAZUMDER M，BANBURY C，YAO X Z，et al. DataPerf：benchmarks for data-centric AI development[C]//Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS 2023). New York：Curran Associates Inc.，2024：5320-5347.

[8] JARRAHI M H，MEMARIANI A，GUHA S. The principles of data-centric AI[J]. Communications of the ACM，2023，66(8)：84-92.

[9] 李国杰．大数据与计算模型[J]. 大数据，2024，10(1)：9-16.
LI G J. Big data and computing models[J]. Big Data Research，2024，10(1)：9-16.

[10] WANG Z G，ZHONG W J，WANG Y F，et al. Data management for training large language models：a survey[J]. arXiv preprint，2024，arXiv：2312.01700v3.

[11] ZHA D C，BHAT Z P，LAI K，et al. Data-centric AI：perspectives and challenges[EB]. arXiv preprint，2023，arXiv：2301.04819.

[12] 杜小勇，陈跃国，范举，等．数据整理：大数据治理的关键技术[J]. 大数据，2019，5(3)：13-22.
DU X Y，CHEN Y G，FAN J，et al. Data wrangling：a key technique of data governance[J]. Big Data Research，2019，5(3)：13-22.

[13] 秦之湄，张会平，王斌，等．智慧治理中的数据质量管理困境及对策研究[J]. 大数据，2024，10(5)：151-167.
QIN Z M，ZHANG H P，WANG B，et al. Research on the difficulties and countermeasures of data quality management in smart governance[J]. Big Data Research，2024，10(5)：151-167.

[14] 刘倩倩，刘圣婴，刘炜．图书情报领域大语言模型的应用模式和数据治理[J]. 图书馆杂志，2023，42(12)：22-35.

2025004-16

LIU Q Q, LIU S Y, LIU W. Data governance and application development of large language models in library and information services[J]. Library Journal, 2023, 42(12): 22-35.

[15] 孟小峰, 王雷霞, 刘俊旭. 人工智能时代的数据隐私、垄断与公平[J]. 大数据, 2020, 6(1): 35-46.
MENG X F, WANG L X, LIU J X. Data privacy, monopoly and fairness for AI[J]. Big Data Research, 2020, 6(1): 35-46.

[16] 杜跃进. 数据安全治理的几个基本问题[J]. 大数据, 2018, 4(6): 85-91.
DU Y J. Several basic questions about data security governance[J]. Big Data Research, 2018, 4(6): 85-91.

[17] 马朝辉, 聂瑞华, 谭昊翔, 等. 大数据治理的数据模式与安全[J]. 大数据, 2016, 2(3): 83-95.
MA C H, NIE R H, TAN H X, et al. Research on data schema and security in data governance[J]. Big Data Research, 2016, 2(3): 83-95.

[18] GAN R Y, WU Z W, SUN R L, et al. Ziya2: data-centric learning is all LLMs need[EB]. arXiv preprint, 2023, arXiv:

作者简介

李继峰（1979– ），男，博士，国务院发展研究中心资源与环境政策研究所气候政策研究室主任、研究员，主要研究方向为能源经济环境系统分析建模、能源战略、碳减排政策分析。

张成龙（1976– ），男，博士，国网能源研究院高级工程师，主要研究方向为能源电力供需预测及预警技术。

刘鑫（1986– ），男，博士，中国农业发展银行总行处长，主要研究方向为数字金融、金融数据治理、人工智能。

陈劲宇（1997– ），男，国网福建省电力有限公司经济技术研究院中级研究员，主要研究方向为能源战略与政策、低碳技术。

张津铭（1994– ），男，中国信息通信研究院人工智能研究所工程师，主要研究方向为人工智能安全与治理。

毕超（1985– ），男，博士，中国农业发展银行总行处长、高级工程师，主要研究方向为大语言模型、数字金融、金融科技。

收稿日期： 2024-05-06

通信作者： 毕超，znufebc@163.com

基金项目： 国家电网有限公司总部管理科技项目（No.1400-202357320A-1-1-ZN）

Foundation Item: Headquarters' Management Science and Technology Project of State Grid Corporation of China (No. 1400-202357320A-1-1-ZN)