可解释、可通用的下一代人工智能方法重大研究计划2022年度项目指南

“可解释、可通用的下一代人工智能方法”重大研究计划面向人工智能发展国家重大战略需求，以人工智能的基础科学问题为核心，发展人工智能新方法体系，促进我国人工智能基础研究和人才培养，支撑我国在新一轮国际科技竞争中的主导地位。

　　基于深度学习的人工智能方法在许多场景取得了重要突破，但仍然存在模型可解释性差、对抗样本鲁棒性差、数据与算力需求大、理论基础薄弱等问题。基于以上问题，本重大研究计划旨在建立规则和学习的有效融合机制，打破现有深度学习“黑箱算法”的现状，建立一套可适用于不同领域、不同场景（语音、图像、视频等）的通用方法体系。

　　一、科学目标

　　本重大研究计划面向以深度学习为代表的人工智能方法鲁棒性差、可解释性差、对数据的强依赖等基础科学问题，挖掘机器学习的基本原理，发展可解释、可通用的下一代人工智能方法，并推动人工智能方法在科学领域的创新应用。

　　二、核心科学问题

　本重大研究计划针对可解释、可通用的下一代人工智能方法的基础科学问题，围绕以下三个核心科学问题开展研究：

　　（一）深度学习的基本原理

　　深入挖掘深度学习模型对超参数的依赖关系，理解深度学习背后的工作原理，建立深度学习方法的逼近理论、泛化误差分析理论和优化算法的收敛性理论。

　　（二）可解释、可通用的下一代人工智能方法

　　通过规则与学习结合的方式，建立高精度、可解释、可通用且不依赖大量标注数据的人工智能新方法。开发下一代人工智能方法需要的数据库和模型训练平台，完善下一代人工智能方法驱动的基础设施。

　　（三）面向科学领域的下一代人工智能方法的应用

　　发展新物理模型和算法，建设开源科学数据库、知识库、物理模型库和算法库，推动人工智能新方法在解决科学领域复杂问题上的示范性应用。

　　三、2022年度资助研究方向

　　（一）培育项目

　　围绕上述科学问题，以总体科学目标为牵引，2022年度对于探索性强、选题新颖的申请项目，将以培育项目方式予以资助。研究方向如下（申报项目须覆盖以下单一方向中列出的部分或全部内容）：

　　1. 深度学习的表示理论和泛化理论

　　研究多层全联接网络、卷积网络（以及其它带对称性的网络）、图神经网络、transformer网络、循环神经网络等模型的逼近性质，发展相应的高维函数空间理论和泛化误差分析理论，并在实际数据集上检验以上理论。

　　2. 深度学习的训练动力学

　　研究深度学习的损失景观，包括但不限于：临界点的分布及其嵌入结构、极小点的连通性等；深度学习中的非凸优化问题、优化算法的正则化理论和收敛行为；神经网络的过参数化和训练过程对于超参的依赖性问题、基于极大值原理的训练方法、训练时间复杂度和训练困难等问题；循环神经网络记忆灾难问题、编码-解码方法与Mori-Zwanzig方法的关联特性等。

　　3. 微分方程与机器学习方法

　　机器学习算法在微分方程正反问题求解方面的应用，需突破传统数值算法的瓶颈，实现正反问题的高效求解；高维微分方程的正则性理论与算法；微分方程解算子的逼近方法（如通过机器学习方法获得动理学方程、弹性力学方程、流体力学方程、Maxwell方程以及其它常用微分方程的解算子）；微分方程经典算法和基于机器学习方法的融合；微分方程方法在机器学习中的应用（如用微分方程设计新的机器学习模型，设计和分析网络结构等）。

　　4. 数据驱动与知识驱动融合的人工智能

　　建立数据驱动的机器学习与知识驱动的符号计算相融合的新型人工智能理论和方法，突破神经网络模型不可解释的瓶颈；研究知识表示与推理框架、大规模隐式表达的知识获取、多源异构知识融合、知识融入的预训练模型、知识数据双驱动的决策推理等；探索不同场景中的应用。

　　5. 安全可靠的下一代人工智能

　　面向数据、模型和算法，构建安全可靠的人工智能方法。研究分布式去中心化学习、联邦学习、密码学等技术，构建隐私保护的数据应用新范式；研究深度学习模型在对抗样本、数据投毒、后门攻击等情况下的鲁棒性和安全性，发展对抗鲁棒和安全的新模型与学习方法；研究存在样本噪声、分布外数据等场景下的可靠机器学习方法、研究因果驱动的鲁棒决策和可靠推理；探索不同场景中的应用。

　　6. 人工智能驱动的下一代科学计算理论及应用

　　将机器学习与电子多体问题相结合，建立薛定谔方程数值解、第一性原理计算、增强采样、自由能计算、粗粒化分子动力学等的机器学习方法，探索机器学习在物质体系研究中的应用。

　　针对典型的物理、化学、材料、生物、燃烧等领域的多尺度问题和动力学问题，通过融合物理模型与机器学习方法，探索复杂物理、化学、材料、生物等体系变量隐含物理关系的挖掘方法，建立构效关系的数学表达，构建具有通用性的跨尺度人工智能辅助计算理论和方法，解决典型复杂多尺度计算问题。

　　（二）重点支持项目

　　围绕核心科学问题，以总体科学目标为牵引，对于前期研究成果积累较好、对总体目标在理论和关键技术上有较大贡献的申请项目，将以重点支持项目方式予以资助。建议研究内容包括，但不限于以下方向：

　　1. 面向复杂数据的、可通用的人工智能算法框架

　　针对多尺度复杂数据处理问题，研究多尺度表示的跨模态人工智能框架，适用于视频、语音、自然语言、点云、地理数据等不同模态的数据，实现小样本（相同精度下降低样本需求一万倍以上）、可解释、跨模态（不少于3个模态）和感知决策一体化方法。

　　2. 新一代非结构化数据管理方法

　　研究海量复杂非结构化数据与人工智能应用一体化系统的构建方法，包括基础数据存储、用户自定义领域数据模型在线构建、自主研发非结构化数据查询语言与优化理论、跨域和跨库非结构化数据的查询融合理论等，支撑下一代人工智能方法在跨领域、多维度（关系、向量、图等）、多粒度数据（不少于3种粒度）的应用。

　　3. 深度学习隐私保护计算新型体系框架与模型

　　针对隐私性需求，研究可证明安全、可实用的人工智能隐私保护方法新框架，包括但不限于：适用于不同场景的多方隐私计算框架，多源异质数据的高效协同建模方法，基于全同态计算的、低内存占用的隐私保护深度学习方法，研究符合《中华人民共和国个人信息保护法》中匿名化要求的、模型精度跌幅可控（精度下降不大于1%）的数据可信发布技术等。

　　4. 面向功能分析的智能化几何造型方法

　　设计能保持几何约束和物理结构、具有可解释性和收敛阶的智能建模方法，构建可解释的、面向功能分析的智能方法以及科学工程计算中几何模型设计与功能分析一体化方法，突破目前工业软件中几何设计与物理性能分析割裂的现状。

　　5. 人工智能驱动的下一代微观科学计算平台建设与应用

　　发展基于人工智能的高精度、高效率的第一性原理方法；面向物理、化学、材料、生物等领域的实际复杂问题，建立多尺度模型，实现高精度、大尺度和高效率的分子模拟；建立人工智能与科学计算双驱动的“软-硬件协同优化”方法和高性能计算专用平台。

　　6. 人工智能框架下的宏观复杂反应流动多尺度建模与应用

　　面向空天发动机等重大需求场景，针对燃烧模型精度低、数值模拟计算效率低等问题，研究从原子尺度到宏观尺度的深度学习算法，发展兼容传统数值模拟和面向超大规模并行的新一代计算方法；发展航空发动机燃烧不稳定性等关键问题的识别、预测和分析的机器学习方法；针对高雷诺数非稳态超声速燃烧的湍流问题，研究湍流与化学反应的时空多尺度相互作用机理，发展机器学习驱动的高精度湍流模拟模型与计算方法。

　　四、项目遴选的基本原则

　　（一）紧密围绕核心科学问题，鼓励基础性和交叉性的前沿探索，优先支持原创性研究。

　　（二）优先支持面向发展下一代人工智能新方法或能推动人工智能新方法在科学领域应用的研究项目。

　　（三）重点支持项目应具有良好的研究基础和前期积累，对总体科学目标有直接贡献与支撑。

　　五、2022年度资助计划

　　2022年度拟资助培育项目20～25项左右，资助直接费用约为80万元/项，资助期限为3年，培育项目申请书中研究期限应填写“2023年1月1日— 2025年12月31日”；拟资助重点支持项目6～8项左右，资助直接费用约为300万元/项，资助期限为4年，重点支持项目申请书中研究期限应填写“2023年1月1日— 2026年12月31日”。

　　六、申请要求

　　（一）申请条件

　　本重大研究计划项目申请人应当具备以下条件：

　　1. 具有承担基础研究课题的经历；

　　2. 具有高级专业技术职务（职称）。

　　在站博士后研究人员、正在攻读研究生学位以及无工作单位或者所在单位不是依托单位的人员不得作为申请人进行申请。

　　（二）限项申请规定

　　执行《2022年度国家自然科学基金项目指南》“申请规定”中限项申请规定的相关要求。

　　（三）申请注意事项

　　申请人和依托单位应当认真阅读并执行本项目指南、《2022年度国家自然科学基金项目指南》和《关于2022年度国家自然科学基金项目申请与结题等有关事项的通告》中相关要求。

　　1. 本重大研究计划项目实行无纸化申请。申请书提交日期为2022年6月16日—6月23日16时。

　　2. 项目申请书采用在线方式撰写。对申请人具体要求如下：

　　（1）申请人应当按照科学基金网络信息系统中重大研究计划项目的填报说明与撰写提纲要求在线填写和提交电子申请书及附件材料。

　　（2）本重大研究计划旨在紧密围绕核心科学问题，将对多学科相关研究进行战略性的方向引导和优势整合，成为一个项目集群。申请人应根据本重大研究计划拟解决的具体科学问题和项目指南公布的拟资助研究方向，自行拟定项目名称、科学目标、研究内容、技术路线和相应的研究经费等。

　　（3）申请书中的资助类别选择“重大研究计划”，亚类说明选择“培育项目”或“重点支持项目”，附注说明选择“可解释、可通用的下一代人工智能方法”，受理代码选择T01，根据申请的具体研究内容选择不超过5个申请代码。

　　培育项目和重点支持项目的合作研究单位不得超过2个。

　　（4）申请人在“立项依据与研究内容”部分，应当首先说明申请符合本项目指南中的资助研究方向，以及对解决本重大研究计划核心科学问题、实现本重大研究计划科学目标的贡献。

　　如果申请人已经承担与本重大研究计划相关的其他科技计划项目，应当在申请书正文的“研究基础与工作条件”部分论述申请项目与其他相关项目的区别与联系。

　　3. 依托单位应当按照要求完成依托单位承诺、组织申请以及审核申请材料等工作。在2022年6月23日16时前通过信息系统逐项确认提交本单位电子申请书及附件材料，并于6月24日16时前在线提交本单位项目申请清单。

　　4. 其他注意事项

　　（1）为实现重大研究计划总体科学目标和多学科集成，获得资助的项目负责人应当承诺遵守相关数据和资料管理与共享的规定，项目执行过程中应关注与本重大研究计划其他项目之间的相互支撑关系。

　　（2）为加强项目的学术交流，促进项目群的形成和多学科交叉与集成，本重大研究计划将每年举办一次资助项目的年度学术交流会，并将不定期地组织相关领域的学术研讨会。获资助项目负责人有义务参加本重大研究计划指导专家组和管理工作组所组织的上述学术交流活动，并认真开展学术交流。

　　（四）咨询方式

　　国家自然科学基金委员会交叉科学部一处

　　联系电话：010-62328382