附件 1

“网络空间安全治理”重点专项
2025年度项目申报指南

为落实“十四五”期间国家科技创新有关部署安排，组织实施好“网络空间安全治理”重点专项，根据本重点专项实施方案部署，现发布 2025 年度项目申报指南。

项目坚持服务国家需求，坚持应用目标导向，聚焦国家和行业网络安全重大风险，解决国家网络安全重大技术问题。建立责任部门机制，每个项目确定一个责任部门，一般为项目需求部门。

本次启动 6 项指南任务，安排国拨经费 3000 万元左右。申报项目的研究内容必须涵盖指南所列的全部研究内容和考核指标，项目以应用为导向设置考核指标，原则上不设置论文指标。除特殊说明外，每个指南支持项目数为 1 项，实施周期不超过 3 年。

一般项目配套经费与国拨经费比例不低于 1:1，项目下设课题数不超过 5 个，项目参与单位总数不超过 5 家，项目设 1 名项目负责人，项目中每个课题设 1 名课题负责人。

青年科学家项目不再下设课题，项目参与单位总数不超过 2 家，项目骨干人数不超过 5 人。青年科学家项目负责人

应为 1985 年 1 月 1 日以后出生，原则上团队其他参与人员年龄要求同上。

具体项目申报指南如下:

1.气象卫星星地一体化网络安全综合防御体系研究

研究内容: 研究星地一体化测控业务中网络安全防御及数据安全管理的需求，研究指令可信生成、安全高效加密、多路径高可信传输等业务测控安全关键技术，提出适应卫星业务测控流程的可信验证模型和方案框架，构建星地一体化网络安全防御体系；研究气象卫星应用的细粒度动态访问控制、跨域访问实体身份可信认证与角色自适应授权、权限冲突检测与消解、安全策略引擎等智能微边界防护关键技术；研究气象卫星数据分类分级标识与访问策略、授权共享、责权界定、风险识别等数据治理技术，研发气象卫星数据安全管理与风险分析平台，支持数据全流程安全防护与精准管控，全面监测气象卫星数据流转过程安全状态；研制气象卫星星地一体化网络安全综合防御系统，满足气象卫星网络海量吞吐需求；研究成果在国家卫星气象中心网络得到试点应用。

考核指标: 研制可信计算安全软硬组件，支持星地一体化业务测控网络的安全防护；支持星地和星间安全加密通信，具备多路径优先级控制、业务指令的带权识别与分发调度、密钥管理能力，支持 SM2/SM3/SM4 等不少于 3 种国家商用密码算法，星地星间加解密速率不低于 10Mbps；研制

基于动态信任评估的气象卫星应用访问控制机制，支持多元实体跨域可信认证与协同动态授权，支持对网络位置、接入途径、操作系统配置等不少于3种因素的实体鉴权和权限冲突消解；研发数据安全管理与风险分析平台，支持气象卫星数据规模≥100PB，支持符合国家气象数据管理要求的气象卫星数据完整分级分类，具备数据可信标识与治理能力，支持数据治理操作不低于5万TPS（次每秒）。支持数据风险种类识别不少于15类、异常行为识别不少于15种，识别准确率均不低于95%（检测样本数量不低于1万个）。支持基于国家商用密码算法的数据全周期全域流转安全管控，形成全局数据风险视图，统一管理数据安全策略，防范违规、越权、滥用数据行为；研制气象卫星星地一体化网络安全综合防御系统，支持智能网络安全策略管理与防御能力调度，支持安全域灵活动态扩展，涵盖静止和极轨2个系列气象卫星、5个卫星地面站，支持气象信息安全服务用户数≥200万；在国家卫星气象中心网络进行试点应用。

有关说明：项目责任部门为中国气象局。项目承担单位遴选方式为公开竞争。

2.电子产品信息清除关键技术与效果验证研究

研究内容：针对电子产品现有信息清除功能清除不彻底、技术标准不统一导致数据泄露等问题，研究典型电子产品、不同存储介质的信息存储和清除技术，研制强制性国家

标准配套技术文件；研究电子产品存储功能组件快速识别、指令清理多系统融合调用、高可靠数据覆写算法等技术，开发混合多种技术方案的国产化信息清除软件；研究信息清除效果验证的形式化方法，开发涵盖“存储器寻址验证、专业恢复软件测试、清除日志审计”三级验证机制的信息清除效果验证公共服务平台；研究电子产品信息清除标识追溯技术与认证机制，在国内主流电子产品厂商、二手交易平台、资源回收企业内开展技术试点与应用推广。

考核指标：信息清除软件具备存储功能组件快速识别、用户数据定向清除、数据覆写定制化设置等不少于3项功能，适配安卓、鸿蒙、Windows、Linux等不少于4类主流操作系统；数据覆写技术适配95%以上的主流存储介质，与全0/全1覆写方式相比覆写效率提升不少于20%，残留数据恢复概率<0.001%；清除效果验证平台万次重复测试误报率<0.01%，支持接入不少于手机、平板、笔记本电脑、办公设备、服务器、智能穿戴等6类主流电子产品；建立电子产品二手流通信息清除标识追溯与检测认证案例库，案例不少于1万项且覆盖不少于30家主流厂商。

有关说明：项目责任部门为中央网信办。项目承担单位遴选方式为公开竞争。

3.网络安全高质量数据集高效构建关键技术（青年科学家项目）

研究内容：针对当前网络安全领域高质量数据集规模严重不足、数据人工标注效率低、缺乏多模态数据编码理论的问题，研究网络安全领域多模态数据采集技术，包括策略配置、恶意样本、漏洞特征方面的文本、日志、流量等多种模态网络安全数据；研究网络安全数据细粒度高效标注技术，降低人工标注依赖，实现大规模数据自动标注；研究网络安全数据集动态迭代机制，实现新数据样本自动化清洗、去重与融合；研究网络安全多模态数据集质量评估技术；研究网络安全多模态数据编码技术，突破网络安全领域异构数据共性表征理论，为大模型训练与评测提供数据基准支撑。

考核指标：建立策略配置、恶意样本、漏洞特征的高质量多模态网络安全数据集，每种类型数据集规模达到 TB 级，支持网络安全运维、漏洞挖掘等领域；建立网络安全多模态数据细粒度标注工具链，细粒度标签准确率≥85%；构建网络安全多模态数据集质量评估和数据编码理论体系，设计网络安全数据集质量评估方法≥3 个，网络安全数据编码方法≥3 个，网络安全数据编码质量指标≥3 个，立项行业以上相关标准≥2 项，在不少于 3 个典型大模型进行应用验证。

有关说明：项目责任部门为中央网信办。项目面向在中央网信办 2024 年、2025 年组织的人工智能技术赋能网络安全应用测试活动相关场景测试中排名前三的单位遴选牵头承担单位。项目成果、技术、代码按照主责单位要求在国内

相关高校、科研机构、企业等一定范围内开源。

4.网络安全运维垂直大模型构建关键技术（青年科学家项目）

研究内容：针对当前安全威胁检测与异常事件响应高度依赖经验丰富的运维人员、安全运维技术智能化程度弱的问题，研究面向安全威胁感知与响应的安全运维垂直大模型构建与安全运维知识注入技术，设计面向安全运维大模型的高效训练机制，提高对于动态化、隐蔽化新型攻击的应对能力；研究基于自学习的安全运维大模型参数训练技术，持续学习新兴威胁特征，实现对威胁态势的实时感知与智能响应处置生成；研究安全运维应用智能体安全治理技术，对实时威胁监测和自动化安全响应流程中的应用智能体行为进行监控和风险评估；研发面向安全威胁感知与响应的安全运维垂直大模型与应用智能体协同系统，并开展应用示范。

考核指标：建立大规模安全运维多模态数据集，其中警报数据样本不少于 10 万，包含网络流量日志、安全事件日志、策略配置等；构建安全运维多模态大模型，大模型参数不少于 100 亿，可快速处理文本、日志、流量等多模态数据，警报误报率≤5%，警报漏报率≤5%；发现面向 Web 安全、C2 通信等隐蔽攻击类型不低于 5 种；针对不少于 3 种典型网络安全应急响应场景，支持面向 DDoS、C2 通信、后门进程、漏洞探测等攻击方式自动生成网络安全响应策略；支持警报

误报消除准确率不低于90%；支持应用智能体自动生成风险处置策略不低于10种。

有关说明：项目责任部门为中央网信办。项目面向在中央网信办2024年、2025年组织的人工智能技术赋能网络安全应用测试活动相关场景测试中排名前三的单位遴选牵头承担单位。项目成果、技术、代码按照主责单位要求在国内相关高校、科研机构、企业等一定范围内开源。

5.网络安全态势感知大模型构建关键技术（青年科学家项目）

研究内容：针对网络流量规模爆炸式增长及网络攻击行为愈发隐蔽的难题，研究构建网络态势感知大模型，具备原始网络流量检测和分析理解能力，并能够同时完成多项网络态势感知子任务；研究面向网络安全态势感知的高效蒸馏技术，能够大幅提高模型的检测及分析效率；研究网络态势感知大小模型协调推理技术，能够提升多项网络安全态势感知子任务的综合处理能力；研究网络态势感知大模型的上下文关联分析方法，能够监测并识别跨部门长时间周期内隐蔽存在的APT攻击;研究网络态势感知大模型的网络攻击行为预测技术，精准预测网络流量未来发展趋势及潜在攻击手段；研制网络态势感知大模型平台，在典型业务场景下应用示范。

考核指标：构建网络态势感知大模型，大模型的参数量

不低于100亿，覆盖网络态势感知子任务不低于10项，包括恶意软件流量检测、匿名网络行为识别、僵尸网络检测、VPN行为检测等；在至少跨2个部门3个月时间以上周期的真实网络数据集中，APT攻击识别率不低于80%；网络态势感知大模型对潜在网络攻击行为的发现准确率不低于80%；实现面向流量检测等子任务的蒸馏小模型，小模型的参数量不高于15亿，子任务的识别准确率不低于90%；在交通、能源等不少于2种关键信息基础设施场景下对网络态势感知大模型应用示范。

有关说明：项目责任部门为中央网信办。项目面向在中央网信办2024年、2025年组织的人工智能技术赋能网络安全应用测试活动相关场景测试中排名前三的单位遴选牵头承担单位。项目成果、技术、代码按照主责单位要求在国内相关高校、科研机构、企业等一定范围内开源。

6.面向人工智能模型安全的密码分析与防护关键技术（青年科学家项目）

研究内容：针对目前人工智能模型安全存在的依赖经验性防御，缺乏数学可解释性和可验证性问题，探索密码技术在人工智能安全领域的应用，基于密码分析技术评估人工智能模型在机密性、鲁棒性、溯源性等方面的安全性风险，并提出针对性的防护技术；研究基于新型密码分析的模型参数高精度恢复攻击技术，提出模型参数恢复攻击新型防御机

制；研究基于密码分析理论的高效黑盒对抗攻击样本生成方法，提出对抗攻击新型防御机制；研究可认证高稳定性大模型生成内容水印技术，提出可证明安全密码学原语；研发基于密码分析的人工智能模型机密性、鲁棒性与溯源性安全评估与防护工具。

研究指标：提出基于密码分析的新型人工智能模型安全评估与防护理论体系：针对不少于 5 种典型人工智能模型，提出基于密码分析的模型参数高精度恢复攻击方法，要求参数最大恢复误差≤0.001（不依赖侧信道信号），并提出针对性的防御方法；提出不少于 3 种基于密码分析的对抗样本攻击方法，相比已有黑盒对抗样本攻击方法成功率相对提高不低于 50%，并提出针对性的防御方法，相比已有黑盒对抗样本防御方法成功率相对提高不低于 20%，并在不少于 5 种典型人工智能模型上进行验证；提出不少于 3 种基于新型密码原语设计可证明安全的模型参数水印以及生成内容溯源方法，在图像压缩、噪声扰动、图像编辑等典型攻击场景下比特恢复率不低于 98%，并在不少于 5 种典型人工智能模型上进行验证。开发基于密码技术的人工智能模型安全评估与防护平台，支持不少于 3 种常用开发框架。

有关说明：项目责任部门为国家密码管理局。项目承担单位遴选方式为公开竞争。项目成果、技术、代码按照主责单位要求在国内相关高校、科研机构、企业等一定范围内开源。