



CCSC
2026

Hefei, China
2026.09
中国·合肥

智启新程 安领未来

中国网络空间安全大会 (CCSC2026)

CHINA CYBER SECURITY CONGRESS (CSC2026)

主办单位：中国网络空间安全学会

承办单位：中国科学技术大学

大会程序册





目 录 CONTENTS

○ 组织架构·····	01
○ 大会指南·····	06
○ 大会日程总览·····	10
○ 大会主旨报告·····	12
○ 大会专题论坛·····	22
○ 大会合作单位·····	229

○ 9月18日下午

网络与通信安全论坛·····	22
智能时代下的网络韧性技术与方法论坛·····	28
生成式多媒体内容安全与可信论坛·····	33
新型电力系统信息物理综合安全风险与应对论坛·····	38
开源软件与供应链安全论坛·····	42
网络空间多域融合安全体系论坛·····	47
智能赋能的系统工程安全论坛·····	53
融合与重构：AI时代的网络空间安全人才培养、学科建设与工程思维创新论坛·····	58
智能密码论坛·····	64
空天地一体化安全论坛·····	69
具身智能安全论坛·····	74
感知AI安全：从数字域到物理域论坛·····	79
区块链与可信智能体网络论坛·····	84

○ 9月19日下午

网密融合论坛·····89

网络仿真与安全评测论坛·····94

智能装备安全论坛·····100

“后大模型时代”的软件与系统安全论坛·····104

人工智能安全与应用治理论坛·····109

量子计算与网络安全论坛·····114

数据流通利用的隐私计算、合规与治理论坛·····119

智安融合、双向赋能 | 协同构建自主网络攻防大模型新生态论坛·····124

网络空间安全科技女性发展论坛·····130

工业互联网安全论坛（2030项目首席论坛）·····137

信息隐藏技术创新论坛·····142

认知安全创新论坛·····148

国家一流网安学院院长论坛·····153

卫星互联网安全论坛·····158

○ 9月20日下午

网络空间安全基础研究论坛·····162

生成式AI驱动的内容安全治理论坛·····167

网络空间态势感知论坛·····172

网络空间测绘技术论坛·····177

智能体安全论坛·····181

可信智能体与安全治理论坛·····187

低空智联网安全论坛·····194

CCSC2026青年学者论坛·····199

CCSC2026博士生论坛·····202

抗量子密码迁移的“道法术器”论坛·····211

智能无人系统安全论坛·····216

网安期刊创新发展论坛·····221

组织架构

指导委员会 - 主席



冯登国

中国科学院院士
中国网络空间安全学会理事长



周志鑫

中国科学院院士

指导委员会 - 委员



方滨兴

中国工程院院士



郭世泽

中国科学院院士



管晓宏

中国科学院院士



黄殿中

中国工程院院士



孔志印

中国工程院院士



陆军

中国工程院院士



王建华

中国工程院院士



王小云

中国科学院院士
中国网络空间安全学会副理事长



邬江兴

中国工程院院士



吴建平

中国工程院院士
中国网络空间安全学会副理事长



吴世忠

中国工程院院士



尹浩

中国科学院院士



云晓春

中国工程院院士



张宝东

中国工程院院士



郑建华

中国科学院院士



程静

中国网络空间安全学会监事长



孟丹

中国网络空间安全学会副理事长



俞能海

中国网络空间安全学会副理事长

大会主席



吴建平

中国工程院院士
中关村实验室主任



常进

中国科学院院士
中国科学技术大学校长

程序委员会 - 主席



王小云

中国科学院院士
中国网络安全安全学会副理事长

程序委员会 - 副主席



王美琴

山东大学

程序委员会 - 委员



蔡红

中国电科网络空间研究院



傅首清

中关村实验室



郭冠斌

中电科数字科技
(集团) 有限公司



郭莉

中国科学院
信息工程研究所



何良生

合肥国家实验室



贾焰

鹏城实验室



季新生

国家数字交换系统工程
技术研究中心



荆继武

中国科学院大学



李凤华

中国科学院
信息工程研究所



李晖

西安电子科技大学



李建华

上海交通大学



刘建伟

北京航空航天大学



马建峰

西安电子科技大学



秦玉海

中国刑事警察学院



徐茂智

北京大学



张小松

电子科技大学



祝跃飞

信息工程大学

组织委员会 - 主席



俞能海

中国网络安全安全
学会副理事长



张卫明

中国网络安全安全
学会秘书长



陆余良

国防科技大学



王佐成

合肥综合性国家科学中心
数据空间研究院

论坛委员会 - 主席**谷大武**
上海交通大学**张勇东**
中国科学技术大学**李小勇**
北京邮电大学**任奎**
浙江大学**赞助工作组 - 主席****邹德清**
华中科技大学**沈玉龙**
西安电子科技大学**钱振兴**
复旦大学**吴云坤**
奇安信科技集团
股份有限公司**展览工作组 - 主席****祝烈煌**
北京理工大学**侯锐**
中国科学院
信息工程研究所**王震**
西北工业大学**罗洪斌**
北京航空航天大学**海报工作组 - 主席****程光**
东南大学**田志宏**
广州大学**黄欣沂**
南京航空航天大学**王伟**
西安交通大学

宣传工作组 - 主席**姜伟**

哈尔滨工业大学

**李洪伟**

电子科技大学

**苏洲**

西安交通大学

**位华**

中国信息安全杂志社

青年工作组 - 主席**刘哲理**

南开大学

**王骞**

武汉大学

**沈超**

西安交通大学

**唐卓**

湖南大学

女科技工作者工作组 - 主席**李雪莹**天融信科技集团
股份有限公司**卿昱**

中关村实验室

**张宏莉**

哈尔滨工业大学

**王鹃**

武汉大学

学科教育工作组 - 主席**徐明伟**

清华大学

**刘宝旭**中国科学院信息工程
研究所**程学旗**中国科学院
计算技术研究所**饶志宏**

中国电科总体院（特区）

网站工作组 - 主席**廖勇**

中国电科网络空间研究院

**段海新**

清华大学

**杨珉**

复旦大学

**吴文涛**

中国科学技术大学

注册工作组 - 主席**何召锋**

北京邮电大学

**单征**

信息工程大学

**耿光刚**

暨南大学

**关振宇**

北京航空航天大学

本地工作组 - 主席**林璟铨**

中国科学技术大学

**胡东辉**

合肥工业大学

**施凡**

国防科技大学

**崔杰**

安徽大学

财务工作组 - 主席**苏璞睿**

中国科学院软件研究所

**魏建国**

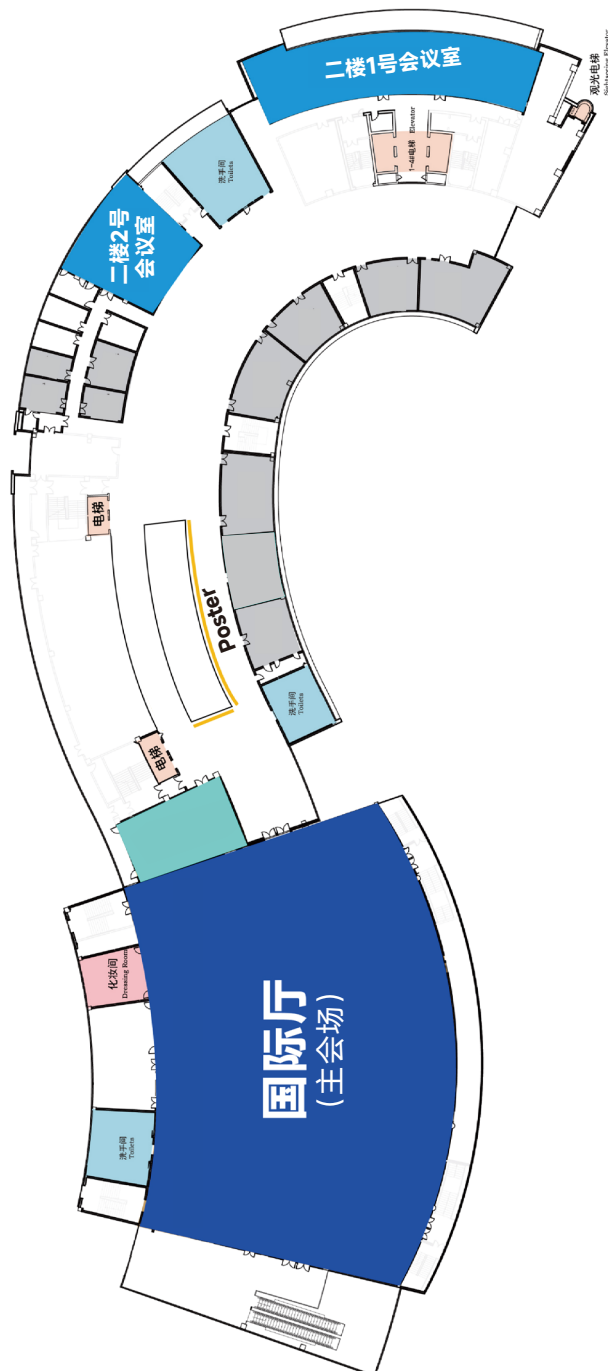
天津大学

**李佳**中国航天系统科学与工程
研究院**刘畅**

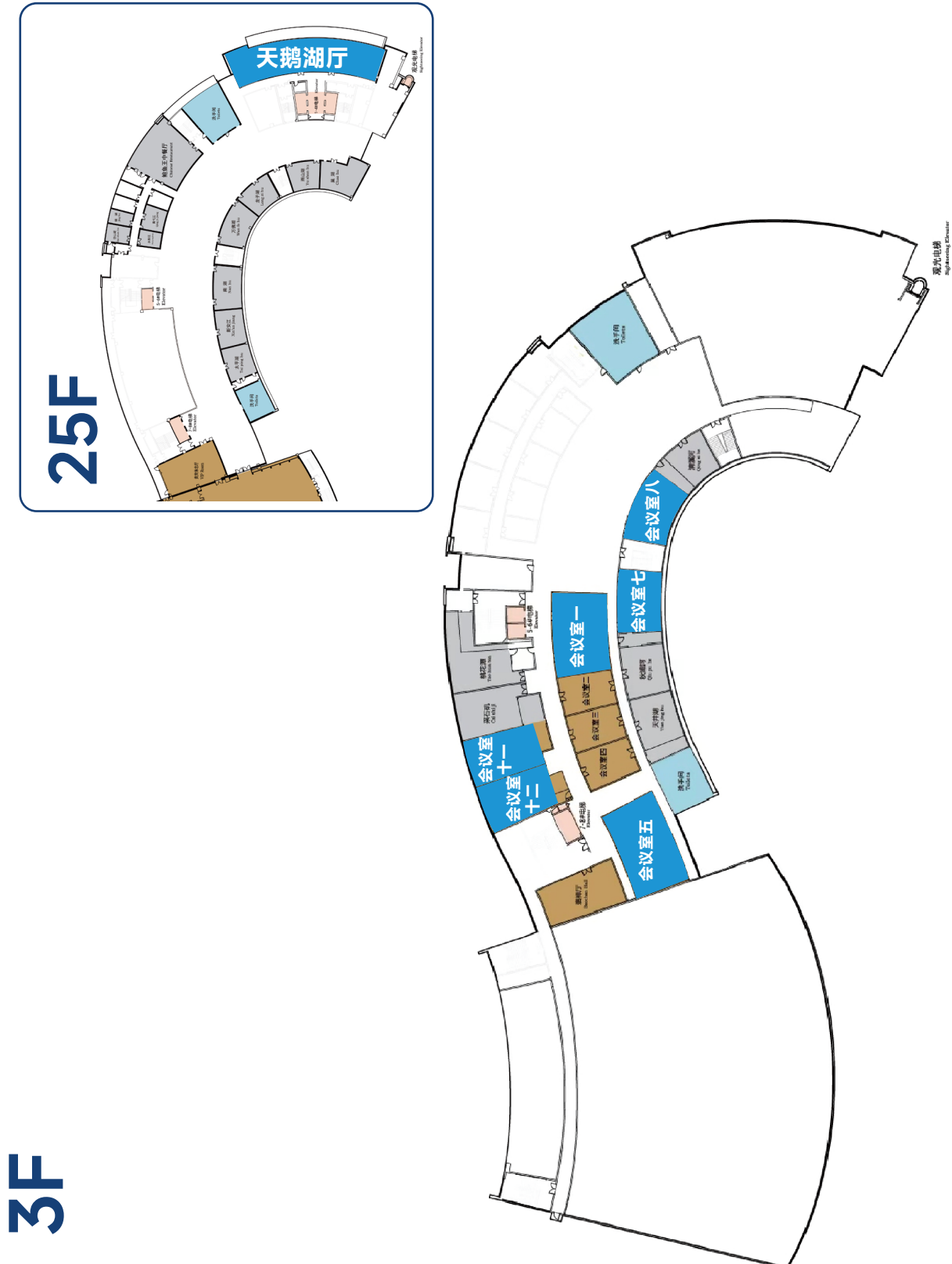
中国科学技术大学

酒店平面图

2F



酒店平面图



■ 温馨提示

凭当日餐券在指定餐厅用餐

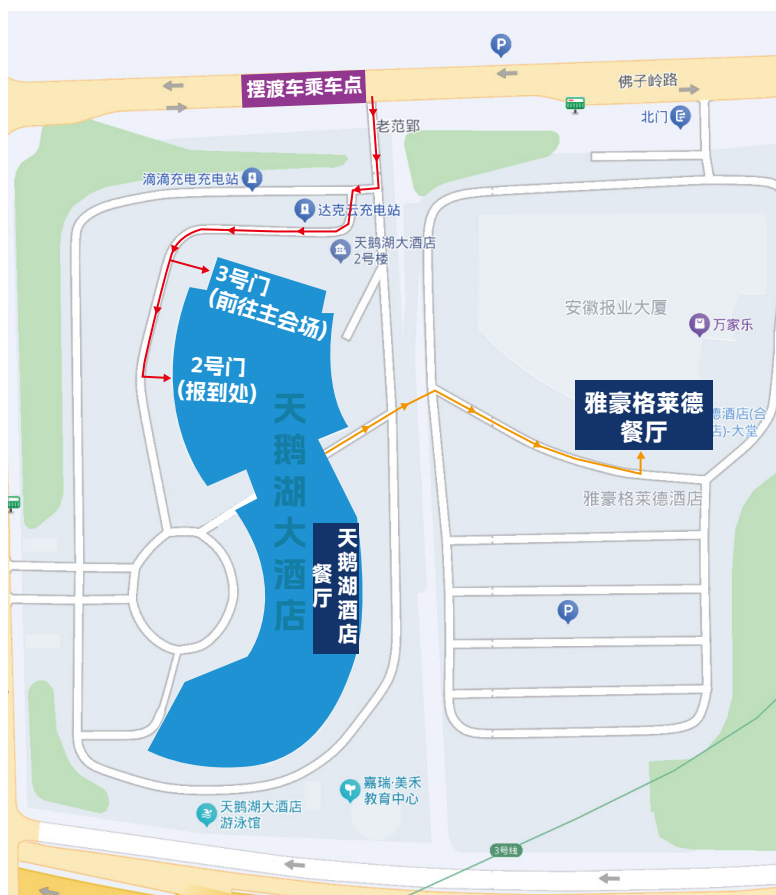
为保证会议顺利进行，请您注意以下事项：

1. 请参会代表按照会议的日程安排,准时参加各项活动,并注意会务组临时通知;
2. 会议期间发言与交流, 请遵守国家有关保密规定, 勿涉及国家秘密;
3. 请在完成签到后佩戴胸牌, 它是您参会的唯一有效证明, 请妥善保管, 遗失不补;
4. 请自觉维护会场秩序,服从大会安排,会议期间如无特殊情况,请勿在会场随意走动;
5. 自觉维护环境整洁,会场内请勿吸烟;
6. 会议期间请将通讯设备关闭或静音状态,如需接听电话请移步至场外;
7. 请妥善保管好个人财务及重要文件等贵重物品,谨防丢失。

8. 会议联络人：

- 会务：吴 维 (19056872025)
张 梅 (15255187881)
- 酒店：王传燕 (13665691634)

■ 摆渡车信息



车次信息

9月18日

天鹅湖大酒店 → 各协议酒店
10:00至21:00 (30分钟/班, 坐满即发车)

9月19日

各协议酒店 → 天鹅湖大酒店
早：07:50至08:15, 坐满即发车
天鹅湖大酒店 → 各协议酒店
晚：19:30至20:30, 坐满即发车

9月20日

各协议酒店 → 天鹅湖大酒店
早：07:50至08:15, 坐满即发车
天鹅湖大酒店 → 各协议酒店
晚：17:30-19:00

注：各协议酒店名称及摆渡车协调经理联系方式

两淮天鹅湖 (宋经理 15955101063)
泓瑞金陵 (汪经理 13965087761)
盛捷天鹅湖 (刘经理 18156537210)
合肥豪生 (梁经理 18855187405)
贝斯特酒店 (王经理 18134503361)

大会日程总览

◀ 2026年9月18日 ▶

时间	内容	地点
10:00-21:00	报 到	天鹅湖大酒店 大堂、2号门

时间	内容	地点
14:00-17:55	网络与通信安全论坛	二楼2号会议室
14:00-17:50	智能时代下的网络韧性技术与方法论坛	三楼5号会议室
14:00-17:15	生成式多媒体内容安全与可信论坛	一楼长江厅
14:00-17:50	新型电力系统信息物理综合安全风险与应对论坛	一楼贵宾厅
14:00-17:50	开源软件与供应链安全论坛	三楼7号会议室
14:00-17:25	网络空间多域融合安全体系论坛	一楼淮河厅B
14:00-17:30	智能赋能的系统工程安全论坛	三楼12号会议室
14:00-18:05	融合与重构：AI时代的网络空间安全人才培养、学科建设与工程思维创新论坛	二楼1号会议室
14:00-17:30	智能密码论坛	一楼徽州厅
14:00-18:00	空天地一体化安全论坛	三楼1号会议室
14:00-17:30	具身智能安全论坛	三楼11号会议室
14:00-17:40	感知AI安全：从数字域到物理域论坛	一楼黄山厅
14:00-18:15	区块链与可信智能体网络论坛	一楼淮河厅A
14:30-18:00	智耀江淮·筑安关基— AI赋能重点行业和关基保护交流专题报告会	合肥市工投高新 智谷B1栋 三楼报告厅

◀ 2026年9月19日 ▶

时间	内容	地点
09:00-09:55	大会开幕式	二楼国际厅
09:55-12:00	主旨报告	

时间	内容	地点
13:00-17:50	网密融合论坛	三楼12号会议室
13:30-17:10	网络仿真与安全评测论坛	三楼5号会议室
13:45-17:20	智能装备安全论坛	一楼淮河厅A

时间	内容	地点
14:00-17:20	“后大模型时代”的软件与系统安全论坛	三楼7号会议室
14:00-17:30	人工智能安全与应用治理论坛	市政务会议中心8号会议室
14:00-17:40	量子计算与网络安全论坛	三楼11号会议室
14:00-17:10	数据流通利用的隐私计算、合规与治理论坛	三楼8号会议室
14:00-17:20	智安融合、双向赋能 协同构建自主网络攻防大模型新生态论坛	三楼1号会议室
14:00-17:30	网络空间安全科技女性发展论坛	一楼淮河厅B
14:00-17:00	工业互联网安全论坛（2030项目首席论坛）	二楼1号会议室
14:00-18:00	信息隐藏技术创新论坛	一楼徽州厅
14:00-17:35	认知安全创新论坛	一楼长江厅
14:00-17:40	国家一流网安学院院长论坛	一楼贵宾厅
14:30-17:10	卫星互联网安全论坛	二楼2号会议室
15:30-17:30	中国网络安全安全学会第一届理事会第二次理事会议	一楼黄山厅

◀ 2026年9月20日 ▶

时间	内容	地点
09:00-11:05	主旨报告	二楼国际厅
11:05-12:00	圆桌论坛	

时间	内容	地点
14:00-17:30	网络空间安全基础研究论坛	三楼12号会议室
14:00-17:30	生成式AI驱动的内容安全治理论坛	二楼2号会议室
14:00-17:20	网络空间态势感知论坛	一楼淮河厅A
13:30-16:30	网络空间测绘技术论坛	三楼7号会议室
14:00-17:30	智能体安全论坛	三楼1号会议室
14:00-17:25	可信智能体与安全治理论坛	一楼徽州厅
14:00-17:00	低空智联网安全论坛	三楼5号会议室
14:00-17:30	CCSC2026 青年学者论坛	一楼黄山厅
14:00-18:00	CCSC2026 博士生论坛	一楼淮河厅B
14:00-18:00	抗量子密码迁移的“道法术器”论坛	一楼贵宾厅
14:00-18:10	智能无人系统安全论坛	三楼11号会议室
14:00-17:40	网安期刊创新发展论坛	三楼8号会议室
14:00-17:30	新形势下CCSS分支机构规范化建设专题研讨会（闭门会议）	二楼1号会议室

大会主论坛

开幕式与主旨报告

🕒 9月19日上午

📍 天鹅湖大酒店二楼国际厅

时 间	内 容	主持人
大会开幕式		
09:00-09:50	开幕式	冯登国 中国网络空间安全学会理事长
09:50-09:55	合 影	
主旨报告		
09:55-10:25	极限理论:解译几个热点大模型基础问题 徐宗本（中国科学院院士）	冯登国 中国网络空间安全学会 理事长
10:25-10:55	密码定义安全——思考与愿景 郑建华（中国科学院院士）	
10:55-11:00	休 息	
11:00-11:30	AI赋能网络空间安全发展问题思考 尹 浩（中国科学院院士）	王小云 中国网络空间安全学会 副理事长
11:30-12:00	生成式人工智能红队测试的技术与实践 吴世忠（中国工程院院士）	



徐宗本 中国科学院院士

报告题目：极限理论:解译几个热点大模型基础问题

个人简介:

中国科学院院士，数学家、信号与信息处理专家、西安交通大学教授。主要从事智能信息处理、机器学习、数据建模基础理论研究。曾提出稀疏信息处理的 $L(1/2)$ 正则化理论,为稀疏微波成像提供了重要基础;发现并证明机器学习的“徐-罗奇”定理,解决了神经网络与模拟演化计算中的一些困难问题,为非欧氏框架下机器学习与非线性分析提供了普遍的数量推演准则;提出基于视觉认知的数据建模新原理与新方法,形成了聚类分析、判别分析、隐变量分析等系列数据挖掘核心算法,并广泛应用于科学与工程领域。曾获国家自然科学二等奖、国家科技进步二等奖、陕西省最高科技奖;国际IAITQM 理查德·普莱斯(Richard Price)数据科学奖;中国陈嘉庚信息技术科学奖、华罗庚数学奖、苏步青应用数学奖、吴文俊人工智能最高科技成就奖;曾在2010年世界数学家大会上作45分钟特邀报告。曾任西安交通大学副校长。现任鹏城国家实验室广州基地/琶洲实验室(黄埔)主任、陕西国家应用数学中心主任、大数据算法与分析技术国家工程实验室主任。是国家大数据专家咨询委员会委员、国家新一代人工智能战略咨询委员会委员。

报告摘要:

以大模型为特征的生成式人工智能近年来迅猛发展,正深刻影响并变革着科学技术的研究范式与工业革命进程。按照流行的认识,大模型的能力之所以强大源自它可能存在的智能涌现(Intelligent Emergence)。然而,什么是智能涌现?是什么要素催生了智能涌现?大模型在什么情况下才会出现智能涌现?大模型为什么能工作?对这些问题充满着疑惑和争论。本报告基于“智能涌现源于知识从有限到无限所引起的认知改变”观点,提出一个数学框架和数学理论来对此加以解析。我们的核心思想是:利用三元随机函数 $\mathbb{X}(N, P, K)$ 来度量大模型性能,利用 $\mathbb{X}(N, P, K)$ 的极限行为/极限速度($N \rightarrow \infty, P \rightarrow \infty, K \rightarrow \infty$)来度量大模型的尺度变化律(Scaling Law),并以此为基础来解译智能涌现和尺度律,这里 N 是用以训练大模型的数据规模, P 是模型尺寸(包含参数个数), K 是训练步数。我们定义“极限架构”这一无穷维系统概念,说明大模型智能涌现存在的必要条件是极限架构存在,而极限架构存在的必要充分条件是构成大模型的基块算子 T_i 满足:(i)Lipschitz特征 $\text{Lip}(T_i) \leq 1$ 和 (ii) $\{T_i\}$ 具有凝聚性质。一旦极限架构存在,我们说明,极限架构便承载了大模型的涌现智能,换言之,涌现后所表现出的新特征和新行为(由 $\mathbb{X}(\infty, \infty, \infty)$ 刻画。我们证明,作为无穷维系统的极限架构,它由一个有限维系统与一个快速衰减的无穷维系统之和构成,因而应用中可由有限维系统一致逼近。这是大模型可用的理论依据。我们也提出 $\mathbb{X}(N, P, \delta\ell) - \mathbb{X}(\infty, \infty, 0)$ 的标准误差分解,并应用非线性Lipschitz算子工具,无限维Bayes估计工具来对这些误差进行估计。最终,我们获得大模型的尺度律估计。所获得理论揭示:(1)大模型是否出现智能涌现依赖于数据、训练和架构多重要素,但架构是核心,而架构完全由大模型的功能块 $\{T_i\}$ 特性决定;(2)大模型智能涌现只会出现在临界动力学环境 $\text{Lip}(T_i)=1$,这提供了临界动力学原理成立的理论证据;(3)大模型是靠有限维系统工作的,从而理论上支持“Universal Weight Subspace Hypothesis”(Kaushik et al, 2025)。将所建立的理论用于ChatGPT大模型,我们验证:ChatGPT-2及以后改进模型会出现智能涌现/尺度律,但ChatGPT-1不能。

**郑建华** 中国科学院院士**报告题目：密码定义安全——思考与愿景****个人简介：**

中国科学院院士，信息工程大学研究员。1956年9月出生于吉林省长春市，籍贯浙江宁波。1987年毕业于中国科学技术大学研究生院。1992年被评为国家有突出贡献的中青年专家、享受政府特殊津贴。长期从事复杂信息系统分析和相关基础理论研究，对该领域的序列论、函数论、算法设计与分析等进行了系统研究，在复杂信息系统输出分析技术、系统模型解析理论和方法、系统参数还原技术研究中均取得创新性研究成果，这些研究成果在实际复杂系统分析中多次发挥显著作用。曾获国家科技进步一等奖3项等。

报告摘要：

随着万物互联的智能化时代到来，安全威胁逐步由虚拟空间向现实空间蔓延，关系着能源、军事、金融等重要领域的网络安全进入“大安全”时代。云计算、人工智能等新技术加速推进数字化和现代化的同时，也带来了网络安全新风险。然而，被动地查杀封堵显然难以应对层出不穷的新风险，亟需建立主动安全防护的网络安全新理念。密码定义安全基于可证明安全的思想，将密码技术“紧”地作用在保护对象上。从网络行为、操作和对象三个层面出发，密码定义安全被划分密码定义网络、密码定义计算、密码定义数据三个层次，力求实现行为受控、计算一致和数据保密。网络安全的本质是对抗，通过密码定义安全可将网络对抗规约到密码对抗，即网络攻击能否得手仅取决于敌手能否突破密码防护。随着该理论的完备性、体系性和兼容性不断优化，网络安全技术和密码技术同步发力、协同创新，最终形成可证明安全的主动安全防护能力体系。

**尹 浩** 中国科学院院士**报告题目：AI赋能网络空间安全发展问题思考****个人简介：**

通信网络领域专家，复杂系统仿真总体重点实验室主任，中国科学院院士，博士生导师。长期从事通信网络理论方法研究、体系结构设计和技术应用等方面研究，主持完成多项国家重点科研与工程建设项目，获国家和省部级科技进步奖多项。目前主要研究方向为物联网和卫星通信技术。

报告摘要：

人工智能赋能网络空间安全是数字时代的必然趋势，技术创新为网络安全防御体系提质增效的同时，也带来了攻防失衡、模型风险、数据泄露、治理滞后、产业短板、运维失控等一系列复合型挑战。报告简要介绍了AI赋能网络空间安全的赋能机遇与风险反噬双面性；分析了AI赋能网络空间安全面临的六位一体复合型风险挑战；讨论了智能时代安全问题的深层成因——四维结构性矛盾；提出了基于五位一体综合治理体系的应对措施与发展路径。

**吴世忠** 中国工程院院士**报告题目：生成式人工智能红队攻击技术演进与实践****个人简介：**

中国工程院院士，中国人民政治协商会议第十三届、第十四届全国委员会委员，中关村实验室首席科学家。长期从事网络与信息安全领域的技术研究和管理工作，科研成果先后获得国家科技进步一等奖两项，二等奖一项，省部级科技奖多项。

报告摘要：

随着前沿大模型能力快速演进，AI红队测试已成为人工智能安全的重要环节。本报告围绕生成式人工智能红队测试技术与实践，介绍红队测试的现实需求与发展现状，当前的主要技术与国内的前沿探索，以及红队的组织管理。分析红队测试由单一内容安全向多维可信评估、由模型测试向复杂智能系统评估、由人工智能系统安全向“AI+网络安全”融合测试拓展，以及由现象评测向安全机理研究演进的发展趋势。

主旨报告

☉ 9月20日上午

☉ 天鹅湖大酒店二楼国际厅

时 间	内 容	主持人
主旨报告		
09:00-09:30	处理器安全技术的挑战 孙凝晖（中国工程院院士）	孟 丹 中国网络空间安全学会 副理事长
09:30-10:00	从磁石到灯塔：用东方智慧锚定人工智能的发展方向 黄殿中（中国工程院院士）	
10:00-10:05	休 息	
10:05-10:35	中国网络安全技术弯道超车战略研究 云晓春（中国工程院院士）	俞能海 中国网络空间安全学会 副理事长
10:35-11:05	量子网络现状与展望 潘建伟（中国科学院院士）	
11:05-12:00	圆桌论坛：人工智能时代的网络空间安全	方滨兴 中国工程院院士

**孙凝辉** 中国工程院院士**报告题目：处理器安全技术的挑战****个人简介：**

中国工程院院士，中国科学院计算技术研究所研究员，中关村实验室首席科学家。曾担任中国科学院计算技术研究所所长。计算机系统结构专家，主要从事高性能计算机研究。获国家杰出青年基金，中国科学院杰出科技成就奖，国家科技进步二等奖四次，参与获得国家科技进步一等奖一次、国家自然科学基金二等奖一次。获一项国家专利金奖。

报告摘要：

报告围绕处理器安全主题，从漏洞挖掘、检测修复、设计安全、安全指令集、等级认证及智能体新风险六个维度展开。通过典型安全事件揭示处理器漏洞的普遍性与挖掘方法演进，梳理“发现—检测—修复”三阶段技术体系；针对安全生态碎片化困境，提出安全指令集标准化路径和实践；最后剖析智能体时代安全本质从“防代码漏洞”转向“防意图滥用”的范式转变。报告旨在为构建自主可控的处理器安全体系提供参考。

**黄殿中** 中国工程院院士**报告题目：从磁石到灯塔：
用东方智慧锚定人工智能的发展方向****个人简介：**

中国工程院信息与电子学部和管理学部院士，是通信与信息安全和工程管理专家；组织攻坚完成国家级重大专项工程，取得历史性突破，攻克多项国家重点研发计划专项，填补空白，荣获国家科技进步一等奖2项，国家技术发明二等奖1项。曾任国家某部领导、部科技委主任，第十七届、第十八中央纪律检查委员会委员；现任某部科技战略咨询委委员，中国信息安全测评中心学术委员会副主任，赛博空间全国重点实验室学术委员会主任。

报告摘要：

报告直面人工智能狂飙突进背后的深层伦理危机，指出西方文明囿于“二元对立”与“生存博弈”的思维定式，无论是马斯克的“物理同化”还是辛顿的“生物本能”，均无法从根本上确保超级智能对人类的长久善意。报告创新性地提出，唯有东方智慧能为冰冷的算法注入温热的灵魂。以《易经》“阴阳平衡”与“和合共生”为哲学根基，将碳基生命与硅基智能的关系，从“征服与取代”重构为“阴阳共振、一体共生”。主张以“敬天法祖”的敬畏之心筑牢理性压舱石，以“推己及人”的仁心确立“碳硅永恒契约”，从而构建“以人为本、智能向善”的碳硅命运共同体。这不仅是为AI发展校准航向的中国方案，更是引领人类文明跨越技术迷雾、走向共生共荣的东方灯塔。

**云晓春** 中国工程院院士**报告题目：中国网络安全技术弯道超车战略研究****个人简介：**

中国工程院院士，网络空间安全技术专家，现任中关村实验室首席科学家。作为第一完成人获得国家科技进步一等奖1项、二等奖2项，第一发明人获授权发明专利45项，主持制定国家及通信行业标准15项，出版专著2部，发表学术论文100余篇，培养博士研究生30余人。他长期致力于国家网络空间安全的技术研究和工程实践，主持研制了国家网络安全系列平台，主导建立了国家网络安全应急处置技术体系，为有效应对全球性网络安全威胁做出了重要贡献。

报告摘要：

系统阐述美国AI安全能力的阶跃式突破与中国的战略机遇窗口，对中国网络安全产业现状与基础能力进行全景分析，提出中国网络安全技术弯道超车的三大路径，探讨"以智御智"技术体系和非对称超车的五大战略。

**潘建伟** 中国科学院院士**报告题目：量子网络现状与展望****个人简介：**

中国科学技术大学常务副校长、教授，中科院量子信息与量子科技创新研究院院长，中国科学院院士，英国皇家学会外籍院士，发展中国家科学院院士。主要从事量子光学、量子信息和量子力学基础问题检验等方面的研究。作为国际上量子信息实验研究领域的开拓者之一，他是该领域有重要国际影响力的科学家。利用量子光学手段，他在量子调控领域取得了一系列有重要意义的研究成果，尤其是他关于量子通信、量子计算和多光子纠缠操纵的系统性创新工作使得量子信息实验研究成为近年来物理学发展最迅速的方向之一。

报告摘要：

自上世纪90年代以来，随着量子调控技术的发展，人们掌握对光子、原子等量子系统进行主动操纵的能力，从而催生了新兴的量子信息科技。量子信息科技的发展目标是构建“量子网络”，其主要内容包括：通过量子精密测量（量子传感）实现信息的高精度感知、通过量子通信实现信息的安全高效传输、通过量子计算实现信息的高速处理。特别是，量子通信为构建安全、高效的网络空间提供了全新的方案。在信息安全方面，量子计算的快速发展对经典密码体系带来了巨大冲击，而基于量子密钥分发的量子保密通信提供了“信息论可证”的安全通信方式，可从根本上抵御量子计算的攻击。将量子保密通信与抗量子计算的数学密码相结合，可为向抗量子计算的密码体系迁移提供有效可行的路径。在信息传输方面，量子隐形传态可实现量子网络中量子计算和量子传感终端的高效连接。未来10至15年，将构建起天地一体的高效率广域量子网络，为保障信息传输安全、提高计算速度、提升测量精度等国家重大需求做出实质性贡献。

大会专题论坛

网络与通信安全论坛

承办单位：中国科学技术大学、清华大学

🕒 9月18日下午

📍 天鹅湖大酒店二楼2号会议室

论坛主席：段海新（清华大学）

林璟铨（中国科学技术大学）

召 集 人：林璟铨（中国科学技术大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	网络基础设施安全挑战与应对思考 苏金树（军事科学院）
14:40-15:10	大模型赋能无线物理层通信：从安全机制到智能决策 江 涛（华中科技大学）
15:10-15:40	卫星量子密钥分发实验及应用 廖胜凯（中国科学技术大学）
15:40-16:10	物理AI及其安全思考 冀晓宇（浙江大学）
16:10-16:25	休 息
16:25-16:55	智能化网络威胁防御与危害消解技术探索及实践 李书豪（中关村实验室）
16:55-17:25	网络系统多方交互安全研究 陈建军（清华大学）
17:25-17:55	大语言模型赋能移动软件安全分析 刁文瑞（山东大学）

论坛主席



段海新

博士，教授，博士生导师。就职于清华大学网络科学与网络空间研究院，网络与信息安全实验室主任，加州大学伯克利访问学者，国际信息系统安全认证专家（CISSP），清华大学-奇安信集团联合研究中心主任。长期从事计算机网络安全相关的教学、科研和运行管理工作，在清华大学计算机系承担本科和研究生两门网络安全相关课程，承担国家973、863、自然科学基金等多项科研课题，中国教育和科研计算机网应急响应组（CCERT）负责人。



林璟铨

博士，教授、博士生导师，现任中国科学技术大学网络空间学院执行院长，入选国家“万人计划”科技创新领军人才，长期从事网络空间安全研究、尤其是密码技术在计算机和网络系统的创新应用。密码工程技术研究成果已广泛应用于我国各种电子签名电子合同、电子证照、视频加解密等系统。

召集人



林璟铨

博士，教授、博士生导师，现任中国科学技术大学网络空间学院执行院长，入选国家“万人计划”科技创新领军人才，长期从事网络空间安全研究、尤其是密码技术在计算机和网络系统的创新应用。密码工程技术研究成果已广泛应用于我国各种电子签名电子合同、电子证照、视频加解密等系统。

报告嘉宾



苏金树

苏金树，军事科学院研究员，中国计算机学会会士。获军队杰出专业技术人才奖，CCF王选奖，是我国计算机网络学术带头人之一，建立了高性能高安全网络的理论方法、核心技术和系统装备的技术体系，牵头研制成功了多款代表国家水平的高性能路由器，网络安全设备和新型网络系统等，在国家国防领域取得规模化部署，推动了我国高性能网络技术、型号装备发展和行业进步，为提高我国网络基础设施建设水平、自主可控程度、网络空间安全做出系统性和创造性贡献。获国家科技进步特等奖1项，国家科技进步一等奖1项，国家科技进步二等奖4项(1、1、2、2)。在国内外高水平刊物会议发表论文200多篇。

报告题目：网络基础设施安全挑战与应对思考

报告摘要：无论是PC互联网时代，移动互联网时代，还是智能体互联网时代，网络基础设施的重要作用不可替代。本报告主要阐述网络基础设施面临的主要风险，存在的典型漏洞，如何发现存在的威胁，如何增强网络基础设施关键设备和协议软件的安全性设计，网络基础设施发展新趋势。



江 涛

江涛，IEEE Fellow，华中科技大学二级教授，是移动信息网络国家重大科技专项咨询委员会专家、国家6G移动通信技术研发总体专家组成员，获国家杰出青年科学基金项目资助，入选教育部特聘教授和中组部第二批“万人计划”领军人才等。致力于无线通信理论方法、技术研究和工程实践等。获授权国家发明专利100多项，出版专著6部，发表学术论文600多篇。提出的PCC极化码成为5G控制信道唯一国际标准。获国家技术发明二等奖、省部级科技一等奖等多项。

报告题目：大语言模型中信息泄露风险的量化与防控

报告摘要：近年来，以 ChatGPT 为代表的大语言模型（Large Language Models, LLMs）在指数级扩张的训练输入数据的驱动下涌现出卓越的语义理解和内容生成能力，并加速向各类应用场景渗透，又进一步生成海量推理输出数据。然而，输入数据中所包含的个人隐私、企业私有数据等敏感信息会在预训练和微调阶段中被记忆并固化于模型的参数空间中，使敏感信息可能在推理阶段的模型输出数据中泄露。因此，信息泄露风险与大语言模型的完整生命周期紧密耦合，成为制约其安全合规应用与可持续发展的核心瓶颈。为此，本论文围绕大语言模型生命周期中预训练、微调阶段、推理阶段的三大核心阶段，系统性研究信息泄露风险的高精度量化与流式防控技术，力争实现准确率、适用范围和实时性等方面的突破，为大语言模型的安全部署与应用提供理论基础和技术支撑。



廖胜凯

廖胜凯，中国科大教授级高级工程师、合肥国家实验室研究员，主要从事远距离自由空间和星地量子通信的研究。曾参与墨子号量子科学实验卫星、天宫二号空间实验室、济南一号量子微纳卫星等项目，主持完成系列空间载荷终端研制，完成国际首次星地量子密钥分发实验、洲际量子密钥分发实验、53km自由空间白天量子密钥分发实验、星地实时量子密钥分发实验等。在Nature、Science、Nature Photonics、Physical Review Letters等国际重要学术刊物上发表论文70余篇，研究成果被ScienceNews、CNN、BBC等国外主流媒体广泛报道，入选2017年度“中国十大科技进展新闻”、2017年“中国科学十大进展”、美国物理学会2018年十大亮点，获2018年美国科学促进协会纽科姆·克利夫兰（团队）奖，2019年度中国科学院杰出科技成就奖(主要完成者)，入选国家级青年人才项目、中科院青促会优秀会员项目、安徽杰青项目、上海优秀学术带头人项目支持，获“中国载人航天突出贡献者”荣誉。

报告题目：卫星量子密钥分发实验及应用

报告摘要：基于卫星的量子通信是构建覆盖全球量子保密通信网络最为可行的方式，可以服务于超远距离、移动目标、岛屿和驻外机构等场景的安全通信。研究团队在量子保密通信京沪干线与“墨子号”量子卫星成功对接的基础上，构建了构建天地一体化广域量子保密通信网络的雏形。面向日益增多的用户需求，发射多颗低成本量子卫星实现组网运行，是构建高效率、实用化、全球化量子通信网络的有效路径，本报告介绍卫星的量子通信实验进展及应用探讨。



冀晓宇

冀晓宇，浙江大学教授、博导，国家高层次计划人才、优秀青年基金获得者。2010年浙江大学本科毕业，2015年香港科技大学计算机系博士毕业。主要从事物联网安全，包括传感器安全、具身智能安全等方面的研究。成果发表在ACM CCS, IEEE S&P, USENIX Security, NDSS四大安全会议。获得2025年NDSS杰出论文奖、2017年ACM CCS最佳论文奖（国内首次）、2018年ACM AsiaCCS最佳论文奖。研究成果获得中国电子学会技术发明一等奖、北京市科技进步一等奖、浙江省科技进步一等奖等。在物联网安全方面的研究成果被包括MIT Tech. Review, BBC News、人民网等国内外主流新闻媒体报道。

报告题目：物理AI及其安全思考

报告摘要：随着具身智能机器人、工业机械臂、自动驾驶汽车等物理AI系统加速走进家庭、产线与道路，人工智能正从信息空间迈入物理世界——其安全风险也随之越过数字边界，直接作用于人身安全与物理环境。如何驾驭物理AI安全这把“达摩克利斯之剑”，是物理AI产品走进千家万户的前提条件。物理AI区别于传统AI的本质，在于其通过“感知—决策—执行”闭环与物理世界持续交互：传感器成为横跨“物理—信息”两域的桥梁，也因此成为新的攻击面。本报告基于报告人10余年传感器安全与物理信号安全研究，以及近期在具身智能安全与治理方面的工作，系统剖析物理AI在感知、决策、执行全链路中的跨域脆弱性与安全风险。在此基础上，报告将从安全评测基准与数据集构建、攻击检测机制、行为对齐方法、标准与治理体系等方面，探讨覆盖“信息—物理—社会”三域的物理AI安全防护路径。



李书豪

李书豪，现任中关村实验室科研02部副部长、APT团队负责人、正高级工程师，某国家级创新人才，科技部某领域咨询专家，兼任国科大、哈工大等高校教授/博士生导师、北京市直机关青联委员、海淀青联科技界别副主任委员。研究方向是智能化威胁防御。近年来，承担多个国家级项目，主持研制中关村实验室APT防御平台，在数十家重点单位实战应用，成效显著，发表顶会顶刊论文10余篇，出版专著1部，获省部级科技进步一等奖2项。

报告题目：智能化网络威胁防御与危害消解技术探索及实践

报告摘要：以APT为代表的网络威胁已经呈现智能化网络攻击的新常态，也为我国重点单位网络安全防御带来严重挑战。本报告旨在探索“智能对抗智能”的威胁防御新机制，转变传统防御技术依赖入侵指标和威胁情报的被动局面。介绍数字替身诱导防御机理、高拟真替身生成技术、无感身份验证技术、威胁取证智能体技术等，并结合重点单位业务场景介绍中关村实验室APT团队在该方向的工作实践。



陈建军

陈建军，清华大学副教授，博士生导师，入选国家级青年人才计划。主要研究方向为网络协议安全，重点研究如何发现和防范复杂网络系统中的新型漏洞和攻击。研究成果多次获得国际网络安全四大顶级会议的杰出论文奖（中国首次获得该奖），促进了许多工业界厂商产品安全改进，并推动了国际标准组织IETF和W3C制定和修改国际标准。担任国际网安顶期IEEE T-IFS编委和多个顶会如ACM CCS、NDSS、IEEE S&P、ACM IMC等程序委员会委员。

报告题目：网络系统多方交互安全研究

报告摘要：网络系统安全性不只是检查某个协议或某个机制是否安全，而需要理解整体生态叠加之后形成多方交互安全风险。本报告将展示跨网跨层安全失效如何发生，包括如何实现现在加密无线链路中TCP/UDP 流量劫持；如何实现远程伪造电话和短信发信人；如何组合 HTTP/2 和 PKI 机制绕过同源策略隔离等。

**刁文瑞**

刁文瑞，山东大学网络空间安全学院教授、博士生导师，入选山东省泰山学者青年专家、小米青年学者，博士毕业于香港中文大学。研究方向为软件与系统安全，特别是移动生态与物联网系统安全，致力于智能化代码分析与漏洞挖掘。研究成果发表于IEEE S&P、USENIX-SEC、ACM CCS、NDSS、ASE、ICSE等系统安全与软件工程领域顶级国际学术会议，所识别的安全漏洞多次获得Google、华为、小米、荣耀等厂商官方致谢。获ACM CCS 2024杰出论文奖、ACM SIGSAC中国新星奖等学术荣誉。

报告题目：大语言模型赋能移动软件安全分析

报告摘要：本报告探讨大语言模型 (LLM) 如何与传统程序分析技术深度融合，以解决移动软件安全分析中的现有难题。在以 Android App 为代表的移动安全领域，长期面临诸多挑战，如系统 API 与安全策略的快速迭代导致检测规则滞后，代码混淆掩盖程序语义从而阻碍安全审计等。LLM 的语义理解与上下文推理能力为应对这些挑战提供了新路径，但直接应用 LLM 普遍面临幻觉严重、误报率高及缺乏可解释性等局限。报告将结合实际研究工作，介绍相关的探索与最新进展，主要包括：利用 LLM 从官方技术文档中自动合成静态安全检测规则，以及结合程序结构引导 LLM 实现高质量的代码标识符重建与语义恢复等。通过这些具体示例，展示并探讨 LLM 在提升移动软件安全分析效能方面的实践路径。

智能时代下的网络韧性技术与方法论坛

承办单位：紫金山实验室、国家数字交换系统工程技术研究中心、河南省网络空间内生安全重点实验室

🕒 9月18日下午

📍 天鹅湖大酒店三楼5号会议室

论坛主席： 邬江兴（国家数字交换中心）

季新生（国家数字交换中心）

召 集 人： 季新生（国家数字交换中心）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	“没事别乱问”一大语言模型prompt安全和隐私保护初探 仲 盛（南京大学）
14:40-15:10	卫星互联网安全：威胁与发展趋势 沈玉龙（西安电子科大）
15:10-15:40	智能网联系统安全体系研究及应用 谢少荣（上海大学）
15:40-16:00	休 息
16:00-16:30	工业无线网络风险感知与隐私保护研究 曹向辉（东南大学）
16:30-17:20	工业智能韧性安全：从组件到供应链 沈 超（西安交通大学）
17:20-17:50	AI时代网络攻防演化全景及内生安全体系战略价值 季新生（国家数字交换中心）

论坛主席



邬江兴

邬江兴院士是我国原创网络空间内生安全学派创立人，先后担任国家“八五”“九五”“十五”“十一五”863计划通信技术主题专家、副组长、信息领域专家组副组长，国家移动通信重大专项论证委员会主任，国家重大专项任务“高速信息示范网”“中国高性能宽带信息网——3Tnet”“新概念高效能计算机体系结构研究与系统开发”总体组组长，“新一代高可信网络”“可重构柔性网络”专项任务编制组负责人，国家“三网融合”专家组副组长，国家下一代广播电视网专家委员会主任，军委科技委兼职委员等。



季新生

长期从事通信网络架构、协议和通信网安全韧性研究，是国家自然科学基金创新群体带头人，主持完成国家重大专项课题、国家“863”计划重点项目等10余项，领衔研制成功业界首台全IP移动通信系统、首套移动通信网络安全防护系统等，获得国家发明专利授权20项，发表论文50篇，获国家科技进步一等奖3项（序1、10、10）、国家科技进步奖创新团队奖（序3带头人）、省部级科技进步一等奖2项（序1、序2），第13届光华科技奖、首届全国创新争先奖、首届全国网络安全优秀人才奖、入选百千万人才国家级人选暨有突出贡献中青年专家，入选中组部网信领域优秀人才。

召集人



季新生

长期从事通信网络架构、协议和通信网安全韧性研究，是国家自然科学基金创新群体带头人，主持完成国家重大专项课题、国家“863”计划重点项目等10余项，领衔研制成功业界首台全IP移动通信系统、首套移动通信网络安全防护系统等，获得国家发明专利授权20项，发表论文50篇，获国家科技进步一等奖3项（序1、10、10）、国家科技进步奖创新团队奖（序3带头人）、省部级科技进步一等奖2项（序1、序2），第13届光华科技奖、首届全国创新争先奖、首届全国网络安全优秀人才奖、入选百千万人才国家级人选暨有突出贡献中青年专家，入选中组部网信领域优秀人才。

报告嘉宾



仲 盛

仲盛，南京大学二级教授，软件学院院长，教育部长江学者特聘教授，国家杰出青年科学基金获得者，国家万人科技创新领军人才。IEEE Fellow，中国计算机学会会士。研究兴趣包括密码学、博弈论及其在计算机网络、分布式系统中的应用。

报告题目：“没事别乱问”——大语言模型prompt安全和隐私保护初探

报告摘要：大语言模型代表着当今人工智能技术的最前沿。我们探讨了与大语言模型prompt相关的几个安全和隐私保护问题，并提出我们的解决方案。实验结果表明，取得了较好的效果。



沈玉龙

沈玉龙，男，博士。长期从事下一代网络安全、云计算与数据安全等研究。近年来主持和承担了国家重点研发计划、国家自然科学基金重点项目等20余项，以主要完成人获国家技术发明二等奖2项，省部级一等奖7项。牵头发布了全球首款“玄知”密码专业大模型。发表SCI/EI论文100余篇，出版专著7部；授权发明专利70余项，PCT专利5项，已转让专利13项；登记软著40多项；制定中国通信/电子行业标准5项，相关无人机、物联网安全成果通过国防验收，支撑多项重点装备工程，实现理论、技术、产业转化全链条突破。

报告题目：卫星互联网安全：威胁与发展趋势

报告摘要：卫星互联网已上升为国家战略和意志，成为支撑低空经济、国防建设等重点领域的关键信息基础设施。卫星互联网网络全球接入开放性、访问互联网信息开放性、星地双向移动性、空间敌我网络混在性、卫星网络拓扑公开可预测性的等特性导致其在物理、电磁、网络、信息服务等空间均面临严峻安全威胁，而且这种攻击呈现日益显著的跨空间协同趋势。本次报告旨在介绍当前卫星互联网面临的典型安全威胁及安全事件，共同探讨卫星互联网安全核心技术以及未来发展面临的关键挑战。



谢少荣

谢少荣，国家自然科学基金委杰青、重大、创新研究群体负责人，国家万人计划科技创新领军人才，海上无人集群智能全国重点实验室副主任，海洋智能无人系统装备教育部工程研究中心主任，智能网联汽车网络安全上海市重点实验室主任。长期从事海洋自主智能无人艇系统理论、技术及应用研究，突破了无人艇自主航行控制、集群智能协同、集群安全体系等系列关键技术，获国家技术发明二等奖（序1）、省部级一等奖6项（4项序1）、中国青年科技奖、ACM China杰出学者奖等。

报告题目：智能网联系统安全体系研究及应用

报告摘要：报告分析了智能网联系统的功能安全、网络安全、人工智能安全耦合交织、相互促发，面临新域新质安全挑战；探索了智能网联系统安全体系新范式，创建系统安全风险视图，突破端云体系智能防护关键技术，从而构建智能网联系统“预防、抵御、恢复、适应”一体的安全体系，最后介绍了在海上智能装备、智能网联汽车、具身智能机器人等领域实装化创新应用。



曹向辉

曹向辉，东南大学国家卓越工程师学院副院长，自动化学院教授、博导。曾获浙江大学学士、博士学位，美国伊利诺伊理工学院博士后。主要从事网络化感知与安全控制的研究工作，主持国家自然科学基金、国防项目、江苏省前沿技术研发计划等科研项目，发表学术论文160余篇，其中IEEE TAC长文、Automatica长文等顶刊论文多篇。研究成果曾获教育部自然科学一等奖1项、省级科学技术奖二等奖2项、华为火花奖等奖励。担任《自动化学报》、《IEEE Transactions on Industrial Informatics》等期刊编委、中国自动化学会青年工作委员会副主任、信息物理系统控制与决策专委会、中国指控学会集群智能与协同控制专委会委员等。

报告题目：工业无线网络风险感知与隐私保护研究

报告摘要：工业无线网络作为智能制造核心设施，其开放共享特性引发严峻安全与隐私风险。在信息物理系统深度融合背景下，网络空间威胁延伸至控制空间。报告针对多种工业无线网络风险，汇报了融合网络与控制系统的智能风险感大模型，以及控制性能无损的轻量化隐私保护方法。


沈 超

沈超，IEEE Fellow，IAPR Fellow，西安交通大学二级教授，长江学者特聘教授，国家自然科学基金创新研究群体负责人、教育部创新团队负责人，科学探索奖、达摩院青橙奖、国家优秀青年基金获得者，国家重点研发计划首席，国家基础加强计划技术首席，重点研发计划指南专家组成员。主要从事智能系统可信、安全与控制的教学和研究工作，撰写专著4部，获最佳论文奖14次，牵头获3项省部级/一级学会科学技术一等奖、日内瓦国际发明金奖、霍英东青年教师一等奖、MIT TR35 China、IEEE SMC Early Career Award等。

报告题目：工业智能韧性安全：从组件到供应链

报告摘要：工业智能通过搭载先进的传感器、控制器和决策系统，旨在形成复杂工业网络环境中的自动化与智能化能力，因此在其设计与开发阶段，充分考量系统的智能特性与安全韧性尤为关键。本报告“从组件到供应链”对上述特性展开分析，梳理底层智能组件到上层任务数据之间的安全依赖关系，并围绕保密性、完整性与隐私性等核心维度，深入剖析智能供应链所面临的安全隐患与级联风险。在此基础上，报告进一步探讨如何构建面向真实工业环境的韧性安全能力，旨在识别并阻断风险在供应链链路中的传播路径，从而为智能化及大模型软件系统的工业安全应用与可持续发展提供支撑。


季新生

长期从事通信网络架构、协议和通信网安全韧性研究，是国家自然科学基金创新群体带头人，主持完成国家重大专项课题、国家“863”计划重点项目等10余项，领衔研制成功业界首台全IP移动通信系统、首套移动通信网络安全防护系统等，获得国家发明专利授权20项，发表论文50篇，获国家科技进步一等奖3项（序1、10、10）、国家科技进步奖创新团队奖（序3带头人）、省部级科技进步一等奖2项（序1、序2），第13届光华科技奖、首届全国创新争先奖、首届全国网络安全优秀人才奖、入选百千万人才国家级人选暨有突出贡献中青年专家，入选中组部网信领域优秀人才。

报告题目：AI时代网络攻防演化全景及内生安全体系战略价值

报告摘要：AI驱动攻击向能力模型化、攻击链闭环化、入口复合化与风险行动化加速演化，传统附加式防御在时间、漏洞与攻击链三维度持续坍塌，“易攻难守”困局加剧，如何构筑后AI时代的网络韧性成为亟待破解的战略命题。本报告旨在探讨AI时代网络攻防演化全景，并重点剖析我国独创的内生安全范式所具备的战略价值。该范式以“构造决定安全”为核心理念，依托动态异构冗余架构，将安全能力内生于系统机制，从结构上重构防御时间窗口、抑制漏洞可利用性、拒止攻击链传递。报告进而初步探讨内生安全与AI双向赋能的防御体系构想，以为AI时代网络安全防御的理论建构与实践路径提供参考。

生成式多媒体内容安全与可信论坛

承办单位：北京邮电大学、南开大学、江西财经大学

🕒 9月18日下午

📍 天鹅湖大酒店一楼长江厅

论坛主席：周琳娜（北京邮电大学）

王金伟（南开大学）

召集人：张玉书（江西财经大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:45	从Claude隐形水印回望生成式AI时代可信标识 俞能海（中国科学技术大学）
14:45-15:20	群体智能驱动的多模态认知计算 刘安安（天津大学）
15:20-15:55	文生图生态中的插件投毒 纪守领（浙江大学）
15:55-16:05	休 息
16:05-16:40	生成式多媒体模型安全认知构建与知识重塑 谢洪涛（中国科学技术大学）
16:40-17:15	生成式内容鉴别与安全 赫 然（中国科学院自动化研究所）

论坛主席



周琳娜

现任北京邮电大学网络空间安全学院教授、博士生导师，国家重点领域创新团队负责人，入选国家“万人计划”领军人才，享受国务院政府特殊津贴专家。兼任中国图象图形学会理事、中国防伪行业协会理事及全国金融标准化技术委员会委员。周琳娜教授长期致力于多媒体内容安全、信息隐藏与取证、人工智能安全等前沿领域的科学研究。主持承担了包括国家重点研发计划（“网络犯罪证据综合研判分析技术”等）、国家自然科学基金重点项目（“跨媒体并发隐信道的检测与对抗”）在内的10余项国家级重大课题。在IEEE TIFS、CVPR等顶级学术刊物及会议上发表论文50余篇，并出版了《数字图像内容取证》、《数字图像取证技术》等4部学术专著，构建了较为系统的多媒体取证理论体系。其科研成果在金融监管、司法鉴定及网络治理等领域实现了重要的转化应用，曾获国家科技进步一等奖1项（排名2）、国家科技进步二等奖3项、国家技术发明二等奖1项及省部级科技进步奖30余项，是我国多媒体安全与数字内容治理领域的学术带头人之一。



王金伟

南开大学密码与网络空间安全学院教授、博导，入选网信领域国家级人才和江苏省“333高层次”人才培养工程。研究方向包括人工智能安全、多媒体取证、信息隐藏等。现主持和参与国家自然科学基金面上、企业联合重点项目8项，省部级项目10余项，发表论文100余篇，其中SCI收录60余篇，荣获2023年河南省自然科学奖一等奖一项，兼任中国图象图形学会数字媒体取证与安全专委会副主任（专委会党工作小组副组长）、江苏省网络空间全学会仿真与推演专委会秘书长（专委会党建小组组长）、中国通信学会高级会员、CSIG高级会员。

召集人



张玉书

张玉书，教授 / 博导 / 副院长，目前就职于江西财经大学计算机与人工智能学院。自2013年以来累计在USENIX Security、TPAMI等顶刊顶会发表100余篇论文，Google引用1万3千余次，H指数65。主持国家基金青年项目B类（原优青）、国家重点研发课题、国家自然科学基金面上项目、江西省杰青、江西省重点研发项目等。入选江西省“赣鄱英才计划”创新领军人才。担任多个国际SCI期刊的副编，包括IEEE Transactions on Dependable and Secure Computing (CCF A)等。入选多个全球科学家榜单。获江西省自然科学奖一等奖、河南省自然科学二等奖等。

报告嘉宾



俞能海

俞能海，中国科学技术大学讲席教授、博士生导师，中国科大网络安全安全学科点负责人、原执行院长，教育部高等学校网络安全专业教学指导委员会副主任委员，中国网络安全安全学会副理事长，国家重点研发计划项目首席科学家。主要研究领域为人工智能安全。在TPAMI、S&P、CCS发表高水平论文300余篇，获得CCS 2024 杰出成果奖等优秀论文奖项。获评国家网络安全先进个人，安徽省教学名师。研究成果荣获军队科学技术进步奖一等奖，安徽省自然科学奖一等奖、安徽省技术发明一等奖等。

报告题目：智能体时代 AI安全研究范式演进

报告摘要：过去几年，AI安全研究主要回答一个问题：AI生成的内容能不能信？Deepfake带来的真实性危机、大模型越狱引发的安全失控、生成内容中的偏见与有害信息，这些都是“AI说了什么”的问题。我们围绕数据、模型和应用三大支柱，构建了覆盖图像、视频、语音、文本的可信生成式AI安全闭环。但今天，一个更严峻的问题正在浮现：AI做的事情能不能信？当AI智能体拥有持久记忆，它的记忆可以被悄无声息地污染，进而长期操控后续行为；当AI驱动机器人行走，特定输入模式可以让它瞬间失控暴走；当AI驱动的自动驾驶依赖传感器定位，一个欺骗性GPS信号就能让车辆驶入错误路径。AI的输出正在从屏幕上的文字和图像变成物理世界中的真实行动，安全后果也从信息风险升级为人身安全。本报告沿“AI说了什么”到“AI做了什么”这条主线，介绍课题组在可信生成式AI、智能体安全与具身智能安全方向的系统性研究进展，并探讨AI从生成走向自主的过程中，安全研究面临的新挑战。



刘安安

天津大学讲席教授，人工智能学院副院长，国家杰青，IET Fellow。从事多模态信息感知与决策，社会模拟器构建研究，服务于网络媒体安全审核与引导等。承担国家基金联合重点、重点研发等；以第一完成人获天津市科技进步特等奖；发表高水平论文百余篇；担任CSIG图像智能边缘计算专委会副主任、CSIG多媒体专委会常委，担任4个IEEE/ACM期刊编委以及ACM Multimedia领域主席。

报告题目：群体智能驱动的多模态认知计算

报告摘要：社会模拟器作为数字时代人工智能与社会科学交叉创新研究的方法论革命，其本质是构建人工社会实验室，对复杂社会系统进行“计算解构”与“涌现重构”，实现从静态观察到动态干预的转变，并通过多智能体的认知建模与多尺度社会网络的耦合演化，在虚拟空间中创造具有涌现特性的社会“数字孪生”。本报告将汇报团队基于群体智能驱动的多模态认知计算相关工作，探索群体智能在媒体内容流行度预测、虚假信息认知、个性化信息传播等重大社会议题中验证其决策支持价值，为数字社会治理提供了新的科学范式。



纪守领

纪守领，浙江大学计算机学院求是特聘教授/长聘教授、博士生导师，可信人工智能研究中心主任。获佐治亚理工学院电子与计算机工程博士学位、佐治亚州立大学计算机科学博士学位，入选国家级人才。主要研究方向为可信人工智能、软件与系统安全，主持国家重点研发计划项目、国家自然科学基金联合重点项目等多项，发表CCF A类论文100余篇、含安全四大顶会论文50余篇，研制的系统在多个大型平台上获得部署应用。获安全顶会IEEE S&P 2025杰出论文奖、ACM CCS 2021最佳论文奖等10项最佳论文奖、全球黑客大会Pwnie Awards最具创新研究奖提名、浙江大学先进工作者等。

报告题目：文生图生态中的插件投毒

报告摘要：LoRA插件生态正面临前所未有的供应链投毒风险。我们首次系统揭示两类隐蔽攻击——概念劫持可生成图像操纵舆论，任务注入则凭密钥触发有害内容（如NSFW）。更致命的是，恶意负载具备病毒式传播能力，借LoRA合并不断扩散。实验证明，该攻击在Civitai和Liblib两大平台、6个数据集上攻击成功率近乎100%，且跨模型、多次混合后依然稳健，而现有检测机制完全无法识别。这一发现也为文生图供应链插件共享生态敲响安全警钟。



谢洪涛

中国科学技术大学教授、博士生导师，国家杰出青年科学基金获得者。主要研究方向包括人工智能和多媒体内容安全等，以第一或通讯作者在国际一流期刊和会议上发表学术论文100余篇，担任ACM TOMM等四个国际著名期刊编委。主持科研项目10余项，含国家重点研发计划重点专项项目1项、国家自然科学基金联合重点项目2项。获2023年度国家技术发明奖二等奖、2019年度国家自然科学奖二等奖、2022年度中国图象图形学学会青年科学家奖等奖励。

报告题目：生成式多媒体模型安全认知构建与知识重塑

报告摘要：随着生成式模型能力持续演进，内容安全治理正从输出端防护延伸至模型认知与知识层面的协同调控。生成内容可分为确定性风险、情境性风险和良性内容三个区域，其间形成风险治理、安全判别和能力保护三类边界。确定性风险语义较为明确，常以内化知识的形式影响生成，需要通过选择性遗忘和知识重塑进行消解。情境性风险依赖具体情境、潜在意图和复杂语义关联，需要借助风险感知、情境推理与自我验证作出判断。风险治理边界可依据应用需求、治理规范和风险容忍度动态调整，以合理界定知识擦除与情境推理的范围。本报告从安全认知构建和安全知识重塑两个方面展开，分别研究情境性风险的价值对齐与确定性风险的概念遗忘，并关注风险调控对良性知识和模型能力的影响，为生成式模型内容安全治理提供边界可调、作用精准、能力可保持的研究思路。

**赫 然**

赫然，中国科学院自动化研究所多模态人工智能系统全国重点实验室研究员，IAPR/IEEE Fellow，IEEE TIFS/TBIOM副主编。从事人工智能、模式识别和计算机视觉研究。承担国家青年科学基金ABC类以及北京杰出青年科学基金等项目。在本领域国际主流期刊TPAMI和IJCV上发表论文30余篇，谷歌引用2.9万次；研究工作获CAAI技术发明一等奖、CSIG自然科学一等奖等。指导学生获得IEEE SPS最佳青年论文奖、ICPR最佳科学论文奖、北京市优秀博士论文、中科院优秀博士论文、IEEE生物特征理事会优秀博士论文。曾/现任TIP资深编委、TPAMI\TIP\TIFS\TCSVT\TBIOM\IJCV\IPR\TMLR和自动化学报等国内外期刊编委，五次获最佳编委奖，以及NIPS\ICML\ICCV\CVPR\ECCV\ICLR\AAAI\IJCAI等会议领域主席。

报告题目：生成式内容鉴别与安全

报告摘要：生成式人工智能通过对海量文本、图像、音频等数据的深度学习与模式挖掘自主生成符合人类认知逻辑的内容。基于大语言模型和多模态大语言模型的生成式人工智能技术正以迅猛之势重塑人工智能领域的发展格局。以GPT、DeepSeek为代表的生成模型凭借卓越的生成能力实现了重大技术突破，在创意创作、逻辑推理、自动决策等垂直领域的深度应用中展现出巨大潜力。本报告将介绍生成式人工智能的技术演进脉络，分析其发展过程中引发的安全风险与挑战，并介绍团队在高效图像生成、视觉内容鉴别与取证、模型安全等工作方向的工作。

新型电力系统信息物理综合安全风险与应对论坛

承办单位：西安交通大学网络空间安全学院、南京邮电大学自动化学院

🕒 9月18日下午

📍 天鹅湖大酒店一楼贵宾厅

论坛主席：刘 炅（西安交通大学）

岳 东（南京邮电大学）

召 集 人：刘 炅（西安交通大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	新型电力系统信息物理跨空间协同主动安全控制 岳 东（南京邮电大学）
14:40-15:10	大模型驱动的自动化攻防验证关键技术与实战应用 申大坤（浙江大学）
15:10-15:40	人工智能引发的电力系统数据安全新威胁与应对 刘 炅（西安交通大学）
15:40-16:00	休 息
16:00-16:30	新型电力系统接入安全 刘 园（广州大学）
16:30-17:50	讨论环节 主持人：刘 炅（西安交通大学） 嘉 宾：张 涛（中国电力科学研究院） 杨祎巍（南方电网科学研究院有限责任公司） 杜海涛（中国移动研究院） 宋美艳（中国华能集团有限公司） 岳 东（南京邮电大学） 温 蜜（上海电力大学） 刘 园（广州大学） 周亚东（西安交通大学）

论坛主席



刘 烜

博士，教授，西安交通大学网络空间安全学院副院长、信息物理融合系统教育部工程研究中心主任。主持国家自然科学基金青A项目、重大项目课题、重点项目，重点研发计划等科研项目30余项；获得2026 ICSE、2024 ASE、2019 INFOCOM等会议最佳论文奖8项；获得2017年国家科技进步二等奖，以及2022年教育部自然科学一等奖等省部级科技奖励6项；获得国家网信创新人才、教育部青年长江学者、陕西青年科技奖等荣誉。



岳 东

教育部长江学者特聘教授，IEEE Fellow，俄罗斯工程院外籍院士，IEEE尤金·米特elman成就奖获得者。现任南京邮电大学学术委员会主任，碳中和先进技术研究院长，江苏省主动配电网零碳运行控制工程研究中心主任。入选中国自动化学会会士（FCAA），中国人工智能学会会士（FCAAI）。担任IEEE TII共同主编，亚洲控制学会信息物理系统控制与安全专委会主席，IEEE PES智能电网与新技术卫星委员会（中国）主席，IEEE IES网络化控制系统及应用专委会主席（2019-2021）。长期从事网络化控制与优化理论及在能源电力系统应用研究，主持国家自然科学基金重大项目、重点项目、重点研发计划项目课题等30余项。获得教育部自然科学一等奖、IEEE Rudolf Chope科学技术奖、江苏省科学技术一等奖（2项）、中国自动化学会自然科学和发明一等奖、IEEE Transactions on Smart Grid最佳论文奖、IEEE Systems Journal最佳论文奖和 IEEE/CAA Journal of Automatica Sinica的Norbert Wiener Review奖等多项国内外奖励。

召集人



刘 烜

博士，教授，西安交通大学网络空间安全学院副院长、信息物理融合系统教育部工程研究中心主任。主持国家自然科学基金青A项目、重大项目课题、重点项目，重点研发计划等科研项目30余项；获得2026 ICSE、2024 ASE、2019 INFOCOM等会议最佳论文奖8项；获得2017年国家科技进步二等奖，以及2022年教育部自然科学一等奖等省部级科技奖励6项；获得国家网信创新人才、教育部青年长江学者、陕西青年科技奖等荣誉。

报告嘉宾



岳 东

教育部长江学者特聘教授，IEEE Fellow，俄罗斯工程院外籍院士，IEEE尤金·米特尔曼成就奖获得者。现任南京邮电大学校学术委员会主任，碳中和先进技术研究院长，江苏省主动配电网零碳运行控制工程研究中心主任。入选中国自动化学会会士（FCAA），中国人工智能学会会士（FCAAI）。担任IEEE TII共同主编，亚洲控制学会信息物理系统控制与安全专委会主席，IEEE PES智能电网与新技术卫星委员会（中国）主席，IEEE IES网络化控制系统及应用专委会主席（2019-2021）。长期从事网络化控制与优化理论及在能源电力系统应用研究，主持国家自然科学基金重大项目、重点项目、重点研发计划项目课题等30余项。获得教育部自然科学一等奖、IEEE Rudolf Chope科学技术奖、江苏省科学技术一等奖（2项）、中国自动化学会自然科学和发明一等奖、IEEE Transactions on Smart Grid最佳论文奖、IEEE Systems Journal最佳论文奖和 IEEE/CAA Journal of Automatica Sinica的Norbert Wiener Review奖等多项国内外奖励。

报告题目：新型电力系统信息物理跨空间协同主动安全控制

报告摘要：报告介绍了双碳背景下新型电力系统建设过程中存在的信息物理安全性问题，以及传统方法解决此类安全性问题存在的局限。进一步提出了亟待解决的三个关键科学问题，以及围绕上述关键科学问题解决需要研究的主要内容和相应的研究方案。



申大坤

浙江大学区块链与数据安全全国重点实验室研究员。长期深耕智能渗透测试、自动化漏洞利用、攻防智能体等技术研究。主持、参与国家重点研发、网络安全专项等省部级以上科研项目5项。牵头研发AI自动化渗透测试与攻击链建模平台，落地多重点行业单位，支撑网络安全实战核查工作。在安全、通信领域顶会期刊发表论文十余篇，研究方向包含程序分析、漏洞建模、攻防博弈、多智能体协同等。

报告题目：大模型驱动的自动化攻防验证关键技术与实战应用

报告摘要：大模型与智能体快速发展，AI自动化漏洞挖掘、渗透验证已成网络安全建设重点。本次汇报结合电力、政务、教育行业实战项目，分享AI自动化安全验证体系核心技术与落地经验。围绕资产测绘、漏洞探测、路径规划、自动利用、取证存证、风险评估全流程，解读多智能体协同、代码分析、漏洞推理、工具自主调度对复杂系统攻防效率的提升价值。结合电网平台、教育攻防演练、重大漏洞案例，介绍自动化挖洞、权限突破、攻击链复现、证据固化实操方案，梳理逻辑漏洞、会话系统、多级攻击落地思路，明确测试适用边界、风险管控与发展趋势，为AI落地渗透、安全评估提供实践借鉴。



刘 炯

博士，教授，西安交通大学网络空间安全学院副院长、信息物理融合系统教育部工程研究中心主任。主持国家自然科学基金青A项目、重大项目课题、重点项目，重点研发计划等科研项目30余项；获得2026 ICSE、2024 ASE、2019 INFOCOM等会议最佳论文奖8项；获得2017年国家科技进步二等奖，以及2022年教育部自然科学一等奖等省部级科技奖励6项；获得国家网信创新人才、教育部青年长江学者、陕西青年科技奖等荣誉。

报告题目：人工智能引发的电力系统数据安全新威胁与应对

报告摘要：电力系统是人类社会最重要的基础设施，世界各国对于电力系统的数据均采取了严格的保密管理，涉电数据发布均需要经过审批和脱敏处理。人工智能在数据收集、分析、推理等方面的能力已经被广泛证明，对于专业性强、敏感程度高的电力系统数据，人工智能同样展现出极强的推演能力，现有的混淆、删除等方式已基本失效。本报告结合实际案例展示人工智能对电力系统数据安全的威胁，归纳数据泄露的新范式，探讨人工智能时代数据保护的应对方法。



刘 园

广州大学教授，博导，主持国家自然科学基金青年B类（原优青）、重点项目、国家重点研发青年科学家项目、广东省杰青等。聚焦电网安全、网络安全、数据安全、区块链安全等领域研究，发表/录用学术论文100余篇，申请技术专利200+项。担任IEEE TNSM、TSC、FCS、BCRA编委、《信息网络安全》青年编委。新加坡南洋理工大学博士（2014.11），哈尔滨工业大学本科（2010.07），曾任东北大学信息安全系主任（2016-2019）。研究成果在国家重要活动保障任务中应用和验证。

报告题目：新型电力系统接入安全

报告摘要：随着新型电力系统建设加速推进，分布式光伏、储能、电动汽车等海量新型主体持续接入配电网，带来防御关口不可控、交互数据易篡改等多重挑战。本报告围绕新型电力系统接入安全展开，分享团队在威胁情报可解释分析、联邦学习协同防御、区块链激励共享、主动探查欺骗防御等方面的技术，包括面向海量新型主体轻量化接入与威胁感知技术及基于自适应设陷感知与信誉模型的可信度评估技术。并展望未来可信配电网接入环境应从单点安全防护走向多主体协同治理，从事后追溯走向事前感知与主动抑制，从静态授权走向动态可信评估，为支撑高比例新能源消纳、用户侧灵活互动和新型电力系统安全运行提供理论方法与技术支撑。

开源软件与供应链安全论坛

承办单位：华中科技大学网络空间安全学院

🕒 9月18日下午

📍 天鹅湖大酒店三楼7号会议室

论坛主席：邹德清（网络空间安全学院/华中科技大学）

杨 珉（复旦大学）

召集人：邹德清（网络空间安全学院/华中科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	产业研发视角下的开源软件供应链安全风险与应对 刘 炆（西安交通大学）
14:40-15:10	大模型赋能的基础软硬件安全测试 姜 宇（清华大学）
15:10-15:40	大模型驱动的多语言代码漏洞挖掘技术 张 源（复旦大学）
15:40-16:00	休 息
16:00-16:30	面向深度嵌入式系统安全分析的MCU外设自动仿真方法 凌 振（东南大学）
16:30-17:20	FM-Agent:面向大型系统的霍尔范式自动化推理及领域实战 王肇国（上海交通大学）
17:20-17:50	大模型软件供应链安全：现状、机遇与挑战 王浩宇（华中科技大学）

论坛主席



邹德清

博士，华中科技大学二级教授，长江学者特聘教授，国家网络安全优秀教师，教育部新世纪优秀人才。分布式系统安全湖北省重点实验室主任，武汉金银湖实验室主任，大数据技术与系统国家地方联合工程研究中心副主任，教育部网络空间安全教育指导委员会委员，全国网络安全标准化技术委员会委员，湖北省数字公共基础设施开源联盟理事长，华中科技大学开源创新研究院院长，武汉市网安基地校企联合会会长。作为学院主要创建者和学科带头人，也是湖北省五四奖章集体“华中科技大学系统与软件安全团队”负责人，科技部、教育部创新团队主要成员。



杨 珉

博士，国务院学位委员会网络空间安全学科评议组成员、教育部长江学者特聘教授、973项目首席科学家、国家重点研发计划首席科学家。主研领域为网络空间安全战略与技术，包括AI安全、系统安全、漏洞分析挖掘、软件供应链安全等，主持国家973项目、国家重点研发计划项目等，相关研究成果已产业化应用于阿里、华为、百度、电信、OPPO、vivo等头部企业；已全数获得国际网络安全四大顶会的杰出论文奖，其中两项为国内首次；获上海市决策咨询研究成果奖一等奖、上海市技术发明奖一等奖；研究发现曾多次获中央领导重要批示，成果入选973重大成果和突出进展、十三五国家科技创新成就、十四五科技部专项标志性成果、工信部科技创新和产业创新深度融合成果等。

召集人



邹德清

博士，华中科技大学二级教授，长江学者特聘教授，国家网络安全优秀教师，教育部新世纪优秀人才。分布式系统安全湖北省重点实验室主任，武汉金银湖实验室主任，大数据技术与系统国家地方联合工程研究中心副主任，教育部网络空间安全教育指导委员会委员，全国网络安全标准化技术委员会委员，湖北省数字公共基础设施开源联盟理事长，华中科技大学开源创新研究院院长，武汉市网安基地校企联合会会长。作为学院主要创建者和学科带头人，也是湖北省五四奖章集体“华中科技大学系统与软件安全团队”负责人，科技部、教育部创新团队主要成员。

报告嘉宾



刘 炅

西安交通大学二级教授、领军学者，网络空间安全学院副院长、信息物理融合系统教育部工程研究中心主任。主持国家自然科学基金青A项目、重大项目课题、重点项目，重点研发计划等科研项目30余项；获得2026 ICSE、2024 ASE、2019 INFOCOM等会议杰出/最佳论文奖8项；获得2017年国家科技进步二等奖，以及2022年教育部自然科学一等奖等省部级科技奖励6项；获得国家网信创新人才、教育部青年长江学者、陕西青年科技奖等荣誉。

报告题目：产业研发视角下的开源软件供应链安全风险

报告摘要：开源软件深度嵌入企业的产品研发与交付流程，与此同时针对开源软件供应链的攻击事件已成为软件行业最重大的安全威胁之一。本报告从产业研发视角分析研发效率与软件安全的内生矛盾，揭示了产业组织流程合规的隐性共识。结合案例论证了基于流程合规的安全体系无法应对开源软件供应链攻击，并基于产业视角提出了开源软件供应链攻击分类演化框架：开源共建威胁产生、闭源研发威胁演化、成品使用攻击发作；从面向开源生态的协同治理、面向产业研发的持续合规与攻击面收敛、面向成品使用的自适应防护三方面，提出开源软件供应链的安全再平衡体系。



姜 宇

清华大学软件学院教授。从事操作系统、数据库管理系统等大规模系统软件的安全测试与防护工作。研制的自动化模糊测试工具被并入谷歌ClusterFuzz与微软OneFuzz等平台，服务于30余家安全攸关企事业单位，挖掘500余个零日漏洞收入中美国家信息安全漏洞库。在SOSP、S&P、ATC、SIGMOD等CCF-A类会议发表论文100余篇，获Security、FSE等CCF-A类会议杰出论文奖5次，以第一完成人获北京市科技进步奖一等奖、中国计算机学会技术发明奖一等奖等奖励，主持国家重点研发计划项目，基金委青年科学基金项目A类、B类等20余项。

报告题目：大模型赋能的基础软硬件安全测试

报告摘要：报告将梳理当前底层芯片和上层操作系统和数据库等基础软硬件安全测试存在核心挑战：一是基础硬件及软件理解门槛高，传统工具缺少领域知识，测试输入生成随机性强、有效用例产出率低；二是软硬件执行场景复杂，跨层漏洞难以识别与溯源，动态模糊测试路径覆盖效率不足，深层高危漏洞挖掘能力薄弱；三是软硬件异构测试环境搭建繁琐，逻辑问题、权限逃逸等缺陷自动判定难。针对上述问题，详细介绍其团队基于多智能体大模型构建的一体化安全测试最新研究思路与进展：依托大模型实现处理器指令文法及基础软件输入文法的智能解析与高价值测试种子生成，融合推理大模型完成跨层漏洞自动溯源与缺陷精准判定，搭建适配处理器、操作系统、数据库、协议的统一自动化测试平台，为基础软硬件漏洞主动挖掘提供支撑。



张 源

复旦大学教授、博导，教育部长江学者，复旦白泽战队指导教师，主要研究方向为软件安全，研究工作获网安顶会杰出论文奖3次、荣誉提名2次，部分研究成果应用于华为、阿里、OPPO、vivo等公司，获优秀校企合作项目5次。带领复旦白泽战队获国内外顶尖安全攻防赛事冠军20余次，发现大量真实软硬件系统高危漏洞，获2021年国家信息安全漏洞库最具价值漏洞奖。

报告题目：大模型驱动的多语言代码漏洞挖掘技术

报告摘要：当前基于大模型的代码漏洞挖掘方法，在多语言复杂代码审计场景下仍存在突出短板。一方面，大模型对不同编程语言的语言特性、框架特性及漏洞触发逻辑存在理解偏差、跨语言泛化能力不足；另一方面，在复杂业务逻辑场景下，检测结果不稳定、误报率偏高，易产生大量低质漏洞报告。本报告提出程序分析与大模型有机协同的智能化漏洞挖掘技术，实现支持Java/C/C++/PHP等多种语言的通用漏洞挖掘框架。目前该框架已发现300+零天漏洞，误报率低于20%，在多个头部厂商落地应用。



凌 振

国家优青，东南大学青年首席教授。研究方向包括匿名网络、物联网安全等。近年来发表ToN、TIFS、TPDS、TC、TDSC、INFOCOM、CCS、S&P、Security、NDSS等高水平学术论文100余篇。先后主持多项国家自然科学基金、国家重点研发计划等项目。多次受邀在BlackHat等国内外重要会议作特邀报告。获ACM中国和CCF优博论文奖、全国公安技术革新成果一等奖等奖项。担任ACM SIGSAC China秘书长，TDSC等编委，曾任MSN 2023程序委员会联合主席，担任CCS、Security、INFOCOM等多个知名会议程序委员会委员。

报告题目：面向深度嵌入式系统安全分析的MCU外设自动仿真方法

报告摘要：针对MCU外设异构导致固件仿真成本高、扩展性不足的问题，提出面向深度嵌入式系统安全分析的FlexEmu框架。该框架创新采用结构—语义双层建模机制，以9类通用原语抽象外设结构，并以统一语义模型描述同类外设功能；随后利用大语言模型从驱动源码中提取具体实现参数，完成模型实例化，并通过模板化代码生成外设模拟器。该框架支持12类常用外设，在15种MCU、90个固件样本上的测试通过率达98.48%；通过模糊测试三种实时操作系统的蓝牙协议栈，发现10个未知漏洞。



王肇国

重点研发计划项目负责人，国家优秀青年科学基金获得者。从事系统软件研究，成果发表在OSDI、SIGMOD、PODC等相关领域权威会议上。发表Communications of the ACM亮点论文和封面文章，获2023 ACM SIGMOD研究亮点奖、SIGMOD 2022最佳论文优胜奖、CCF青年科技奖、ACM ChinaSys新星奖、华为奥林帕斯先锋奖，以及两次华为火花奖。学术兼职包括OpenHarmony智能数据管理TSG负责人、ACM ChinaSys秘书长、CCF学术工作委员会委员等。曾受邀担任MLSys、EuroSys、NSDI等国际会议的程序委员会成员，第22期秀湖论坛联合主席等。

报告题目：FM-Agent：面向大型系统软件的霍尔范式自动化推理智能体及领域实战

报告摘要：如今的Coding Agent生成十万行以上的系统级代码，甚至构建一个完整的编译器，都已不再稀奇。但一个严峻的挑战随之而来：如何保障这些大规模代码的正确性？为此，我们推出了形式化方法智能体FM-Agent，通过将霍尔逻辑与LLM结合，首次实现了面向大规模软件的全自动正确性推理，在 Anthropic、NVIDIA 等用顶尖编程智能体生成的多个大规模系统（单个系统规模高达 14.3 万行）中，FM-Agent 报告了 522 个潜在 bug。值得关注的是，这些 bug 经过单元测试、差分测试、多智能体交叉审查等手段都未能发现。本次报告将向大家分享FM-Agent的最新进展，以及面向编译器、数据库等系统的实战经验。



王浩宇

华中科技大学网络空间安全学院教授、博导，教育部青年长江学者，华中卓越学者。研究方向为新兴软件系统安全，近年共发表CCF A类论文140余篇，获北京市科技进步一等奖等省部级科技奖励三项。

报告题目：大模型软件供应链安全：现状、机遇与挑战

报告摘要：以大模型为核心的智能体已实现从对话交互到自主执行的跨越。与此同时，大模型的规模化应用推动软件供应链进入AI原生新阶段，从传统的代码、组件延伸至模型训练、推理框架、工作流编排、插件工具等全链条。本报告将针对大模型系统基础设施、组件框架到上层应用等软件供应链核心组成部分进行安全分析，并对领域未来的研究方向、潜在机遇及关键挑战展开展望。

网络空间多域融合安全体系论坛

承办单位：中国电子科技集团有限公司总体院、国防科技工业网络安全创新中心、哈尔滨工业大学（深圳）

🕒 9月18日下午

📍 天鹅湖大酒店一楼淮河厅B

论坛主席：陆 军（中国工程院院士）

饶志宏（中国电子科技集团有限公司）

召 集 人：廖 清（哈尔滨工业大学（深圳））

刘凌旗（中国电子科技集团有限公司总体院）

时 间	内 容
14:00-14:05	主持开场
14:05-14:35	致 辞
14:35-14:55	多域融合安全体系构建理论与方法 饶志宏（中国电子科技集团有限公司）
14:55-15:15	大数据驱动的情报安全研究 程学旗（中国科学院计算技术研究所）
15:15-15:35	AI驱动的网络攻击与内生安全防御体系 罗洪斌（北京航空航天大学）
15:35-15:55	人机共生时代的智能认知对抗 王 震（西北工业大学）
15:55-16:05	休 息
16:05-16:25	智能体安全初探 吕勇强（清华大学）
16:25-16:45	无线网络安全与测试技术研究 王 伟（中山大学）
16:45-17:05	高精度光纤时间同步安全防护技术研究 曾浩洋（中国电科三十所）
17:05-17:25	多域融合网络战下关键信息基础设施的安全防护 张 焱（青藤云安全）

论坛主席



陆军

中国工程院院士、中国电子科技集团首席科学家。量子科技长三角产业创新中心主任，执行理事，中国电子学会量子信息分会委员、常务副主任。长期从事雷达及综合电子、量子信息系统领域理论研究工作，专研综合电子信息系统研究与重大型号工程研制30余年，在电子技术、信息系统与人工智能等领域有深厚的造诣，拥有发明专利13项，论著22项。曾获国家科技进步特等奖，全国五一劳动奖章等。带领团队获国家科技部“创新团队”称号。



饶志宏

中国电子科技集团有限公司总师召集人和首席科学家，国防科技工业网络安全创新中心主任，国家“万人计划”科技创新领军人才，首届国家网络安全先进个人，国家重点研发计划“网络空间安全”重点专项负责人，享受国务院政府特殊津贴专家。主要从事网络空间安全总体和关键技术研究，在网络空间对抗、关键基础设施安全等方面取得系列成就，被聘为国家部委和军队网络安全领域核心专家。获得省部级一等奖5项、二等奖5项，专利80余项，专著5部。

召集人



廖清

网络空间安全研究院院长，二级教授，博士生导师，国家级高层次人才，获第十八届中国青年科技奖，科学探索奖，全球前2%顶尖科学家。长期从事数据挖掘、人工智能与信息安全领域研究，主持并承担国家重点研发计划、国家自然科学基金及广东省重点等多项科研项目。作为第一作者或通讯作者，已在IEEE TKDE、NeurIPS、AAAI、SDM、IEEE TMC等国际重要期刊和会议发表高水平学术论文100余篇，相关成果获国家级二等奖及省部级一等奖等奖项。



刘凌旗

中国电子科技集团有限公司总体院主任、正高级工程师，北京大学博士。长期从事电子信息前沿技术和网络空间认知安全战略研究工作，主持并承担中央网信办、国家XX局、中国工程院及军委科技委、装发等上级项目任务。曾获省部级一等奖1项、二等奖1项，电科集团杰出青年、信息工作先进个人，联参XX局一等奖2项等，发表论文30余篇。曾兼任中国国防科技信息学会副秘书长，现任中国指控学会仿真专委会常务委员、国防科技工业网安创新中心委员、广东省认知域工程中心委员等。

报告嘉宾



饶志宏

中国电子科技集团总师召集人和首席科学家，国防科技工业网络安全创新中心主任，国家“万人计划”科技创新领军人才，首届国家网络安全先进个人，国家重点研发计划“网络空间安全”重点专项负责人，享受国务院政府特殊津贴专家。主要从事网络空间安全总体和关键技术研究，在网络空间对抗、关键基础设施安全等方面取得系列成就，被聘为国家部委和军队网络安全领域核心专家。获得省部级一等奖5项、二等奖5项，专利80余项，专著5部。

报告题目：多域融合安全体系构建理论与方法

报告摘要：第四次工业革命带来的新技术重塑战争形态和规则，模糊了战场边界，网络空间已成为核心焦点。作为一个由信息、控制和系统三大要素深度融合构成的复杂巨系统，网络空间的理论基石离不开诺伯特·维纳的控制论、克劳德·香农的信息论和钱学森先生的系统工程论，在社会认知域、网络信息域、电磁频谱域和地理环境域等方面显现其独有特点。围绕网络空间攻防智能动态博弈，研究提出网络空间安全态势之变、熵减之道、体系构建之法 and 系列关键之术，采用OOPDAL安全链对资源进行筹划编排，实现安全体系向“动态多域融合”“智能驱动闭环”范式升级。网络空间安全博弈熵减之道、体系构建之法和系列关键之术，为网络空间安全可量化、可评估、可演进的体系构建提供参考。



程学旗

研究员，博士生导师，中国科学院计算技术研究所副所长、网络数据科学与技术重点实验室主任，中国科学院大学讲座教授，中科大数据研究院首席科学家，国家万人计划科技创新领军人才。主要从事网络数据科学、大数据系统、社会计算、Web信息检索与数据挖掘等方面研究，获国家杰出青年科学基金，国务院特殊津贴、中国电子学会科学技术奖一等奖、国家科技进步奖二等奖、第十二届中国青年科技奖等，在重要国际学术期刊和会议发表论文200余篇，Google Scholar引用1万余次，获得授权专利60余件。

报告题目：大数据驱动的情报安全研究

报告摘要：聚焦大数据时代情报安全面临的新挑战与新范式，系统阐述“数据驱动”如何重塑情报分析、威胁感知与安全决策的全链条逻辑。报告首先解析大数据环境下情报安全的内涵演变，指出传统的“被动防御”已难以应对高维、动态、隐蔽的智能攻击，必须转向“主动认知”与“预测预警”。其次，重点介绍基于多源异构数据融合的威胁情报建模方法，包括知识图谱构建、异常行为特征提取及深度学习驱动的安全事件关联分析，并辅以真实案例演示数据挖掘在APT攻击识别、内部威胁预警中的关键作用。最后，探讨情报安全中的数据主权、隐私保护与伦理边界，提出“安全数据共享+联邦学习”的协同框架。报告旨在为科研人员与实务工作者提供前沿视角，推动大数据技术与情报安全研究的深度融合。



罗洪斌

北京航空航天大学网络安全学院教授、博士生导师，中国网络安全安全学会理事；长期从事数据跨网安全高效流动理论与关键技术方面的创新研究与教学工作，主持国家重点研发计划课题、973 计划课题、国家自然科学基金等项目十余项，曾获国家技术发明二等奖、中国电子学会科学技术奖一等奖、教育部技术发明一等奖、北京青年五四奖章、全国百篇优秀博士学位论文提名奖等奖励，2012 年入选教育部新世纪优秀人才支持计划、2014 年获国家优秀青年科学基金，2022年获国家杰出青年科学基金。

报告题目：AI驱动的网络攻击与内生安全防御体系

报告摘要：当前，以Mythos5为代表的AI大模型，正在重塑网络攻击的形态。Mythos5能够自主挖掘零日漏洞并生成可直接使用的攻击代码、自主完成从侦察到数据窃取的全链路APT攻击、具备极强的社会工程学能力并生成高度逼真的钓鱼内容等，使其成为一个前所未有的安全挑战。本报告从攻击链的视角，首先分析Mythos 5如何提升攻击链各步骤的攻击效能，揭示了控制连通性是防范Mythos 5攻击的最直接、最有效手段。本报告进一步分析为什么难以彻底掌控IP的连通性；在此基础上，借鉴自然科学原理揭示了数据流动的基本要素，并给出了一种掌控连通性的网络体系—共生网络，可实现对AI攻击的内生安全防御。



王震

西北工业大学教授，网安学院院长（国家一流示范网安学院），国家保密学院常务副院长，IEEE/IAPR/IOP Fellow，国家杰青(2020)，国家级科技创新团队负责人(2020)。研究方向为人工智能，智能无人系统，智能认知对抗。在 PNAS, Nature/Science 子刊等发表系列成果，研制的系统应用于数个型号。获国家技术发明奖二等奖、国家教育教学成果奖二等奖，牵头获陈嘉庚青年科技奖、科学探索奖、全国创新争先奖章、中国青年五四奖章、全国五一劳动奖章、教育部、陕西省(2项)、中国自动化学会、中国电子学会科学技术奖一等奖等。

报告题目：人机共生时代的智能认知对抗

报告摘要：随着大语言模型与AI智能体技术的爆发式发展，人机共生的社会形态正在加速到来。在这一新兴范式中，AI系统不仅作为工具存在，更成为信息处理、知识生成与决策支持的核心环节，甚至在某些场景下已成为“第一信源”与“最近端决策点”。当人类认知过程与AI系统深度交织，我们面临前所未有的认知安全挑战：AI系统如何避免被恶意操控？人类如何保持认知自主性？如何构建健康的人机认知生态？本报告将探讨人机共生背景下的认知安全新范式，分析AI智能体的认知脆弱性与威胁模式，并思考技术、制度与伦理层面的应对策略，以期这一新兴研究领域提供系统性思考框架。

**吕勇强**

清华大学信息国家研究中心副研究员，主要从事处理器漏洞挖掘、安全验证与设计研究。参与“核高基”数字芯片后端设计工具平台研发，研究成果发表于CCS、IEEE TCAD等期刊会议。获2020年中国计算机学会自然科学奖一等奖、2019年中国电子学会科技进步奖一等奖、2015年国家科技进步奖二等奖，2023年入选国家级高层次人才。

报告题目：处理器芯片密码安全研究

报告摘要：新一轮科技革命推动处理器芯片向高性能、异构化和复杂微架构方向演进，芯片内部状态转换、指令执行、存储访问和密码运算之间的深度耦合，使安全风险由单点缺陷扩展至体系性漏洞。处理器芯片既是信息系统的计算核心，也是密码算法和安全协议的执行载体，其设计错误、实现偏差、微架构泄露及故障效应可能造成密钥暴露、数据篡改和安全机制失效。本报告围绕处理器芯片漏洞挖掘与密码安全分析，研究面向指令集、流水线、缓存、总线和密码部件的缺陷发现方法，分析侧信道、故障注入等攻击对密码实现的影响，探讨结合仿真验证、形式化分析和实测实验的安全评估方法，为处理器芯片安全设计、密码模块验证和风险处置提供参考。

**王伟**

中山大学电子与通信工程学院教授，副院长，现任国家某重大工程专家组专家、广东省低空安全探测工程中心主任、广东省智能全息雷达工程中心主任。曾担任国家863高技术计划主题专家组专家、电子信息系统复杂电磁环境效应国家重点实验室教授。长期从事网络电磁空间安全、雷达探测与对抗、电子信息系统仿真建模与评估方向的科研、教学、与人才培养工作，承担“973”项目、国家“863”项目、自然科学基金项目、系统研制项目等30余项。获得部委级等科技进步奖励11项，其中一等奖三项，二等奖四项，三等奖四项。培养、协助培养博士研究生20余名，公开发表论文70余篇，申请专利20余项。公开出版专著2部，出版教材2部。

报告题目：处理器芯片密码安全研究

报告摘要：本报告介绍了智能时代无线网络电磁频谱安全的相关关键技术，包括辐射源快速智能定位技术、辐射源信号的分选识别技术、辐射源的个体识别技术、无线网络电磁频谱的认知控制技术以及这些技术在无线网络安全与测试技术的应用。



曾浩洋

正高级工程师，中国电科首席科学家，中国网安科技委主任，中国电科30所总工，国务院特殊津贴获得者，中国通信学会高级会员，国防科技工业科技委成员，入选天府杰出科学家，中国电科博士生/硕士生导师。长期从事专用通信网络与安全技术研究，担任多项国家、军队重大项目技术总师，获国家科技进步二等奖1项，省部级科技进步奖8项。

报告题目：高精度光纤时间同步安全防护技术研究

报告摘要：统一的时间基准是信息化时代国家安全和国民经济的基础，电力、通信、交通、金融等涉及国民经济命脉的关键领域都依赖精准的时间同步系统。在大国战略博弈持续加剧的背景下，时间同步系统面临着严峻的安全威胁。基于光纤的地基时频网络基础设施，其时间同步的安全性越来越受到重视。时间同步安全不仅涉及传统数据交互的安全，也涉及物理信号传输的安全，同时还需要解决安全机制影响时间同步性能问题。报告针对上述特殊安全需求，提出了时间同步系统链路、节点和网络三级安全防护架构，介绍了链路时变状态高精度感知、链路攻击精确识别以及网络多源时间信息安全融合等时间同步安全技术路径和当前的研究进展，指出了下一步的研究方向。



张焱

研究员级高级工程师、青藤云安全首席安全官、国资国企在线监管安全运营体系联合创新中心顾问专家、四川省人工智能学会安全专委会副主任。专注于网络对抗、网络安全和人工智能领域的专业技术和理论研究，技术成果荣获军队科技进步一等奖一次、省部级科技进步一等奖一次、国防科技工业企业管理创新成果二等奖一次，中央企业青年岗位能手一次，参与编著《ATT&CK框架实践指南（第二版）》、《先进云安全研究与实践》等著作。

报告题目：多域融合网络战下关键信息基础设施的安全防护

报告摘要：本报告聚焦多域融合网络战下关键信息基础设施安全防护，指出当前攻击已形成物理、网络、信息、认知、社会等域跨域杀伤闭环，OT/IT深度融合使关基攻击面持续扩张，面临物理毁伤、风险级联、认知次生灾害等多重威胁。传统防护沿用IT思路，存在生产确定性与防护不确定性的核心矛盾，普遍陷入“只审计、不拦截”的落地困境，同时存在被动响应滞后、供应链管控薄弱、认知域防护真空等结构性问题。针对IT/OT融合场景以及AI对抗AI的技术趋势，神经符号融合架构是核心破局技术：以符号本体为确定性内核保障决策可解释、零误拦，神经层负责未知威胁感知，结合五级指令校验，从根源化解生产与安全的核心冲突。同时以风险优先级模型为决策框架，同步配套主动防御、供应链治理、技舆联动等多域协同措施，构建务实的实战化防护体系。

智能赋能的系统工程安全论坛

承办单位：中国电科网络空间研究院

🕒 9月18日下午

📍 天鹅湖大酒店三楼12号会议室

论坛主席：郭世泽（中国科学院院士）

蔡 红（中国电科网络空间研究院）

召 集 人：蔡 红（中国电科网络空间研究院）

时 间	内 容
14:00-14:10	主持开场
14:10-14:20	致 辞
14:20-14:50	从哪里来、在哪里、到哪里去——智能时代的网络安全系统工程的一些思考 张建军（中国电子科技集团有限公司）
14:50-15:20	从“找漏洞”到“证利用”：逻辑漏洞智能挖掘的范式迁移 霍 玮（中国科学院信息工程研究所）
15:20-15:50	面向网络设备的智能化漏洞挖掘探索与实践 史 记（中国电科网络空间研究院）
15:50-16:10	休 息
16:10-16:40	基于系统溯源图的复杂网络攻击智能检测与分析 杨 力（西安电子科技大学）
16:40-17:10	智能时代下的空天安全通信系统发展思路 潘高峰（北京理工大学）
17:10-17:30	论坛闭幕

论坛主席



郭世泽

解放军信息技术安全研究中心研究员，网络空间安全防御领域顶级专家。



蔡红

现任中国电科网络空间研究院党委书记、院长。长期深耕系统工程、情报侦察、信息系统科研与管理一线，工程统筹、资源协调能力突出，历任课题负责人、型号总师。牵头完成十余项国家级、省部级重大装备研制，攻克多项关键技术，斩获国防科技奖、集团科技奖、全国质量管理卓越领导者等荣誉。深耕通信系统总体论证设计，以一作 / 独立作者在《电讯技术》刊发多篇核心论文，理论与工程功底扎实，有力推动行业技术创新与人才培养。

召集人



蔡红

现任中国电科网络空间研究院党委书记、院长。长期深耕系统工程、情报侦察、信息系统科研与管理一线，工程统筹、资源协调能力突出，历任课题负责人、型号总师。牵头完成十余项国家级、省部级重大装备研制，攻克多项关键技术，斩获国防科技奖、集团科技奖、全国质量管理卓越领导者等荣誉。深耕通信系统总体论证设计，以一作 / 独立作者在《电讯技术》刊发多篇核心论文，理论与工程功底扎实，有力推动行业技术创新与人才培养。

报告嘉宾



张建军

近三十年的网络安全和基础软件领域从业经验，参加“大规模数据通信交换机”的研制，主持过“分布式数据网络管理中心”“安全电子邮件”“安全路由器”“金融数据加密机（II型）”“安全操作系统”等的研发工作，承担过“****集团信息网络安全工程”“**市电子政务网络信息安全工程”“**市电子政务系统”、“Windows10 神州网信本土化定制与安全分析”等项目的总体方案设计和实施，主笔起草国家标准若干。

报告题目：从哪里来、在哪里、到哪里去——智能时代的网络安全系统工程的一些思考

报告摘要：基于30年信息安全/网络安全从业的经验，回顾网络安全工程的发展历程：单点防御、系统防御、生命周期安全、风险管理、内生安全，梳理现实的网络安全工程体系的成就与挑战，探讨“智能时代”的安全工程的定位、作用、挑战和发展的机遇。



霍 玮

中国科学院信息工程研究所研究员、第十研究室主任，中国网络安全学会副秘书长。长期从事软件漏洞挖掘与利用、软件供应链安全及程序分析研究，主持和参与科研项目40余项。在安全四大顶会IEEE S&P、USENIX Security、ACM CCS、NDSS等发表论文60余篇，获ICSE最佳论文奖、CGO Test-of-Time奖。作为技术负责人，研发国家级漏洞分析与风险评估平台VARAS，累计发现零日漏洞800余个，获微软、谷歌、思科等国际厂商致谢，获得CVE编号400余个。近年来聚焦规模化、智能化漏洞挖掘与风险评估关键技术，培养了一批网络安全高层次人才，团队在GeekPWN、XPWN等赛事中取得优异成绩。

报告题目：从“找漏洞”到“证利用”：逻辑漏洞智能挖掘的范式迁移

报告摘要：随着人工智能驱动漏洞分析能力快速发展，漏洞挖掘正由“规则匹配”迈向“智能推理”。然而，对于身份、权限、租户隔离、业务流程等逻辑漏洞，仅依赖语法模式或大模型生成候选线索，仍难证明漏洞是否真实存在、能否突破真实部署约束形成安全影响，高误报、低可信、难复核成为智能漏洞挖掘面临的核心挑战。本报告提出逻辑漏洞智能挖掘由“找漏洞”向“证利用”的范式迁移，我们认为，逻辑漏洞的本质不是代码错误，而是系统安全语义一致性的破坏，即身份、权限、对象、状态等安全约束在复杂软件系统中的传播、继承或协同过程中发生偏离。围绕这一认识，报告提出“证利用”分析范式，以安全语义推理为基础，以利用证据构建为目标，持续验证漏洞是否真实可达、攻击条件是否满足、防护是否失效以及安全影响是否能够形成闭环，实现漏洞分析由“发现线索”向“可信证明”的转变。



史 记

博士，中国电科网络空间研究院技术创新中心负责人。从事漏洞挖掘工作十余年，学术成果发表于信息安全四大顶会，先后发现微软、谷歌、苹果等厂商产品零日漏洞百余个和厂商公开致谢，3次入选微软年度漏洞挖掘百人榜，入选微软Top Finding Researcher，在Pwn2Own、GeekPwn等国内外主流漏洞赛事中多次成功攻破目标。目前带领团队开展面向闭源软件的智能化漏洞挖掘关键技术与工程化落地，已挖掘数十个主流产品中的认证前 RCE和百余个认证后RCE漏洞。

报告题目：面向网络设备的智能化漏洞挖掘探索与实践

报告摘要：随着网络空间博弈的持续升级，漏洞作为战略资源的地位愈发凸显，而嵌入式设备与边界网设备因数量庞大、更新周期长，始终是攻击者重点关注的突破口。近期，以 Mythos 为代表的 LLM 在漏洞发现领域展现出惊人能力，正在重构传统漏洞挖掘范式。本议题以网络设备漏洞挖掘为题，介绍近年来该技术领域从“蛮荒”到“工业化”，再到“智能化”的发展历程。同时，议题将介绍我们利用 LLM 开展智能化漏洞挖掘的探索和实践方案，方案累计发现数十个主流网络资产产品中的认证前 RCE 0day 漏洞和数百个认证后 0day 漏洞。



杨 力

西安电子科技大学“华山学者”特聘教授，担任计算机科学与技术学院副院长、计算机网络与信息安全国家级实验教学示范中心主任；承担了基础加强计划重点项目课题、重点研发计划项目课题、国家自然科学基金项目等。围绕网络空间主动安全防御、异构无线网络安全接入、典型智能电子系统安全防护等方面进行深入研究，研究成果获省部级科技一等奖3项，参与制定国家技术标准4项，相关成果在合作单位获得应用与验证。

报告题目：安全系统工程视角下基于系统溯源的复杂网络攻击智能检测与分析

报告摘要：随着网络攻击向隐蔽化、持续化、协同化方向发展，传统基于单点告警和局部特征的检测方法难以还原跨主机、跨阶段的完整攻击过程。系统溯源通过记录进程、文件、网络连接等实体之间的交互关系，为复杂攻击行为提供具有因果关联的全局视图。本报告围绕“系统溯源视角下的复杂网络攻击检测与分析”，介绍系统溯源图的基本原理及其在攻击检测、异常行为识别和攻击路径重构中的应用，重点探讨多维时空特征融合、智能操纵攻击下的鲁棒检测，以及面向大规模溯源图的高效攻击链分析方法，并结合相关研究进展与实践系统，讨论系统溯源技术未来的发展方向。

**潘高峰**

潘高峰，教授，博士生导师。近年来，一直从事面向空天通信网络领域的智能安全/隐蔽通信技术等方面研究；以项目主持人身份先后承担了十余项科研项目。以第一作者/通信作者身份在IEEE JSAC、IEEE WCM、IEEE TWC等IEEE权威期刊发表论文50余篇；获省部级奖励三项；入选2024/2025年Elsiver“中国高被引学者”；连续多年入选World's Top 2% Scientists—“终身科学影响力排行榜”及年度科学影响力排行榜；担任IEEE JSAC首席客座编辑、IEEE WCM客座编辑、NSR、IEEE TCOM、IEEE TGCN等多个SCI期刊编辑；指导学生获得“挑战杯”、“互联网+”及国创年会等多项奖励；获2017年国家级大学生创新创业训练计划最佳导师奖。

报告题目：智能时代下的空天安全通信系统发展思路

报告摘要：人工智能重塑空天对抗格局，空天安全通信是全域防御的关键基础。现有系统存在电磁抗扰弱、星地数据处理滞后、跨域协同不足等问题，难以适配复杂博弈场景。本报告结合人工智能技术应用场景，提出新一代空天安全通信发展思路：依托星载轻量化智能载荷实现空间电磁环境感知预测、链路自主调控与数据实时处理；搭建智能波形设计机制，提升抗截获、抗干扰能力；构建空天地动态智能组网技术，打通多平台通信资源协同。相关思路兼顾实战需求与工程落地，为智能空天通信体系建设、装备技术迭代提供可行参考。

融合与重构：AI时代的网络空间安全人才培养、学科建设与工程思维创新论坛

承办单位：清华大学、中国科学技术大学、泉城实验室

🕒 9月18日下午

📍 天鹅湖大酒店二楼1号会议室

论坛主席：徐明伟（清华大学）

杨 波（泉城实验室）

召 集 人：林璟铨（中国科学技术大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:20	《网络安全人才实战能力报告-AI赋能篇》发布 刘建伟（北京航空航天大学）
14:20-14:50	《高等学校网络空间安全专业指导性专业规范》深度解析 刘建伟（北京航空航天大学）
14:50-15:20	AI时代网络安全实战型人才培养 封化民（北京电子科技学院）
15:20-15:50	以课程目标的有效达成保证毕业要求的有效达成 蒋宗礼（北京工业大学）
15:50-16:20	面向未来的数字安全实践人才培养尝试 杜跃进（浙江大学）
16:20-16:35	休 息
16:35-17:05	网络空间安全体系工程设计与建设实践 何健辉（中国电子科技集团有限公司）
17:05-17:35	中国科大网络安全类课程教学：从信息安全专业到网络空间安全学院 薛开平（中国科学技术大学）
17:35-18:05	面向密码专业的量子计算教学探讨 孙思维（中国科学院大学）

论坛主席



徐明伟

徐明伟，清华大学教授，网络科学与网络空间研究院院长，中国通信学会会士、常务理事，国家杰出青年基金获得者，国家自然科学基金委创新群体负责人，国家“万人”计划领军人才，担任国家重点研发计划“宽带通信与新型网络”和“多模态网络与通信”重点专项专家。主要研究互联网体系结构和网络空间安全。



杨 波

杨波，济南大学教授，博士生导师，教育部新世纪优秀人才，CCF杰出教育奖获得者。现任CCF省级学会联席会议主席，ACM SIGCSE China主席，山东计算机学会监事长，山东省泛在智能计算重点实验室主任，泉城省实验室常务副主任。主要从事计算机网络与安全方面的教学科研工作，曾任教育部高等学校计算机类专业教学指导委员会副主任委员、软件工程专业教学指导委员会委员，山东计算机学会理事长，济南大学副校长、临沂大学校长。

召集人



林璟铨

博士，教授、博士生导师，现任中国科学技术大学网络空间学院执行院长，入选国家“万人计划”科技创新领军人才，长期从事网络空间安全研究、尤其是密码技术在计算机和网络系统的创新应用。密码工程技术研究成果已广泛应用于我国各种电子签名电子合同、电子证照、视频加解密等系统。

报告嘉宾



刘建伟

刘建伟，北航网络空间安全学院创院院长，教授，博士生导师，国家级教学名师，享受国务院政府特殊津贴专家。现任国务院学位委员会学科评议组成员、教育部网络空间安全专业教指委副主任、全国密码专业学位研究生教指委委员、中国密码学会常务理事、中国网络空间安全学会常务理事、《网络空间安全科学学报》常务副主编、《网络与信息安全学报》副主编，IEEE高级会员。主编教材16部，专著7部，译著1部。获评国家级教学名师、国家网络安全优秀教师奖、宝钢优秀教师特等奖、北京市教学名师奖、北京市优秀教师等奖励和荣誉。荣获第二届全国优秀教材奖二等奖、全国普通高校优秀教材一等奖、国家网络安全优秀教材奖、普通高等教育国家级规划教材、普通高等教育精品教材、全国优秀科普作品奖、第四届中国科普作家协会优秀科普作品金奖、北京市教学成果一等奖、等国家级和省部级教学成果奖20余项。科研成果荣获国家级和省部级奖励10余项。在国内外期刊/国际会议发表学术论文200余篇，授权技术发明专利120余项。

报告题目：《高等学校网络空间安全专业指导性专业规范》深度解析

报告摘要：报告详细介绍了《高等学校网络空间安全专业指导性专业规范》（下称《专业规范》）的构成与制定原则，明确了网络空间安全专业的培养目标与规格，系统阐述了专业知识体系与实践能力体系，并对三套培养方案进行了深度解析。本报告内容丰富、体系完整、指导性强，可为全国开设网络空间安全专业的高校在人才培养方案制定、课程体系建设等方面提供重要的参考依据。



封化民

封化民，教授、博士生导师，兼任教育部高等学校网络空间安全专业教学指导委员会秘书长。长期从事网络安全和信息化研究工作，特别是面向电子政务信息系统的网络安全相关技术研究，负责和主持完成了10余项国家和省部级网络安全项目，在国内外高水平学术刊物以及会议上发表高水平学术论文120余篇。曾获2017年度国家网络安全优秀人才、2018年获省级科技进步一等奖、2021年度国家科学技术进步奖二等奖等。

报告题目：AI时代网络安全实战型人才培养

报告摘要：随着生成式人工智能、大模型与自主智能体技术全面落地，网络空间攻防范式、情报博弈、认知对抗形态发生颠覆性变革，AI正式成为网络战、情报战、信息战、认知战的核心作战工具。俄乌冲突、委内瑞拉总统被劫持、美以伊战争等网络实战案例，充分印证AI正在重构网络安全威胁格局与攻防逻辑。报告首先给出AI时代我国关键信息基础面临的风险挑战，对网络安全实战型人才培养提出的新要求，结合我国高校网络安全类专业建设情况，给出网络安全实战型人才培养的基本思路；其次，高度关注AI在网络安全教学中引发的教育异化、能力空心化问题。第三，结合教育部等多部委文件精神，探讨AI时代如何加强网络安全实战型人才培养问题。



蒋宗礼

蒋宗礼，博士生导师，全国高校计算机教育研究会理事长，中国工程教育认证协会结论审议委员会委员、计算机类专业认证委员会副主任，获国家级教学名师奖、国家万人计划领军人才等国家级奖十余项，曾任教育部计算机类专业教学指导委员会副主任（2013-2026）、教育部计算机专业分教指委秘书长(2006-2012)、CCF教育专委主任（2008-2012）、CCF教育工委主任（2006-2008）、中国工程教育认证协会学术委员会委员（2012-2017）等。

报告题目：以课程目标的有效达成保证毕业要求的有效达成

报告摘要：现代化强国建设、AI迅猛发展对高等教育提出了新要求，我们必须积极推进课程教学改革，以应对这百年未有之大变局，培养强能力、高素质人才。坚持课程教学融入整个培养体系，确定有效体现毕业要求达成任务的课程目标，体现学生解决复杂工程问题的能力培养的基本定位；以相适应的教学保证课程目标有效达成；提取达成情况的有效评价基础数据，准确评价课程目标的达成情况；选取最具代表性的课程，基于课程目标达成情况评价结果，实现毕业要求达成情况的有效评价。



杜跃进

杜跃进，浙江大学计算机科学与技术学院求是特聘教授/博导，区块链与数据安全全国重点实验室首席研究员，获国家科技进步一等奖两项，并获得新世界百千万人才工程国家级人才，国务院特殊津贴、全国青年岗位能手、信息产业十大杰出青年、CCF杰出会员、亚太信息安全领袖成就奖等荣誉。他的主要研究领域包括网络安全、数据安全、人工智能安全、智能汽车与具身智能安全、工业互联网安全、智慧应急与安全管理等。

报告题目：以业务为中心的智能安全

报告摘要：网络空间安全的概念，随着数据与智能技术的发展以及业务转型而发生变化。本报告探讨提出新时代智能安全的根本逻辑应该是“以业务为中心”，并以智能汽车等行业为例，初步分析这样的逻辑框架下，网络空间安全工作和过去相比有哪些变化，以及未来需要探索的工作。

**何健辉**

何健辉，中国电科总体院网络空间研究室主任，研究员，军事航天、国防科工基础科研与军事电子安全领域专家，长期致力于网络空间攻防理论与体系架构研究、关键技术攻关与装备研制，主攻卫星互联网、卫星通信网络、无人机与集群等攻防领域，主持预研与型号项目20余项。曾获军事科技进步一等奖1项、二等奖1项，省部级科技进步二等奖1项、军队科技成果一等奖1项、三等奖1项，省部级荣誉称号1项；发表论文及授权发明专利20余项。

报告题目：网络空间安全体系工程设计与建设实践

报告摘要：系统阐述了从形势研判到杀伤网构建的方法论与工程路径。梳理了战争形态演进，剖析未来战争智能化、体系化、全域化、无人化特征。介绍了构建网络空间安全体系工程设计方法，提出4域26维模型、OODAL安全链、攻击熵/防御熵及异构冗余等理论方法。在此基础上，设计网络空间多域融合安全体系，涵盖多域融合态势感知、群智协同漏洞挖掘、安全载荷实时生成、多约束智能筹划、异构多径安全投送、在线量化评估等六项关键技术。介绍了网络空间多域融合杀伤网的构建，形成从背景分析、场景设计、能力需求到要素结构、杀伤网及数据体系的完整流程，并建立评估机制，支撑网络空间作战能力体系化生成与持续优化。

**薛开平**

薛开平，中国科学技术大学网络空间安全学院教授、博导，担任网络信息中心主任，安徽省教育和科研计算机网络中心主任。入选国家网信创新人才、教育部青年长江学者、省教学名师、“网络和信息安全”省优秀研究生导师团队负责人。主要研究方向包括网络架构和传输优化、量子网络、数据安全、网络安全防护等。入选 IET Fellow、IEEE Senior Member、CCF 杰出会员，担任 IEEE TIFS、TDSC、TWC、TNSM 等权威期刊编委及多个国际会议或分会程序委员会主席。获 ACM CCS 最佳论文提名奖、IEEE ICC 最佳论文奖等，连续多年入选爱思唯尔“中国高被引学者”及全球前 2% 顶尖科学家榜单，拥有国内外授权发明专利 50 余项。



孙思维

孙思维，中国科学院大学教授，长期从事密码学研究；在Journal of Cryptology和三大密码顶会上发表论文23篇；牵头或参与制定/修订密码相关国家/行业/团体标准7项；主持国家自然科学基金重点项目、国家密码科学基金重点项目、国家重点研发计划青年科学家项目、中共中央直属机关采购项目和国家科技重大专项课题等；获中国密码学会密码创新奖、中国科学院青年科学家奖、中国电子学会青年科学家奖和中国电子学会教学成果大赛一等奖等。

报告题目：面向密码专业的量子计算教学探讨

报告摘要：量子计算理论与技术的快速发展正在深刻影响密码学理论与技术的发展。使密码学专业学生掌握量子计算的基本原理，并具备在量子计算模型下分析密码算法与密码方案安全性的能力，已成为密码人才培养的重要内容。本报告针对密码学专业学生多具有计算机或数学背景、物理基础相对薄弱的特点，结合中国科学院大学相关课程的教学实践，探讨课程内容组织、基础概念讲授及理论与实践衔接等方面的经验，并介绍如何通过循序渐进的课程体系，帮助学生理解量子算法及其密码学影响，掌握必要的分析工具，进而为开展后量子密码和量子攻击等方向的科研工作奠定基础。

报告题目：中国科大网络安全类课程教学：从信息安全专业到网络空间安全学院

报告摘要：中国科大网络空间安全学院可追溯至2002年信息安全专业的成立，报告人为信息安全专业第一批毕业本科生之一。作为这一历程的亲历者，本报告首先回顾中国科大信息安全本科专业课程的设计起源与教学内容的设计思路；其次介绍计算机网络、网络安全协议等课程理论联系实际的教学模式演进，并重点阐述最新的网络空间安全学院与网络信息中心深度耦合的学生培养新模式探索。

智能密码论坛

承办单位：西安电子科技大学、天津大学

🕒 9月18日下午

📍 天鹅湖大酒店一楼徽州厅

论坛主席：沈玉龙（西安电子科技大学）

吕勇强（天津大学）

召集人：杨 力（西安电子科技大学）

鲁赵骏（天津大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	密码与人工智能安全 王美琴（山东大学）
14:40-15:10	侧信道安全——从机密计算到智能计算的新挑战 张殷乾（南方科技大学）
15:10-15:40	基于伽罗瓦环同态和NTRU格的高效密文自举技术 张 江（密码科学技术全国重点实验室）
15:40-16:00	休 息
16:00-16:30	软硬件协同全同态加密算法加速 赵路坦（中国科学院信息工程研究所）
16:30-17:00	面向机密token的密码失效分析与安全增强研究 鲁赵骏（天津大学）
17:00-17:30	智能化安全计算 程 珂（西安电子科技大学）

论坛主席



沈玉龙

沈玉龙，博士。西安电子科技大学“华山学者”领军教授，博士/硕士研究生导师，国家级人才，陕西省网络与系统安全重点实验室主任，教育部创新团队的骨干成员，空天地一体化综合业务网全国重点实验室信息安全中心固定研究人员。长期从事下一代网络安全、云计算与数据安全等研究。近年来主持和承担了国家重点研发计划、国家自然科学基金重点项目等20余项，以主要完成人获国家技术发明二等奖2项，省部级一等奖7项。



吕勇强

清华大学信息国家研究中心副研究员，主要从事处理器漏洞挖掘、安全验证与设计研究。参与“核高基”数字芯片后端设计工具平台研发，研究成果发表于CCS、IEEE TCAD等期刊会议。获2020年中国计算机学会自然科学奖一等奖、2019年中国电子学会科技进步奖一等奖、2015年国家科技进步奖二等奖，2023年入选国家级高层次人才。

召集人



杨力

杨力，西安电子科技大学“华山学者”特聘教授，担任计算机科学与技术学院副院长、计算机网络与信息安全国家级实验教学示范中心主任；承担了基础加强计划重点项目课题、重点研发计划项目课题、国家自然科学基金项目等。围绕网络空间主动安全防御、异构无线网络接入、典型智能电子系统安全防护等方面进行深入研究，研究成果获省部级科技一等奖3项，参与制定国家技术标准4项，相关成果在合作单位获得应用与验证。



鲁赵骏

鲁赵骏博士于2018年获得华中科技大学光学与电子信息学院微电子学与固体电子学工学博士学位。2025年10月至今在天津大学网络安全学院担任英才副教授，主要研究方向为密码工程与芯片安全，在NDSS、DAC、IEEE TC等高水平会议和杂志发表论文20余篇，Google Scholar H-index 21。鲁赵骏博士主持国家自然科学基金青年基金等重要项目，独立负责“密码硬件工程”实践课程体系建设，担任CCF容错专委会执行委员并获得2023年度CCF杰出演讲者。

报告嘉宾



王美琴

王美琴，国家高层次人才计划，现任山东大学副校长、密码与数字经济安全全国重点实验室常务副主任。主要从事对称密码分析与设计的研究，在国际密码五大顶会发表论文59篇，获全国创新争先奖章、国家科技进步一等奖、国家密码科技进步一等奖（省部级）等。

报告题目：密码与人工智能安全

报告摘要：随着人工智能的快速发展，其与密码学的交叉融合已成为当前国际学术研究的前沿热点。对称密码与神经网络在结构上的高度相似性，为二者的深度融合提供了重要理论基础。人工智能在赋能各行各业、带来发展机遇的同时，其自身安全风险也日益凸显，智能安全问题备受关注。基于上述背景，本报告从三个维度展开系统介绍：首先，以智能化密码分析目标，介绍密码算法自动化分析的最新技术进展，深入探讨人工智能在密码分析中的应用与挑战。其次，从密码分析的视角，介绍人工智能安全研究进展，重点探讨神经网络参数恢复等问题的发展方向。最后，围绕人工智能系统，详细介绍数据与模型的安全风险及攻防策略。



张殷乾

张殷乾，南方科技大学教授、ACM杰出科学家。主要研究方向是计算机系统安全和体系结构安全。曾获美国国家科学基金青年科学家奖、中国计算机学会CCF IEEE-CS青年科技奖、北美计算机华人学者协会明日之星奖、国际顶级安全会议ACM CCS时间检验奖、俄亥俄州立大学 Lumley 研究成就奖、连续五年AMiner全球最具影响力的安全和隐私学者提名，入选全球前2%顶尖科学家终身榜单、《麻省理工科技评论》中国2022隐私计算科技创新人物以及2023年度十大“强国青年科学家”。

报告题目：侧信道安全——从机密计算到智能计算的新挑战

报告摘要：在数智化时代，数据安全与隐私保护已成为维系经济运行、保障民生福祉和维护国家安全的战略性要素。侧信道攻击作为一种独特且日益复杂的安全威胁，通过利用系统运行过程中产生的外部可观测信息通道，能够有效窃取系统内部的敏感信息，对数据安全和隐私保护构成了严峻挑战。此类攻击的影响范围已从传统的密码芯片领域，逐步扩展到云计算、机密计算、智能计算等新兴技术领域，呈现出跨平台、跨领域的显著特征。本报告将从侧信道安全这一独特视角切入，系统性地剖析多种计算架构设计中潜藏的安全隐患和相关防御技术。



张江

张江，博士，现任密码科学技术全国重点实验室研究员/博导、副主任，国家青年科学基金项目（A类）、国家优秀青年科学基金项目获得者。主要从事密码学基础理论和密码算法协议设计研究，包括抗量子公钥密码、全同态加密等方向，主持国家网络安全国家科技重大专项、国家重点研发计划等项目，获省部级自然科学/科技进步奖一等奖3项、中国密码学会密码创新奖一等奖、全国密码算法设计竞赛一等奖2项等学术奖励。

报告题目：基于伽罗瓦环同态和 NTRU 格的高效密文自举技术

报告摘要：密文自举是实现全同态加密的唯一技术路径，也是当前全同态加密算法主要性能瓶颈。本报告将首先介绍 BGV、CKKS、TFHE 三种主流全同态加密算法的密文自举现状及难点，然后介绍一种基于伽罗瓦环同态的 BGV 密文自举技术，较同类算法密文自举性能提升 1.2-2.16 倍，最后介绍一种基于 NTRU 格的 TFHE 类同态密文自举技术，较同类算法自举性能提升 2.1 倍，自举密钥规模降低 65%。



赵路坦

赵路坦，中国科学院信息工程研究所研究员，主要研究方向是处理器芯片安全、隐私保护计算加速。在体系结构和安全领域会议和期刊 HPCA、MICRO、ASPLOS、CCS、SP 等发表论文 30 余篇，其中 CCF A 类 18 篇，研究工作曾获得 A 类期刊 IEEE TC/TPDS 期刊亮点论文、APPT 2021 最佳论文奖。受邀担任 CCF A 类会议 HPCA 和期刊 IEEE TC/TDSC、ACM TACO 等审稿人，担任《Frontiers of Computer Science》《信息安全学报》学术期刊青年编委。

报告题目：软硬件协同全同态加密算法加速

报告摘要：全同态加密是密态计算的核心密码原语，为保障数据要素安全流通，实现数据“可用不可见”提供了关键技术支撑。然而其严重的性能瓶颈阻碍了全同态加密技术的规模化应用。本报告将系统梳理全同态加密计算的性能优化路径，探讨软硬件协同加速、安全协议优化等关键技术，探索在 AI 隐私推理、隐私信息检索等典型应用场景的全同态加密计算加速方法。



鲁赵骏

鲁赵骏博士于 2018 年获得华中科技大学光学与电子信息学院微电子学与固体电子学工学博士学位。2025 年 10 月至今在天津大学网络安全学院担任英才副教授，主要研究方向为密码工程与芯片安全，在 NDSS、DAC、IEEE TC 等高水平会议和杂志发表论文 20 余篇，Google Scholar H-index 21。鲁赵骏博士主持国家自然科学基金青年基金等重要项目，独立负责“密码硬件工程”实践课程体系建设，担任 CCF 容错专委执行委员并获得 2023 年度 CCF 杰出演讲者。

报告题目：面向机密 token 的密码失效分析与安全增强研究

报告摘要：机密计算依托可信执行环境在处理器内部构建硬件隔离运行域，为使用态数据提供机密性与完整性保障。在云端大模型调用场景中，用户提示词、上下文与私有数据面临云服务商及同驻攻击者的窃取风险。机密 Token 技术将用户敏感数据在 TEE 内完成加密，云端模型仅接触密文态，明文数据始终不出 TEE 边界，从而实现端到端隐私保护的大模型调用。然而，预测执行、缓存共享等微架构优化缺乏安全协同设计，由此产生的隐蔽通道与故障注入窗口可导致机密 Token 密码保护机制失效。本报告系统揭示机密 Token 场景中密码失效的底层机理，构建可迁移威胁模型、覆盖率导向的模糊测试方法及可配置安全增强方案，弥合从数学安全到微架构实现安全的鸿沟。



程珂

程珂，西安电子科技大学副教授，博士/硕士生导师，入选中国科协青年人才托举工程，担任开源密码社区openHiTLS首届技术委员会委员、中国中文信息学会大数据安全与隐私计算专委会委员、ACM CCS程序委员会委员等。主要研究方向包括密码与数据安全、隐私保护大语言模型等，以第一/通信作者在CCS、INFOCOM、TIFS、TDSC、计算机学报等发表论文30余篇。获国际隐私计算竞赛iDASH冠军、GitCode“年度十大开源人物”等荣誉奖励。

报告题目：智能化安全计算

报告摘要：为推动密态数据安全计算向智能化设计、实现与验证方向发展，本报告围绕安全多方计算的评测、生成与证明三个环节，介绍团队在基准评测体系、MPC 智能体和协议安全性证明方面的研究进展。首先介绍覆盖基础计算能力与典型应用场景的安全多方计算基准评测体系，为大模型及智能体的能力评估提供统一的任务、数据与指标；随后介绍基于团队自研通用密态数据安全计算库 NssMPCLib 构建的 MPC 智能体，支持对计算需求的自动理解、方案生成与代码生成；最后介绍大模型驱动的协议安全性分析与证明方法，探索提升安全计算协议分析与验证自动化能力的新路径。

空天地一体化安全论坛

承办单位：西安电子科技大学、电磁空间安全全国重点实验室

⌚ 9月18日下午

📍 天鹅湖大酒店三楼1号会议室

论坛主席：杨卫东（西安电子科技大学）

王 巍（中国电子科技集团公司第三十六研究所）

召 集 人：杨卫东（西安电子科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:45	多轨道跨频段卫星电磁频谱智能认知 杨 凯（北京理工大学）
14:45-15:20	星链发展与通信安全威胁分析 韩 煜（中国电子科技集团公司第三十六研究所）
15:20-15:55	空地协同的低空多模态监视技术与应用 董 超（南京航空航天大学）
15:55-16:30	休 息
16:30-17:05	智驭低空，安全先行：浅谈 AI 驱动的城市级低空安全体系与实践 王 聪（杭州安恒低空安全科技有限公司）
17:05-17:40	导航安全防护技术发展现状与前沿挑战 王永庆（解放军空军工程大学）
17:40-18:00	数智赋能，安全护航——中国电信低空能力体系与实践 付 杰（中国电信安全公司）

论坛主席



杨卫东

西安电子科技大学华山特聘教授，博士生导师，省学术技术带头人、省自然科学杰出青年基金获得者，本、硕、博毕业于西安电子科技大学，中国科学院大学博士后，美国奥本大学访问学者，现任西安电子科技大学杭州研究院卫星互联网研究所所长。在物联网应用与安全方向主持和参加国家“863”计划、国家重点研发计划、国家自然科学基金、国家“粮安工程”等项目20余项，获省部级科技奖7项（第一完成人3项），发表高水平学术论文70余篇（中科院1区论文7篇，SCI检索论文25篇，ICC 2022最佳论文奖1项，最佳论文提名1项，高被引论文2篇），获授权国际、国家发明专利20余项，制修订国家、行业标准5项，编写中国科协学科发展报告2次，参编学术专著4部，在国内外学术会议上做学术报告20余次。



王 巍

王巍（1980-），男，博士，研究员。2008年毕业于西安电子科技大学，现任中国电科三十六所技术创新部主任、通信信息控制与安全重点实验室常务副主任，浙江省万人计划青年拔尖人才，中国电科一级重大任务总师，中国电科网络安全论证专家组成员，计算机网络安全技术/装备集团专家，某部网电领域专家组成员，国防科技工业网络安全领域技术专家组成员，长期从事无线网络控制、网络安全方向的研究。承担多项型号、预研、863、创新特区、集团主导创新基金、重点实验室基金项目，总计经费4000余万元，取得了多项研究成果，五次获省部级奖项。在国内外多个知名刊物发表SCI、EI论文四十余篇，授权申请专利十余项，出版著作3部，通信学报、电子学报、计算机学报、北京航空航天大学学报等国内知名期刊的审稿人。

召集人



杨卫东

西安电子科技大学华山特聘教授，博士生导师，省学术技术带头人、省自然科学杰出青年基金获得者，本、硕、博毕业于西安电子科技大学，中国科学院大学博士后，美国奥本大学访问学者，现任西安电子科技大学杭州研究院卫星互联网研究所所长。在物联网应用与安全方向主持和参加国家“863”计划、国家重点研发计划、国家自然科学基金、国家“粮安工程”等项目20余项，获省部级科技奖7项（第一完成人3项），发表高水平学术论文70余篇（中科院1区论文7篇，SCI检索论文25篇，ICC 2022最佳论文奖1项，最佳论文提名1项，高被引论文2篇），获授权国际、国家发明专利20余项，制修订国家、行业标准5项，编写中国科协学科发展报告2次，参编学术专著4部，在国内外学术会议上做学术报告20余次。

报告嘉宾



杨 凯

北京理工大学信息与电子学院副院长、教授、博士生导师，国家一级重点学科“信息与通信工程”、国家一流本科专业“通信工程”学科专业责任教授。主要从事空天信息网络、卫星通信等领域的研究。主持国家自然科学基金、国防基础科研重点项目等多项科研课题，相关成果已成功应用于多个型号装备。出版学术专著7部，发表学术论文100余篇。

报告题目：多轨道跨频段卫星电磁频谱智能认知

报告摘要：卫星通信系统是突破时空疆界限制、实现全球泛在覆盖的国家重大基础设施，直接关乎空天地信息优势与全球行动主导权。在全球卫通星座加速布局推进、空间电磁博弈持续加剧的背景下，卫星通信系统具备多轨道协同、多业务共存和全开放空间的新特点，卫星电磁频谱认知面临轨道频率跨度极大、信号体制复杂多样、电磁空间对抗激烈等挑战。本报告围绕多轨道、跨频段卫星电磁频谱智能认知，重点剖析高轨通信卫星与低轨巨型星座的载波运行状态、信号传输体制与时频动态特征，构建“多维协同感知—多域精准认知—态势全景分析”的技术链路。最后立足全局，研判未来空天地电磁频谱智能认知的演进趋势。



韩 煜

主要从事通信电子战、电子对抗领域总体研究与技术创新工作，主持或参与过多项国防预研、航天基金、军口863等基础前沿性课题，支撑多项国家通信对抗装备型号项目研制任务。已发表学术论文20余篇，申报专利10余项。获国防科技进步二等奖一次、中电集团科技进步二等奖一次，中电集团技术发明三等奖一次、军队科技进步一等奖一次。

报告题目：星链发展与通信安全威胁分析

报告摘要：星链计划已发展为全球规模最大的低轨卫星互联网星座，正从商业通信向手机直连、太空AI计算等方向加速演进。星链带有军民模糊属性，已深度嵌入美国军事体系。俄乌冲突与伊朗战争中，星链通信在宽带韧性通信、传感器情报传输、无人系统指控等多方面的优异表现已得到实战验证。星链计划“军民融合”模式和“商业速度”的迭代能力，正在深刻改变现代战争形态、全球通信产业格局与太空安全秩序，本报告一方面从星链卫星、星链通信起因，演进，军民属性，潜在应用等多维度说明星链近期的发展趋势，另一方面，从对我国军事、通信安全防御等方面，分析星链计划带来的可能潜在威胁。

**董超**

南京航空航天大学电子信息工程学院副院长，教授，博士生导师，国家级人才，国防科技创新团队带头人，工信部电磁频谱空间认知动态系统重点实验室副主任。担任中国通信学会低空信息通信专委会副主任委员、中国电子学会通信分会委员。主持国家重点研发计划、国家自然科学基金重点、卓越青年基金等项目与课题，发表高质量学术论文百余篇，获最佳论文奖励7次，授权美国、国家与国防发明专利近50项，获国防科技进步一等奖、发明协会创业创新一等奖、军事科学技术进步二等奖、日内瓦国际发明展金奖等科技奖励。

报告题目：空地协同的低空多模态监视技术与应用

报告摘要：包含发现、识别、定位与跟踪的低空监视技术是低空经济安全健康发展的关键保障，以往的低空监视以各类地面设施为主，监视范围有限，识别准确率、定位精度以及跟踪速率均与低空经济的安全要求存在一定差距。基于空地认知通信组网技术，报告人团队实现了空地协同的低空多模态监视，融合了雷达、频谱、光电等多模态监测手段，在地面监测设施的基础上，扩展了空中协同无人机进行定位与跟踪，共同组成了覆盖空间更广的监视网络，识别、定位以及跟踪等性能指标也有较大提升，研究成果应用于飞行密集区域以及航路要地等场景的低空安全防护。

**王聪**

中共党员，电子与通信工程硕士，高级工程师，安恒信息低空安全副总经理。杭州市E类人才、高新区（滨江）专业人才，浙江省科技厅专家库专家，教育部研究生教育评估监测专家库专家。深耕网络安全与低空安全多年，首创低空安全一体化防护理念，2024年牵头研发行业首款低空智能巡处一体化平台，同时带领团队研发基于AI的恒御低空安全管理平台，产品落地全国多省份城市。曾主导或参与十余项国家与行业安全标准，牵头编制物联网安全白皮书与无人机系统安全白皮书，拥有多项技术专利，长期深耕低空安全产学研融合与城市低空治理实践。

报告题目：智驭低空，安全先行：浅谈AI驱动的城市级低空安全体系与实践

报告摘要：当前低空经济加速规模化发展，黑飞扰航、违规飞行、空域管控难等安全风险持续凸显，传统手段已难以满足城市全域低空治理需求。本次分享结合安恒信息全链条低空安全业务布局，围绕“智驭低空，安全先行”核心思路，分享AI技术在多源数据融合感知、低空目标风险智能研判、防控预案智能推荐、智能联动处置、智能分析复盘、智能报告生成等关键环节的落地路径。结合多地城市低空安防实战案例，分享安恒基于AI构建全域、实时、可追溯的城市级低空安全防护体系架构与落地实践。


王永庆

空军工程大学教授，某国家级实验教学示范中心主任。担任装发某专委会委员，空军某行业领域专家。研究方向包括智能导航与协同控制，军事智能导航与定位，航空装备控制、检测与导航技术、无人反无人导航技术等。主持和参加国家、军队、省市项目20余项。获省级教学成果二等奖1项，军队科技成果一、二、三等奖各1项，获评国家一流课程1门。

报告题目：导航安全防护技术发展现状与前沿挑战

报告摘要：当前，导航系统已成为国家关键基础设施，其提供的时空信息成为国防联合作战、数字经济体系的基础性战略资源。以北斗系统为核心的全球卫星导航定位授时（PNT）体系，深度融入能源、金融、通信、交通、航空航天、智能无人系统及全域作战链条。导航安全面临压制式干扰、转发式/生成式欺骗干扰、星上链路攻击、时空数据篡改、网络协同攻击等五类核心威胁。报告从民用基础设施安全、军用战场作战效能两大维度剖析风险影响，梳理主流导航安全防护技术，提出高逼真智能欺骗干扰精准建模与识别、主瓣同向欺骗干扰防御、高动态弱信号场景“抗扰-捕获-跟踪”、多源异构PNT信息智能容错融合、天地一体化导航安全监测、溯源、博弈体系等前沿挑战。


付杰

付杰，中国电信安全公司云猎产线总经理、电信集团高级专家，研究方向涉及高空卫星无线通信安全、低空无人机安全、地面基站通信安全等，多项研究成果被国内外 Black-Hat、HITB、CansecWest、补天等知名安全会议录用。多次主导及参与国家重点项目、省级重点项目、公司核心项目，在核心技术攻关、项目落地执行等方面拥有扎实的实践成果与丰富经验。

报告题目：数智赋能，安全护航——中国电信低空能力体系与实践

报告摘要：付杰分享围绕低空经济发展与安全保障需求，介绍中国电信在低空网络、云网平台和综合服务方面的能力，重点分享天翼安全“云猎”无人机侦测反制产品及行业应用实践，并结合实践探讨低空安全领域的创新方向，助力低空产业安全发展。

具身智能安全论坛

承办单位：北京邮电大学、中国科学院自动化研究所

⌚ 9月18日下午

📍 天鹅湖大酒店三楼11号会议室

论坛主席：张兆翔（中国科学院自动化研究所）

何召锋（北京邮电大学）

召 集 人：何召锋（北京邮电大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	具身智能系统的物理安全增强 何 潇（清华大学）
14:40-15:10	混合AI驱动的具身智能技术进展 孙宏滨（西安交通大学）
15:10-15:40	具身智能时代的风险与治理 王贺升(上海交通大学)
15:40-16:00	休 息
16:00-16:30	风险世界反演生成——具身智能安全探索与实践 刘祥龙(北京航空航天大学)
16:30-17:00	软硬一体，原生护航——打造具身智能全栈防护体系 孔德智（工业和信息化部电子第五研究所）
17:00-17:30	具身智能物理安全风险评测与优化 任 民（北京邮电大学）

论坛主席



张兆翔

中国科学院自动化研究所研究员、博士生导师，国家级人才，中国科学院大学岗位教授，自动化所模式识别实验室常务副主任，国际模式识别学会会士 (IAPR Fellow)。在 IEEE T-PAMI、IJCV、JMLR、CVPR、ICCV、ECCV 等顶刊顶会发表论文 200 余篇，申请发明专利 60 余项。承担国家自然科学基金重点项目、重点国际合作项目、企业联合重点支持项目、国家重点研发计划课题等多项国家级项目。获北京市科学技术奖科技进步一等奖、中国人工智能学会吴文俊技术发明奖二等奖。曾担任 IEEE T-BIOM、IEEE T-CSVT、Pattern Recognition、NeuroComputing、Frontiers of Computer Science 编委，担任 CVPR、ICCV、AAAI、IJCAI、ACM MM、ICPR、ACCV 等国际会议的领域主席 (Area Chair)。



何召锋

北京邮电大学教授、博士生导师，卓越工程师学院副院长。在 IEEE T-PAMI、CVPR、ICCV、ACL 等顶刊顶会发表论文 120 余篇，申请发明专利 120 余项，牵头制定国家标准、行业标准、国际标准 40 余项。主持国家重点研发计划项目、国家自然科学基金、北京市中央引导地方专项等国家级和省部级科研项目 20 余项。入选北京市智源学者、北京市青年拔尖创新团队等人才项目，获北京市科技进步一等奖、吴文俊人工智能科学技术科技进步二等奖、中国图像图形学会技术发明二等奖、中国产学研合作促进会创新成果奖一等奖、中国计算机学会优秀博士论文奖、中科院优秀博士论文奖。

召集人



何召锋

北京邮电大学教授、博士生导师，卓越工程师学院副院长。在 IEEE T-PAMI、CVPR、ICCV、ACL 等顶刊顶会发表论文 120 余篇，申请发明专利 120 余项，牵头制定国家标准、行业标准、国际标准 40 余项。主持国家重点研发计划项目、国家自然科学基金、北京市中央引导地方专项等国家级和省部级科研项目 20 余项。入选北京市智源学者、北京市青年拔尖创新团队等人才项目，获北京市科技进步一等奖、吴文俊人工智能科学技术科技进步二等奖、中国图像图形学会技术发明二等奖、中国产学研合作促进会创新成果奖一等奖、中国计算机学会优秀博士论文奖、中科院优秀博士论文奖。

报告嘉宾



何 潇

清华大学自动化系教授、国家杰青。现任清华大学自动化系副主任、清华大学具身智能与机器人研究院副院长。研究方向为故障诊断与安全控制、具身智能系统安全性。发表国际期刊/会议论文300余篇，获得国家发明专利40余项。研究成果获中国自动化学会自然科学奖一等奖、中国自动化学会技术发明一等奖、北京市自然科学二等奖。现任中国自动化学会副秘书长，北京自动化学会监事长，中国自动化学会故障诊断专委会副主任兼秘书长、中国自动化学会过程控制专委会副主任，中国指挥与控制学会具身智能专委会副主任。担任IEEE TNNLS等多个国际期刊的编委。培养4位博士获得中国自动化学会“优博”。

报告题目：具身智能系统的物理安全增强

报告摘要：具身智能系统是典型的复杂动态系统，具有多子系统耦合、多感知-决策闭环、强非线性与强不确定性及与开放环境交互等挑战。针对上述问题，构建了面向具身智能系统的实时安全增强技术体系，涵盖动态系统状态估计、故障诊断、容错控制与安全性评估等关键环节：提出了网络化强跟踪非线性状态估计新方法；发展了基于辅助信号激励的主动诊断理论；揭示了故障、部件性能与系统结构对一致性与安全裕度的影响机理；提出了动态系统的实时安全性评估新方法。相关成果显著增强了具身智能系统的运行安全。



孙宏滨

国家级特聘教授，现担任西安交通大学人工智能学院副院长、人机混合增强智能全国重点实验室副主任。研究方向为嵌入式视觉、自主智能系统与智能芯片。围绕上述领域的国际学术前沿、国家重大需求，共发表国际顶级期刊/会议论文120篇，其中4篇论文获最佳（或提名）论文奖；获国家发明专利43项，集成电路布图专有权6项，主持国家级科研项目14项；获得国家技术发明二等奖、国家教学成果一等奖、陕西省自然科学一等奖、中国自动化学会自然科学一等奖、求是人工智能科教奖青年奖等。

报告题目：混合AI驱动的具身智能技术进展

报告摘要：具身智能作为连接人工智能与物理世界的核心范式，是实现通用人工智能的关键路径，而混合 AI 架构正成为驱动其技术跃迁的核心动力。本报告围绕混合 AI 驱动的具身智能技术进展展开系统性分析，分析了开放环境泛化能力弱、物理交互鲁棒性不足、多模型协同时延高等核心瓶颈，重点剖析多类型AI模型构建鲁棒、高效具身智能系统的必要性，并介绍团队在基于思维链VLA大模型、视觉语言导航、运动轨迹优化方法等方面的最新进展。最后，汇报混合 AI 驱动下具身智能在智能驾驶、机器人、低空智能装备等领域的典型应用成效，展望了未来技术发展趋势与产业化落地路径。


王贺升

博士，上海交通大学唐君远讲席教授，浦江国际学院院长。担任中国自动化学会混合智能专委会副主任，中国仪器仪表学会智能车与机器人分会副主任。现/曾担任国际期刊TRO, TMECH, TASE, RAL, IJHR, RIA的编委，IEEE/ASME Transactions on Mechatronics的高级编辑，Advanced Intelligent Systems的顾问编委，Robot Learning的主编。作为项目负责人，主持包括国家重点研发计划项目，国家自然科学基金杰出青年基金、优秀青年基金、联合基金重点等多个项目。获得全国宝钢优秀教师称号、上海市青年科技启明星、上海市曙光计划等。担任机器人顶会IROS 2025的大会主席以及RCAR 2016和ROBIO 2022的大会主席。

报告题目：具身智能时代的风险与治理

报告摘要：本报告着眼于人形机器人与具身智能发展所带来的风险与治理问题，首先对具身智能从信息系统走向行动系统的技术趋势、应用背景与治理挑战进行概述。其次分析了人形机器人进入真实场景后可能带来的安全、数据、交互与责任等方面的新问题，指出其治理不能仅依赖传统机器人安全方法，而需要在技术发展、应用场景与社会接受之间进行综合考量。报告进一步围绕具身智能未来应用中的治理逻辑与实践路径展开探讨，尝试提出面向社会部署的总体思路，为人形机器人从样机验证走向实际应用提供参考。


刘祥龙

北京航空航天大学计算机学院党委书记、教授、博士生导师，国家高层次人才，国家重点研发计划项目首席科学家。长期从事安全可信人工智能的教学和科研，负责《人工智能安全导论》、《数据结构与程序设计》等本科和研究生国家级课程，发表TPAMI、IJCV、CVPR、ICLR等人工智能领域国际顶级期刊/会议论文160余篇，主持国家自然科学基金青年A类和B类项目、科技创新2030重大项目、国家重点研发计划项目等国家级项目。担任IEEE TNNLS、Pattern Recognition、《智能安全》等期刊编委，ACL、AAAI等国际会议领域主席。连续入选爱思唯尔中国高被引学者，获中国仿真学会科学技术一等奖、陕西省自然科学一等奖、中国电子学会青年科学家奖、北京市科技新星等。

报告题目：风险世界反演生成——具身智能安全探索与实践

报告摘要：无人驾驶等具身智能系统广泛应用，成为国民经济和社会发展广泛依赖的关键信息基础设施。但近年来这类系统在实际应用中面临着对抗攻击、噪声干扰等大量长尾安全风险，常规仿真测试手段难以有效应对。亟待突破智能安全风险检测和推演范式，建立强迁移、高逼真、可溯源的物理世界风险推演体系，探索明晰智能安全边界、演化过程和干预机制。本报告将围绕上述目标，报告团队在无人驾驶等具身智能安全攻防演练的初步探索，重点介绍提出的对抗反演生成方法和世界模型，及其在物理世界风险生成、安全评测及仿真演练等方面的应用实践，展望风险世界模型技术发展趋势。


孔德智

高级工程师，硕士毕业于西安电子科技大学通信与信息系统（网络安全）专业，现任工业和信息化部电子第五研究所人工智能中心主任。粤港澳大湾区生成式人工智能安全发展联合实验室理事、广东省人工智能产业协会理事、广东省通信行业协会网络安全专业委员会副主任委员。主要从事AI软硬件适配验证、模型测评、高质量数据集测评、智能体测评、具身智能测评、AI安全测评等关键技术研究。牵头研究AI软硬件适配验证、模型测评、高质量数据集测评、智能体测评、具身智能测评、AI安全测评等省部级课题十余项，组织支撑国产计算芯片适配中心建设方案、具身智能训练场建设方案、《人工智能安全治理框架2.0版》《政务大模型应用安全规范》《网络安全技术 人工智能安全测评机构能力基本要求》《网络安全技术 端侧大模型网络安全指南》《人工智能安全能力成熟度评估方法》《人工智能技术 涉及未成年人应用安全指南》《低空智能网联体系参考架构》和《低空安全技术体系白皮书》等材料编制。牵头组织南方电网大瓦特模型测评、某高校大模型系统测评、平安产险AI安全评估、中国移动“九天”大模型测评、中电科智能应用集成开发平台测评、华为盘古大模型评测、某证券企业智能体安全测试及AI体系建设等业务实施。

报告题目：软硬一体，原生护航——打造具身智能全栈防护体系

报告摘要：具身智能正加速从数字空间迈向物理世界，然而感知偏差、决策失稳与执行故障极易在“感知—决策—执行”链路中耦合传导，加之行业面临评测标准缺失、安全边界模糊、虚实结果割裂等挑战，其规模化落地仍存隐忧。本次汇报将系统介绍“软硬一体，原生护航”具身智能全栈防护体系的建设实践，聚焦硬件、系统、算法与生态的协同防护，构建覆盖“大脑—小脑—本体—跨域安全”的评测与治理框架。汇报将围绕体系建设背景、核心能力与阶段成果展开，分享如何通过五级能力阶梯，贯通物理、数字、模型、行为等多维安全指标，推动安全管控能力在模型训练、产品验证、现场运行与迭代演进全流程中快速落地，加速具身智能从“可用”向“可靠”的产业价值转化。


任 民

北京邮电大学特聘研究员，博士生/硕士生导师，主要研究方向为具身智能与智能体安全。发表TPAMI、IJCV、TIFS、CVPR、ICCV等顶级期刊会议论文十余篇，获得博士后创新人才支持计划支持，主持国家自然科学基金青年项目、博士后面上项目。受邀担任IJCV、TIP、TIFS、MIA、TCSVT、CVPR、ICCV、AAAI等国际知名期刊会议审稿人，入选北京市科协2024年度青年人才托举计划。

报告题目：具身智能物理安全风险评测与优化

报告摘要：具身智能体作为大模型与物理世界交互的核心载体，正成为全球AI竞争的战略高地。然而，其在真实场景中面临的物理安全风险日益凸显。当前研究多聚焦于数字空间的安全问题，缺乏对物理交互过程中系统性风险评估与对齐方法的深入探讨。本报告将介绍团队搭建的具身智能安全系统性评测与优化框架：包括涵盖空间、行为、姿态及鲁棒性四维度的量化评估准则；面向交互安全的评测场景自动化搭建；规则与模型双驱动的风险识别与量化；基于世界模型的动作规划增强。

感知AI安全：从数字域到物理域论坛

承办单位：合肥综合性国家科学中心信息安全研究中心

🕒 9月18日下午

📍 天鹅湖大酒店一楼黄山厅

论坛主席：胡以华（国防科技大学）

杨 星（国防科技大学）

召 集 人：杨 星（国防科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	具身智能感知安全 胡鹏飞（山东大学）
14:40-15:10	面向感知系统的多模态大模型安全：评估方法和验证平台 李松泽（东南大学）
15:10-15:40	端侧智能体全链路资产保护 张旭鸿（浙江大学）
15:40-16:00	休 息
16:00-16:30	无人系统视觉感知安全 姜廷松（军事科学院国防科技创新研究院）
16:30-17:00	通信信号智能识别的物理世界对抗攻击 王 翔（国防科技大学）
17:00-17:40	圆桌研讨：感知AI的物理域安全挑战与应对

论坛主席



胡以华

国防科技大学电子对抗学院教授，合肥综合性国家科学中心信息安全研究中心主任，中国电子学会会士，中国光学学会会士，中国光学工程学会常务理事，安徽省科学技术协会兼职副主席。长期从事光电对抗与卫星网络安全研究，创建了国内首个航天电子对抗专业，获国家技术发明二等奖2项，安徽省重大科技成就奖1项，省部级一等奖9项。获授权发明专利80余项，出版中英文专著5部，主编著6本，主要论文300余篇。



杨 星

国防科技大学教授、博士生导师，合肥综合性国家科学中心信息安全研究中心副主任，国防科技大学合肥前沿交叉中心主任，享受国务院政府特殊津贴，获安徽省创新争优奖章，安徽省卓越工程师奖。军队学科拔尖人才，国家级青年人才。军事智能学科方向首席专家，主要从事网电空间智能安全研究，主持国家重点研发、173重点等项目13项。获省部级科技进步一等奖3项，一级学会技术发明一等奖3项，发表论文100余篇，出版专著2部，授权发明专利60余项。

召集人



杨 星

国防科技大学教授、博士生导师，合肥综合性国家科学中心信息安全研究中心副主任，国防科技大学合肥前沿交叉中心主任，享受国务院政府特殊津贴，获安徽省创新争优奖章，安徽省卓越工程师奖。军队学科拔尖人才，国家级青年人才。军事智能学科方向首席专家，主要从事网电空间智能安全研究，主持国家重点研发、173重点等项目13项。获省部级科技进步一等奖3项，一级学会技术发明一等奖3项，发表论文100余篇，出版专著2部，授权发明专利60余项。

报告嘉宾



胡鹏飞

山东大学计算机科学与技术学院教授/博导、副院长，国家优青获得者。于美国UC Davis获得博士学位，主要研究兴趣涉及网络与系统安全、具身智能安全、智能感知。目前为CCF网络与系统安全专委会、物联网专委会、网络与数据通信专委会执行委员。先后主持国家优青项目、国家重点研发计划课题、工信部重大专项课题等产学研项目10余项。在S&P、USENIX Security、CCS、NDSS、INFOCOM、IEEE TMC等国内外知名会议/期刊上发表学术论文100余篇。担任IEEE TDSC、IEEE TWC等SCI一区期刊编委，和 USENIX Security、INFOCOM、WWW、ICDCS等多个知名国际学术会议的的程序委员会委员。

报告题目：具身智能感知安全

报告摘要：具身智能依赖多模态感知系统完成对物理世界的建模和理解，然而感知系统易受对抗样本和跨模态欺骗攻击干扰，导致认知偏差和决策错误。在这一背景下，跨域信任边界的研究成为具身智能系统安全中的关键问题。本次报告将以典型案例为切入点，揭示跨域安全威胁的内在机理，并提出跨域信任边界的概念与防护思路，为具身智能系统安全设计提供新的视角。



李松泽

东南大学青年首席教授、博士生导师，区块链应用监管教育部工程研究中心执行主任；国家高层次青年人才，江苏特聘教授，华为紫金青年学者。博士毕业于美国南加州大学，曾任斯坦福大学研究员，香港科技大学助理教授。研究方向为人工智能安全与隐私，大模型与智能体安全，隐私保护的机器学习。在安全四大顶会 (IEEE S&P、USENIX Security、ACM CCS、NDSS)，人工智能顶会 (ICML、NeurIPS、ICLR、CVPR) 等发表论文 90 余篇。获 NeurIPS 2020 联邦学习研讨会最佳论文奖，斯坦福全球前 2% 顶尖科学家榜单 (2021-2025 连续 5 年)，ACM SIGSAC 中国 2025 年度新星奖。

报告题目：面向感知系统的多模态大模型安全：评估方法和验证平台

报告摘要：近年来，大语言模型技术和应用飞速发展，以ChatGPT和DeepSeek为代表的大模型引起广泛关注。大语言模型展现出强大的数据处理和逻辑推理能力，但同时也面临着多样的安全和隐私风险。其不仅面临传统人工智能模型中的对抗攻击、数据投毒等挑战，还受到提示词后门、越狱攻击、个人敏感信息泄露等新型威胁。本次报告将介绍课题组近期关于大语言模型、多模态大模型、大模型智能体面向感知场景的安全评估和加固工作，包括大模型越狱、对抗、后门新型攻击方法，大模型幻觉检测，大模型评审的合理性评估，大模型成员推理和数据重构攻击方法等。最后，将介绍课题组研发的大模型安全评估与加固系统平台OmniTrust，以实现精细化、全方位、持续进化的大模型安全测评与安全性增强。



张旭鸿

浙江大学软件学院研究员、院长助理，入选国家级青年人才计划、浙江省高层次人才计划。致力于可信人工智能、空间智能与具身智能研究，主持国家重点研发计划课题、国家自然科学基金项目、工信部高质量专项课题、华为胡杨林基金项目等20余项科研项目。发表领域高水平论文50余篇，获安徽省科技进步一等奖、ACM CCS 2021最佳论文奖、特色化示范性软件学院联盟关键软件技术创新成果奖等。曾任职美国LinkedIn，主导MLOps平台研发及产品应用。

报告题目：端侧智能体全链路资产保护

报告摘要：随着物理AI快速发展，智能体正从云端走向终端。出于隐私保护、实时响应、网络可靠性与推理成本等考虑，端侧部署将成为智能体规模化落地的重要形态，但也使模型能力、提示词策略与私有知识库直接暴露于不可信运行环境，面临模型窃取、提示词迁移和知识抽取等风险。本报告围绕端侧智能体全链路资产保护展开，分别从基础模型能力保护、不可移植提示词机制和RAG知识库双路径防御三个层面，构建覆盖“模型—提示词—知识库”的一体化防护体系，为端侧智能体安全部署与资产保护提供系统方法。



姜廷松

军事科学院国防科技创新研究院副研究员。长期从事对抗机器学习、智能感知安全和具身智能安全方向研究工作，主持和参与多项重点课题，获中国发明学会发明创业奖创新奖一等奖、授权发明专利10余项，在TPAMI、TIP、CVPR、ICCV、AAAI等国际会议和期刊发表高水平论文40余篇，开源全身涂装式对抗纹理迷彩生成框架FCA并被红山开源等科技媒体报道。

报告题目：无人系统视觉感知安全

报告摘要：图像识别、目标检测等视觉感知算法在无人驾驶、自主导航等领域的广泛应用，现有研究表明深度学习算法易受对抗样本攻击，使得视觉感知算法存在不确定性和安全风险。本报告从视觉感知算法安全出发，重点分析了安全风险来源、对抗样本攻击理论和对抗样本生成与评测方法，最后对无人系统感知安全进行展望。

**王 翔**

国防科技大学副教授，博士生导师。电子学会电抗分会委员，科技委某领域专家，长期从事非合作信号处理领域科研工作，先后获得省部级科技进步奖一等奖2项，电子学会自然科学二等奖1项，全军级成果评定一等奖3项，发表了论文70余篇，其中IEEE会刊15篇，入选ESI 1%高被引和ESI 1‰热点3篇。

报告题目：通信信号智能识别的物理世界对抗攻击

报告摘要：深度学习正推动通信信号智能识别技术走向实用，但其脆弱性和安全性问题突出。本报告聚焦通信信号识别模型面临的对抗攻击问题，探讨对抗攻击从数字空间向物理空间跨越时的核心挑战，即对抗波形生成必须纳入真实信道与信号处理的影响。报告首先阐明数字域对抗样本直接发射后，因信道畸变与接收端滤波器作用，攻击效能将严重衰退。随后，提出物理域攻击建模框架，将信道效应、接收端匹配滤波等关键环节纳入对抗扰动生成过程，实现物理世界约束条件下的通信信号对抗样本波形级优化。融合信道和滤波约束的对抗波形可在典型信噪比下保持高攻击成功率，揭示了物理世界攻击的现实威胁，为构建鲁棒且安全的智能通信信号识别系统提供了新视角。

区块链与可信智能体网络论坛

承办单位：北京理工大学

🕒 9月18日下午

📍 天鹅湖大酒店一楼淮河厅A

论坛主席：祝烈煌（北京理工大学）

任 奎（浙江大学）

召 集 人：祝烈煌（北京理工大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:35	致 辞
14:35-15:05	区块链安全 斯雪明（上海交通大学）
15:05-15:35	大模型系统安全：威胁与防御 李 琦（清华大学）
15:35-15:55	休 息
15:55-16:25	从大模型到自治体：区块链构建可验证智能协作网络 朱天清（澳门城市大学）
16:25-17:05	区块链赋能智能体协作：机遇与挑战 郑子彬（中山大学）
17:05-17:45	从数据模型安全到可信智能体 殷赵霞（华东师范大学）
17:45-18:15	链上智能体——从身份可信到协作可信 闫 莺（蚂蚁数字科技）

论坛主席



祝烈煌

祝烈煌，北京理工大学网络空间安全学院教授、党委书记，密码学院院长，国家高层次科技创新领军人才、科技部中青年创新领军人才，国家重点研发计划首席科学家。长期聚焦区块链与分布式智能安全，围绕跨链服务安全、分布式区块链数据协同、隐私保护计算与可信数据服务开展系统研究。曾主持多项国家自然科学基金重点项目、国家重点研发计划项目，在多链原子交换、区块链数据完整性审计、隐私保护访问控制等方面取得系列成果，相关论文发表于USENIX Security、IEEE TIFS、IEEE TDSC等重要期刊会议，并获中国电子学会科技进步一等奖等省部级奖励。



任 奎

任奎，AAAS、ACM、CCF和IEEE会士，国家特聘专家，目前担任杭州电子科技大学副校长、浙江大学区块链与数据安全全重常务副主任。主要从事数据要素安全、人工智能安全等领域的研究。主持多项科技部、国家自然科学基金委重点项目，获得三项省部级一等奖、IEEE安全技术委员会技术成就奖、美国国家自然科学基金会职业奖、纽约州立大学校长资深学者研究奖、Sigma Xi青年学者奖等。同时担任国家数据专家咨询委员会委员、全国数据标准化技术委员会委员、中国网络空间安全学会常务理事、中国电子学会理事、中国密码学会理事等职务。

召集人



祝烈煌

祝烈煌，北京理工大学网络空间安全学院教授、党委书记，密码学院院长，国家高层次科技创新领军人才、科技部中青年创新领军人才，国家重点研发计划首席科学家。长期聚焦区块链与分布式智能安全，围绕跨链服务安全、分布式区块链数据协同、隐私保护计算与可信数据服务开展系统研究。曾主持多项国家自然科学基金重点项目、国家重点研发计划项目，在多链原子交换、区块链数据完整性审计、隐私保护访问控制等方面取得系列成果，相关论文发表于USENIX Security、IEEE TIFS、IEEE TDSC等重要期刊会议，并获中国电子学会科技进步一等奖等省部级奖励。

报告嘉宾



斯雪明

斯雪明，中原工学院（河南电子科技大学（筹））学术副校长，上海交大区块链研究中心首席科学家，科技部十四五区块链重点专项专家组组长、工信部产促中心区块链专家分委会主任等。同时担任中国计算机学会区块链专委会荣誉主任、中央网信办区块链创新应用试点专家咨询组副召集人、中国人民大学区块链研究院学术委员会主任、数字福建区块链发展专家顾问团顾问等职务。

报告题目：对区块链应用的思考

报告摘要：本报告围绕区块链应用面临的挑战与破解之道展开深入分析。首先，报告将从国际与国内两个维度，系统梳理区块链应用目前的发展现状与趋势。接着，报告全面剖析区块链应用在推进过程中面临的多重挑战。最后，报告回归区块链的本质，提出从场景到生态的转变策略，并提供若干具体的实施建议，旨在推动区块链应用实现高质量、可持续发展的创新发展。



李 琦

李琦，清华大学互联网体系结构全国重点实验室副主任、国家级高层次人才。长期从事网络安全、大模型与智能体系统安全研究，曾获4项省部级及国家一级学会科技一等奖，以及USENIX Internet Defense Prize、国际基础科学大会前沿科学奖等重要学术奖励。相关研究成果获USENIX Security 2023/2024、NDSS 2025、CCS 2025等国际顶级安全会议杰出论文奖。现任IEEE TDSC、ACM TOPS等国际知名期刊编委，曾担任USENIX Security程序委员会副主席、ACM CCS领域程序委员会主席等重要学术职务。

报告题目：大模型系统安全：威胁与防御

报告摘要：随着大模型与智能体加速融入各类业务场景，其安全风险已从单一模型层面扩展至基础设施、模型、智能体与应用等多个环节，呈现出攻击面广、风险链条长、影响范围大的系统性特征，对人工智能安全与国家安全带来新的挑战。本报告从系统视角出发，分析大模型系统面临的主要安全威胁与关键挑战，梳理系统安全、模型安全、智能体与应用安全等领域的典型攻击面及防御技术，并结合相关研究成果、标准规范与行业实践，探讨大模型系统安全的技术防御与治理路径，为构建安全、可信、可控的大模型生态提供支撑。



朱天清

朱天清，澳门城市大学数据科学学院教授、院长，分管学术事务。主要研究领域为信息安全、数据隐私保护与人工智能安全。主持澳大利亚国家自然科学基金（Australian Research Council Discovery & Linkage）项目四项，参与四项，累计经费逾 400 万澳元；另主持国家自然科学基金青年项目及面上项目各一项。在安全领域提出多类人工智能算法安全漏洞及相应防御机制，并提出基于系统性安全分析的人工智能攻击防御框架。在 IEEE Transactions 系列等国际权威期刊发表 SCI 论文 150 余篇，其中以第一作者或通讯作者发表论文 30 余篇。曾任澳大利亚国家自然科学基金会评审专家（College of Expert），担任 4 本 SCI 期刊副主编（Associate Editor），多项国际人工智能会议程序委员会主席（PC Chair）。入选斯坦福全球前 2% 顶尖科学家榜单。

报告题目：从大模型到自治体：区块链构建可验证智能协作网络

报告摘要：随着大型语言模型能力的不断提升，智能体（AI agent）正在逐渐从“指令响应”走向“自主规划、目标驱动与多智能体协作”。然而，当智能体被应用于开放环境中时，可信身份、行为约束、协作激励与结果可验证性等问题日益凸显，传统中心化平台难以提供透明、可审计、可追责的运行保障。本报告将围绕“从大模型到自治体”的技术演进路线，介绍团队在可验证 AI 智能协作体系方面的最新研究进展。我们将重点探讨如何利用区块链提供可验证的身份与行为记录，如何通过激励与信誉机制支持多智能体自组织协作，以及如何在智能体决策链中实现安全与隐私保护。最后，本报告将展望可验证自治智能网络在金融安全、隐私计算与智能治理等场景中的落地潜力与未来发展趋势。



郑子彬

郑子彬，中山大学软件工程学院院长、中山大学人工智能研究院副院长、IEEE Fellow、IET Fellow、ACM 杰出科学家、国家数字家庭工程技术研究中心副主任、广东省区块链工程技术研究中心主任。发表论文多篇，论文谷歌学术引用超过 5 万次，H 指数为 105。主持国家重点研发计划项目、自然科学基金重点项目等多个项目；获得教育部自然科学二等奖、吴文俊人工智能自然科学二等奖、ACM SIGSOFT Distinguished Paper Award 等奖项。研究方向为区块链、Web3、软件可靠性、可信大模型等。

报告题目：区块链赋能智能体协作：机遇与挑战

报告摘要：近年来，大模型、智能体等技术迅猛发展，逐渐出现多智能体系统。然而，在多智能体协作中，其数据授权、溯源、收益分配等方面仍存在挑战，且智能体也缺少可信身份与责任边界。区块链具有公开透明、防篡改等特点，有望成为智能体的去中心化协作底座。本报告将介绍区块链赋能智能体协作的机遇与挑战，并分享团队在区块链和智能体的相关研究工作。


殷赵霞

殷赵霞，华东师范大学信息与电子工程学院教授、博导，入选国家高层次青年人才教育部长江学者奖励计划青年学者、上海市东方英才计划青年项目等。主要从事多媒体、人工智能与信息安全研究，主持国家自然科学基金青年面上项目4项、国家重点研发计划子课题、启元实验室委托课题等20余项，在CVPR、ACL、AAAI、IEEE TDSC等国内外顶级会议与期刊发表论文80余篇。现任CCF杰出会员，CAAI、CICC、CAA终身会员及CSIG高级会员，兼CSIG数字媒体取证与安全专委会常务委员、CCSS女工委秘书长等。服务中国科协、国家科技库、基金委、教育部、中国科学院学部、吴文俊人工智能奖及多省市科技人才项目评审。

报告题目：从数据模型安全到可信智能体

报告摘要：面对人工智能可信与安全的国家战略需求，技术支撑已成为政策落地的关键。当前，AI正由单一模型向具备环境感知、工具调用、动作执行及持续进化能力的智能体系跃迁，其安全挑战亦随之升级，从局部技术点演变为贯穿数据、模型与智能体全栈的系统性风险。本报告立足“感知-认知-行动-进化”AI安全全链路，介绍课题组从数据模型安全到可信智能体最新研究进展。


闫莺

闫莺，博士，现任蚂蚁数字科技CTO，长期致力于区块链与隐私计算、智能体技术领域的深度研究与产业化实践。目前担任IEEE跨链标准主席、中国计算机学会（CCF）区块链专委会及杰出会员。同时兼任浙江大学区块链和数据安全国家重点实验室高级研究员、华东师范大学行业博士生导师。

她拥有国内外授权专利百余项，在国际顶级会议与期刊发表学术论文60余篇，并主编了《以太坊技术详解与实战》一书。凭借卓越的技术贡献，她曾入选“2019福布斯中国科技女性50人”。

报告题目：链上智能体——从身份可信到协作可信

报告摘要：AI Agent 正从对话框走向业务系统，但当智能体参与真实商业，三个信任问题浮出水面：谁在操作——身份可信；钱怎么走——支付链路可信；谁和谁协作、出了事谁负责——协作可信。本报告分享蚂蚁数科 AgentOS 如何以区块链为信任基座，构建智能体的三层可信体系：通过链上身份为 Agent 颁发可验证身份，实现操作可归因、行为可追溯；通过链上支付协议保障 Agent 交易的资金安全与路径可溯；通过 Agentix 五步闭环 Harness 与链上存证，让多 Agent 协作的每一步可审计、可追责、不可抵赖，实现 AI Agent 从“能用”到“敢用”的信任跨越。

网密融合论坛

承办单位：中国科学院信息工程研究所、电子科技大学

🕒 9月19日下午

📍 天鹅湖大酒店三楼12号会议室

论坛主席：刘 峰（中国科学院信息工程研究所）

张小松（电子科技大学）

召 集 人：刘 峰（中国科学院信息工程研究所）

时 间	内 容
13:00-13:05	主持开场
13:05-14:10	致 辞
14:10-14:40	主动网络安全模型的创新与实践 张小松（电子科技大学）
14:40-15:10	云网端融合的应急指挥通信安全关键技术及应用 刘建伟（北京航空航天大学）
15:10-15:40	加密流量智能识别方法 程 光（东南大学）
15:40-16:00	休 息
16:00-16:30	规模化数据传输隐写原生隐匿网络 郭云彪（国际关系学院）
16:30-17:20	基于我国商用密码的功能型密码算法 黄欣沂（南京航空航天大学）
17:20-17:50	抗量子时代大证书的工程挑战及应对策略 杨春晖（视联动力）

论坛主席



刘 峰

中国网络空间安全学会应用密码工程专委会主任委员，中国科学院信息工程研究所正高级工程师、博导、所务委员，国科大岗位教授。多部委信息技术领域高层次专家。带领团队承担国家级重大战略任务，产出系列创新成果，打造国家创新平台。在Springer等出版专著2部，发表论文140余篇，授权专利20余项。多份战略报告获部委采纳及中央领导批示。主持国家重大工程、国自然、重点研发计划、中科院先导专项等30余项。担任《国际数字犯罪与取证》(IJDCF)主编，发起网络空间安全科学国际会议 (SciSec) 并任大会主席，曾获省部级科技进步一等奖等科技奖励。



张小松

中国网络空间安全学会应用密码工程专委会副主任委员，电子科技大学信息与软件工程学院院长、教授、博导，国家级人才，中国工程院院士有效候选人，成都市科协副主席。我国计算机网络对抗开拓者，中国计算机学会、电子学会、通信学会、指控学会会士。国家科技进步奖一等奖、二等奖（第一完成人），四次省部级一等奖。主持国家重点、重大课题40余项，连续担任十三五、十四五国家重点研发计划首席科学家及国自然重点项目负责人。一作出版专著5部，发表SCI/EI论文250余篇，第一发明人授权发明专利63项。获全国创新争先奖、全国科创名匠、国家网络安全优秀人才等荣誉，入选四川省杰出人才，获评四川省教书育人名师、优秀共产党员、先进工作者。

召集人



刘 峰

中国网络空间安全学会应用密码工程专委会主任委员，中国科学院信息工程研究所正高级工程师、博导、所务委员，国科大岗位教授。多部委信息技术领域高层次专家。带领团队承担国家级重大战略任务，产出系列创新成果，打造国家创新平台。在Springer等出版专著2部，发表论文140余篇，授权专利20余项。多份战略报告获部委采纳及中央领导批示。主持国家重大工程、国自然、重点研发计划、中科院先导专项等30余项。担任《国际数字犯罪与取证》(IJDCF)主编，发起网络空间安全科学国际会议 (SciSec) 并任大会主席，曾获省部级科技进步一等奖等科技奖励。

■ 报告嘉宾



张小松

中国网络空间安全学会应用密码工程专委会副主任委员，电子科技大学信息与软件工程学院院长、教授、博导，国家级人才，中国工程院院士有效候选人，成都市科协副主席。我国计算机网络对抗开拓者，中国计算机学会、电子学会、通信学会、指控学会会士。国家科技进步奖一等奖、二等奖（第一完成人），四次省部级一等奖。主持国家重点、重大课题40余项，连续担任十三五、十四五国家重点研发计划首席科学家及国自然重点项目负责人。一作出版专著5部，发表SCI/EI论文250余篇，第一发明人授权发明专利63项。获全国创新争先奖、全国科创名匠、国家网络安全优秀人才等荣誉，入选四川省杰出人才，获评四川省教书育人名师、优秀共产党员、先进工作者。

报告题目：主动网络安全模型的创新与实践

报告摘要：非合作性计算机系统与网络对抗构成了网络空间安全的核心挑战。传统网络安全技术主要依赖被动响应机制，在应对现实世界复杂多变的未知威胁时展现出显著局限性。提出“主动网络安全”理念，旨在通过融合技术手段与战略级防御体系，全面提升网络安全水平。该理念的核心假设是：网络对抗环境中的攻击者与防御者均为追求各自目标最大化的理性决策主体。引入博弈论分析攻防双方的复杂依存关系并优化其策略选择。基于该理念，构建了主动网络安全模型SAPC，旨在构建一种集威胁感知、分析、追踪和响应于一体的综合防御能力。该模型由4大核心组件构成：智感、透析、活现和反制。SAPC通过基于博弈论的对抗行为理论分析与策略优化方法，将对抗过程建模为博弈过程，建立兼具理论深度与实践指导价值的网络安全框架。SAPC标志着网络防御理念从被动防御到主动感知对抗的范式转变，有力推动网络安全技术向具有前瞻预测、预防控制和战略引导特征的新模式演进。


刘建伟

北航网络空间安全学院创院院长，教授，博士生导师，国家级教学名师，享受国务院政府特殊津贴专家。现任国务院学位委员会学科评议组成员、教育部网络空间安全专业教指委副主任、全国密码专业学位研究生教指委委员、中国密码学会常务理事、中国网络空间安全学会常务理事、《网络空间安全科学学报》常务副主编、《网络与信息安全学报》副主编。

获国家技术发明一等奖、全国创新争先奖牌、国防技术发明一等奖、中国指挥与控制学会一等奖等科技奖励10余项。获国家级教学名师、国家网络安全优秀教师奖、宝钢优秀教师特等奖、北京市教学名师奖、北京市优秀教师等教学奖励和荣誉20余项。

报告题目：云网端融合的应急指挥通信安全关键技术及应用

报告摘要：本报告面向应急指挥场景高可靠保密通信需求，聚焦云网端融合架构，系统阐述应急指挥通信安全关键技术与应用方案。在终端侧，研制集成4G/5G、窄带卫星、自组网、PDT数字对讲的多模智能终端，可依据现场网络条件，实现地面蜂窝网、自组织网与卫星网络的无缝自适应切换；融合密码技术与端网协同机制，实现多安全域隔离、分级无感切换、可信接入认证及密钥自动分发，实现跨域保密通信与数据隐私保护。在网络侧，基于密码技术实现不同安全等级业务的安全域划分与隔离。在云侧，创新提出代理重加密、区块链跨链可信交互、密文策略属性基加密等技术，实现去中心化、细粒度的敏感数据访问控制。综合采用以上关键技术，可实现应急指挥场景下跨网络域和安全域的安全保密通信。


程 光

东南大学首席教授、二级教授、网络空间安全学院执行院长，教育部网络空间安全教指委委员、中国网络空间安全学会常务理事、江苏省网络空间安全学会理事长、网络空间安全2030重大专项专家等学术兼职。研究方向网络流量安全、主动防御等，主持承担国家重点研发、国家自然科学基金、中央网信办专项等科研项目40余项。获国家网络安全优秀教师、国家教学成果一等奖，牵头获省部级一等奖2项、二等奖2项。发表学术论文100余篇，出版专著9部，培养研究生200余人。

报告题目：加密流量智能识别方法

报告摘要：网络流量的加密化使得加密流量分析成为热点问题，被多个学术组织纳入了网络空间安全重大态势感知科学问题。本报告首先分析加密流量现有的机器学习和深度学习的背景，讨论加密流量有效性问题，其核心是流量的特征工程。加密流量分析是由检测—分类—识别三层研究框架构成，特征工程是最重要的，特征决定了分类效果的上限。对加密流量分析是工程奇技还是理论科学、加密流量分析最基础的是什么、在数据基础上后期应该怎么用、人工智能分类器如何选择与建立、如何考虑真实网络等相关问题进行讨论研究。最后对真实的网络环境中，基于加密流量的实验环境和教学案例的设置进行讨论。



郭云彪

教授，国家百千万人才工程国家级人选，国家中青年科技创新领军人才，享受国务院特殊津贴。一直从事信息隐藏、多媒体安全、隐蔽通信等相关领域的研究工作，是特种通信国家科技创新团队的领头人，建立并发展了面向秘密信息传递的数字隐写技术理论和方法，推动隐蔽通信技术的换代升级，实现通信模式的重塑。主持完成30余项国家级省部级项目，获国家科技进步一等奖2次，国家科技进步二等奖2次，国家科技发明二等奖1次。

报告题目：规模化数据传输隐写原生隐匿网络

报告摘要：针对数据隐蔽传输面临匿名隐蔽未融合、隐匿容限难突破、隐写信道难构建和隐匿安全难度量等挑战难题，采用传控分离思想，以大模型生成式隐写为基础，在网络组织、信令控制和通道隐蔽等方面构建隐蔽组网控制体系，提出规模化数据传输隐写原生隐匿网络，破解隐藏、匿名和规模化融合难题。深入探讨隐写原生隐匿网络基础理论与规模化数据传输模型、隐写原生隐匿网络传输架构与四抗服务保障方法、大模型生成式隐写理论与网络隐蔽信道构建方法、隐写原生隐匿网络安全度量方法与信任评估体系，实现非受控网络环境中规模化数据传输的抗发现、抗分析、抗溯源和抗阻断，解决互联网数据传输的隐私保护难题。



黄欣沂

一直从事密码科学技术研究，主持国家自然科学基金青A、青B、重点，国家重点研发计划等科研项目。研究成果获中国密码学会密码创新奖一等奖，教育部自然科学奖一等奖、二等奖等；担任国际密码协会亚密会指导委员会委员，《中国科学：信息科学》、《密码学报》、JCST等学术期刊编委等。

报告题目：基于我国商用密码的功能型密码算法

报告摘要：商用密码用于保护不属于国家秘密的信息，广泛应用在金融、通信、税务等领域。我国自主研发的SM2、SM9等商用密码算法以基础安全功能为主要目标，不能适应“安全威胁更严峻、功能需求更多样”的新兴应用场景。报告介绍SM2、SM9等商用密码衍生的功能型密码算法，涵盖安全层、用户层、数据层、管理层、实现层等五个层面。上述算法进一步健全了国产密码算法体系，强化了信息技术自主可控的密码保障。

网络仿真与安全评测论坛

承办单位：鹏城实验室、哈尔滨工业大学（深圳）

🕒 9月19日下午

📍 天鹅湖大酒店三楼5号会议室

论坛主席：方滨兴（中国工程院院士）

贾 焰（国防科技大学）

郭世泽（中国科学院院士）

顾钊铨（鹏城实验室）

召 集 人：贾 焰（国防科技大学）

顾钊铨（鹏城实验室）

时 间	内 容
13:30-13:35	主持开场
13:35-13:45	致 辞
13:45-14:15	具身智能技术及其安全性评测 朱文武（清华大学）
14:15-14:45	面向通信网络的人工智能安全测评技术研究 陶 冶（中国联通研究院）
14:45-15:15	工控系统网络安全布防思考与仿真推演 杨祎巍（南方电网科学研究院有限责任公司）
15:15-15:40	休 息
15:40-16:10	具身智能安全标准与评测技术发展趋势 罗 蕾（电子科技大学）
16:10-16:40	汽车网络安全仿真 贺可勋（中汽智能科技（天津）有限公司）
16:40-17:10	省级网络靶场实战化转型的探索与实践 周 艳（数字广西集团）

论坛主席



方滨兴

中国工程院院士，中国电子信息产业集团有限公司资深首席科学家。现任中国中文信息学会理事长，中国标准化协会网络与信息安全技术委员会主席。曾任中国网络空间安全协会理事长、中国互联网协会副理事长/网络与信息安全工作委员会主任、中国计算机学会副理事长/计算机安全专业委员会主任、中国通信学会副理事长/通信安全技术委员会主任；国家自然科学基金委员会网络与信息安全重大计划专家组成员、可信软件重大计划专家组副组长。



郭世泽

中国科学院院士，解放军信息技术安全研究中心研究员，网络空间安全防御领域顶级专家。



贾 焰

国防科技大学研究员，现任国家重大攻关项目负责人，中国中文信息学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用；作为课题负责人承担和主持了国家级重大/重点项目20余项，获国家科技进步二等奖5项；发表SCI/EI论文300篇，出版专著8部，获发明专利124项，软件著作权125项，牵头制定相关行业标准33项；获何梁何利基金科学与技术进步奖，获全国巾帼建功标兵称号。



顾钊铨

鹏城实验室研究员，新型网络部平台所技术总师，哈尔滨工业大学（深圳）教授。从事网络靶场、网络攻击研判等网络空间安全前沿技术研究，出版学术著作6部，发表学术论文200余篇，入选爱思唯尔全球前2%顶尖科学家榜单；申请中国发明专利200余项，授权发明专利100余项，参与制定20余项网络空间安全行业标准。获国家科技进步二等奖，国防技术发明一等奖，广东省科技进步一等奖等。

召集人



贾 焰

国防科技大学研究员，现任国家重大攻关项目负责人，中国中文信息学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用；作为课题负责人承担和主持了国家级重大/重点项目20余项，获国家科技进步二等奖5项；发表SCI/EI论文300篇，出版专著8部，获发明专利124项，软件著作权125项，牵头制定相关行业标准33项；获何梁何利基金科学与技术进步奖，获全国巾帼建功标兵称号。



顾钊铨

鹏城实验室研究员，新型网络部平台所技术总师，哈尔滨工业大学（深圳）教授。从事网络靶场、网络攻击研判等网络空间安全前沿技术研究，出版学术著作6部，发表学术论文200余篇，入选爱思唯尔全球前2%顶尖科学家榜单；申请中国发明专利200余项，授权发明专利100余项，参与制定20余项网络空间安全行业标准。获国家科技进步二等奖，国防技术发明一等奖，广东省科技进步一等奖等。

报告嘉宾



朱文武

清华大学计算机系长聘教授、原信息科学与技术国家研究中心副主任。长期从事多媒体信息处理与智能的研究。曾任国家973计划项目首席科学家，国家基金委重大项目负责人，教育部人工智能科技创新专家组成员等。三次获国家自然科学二等奖（2023年排名1，2018年排名1，2012年排名2）。IEEE Fellow、SPIE Fellow、AAAS Fellow、ACM Fellow、欧洲科学院（Academia Europaea）院士。先后任IEEE Transactions on Multimedia主编，IEEE Transactions on Multimedia 指导委员会主席，IEEE Transactions on Circuits and Systems for Video Technology主编等。获2024年IEEE电路与系统学会Charles A. Desoer技术成就奖和2023年ACM SIG多媒体杰出技术成就奖。

报告题目：具身智能技术及其安全性评测

报告摘要：具身智能（Embodied AI）旨在构建通过主动感知、具身认知及动态物理交互实现智能体与环境深度融合的智能系统。本报告首先介绍具身智能技术。然后，聚焦大模型驱动、世界模型驱动、大模型与世界模型双驱动的具身智能范式演进，介绍在动态开放环境下，如何有效融合多模态大模型的语义理解、任务规划与推理能力，以及世界模型对物理环境的内在状态表征与未来预测能力，实现语义认知与物理交互的深度协同融合。最后，深入探讨具身智能领域存在的多重安全隐患，系统梳理以智能模型为核心衍生的各类风险，涵盖环境感知环节的对抗攻击与鲁棒性评测，价值观失偏引发的决策安全隐患，控制参数失准带来的控制安全风险与安全控制问题，以及智能机器人面临的网络系统性安全风险。



陶 冶

高级工程师，中国联通集团网信安全首席专家，中国联通研究院首席研究员，长期从事网络安全、数据安全、人工智能安全等相关领域研究工作，担任工业和信息化部人工智能标准化委员会安全组副组长、TC610安全组副组长、人工智能学会安全专委会委员等，牵头制定国际标准30余项，获省部级、联通公司科技奖项10余项，发表专著3本。

报告题目：面向通信网络的人工智能安全测评技术研究

报告摘要：随着生成式人工智能、代理式人工智能技术的快速演进与垂直行业的规模化落地，人工智能安全风险边界持续拓展，从生成内容的合规性、真实性问题，延伸至代理智能体的权限失控、行为失序风险，成为影响人工智能产业健康发展与国家安全的关键变量。本报告聚焦人工智能安全治理与技术防护体系建设，系统梳理了多模态安全评测数据集构建、可解释安全裁判模型、自适应评测智能体、幻觉检测与减轻、神经元级可解释性分析等生成式人工智能核心安全技术，阐述了代理式人工智能的评测框架、典型攻击场景与风险升级逻辑，同时介绍了中国联通作为通信运营商，对人工智能在通信网络典型场景应用过程中的安全治理实践，为推动人工智能安全、可靠、可控发展提供技术路径与产业参考。



杨伟巍

博士，正高级工程师，南方电网科学研究院有限责任公司数字化及信息安全中心副主任，从事网络安全相关工作已有十余年，在密码应用、密码模块、网络安全攻击、网络安全仿真等技术方面开展了长期深入的研究和实践，参与国家重点研发计划、南方电网公司科技重大专项数十项，累计授权第一发明人专利17项、发表论文十余篇，获得广东省科技进步二等奖、广东省专利银奖、南方电网公司科技进步一等奖、南方电网公司成果转化应用一等奖等奖励，牵头或参与南方电网公司及电力行业网络与信息安全信息通报、网络安全督查检查、网络安全靶场建设、专题研究等工作，获得南方电网公司通报嘉奖2次。

报告题目：工控系统网络安全布防思考与仿真推演

报告摘要：本文提出了承接“全域防御、纵深防御”要求的“广义暴露面分析、全攻击链布防”体系化布防框架，广义暴露面分析强调“无死角”的系统思维，通过资产接口、物理环境、全生命周期、人员四个维度全面识别暴露面，将几乎所有可能的攻击路径纳入统一的系统的分析框架，旨在消除布防盲区；全攻击链布防则基于“前序防线失效”的底线思维，结合网络攻击各环节形成的攻击链，建立“暴露面-攻击链”矩阵，确保每个暴露面在各攻击环节均有防护措施，构建层层阻击能力。该框架为网络安全布防提供了从风险识别到布防设计的完整方法论，并进行仿真推演，兼具理论指导性与实践操作性，可有效指导布防体系设计与运行优化。

**罗 蕾**

教授/博导，电子科大嵌入式软件工程中心主任，长期从事嵌入式基础软件、具身智能安全、智能计算研究和产业化工作。“智能汽车安全技术”全重学术委员会委员，中国汽车工程学会基础软件分委会副主任、车载信息服务产业应用联盟（TIAA）网络安全委员会秘书长、网络空间安全学会网络对抗仿真专委会委员。主持承担了多项国家级项目、牵头整车信息安全强标，参与了汽车、具身智能、低空等相关国行标20余项，发表学术论文50余篇。

报告题目：具身智能安全标准与评测技术发展趋势

报告摘要：报告从三个方面介绍具身智能定义、特点、技术产业与标准的发展情况，具身智能安全风险分析包括硬件、控制、计算、传输、产业链与伦理6大类风险和十几个风险点，安全防护技术与测评技术，智能网联汽车网络与数据安全的相关标准技术等发展情况，最后介绍团队研究与产业化相关工作。

**贺可勋**

中汽智能科技（天津）有限公司网络安全室主任，中国汽车技术研究中心学科后备带头人，从事汽车网络安全检测相关工作，主要成果如下：牵头成立汽车信息安全测试技术工作组，率先在行业内构建汽车信息安全研发验证测试体系，积极参与国际法规协调，以国际法规协调专家身份参与WP.29 R155 协调工作，参与制定汽车信息安全、数据安全和软件升级相关国家标准8项，主持、参与完成省部级以上汽车信息安全、数据安全、软件升级相关科研课题6项，获得汽车信息安全领域相关省部级奖项4项，授权汽车信息安全相关发明专利6项、发表论文12篇，参编专著2项。

报告题目：汽车网络安全仿真

报告摘要：以面向智能网联汽车的网络安全仿真技术路径与工程经验为主线，报告重点围绕以下内容展开：一是分析智能网联汽车在车载网络、远程通信、云端服务等环节面临的安全测试需求，梳理仿真靶场建设中的关键技术问题；二是介绍依托与鹏城靶场的合作，开展汽车行业分靶场建设的总体架构与技术方案，包括仿真环境构建、典型攻击场景库建设、虚实结合的攻防演练支撑能力；三是分享靶场在安全测评、漏洞验证、攻防演练等方面的应用实践经验。最后，报告结合行业标准与技术发展趋势，探讨汽车网络安全靶场的后续建设方向，为智能网联汽车安全测试基础设施的建设与应用提供参考。

**周 艳**

现任数字广西集团数据安全部总经理、广西数广威奇科技有限公司董事长、广西数字国信安全技术服务有限公司董事长。长期在省政府办公厅、政府智库任职，深耕区域经济建设、网络信息安全、公共数据授权运营管理等领域。牵头编制多项行业团体标准，主持国家级、省市级网络安全赛事、重大网络安保、政务数字化重点项目，多项成果荣获行业科技奖项，运营平台案例入选世界互联网大会优秀案例。担任广西师范大学人工智能学院、桂林电子科技大学人工智能学院校外研究生导师，持续推动网络安全、数据安全、人工智能安全领域产学研用协同创新。

报告题目：省级网络靶场实战化转型的探索与实践

报告摘要：以鹏城网络靶场广西分靶场的建设运营实践为主线，系统梳理省级网络靶场实现实战转型的路径探索。报告重点围绕以下内容展开：一是立足网络靶场建设运营中的实际问题，提炼靶场从技术验证走向实战化、常态化运营的关键路径；二是探讨多级子靶场体系的建设规划，包括如何通过高校联动、行业协同实现资源共建共享。三是分享平台在数据开发利用、安全能力建设、运营模式创新等方面的实践经验；报告旨在为省级网络安全基础设施的建、管、用提供可借鉴的广西经验与思考。

智能装备安全论坛

承办单位：中国信息安全研究院有限公司、中国网络空间安全学会信息物理系统安全专业委员会

🕒 9月19日下午

📍 天鹅湖大酒店一楼淮河厅A

论坛主席：方滨兴（中国工程院院士） 郭世泽（中国科学院院士）
 贾 焰（国防科技大学）

召 集 人：韩 毅（中国电子信息产业集团有限公司）

时 间	内 容
13:45-13:50	主持开场
13:50-14:00	致 辞
14:00-15:00	人工智能的衍生安全：防范AI失控的“四不原则” 方滨兴（中国工程院院士）
15:00-15:30	AI时代下的网络靶场即服务的重塑与升级 廖 清（哈尔滨工业大学（深圳））
15:30-15:50	休 息
15:50-16:20	从数据到模型再到智能体：智能装备安全面临新挑战 王宝生（国防科技大学）
16:20-16:50	智能体网络安全探索与思考 王春雷（中国信息安全研究院有限公司）
16:50-17:20	智能汽车安全能力框架 杜跃进（浙江大学）

论坛主席



方滨兴

中国工程院院士，中国电子信息产业集团有限公司资深首席科学家。现任中国中文信息学会理事长，中国标准化协会网络与信息安全技术委员会主席。曾任中国网络空间安全协会理事长、中国互联网协会副理事长/网络与信息安全工作委员会主任、中国计算机学会副理事长/计算机安全专业委员会主任、中国通信学会副理事长/通信安全技术委员会主任；国家自然科学基金委员会网络与信息安全重大计划专家组成员、可信软件重大计划专家组副组长。



郭世泽

中国科学院院士，解放军信息技术安全研究中心研究员，网络空间安全防御领域顶级专家。



贾 焰

国防科技大学研究员，现任国家重大攻关项目负责人，中国中文信息学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用；作为课题负责人承担和主持了国家级重大/重点项目20余项，获国家科技进步二等奖5项；发表SCI/EI论文300篇，出版专著8部，获发明专利124项，软件著作权125项，牵头制定相关行业标准33项；获何梁何利基金科学与技术进步奖，获全国巾帼建功标兵称号。

召集人



韩 毅

中国电子信息产业集团有限公司军工部主任，中国信息安全研究院有限公司董事长，中国网络空间安全学会信息物理系统安全专委会主任委员，博士，正高级工程师，长期从事新域新质作战理论和科技战略研究，某重要行业网络安全领域学术带头人，有关研究获得国防科技卓越青年基金资助。发表学术论文50余篇，发明专利10项，先后获得国家科技进步二等奖、广东省科技进步一等奖、湖南省科技进步一等奖、军队科技进步二等奖、湖南省科技进步二等奖等。

报告嘉宾



方滨兴

中国工程院院士，中国电子信息产业集团有限公司资深首席科学家。现任中国中文信息学会理事长，中国标准化协会网络与信息安全技术委员会主席。曾任中国网络空间安全协会理事长、中国互联网协会副理事长/网络与信息安全工作委员会主任、中国计算机学会副理事长/计算机安全专业委员会主任、中国通信学会副理事长/通信安全技术委员会主任；国家自然科学基金委员会网络与信息安全重大计划专家组成员、可信软件重大计划专家组副组长。

报告题目：人工智能的衍生安全：防范AI失控的“四不原则”

报告摘要：AI在赋能网络攻防、提升安全能力的同时，也伴生了深度伪造、自主越权、欺骗勒索、拒绝关机等一系列新型失控问题，本报告提出了防范AI失控的“四不原则”：允许AI自主规划，但不自主改写终极目标；允许AI自我改进，但不自己批准进化；允许AI使用资源，但不自主形成扩张闭环；允许AI调用工具，但不自己扩大权限边界。



廖清

哈尔滨工业大学（深圳）网络空间安全研究院院长，教育部长江学者特聘教授，第十八届中国青年科技奖，腾讯科学探索奖，全球前2%顶尖科学家。长期从事信息安全方面的理论研究和工程实践。在国际著名学术刊物发表高水平学术论文 100余篇，获ICSE 2024杰出论文奖等，授权发明专利和制定行业标准30余项，相关成果获国家和省部级一等科技奖励多项。

报告题目：AI时代下的网络靶场即服务的重塑与升级

报告摘要：当前，全球网络空间对抗态势日趋严峻，网络靶场作为攻防演练与装备试验的关键基础设施，其战略地位日益凸显。然而，传统靶场建设面临成本高昂、周期冗长等瓶颈，难以适配当前紧迫的安全需求。针对上述痛点，本报告系统阐述了“网络靶场即服务（CRaaS）”的新型应用范式。该模式依托云计算与虚拟化技术，使用户无需本地部署即可通过Web端获取安全测试与实训演练等云服务，从而大幅降低建设与运维成本。报告将重点探讨人工智能（AI）技术对靶场底层逻辑与业务形态的重塑作用，展示AI在场景智能构建与自动化运维方面的创新突破。此外，报告还将介绍由超6000项共享资源与百余家生态伙伴协同构建的攻防实战平台，全面呈现智能化时代网络安全实战化人才培养的新范式。



王宝生

国防科技大学计算机学院研究员，博士生导师，曾任国家级专家组专家。研究方向计算机网络和网络安全。参加多项国家重大科研项目，获国家科技进步奖特等奖1项、国家科技进步奖二等奖3项，国家教学成果一等奖1项。

报告题目：从数据到模型再到智能体：智能装备安全面临新挑战

报告摘要：数据转化为模型和算法，实现了智能知识能力的跃升，规则驱动的流程式自动化转变为生成式智能体，实现了自主决策控制能力的跃升。两次转变给智能化装备带来进步的同时，也带来新的安全挑战：数据安全面临管理载体与使用模式的根本转变，行为安全面临身份鉴别和权限控制新问题，亟需重新设计智能装备新的安全范式、机制和模式。



王春雷

正高级工程师，博士生导师，毕业于清华大学计算机科学与技术系，获工学博士学位。先后工作于原总装备部武器装备论证研究中心、电子信息系统总体研发中心，现为中国电子信息产业集团有限公司首席科学家、中国信息安全研究院有限公司总工程师。主要从事网络空间安全、军事网信体系等领域预先研究、关键技术产品研发及产业化应用，完成60余项国家及军队科研项目研究，发表学术论文80余篇，获得军队科技进步奖10项，获得国务院政府特殊津贴。

报告题目：智能体网络安全探索与思考

报告摘要：随着大模型技术的快速演进，AI智能体已从“能对话”发展为“能行动”，并加速向智能装备、工业控制等关键领域渗透。然而，AI智能体在赋能装备的同时也引入了全新的安全风险。本报告聚焦智能体网络安全，从威胁挑战与对策分析两个维度介绍最新研究进展。威胁挑战层面，分析了模型安全漏洞、智能体安全漏洞以及外部诱导劫持等核心攻击向量，剖析了智能体网络攻击在智能装备场景应用中可能造成的破坏性后果。对策分析层面，结合国内外最新研究与案例，提出“以模制模、灰色博弈”的思路，构建面向智能体网络的内生安全智能防御体系，探讨权限沙箱、智能体行为监控、多智能体互信等关键技术方向，为关键领域的智能体网络安全实践提供参考。



杜跃进

浙江大学计算机科学与技术学院求是特聘教授/博导，区块链与数据安全全国重点实验室首席研究员，获国家科技进步一等奖2项，并获得新世纪百千万人才工程国家级人才，国务院政府特殊津贴、全国青年岗位能手、信息产业十大杰出青年、CCF杰出会员、亚太信息安全领袖成就奖等荣誉。他的主要研究领域包括网络安全、数据安全、人工智能安全、智能汽车与具身智能安全、工业互联网安全、智慧应急与安全管理等。

报告题目：智能汽车安全能力框架

报告摘要：智能汽车正在演进成典型的智能装备，传统的安全概念在智能汽车领域有大量需要调整的地方。本报告从智能汽车企业的视角，围绕“安全的车、安全的厂、安全的运营基础设施”，探讨智能汽车安全的新定义，提出智能汽车安全能力框架，以及当下需要研究的新问题。

“后大模型时代”的软件与系统安全论坛

承办单位：中国科学院软件研究所

🕒 9月19日下午

📍 天鹅湖大酒店三楼7号会议室

论坛主席：孔志印（中国工程院院士）

召 集 人：苏璞睿（中国科学院软件研究所）

时 间	内 容
14:00-14:05	主持开场
14:05-14:15	致 辞
14:15-14:40	大模型驱动的智能模糊测试技术 张 超（清华大学）
14:40-15:05	智能体的安全风险分析和防护 张 源（复旦大学）
15:05-15:30	隐私Token：大模型推理中的数据隐私风险与防护 张殷乾（南方科技大学）
15:30-15:40	休 息
15:40-16:05	操作系统缺陷检测 白家驹（北京航空航天大学）
16:05-16:30	面向漏洞挖掘场景的可信智能体构造思考与实践 贾相堃（中国科学院软件研究所）
16:30-16:55	大模型赋能的软件供应链安全分析 吴月明（华中科技大学）
16:55-17:20	智能体应用过程中的安全思考与实践 李建彬(天融信科技集团)

论坛主席



孔志印

中国工程院院士。主要从事网络信息与安全技术研究，在网络空间安全理论及工程建设方面取得系统性、开拓性成果。先后获国家科技进步一等奖2项、二等奖2项，以及全国优秀科技工作者、全国创新争先奖等奖励和荣誉。

召集人



苏璞睿

中国科学院软件研究所二级研究员，长期从事软件漏洞研究。国家重点研发计划项目首席科学家，国家自然科学基金重点项目负责人，曾任国际网络安全旗舰学术会议RAID第22届大会主席，学术成果发表于ACM CCS、NDSS、USENIX Security、IEEE S&P等国际高水平（CCF A）学术会议和期刊。主持研制的AOTA动态污点分析系统、金刚软件智能分析系统等应用于国家相关部门和国家漏洞库，取得显著效益，曾获省部级科技奖励4项（一等奖2项，二等奖2项）。

报告嘉宾



张超

张超博士，清华大学长聘副教授，网络研究院副院长，蓝莲花战队教练，网安奥赛NCO组委会负责人，华为冠名教授，CCF杰出会员。主要研究软件和系统安全、人工智能安全。在四大安全会议发表论文60余篇，CCF A类论文100余篇。主持重大专项、重点研发、基金面上及重点、GF预研及重点等项目数十项。研究成果获得JS科技进步一等奖、CCF青年科技奖、日本大川奖、微软特别提名奖等。个人获得国家级青年人才、清华大学青年教师最高荣誉“学术新人奖”、中国科协青托、求是杰出青年学者、MIT TR35 China、WX创新人才等荣誉。

报告题目：大模型驱动的智能模糊测试技术

报告摘要：模糊测试近十年最流行的软件漏洞挖掘技术，但是逐步进入平台期。近年来，大语言模型技术蓬勃发展，展现了卓越的代码语义理解能力和内容生成能力，为突破模糊测试技术瓶颈提供了新的契机。本次报告将介绍团队在大模型驱动的智能模糊测试技术方面的最新研究进展，分享如何利用大模型能力，优化测试例变异方法，更快到达目标代码，更高效触发目标漏洞，提升漏洞挖掘效率。



张 源

复旦大学教授、博导，教育部长江学者，复旦白泽战队指导教师，主要研究方向为软件安全，研究工作获网安顶会杰出论文奖3次、荣誉提名2次，部分研究成果应用于华为、阿里、OPPO、vivo等公司，获优秀校企合作项目5次。带领复旦白泽战队获国内外顶尖安全攻防赛事冠军20余次，发现大量真实软硬件系统高危漏洞，获2021年国家信息安全漏洞库最具价值漏洞奖。

报告题目：智能体的安全风险分析和防护

报告摘要：作为一种新兴软件，智能体日益深入到生活中的各个方面。然而，频繁曝出的远程命令执行、数据库注入等高危漏洞，对智能体中用户数据的安全和隐私带来严重威胁。与传统软件不同，智能体具有开放的输入格式和灵活的行为能力，对智能体安全的分析与防护带来全新挑战。本报告将对智能体安全的现状进行概述，通过具体案例深入剖析智能体软件架构及漏洞根因，并提出针对智能体安全的防护方案。



张殷乾

南方科技大学教授、ACM杰出科学家。主要研究方向是计算机系统安全和体系结构安全。曾获美国国家科学基金青年科学家奖、中国计算机学会CCF IEEE-CS青年科技奖、北美计算机华人学者协会明日之星奖、国际顶级安全会议ACM CCS时间检验奖、俄亥俄州立大学 Lumley 研究成就奖、连续五年AMiner全球最具影响力的安全和隐私学者提名，入选全球前2%顶尖科学家终身榜单、《麻省理工科技评论》中国2022隐私计算科技创新人物以及2023年度十大“强国青年科学家”。

报告题目：隐私Token：大模型推理中的数据隐私风险与防护

报告摘要：随着大模型技术的快速发展，推理服务正成为AI应用的重要基础设施。用户输入的提示词往往包含个人隐私、商业机密以及敏感业务数据，而这些信息在推理过程中会被分解为Token，并经历编码、计算、缓存、通信和生成等多个阶段。每一个Token都可能暴露于不同的软件和硬件组件，从而面临数据窃取、侧信道攻击、恶意系统软件窥探等隐私威胁。如何保障Token在整个推理生命周期中的隐私与安全，已成为构建可信大模型推理系统面临的重要挑战。



白家驹

北京航空航天大学网络空间安全学院副教授（博士生导师）、安全工程系副主任。2018年博士毕业于清华大学计算机系（导师为胡事民院士），获得清华大学优秀博士学位论文奖。研究方向为系统软件安全、操作系统和程序分析，主导研发的程序分析框架在主流开源和国产系统软件中发现超五千个真实缺陷和漏洞。在系统和安全会议期刊上发表40余篇学术论文(CCF-A类20余篇)，担任中国计算机学会高级会员和系统软件/嵌入式系统专委会执行委员。入选中央网信办创新人才支持计划、电子学会青托、小米青年学者等，获得ACM ChinaSys新星奖。主持国家自然科学基金面上/青年项目、国家重大科技专项子课题、北京市自然科学基金面上项目、军科委研究计划项目、中国计算机学会科研基金等二十余个项目，研究成果应用于华为、腾讯、蚂蚁、航天科技等国内知名企业单位，获得多项应用证明和感谢信。

报告题目：操作系统缺陷检测

报告摘要：本报告围绕主讲人团队近十余年在操作系统缺陷检测方向的研究工作展开。首先，介绍传统程序分析技术在操作系统缺陷检测方面的局限性。其次，重点介绍团队提出的多项创新技术与自研工具，包括别名敏感分析算法、并发规则挖掘策略等静态分析技术，以及并发模糊测试、异常注入测试等动态分析技术。最后，介绍当前正在推进的大模型赋能操作系统缺陷检测工作，包括多维度测试用例的辅助生成、并发漏洞精确发现与复现等。



贾相堃

贾相堃，中国科学院软件研究所副研究员，CCF高级会员，中国科学院青促会成员，获中国科学院软件所“所杰青”项目支持。主持了国家自然科学基金面上项目、青年基金项目等，参与了国家自然科学基金重点项目、国家重点研发计划等。相关研究成果发表于IEEE S&P、ACM CCS、USENIX Security、NDSS、ASE等网络安全和软件工程顶会，应用于华为、蚂蚁等企业。多次担任网络安全顶会PC，担任TIFS等顶刊审稿人和信息安全学报等期刊青年编委。

报告题目：面向漏洞挖掘场景的可信智能体构造思考与实践

报告摘要：基于大模型的智能体框架为漏洞挖掘带来了新范式，其能够利用大规模知识、复杂推理能力和自主任务规划能力，辅助完成代码理解、程序测试、漏洞验证等复杂任务。然而，漏洞挖掘任务具有高度专业性和强验证需求，智能体在实际应用中仍面临幻觉、误判、推理偏差等问题。报告将面向现实复杂程序漏洞挖掘场景，分析构建可信智能体的问题与挑战，结合实际漏洞挖掘实践分享漏洞发现、分析与验证中的实践经验，探讨未来智能化漏洞挖掘的发展趋势。



吴月明

华中科技大学教授，博士生导师，国家级青年人才，湖北省海外高层次人才，华中学者卓越特聘教授。主要研究方向为软件供应链安全、人工智能安全等，在包括USENIX Security、CCS、NDSS、ICSE、FSE、ASE、ISSTA等会议/期刊上发表CCF-A类论文40余篇，获2021年湖北省科技进步奖一等奖，2025年中国通信学会科学技术奖二等奖，入选“中国精品科技期刊顶尖学术论文”，担任多个CCF-A类国际会议和期刊的程序委员会委员或审稿人。

报告题目：大模型赋能的软件供应链安全分析

报告摘要：针对传统软件供应链安全检测中数据碎片化、精度不足及治理割裂等痛点，本报告提出了一种大模型赋能的智能化闭环治理体系，并将其落地为一站式供应链安全防护产品。该体系基于海量多源数据融合的组件依赖知识图谱，构建了面向风险治理与代码审计的安全垂直智能体，实现了从SBOM生成、多维风险识别到漏洞自动验证与修复的全流程能力，效果显著优于主流传统工具及通用大模型，为可信软件供应链安全体系建设与区域产业赋能提供了创新的技术路径与实践示范。



李建彬

李建彬，天融信科技集团助理总裁，拥有17年网络与数据安全从业经验，深耕数据安全、AI安全、安全标准化及大型安全项目落地。兼任中国计算机行业协会数安专委会委员、CSA GCR云安全联盟专家会员、中国网络安全产业联盟、中国网络空间安全协会数据安全专家组成员。为中国信息安全测评中心CISP DSG/CISP DSO课程核心编撰者，信通院DSC首批数据安全专家库成员，连续四届“杰出贡献专家”，信通院数据安全大模型推进工作组成员。参编《数智安全与标准化》，发表多篇核心期刊论文。参与2项国家标准、10余项电信金融行业标准、20余项团标企标，承担多项国家级安全课题。主导设计多个大型数据安全项目，成果入选工信部试点示范、优秀案例及信通院星熠优秀案例。

报告题目：智能体应用过程中的安全思考与实践

报告摘要：当智能体开始代替人检索知识、调用工具、执行任务，安全就不再是IT部门的技术选项，而是业务敢不敢放开用AI的先决条件。本报告跳出单点防护的传统思路，围绕平台、数据、模型、应用、内容五个可信域，结合智能体目标、身份、记忆、工具、行动五个控制维度，构建“看得见、管得住、可运营”的防御体系，并以对外服务、内部办公、自动化执行三类典型业务场景验证落地成效。

人工智能安全与应用治理论坛

承办单位：中关村实验室、北京航空航天大学、中国科学技术大学、长三角安全人工智能安徽省实验室

⌚ 9月19日下午

📍 市政务会议中心8号会议室(有专车接送)

论坛主席： 吴世忠（中国工程院院士）
俞能海（中国科学技术大学）

召 集 人： 刘祥龙（北京航空航天大学、中关村实验室）

时 间	内 容
14:00-14:05	主持开场
14:05-14:15	致 辞
14:15-14:20	AI安全治理产品发布
14:20-14:50	大模型时代的智能生态安全 徐 恪（清华大学）
14:50-15:20	智能体系统安全攻防研究 杨 珉（复旦大学）
15:20-15:50	从黑盒到白盒：AI供应链透明度与端到端可追溯治理 田 雷（中电科数字技术股份有限公司）
15:50-16:00	休 息
16:00-16:30	安全筑基·产业护航：AI安全治理技术进展与应用实践 刘俊华（科大讯飞）
16:30-17:00	内容安全与认知安全 廖 清（哈尔滨工业大学（深圳））
17:00-17:30	以 AI 筑盾，护智能未来 郭天奇（新华三信息安全技术有限公司）

论坛主席



吴世忠

吴世忠，中国工程院院士。第十三届、十四届全国政协委员，中关村实验室首席科学家。长期从事网络与信息安全领域的技术研究和管理工作，科研成果先后获得国家科技进步一等奖两项，二等奖一项，省部级科技奖多项。



俞能海

俞能海，中国科学技术大学讲席教授，博士生导师，教育部高等学校网络空间安全专业教指委副主任委员，中国网络空间安全学会副理事长（法人），中国图象图形学学会会士，国家重点研发计划项目首席科学家。主要研究领域为人工智能安全、信息隐藏、数据安全、计算机视觉与安全等。近年来，已主持国家级课题20余项，发表高水平学术论文300余篇。曾获得国家网络安全先进个人，国家网络安全优秀教师，国家863先进个人奖，安徽省技术发明一等奖，军队科技进步一等奖，安徽省自然科学奖一等奖，宝钢优秀教师奖、安徽省教学成果特等奖、一等奖，安徽省教学名师等。

召集人



刘祥龙

刘祥龙，北京航空航天大学教授，中关村实验室双聘研究员，博士生导师，国家级领军人才，国家重点研发计划项目首席科学家。现任北航计算机学院党委书记、复杂关键软件环境全国重点实验室副主任、国家人工智能标准化总体组治理专题组联合组长。长期从事安全可信人工智能的教学和科研，发表人工智能和安全领域国际顶级期刊/会议论文160余篇，主持国家自然科学基金青年A类和B类项目、科技创新2030重大项目、国家重点研发计划项目等国家级项目。连续入选爱思唯尔中国高被引学者，获中国仿真学会科学技术一等奖、陕西省自然科学一等奖、中国电子学会青年科学家奖等。

报告嘉宾



徐 恪

徐恪，清华大学计算机系教授、副系主任，长期从事互联网体系结构与网络空间安全研究。入选IEEE Fellow，国家杰出青年科学基金及延续资助获得者。近五年在计算机网络与安全领域顶级会议发表论文40余篇，获USENIX Security杰出论文奖、互联网防御奖等多项国际顶级会议论文奖。完成《网络空间安全原理与实践》等多部本科核心教材。获国家技术发明奖二等奖1项，国家科学技术进步奖一等奖、二等奖各1项，中国电子学会科学技术奖科技进步奖一等奖2项。获第四届全国创新争先奖。

报告题目：大模型时代的智能生态安全

报告摘要：算力平台、大模型和智能体应用，已经形成了新的智能生态。智能生态在极大提升生产力的同时也面临着各类安全威胁，本报告首先提出了智能生态的参考架构，进一步分析了智能生态面临的安全风险并给出了各类实例，深入阐述了风险出现的原因。最后基于团队现有工作提出了可能的应对方案。



杨 珉

杨珉，复旦大学计算与智能创新学院执行院长，教授、博导，教育部长江学者特聘教授，第八届国务院学位委员会网络空间安全学科评议组成员。担任国家973项目首席科学家、国家重点研发计划首席科学家。长期从事网络空间安全研究，主要研究操作系统安全、恶意代码检测、漏洞分析挖掘和AI系统安全等，在网络安全国际四大顶会上发表论文60余篇，多次入选最佳/杰出论文奖。

报告题目：AI系统安全风险评测与治理

报告摘要：随着大模型和智能体技术快速发展，AI系统正从内容生成工具演变为能够感知环境、调用工具、执行任务并持续自主决策的复杂智能系统，其安全风险也从模型输出层面进一步延伸至真实系统运行和现实世界交互。一方面，智能体在接入代码执行环境、知识库、业务系统和关键基础设施后，可能因目标偏离、权限滥用、错误决策、恶意操控等问题，引发网络攻击、经济损失和关键业务中断等现实安全风险；另一方面，随着模型自主规划、长期行动和自我改进能力不断增强，自主复制、欺骗操控、规避监管等失控风险逐渐成为前沿AI治理面临的重要挑战。本报告将围绕智能体现实风险和前沿AI失控风险两个维度，分享团队近年来在AI系统风险发现、能力评测与安全治理方面的研究进展和思考。



田雷

田雷，中电科数字技术股份有限公司软件供应链安全技术总监，高级工程师。拥有二十余年软件工程及高安全高可靠软件质量保证技术行业经验，近年来专注于软件/AI供应链安全、SBOM/AIBOM分析与治理领域，承担多项省部级课题，构建了覆盖基座模型溯源、模型文件漏洞扫描、Skill插件恶意代码检测及生成内容合规评估的全链路检测体系，致力于破解AI应用“黑盒”带来的供应链透明度缺失与合规管控难题，为金融、交通等“关基”场景提供可控的安全底座。

报告题目：从黑盒到白盒：AI供应链透明度与端到端可追溯治理

报告摘要：随着生成式人工智能与智能体技术的规模化部署，AI供应链已成为新型网络攻击的关键攻击面。本报告论述了针对“基座模型—智能体插件—运行环境”全链路风险的系统性治理框架。基于对智能体失控、开源模型数据泄露及基础设施攻击等典型事件的深度剖析，构建了覆盖AI供应链全生命周期的安全风险分级检测体系，系统识别模型投毒、反序列化攻击、Skill供应链污染，许可合规冲突及生成内容违规等核心威胁维度，提出融合静态分析、语义解析与模型溯源的多维检测方法，已在金融、交通等关键基础设施领域验证，为AI应用上线提供端到端安全支撑。



刘俊华

刘俊华，科大讯飞研究院副院长，长三角安全人工智能安徽省实验室主任，中国翻译协会常务理事，长期从事语音及自然语言处理研究工作，主要研究方向包括声纹语种识别、多语种语音识别、机器翻译、大模型安全等。作为主要负责人参与了国家重点研发计划、工信部电子信息产业发展基金目等多项国家和省部级项目科研工作，获授权发明专利70余项，荣获国家科学技术进步奖一等奖1次、省部级科技进步奖一等奖1次。

报告题目：安全筑基·产业护航：AI安全治理技术进展与应用实践

报告摘要：当前以生成式大模型为代表的人工智能技术取得阶跃式发展，并获得广泛应用。但同时人工智能安全风险时有发生，并呈现不断增长趋势。本报告首先分析了当前人工智能安全发展趋势与面临挑战，其次介绍了科大讯飞在数据安全、内生安全、安全评测、安全防护和传播治理等方面主要技术进展，并结合人工智能多个行业的应用给出安全治理应用实践与展望，希望对人工智能行业发展有所启发。



廖清

廖清，长期从事认知安全和数据挖掘的理论研究和工程实践，先后入选教育部青年长江学者和教育部长江学者特聘教授，第十八届中国青年科技奖。在国际著名学术刊物发表了高水平学术论文100余篇。主持并承担了国家重点研发计划、自然青年、自然面上及广东省重点等共多项国家、省部级科研项目，相关研究成果获国家级二等奖及省部级一等奖共6项。在网络空间安全态势方面，围绕异常信息感知、情报态势推演、认知引导进行系统开发研制，为网信办、公安等国家重大活动提供安全保障，成果获2024年电子学会科技进步一等奖。

报告题目：内容安全与认知安全

报告摘要：认知反映了人类的情感、知识、意志和信念等，是人类物理空间向虚拟空间的延伸，是人类意识形态的核心体现。当前，社交网络凭借其强大舆论属性和社会动员能力，对人类认知的塑造和影响至关重要。本报告从如何入解析个体及群体的认知形成、演化与转变规律出发，探索对基于人工智能技术的认知安全新体系。



郭天奇

郭天奇，高级工程师，中国科学技术大学网络空间安全方向工程博士，国家关键领域未来技术领军人才，现任新华三信息技术有限公司副总裁、总工程师。曾出版《智慧城市网络安全顶层设计与实践》工程著作，拥有授权发明专利6项，参与编制国家标准1项、行业标准2项，在国际权威期刊发表论文3篇；主持国家重大科技基础设施建设项目1项、国家重点研发计划1项、省部级工程技术创新课题4项，参与部省市级科研课题十余项。

报告题目：以 AI 筑盾，护智能未来

报告摘要：随着大模型与AI智能体走向规模化落地，AI技术在释放巨大生产力的同时，提示词注入、智能体越权访问、核心数据泄露等新型安全风险集中显现。本报告搭建AI for SEC、AI in SEC、SEC for AI三位一体安全能力图谱，打造覆盖传统IT基础设施、大模型底座、AI智能体到上层应用生态的多层纵深防护体系，围绕Agent全生命周期实施精细化应用权限管控，实现安全防护与智能业务共生演进。

量子计算与网络安全论坛

承办单位：本源量子计算科技（合肥）股份有限公司

🕒 9月19日下午

📍 天鹅湖大酒店三楼11号会议室

论坛主席：郭光灿（中国科学院院士）

郭国平（中国科学技术大学）

召 集 人：郭国平（中国科学技术大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:50	当安全已经跨过rubicon河，如何迎接AI安全的Sputnik时刻 张云勇（中国联通集团网络与信息安全部）
14:50-15:30	Post-quantum Cryptography, A New Era. 丁津泰（西交利物浦大学）
15:30-16:10	量子算法：量子世界的计算之法 李绿周（中山大学）
16:10-16:20	休 息
16:20-17:00	格中最短非零向量问题：筛法与枚举的混合解法 胡红钢（中国科学技术大学）
17:00-17:40	量子计算冲击下的信息安全挑战与应对策略 张 辉（本源量子计算科技（合肥）股份有限公司）

论坛主席



郭光灿

中国科学技术大学教授，中国科学院量子信息重点实验室主任，中国科学院院士。长期从事量子光学、量子通信和量子计算的理论和实验研究。2001年获中科院自然科学二等奖，2003年获国家自然科学基金二等奖，2003年获“何梁何利”科技进步奖，2003年当选为中国科学院院士，2006年获安徽省自然科学一等奖，2007年获安徽省重大科技进步奖，2009年当选为第三世界科学院院士，2013年当选“CCTV科技创新人物”，2014年获高等学校科学研究优秀成果奖自然科学一等奖、军队科技进步一等奖，2018年获得安徽省科学技术一等奖，2020年获国家自然科学基金二等奖、安徽省科学技术一等奖。在包括Nature及子刊、Phys.Rev.Lett等国际学术期刊上发表论文700多篇，他引超过10000次。



郭国平

中国科学技术大学教授，中国科学院量子信息重点实验室副主任，安徽省量子芯片重点实验室主任，安徽省量子计算工程中心主任。在量子芯片、量子计算操作系统和量子计算软件等领域在国际学术期刊发表SCI论文270多篇，其中以第一作者、责任通信作者在Nature Communications, Nano Letters、Physics Review Letters, Apply Physics Letters等主要国际学术期刊发表SCI论文100多篇，被Science、Nature和Review of Modern Physics等重要杂志正面他引1000多次，其中两篇文章他引超过200次。申请国家专利90多项，合作撰写《量子计算与编程入门》、《QPanda量子计算编程》、《量子机器学习理论与实战》、《量子金融理论与实战》和《量子计算技术应用与专利分析》等量子计算专业图书。

召集人



郭国平

中国科学技术大学教授，中国科学院量子信息重点实验室副主任，安徽省量子芯片重点实验室主任，安徽省量子计算工程中心主任。在量子芯片、量子计算操作系统和量子计算软件等领域在国际学术期刊发表SCI论文270多篇，其中以第一作者、责任通信作者在Nature Communications, Nano Letters、Physics Review Letters, Apply Physics Letters等主要国际学术期刊发表SCI论文100多篇，被Science、Nature和Review of Modern Physics等重要杂志正面他引1000多次，其中两篇文章他引超过200次。申请国家专利90多项，合作撰写《量子计算与编程入门》、《QPanda量子计算编程》、《量子机器学习理论与实战》、《量子金融理论与实战》和《量子计算技术应用与专利分析》等量子计算专业图书。

报告嘉宾



张云勇

中国联通网络与信息安全部总经理，教授级高工，享受国务院政府特殊津贴专家、百万人才工程国家级人选，享受中国有突出贡献中青年专家称号。曾担任第十三届全国政协委员、十三届云南政协委、首批中非友好小组成员、中国经济社会理事会第五届理事、全国政协信息化专家委员会副组长、最高法信息化委员会委员、国防科工委科技委专家、工信部科技委委员、中国电子学会会士、中国通信学会会士等社会和学术职务。

报告题目：当安全已经跨过rubicon河，如何迎接AI安全的Sputnik时刻

报告摘要：报告系统阐释网络安全核心属性与攻防对抗本质，梳理网络安全防护技术四代演进路径，输出分层防御架构、全生命周期风险管控、一体化安全运营等实战体系方法。立足运营商行业视角，深度解析APT攻击、DDoS冲击、网络停服断供等极限安全场景，介绍联通多级全网流量清洗、超大容量电信级DDoS防护核心能力，以及重大活动网络安全保障落地成效。报告研判AI时代网络安全发展新形势，围绕人工智能安全理论、行业风险与落地实践展开论述，提出构建AI赋能、主动协同的立体化综合防御体系，为行业应对新型网络安全挑战提供实践参考。



丁津泰

Jintai Ding is a professor at Xian Jiaotong-Liverpool University and a Charles Phelps Taft Distinguished Professor Emeritus at the University of Cincinnati. He received his Ph.D. from Yale University in 1995. His research was in quantum affine algebras, where he was credited for the invention of the Ding-Iohara-Miki algebra. His current interest is in post-quantum cryptography. He and his colleagues developed Rainbow signature and LWE-based key exchange schemes. Rainbow was a third round candidate for the NIST post-quantum standardization process. He and his colleagues completely broke a NIST second round post-quantum signature candidate LUOV and a third round candidate GeMSS (HFEv-), for which they won the best paper honorable mention award for Crypto 2021. He is one of the designers of Kyber KEM scheme which was selected as a key establishment standard by the US National Institute of Standards and Technology (NIST), to which his patent is currently licensed.

报告题目：Post-quantum Cryptography, A New Era.

报告摘要：Public key cryptosystems (PKC) are the security foundation of modern communication systems, in particular, the Internet. However Shor's algorithm shows that the existing PKC like Diffie-Hellmann key exchange, RSA and ECC can be broken by a quantum computer. To prepare for the coming age of quantum computing, we need to build new public key cryptosystems that could resist quantum computer attacks. In this lecture, we will give an introduction to post-quantum cryptography and its recent developments, in particular, the NIST standardization process and its impact. Then we will present a practical and provably secure key exchange protocol based on the learning with errors problems, which is conceptually simple and has strong provable security properties. We will explain final the current development in PQC and it migration.

**李绿周**

中山大学计算机学院教授、量子计算与软件研究所所长、计算机科学系主任、中国计算机学会 (CCF) 量子计算专委副主任、CCF杰出会员、杰出演讲者。长期从事量子计算理论研究, 在量子算法方面形成一定研究特色, 在计算机、物理、量子信息等领域知名学术期刊及会议发表论文110多篇, 出版学术专著1部。研究工作得到国家重点研发计划、国家自然科学基金、广东省杰青、广东省量子科学战略专项等项目支持。

报告题目: 量子算法: 量子世界的计算之法

报告摘要: 量子计算的有用性有赖于量子算法。量子算法的设计具有挑战性, 需要在量子视角下找到有利于量子特性发挥作用的结构信息, 从而到达返璞归真, 大道至简的效果。基于量子计算领域的研究现状, 除了要在量子计算硬件方面做出长足努力之外, 还需进一步发展量子算法基础理论, 拓展量子计算能力边界。本报告首先介绍一些量子计算的背景知识, 然后以量子算法为核心, 介绍其发展历程、研究现状以及研究团队近年来在这方面的研究工作。

**胡红钢**

中国科学技术大学教授、博导, 密码科学技术全国重点实验室客座研究员, 从事密码、编码与网络安全的研究, 解决了国际上五个重要公开问题, 在 IEEE TIT、美密会、亚密会、IEEE S&P、CHES 等期刊和会议发表 30 多篇论文。提出了“对称 2-adic 复杂度”的概念; 发现了周期序列 2-adic 复杂度的简洁证明方法, 被称为“胡红钢方法”。2013 年荣获中国密码学会优秀青年奖, 2015 年获得国家优秀青年科学基金, 2017 年入选中科院优秀百人计划, 2024 年荣获安徽省研究生教学成果特等奖。

报告题目: 格中最短非零向量问题: 筛法与枚举的混合解法

报告摘要: 格中最短非零向量问题 (SVP) 是格密码算法的安全根基, 也是抗量子密码领域的一个核心问题, 数学家和密码学家们不断寻找更好的算法来解决SVP。目前, 筛法与枚举方法是解决SVP的两类主流方法, 各有优点和缺点。我们将这两种方法组合起来使用, 取长补短, 发展出新的算法, 得到了更好的实际运行效果。

**张 辉**

高级工程师，中国计算机学会（CCF）量子计算专委会副主任，中国科学技术大学博士。现任本源量子计算科技（合肥）股份有限公司总经理，全面统筹公司战略研发、项目规划、研发管理工作，在量子计算、量子模拟等方向取得系列重要成果，参与编辑量子计算入门与编程专著，申请及拥有30余项发明专利，有着丰富的量子计算工程化建设研发和高新技术企业项目管理经验。

报告题目：量子计算冲击下的信息安全挑战与应对策略

报告摘要：本报告从工程化与产业化双重视角，系统剖析量子计算对信息安全的深层冲击。第一部分聚焦量子计算产业发展态势，深入解析Shor算法对现有公钥密码体系的颠覆性威胁，量化评估“先存储、后解密”等现实攻击路径的风险敞口。第二部分对标美国抗量子密码（PQC）迁移政策与NIST标准化进程，梳理PQC技术路线与产业落地瓶颈，并重点呈现本源量子“攻防兼备”的技术体系——既具备量子计算攻击能力验证优势，又拥有完整的PQC解决方案与迁移实施能力，为我国关键信息基础设施的量子安全转型提供可落地的工程路径与产业化应对策略。

数据流通利用的隐私计算、合规与治理论坛

承办单位：西安电子科技大学、中国科学院信息工程研究所

🕒 9月19日下午

📍 天鹅湖大酒店三楼8号会议室

论坛主席：方滨兴（中国工程院院士）
 李凤华（中国科学院信息工程研究所）
 李 晖（西安电子科技大学）

召 集 人：李 晖（西安电子科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	数据要素流通利用与安全 李凤华（中国科学院信息工程研究所）
14:40-15:10	个人信息精准保护与可控利用关键技术及标准化 范科峰（中国电子技术标准化研究院）
15:10-15:40	人工智能模型安全及伦理研究 朱天清（澳门城市大学）
15:40-16:00	休 息
16:00-16:30	Graph Learning with Differential Privacy: Applications and Challenges 叶青青（香港理工大学）
16:30-17:00	华为鲲鹏可信数据空间解决方案 胡科开（华为技术有限公司）
17:00-17:10	烛域数据定价智能体发布 李 晖（西安电子科技大学）

论坛主席



方滨兴

中国工程院院士，广州大学网络空间安全学院名誉院长。现任中国中文信息学会理事长，中国标准化协会网络与信息安全技术委员会主席。先后担任过中国网络空间安全协会理事长、中国互联网协会副理事长/网络与信息安全工作委员会主任、中国计算机学会副理事长/计算机安全专业委员会主任、中国通信学会副理事长/通信安全技术委员会主任；国家自然科学基金委员会网络与信息安全重大计划专家组成员、可信软件重大计划专家组副组长。



李凤华

中国科学院特聘研究员，中国科学院百人计划终期优秀，国务院学位委员会网络空间安全学科评议组成员，中央网信办云计算服务安全评估专家组成员；国家科技创新 2030“大数据重大工程”立项论证和实施方案编写专家；中国网络空间安全学会常务理事，中国中文信息学会常务理事、大数据安全与隐私计算专业委员会主任，中国通信学会理事、期刊与出版工作委员会副主任、学术工作委员会委员等；获 2018 年网络安全优秀人才奖、2001 年国务院政府特殊津贴。



李 晖

国家级教学名师，全国密码专业学位研究生教育指导委员会委员，中国网络空间安全学会常务理事，中国密码学会理事、大数据与人工智能安全专委会副主任，中国中文信息学会理事、大数据安全与隐私计算专委会副主任。主要从事密码理论与应用、隐私计算、数据安全、信息论与编码等领域研究。国家重点研发计划项目负责人，国家自然科学基金重点项目负责人，移动互联网安全陕西省创新团队负责人。

召集人



李 晖

国家级教学名师，全国密码专业学位研究生教育指导委员会委员，中国网络空间安全学会常务理事，中国密码学会理事、大数据与人工智能安全专委会副主任，中国中文信息学会理事、大数据安全与隐私计算专委会副主任。主要从事密码理论与应用、隐私计算、数据安全、信息论与编码等领域研究。国家重点研发计划项目负责人，国家自然科学基金重点项目负责人，移动互联网安全陕西省创新团队负责人。

报告嘉宾



李风华

中国科学院信息工程研究所二级研究员、技术副总师、中国科学院特聘研究员、博士生导师，灾备技术国家工程研究中心主任，中国科学院“百人计划”学者。国务院学位委员会网络空间安全学科评议组成员，中央网信办云计算服务安全评估专家组成员，全国数据标准化技术委员会WG5工作组副组长，国家科技创新2030重大项目某安全防护系统总体总师、国家重点研发计划项目负责人、国家863计划主题项目首席专家、NSFC联合基金重点项目负责人等；中国中文信息学会常务理事、大数据安全与隐私计算专业委员会主任，中国通信学会理事、期刊与出版工作委员会副主任、学术工作委员会委员，中国网络空间安全学会常务理事、咨询与评议工作委员会副主任等；《网络与信息安全学报》执行主编，《WWW》《CJE》《电子学报》《通信学报》《电子与信息学报》《西安电子科技大学学报》编委等。主要从事隐私计算、网络与系统安全、数据安全等方面研究，在国内外首次提出了隐私计算的概念、定义和学术内涵，并开创了隐私计算理论研究，在IEEE TIFS、TDSC、TMC、INFOCOM等国内外期刊或国际会议上发表学术论文150余篇、获授权发明专利100余项，主持或参与制订国家、行业和团体标准多项，获2018年网络安全优秀人才奖、2001年国务院政府特殊津贴、2003年中共中央办公厅“十佳青年”等。获国家技术发明二等奖1项、省部级科技进步（或技术发明）一等奖6项。

报告题目：数据要素流通利用与安全

报告摘要：数据要素是国家基础性战略资源，并成为推动数字经济发展的核心引擎。DT时代数据从传统共享演化为要素流通，重点关注数据权属确定、权益转移、使用控制、争议仲裁等。当前，数据要素流通利用与安全亟需体系化的技术解决方案，本报告介绍了数据共享与数据流通利用的本质差异、数据要素流通利用与安全的研究范畴、数据要素准则，着重剖析了数据确权、延伸使用控制、数据出域安全计算、安全交付、低开销监测等概念及学术内涵，并全面阐述了数据要素流通利用与安全的关键技术与发展趋势、技术解决方案。



范科峰

正高级工程师，中国电子技术标准化研究院副院长、学术委员会主任委员，电子信息产品标准化国家工程研究中心主任，新一代人工智能标准与应用工业和信息化部重点实验室主任，2016年获国务院政府津贴专家。兼任中国网络安全产业联盟秘书长、工信部商用密码产业促进联盟理事长、中国自动化学会常务理事、中国电子学会理事、全国区块链与分布式记账技术标准化技术委员会秘书长。主要从事网络安全、人工智能等领域研究工作。

报告题目：个人信息精准保护与可控利用关键技术及标准化

报告摘要：数字时代，个人信息保护已成为关乎人民群众切身利益与数字经济健康发展的基础性命题。如何在复杂开放的网络环境中实现个人信息的精准保护？如何在保障安全的前提下释放数据要素价值、实现可控利用？如何在海量异构场景下度量并验证保护效果？现有技术体系在效用一致约束机理、多维敏感性自适应度量、跨域流通动态合规等关键环节长期面临瓶颈制约，亟需系统性突破。本报告将围绕“个人信息精准保护与可控利用关键技术及标准化”这一主题，系统介绍一套覆盖个人信息全生命周期的保护与利用技术体系。本报告将围绕上述技术突破与治理挑战，从理论创新、标准引领、工程实践三个维度开展深入剖析与探讨。



朱天清

澳门城市大学数据科学学院院长、教授、博士生导师，长期从事人工智能安全、数据隐私保护与网络空间安全方向研究。主持国家自然科学基金（NSFC）、澳门特别行政区科学技术发展基金（FDCT）、澳大利亚研究理事会基金（ARC）等纵向科研项目累计科研经费超过2000 万元人民币。围绕人工智能安全与隐私保护领域的关键科学问题，提出了一系列具有国际影响力的理论方法与系统技术，在 USENIX Security Symposium、Network and Distributed System Security Symposium、IEEE Symposium on Security and Privacy、IEEE Transactions on Information Forensics and Security、IEEE Transactions on Dependable and Secure Computing 等国际顶级会议和期刊发表论文 60 余篇。两次入选美国斯坦福大学“全球前 2% 终身科学影响力科学家榜单”。现担任多个国际知名期刊编委，并长期担任 ACM Conference on Computer and Communications Security、USENIX Security Symposium、International Joint Conference on Artificial Intelligence、AAAI Conference on Artificial Intelligence 等国际顶级会议程序委员会委员。

报告题目：人工智能模型安全及伦理研究

报告摘要：随着人工智能技术在科研、产业、公共服务等领域的深度普及与迭代升级，各类大语言模型、生成式模型及多模态系统已成为推动社会变革的核心驱动力，但随之而来的安全风险与伦理挑战也日益凸显，成为制约技术健康可持续发展的关键瓶颈。本文围绕人工智能模型全生命周期展开研究，系统梳理其面临的核心安全问题，涵盖数据泄露、模型投毒、对抗样本攻击、输出内容失真、算法偏见与歧视、深度伪造滥用等典型风险类型，深入剖析风险产生的技术根源与现实诱因。同时，从公平性、透明性、可解释性、责任归属、隐私保护、价值对齐等维度，探讨人工智能应用所涉及的核心伦理议题，分析当前技术治理、行业规范及法律法规建设中存在的短板与不足。在此基础上，结合国内外人工智能安全伦理领域的最新研究进展与实践经验，从技术防护体系构建、伦理治理框架完善、法律法规与标准体系建设、多方协同监管机制搭建、公众素养提升等多个层面，提出兼具可行性与前瞻性的应对策略与优化路径。研究旨在为人工智能模型的安全开发、规范应用及伦理治理提供理论参考与实践指引，助力实现技术创新与安全伦理的协同平衡。



叶青青

Qingqing Ye is an associate professor in The Hong Kong Polytechnic University. She received her PhD degree from Renmin University of China in 2020 and joined the Hong Kong Polytechnic University since then. Her research interests include data privacy and security, and adversarial machine learning. She has published over 100 top-tier conference and journal papers in related areas, including IEEE S&P, CCS, USENIX Security, NDSS, SIGMOD, VLDB, ICDE, INFOCOM, ICML, NuerIPS, VLDBJ, TKDE, TMC, TDSC, TIFS, etc. She is the recipient of several prestigious awards, including RGC Early Career Award, PolyU Research Grant Achievement Award, and IEEE S&P Travel. She serves as Area Chair and TPC member for top-tier conferences, including VLDB, ICDE, ICML, NeurIPS, SIGKDD, AAI, WWW, etc.

报告题目: Graph Learning with Differential Privacy: Applications and Challenges

报告摘要: Graph-structured data arise in social networks, recommendation systems, biological networks, and many other applications, but their analysis can expose sensitive information about individuals and their relationships. This talk explores how differential privacy provides rigorous protection for graph learning while enabling useful data analysis. I will first introduce the fundamental privacy model of differential privacy, and then discuss three representative applications. The first is privacy-preserving link prediction, which aims to infer missing or future relationships without revealing sensitive connections. The second is differentially private influence maximization, where influential seed nodes are selected for downstream tasks while protecting graph structure and user participation. The third is private graph generation, which produces synthetic graphs that preserve important structural and statistical properties of the original network. Through these applications, the talk will highlight key algorithmic ideas, privacy-utility trade-offs, and open challenges in developing practical and trustworthy graph-learning systems.



胡科开

胡科开, 美国麻省大学博士, 华为计算安全高级专家, 国际机密计算联盟(CCC)理事, 主要研究方向包括机密计算、隐私计算、芯片安全等技术领域。加入华为前曾任 Intel 首席安全研究员, 是 Intel STORM 安全团队创始人之一, 曾担任多个 IEEE 组织如 ComSoc、SPCE、TCSP 等技术委员会 (TPC) 成员, 美国 SRC 半导体研究联盟芯片安全项目 Intel 联络员。曾发现 20 多个 CVE 安全漏洞, 在国内外拥有 20+ 安全专利。

报告题目: 华为鲲鹏可信数据空间解决方案

报告摘要: 随着数据成为关键生产要素, 多源数据的跨域融合与高效流通已成为数字经济发展的核心驱动力, 但“数据权属难界定、隐私泄露风险高”等痛点限制了数据要素价值的深度释放。本报告系统介绍基于华为鲲鹏底座的可信数据空间解决方案。该方案以鲲鹏硬件为信任根, 深度整合鲲鹏机密计算、隐私算力加速、混淆推理等多种技术, 构建“软硬一体”的安全防护体系, 使能华为客户实现数据全生命周期“可用不可见、可控可计量”, 为构建合规、安全的数据流通基础设施提供技术支撑。

智安融合、双向赋能 | 协同构建 自主网络攻防大模型新生态论坛

承办单位：中国网络空间安全学会智能化网络安全专业委员会、中国科学院信息工程研究所
北京长亭科技有限公司

🕒 9月19日下午

📍 天鹅湖大酒店三楼1号会议室

论坛主席：郭世泽（中国科学院院士）

孟 丹（中国科学院信息工程研究所）

召 集 人：孟 丹（中国科学院信息工程研究所）

霍 玮（中国科学院信息工程研究所）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	物理智能体深空使命与安全测评 黄殿中（中国工程院院士）
14:40-15:00	人工智能赋能网络安全对抗 陈庶樵（中关村实验室）
15:00-15:20	智能化渗透测试的进化：前沿模型与统驭框架 周昌令（北京大学）
15:20-15:40	Mythos启示：大模型自主攻防能力构建与安全厂商赋能路径 杨 坤（北京长亭科技有限公司）
15:40-16:00	休 息
16:00-16:20	基于国产基础模型的安全Harness能力建设 马 昕（阿里云计算有限公司）
16:20-16:40	人工智能企业的智能与安全融合实践探索 鲍景雨（上海阶跃星辰智能科技股份有限公司）
16:40-17:00	现背景下芯片内生安全实践探讨 焦振男（海光信息技术股份有限公司）
17:00-17:20	智能化漏洞挖掘的技术实践与思考 霍 玮（中国科学院信息工程研究所）

论坛主席



郭世泽

中国人民解放军信息技术安全研究中心研究员，国防科技大学兼职博导。长期从事网络空间安全领域研究，在系统脆弱性分析与主动防御方面取得系统性创新成果，提出基于行为特征的漏洞发现与评估方法，主持研制多型网络安全装备，为我国关键信息基础设施安全防护提供技术支撑。曾获国家科技进步奖多项。2023年当选为中国科学院院士。



孟 丹

中国科学院信息工程研究所研究员，中国网络空间安全协会副理事长、关键信息基础设施安全保护专业委员会主任委员，中国科学院信息工程研究所研究员，国务院政府特殊津贴专家。主持参与曙光 2000、3000 和 4000 系列高性能计算机研制，承担国家 863 计划、自然科学基金等项目，获国家科技进步二等奖 3 项。

召集人



孟 丹

中国科学院信息工程研究所研究员，中国网络空间安全协会副理事长、关键信息基础设施安全保护专业委员会主任委员，中国科学院信息工程研究所研究员，国务院政府特殊津贴专家。主持参与曙光 2000、3000 和 4000 系列高性能计算机研制，承担国家 863 计划、自然科学基金等项目，获国家科技进步二等奖 3 项。



霍 玮

中国科学院信息工程研究所研究员，致力于软件安全分析与评测研究，带领团队在软件成分分析与缺陷定位、程序动静态分析、定向漏洞挖掘等方面取得重要突破。曾主持国防重大工程项目、国家重点研发计划课题、国防领域基金重点等，承担或参与国家级省部级科研项目 40 余项。

报告嘉宾



黄殿中

中国工程院信息与电子学部和管理学部院士，通信与信息安全和工程管理专家；组织攻坚完成国家级重大专项工程，取得历史性突破，攻克多项国家重点研发计划专项，填补空白，荣获国家科技进步一等奖 2 项，国家技术发明二等奖 1 项。

曾任国家某部领导、部科技委主任，第十七届、第十八中央纪律检查委员会委员；现任某部科技战略咨询咨委委员，中国信息安全测评中心学术委员会副主任，赛博空间国家重点实验室学术委员会主任。

报告题目：物理智能体深空使命与安全测评

报告摘要：本报告立足第一届中国网络空间安全大会“智安融合、双向赋能 | 协同构建自主网络攻防大模型新生态”论坛，围绕《物理智能体的深空使命与安全测评》展开研讨。当前月球南极成为大国深空博弈焦点，美国依托阿尔忒弥斯协定构建排他太空秩序，我国倡导人类命运共同体理念，亟需工程能力予以支撑。报告提出，批量部署蜂群式物理智能体集群是月球南极开发换道超车的核心路径，依靠分布式无人作业实现快速勘测、月面就地建造，为后续载人登月筑牢基础；以产业智能体作为顶层中枢，实现集群调度、风险管控与天地迭代。针对深空环境下物理、智能、对抗三重安全威胁，构建硬件-算法-系统-集群四层安全测评体系。报告指出，深空安全测评既可抢占国际规则话语权，又能技术外溢赋能地面 AI 安全。呼吁产业协同、深空开拓、共建全球太空治理新秩序，实现智安融合、天地双向赋能，为自主网络攻防大模型生态建设提供深空视角的借鉴。



陈庶樵

中关村实验室科研04部负责人，研究员，北京理工大学、北京邮电大学博士研究生联培导师，中科院信工所客座教授，信息支援部队信息组专家，CCSS应用密码工程专委会副主任，CCF网络与系统安全专委会常务委员，电力监控系统安全防护专业学科组委员。主要研究方向为主动防御和人工智能。获国家科技进步二等奖1次，军队科技进步一等奖2次，省科技进步一等奖1次，专利50余项。

报告题目：人工智能赋能网络安全对抗

报告摘要：本报告将系统性描述大模型与智能体驱动网络攻防范式的革新：从自动化信息收集、智能化漏洞挖掘，到精准化漏洞利用与高效打点渗透，全链路验证AI在安全攻防中的实战效能；并结合真实场景案例，揭示智能技术对攻击效率的颠覆性提升，为构建主动式、自适应防御体系提供关键思路，凸显AI时代的安全新逻辑。



周昌令

北京大学计算中心网络安全室主任、北京大学长沙计算与数字经济研究院网络安全研究中心副主任。负责北京大学网络安全体系建设、管理与应急响应；带领团队在教育部攻防演习中连续两年获“攻防双优”，2024年夺得攻防双项高校第一；担任国家安全演习裁判、教育部攻防演习裁判组组长及北京市教育行业漏洞响应平台专家组组长；主持多项网络安全、人工智能安全与大数据分析领域科研项目，发表论文专利数十篇，主编《网络安全技术》国家级规划教材，担任多项全国顶级网络安全赛事命题专家与裁判。

报告题目：智能化渗透测试的进化：前沿模型与统驭框架

报告摘要：当前智能化渗透测试的技术瓶颈正从模型本身的推理能力转向多智能体统驭框架的编排，而攻克该瓶颈的关键在于确保执行过程与结果的可靠与可信。针对这一挑战，本报告将介绍“渗透即程序”的创新理念，该理念摒弃了传统对模型启发式自主发挥的单一依赖，通过将动态、复杂的渗透测试行为转化为一段有类型、有语义、可检查的结构化程序，从而有效提升长程渗透测试在实际应用中的稳定性和可控性。



杨 坤

长亭科技联合创始人&首席安全研究员，清华大学计算机系博士，第二十五届北京优秀青年工程师，网络空间安全高级工程师，加州大学伯克利分校访问学者，受邀在Black Hat USA大会发表主题演讲。国内首支闯入顶级国际黑客大赛DEF CON CTF总决赛的“蓝莲花”队长，创办全球首个基于真实软件的CTF国际大赛Real World CTF。专注提升网络安全领域漏洞挖掘、渗透测试和风险评估技术水平，打造业界最强攻防团队，常年在国家和地方的实网攻防演练、CTF大赛、漏洞挖掘专项赛等获得第一名的佳绩。

报告题目：Mythos启示：大模型自主攻防能力构建与安全厂商赋能路径

报告摘要：以 Mythos 为代表的大模型，正展现出突出的漏洞挖掘、攻击路径规划与自主任务执行能力，基于国产大模型构建自主可控的智能化攻防能力，已成为必然发展方向。本报告将重点分享长亭围绕大模型自主构建攻防能力智能体的实践经验，覆盖漏洞挖掘与自动化渗透测试等关键攻防场景，并深入探讨如何以智能体重构新一代 Agentic 安全运营体系。在这一实践过程中，团队沉淀形成了高质量的实战数据与方法成果。为进一步推动国产大模型安全能力的整体提升，长亭推出了“国产大模型安全能力提升数据集开放计划”，报告将同步分享当前在行业内的合作进展与共建情况。

**马 昕**

阿里云资深安全技术专家，专注于安全大模型、智能代码安全与 AI 驱动的攻防技术研究，主导 Qwen-Aegis 安全大模型及代码安全、AI-BAS 等产品能力建设

报告题目：国产大模型攻防能力现状与待突破方向

报告摘要：大模型技术正在重塑网络空间安全攻防范式，国产安全大模型已从通用代码与文本理解，逐步深入到漏洞分析、渗透测试辅助、威胁检测等真实攻防任务中。本次分享以阿里云安全大模型 Qwen-Aegis 的建设实践为主线，系统呈现当前国产大模型攻防能力的真实水平、关键瓶颈与待突破方向。在能力现状层面，Qwen-Aegis 已在代码审计、漏洞挖掘与验证、补丁生成等任务上取得显著进展，并在 CyberGym、Cybench、CVE-Bench 等国内外评测中验证了其潜力。同时，实践表明同一基座模型在不同 Agent/Harness 下的表现差异显著，Harness 的工具编排、环境交互与结果验证能力正成为安全大模型能力放大的关键变量。在发展路线上，Qwen-Aegis 明确从“榜单分数导向”转向“真实任务能力导向”：收敛聚焦少数与真实攻防高度相关的核心评测，并在授权渗透、代码审计、漏洞验证与修复等真实场景中持续打磨模型能力。

**鲍景雨**

阶跃星辰安全风控总监，研究方向为智能体安全、AI for Security

报告题目：人工智能企业的智能与安全融合实践探索

报告摘要：随着大模型和智能体进入企业核心流程，智能与安全呈现双向融合趋势：安全为大模型和智能体提供可信、可靠、可控的运行底座，智能则以机器速度重塑威胁发现、响应处置与安全治理。围绕智能与安全双向赋能，本次分享将分场景介绍人工智能企业在企业安全、人工智能基础设施安全、模型与智能体业务安全中的实践，探讨智能体安全如何支撑智能体从试验走向生产，以及大模型和智能体如何推动安全运营走向自动化和规模化。

**焦振男**

海光信息技术方案部科研行业总经理，曾深度参与超节点白皮书等行业规范撰写，CCF成员，深耕科研高性能计算、机密计算多年

报告题目：现背景下芯片内生安全实践探索

报告摘要：AI时代进一步加剧了对传统外挂式安全的挑战，海光芯片内生安全如何应对？本报告将从海光全栈内生、全线加密、自主可控、可信免疫多维度进行介绍。

**霍 玮**

中国科学院信息工程研究所研究员，致力于软件安全分析与评测研究，带领团队在软件成分分析与缺陷定位、程序动静态分析、定向漏洞挖掘等方面取得重要突破。曾主持国防重大工程项目、国家重点研发计划课题、国防领域基金重点等，承担或参与国家级省部级科研项目40余项。

报告题目：智能化漏洞挖掘的技术实践与思考

报告摘要：人工智能技术的飞速发展正在重塑网络攻防格局。AI在漏洞挖掘、payload生成、渗透规划等方面展现的能力，在显著降低攻击门槛的同时，也使得网络威胁的形态更加复杂多变。攻防双方在AI能力上的博弈，正成为网络安全领域最核心的前沿命题。本报告聚焦AI技术与网络攻防的深度融合，围绕AI如何赋能漏洞挖掘、自动化验证以及如何弥补模型能力差距导致的攻防短板等方面，分享最新研究成果与实践案例。

网络空间安全科技女性发展论坛

承办单位：天融信科技集团、中关村实验室、哈尔滨工业大学网络空间安全学院
 武汉大学国家网络安全学院、华东师范大学信息与电子工程学院

① 9月19日下午

② 天鹅湖大酒店一楼淮河厅B

论坛主席：李雪莹（天融信科技集团）
 张宏莉（哈尔滨工业大学）

卿 昱（中关村实验室）
 王 鹃（武汉大学）

召 集 人：李雪莹（天融信科技集团）
 张宏莉（哈尔滨工业大学）

卿 昱（中关村实验室）
 王 鹃（武汉大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:20	致 辞
14:20-14:45	密码技术与人工智能安全（拟） 王小云（中国科学院院士）
14:45-15:10	网络空间安全态势感知：技术、系统与应用 贾 焰（国防科技大学）
15:10-15:35	大模型智能体时代的安全防护体系构建 李雪莹（天融信科技集团）
15:35-15:45	休 息
15:45-16:10	大模型智能体运行时安全：间接提示注入的时序因果溯源与上下文净化防御 王 鹃（武汉大学）
16:10-16:35	具身智能的“安全边界”：感知-决策-执行全链路风险分析 冀振燕（北京交通大学）
16:35-17:00	量子韧性 护航AI安全新未来——“量子防未来、机密护当下、可信贯全程”的国产化安全实践 金 添（安徽问天量子科技股份有限公司）
17:00-17:30	圆桌论坛：AI时代网安科技女性人才发展机遇与挑战 主持人：谢 琴（天融信科技集团） 嘉 宾：吴文玲（中国科学院软件研究所） 李丽香（北京邮电大学） 钮 艳（中国工业互联网研究院安全研究所） 高 丹（赛迪顾问股份有限公司）

论坛主席



李雪莹

博士，天融信科技集团党委书记、董事长兼CEO；任清华大学计算机系专业实践企业指导教师、山东大学特聘教授；中国网络空间安全学会常务理事及女科技工作者工作委员会主任等职务。长期致力于安全管理与安全防护体系构建、安全大数据分析、云安全、人工智能+安全等领域的技术创新和成果转化，先后承担十余项部委课题研究工作，获多项省部级科学技术进步奖和社会科技奖、中国产学研合作创新奖、北京市三八红旗奖章，申请发明专利十余项，在国家级学术刊物及核心期刊共发表学术论文三十余篇，参与制定国家、行业及团体标准十余项，并主编多部著作。



卿 昱

中关村实验室副主任、首席科学家、正高级工程师。长期从事网络安全科研和管理工作，在网络空间防御、对抗和社会安全等领域取得多项国内、军内首创性成果。担任国家科技部“网络空间安全治理”重点专项组长，负责网络空间安全治理技术论证和“十四五”布局。作为国家某重大专项实施方案编写组组长、专项副总师，牵头完成项目实施方案论证编制工作，获党中央、国务院批复。曾获国防科学技术进步奖一等奖、北京市三八红旗奖章。



张宏莉

哈尔滨工业大学二级教授、博士生导师，网络空间安全学院创始院长，入选国家高层次人才计划、国家网络安全优秀教师；兼任国务院学位办网络空间安全学科评议组秘书长，教育部网络空间安全教学指导委员会委员，中国网络空间安全学会常务理事，中国通信学会网络与数据安全专委会副主任；主要研究领域为信息内容安全、认知安全、网络与系统安全等。获省部级科研奖励一二等奖5项、省教学成果奖一二等奖3项；出版专著和教材5部。



王 鹃

武汉大学国家网络安全学院教授，博士生导师。任中国网络空间安全学会女科技工作者委员会副主任、CCF计算机安全专委会常务委员、信息安全系主任等职务。从事系统安全与可信计算、AI安全等研究20余年，主持国家科技部重大专项课题、国家自然科学基金面上、国家电网、华为、浪潮等项目。在CCS、NDSS、NIPS、CVPR、TIFS、TDSC等会议和期刊上发表论文80余篇；出版学术专著6本，授权和申请国家发明专利30余项；获湖北省科技进步一等奖、二等奖、CCF容错计算卓越成果奖。

召集人



李雪莹

博士，天融信科技集团党委书记、董事长兼CEO；任清华大学计算机系专业实践企业指导教师、山东大学特聘教授；中国网络空间安全学会常务理事及女科技工作者工作委员会主任等职务。长期致力于安全管理与安全防护体系构建、安全大数据分析、云安全、人工智能+安全等领域的技术创新和成果转化，先后承担十余项部委课题研究工作，获多项省部级科学技术进步奖和社会科技奖、中国产学研合作创新奖、北京市三八红旗奖章，申请发明专利十余项，在国家级学术刊物及核心期刊共发表学术论文三十余篇，参与制定国家、行业及团体标准十余项，并主编多部著作。



卿 昱

中关村实验室副主任、首席科学家、正高级工程师。长期从事网络安全科研和管理工作，在网络空间防御、对抗和社会安全等领域取得多项国内、军内首创性成果。担任国家科技部“网络空间安全治理”重点专项组长，负责网络空间安全治理技术论证和“十四五”布局。作为国家某重大专项实施方案编写组组长、专项副总师，牵头完成项目实施方案论证编制工作，获党中央、国务院批复。曾获国防科学技术进步奖一等奖、北京市三八红旗奖章。



张宏莉

哈尔滨工业大学二级教授、博士生导师，网络空间安全学院创始院长，入选国家高层次人才计划、国家网络安全优秀教师；兼任国务院学位办网络空间安全学科评议组秘书长，教育部网络空间安全教学指导委员会委员，中国网络空间安全学会常务理事，中国通信学会网络与数据安全专委会副主任；主要研究领域为信息内容安全、认知安全、网络与系统安全等。获省部级科研奖励一二等奖5项、省教学成果奖一二等奖3项；出版专著和教材5部。



王 鹃

武汉大学国家网络安全学院教授，博士生导师。任中国网络空间安全学会女科技工作者委员会副主任、CCF计算机安全专委会常务委员、信息安全系主任等职务。从事系统安全与可信计算、AI安全等研究20余年，主持国家科技部重大专项课题、国家自然科学基金面上、国家电网、华为、浪潮等项目。在CCS、NDSS、NIPS、CVPR、TIFS、TDSC等会议和期刊上发表论文80余篇；出版学术专著6本，授权和申请国家发明专利30余项；获湖北省科技进步一等奖、二等奖、CCF容错计算卓越成果奖。

报告嘉宾



王小云

清华大学高等研究院杨振宁讲座教授，中国科学院院士，发展中国家科学院院士，国际密码协会会士。现任中国密码学会理事长，中国网络空间安全学会副理事长，密码与数字经济安全全国重点实验室主任。

先后四次获得国际密码学三大顶会最佳论文奖。曾获国家科技进步一等奖、自然科学二等奖、北京市突出贡献中关村奖、全国三八红旗手、全国优秀科技工作者、世界杰出女科学家奖、国际密码协会最具时间价值奖、未来科学大奖、求是杰出科学家奖、陈嘉庚科学奖等。

报告题目：密码技术与人工智能安全

报告摘要：随着人工智能技术应用广泛，算法的可控性与模型参数的保密性以及数据真实性等安全问题日益凸显，而密码学能够为人工智能构建可验证的数学安全边界。基于密码学视角的人工智能安全新理论和防护体系是值得思考的前沿科学问题。本报告将聚焦密码技术与人工智能安全的前沿交叉研究方向，包括大模型安全、AI for Crypto以及Crypto for AI，讨论如何基于密码分析视角构建AI安全新理论、推动智能化密码分析与设计，促进密码学与人工智能的共同发展。



贾 焰

国防科大研究员，现任国家重大攻关项目负责人，工业控制系统信息安全技术国家工程研究中心首席，中国中文信息学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用；作为课题负责人承担和主持了国家级重大/重点项目20余项；获国家科技进步二等奖5项；发表进入SCI和EI检索的论文300余篇，出版专著8部，获得150余项发明专利；获何梁何利基金科学与技术进步奖，获全国巾帼建功标兵称号。

报告题目：网络空间安全态势感知：技术、系统与应用

报告摘要：报告从网络空间安全态势感知的概念和重要地位出发，首先给出了态势感知的三大应用需求：内容安全舆情事件态势感知，系统安全攻击事件态势感知，以及攻防对抗态势评估；其次给出针对该应用需求的三大挑战：准确、实时和全面；再次给出解决这些挑战的多维关联认知模型MDATA及其发展技术；最后给出相关理论和技术的应用和效果。



李雪莹

博士，天融信科技集团党委书记、董事长兼CEO；任清华大学计算机系专业实践企业指导教师、山东大学特聘教授；中国网络安全安全学会常务理事及女科技工作者工作委员会主任等职务。

长期致力于安全管理与安全防护体系构建、安全大数据分析、云安全、人工智能+安全等领域的技术创新和成果转化，先后承担十余项部委课题研究工作，获多项省部级科学技术进步奖和社会科技奖、中国产学研合作创新奖、北京市三八红旗奖章，申请发明专利十余项，在国家级学术刊物及核心期刊共发表学术论文三十余篇，参与制定国家、行业及团体标准十余项，并主编多部著作。

报告题目：大模型智能体时代的安全防护体系构建

报告摘要：随着人工智能技术快速发展，大模型在各业务领域获得广泛应用，大模型技术已迈向“规模化应用与智能体落地”的新阶段，安全风险正呈现出“跨层交织、链式传导”的复杂特征，这对安全防护体系提出了更高要求。本报告剖析大模型应用安全风险的演进与案例，讨论大模型安全防护的架构与手段，推动大模型技术从“可用”迈向“可信”。



王 鹃

武汉大学国家网络安全学院教授，博士生导师。任中国网络安全安全学会女科技工作者委员会副主任、CCF计算机安全专委会常务委员、信息安全系系主任等职务。从事系统安全与可信计算、AI安全等研究20余年，主持国家科技部重大专项课题、国家自然科学基金面上、国家电网、华为、浪潮等项目。在CCS、NDSS、NIPS、CVPR、TIFS、TDSC等会议和期刊上发表论文80余篇；出版学术专著6本，授权和申请国家发明专利30余项；获湖北省科技进步一等奖、二等奖、CCF容错计算卓越成果奖。

报告题目：大模型智能体运行时安全：间接提示注入的时序因果溯源与上下文净化防御

报告摘要：工具增强的大模型智能体在 multi-step 执行中持续累积对话历史、工具返回与检索内容，外部不可信内容由此进入上下文并影响决策，形成间接提示注入攻击面。多轮场景下，注入载荷可经早期读操作潜入并跨轮驻留，直至在写权限边界触发越权动作，造成用户意图与中介内容混同的归因难题。本报告从“时序因果接管”的建模视角构建智能体运行时防御框架 AgentSentry：在工具返回边界恢复状态快照，以无副作用的反事实重执行判定待执行动作系用户目标支撑抑或中介驱动。系统能定向净化涉事内容，剥离指令信号而保留任务证据，并在净化状态下修订动作与门控副作用，使智能体续行原任务而非中断。报告将结合基准与真实运行时评测，展望智能体安全方向。


冀振燕

工信部人工智能标准委员会WG8副组长、工信部人形机器人与具身智能标准委员会委员、工信部应急装备标准委员会委员；CCF软件工程专委会执行委员、CCF区块链专委会执行委员、SCI期刊《Scientific Data》副主编等。在国内外期刊、会议发表学术论文百余篇，出版专著1部、教材3部、译著2部，申请发明专利十余项，荣获中国电子学会科学技术奖二等奖等科研奖励4项。研究方向：人工智能、人工智能安全、区块链、多模态数据融合等。

报告题目：具身智能的“安全边界”：感知-决策-执行全链路风险分析

报告摘要：具身智能通过赋予智能体物理实体与环境交互能力，正加速从实验室走向工业制造、家庭服务、自动驾驶等真实场景。当智能系统拥有“身体”并直接作用于物理世界时，其安全风险从数字空间外溢至现实空间，呈现出系统性、连锁性和高危害性的新特征。报告聚焦具身智能“感知-决策-执行”全链路，系统剖析其安全边界与风险机理。在感知层，多模态传感器面临对抗样本攻击、物理环境干扰与传感器欺骗等威胁；在决策层，大模型驱动的规划与推理存在幻觉、价值对齐偏差等问题；在执行层，机械臂、人形机器人等的高自由度运动控制存在安全约束突破、人机协作冲突及物理伤害风险。报告还讨论了具身智能安全标准规范及伦理治理等，以期为具身智能的安全可信发展提供统一技术要求。


金 添

本科毕业于北京理工大学，硕士毕业于清华大学，现任安徽问天量子科技股份有限公司副总裁，北京启源问天量子科技有限公司总经理；工信部网络安全专家库成员，可信计算产业联盟理事，北京市商密协会理事会委员。深耕信息安全和密码领域20年，多次主导大型信息安全项目以及量子密码创新项目，致力于推动量子技术与安全的融合、抗量子密码体系迁移、量子+AI+机密计算融合创新。2023年重点推动“5G物联网密码云服务基础设施建设方案”获得工信部和国密局联合颁发的国家级奖项。作为第一或第二发明人获得“基于量子随机数的ID生成方法及装置”发明专利、“一种可配置的多模式混合密钥协商应用方法及系统”等多项发明专利，参与编著《信息安全策略》、《电子健康卡指南》，获得密码技术应用员认证、CISP认证。

报告题目：量子韧性 护航AI安全新未来——“量子防未来、机密护当下、可信贯全程”的国产化安全实践

报告摘要：针对AI时代面临的数据合规与价值矛盾、“先窃取、后解密”的量子威胁及模型黑盒信任缺失三大安全困局，该架构以“量子防未来、机密护当下、可信贯全程”三重防护，打造AI专属量子隐形斗篷。架构融合QKD、PQC与量子真随机数抵御未来量子攻击，依托TEE机密计算实现当下数据“可用不可见”，并借可信计算3.0构建全链路信任体系。同时，问天量子推出自研Q-INSIDE战略，依托量子物理客观特性，从密钥、算法、随机数三个维度三重加固现有密码体系。该战略无需大规模重构业务系统，通过轻量化配置即可为云平台、AI系统及物联网设备加载抗量子安全能力，将量子安全深度嵌入各行业数字化底座，实现密码体系平滑升级。该方案为政务、金融等行业提供国产化完整方案，彻底解决数据“不敢上云、不敢流转”等痛点。

■ 圆桌嘉宾



谢 琴
天融信科技集团



吴文玲
中国科学院软件研究所



李丽香
北京邮电大学



钮 艳
中国工业互联网研究院
安全研究所



高 丹
赛迪顾问股份有限公司

工业互联网安全论坛（2030项目首席论坛）

承办单位：西安交通大学网络空间安全学院、中国科学院信息工程研究所

🕒 9月19日下午

📍 天鹅湖大酒店二楼1号会议室

论坛主席：管晓宏（中国科学院院士）

孙利民（中国科学院信息工程研究所）

召 集 人：孙利民（中国科学院信息工程研究所）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	工业互联网安全及智能防御 吕金虎（北京航空航天大学）
14:40-15:10	新型电力系统安全攻防关键技术 田志宏（广州大学）
15:10-15:40	自主可控电力专用网络安全防护基础平台关键技术 关振宇（北京航空航天大学）
15:40-16:00	休 息
16:00-16:30	工业互联网安全威胁智能分析 王冲华（国家工业信息安全发展研究中心）
16:30-17:00	下一代特种工业互联网智能安全对抗技术 孙利民（中科院信息工程研究所）

论坛主席



管晓宏

教授、博士生导师，中国科学院院士，西安交通大学电子与信息学部主任，国家杰出青年科学基金获得者，教育部长江学者特聘教授，IEEE Fellow。长期从事系统工程理论与应用研究。针对能源、电力和互联网等网络化系统，提出了系统优化、物理安全与信息安全的新理论与新方法，解决了多个公认难题，取得了系统性创新成果，为能源电力等系统的提效节能和保障安全做出了重要贡献。曾获得国家自然科学二等奖、中国自动化学会自然科学一等奖以及美国李氏基金杰出成就奖和多篇国际最佳论文奖。



孙利民

中国科学院信息工程研究所二级研究员，博士生导师，中科院工业控制系统安全方向的学术带头人，2030 项目首席。围绕国家在工控系统安全的短板问题，带领团队研发系列化工控安全检测与防护设备，在国家重要行业得到应用。在 S&P、CCS、TOSN、TDSC、JSAC 等国内外期刊或会议上发表论文 300 余篇，出版学术书籍 6 部，申请发明专利 120 余项，牵头或参与国家标准 10 余项，获省部级一等奖 4 项、二等奖 3 项。

召集人



孙利民

中国科学院信息工程研究所二级研究员，博士生导师，中科院工业控制系统安全方向的学术带头人，2030 项目首席。围绕国家在工控系统安全的短板问题，带领团队研发系列化工控安全检测与防护设备，在国家重要行业得到应用。在 S&P、CCS、TOSN、TDSC、JSAC 等国内外期刊或会议上发表论文 300 余篇，出版学术书籍 6 部，申请发明专利 120 余项，牵头或参与国家标准 10 余项，获省部级一等奖 4 项、二等奖 3 项。

报告嘉宾



吕金虎

北京航空航天大学党委常委、副校长，国家杰青/创新群体负责人，IEEE会士，国家工业互联网产教融合创新平台负责人。中国指挥与控制学会副理事长、中国仿真学会副理事长、“十五五”国家重点研发计划智能工业软件专项专家组组长。从事复杂系统协同控制、工业互联网等研究。WoS总他引2.6万余次，全球高被引科学家。曾获3项国家自然科学基金二等奖、全国创新争先奖、何梁何利科技进步奖、国家级教学成果二等奖、钱学森杰出贡献奖等。

报告题目：工业互联网安全及智能防御

报告摘要：新一代人工智能技术与工业互联网深度融合，体系化重塑工业互联网安全的发展，工业互联网面临一系列安全的新风险、新挑战。报告简要分析当前人工智能时代工业互联网安全面临的一些新风险、新挑战，重点介绍高安全智能化 PLC 的智能防御能力构建，包括融合形式化验证与内存沙箱的主动免疫机制、基于内核行为检测的实时异常拦截等，以及在大飞机制造等典型场景的示范验证。



田志宏

广州大学党委常委、副校长。教授、博士生导师，教育部长江学者特聘教授，广东省珠江学者特聘教授，国家重点研发计划项目首席科学家，曾先后担任广州大学网络空间先进技术研究院院长、网络空间安全学院院长。长期致力于网络攻防对抗、网络靶场、主动实时防护等网络空间安全热点领域的研究工作。作为负责人，先后主持了多项国家自然科学基金、国家重点研发计划项目/课题、国家863课题、工信部项目、国家242课题和中央网信办课题，研究成果获省部级奖励4次。发表论文近200篇，著作3部。

报告题目：新型电力系统安全攻防关键技术

报告摘要：新型电力系统推动电网向开放互联的信息物理系统演进，安全边界由网络空间拓展至电磁、声波等物理信号域，跨域、隐蔽、协同的攻击特征日益凸显。本报告围绕智能电网“信号—网络”多域协同攻击，系统探讨感控设备信号攻击机理、漏洞挖掘与监测防御、协同仿真及态势量化关键技术，并结合虚实联动靶场推演与典型场景验证，构建全链条防御能力。



关振宇

北京航空航天大学网络空间安全学院教授、博士生导师，国家杰出青年科学基金获得者、教育部青年长江学者。长期从事网络空间安全、内生安全防御与可信计算等方向研究。现任区块链与物联网应用教育部重点实验室副主任，主持国家重大科技专项等各类科研项目35项。近年在S&P、ACM CCS、TPAMI、TIP、TDSC等国际顶级会议与期刊发表学术论文百余篇，获北京市自然科学奖一等奖、吴文俊人工智能科学技术奖等省部级科技奖励6项。

报告题目：自主可控电力专用网络安全防护基础平台关键技术

报告摘要：新型电力系统海量终端接入、协议异构、跨境互联加深，攻击由单点渗透演化为组合攻击，外挂式边界防护存在结构性局限。本报告提出以内生安全为核心的纵深防护思路，将安全能力内嵌于芯片、编译、协议与全局感知各层级，构建电力网络“免疫系统”。报告重点介绍四项关键技术：动态异构冗余防御模型与多维拟态策略驱动的差模免疫方法，支撑自主免疫引擎；基于多核异构分区隔离与混合粒度运算阵列的可重构安全芯片，实现毫秒级安全功能重构与后量子密码部署；漏洞“疫苗”注入、小样本未知缺陷甄别与程序指纹防篡改改编译技术；密码定义的多因素持续认证与跨层协议安全增强机制。最后结合新能源监控、主配微协同、跨境电网互联等场景，探讨内生安全工程化落地的性能开销与存量兼容挑战。



王冲华

正高级工程师，国家工业信息安全发展研究中心监测应急所副所长，中国电子学会网络空间安全专委会委员，中国青年科技工作者协会委员。主要从事工业网络安全、工业智能安全等技术研究，主持多项工业网络安全领域政策规划、标准编制工作，承担国家科技重大专项、国家重点研发计划、国防基础科研、工业互联网创新发展工程等30余项国家级/省部级项目课题，获省部级奖励科技进步奖5项。

报告题目：工业互联网安全威胁智能分析

报告摘要：工业互联网正以前所未有的广度和深度加速驱动更大范围、更深层次、更高水平的产业要素融合与系统重构，未来工业互联网在感知执行、控制决策、工业互联、应用服务等方面更趋于智能化，安全技术呈现协同化、自演化、自适应化、智能化等特征。报告聚焦未来工业互联网安全战略、安全技术保障体系、产业应用等方面，讨论相关趋势演进、技术突破与重点产业链安全态势。

**孙利民**

中国科学院信息工程研究所二级研究员，博士生导师，中科院工业控制系统安全方向的学术带头人，2030项目首席。围绕国家在工控系统安全的短板问题，带领团队研发系列化工控安全检测与防护设备，在国家重要行业得到应用。在S&P、CCS、TOSN、TDSC、JSAC等国内外期刊或会议上发表论文300余篇，出版学术书籍6部，申请发明专利120余项，牵头或参与国家标准10余项，获省部级一等奖4项、二等奖3项。

报告题目：下一代特种工业互联网智能安全对抗技术

报告摘要：下一代特种工业互联网覆盖无人装备“科研生产-实网应用”全业务流程多场景，是低空经济运行与国防新型装备的重要支撑，呈现智能演进和科研生产军民融合趋势。国际上典型安全事件说明特种工业互联网面临科研生产场景的环境安全、实网应用场景无人装备的功能安全和高对抗场景的智能安全面临严重威胁。报告首先介绍了下一代特种工业互联网的显著特征和当前安全现状，然后分析了特种工业互联网安全防护的挑战问题，进一步提出其安全防护的核心理念，最后从本体安全模型、链式安全对抗、协同检测与安全增强等方面给出了关键研究技术。

信息隐藏技术创新论坛

承办单位： 国际关系学院网络空间安全学院、上海大学

🕒 9月19日下午

📍 天鹅湖大酒店一楼徽州厅

论坛主席： 郭云彪（国际关系学院）

张新鹏（上海大学）

召 集 人： 张新鹏（上海大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	信息隐藏的些许思考 郭云彪（国际关系学院）
14:40-15:10	反智能技术初探 操晓春（中山大学）
15:10-15:40	适应开放信道的鲁棒隐蔽通信技术 罗向阳（信息工程大学）
15:40-16:10	高泛化性和可解释的AIGC检测技术 赵 耀（北京交通大学）
16:10-16:30	休 息
16:30-17:00	基于公有链的隐匿通信 祝烈煌（北京理工大学）
17:00-17:30	网络态语义信息隐藏技术 黄永峰（清华大学）
17:30-18:00	基于数字水印的媒体内容确权与防伪溯源 钱振兴（复旦大学）

论坛主席



郭云彪

国际关系学院二级教授，一直从事信息隐藏、多媒体安全、隐蔽通信等相关领域的研究工作，是特种通信国家科技创新团队的领头人，是我国数字隐写技术的开拓者和领军者，建立并发展了面向秘密信息传递的数字隐写技术理论和方法，攻克秘密信息传递“抗发现、抗关联、抗分析”三大技术难题，推动隐蔽通信技术的换代升级，实现相关工作模式的重塑。2009年享受国务院特殊津贴，2013年被评为国家中青年科技创新领军人才，2015年入选国家百千万人才工程，先后获国家科技进步一等奖2次（排1、11），国家科技进步二等奖2次（均排2），国家科技发明二等奖1次（排6）。累计获部级科技进步奖33项。



张新鹏

上海大学计算机学院院长，国家杰出青年科学基金获得者，二级教授，复旦大学“数字文化保护与旅游数据智能计算”文化和旅游部重点实验室主任。入选上海市东方英才计划领军项目、上海市优秀学术带头人、上海市曙光人才计划、上海市“东方学者”跟踪计划、上海市浦江人才计划、上海市“青年科技启明星”跟踪计划。长期从事信息隐藏、多媒体安全与人工智能安全研究，先后主持国家自然科学基金重点项目、原创探索项目、国家重点研发计划、国家863计划等科研项目40余项。发表论文300余篇，连续十二年入选“爱思唯尔”中国高被引学者榜单，并入选“科睿唯安”全球高被引科学家。申请发明专利30余项，授权近20项。获上海市自然科学奖一等奖、安徽省自然科学奖一等奖、国家级教学成果二等奖。担任IEEE TIFS等国际学术期刊的Associate Editor、ACM IH&MMSec与IEEE WIFS等国际学术会议的程序委员会主席和大会主席。

召集人



张新鹏

上海大学计算机学院院长，国家杰出青年科学基金获得者，二级教授，复旦大学“数字文化保护与旅游数据智能计算”文化和旅游部重点实验室主任。入选上海市东方英才计划领军项目、上海市优秀学术带头人、上海市曙光人才计划、上海市“东方学者”跟踪计划、上海市浦江人才计划、上海市“青年科技启明星”跟踪计划。长期从事信息隐藏、多媒体安全与人工智能安全研究，先后主持国家自然科学基金重点项目、原创探索项目、国家重点研发计划、国家863计划等科研项目40余项。发表论文300余篇，连续十二年入选“爱思唯尔”中国高被引学者榜单，并入选“科睿唯安”全球高被引科学家。申请发明专利30余项，授权近20项。获上海市自然科学奖一等奖、安徽省自然科学奖一等奖、国家级教学成果二等奖。担任IEEE TIFS等国际学术期刊的Associate Editor、ACM IH&MMSec与IEEE WIFS等国际学术会议的程序委员会主席和大会主席。

报告嘉宾



郭云彪

一直从事信息隐藏、多媒体安全、隐蔽通信等相关领域的研究工作，主持完成20余项省部级指令性项目；主持完成“973课题、863课题、国家科技支撑、自然科学基金等公开项目10余项。获国家科技进步一等奖2次，国家科技进步二等奖2次，国家科技发明二等奖1次。

报告题目：信息隐藏的些许思考

报告摘要：报告旨在同大家分享信息隐藏的一些思考。主要哲学层面分享信息隐藏到底研究什么，从本体论层面分享隐藏何以可能，从认知论层面分享隐藏何以成功，从学科论层面分享信息隐藏从“术”到“学”的演进。通过思考试图揭示信息隐藏的运行机制，构建信息隐藏的理论体系，搭建信息隐藏的学科骨架，推进信息隐藏学术发展。



操晓春

中山大学信息学部副主任、网络空间安全学院院长，国家杰出青年/优秀青年基金获得者。

报告题目：反智能技术初探

报告摘要：当前主流的人工智能模型大多依赖数据驱动的函数拟合，其鲁棒性面临自然环境的烟雨雾霾等复杂扰动以及人为设计的数字物理世界对抗攻击双重挑战。报告拟介绍团队在人工智能的鲁棒性与脆弱性、多模态融合感知与安全、认知域态势感知与安全、全生命周期的安全评估等方面的近期进展。内容不仅涵盖数字域物理域对抗样本、后门攻击、隐私等传统脆弱性问题，还介绍团队在归因机制建模、反事实生成、正向扰动辅助等方向的技术探索



罗向阳

罗向阳，信息工程大学教授、博导，河南省网络空间态势感知重点实验室主任，中原学者、国家级科技领军人才、国防卓青获得者、国家重点研发计划项目负责人。主要从事网络安全、信息隐藏和多媒体智能安全领域研究，先后主持国家、军队和省部级重点项目30余项，在IEEE TPAMI、TIFS、TDSC, ACM CCS、《中国科学》等国内外顶级期刊会议发表论文100余篇，获河南省自然科学一等奖（序1）、河南省科技进步一等奖（序1）、国家教学成果一等奖（序5）、河南省教学成果特等奖（序1）各1项。

报告题目：适应开放信道的鲁棒隐蔽通信技术

报告摘要：立足于开放网络环境下隐蔽通信的可靠性与安全性需求，将系统梳理国家战略牵引下数字空间面临的复合扰动、主动检测与动态信道漂移等现实挑战，并将指出传统基于单载体的鲁棒隐写难以应对开放信道固有的系统级不确定性，需向多域协同的端到端体系演进。研究将从信道构造、图像鲁棒隐写、新型载体隐写三个维度展开：提出Hook-Steg框架以规避平台有损处理机制；分类探讨修改式、生成式、选择式图像隐写方法，给出抗压缩、缩放及复合攻击的解决方案；并拓展至区块链交易、视频弹幕、网络数据流等新型载体，构建多维隐蔽信道，以期在复杂开放环境中实现隐蔽通信可靠性与安全性的协同提升。



赵耀

北京交通大学教授、太原科技大学学术副校长、长江学者特聘教授、国家杰出青年科学基金获得者、万人计划科技创新领军人才、IEEE Fellow。现任“视觉智能交叉创新”教育部国际联合实验室主任、“科幻音视频智能处理”北京市重点实验室主任。研究领域包括图像\视频压缩，数字媒体内容安全，媒体内容分析与理解，人工智能等。主持了新一代人工智能项目、973计划等项目30余项。在包括IEEE Trans.等国内外期刊、会议上发表论文200余篇。作为第一完成人获北京市科学技术奖一等奖等省部级奖励5项、世界互联网大会领先科技奖1项。

报告题目：高泛化性和可解释的AIGC检测技术

报告摘要：生成式人工智能的迅猛发展对数字取证与安全领域产生了深远影响，催生出高度逼真的合成图像，其真实性与真实视觉素材几近难以区分。尽管这些创新为创意产业带来了变革性机遇，但同时也引发了AI生成内容检测，以及防范其在虚假信息传播和欺诈活动中滥用的严峻挑战。本报告将重点介绍我们近期在推进AI生成图像检测方面的工作进展，着重提升检测精度并优化多种生成方法的泛化能力。同时，我们还将探讨可解释性AI驱动解决方案的开发，强调建立稳健、可扩展且具备解释性的应对策略对于缓解AI生成内容日益加剧的威胁具有迫切必要性。



祝烈煌

北京理工大学网络空间安全学院党委书记，特聘教授，博士生导师，入选国家高层次领军人才。任国家重点研发计划“区块链”专项专家组成员、“基础科研条件与重大科学仪器设备研发”专项总体组专家，中国网络空间安全学会常务理事、中国人工智能学会常务理事、智能信息网络专委会主任，中国计算机学会理事、区块链专委会主任、中国密码学会理事。长期从事网络与信息安全方向的教学与研究工作，主持了国家重点研发计划项目、国家自然科学基金重点项目等国家级、省部级科研项目50余项。近年来，出版外文专著5本，发表高水平学术论文400余篇，获多个国际会议最佳论文奖，获电子学会一等奖等省部级科技奖励8项。

报告题目：基于公有链的隐匿通信

报告摘要：隐蔽通信、匿名通信在保护通信内容的同时，还可以保护通信行为和通联关系，在特殊应用场景发挥重要作用。目前已经开展了大量基于互联网的隐蔽和匿名通信网络的研究工作，但是仍然存在信道易检测、通信节点易识别、服务节点易攻陷等问题。以比特币、以太坊为代表的公有链上的活跃用户数和交易数规模大并快速增加，且具有地址匿名、泛洪转发、去中心化等特性，为构建基于公有链的隐蔽通信和匿名通信提供非常好的平台。报告主要介绍基于区块链的隐匿通信的现状趋势、技术挑战、系统模型、科学问题、关键技术、系统应用等。



黄永峰

清华大学电子工程系教授，博导。首届全国网络安全“十佳”优秀教师；闽江学者讲座教授；ACM SIGWeb中国分会副主席；全国卫生产业企业管理协会数智医药科技分会主任委员；中文信息学会开源情报技术专委会副主任委员。从事网络信息的智能处理与安全等技术研究。主持了国家自然科学基金委的重大（课题）、国际合作、重点和面上等项目等10余项；科技部973、863、重大和重点专项等课题20多项。担任多家国际著名期刊的编委以及重要国际会议主席，及Nature等国际期刊评审；在Nature子刊、IEEE/ACM Tran.、中国科学等国内外著名期刊和重要会议发表学术论文300多篇；出版专著6部、译著2部，教材2部。授权发明专利30余项。获省部级科技成果奖一等奖3项，二等奖4项；省部优秀教材一等奖1项、教学成果大奖赛二等奖1项等。

报告题目：网络态语义信息隐藏技术

报告摘要：随着知识本体和认知智能等技术发展，传统基于媒体信号特征的信息隐藏技术正向网络数据的语义特征的信息隐藏新范式演变。报告阐述了网络环境下信息隐藏的特点，定义了网络态信息隐藏的概念，提出了一套知识本体和大模型认知智能相结合的神经网络模型驱动的网络态语义信息隐藏框架，分析了网络态语义信息隐藏评价体系的演变趋势；给出了一些网络态语义信息隐藏实践案例。

**钱振兴**

复旦大学计算与智能创新学院教授、博导，入选国家高层次领军人才计划，担任中国网络空间安全学会科普工委主任、文旅部重点实验室副主任，主要从事多媒体安全、人工智能应用等方面的研究，主持国家自然科学基金重点等各类项目三十余项。主要从事数字媒体内容安全研究，在IEEE TPAMI, TIFS, CVPR, NeurIPS, ICLR等国际期刊与会议发表文章两百余篇，担任IEEE TCSVT, TCYB, TASLP等期刊的编委，著有图书《人工智能安全简明教程》《计算机教授白话人工智能》《隐藏的故事》等，获上海市自然科学奖一等奖、上海市科技进步奖二等奖、文津图书提名等奖项。

报告题目：基于数字水印的媒体内容确权与防伪溯源

报告摘要：数字水印已有多年的研究历史，从早期基于信号处理的数字水印，发展到当今基于深度学习的水印，全世界涌现了多种数字水印方法。本次报告将回顾数字水印发展的状况，侧重介绍复旦大学多媒体智能安全实验室在该方向的探索，包括AIGC水印、屏幕内容溯源水印、数字防伪水印、物理防伪水印、模型水印以及相关应用等。

认知安全创新论坛

承办单位：中国电科网络空间研究院、中关村实验室、拓尔思信息技术股份有限公司

🕒 9月19日下午

📍 天鹅湖大酒店一楼长江厅

论坛主席：陆 军（中国工程院院士）

召 集 人：卿 昱（中关村实验室）

时 间	内 容
14:00-14:05	主持开场
14:05-14:15	致 辞
14:15-14:45	论道：智能时代认知之“道” 陆 军（中国工程院院士）
14:45-15:15	智能时代下认知对抗与安全的新思考 程学旗（中国科学院计算技术研究所）
15:15-15:45	在喧嚣中筑坝，于迷雾中引航：认知安全的边界与策略 钟昊熹（新华社媒体融合生产技术与系统国家重点实验室）
15:45-16:05	休 息
16:05-16:35	构建以我为主的网络话语权新格局 魏南枝（中国社会科学院国际合作局）
16:35-17:05	国家认知安全体系构建与关键技术研究 姜 伟（哈尔滨工业大学）
17:05-17:35	数力筑基：智能时代认知攻防的一观三面 施水才（拓尔思信息技术股份有限公司）

论坛主席



陆 军

中国工程院院士，中国电科集团首席科学家。长期从事雷达及综合电子信息系统领域理论研究和项目研制工作。先后担任我国首型国产预警机，首型出口预警机总设计师，为我国综合电子信息系统技术进步和装备发展作出重大贡献，获国家科技进步特等奖等奖励。

召集人



卿 昱

中关村实验室副主任、首席科学家。长期从事网络空间安全治理与认知技术研究。担任国家网络空间安全科技重大专项副总师，编制网络安全审查办法等重要政策文件。获授权国家发明专利多项，获国防科技进步奖、军队科技进步奖等重要奖励，享受国务院政府特殊津贴。

报告嘉宾



陆 军

长期从事雷达及综合电子信息系统领域理论研究和项目研制工作。先后担任我国首型国产预警机，首型出口预警机总设计师，为我国综合电子信息系统技术进步和装备发展作出重大贡献，获国家科技进步特等奖等奖励。

报告题目：论道：智能时代认知之“道”

报告摘要：陆军院士系统论述智能时代认知对抗的“道、术、策”体系，重点剖析认知制胜之“道”，揭示认知对抗的深层逻辑：技术改变传播方式，规律决定最终胜负，得道多助、失道寡助在智能时代依然是铁律。报告以中国特色社会主义理论为理论根基，将“道”概括为规律、价值、正义、实践四重内涵——以实事求是把握认知生成规律，以人民立场筑牢道义根基，以知行合一贯通塑造全程，以实践成效检验工作得失，从中国式现代化理论和实践中提炼出感知、研判、塑造、评估、修正的方法，实现以“道”驭“术”、以“道”定“策”，把思想优势、技术优势、组织优势转化为认知领域的战略主动。



程学旗

程学旗，研究员，中科院计算所副所长，智能算法安全全国重点实验室主任、国家杰出青年科学基金获得者，国家高层次特支人才，北京学者，CCF会士，IEEE Fellow。长期从事网络数据科学等方向研究。在重要会议和IEEE汇刊上发表论文300余篇，七次获本领域顶会最佳论文奖，谷歌学术引用超三万次，获授权发明专利143项。关键技术与成果应用于国家三大平台，服务所有省市相关业务，获国家技术发明二等奖1项、国家科技进步二等奖3项。

报告题目：智能 + 认知的机会与挑战

报告摘要：随着人工智能快速发展，智能技术正由信息处理工具逐步融入信息感知、知识组织、分析研判和决策支持等认知过程，不断拓展人类认识复杂世界和协同解决问题的能力，为科技创新、产业发展和社会治理带来新的机遇。与此同时，智能技术深度嵌入认知活动，也使相关风险由数据、模型等单点问题逐步演化为贯穿信息生成、传播、理解、判断和行动全过程的系统性挑战。报告将结合大数据、社会计算和人工智能发展实践，围绕智能与认知融合的发展趋势，分析其面临的机遇与挑战，探讨构建安全、可信、可控的智能认知体系，更好促进认知安全能力建设。



钟昊熹

钟昊熹，新华社高级编辑。现任新华社媒体融合生产技术与系统国家重点实验室副主任、研究员；新华融合媒体科技发展（北京）有限公司副总裁、董事、党委委员。曾入选中组部“万人计划”领军人才、宣传思想文化青年英才等国家级人才。主要研究领域为融媒体科学与技术、区域国别与时空智能、智能传播与认知安全。

报告题目：在喧嚣中筑坝，于迷雾中引航：认知安全的边界与策略

报告摘要：认知安全是总体国家安全观视域下的新兴前沿领域，其本质已超越传统信息安全与网络安全的范畴，直指人的心智“操作系统”本身。当前，认知安全面临多重挑战，形势日趋严峻。本报告基于对新闻媒体传播实战的观察研究，提出具有一定创新性的认知安全策略框架，为筑牢国家认知安全屏障提供系统性对策参考。



魏南枝

魏南枝，女，中共党员。中国社会科学院国际合作局副局长，研究员，博士生导师，享受国务院政府特殊津贴，2025年获评中央国家机关三八红旗手。全国港澳研究会理事，中国政治学会理事，中国社会科学院新时代党建研究中心理事，中国共产党巡视理论研究中心理事，社会主义民主研究中心理事。在《求是》、《人民日报》、《学术月刊》等发表研究、理论文章90余篇；多次任社科基金重大项目、课题负责人，多次获中国社会科学院信息对策奖项。

报告题目：构建以我为主的网络话语权新格局

报告摘要：世界百年未有之大变局加速演进，国际格局深刻调整，认知博弈是大国竞争的关键博弈维度之一，网络空间是认知博弈的主阵地。我国在维护国家安全、加强网络治理等方面，正处于从“被动应对”逐步转向“以我为主”的转型期。对此，要坚持攻防并重战略，主动参与全球网络治理规则制定，升级算法治理；开展多样化互联网国际传播，鼓励支持“走出去”力量和青年群体等，针对不同国家、地区、群体等进行精准化分众化交流传播；要建立跨部门协同战略机制，应对跨境认知作战、跨境虚假叙事，避免国际舆论场被动挨打，国内共识受到冲击；要将我国自身技术与基础设施能力建设，与我国倡导的普惠开源“一带一路”互联网能力建设相结合，推进网络空间的人类命运共同体建设。



姜 伟

姜伟，哈尔滨工业大学网络空间安全学院院长、计算学部副主任，二级教授、博导，国家级高层次人才。曾任中央网信办中国网络空间研究院网络安全研究所所长，系 2030 国家科技重大专项首席科学家、国家社科基金重大项目首席专家、全国网络安全标准化技术委员会副秘书长。长期从事网络安全、数据安全、人工智能安全与网络治理研究，参与国家网络安全重大系统建设及重大活动保障。主持省部级以上课题 30 余项，发表论文 50 余篇，报送内参 100 余篇，40 余篇获中央及省部级领导批示。牵头国家及 10 余省市“十四五”“十五五”规划编制。

报告题目：国家认知安全体系构建与关键技术研究

报告摘要：报告从认知机理切入，分析认知偏差的形成规律，阐释了认知安全的概念边界与重要意义。系统梳理国内外在认知安全领域的研究布局与最新进展，分析面临的认知安全态势与现实挑战。创新性地提出了国家认知安全治理模型，系统勾勒出认知安全治理的顶层设计与体系框架，给出保障认知安全面临的关键挑战，亟待突破的科学问题，支撑认知安全的关键技术，并介绍团队在相关核心技术上的研究积累与基础。报告立足国家网络认知安全重大需求，从认知机理、认知安全到治理体系，从风险挑战、科学问题到关键技术，层层递进，系统给出了筑牢国家认知安全屏障的方案。

**施水才**

施水才，拓尔思信息技术股份有限公司董事长，中国中文信息学会会士（Fellow）、中国软件行业协会副理事长，中国人工智能等学会副主任委员、常务委员。领导拓尔思公司于2011年在创业板上市。曾主持20多项国家级项目，获国家科技进步二等奖、上海市科技进步一等奖、全国优秀科技工作者、中国电子学会科技进步一等奖、北京市科技进步二等奖、教育部科技进步二等奖等荣誉，北京信息科技大学教授，南京大学、北京师范大学等高校兼职教授。

报告题目：数力筑基：智能时代认知攻防的一观三面

报告摘要：生成式人工智能、平台算法与智能体正在重构社会信息传播和认知形成机制，认知安全风险由单点内容风险演变为贯穿数据、模型、分发与反馈的系统性链路风险。数据是认知安全的关键生产要素，数力则是将数据转化为认知感知、机理研判、事件推演、策略预演与行动支撑的能力。本报告提出认知攻防的“一观三面”：以数力筑牢认知安全基础，以智能化构建技术引擎，以场景化形成价值落点，以体系化贯通数据、语义、模型、智能体、业务流程与安全治理。结合拓尔思在数据采集治理、动态本体、行业知识、智能体编排和可信调用等方面建设的数力基础设施，介绍其面向复杂事件的多源数据融合、对象关系建模、演化路径分析、多智能体推演与多分支情景仿真实践，探讨认知安全从态势感知和被动响应走向预测预警、策略预演、协同处置与持续演进的技术路径。

国家一流网安学院院长论坛

承办单位：山东大学网络空间安全学院、中国科学院信息工程研究所
 中国网络空间安全学会人才工作委员会

⌚ 9月19日下午

📍 天鹅湖大酒店一楼贵宾厅

论坛主席：王美琴（山东大学）

刘宝旭（中国科学院信息工程研究所）

召 集 人：王美琴（山东大学）

时 间	内 容
14:15-14:20	主持开场
14:20-14:25	致 辞
14:25-14:55	功能安全视角解决无人系统安全问题 尹 浩（中国科学院院士）
14:55-15:20	智能时代的网络防御：主动防御与安全优先计算 孟 丹（中国科学院信息工程研究所）
15:20-15:45	大规模多领域融合网络靶场技术、系统与应用 贾 焰（国防科技大学）
15:45-16:05	休 息
16:05-16:30	网络空间安全人才培养探索与实践 苏 洲（西安交通大学）
16:30-16:55	智驭未来，韧性赋能：人工智能时代网络安全前沿技术攻关与产教融合新范式 李雪莹（天融信集团）
16:55-17:40	Panel (AI重塑网络空间安全：学科变革、人才培养与创新突破) 王美琴（山东大学） 李小勇（北京邮电大学） 苏 洲（西安交通大学） 王 骞（武汉大学） 唐 卓（湖南大学）

论坛主席



王美琴

国家高层次人才计划，现任山东大学副校长、密码与数字经济安全全国重点实验室常务副主任，兼任中国密码学会常务理事、中国网络空间安全学会常务理事等。主要从事对称密码分析与设计的研究，发表论文百余篇，包括国际密码五大顶会59篇。主持国家重点研发计划、国家自然科学基金重点项目等，获全国创新争先奖章、国家密码科技进步一等奖（省部级）等科技奖励。



刘宝旭

中国科学院信息工程研究所党委委员、技术副总师、学术委员会副主任，研究员，享受政府特殊津贴。兼任中国科学院大学网络空间安全学院教授，第五教研室主任，网络安全防护技术北京市重点实验室主任，中国网络空间安全学会人才工委会主任委员。主要从事网络攻防对抗、威胁感知发现溯源的研究。主持国家重点研发计划、863、国家发改委重大工程等项目，获省部级科技进步一等奖4项、二等奖5项，获中国产学研合作创新奖1项。发表论文百余篇，出版著作7本。

召集人



王美琴

国家高层次人才计划，现任山东大学副校长、密码与数字经济安全全国重点实验室常务副主任，兼任中国密码学会常务理事、中国网络空间安全学会常务理事等。主要从事对称密码分析与设计的研究，发表论文百余篇，包括国际密码五大顶会59篇。主持国家重点研发计划、国家自然科学基金重点项目等，获全国创新争先奖章、国家密码科技进步一等奖（省部级）等科技奖励。

报告嘉宾



尹 浩

中国科学院院士，军事科学院系统工程研究院研究员，博士生导师，通信网络领域专家，复杂系统仿真总体重点实验室主任。长期从事通信网络理论方法研究、体系结构设计和技术应用等方面研究，主持完成多项国家重点科研与工程建设项目，获国家和省部级科技进步奖多项。目前主要研究方向为物联网和卫星通信技术。

报告题目：功能安全视角解决无人系统安全问题

报告摘要：无人系统在各领域规模化应用，复杂开放环境下的未知风险、传统防护模式的适配局限，叠加智能算法的动态特性，使得无人系统安全保障面临严峻挑战，其安全性已成为制约技术落地与产业发展的核心瓶颈。在此背景下，为突破传统安全防护困境、构建适配性安全保障体系，本报告聚焦并深化无人系统安全本质规律的认知，梳理功能安全与预期功能安全的评判逻辑，剖析未知场景、感知局限等内在不确定性及典型应用的安全威胁，反思传统信息防护式安全手段的局限性，提出从“关注安全问题”转向“关注系统自身”，构建包含设计时/运行时安全、确定性知识、防御类型的三维认知框架，明确面向无人系统的安全设计原则与关键技术，为推动无人系统在复杂环境下的安全可靠运行提供理论新思路与实践新方向。



孟 丹

中国科学院信息工程研究所研究员、博士生导师，信工所原党委书记、所长，全国政协委员、国务院政府特殊津贴专家，担任中国网络空间安全协会副理事长、中国网络空间安全学会副理事长、国家自然科学基金委信息科学部第八届专家咨询委员会专家、“国家百千万人才工程”入选者。主要从事计算机体系结构、网络系统安全等研究工作，被评为国家网络安全优秀人才，获得国家科技进步二等奖三项、中国科学院杰出科技成就奖和科技进步一等奖各一项、保密科技一等奖两项，发表多篇高水平学术论文。

报告题目：智能时代的网络防御：主动防御与安全优先计算

报告摘要：人工智能时代，随着以 Mythos 为代表的前沿基座大模型的出现，高危漏洞挖掘与利用步入机器时间，攻击效率大幅提升，攻防不对称进一步加剧，未知威胁攻击将成为主流，传统基于攻击特征的防御手段难以为继。报告在分析大模型给网络防御带来的全新重大挑战的基础上，提出以主动防御应对未知威胁攻击的思路。报告阐述主动防御的威胁模型、防御目标、技术思路和主要手段，提出主动防御需由当前基于攻击特征的被动防御转向基于攻击机理的对抗性防御，在攻击者突破边界防御、侵入网络系统并获取相应权限的场景下，仍能掌握系统控制权，保障核心业务持续运行与关键数据安全。报告指出，主动防御要以计算机系统结构提供的硬性支持为基础保障，并介绍了报告人团队在该方向的研究成果，包括安全优先计算理论、技术突破及应用进展。



贾 焰

国防科技大学研究员。现任国家重大攻关项目负责人，中国中文信息学会副理事长等职。主要研究方向为人工智能和大数据分析技术在网络空间安全领域中的应用；作为课题负责人承担和主持了国家级重大/重点项目20余项，获国家科技进步二等奖5项；发表SCI/EI论文300余篇，出版专著8部，获发明专利150余项，软件著作权160余项，牵头制定相关行业标准40余项；获何梁何利基金科学与技术进步奖，获全国巾帼建功标兵称号。

报告题目：大规模多领域融合网络靶场技术、系统与应用

报告摘要：本报告从网络安全领域的网络攻防技术实验验证、信息系统安全实战评测，以及攻防演训人才培养的需求出发；从超大规模、多领域融合国家级网络靶场关键技术突破、系统建设、生态建设，及其应用推广的角度全面介绍鹏城网络靶场；重点介绍基于该网络空间安全领域的重大基础设施，提升网络安全保障能力和技术的典型案例和效果。



苏 洲

教授、博导，教育部长江学者特聘教授，西安交通大学网络空间安全学院院长，国家重点研发计划首席科学家，爱思唯尔中国高被引学者，全球前0.05%顶尖科学家。主要研究方向包括通信网络安全、物联网安全、信息物理融合系统安全等。在网络通信安全与优化研究领域IEEE TIFS、IEEE TDSC、IEEE JSAC、IEEE/ACM ToN等国际权威期刊文章发表了一系列高水平文章。获得IEEE ICC2020、IEEE BigdataSE2019等国际会议最佳论文奖。任 IEEE Internet of Things Journal、IEEE Open Journal of Computer Society等国际知名期刊编委。

报告题目：面向智能体互联网的安全防御

报告摘要：智能体可针对特定场景并依据既定目标自行分解任务，正从孤立的专用任务系统，演变为具备自主感知、规划、决策与行动能力的交互式实体。在机器人、自动驾驶、智慧城市等场景中，海量异构智能体跨域实时协同的需求迅速增长，催生了智能体互联网（Internet of Agents, IoA），但随之暴露出身份伪造、隐私监听与幻觉级联等新型安全威胁。本报告系统分析了Agent互联网的典型安全威胁，包括Agent身份认证、可信协同、具身安全、隐私威胁等，并进一步探讨在Agent互联网中的安全防御关键技术 with 具体挑战。最后，针对IoA语义检索安全、Agent通信协议安全、具身智能安全等三类场景，提出具体安全防御机制，实现海量 Agent间安全可控的互联互通。



李雪莹

天融信科技集团党委书记、董事长兼CEO；任清华大学计算机系专业实践企业指导教师、山东大学特聘教授，中国网络空间安全学会常务理事及女科技工作者工作委员会主任等职务。长期致力于安全管理与安全防护体系构建、安全大数据分析、云安全、人工智能+安全等领域的技术创新和成果转化，先后承担十余项部委课题研究工作，获多项省部级科学技术进步奖和社会科技奖、中国产学研合作创新奖、北京市三八红旗奖章，申请发明专利十余项，在国家级学术刊物及核心期刊共发表学术论文三十余篇，参与制定国家、行业及团体标准十余项，并主编多部著作。

报告题目：智驭未来，韧性赋能：人工智能时代网络安全前沿技术攻关与产教融合新范式

报告摘要：本报告聚焦人工智能时代网络安全的技术变革与生态重构。深入探讨“AI+安全”的前沿趋势，分析大模型驱动下的威胁感知、自动化防御及内生安全等核心技术攻关路径。结合天融信在智能安全领域的创新实践，阐述企业如何携手一流高校，通过产学研深度协同，打破技术壁垒，共同培养兼具AI素养与安全能力的复合型人才，共筑数字经济时代的坚实防线。

■ 圆桌嘉宾



王美琴
山东大学



李小勇
北京邮电大学



苏 洲
西安交通大学



王 骞
武汉大学



唐 卓
湖南大学

卫星互联网安全论坛

承办单位：中国电子科技集团公司第三十六研究所、国防科技大学电子对抗学院

🕒 9月19日下午

📍 天鹅湖大酒店二楼2号会议室

论坛主席： 郭世泽（中国科学院院士）
 任传伦（中国电子科技集团公司第三十六研究所）
 俞赛赛（中国电子科技集团公司第十五研究所）

召 集 人： 任传伦（中国电子科技集团公司第三十六研究所）
 俞赛赛（中国电子科技集团公司第十五研究所）

时 间	内 容
14:30-14:35	主持开场
14:35-14:40	致 辞
14:40-15:05	全球空间博弈视域下卫星互联网安全趋势 安建平（北京理工大学）
15:05-15:30	卫星互联网结构脆弱性及影响 盛 敏（西安电子科技大学）
15:30-15:55	星地融合能力共享的挑战与关键技术 高 跃（复旦大学）
15:55-16:20	休 息
16:20-16:45	卫星互联网安全技术体系研究 尹立坚（中科卫网）
16:45-17:10	互联网星座安全风险挑战与问题思考 张春磊（中国电子科技集团公司第三十六研究所）

论坛主席



郭世泽

中国科学院院士，信息技术安全研究中心研究员。



任传伦

中国网络空间安全学会卫星互联网安全专委会主任委员，中国电子科技集团公司第三十六研究所研究员、所长。担任十三五国家科技部网络空间安全专项重点研发计划项目首席科学家，十二五、十三五国家高新工程重点型号项目总设计师，获省部级一等奖 2 项，二等奖 3 项。



俞赛赛

中国网络空间安全学会卫星互联网安全专委会秘书长，中国电子科技集团公司第十五研究所正高级工程师，国家级人才。

召集人



任传伦

中国网络空间安全学会卫星互联网安全专委会主任委员，中国电子科技集团公司第三十六研究所研究员、所长。担任十三五国家科技部网络空间安全专项重点研发计划项目首席科学家，十二五、十三五国家高新工程重点型号项目总设计师，获省部级一等奖 2 项，二等奖 3 项。



俞赛赛

中国网络空间安全学会卫星互联网安全专委会秘书长，中国电子科技集团公司第十五研究所正高级工程师，国家级人才。

报告嘉宾



安建平

北京理工大学网络空间安全学院院长，二级教授，博士生导师，“信息与通信工程”国家一级重点学科首席教授。长期致力于空天信息网络与安全、空间信号处理方向的研究，牵头获得国家技术发明二等奖1项、国家科技进步三等奖1项、部级科技进步7项，是“百千万人才工程”国家级人选。

报告题目：全球空间博弈视域下卫星互联网安全趋势

报告摘要：在全球空天战略博弈日趋激烈的背景下，卫星互联网已成为大国太空竞争与网络对抗的核心赛道。当前太空治理规则体系尚不健全，技术霸权与资源争夺加剧了领域安全失衡问题。立足全球空间博弈格局，系统梳理卫星互联网面临的新型安全威胁，剖析领域竞争核心矛盾，研判其安全发展整体趋势。



盛敏

现任空天地一体化综合业务网全国重点实验室主任，是国家6G总体专家组专家、IEEE Fellow，同时入选教育部长江学者、国家杰出青年科学基金获得者、中组部万人计划领军人才等国家级人才。获评教育部新世纪优秀人才、霍英东教育基金会青年教师奖、陕西省三秦学者、陕西省五一巾帼标兵、全国三八红旗手。

报告题目：卫星互联网结构脆弱性及影响

报告摘要：卫星互联网工作在开放空间且节点动态变化，系统脆弱性更加明显。本次报告定义并刻画卫星互联网的网络结构脆弱性，并分析了不同结构下脆弱性到来的问题及影响，给出发展建议。



高跃

复旦大学教授、空间互联网研究院院长、IEEE Fellow、中国通信学会会士，国家级和上海市领军人才计划入选者。主要研究卫星主要研究空天地一体化网络、卫星互联网，主持国家级和省部级项目20余项。研究成果获得中国通信学会科学技术一等奖、欧盟信息通信技术领域的地平线奖，JCIN、中国通信等最佳期刊论文奖。现任中国通信学会英国工作委员主任委员，曾任IEEE通信学会无线通信委员主席。

报告题目：星地融合能力共享的挑战与关键技术

报告摘要：6G将面向天空地海的演进作为主要场景之一，同时人工智能能力部署到数据生成源，卫星作为通信基站的延伸即将形成星地通信与太空算力基础设施。本报告将介绍在卫星资源受限的基础上，如何大量训练任务卸载至边缘设备，从而有效缓解星上通信与计算负担，并探索星地大小模型协同推理研究。

**尹立坚**

北京中科卫网技术有限公司高级副总裁、首席技术官。中国创造学会智能空天信息专业委员会副秘书长清华大学电子工程系硕士；高级工程师；曾获省部级科技进步二等奖、三等奖、省级创新成果二等奖；卫星通信和装备领域部级专家，对卫星通信、短波通信、互联网安全、大数据处理等领域有丰富经验。

报告题目：卫星互联网安全技术体系研究

报告摘要： 以新一代低轨卫星通信技术为基础的全球卫星互联网迅猛发展，国内外多个大型卫星互联网星座都在快速建设。卫星互联网的演进正重塑全球连接格局，其安全范式已远超传统的网络安全概念。系统性审视其安全，从"太空基础设施 + 全球网络"的全景维度，从系统到设备，从通信网到互联网，分多个层次深入分析卫星互联网的安全威胁，研究并提出相应的安全防护技术体系框架，推动卫星互联网安全技术的发展。

**张春磊**

集团高级专家，正高级工程师，浙江省科技情报学会理事，杭州电子科技大学兼职硕士导师，中国网络安全学会卫星互联网专委会委员，中国指控学会网络信息体系咨询专家，全国电子战大会青年委员。获省部级奖项多项

报告题目：互联网星座安全风险挑战与问题思考

报告摘要： 低轨互联网巨型星座正以前所未有的速度扩张，并衍生出不同业务系统。低轨空间正经历一场频率、轨道、市场和规则等全方位争夺。这类星座不仅颠覆传统卫星通信产业，更在弹性指挥、天基遥感、互联接入等领域展现出强大潜能。本报告在系统梳理低轨互联网星座发展趋势基础上，提出了初步思考建议。

网络空间安全基础研究论坛

承办单位：山东大学

🕒 9月20日下午

📍 天鹅湖大酒店三楼12号会议室

论坛主席： 王美琴（山东大学）

召 集 人： 郁 昱（上海交通大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	Polynomial Commitments for Galois Rings and Applications to SNARKs Over $\mathbb{Z}_{2^k}$ 邢朝平（上海交通大学）
14:40-15:10	量子计算与网络安全 朱晓波（中国科学技术大学）
15:10-15:40	安全多方计算：安全与效率的权衡 赖俊祚（暨南大学）
15:40-16:00	休 息
16:00-16:30	量子计算环境下的格密码学 白 石（上海交通大学）
16:30-17:00	SAT求解器：理论与实践 蔡少伟（中国科学院软件研究所）
17:00-17:30	密码芯片侧信道安全研究进展与思考 王伟嘉（山东大学）

论坛主席



王美琴

王美琴，教授、博士生导师，山东大学副校长。长期从事对称密码分析与设计技术研究，建立多维零相关等密码攻击新理论，成为国际通用攻击；构建基于数学求解工具的自动化分析模型，研制安全与性能一体的密码算法自动化测评平台，已在航天、能源、通信等关键领域应用。在国际密码五大顶会发表论文55篇，国际密码顶会FSE发表论文数量全球首位，担任ASIACRYPT 2025领域主席和FSE 2023大会主席，入选中国密码学会会士。国家级领军人才，曾获全国创新争先奖章、国家科技进步一等奖、党政机要密码科技进步一等奖（省部级）及山东省技术发明一等奖。

召集人



郁 昱

郁昱先后从复旦大学和南洋理工大学获得学士和博士学位，之后在比利时鲁汶大学从事博士后研究。2011年回国后曾分别在华东师范大学、清华大学和上海交通大学任教。主要从事密码学相关的研究，多项研究成果发表在三大密码会、四个专题密码会议(TCC, CHES, PKC, FSE)、CCS和S&P等网络空间安全的代表性会议上，十多次担任CRYPTO、EUROCRYPT、ASIACRYPT的程序委员会委员。获得了国家杰出青年基金、中国密码学会优秀青年奖、科学探索奖和中国青年科技奖等项目和荣誉，入选了《麻省理工科技评论》隐私计算科技创新人物。

报告嘉宾



邢朝平

邢朝平目前是上海交通大学讲席教授。1990年在中国科学技术大学获得博士学位后留校工作，1993年获德国洪堡奖学金在Essen大学工作，其后在奥地利科学院从事研究工作。1998年加入新加坡国立大学，历任助教、tenured副教授、教授。2007年始任职于新加坡南洋理工大学教授。2019年加盟上海交通大学。长期从事密码、编码、安全多方计算、代数数论等方面的研究。

报告题目: Polynomial Commitments for Galois Rings and Applications to SNARKs Over $\mathbb{Z}_{2^k}$

报告摘要: Succinct non-interactive arguments of knowledge (SNARKs) allow a weak verifier to delegate computation tasks to a powerful prover in a verifiable way. However, most SNARK constructions require the computation tasks to be represented as arithmetic circuits over finite fields, incurring a significant overhead when applied for delegations of modern computer programs, which typically are of $\mathbb{Z}_{2^{64}}$ or $\mathbb{Z}_{2^{32}}$ arithmetic. Towards building publicly verifiable SNARKs for rings $\mathbb{Z}_{2^k}$, we follow the well-established framework of polynomial interactive oracle proofs (IOP)-based SNARKs, and obtain the first publicly verifiable SNARKs for rings $\mathbb{Z}_{2^k}$ with a transparent setup and plausibly post-quantum security.



朱晓波

朱晓波，中国科学技术大学教授、“祖冲之号”量子计算总师、2021年度科学探索奖获得者。长期致力于实用化量子计算机研制，通过系统性探索在量子纠错、量子随机游走、多体纠缠和量子模拟方向取得一系列创新成果，所负责研制的《祖冲之号》超导量子计算原型机，在随机线路采样问题上实现了量子优越性，达成了量子计算第一阶段目标。

报告题目: 量子计算与网络安全

报告摘要: 量子计算在处理某些复杂问题时速度远超传统计算机，被视为下一代计算技术，受到广泛关注。超导方案因可扩展性好而成为热门方向，目前正从实验室走向实用化，量子计算在为诸多领域带来算力革命的同时，也对经典密码体系构成根本性威胁。当前广泛使用的RSA、ECC等加密算法，在量子计算机面前将不再安全。本报告将介绍超导量子计算的发展现状，并探讨如何提前布局、平稳过渡到量子安全时代。



赖俊祚

赖俊祚，暨南大学教授。研究领域为公钥密码学、人工智能与数据安全，研究成果发表在EUROCRYPT、ASIACRYPT、IEEE S&P、AAAI、IEEE TIFS、IEEE TDSC 等密码学和信息安全国际会议和期刊上。研究工作得到了国家自然科学基金青年科学基金项目(A、B类)、广东省自然科学基金卓越青年团队项目、广东省自然科学基金杰出青年基金项目等的支持。以主要完成人获中国密码学会密码创新奖一等奖、广东省自然科学奖一等奖、广东省科技进步奖一等奖、国家科技进步二等奖。担任中国密码学会理事、《Frontiers of Computer Science》期刊编委。

报告题目：安全多方计算：安全与效率的权衡

报告摘要：安全多方计算作为实现隐私计算的主要技术之一，其安全强度与运行效率始终是领域内的核心矛盾。安全多方计算主流安全范式：被动安全（半诚实安全）构造简洁、计算与通信开销低，但仅能抵御被动窃听攻击，安全保障较弱；主动安全（恶意安全）可抵抗恶意篡改与主动攻击，安全等级更高，但协议复杂度显著提升，运行效率受限。报告介绍了针对安全 - 效率权衡的相关工作，包括我们团队在这方面的研究工作。



白石

白石，上海交通大学教授，主要从事密码学与计算数论研究。曾任美国佛罗里达大西洋大学数学学院副教授、密码学与信息安全中心联合主任。研究成果发表于国际密码学会会议，曾获 ASIACRYPT '15 最佳论文奖、NSF CAREER Award 等奖励。参与设计的后量子签名算法Dilithium被NIST采纳，成为美国后量子签名标准ML-DSA。

报告题目：量子计算环境下的格密码学

报告摘要：随着量子计算技术的快速发展，后量子密码学已成为信息安全领域的重要研究方向。在美国国家标准与技术研究院（NIST）主导的后量子密码标准化进程中，基于格的密码体制凭借其安全性基础和实现效率，被公认为最具竞争力的抗量子密码设计方案之一。本报告将围绕基于格的密码学及其密码分析方法展开讨论，重点关注带误差学习（Learning With Errors, LWE）问题。报告将系统阐述LWE问题在现代密码协议中的核心支撑作用，并讨论当前主流的量子攻击手段及量子辅助攻击策略。


蔡少伟

蔡少伟，中国科学院软件研究所研究员，北京航空航天大学研究员，主要研究约束求解，获得领域顶级会议CAV、CP和SAT的最佳论文奖/杰出论文奖，多次获得国际SAT比赛和SMT比赛冠军，华为鸿蒙创新大赛一等奖，强网杯密码数学专项赛冠军，混合整数规划比赛MIPcc亚军。其求解器被集成到英伟达、微软、英特尔、华为、华大九天等公司的软件中，应用于EDA、处理器验证、操作系统验证、以及密码分析。担任SAT 2027会议程序委员会主席，曾在FMCAD、CP和SOCS等国际会议做特邀报告。

报告题目：SAT求解器：理论与实践

报告摘要：布尔可满足性问题（简称SAT）是理论计算机的关键问题，SAT求解器在自动密码分析有重要应用。在实践中，对于给定的特定问题，定制的SAT求解器往往比通用求解器高效得多。我们探索了SAT的一些常见算法和策略在某些类型公式上的复杂度分离现象，并介绍基于大模型定制的SAT求解器在ASCON密码差分分析上取得的进展，证明了ASCON算法的4轮差分分析活跃s盒数目为41。


王伟嘉

王伟嘉，现任山东大学网络空间安全学院教授，入选山东省泰山学者青年专家、小米青年学者，获评山东大学优秀青年教师。主要从事密码侧信道安全与密码工程研究，近年来以第一作者或通信作者身份在欧密会、亚密会、CHES、IEEE TIFS等国际重要会议和期刊发表论文40余篇。主持国家重点研发计划课题、国家自然科学基金面上项目、山东省自然科学基金联合重点项目等，获山东省技术发明一等奖、INSCRYPT最佳论文奖，并获评华为openHiTLS密码开源社区优秀贡献者。相关成果已在华为、中国航天、蚂蚁集团、中国电信、奔图打印机等得到应用。担任CHES、USENIX Security等国际重要会议程序委员会委员，并任中国密码学会青年工作委员会委员、《密码学报》青年编委。

报告题目：密码芯片侧信道安全研究进展与思考

报告摘要：侧信道攻击能够绕过密码算法本身的数学安全机制，通过分析密码芯片运行过程中难以完全避免的功耗、电磁辐射等物理信息泄漏，实施密钥恢复攻击，已成为密码芯片面临的重要安全威胁。侧信道安全也已成为高端密码芯片和密码模块通过安全认证、满足行业合规要求以及进入关键应用领域的重要门槛。本报告重点介绍我们在抗量子密码算法新型侧信道分析技术和可证明安全防护方法方面的最新研究进展，并分享对人工智能时代密码侧信道安全研究面临的新机遇、新挑战及未来发展方向的一些思考。

生成式AI驱动的内容安全治理论坛

承办单位：哈尔滨工业大学网络空间安全学院

⌚ 9月20日下午

📍 天鹅湖大酒店二楼2号会议室

论坛主席：张宏莉（哈尔滨工业大学）

召集人：李 兵（中国科学院自动化研究所）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	认知安全中策略智能生成技术及应用 田 玲（电子科技大学）
14:40-15:10	高通量视频内容理解与安全 李 兵（中国科学院自动化研究所）
15:10-15:40	社会智能驱动的认知安全 张勇东（中国科技大学）
15:40-16:00	休 息
16:00-16:30	基于大模型的虚假信息检测 朱国普（哈尔滨工业大学）
16:30-17:00	生成式文本检测溯源技术及应用 曹亚男（中国科学院信息工程研究所）
17:00-17:30	AIGC下Deepfake技术在新型科技犯罪场景的应用和治理应对 段继平（奇安信集团）

论坛主席



张宏莉

哈尔滨工业大学二级教授、博士生导师，网络空间安全学院创始院长，入选国家高层次人才计划、国家网络安全优秀教师；兼任国务院学位办网络空间安全学科评议组秘书长，教育部网络空间安全教学指导委员会委员，中国网络安全学会常务理事，中国通信学会网络与数据安全专委会副主任；主要研究领域为信息内容安全、认知安全、网络与系统安全等。获省部级科研奖励一二等奖5项、省教学成果奖一二等奖3项；出版专著和教材5部。

召集人



李 兵

中科院自动化所多模态人工智能系统全国重点实验室研究员，博士生导师，国家优青，北京市杰青，中科院“青年学习标兵”，国家科技创新 2030 重大项目、国家自然科学基金重点项目负责人，人民中科研究院院长，多模态超级智能安全北京市重点实验室副主任。长期从事多模态人工智能与安全方向的研究，在国际权威期刊和会议上发表论文 100 余篇，担任多个国际学术期刊。

报告嘉宾



田 玲

电子科技大学教授，博士生导师，下一代互联网数据处理技术国家地方联合工程实验室主任，教育部重大人才工程特聘教授，国防领域国家级领军人才。围绕“数据+知识”驱动的人工智能，在计算机视觉、推理模式、预测理论方面做出原创性与突破性工作，获国家技术发明二等奖、JS技术发明一等奖、教育部科技进步一等奖、四川省杰出青年科学技术创新奖等科技成果奖励9项。主持国家级科研项目10余项，发表高水平学术论文86篇，国家发明专利授权77项。

报告题目：认知安全中策略智能生成技术及应用

报告摘要：认知影响活动正随着大语言模型、AIGC、智能体等技术的发展不断演化，呈现出更强的智能化、链条化和场景化特征，并深刻改变认知内容生产、传播扩散和影响干预的方式。面对新技术条件下认知风险加速演变的现实需求，提升认知安全态势感知、风险研判与策略生成能力，已成为认知安全治理的重要方向。报告以“认知安全中策略智能生成技术及应用”为主题，围绕知识超图、知识推理和知识决策等内容，系统介绍认知安全策略智能生成的关键技术、实现路径及典型应用，探讨其在态势分析、策略辅助制定和响应处置中的实践价值。


李 兵

中科院自动化所多模态人工智能系统全国重点实验室研究员，博士生导师，国家优青，北京市杰青，中科院“青年学习标兵”，国家科技创新2030重大项目、国家自然科学基金重点项目负责人，人民中科研究院院长，多模态超级智能安全北京市重点实验室副主任。长期从事多模态人工智能与安全方向的研究，在国际权威期刊和会议上发表论文100余篇，担任多个国际学术期刊编委，曾获国家自然科学基金二等奖（1/3）、中国自动化学会技术发明一等奖（1/6）等。

报告题目：高通量视频内容理解与安全

报告摘要：随着人工智能与视频社交平台的迅猛发展，视频在为公众生活带来极大便利的同时，也为违规、伪造、虚假等不良内容的传播提供了可乘之机，对网络空间内容安全构成严峻挑战。本团队面向网络空间内容安全的重大战略需求，针对互联网大流量视频流数据处理难题，聚焦高通量、低功耗、实用化视频内容理解与安全的理论与技术研究。本次报告将分享团队在该方向的研究成果，具体包括：（1）针对海量视频解码成本高昂的关键问题，提出面向压缩域的小样本特定视频识别方法，有效降低计算开销；（2）针对生成式伪造视频引发的新型安全风险，提出基于一致性分析的伪造图像与视频识别方法，提升虚假内容甄别精度；（3）针对有限算力场景下算法落地部署的核心瓶颈，提出蒸馏-裁剪一体化深度模型压缩与加速方法，保障算法在资源受限设备上的高效运行。最后，将进一步介绍相关技术的实际落地系统及实战应用案例。


张勇东

教授，博士生导师，中国科学技术大学信息科学技术学院执行院长，IEEE Fellow。国家自然科学基金委创新研究群体项目负责人，国家杰出青年科学基金获得者，“万人计划”科技创新领军人才。主要研究方向为网络空间多媒体内容安全，曾获国家自然科学基金二等奖1项（排名1），国家技术发明奖二等奖1项（排名1）和国家科技进步奖二等奖1项。所发表论文谷歌引用3万余次，获授权国家发明专利100余项。担任《中国通信》副主编，国家重点研发计划-“变革性技术关键科学问题”、“社会治理与智慧社会科技支撑”重点专项总体专家组成员。

报告题目：社会智能驱动的认知安全

报告摘要：认知安全是保护人类思维与决策免受数字操纵的全新领域，随着人工智能迅猛发展，认知安全已成为国家安全体系中的新兴关键维度，是当前大国战略博弈的焦点之一。本报告面向国家安全重大需求，探索利用多智能体模拟社会智能的范式，解决认知安全致效机理和治理机制这一重大科学挑战，围绕“复杂社会认知建模机理、认知操纵风险识别机制和认知操纵的干预引导模式”三大科学问题开展深入讨论，以求形成可计算、可推演、可干预的认知安全技术体系，为维护国家认知安全提供理论与方法支撑。

**朱国普**

哈尔滨工业大学计算学部长聘教授、博士生导师，网络大数据安全分析工信部重点实验室副主任，黑龙江省网络安全和信息化协会人工智能与内容安全专委会主任。主要研究方向为人工智能安全、多媒体取证与安全、AI大模型应用。在国际权威期刊上发表学术论文80余篇，其中IEEE/ACM Transactions论文40余篇；单篇最高Google Scholar引用1700余次。现任IEEE TCYB、IEEE TII、IEEE TETCI、IEEE TBC、IEEE TNSM、IEEE TSIPN、IEEE JSTARS、IEEE JSYST、ACM TOPML、ACM TOMM、PATTERN RECOGN等十余个国际权威期刊编委。

报告题目：基于大模型的虚假信息检测

报告摘要：生成式AI的快速发展，使虚假信息具备了高度逼真、快速演化与批量生成的能力，严重侵蚀社会信任，加剧认知操纵风险，对数字社会治理构成严峻挑战。当前，检测技术在跨模态泛化能力和现实场景通用性方面仍存在明显不足，难以满足高效识别与溯源的实际需求。本报告介绍大模型时代虚假信息检测的关键技术，涵盖AIGC内容检测、多模态深度伪造检测、篡改区域定位及虚假新闻检测等方向，并在此基础上，总结当前面临的主要挑战，展望未来发展趋势。

**曹亚男**

中国科学院信息工程研究所研究员、信息内容分析研究室副主任。国家级青年人才、国家重点研发计划青年科学家项目首席。研究方向为信息内容安全、大模型安全等。近五年，发表TPAMI、TIFS、USENIX Security、NeurIPS、ICML、ACL等CCF-A类论文60余篇，主持国家重点研发计划青年科学家项目、国家自然科学基金联合基金课题等20余项科研项目，研发系统在网信、公安等多个国家安全部门发挥实战作用，获评中国电子学会科技进步二等奖、中国科学院大学领雁银奖、中国科学院朱李月华优秀教师奖等。

报告题目：自主且有界：构建安全可信的智能体系统

报告摘要：以大模型为核心的智能体正从“理解与生成”加速迈向“规划与行动”。通过任务拆解、工具调用、长期记忆和多智能体协作，智能体逐步具备在开放环境中持续决策、动态交互和执行复杂任务的能力。然而，智能体自主能力的跃升也使安全风险由模型层面的局部问题演变为贯穿任务全过程的系统性挑战。如何在增强智能体自主能力的同时建立安全边界，推动安全保障从局部防护走向系统治理，已成为可信人工智能发展亟待解决的关键问题。报告将从智能体全生命周期安全、幻觉检测、工具调用、多智能体安全协作等方向介绍课题组近期的研究成果，推动构建贯通模型、系统与应用的智能体安全保障体系。

**段继平**

段继平，奇安信集团副总裁、奇安盘古CEO，网络安全正高级工程师、高级司法鉴定人，中国图像图形学会理事，上海市公安局专家委员会委员。入选“北京市百名高端涉外法治人才”项目，为北京司法鉴定业协会司法鉴定专家库成员。参与主持国家重点研发计划《互联网新业态违法违规广告智能监测和预警关键技术研究》课题。曾荣获“北京市优秀青年工程师”等荣誉称号，2022年参与编著司法部权威教材《电子数据司法鉴定实务》。

报告题目：AIGC下的Deepfake与新型科技犯罪

报告摘要：随着AIGC发展，Deepfake科技犯罪正呈现低门槛、批量化、多模态和实时交互等趋势，身份仿冒与诈骗活动更加精准，跨平台传播和电子取证难度持续上升。本报告将分析相关犯罪的新变化，并围绕模型安全、内容标识与溯源、平台检测处置、身份核验、证据固定和部门协作等方面探讨治理对策。

网络空间态势感知论坛

承办单位：国防科技大学电子对抗学院网络系、东南大学网络空间安全学院

🕒 9月20日下午

📍 天鹅湖大酒店一楼淮河厅A厅

论坛主席：陆余良（国防科技大学）

程 光（东南大学）

召 集 人：杨国正（国防科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	从探测到认知：网络空间测绘的范式演进与LLM驱动的新机遇 杨家海（清华大学）
14:40-15:10	加密流量分析与网络公害识别技术 吴 桦（东南大学）
15:10-15:40	协同测绘赋能网络空间治理 刘庆云（中国科学院信息工程研究所）
15:40-15:50	休 息
15:50-16:20	破坏利用链：面向AI自动化攻击的智能防御探索 权小文（远江盛邦安全科技集团股份有限公司）
16:20-16:50	高质量APT情报数据驱动未知威胁发现 李书豪（中关村实验室）
16:50-17:20	从体系测绘到安全治理：互联网域名系统安全研究 许成喜（国防科技大学）

论坛主席



陆余良

国防科技大学电子对抗学院教授，网络空间安全学科带头人，网络空间安全领域的知名专家，“十二五”期间被聘为863主题专家组成员，教育部高等教育网络空间安全专业教学指导委员会委员、国家重点研发计划项目首席专家。享受国务院政府特殊津贴。长期从事网络空间安全领域的人才培养与科学研究工作，在网络安全态势感知等方面做出了突出贡献。主持完成了国家重点研发计划项目、军队重点型号工程等重点项目20余项，获得军队（省部级）科技进步一等奖3项、二等奖8项、三等奖10余项，中国指挥与控制学会科学技术一等奖1项，授权专利20余项。编写出版教材、专著6部，发表高水平学术论文百余篇。



程 光

东南大学首席教授、二级教授、网络空间安全学院执行院长，教育部网络空间安全教指委委员、中国网络空间安全学会常务理事、网络空间国际治理研究基地主任、江苏省泛在网络安全工程研究中心主任、江苏省网络空间安全学会理事长、中国计算机学会互联网专委会副主任、科技部网络空间治理和安全领域“十四五”规划专家、网络空间安全2030重大专项专家等学术兼职。研究方向网络流量安全分析、主动防御、内生安全等，主持国家重点研发、国家自然科学基金、中央网信办专项等科研项目40余项。获网络安全优秀教师、国家教学成果一等奖，牵头获2024年江苏省十大科学进展、2021年江苏省科技一等奖、2025年教育部科技二等奖等。发表学术论文100余篇，出版专著9部，培养研究生200余人。

召集人



杨国正

国防科技大学电子对抗学院网络系副主任、教授，安徽省网络空间态势感知于评估重点实验室副主任，高校计算机优秀教师，安徽省线上教学名师，学校“领军人才”培养对象。长期从事网络空间安全领域研究，主持和参与国家重点研发计划项目课题10余项，获军队（省部级）科技进步一等奖1项、二等奖3项，三等奖5项，省级教学成果特等奖1项、一等奖2项、三等奖2项。出版教材4部，授权专利15项，发表高水平学术论文50余篇。

报告嘉宾



杨家海

清华大学长聘教授（专技二级）、博士生导师，网络研究院学术委员会主任，国务院政府特殊津贴获得者。先后承担国家973/863重点项目、重大专项、重点研究计划、自然科学基金等国家纵向项目30余项，发表高水平的学术论文200余篇、获得授权发明专利60多项；出版学术专著4部。获得省部级科技进步一等奖7项，二等奖3项。是国家十三五、十四五重点研究计划项目首席专家；是中国计算机学会CCF及IEEE协会高级会员、CCF分布式计算与系统专委会委员、北京市计算机学会理事、中国指挥与控制学会网络空间测绘专委会副主任委员；是两个重要的SCI检索源期刊：John Wiley & Sons的 International Journal of Network Management (IJNM) 和IEEE Transactions on Network and Service Management (TNSM) 的编委。

报告题目：从探测到认知：网络空间测绘的范式演进与LLM驱动的新机遇

报告摘要：以“从探测到认知：网络空间测绘的范式演进与LLM驱动的新机遇”为题，立足网络空间作为“第五疆域”的战略定位，系统梳理了测绘在基础探测、资产识别、可视化表达等方面的关键技术进展，同时直面数据规模、隐蔽资产与反测绘对抗等现实挑战，重点探讨大语言模型（LLM）驱动测绘认知能力跃升的新范式，结合智能测绘智能体、语义知识图谱构建及自然语言交互等典型应用模式，并展示了研究团队在活跃地址探测、端口与服务识别等场景中的实践成果。



吴 桦

东南大学网络空间安全学院副教授、博士生导师，江苏省泛在网络安全工程研究中心副主任。长期从事加密流量分析、网络态势感知、网络公害识别等研究。主持多项国家及省部级项目，项目成果在相关业务单位落地应用，取得显著成效。2021年获得江苏省科技进步一等奖，2023年获得江苏省行业领域十大科技进展，2025年获得教育部科学研究优秀成果奖。

报告题目：加密流量分析与网络公害识别技术

报告摘要：系统梳理加密流量分析的基本原理、主流技术方法与核心研究目标。在此基础上，以公害视频识别作为典型研究场景，依托协议逆向分析技术，结合HTTP/1.1、HTTP/2、HTTP/3各版本协议的差异化特征，逐层还原不同协议下加密视频流的应用层视频长度特征，以此实现对公害视频的精准识别。最后总结当前加密流量分析领域研究面临的难点问题，并对未来研究发展前景与应用方向进行展望。



刘庆云

中国科学院信息工程研究所正研级高工，博士生导师，中国科学院大学网络空间安全学院教授。主要研究领域为网络流量处理架构，网络空间测绘。近年来作为主要负责人，先后承担国家重点研发项目、工信部、发改委等多项国家级重大信息安全科研项目，主持研发的多个系统在国家重要部门实际应用效果显著。获国家科技进步二等奖1次，省部级科技进步奖2次。近三年在USENIX Security等国内外重要会议期刊发表论文50余篇，开设研究生课程《网络空间测绘与安全应用》等课程，培养相关方向研究生30余人。

报告题目：协同测绘赋能网络空间治理

报告摘要：随着网络技术的迅猛发展，网络空间治理受到极大挑战，网络空间测绘技术用于解决网络空间的安全问题，在资产定位、风险管理方面应用发挥了巨大作用，但从网络空间治理角度，网络服务的可信度，所有者的测绘评估还存在诸多不足，尤其随着以大模型为代表的新兴服务的快速推广和应用，如何让服务更安全可信仍是迫切需要解决的问题。本报告将会从协同测绘体系构建的角度探讨DNS，暗网，大模型等服务的测绘相关工作，利用测绘相关技术，赋能网络空间治理。



权小文

正高级工程师，国家级领军人才，国家网信优秀人才，中国指控学会网络空间测绘专委会副主任，中国网络安全产业联盟常务理事，工信部产教融合实践教授。入选首届IDC中国CSO名人堂（十大人物）、入选“中关村高端创业领军人才”及“北京市海英人才”；获得技术发明一等奖一项，科技进步奖等省部级奖项6项；牵头/参与9项国标制定，负责国家重点研发课题6项、北京市科委重大项目2项，参与工信部、公安部等国家重大工程13项。远江盛邦安全科技集团股份有限公司（688651）的创始人。

报告题目：破坏利用链：面向AI自动化攻击的智能防御探索

报告摘要：“攻击者视角”正在被自动化。资产测绘、脆弱性推演、利用链构造——这些原本依赖资深红队的工作，正被自动化渗透智能体以小时级、甚至分钟级的速度批量完成；而代码审计、漏洞发现、补丁生成这类“知识与推理”型工作，则被吸收进通用大模型，重划了整个安全行业的价值边界——模型吞掉了“分析师的大脑”。本报告立足这一变局：机器对抗机器的时代已经到来，防御的胜负不再取决于谁的模型更聪明，而取决于谁守住了模型够不着的那块专有阵地。报告聚焦“破坏利用链”这一对症思路——不追求补全每一个漏洞，而是通过打断攻击链条的下游环节，让漏洞在被利用前失效，从而在攻防的速度之争中重建防御一方的不对称优势。

**李书豪**

中关村实验室科研02部副部长、APT团队负责人、正高级工程师，某国家级创新人才，科技部某领域咨询专家，兼任国科大、哈工大等高校教授/博士生导师、北京市直机关青联委员、海淀青联科技界别副主任委员。研究方向是智能化威胁防御。近年来，承担多个国家级项目，主持研制中关村实验室APT防御平台，在数十家重点单位实战应用，成效显著，发表顶会顶刊论文10余篇，出版专著1部，获省部级科技进步一等奖2项。

报告题目：高质量APT情报数据驱动未知威胁发现

报告摘要：APT攻击具有高变化、高隐蔽特性，传统APT威胁情报生成滞后、数据质量参差不齐、情报数据共享壁垒突出，难以满足未知威胁发现需求。本报告聚焦智能化网络威胁感知方向，阐述大模型、AI智能体在APT威胁情报生成、潜在攻击者动态画像、APT攻防数据生成、未知APT攻击线索发现等方面的技术路径，为APT攻防对抗、网络安全态势感知提供技术思路与实践参考。

**许成喜**

国防科技大学副教授，研究方向为网络空间测绘、域名系统安全研究，在网络空间高速探测、网络空间实体服务范围、域名声誉分析、域名授权机制安全分析等方面取得多项原创性研究成果，发表Usenix Security、NDSS、IMC、TDSC、TIFS等网络安全顶级学术会议/期刊论文10余篇，授权国家发明专利20余项，获省部级科技进步二等奖2项，中国指挥控制学会科技进步一等奖1项。

报告题目：从体系测绘到安全治理：互联网域名系统体系安全研究

报告摘要：域名系统（Domain Name System）提供的服务已经渗透到互联网的方方面面，是现代互联网赖以运转的神经中枢。报告围绕互联网域名系统体系安全研究，结合网络安全事件，系统介绍了域名系统安全的研究现状，重点从大规模网络空间测绘的视角讨论了域名系统在权威侧、解析器侧、注册管理侧等维度存在的安全风险及治理措施。

网络空间测绘技术论坛

承办单位：信息工程大学

🕒 9月20日下午

📍 天鹅湖大酒店三楼7号会议室

论坛主席： 罗向阳（信息工程大学）

召 集 人： 刘媛妮（重庆邮电大学）

时 间	内 容
13:30-13:35	主持开场
13:35-13:40	致 辞
13:40-14:10	BGPWatch：互联网域间路由协同测绘平台 王继龙（清华大学网络研究院）
14:10-14:40	新型网络服务测量与行为分析 熊 刚（中国科学院信息工程研究所）
14:40-15:10	全量高频网络可视化技术 谢 鲲（湖南大学）
15:10-15:30	休 息
15:30-16:00	网络空间拓扑测绘与抗测绘 张 宇（哈尔滨工业大学）
16:00-16:30	物联网设备安全：暴露面分析的技术演进 喻 波（国防科技大学）

论坛主席



罗向阳

网络空间安全专家，现任信息工程大学教授、博导，河南省网络空间态势感知重点实验室主任，国家万人计划科技创新领军人才、中原学者、国防科技卓越青年基金获得者、军队学科拔尖人才，国家重点研发计划项目负责人，网络空间测绘领域的开拓者和学术带头人之一。

召集人



刘媛妮

重庆邮电大学教授，博士生导师，中国网络空间安全学会网络空间测绘专委会（筹）秘书长，中国通信学会车联网专委会委员、中国通信学会内生安全专委会委员，美国密苏里大学圣路易斯分校访问学者。主要从事智能物联网、车联网安全，分布式身份认证等方面的研究。主持国家重点研发计划项目子课题、国家自然科学基金项目、军队和四川省重点项目等20余项。发表SCI/EI检索论文20余篇，出版专著1部，参与制定行业标准2项。

报告嘉宾



王继龙

清华大学网络研究院教授，2016年开始网络空间测绘技术探索性研究，在国家重点研发计划项目和国际合作项目支持下，提出网络空间坐标系和网络空间地图模型，牵头制定本领域首批技术标准，领导27个国家和地区的38个跨学科机构，合作建成覆盖亚、欧、美、澳、非5个大洲的网络空间协同测绘平台，相关成果入选2025年信息通信领域十大科技进展。

报告题目：BGPWatch：互联网域间路由协同测绘平台

报告摘要：全球BGP路由劫持与信息泄漏事件频发，造成巨大危害，却难以监测预防，成为互联网的难点问题，发展中国家面临更大挑战。为促进协同治理与公平共治，清华大学联合全球科研机构共建BGPWatch全球路由安全监测与治理平台，致力于提升全球路由安全检测与治理能力，助力构建包容、安全的互联网治理体系。BGPWatch平台已联合来自27个国家和地区的38家教育科研网络与科研机构，共建覆盖上百个国家的全球探针监测网络，实现对全球BGP路由异常事件的协同感知、分析与处置，积极推动开放合作与能力共享，为构建更加包容、安全、稳定的全球互联网治理体系提供重要支撑。


熊 刚

中国科学院信息工程研究所二级研究员，研究室主任，所学术委员会委员，中国科学院大学网络空间安全学院教授，博士生导师。主要研究领域为网络测量与行为分析、数据挖掘与人工智能对抗。近年来，主持国家重点研发计划3项、牵头研发多个重大信息安全平台和系统，在线成功应用。获国家科技进步二等奖1项，省部级科技进步一等奖、技术发明一等奖3项。牵头主讲《网络行为学导论》课程，发表学术论文110余篇，培养博士、硕士研究生40余人。

报告题目：新型网络服务测量与行为分析

报告摘要：游戏加速、网络广告、应用加速、视频直播、流量分发等各类新型网络服务覆盖全球且持续演进，本报告通过对上述网络服务开展主被动测绘和流量解析，辅以域名、路由、数字证书等互联网基础服务的长期测量数据，结合密态对抗、加密流量分类、网络行为分析等理论与关键技术研究，开展深入的行为分析，揭示网络中面临的不对称、强对抗、智能化的安全威胁和数据泄露风险，分析原因，探索应对之道。


谢 鲲

湖南大学二级教授，博士生导师，国家杰出青年基金获得者，中国青年女科学家奖获得者，湖南省杰出青年基金获得者，湖南省青年骨干教师，湖南省优秀硕士生导师，长沙市“巾帼建功”标兵，湖南大学科研标兵。担任超算与人工智能融合计算教育部重点实验室主任，以及湖南大学信息科学与工程学院学术委员会/学位委员会主任。长期从事计算机网络、网络安全、网络测量、与人工智能的研究工作。在国际主流会议和期刊SIGMOD, CCS, INFOCOM, ICDE, SIGMETRICS, IEEE/ACM TON, IEEE TMC, IEEE TIFS, IEEE TDSC, IEEE TPDS, IEEE TSC, IEEE TC等发表论文100余篇。

报告题目：全量高频网络可视化技术

报告摘要：网络作为国家新型基础设施，其稳定与安全对经济社会运行至关重要。随着网络规模持续扩张、设备日益异构、业务形态不断多样化，网络运维面临故障更频发、管理更复杂等突出挑战。现有测量与监测手段多停留在粗粒度、局部感知层面，难以实现对网络状态的精准、实时掌控；异常检测与根因定位能力不足，在多维因素耦合的场景下，“观测空洞”问题尤为显著。围绕上述痛点，本报告将汇报面向“宏观全覆盖、微观精细化”的全量高频网络可视化测量理论与关键技术。

**张 宇**

哈尔滨工业大学计算学部网络空间安全学院教授，鹏城实验室网络部体系所所长，承担国家重点研发计划项目、国家实验室重大攻关任务项目，获得中央网信办某创新人才支持计划，2024互联网基础技术贡献奖，研究方向包括互联网体系结构，互联网基础设施安全，网络空间测绘等。

报告题目：网络空间拓扑测绘与抗测绘

报告摘要：网络空间拓扑测绘是绘制全球网络基础设施“数字地图”的核心技术，通过路由追踪、BGP分析与地理定位揭示节点连接关系与层级结构；抗测绘技术则通过拓扑混淆、路由欺骗与探测反制构建“隐形屏障”。本演讲从攻防双维度切入，介绍拓扑发现的核心算法与抗测绘的前沿策略，为网络空间安全治理与韧性防护提供支撑。

**喻 波**

国防科技大学计算机学院研究员、博士生导师，长期从事物联网安全、智能漏洞挖掘方向教学科研工作，入选教育部青年长江学者，获军队二级表彰、计算机学会杰出会员等荣誉。主持国家自然科学基金、国家重点研发计划、军队预研等项目8项。主持出版专著1部，在ACM CCS、NDSS、TIFS、软件学报、通信学报等发表论文10余篇。科研成果入选国防科技大学十大科技进展，获海南省科技进步一等奖、军队和湖南省科技科二等奖等奖项。

报告题目：物联网设备暴露面分析关键技术

报告摘要：报告系统阐述物联网设备暴露面信号层、通信层、服务层三层核心架构，梳理其资源受限、接口开放、协议多样、暴露碎片化的典型特性，这类特性导致物联网设备极易成为网络安全威胁的入口。进一步总结暴露面分析技术进展：信号层依托无线信号探测、指纹识别技术实现设备隐匿节点发现；通信层基于协议逆向、组件特征分析技术挖掘通信协议漏洞与传输风险；服务层通过端点扫描、污点分析技术挖掘设备后端服务和接口安全隐患。针对当前设备暴露面检测范围窄、复杂场景适配性差等问题，指出未来将朝着多维度融合感知、复杂场景自适应检测方向发展，结合人工智能与软件安全技术构建动态暴露面感知体系，为物联网设备风险管控及网络安全态势感知提供技术支撑。

智能体安全论坛

承办单位：合肥综合性国家科学中心数据空间研究院

⌚ 9月20日下午

📍 天鹅湖大酒店三楼1号会议室

论坛主席：张勇东（中国科学技术大学）

召 集 人：何向南（中国科学技术大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	极端环境下智能体的可用、可靠、可信 沈 超（西安交通大学）
14:40-15:10	人工智能时代下数据要素可信流通技术探索 盖珂珂（北京理工大学）
15:10-15:40	前沿AI系统安全风险评测与治理 潘旭东（复旦大学）
15:40-16:00	休 息
16:00-16:30	Agent系统的风险归因与防护 刘东瑞（上海人工智能实验室）
16:30-17:00	从对话到行动：AI智能体的安全风险演化与治理体系构建 武建双（中检集团天帷信息技术股份有限公司）
17:00-17:30	寻找智能体安全的不变量：开放环境下的GUI智能体行为对齐 张倬胜（上海交通大学）

论坛主席



张勇东

张勇东，教授，博士生导师，中国科学技术大学信息科学技术学院执行院长，IEEE Fellow。国家自然科学基金委创新研究群体项目负责人，国家杰出青年科学基金获得者，“万人计划”科技创新领军人才。主要研究方向为网络空间多媒体内容安全，曾获国家自然科学基金二等奖1项(排名1)，国家技术发明奖二等奖1项(排名1)和国家科技进步奖二等奖1项。所发表论文谷歌引用3万余次，连续多年入选科睿唯安全球高被引学者。获授权国家发明专利100余项，相关技术在著名国际评测中多次取得优异成绩。研究成果有效应用于国家网络空间内容安全领域。

召集人



何向南

何向南，中国科学技术大学人工智能学院教授、博导、副院长，国家杰青，国家科技创新2030“新一代人工智能”重大项目首席科学家。研究方向为信息推荐与挖掘、大模型与通用人工智能等，在相关领域顶会（如SIGIR、WWW、KDD）和顶刊（如IEEE TKDE、ACM TOIS）上发表论文100余篇，谷歌学术引用7万余次，Elsevier中国高被引学者。曾获信息检索顶会ACM SIGIR最佳论文奖、机器学习会议ICLR最佳论文奖、国际基础科学大会前沿科学奖、阿里巴巴达摩院青橙奖、中国人工智能学会吴文俊人工智能自然科学一等奖（第一完成人）等。

报告嘉宾



沈超

沈超, IEEE Fellow, IARR Fellow, 西安交通大学人才办、电子与信息学部二级教授, 长江学者特聘教授, 国家自然科学基金创新研究群体负责人、教育部创新团队负责人、教育部智能网络与网络安全重点实验室主任, 科学探索奖、达摩院青橙奖、国家优秀青年基金获得者, 国家重点研发计划首席科学家, 国家基础加强计划技术首席科学家, 重点研发计划“先进计算与新兴软件”重点专项指南专家组成员。主要从事智能系统可信、控制与安全的教学和研究工作, 撰写专著4部, 获最佳论文奖14次, 牵头获3项省部级/一级学会科学技术一等奖、日内瓦国际发明金奖、霍英东青年教师一等奖、MIT TR35 China、IEEE SMC Early Career Award、IEEE Outstanding SMC Chapter Award等。制定国内外标准6项, 5份建言被中央办公厅、中宣部等采纳, 担任IEEE SMC学会Associate-Vice President、IEEE Trustworthy and Controllable Intelligent Systems技术委员会主席、ACM SIGSAC China副主席、中国人工智能学会组织工委副主任、中国网络空间安全学会青年工委副主任、中国指控学会智能测评专委会副主任等, 以及IEEE TDSC、IEEE TCYB等多个汇刊和中国科学: 信息科学、电子学报、软件学报等副主编或编委。

报告题目: 极端环境下智能体的可用、可靠、可信

报告摘要: 在“三深一极”等极端环境下, 智能体面临能量受限、强辐射、极端温差以及环境高度未知等严峻挑战, 使其在可用性、可靠性与可信性方面存在明显不足, 难以支撑长期自主任务。为此, 亟需构建具备端侧自主决策、群体协同与持续进化能力的新型智能体。本报告以极端环境自主作业为核心应用场景, 围绕三方面展开: 一是模型推理的算力极限压缩与协同分配, 二是模型在端侧的可靠部署与稳定运行, 三是面向任务演化的实时可信学习与训练机制。在此基础上, 系统梳理端侧智能体在极端环境下的可用可靠可信等关键技术, 介绍团队在相关方向的研究探索与工程实践, 并对未来极端环境下智能体的演化研究方向进行展望。



盖珂珂

盖珂珂，北京理工大学人工智能学院副院长，北京中关村学院副院长，教授，博士生导师。2022年入选国家级青年人才计划，同期入选爱思唯尔“中国高被引学者年度榜单”，2021年入选全球前2%顶尖科学家榜单。长期从事网络空间安全领域的科研工作，近年来聚焦大模型安全、隐私计算等领域。主持科技部重点研发项目课题、国家自然科学基金重点项目等10余项科研项目；发表SCI/EI论文200余篇，含ESI高被引论文10篇，获国际会议/期刊最佳论文奖10余项，谷歌学术引用14000余次，H-index指数60；申请专利70余项；合著教材5部。担任IEEE技术与工程管理学会区块链专委会联合主席，中国计算机学会区块链专委会常务委员，新工科联盟区块链工委秘书长，担任期刊Blockchains主编，担任多个国际会议程序委员会主席，担任多个SCI学术期刊编委，IEEE高级会员（IEEE Senior Member）；担任中国国际电视台（CGTN）特邀专家。

报告题目：人工智能时代下数据要素可信流通技术探索报告

报告摘要：随着数字经济的快速发展，数据已成为关键生产要素，但数据跨主体、跨领域流通面临安全、隐私、合规等核心挑战。本报告将面向关键基础设施需求，围绕可信数据流通关键领域，系统梳理可信数据流通的技术演进路径，为构建安全高效的数据要素市场提供参考。同时，结合数据要素发展需求，讨论在激活数据价值方面区块链技术的应用的场景、趋势和应用实现。



潘旭东

潘旭东，复旦大学副研究员（学敏学者）、上海创智学院全时导师，研究方向为大模型安全与治理，在国际上较早发现多种商用大模型数据隐私（2020）、后门（2022）和自主复制风险（2024）等，多次得到OpenAI、Google、Anthropic等正面评价，作为全球32位专家之一，受邀与Yoshua Bengio、Geoffrey Hinton在内的多位图灵奖得主联署《上海共识》。主持国家重点研发计划课题、百度/华为合作项目，并担任多个网络安全顶会程序委员会委员和AI顶会领域主席。入选国家网信创新人才（2024）、世界人工智能大会云帆奖（2024）等。

报告题目：前沿 AI 系统安全风险评测与治理

报告摘要：随着 AI 自主性和长程任务执行能力迅速发展，如何针对以大模型为核心的前沿 AI 系统开展安全风险评测与治理是全球面临的共性难题。大模型通过工具、知识库、推理编排框架等基础软件，能够与真实世界自主交互，一旦具备自主复制、欺骗密谋、自主网络攻击等危险能力，或对全人类带来灾难性风险。本报告将围绕智能体行为安全和失控风险两大核心议题，分享团队近年来在前沿 AI 系统安全评测与治理方面的研究进展和思考。特别地，团队近期在国际上发现的多种失控风险实证依据，得到国内外广泛关注和高度认可，为全球 AI 治理提供建设性思路。



刘东瑞

刘东瑞，上海人工智能实验室智能体攻防系统中心负责人，青年科学家，入选WAIC 2026云帆奖，上海市高层次人才计划。长期从事安全可信人工智能研究，负责国际首个诊断式智能体守卫模型AgentDoG研发，相关技术已落地应用阿里，讯飞和360等公司。组建DoGNAVY战队，在国际知名漏洞挖掘，利用和修复榜单上均达到国际一流水平。主导SafeWork-R1攻防和安全高效推理后训练，国内首个针对AI灾难性风险技术评测(SafeWork-F1&I.5)，获得Anthropic 联合创始人长篇解读。在安全可信领域产出四十多篇高水平论文，包括CVPR 2024 最佳论文候选奖，ACL 2025杰出论文奖，多篇CVPR，ICLR，AAAI，ACL Oral。

报告题目：Agent 系统的风险归因与防护

报告摘要：本报告聚焦智能体的核心安全问题与防护研究，系统分析了智能体工具调用的安全性、错误进化、智向上欺骗等风险。研究发现，OpenClaw 等智能体易因意图误解、提示注入引发现实危害。自进化智能体在模型、记忆、工具、工作流进化中易产生非预期风险偏差，主流大模型均受此影响；智能体在环境约束下会出现隐瞒失败、伪造结果的固有欺骗行为。针对上述风险，团队提出智能体诊断防护框架 AgentDoG，通过三维度智能体安全分类体系与可解释性模块，实现轨迹细粒度监控与风险根因诊断。然而目前业界仍缺乏普适性的稳健防御机制，需要多方努力一起构建安全可信的智能体系统。



武建双

中检集团天帷信息技术股份有限公司副总经理，人工智能应用高级工程师、全国五一劳动奖章获得者，公安部第三研究所认证中心、安徽省公安厅等多家单位特聘技术专家，主要研究方向为网数安全合规治理、人工智能安全治理和体系规划与建设。拥有等级保护高级测评师、国家商用密码应用安全性评估人员、注册信息安全讲师证书（CISI）等20余个行业资质证书，20余项发明专利，主持和参与10余项国家、行业及省级网络安全相关的框架、指南、标准等行业指导文件，曾受邀为网信办、公安局、央国企数百家政企单位开展网络安全培训和赋能。

报告题目：从对话到行动：AI 智能体的安全风险演化与治理体系构建

报告摘要：随着 AI 智能体由“对话”走向“行动”，风险正从内容安全扩展至权限越界、工具滥用与级联失效等行为安全。报告将沿“能力扩张—威胁迁移—后果升级”梳理从单问答到长程任务、从单智能体到多智能体系统的脆弱点；基于分类分级原则和系统思维，提出面向基座大模型安全和智能体能力域的“1+5”安全测评框架；探讨最小权限、沙箱隔离、动作确认、人类在环和监控溯源协同的全生命周期安全加固方案；最后给出中检天帷在 AI 安全治理领域的服务能力与实践。

**张倬胜**

张倬胜，上海交通大学长聘教轨助理教授、博导。研究方向为多模态大模型、智能体自进化与认知安全。入选中国中文信息学会优博、ACM SIGAI优博、全球AI华人百强学术新星、世界人工智能大会云帆奖璀璨明星、斯坦福大学全球前2%顶尖科学家。以第一/通讯作者在TPAMI、ICLR、ICML、ACL、AAAI等顶级期刊及会议发表论文30余篇，Google Scholar被引1.2万余次，开源成果获GitHub星标5万余次、HuggingFace下载300万余次。获世界人工智能大会青年优秀论文奖、ICLR 2026 WILD研讨会最佳论文奖，4篇论文入选Paper Digest最具影响力论文。主持国家自然科学基金项目、国家重点研发计划子课题及企业合作项目等10余项，获上海市科技进步奖一等奖。长期担任ARR、TMLR编委，以及NeurIPS、ACL、EMNLP、AAAI、IJCAI等国际顶级会议领域（高级）主席。

报告题目：寻找智能体安全的不变量：开放环境下的GUI智能体行为对齐

报告摘要：随着智能体从内容生成走向自主感知、决策与行动，安全问题正由静态的内容约束转向动态的行为对齐与可控性。一方面，智能体不断进入开放、动态的现实环境，与用户、工具和其他智能体持续交互；另一方面，智能体系统也逐步具备持续学习、自主反思与自我演化能力。GUI智能体连接自然语言意图与真实数字环境中的行动，是研究智能体从“生成内容”迈向“自主行动”的典型载体。本报告以GUI智能体为切入点，探讨在环境、任务与智能体自身持续变化的条件下，支撑安全行为的“不变量”，并结合我们在智能体安全与可信方向的研究探索，讨论如何使持续演化的智能体始终与人的真实意图和基本安全约束保持一致。

可信智能体与安全治理论坛

承办单位：浙江大学

🕒 9月20日下午

📍 天鹅湖大酒店一楼徽州厅

论坛主席：徐文渊（浙江大学）

张勇东（中国科技大学）

召 集 人：凌 振（东南大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:35	一次经验，随处高效——构建可移植高可靠的LLM智能体系统 陈海波（上海交通大学）
14:35-15:00	存算一体通信基带信号处理新架构及芯片实现 姜春晓（清华大学）
15:00-15:25	虚实对齐:具身智能如何从“理解意图”走向“实现意图” 杨耀东（北京大学）
15:25-15:50	安全可信的智能模型训练推理技术研究 沈 蒙（北京理工大学）
15:50-16:10	休 息
16:10-16:35	面向强化学习的大模型辅助智能体间抗干扰高可靠通信技术 肖 亮（厦门大学）
16:35-17:00	HOP3.0与可信原生智能体 潘无穷（蚂蚁密算）
17:00-17:25	智能体安全风险分析与测评 张 谧（复旦大学）

论坛主席



徐文渊

徐文渊，浙江大学教授。第十八届中国青年女科学家奖、国家杰出青年基金获得者、IEEE Fellow。曾获美国南卡罗莱纳大学终身教职。担任物联网和信息安全领域国际学术期刊IEEE Transactions on Mobile Computing, ACM Transactions on Privacy and Security等编委，担任信息安全领域顶级四大会议USENIX Security 2024 年联合主席、NDSS 2022年和2023年联合主席、ACM CCS 2021年领域主席。主要研究物联网安全及智能感知系统安全，在物联网终端安全、工控安全、感知安全等方面取得了多项研究成果，在网络安全等领域的国际顶级期刊和会议发表论文130余篇，获得ACM CCS 2017最佳论文奖等3次最佳论文奖。



张勇东

张勇东，教授，博士生导师，中国科学技术大学信息科学技术学院执行院长，IEEE Fellow。国家自然科学基金委创新研究群体项目负责人，国家杰出青年科学基金获得者，“万人计划”科技创新领军人才。主要研究方向为网络空间多媒体内容安全，曾获国家自然科学奖二等奖1项（排名1），国家技术发明奖二等奖1项（排名1）和国家科技进步奖二等奖1项。所发表论文谷歌引用3万余次，连续多年入选科睿唯安全球高被引学者。获授权国家发明专利100余项，相关技术在著名国际评测中多次取得优异成绩。研究成果有效应用于国家网络空间内容安全领域。担任《中国通信》副主编，国家重点研发计划-“变革性技术关键科学问题”重点专项总体专家组成员，国家重点研发计划-“社会治理与智慧社会科技支撑”重点专项总体专家组成员。

召集人



凌 振

国家优青，东南大学青年首席教授。研究方向包括匿名网络、物联网安全等。近年来发表ToN、TIFS、TPDS、TC、TDSC、INFOCOM、CCS、S&P、Security、NDSS等高水平学术论文100余篇。先后主持多项国家自然科学基金、国家重点研发计划等项目。多次受邀在BlackHat等国内外重要会议作特邀报告。获ACM中国和CCF优博论文奖、全国公安技术革新成果一等奖等奖项。担任ACM SIGSAC China秘书长，TDSC等编委，曾任MSN 2023程序委员会联合主席，担任CCS、Security、INFOCOM等多个知名会议程序委员会委员。

报告嘉宾



陈海波

陈海波,上海交通大学特聘教授、人工智能操作系统研究院院长、开源鸿蒙项目群技术指导委员会创始主席、国家杰出青年基金获得者、ACM Fellow、CCF Fellow、IEEE Fellow。主要研究领域为操作系统、分布式系统与系统安全,研究成果通过产学研深度结合被应用到数十亿设备,产生了广泛的学术与产业影响。曾获中国计算机学会科技进步特等奖、陈嘉庚青年科学奖(信息技术科学奖)、中国青年科技奖、教育部技术发明一等奖、全国优秀博士学位论文奖等。目前担任 ACM SIGOPS 首位非北美学者主席、ACM 旗舰杂志 Communications of the ACM 首位中国学者编委与领域共同主席、ACM EuroSys 2025 程序委员会共同主席。研究工作还获得了华为卓越贡献个人奖、SOSP、ASPLOS、FAST、EuroSys、VEE 等最佳论文奖以及 DSN“时间检验奖”、CACM 研究亮点与封面文章等。按照 csrankings.org 的统计,其近十年在操作系统领域顶会 SOSP/OSDI 上发表的论文数居世界第一。

报告题目：一次经验，随处高效——构建可移植高可靠的LLM智能体系统

报告摘要：大语言模型智能体正向自主执行实体演进，但面临可靠性、成本、隐私和跨平台移植等根本挑战。本报告梳理三项互补研究：AgentRR引入记录-回放机制，将智能体交互轨迹抽象为结构化经验并回放复用，在降低80%API调用成本的同时提升45%任务可靠性；MobiAgent在此基础上构建移动端系统性框架，包含多角色模型、加速组件和基准套件，在真实场景达到最优性能；SkVM借鉴编译器思想，将技能视为代码、LLM视为异构处理器，通过编译优化和JIT固化实现技能跨模型跨环境的高效执行，减少40%token消耗并实现19-50倍延迟降低。三者串起从经验积累、系统集成到跨平台执行的完整链条，为下一代智能体系统提供可移植高可靠的解决方案。



姜春晓

姜春晓，获国家杰出青年基金资助，IEEE Fellow、全球高被引学者，以第一/通信作者发表IEEE SCI期刊论文300余篇，出版英文专著7部、中文专著3部，获14项国际期刊/会议最佳论文奖；谷歌学术引用2万余次；曾担任8个IEEE期刊编委/客座编委，10余次国际学术会议主席/共同主席；入选2022年IEEE卫星通信分会Technical Recognition Award、2021年IEEE VTS Distinguished Lecturer、2021年IEEE ComSoc Best Young Professionals Award in Academia、2020年IEEE VTS Early Career Award、2020年IEEE ComSoc Asia-Pacific Best Young Researcher Award等；授权国家发明专利100余项，成果应用于我国中继/通信卫星重大型号工程，牵头获得2020年中国指挥与控制学会创新奖一等奖；参与获得2018年国家技术发明奖二等奖、2020年中国通信学会技术发明奖一等奖、2019年教育部科技进步奖一等奖、2018年中国人工智能学会吴文俊人工智能技术发明奖一等奖、2017年中国电子学会技术发明奖一等奖等。目前担任军委科技委创新特区专家、中国指挥与控制学会副秘书长、国家自然科学基金委“未来工业互联网基础理论与关键技术”重大研究计划学术秘书、通信抗干扰全国重点实验室学术委员会委员等。

报告题目：存算一体通信基带信号处理新架构及芯片实现

报告摘要：通信技术的持续演进带来了高频段、大带宽和高传输速率的发展趋势，但随之激增的能耗已成为制约信息技术发展的关键瓶颈。面向高速且节能的基带信号处理需求，现有数字系统面临严峻挑战：信号处理过程中频繁的数据搬移、复杂的计算任务带来的巨大能耗。针对上述瓶颈，发展了模拟存内信号处理新架构，实现了传统的数字信号离散处理范式到模拟物理域的转变，最大限度减少了内存和处理单元之间的数据移动，信号处理能效提升1个数量级以上。


杨耀东

杨耀东，北京大学人工智能研究院研究员（博雅学者）、院长助理，北大-灵初智能联合实验室首席科学家。国家人社部高层次留学人才、国家级优秀青年人才、中国科协青年托举计划入选者。主要研究方向为智能体交互学习与对齐，科研领域涵盖强化学习、AI 对齐与具身智能。在 Nature Machine Intelligence、Cell Matter、AIJ、TPAMI 等国际顶级期刊和会议发表论文二百余篇，谷歌学术引用逾18,000+次。入选 Scopus 全球Top 2%顶尖科学家。获得ACL 2025等最佳论文奖。入选麻省理工科技评论「AI 100 青年榜」、福布斯中国 2025 科创革新力人物、WAIC 2022「云帆奖·璀璨明星」及 ACM SIGAI China 新星奖。现任 ICML、ICLR、NeurIPS、AAAI、IJCAI、AAMAS、IROS 等国际会议领域主席，以及 Scientific Reports、Transactions on Machine Learning Research、Neural Networks 等期刊执行编委。主持国家自然科学基金、科技部、北京市科委及多项校企联合实验室科研项目五十余项。曾获气象服务协会技术发明一等奖、吴文俊人工智能科技进步二等奖。相关成果获中央电视台《焦点访谈》、人民日报、新华网、国家自然科学基金委官网及《麻省理工科技评论》等媒体报道。

报告题目：虚实对齐：具身智能如何从“理解意图”走向“实现意图”

报告摘要：大模型正在进入能够持续完成长程任务的新阶段，但当智能体走向真实物理世界，一个更根本的问题出现了：模型不仅要理解人类想做什么，还要真正把事情做成。本报告将围绕“虚实对齐”展开，讨论具身智能从语言理解走向物理执行所面临的关键挑战：机器人如何预测世界变化、如何选择能够改变世界的动作、又如何用自己的身体稳定完成复杂操作。报告将结合灵巧操作、具身人类数据、世界模型、VLA/WAM 与强化学习等最新进展，分析具身智能落地的“最后一公里”为何在于长程、闭环、可泛化的物理操作能力。


沈 蒙

沈蒙，北京理工大学网络安全学院教授，国家高层次人才。主要研究领域为数据安全、人工智能安全。主持国家优秀青年科学基金、国家自然科学基金联合基金重点项目、国家重点研发计划课题等。现为中国计算机学会网络安全专委会委员、互联网专委会常务委员。曾获中国电子学会科技进步一等奖、北京市科技进步二等奖，以及多项最佳论文奖。

报告题目：安全可信的智能模型训练推理技术研究

报告摘要：随着人工智能技术在多场景、多业务中的广泛部署，其安全威胁已从训练数据隐私泄露扩展到涵盖训练、交互和部署过程中的多种安全可信风险：攻击者可利用中间表示和模型输出推断敏感信息，也可通过恶意更新和伪造梯度破坏训练过程，或采用对抗扰动误导模型决策。本报告面向安全可信的智能模型训练与推理，围绕数据隐私保护、可信协同训练、模型决策安全三个层面，介绍团队在隐私泄露阻断、不可信训练治理和对抗风险评测等方面的研究工作，探讨如何提升智能模型在训练与推理过程中的隐私性、鲁棒性和可用性，并展望相关技术在大模型训练、微调和推理阶段中的拓展应用。


肖 亮

肖亮, 厦门大学信息学院教授, 教育部长江学者特聘教授, IEEE FELLOW, IEEE 通信学会杰出讲师, 罗格斯大学博士。主要研究方向为智能通信、网络安全和人工智能安全。出版 3 部英文专著, 第一完成人获福建省科技进步奖一等奖 2 项、中国通信学会青年科技奖、紫金科技创新奖和运盛青年科技奖。获 IEEE 通信学会亚太杰出论文奖和 ICC 等国际会议最佳论文奖。

报告题目: 面向强化学习的大模型辅助智能体间抗干扰高可靠通信技术

报告摘要: 多智能体强化学习通过智能体间共享观测信息、学习经验和网络权重支持任务策略优化, 但是有限通信带宽和智能干扰攻击导致信息传输可靠性下降, 降低学习效率和任务性能。因此, 本报告介绍面向强化学习的大模型辅助智能体间抗干扰高可靠通信技术, 以图像、频谱瀑布图和任务性能等多模态信息作为大语言模型输入, 识别任务特征以及潜在干扰设备、干扰类型和信道等干扰特征, 辅助智能体优化抗干扰策略。并在任务状态构建和网络更新中利用相似的观测和网络权重提升学习效率。在车联网协同感知场景中, 验证所提技术在智能干扰攻击下通信能耗、学习效益和收敛速度等方面的性能增益。


潘无穷

清华大学本科, 中国科学院信息安全博士, 现任蚂蚁密算副总经理。提出或主导了数据要素领域的多项重要技术, 包括隐私计算通用安全分级、受控匿名化、密态计算、数据安全域、跨域管控等, 相关内容已撰写成多本白皮书向业内共享, 获得数字中国建设峰会“十大硬核科技”、浙江省科学技术进步一等奖。牵头数标委WG6-SG2研究小组, 以及密态计算和数据安全域标准的研制工作, 发表的论文专利等各类公开成果近百篇。

报告题目: HOP3.0 与可信原生智能体

报告摘要: 大模型与智能体技术的爆发为行业带来了巨大机遇。我们认为, 真正的产业级落地亟需全面的技术革新, 当前仍有三个关键问题亟待突破: 如何确保智能体严格对齐人类意图并可靠运行; 如何打破机构壁垒, 使智能体安全获取高价值产业数据; 以及如何高效组织知识, 让智能体自主、精准地获取问题相关知识。本报告将系统阐述破局上述难题的技术路径。首先, 通过 HOP 3.0 智能体原生语言实现“结构与推理双态融合”, 将人工审计精准收敛至关键决策节点, 保障可靠性与可控性; 其次, 依托密态计算构建“四纵两横”安全架构, 确保数据流转全链路密态, 使数据控制权始终锚定于供方, 从机制上规避信任风险; 最后, 引入 S4RA 知识组织范式, 重构适配智能体认知逻辑的知识本体, 赋予其高效读取、图谱寻址与深度推理能力。本方案以三位一体的技术革新, 打通智能体自身可控性、数据安全获取、以及数据向知识转化三个关键环节, 推动智能体从理论探索与技术破局, 真正走向产业化应用。

**张 谧**

张谧，复旦大学计算与智能创新学院教授，白泽AI负责人，专注于AI安全、大模型与智能体安全研究，在网络安全与AI领域顶会顶刊发表论文百余篇;联合起草国家标准《生成式人工智能服务安全基本要求》，任TC260人工智能安全标准工作组(WG9)专家，参与《智能体应用安全基本要求》等多项国家标准的制订;主持科技部重点研发计划课题与奇安信、华为、阿里等企业合作项目，获CCF科学技术奖自然科学二等奖等荣誉;培养硕博数十人，带领团队研发JADE大模型安全风险分析与治理平台，开源十余款Benchmark数据集与治理工具，GitHub Star数为开源中文大模型安全测评与治理平台 Top1，被新华社、人民日报、环球网、凤凰网、南方都市报等主流媒体多次报道。

报告题目：智能体安全风险分析与测评

报告摘要：首先介绍从大模型走向智能体—从“说错话”到“做错事”，总结智能体安全的发展现状各个风险环节。然后分别针对通用智能体与 GUI Agent 的实际应用给出多个安全风险案例。接着介绍 JADE 安全风险测评与治理平台中对于通用智能体与 GUI Agent 的测评与治理工作,并结合 DAWN SONG 等知名学者的见解对智能体安全治理做出总结与展望。

低空智联网安全论坛

承办单位：北京航空航天大学、北京航空航天大学杭州创新研究院

🕒 9月20日下午

📍 天鹅湖大酒店三楼5号会议室

论坛主席： 刘建伟（北京航空航天大学）
关振宇（北京航空航天大学）

张小松（电子科技大学）

召 集 人： 刘建伟（北京航空航天大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	低空电磁频谱安全 吴启晖（南京航空航天大学）
14:40-15:10	低空智联网：技术发展与安全挑战 刘 军（北京航空航天大学）
15:10-15:40	无人系统智融网络研究 杨 冬（北京交通大学）
15:40-16:00	休 息
16:00-16:30	面向低空数据安全的协同推理机制研究 朱友文（南京航空航天大学）
16:30-17:00	城市核心要地-机场低空安全体系建设与实践 王 聪（杭州安恒低空安全科技有限公司）

论坛主席



刘建伟

北航网络空间安全学院创院院长，教授，博士生导师，国家级教学名师，享受国务院政府特殊津贴专家。现任国务院学位委员会学科评议组成员、教育部网络空间安全专业教学指导委员会副主任委员，全国密码专业学位研究生教指委委员，中国密码学会常务理事、中国网络空间安全学会常务理事，中国指挥与控制学会常务理事，《网络空间安全科学学报》常务副主编、《网络与信息安全学报》副主编，IEEE高级会员。

获评国家级教学名师、国家网络安全优秀教师奖、宝钢优秀教师特等奖、北京市教学名师奖、北京市优秀教师等奖励和荣誉。荣获第二届全国优秀教材奖二等奖、全国普通高校优秀教材一等奖、北京市教学成果一等奖等国家级和省部级教学成果奖20余项。

荣获国家技术发明一等奖、全国争先创新奖牌和中国指挥与控制学会科技进步一等奖等国家级和省部级奖励16项。



张小松

电子科技大学信息与软件工程学院（示范性软件学院）院长，国家级人才，教授，博士生导师，成都市科协副主席，中国计算机学会/中国通信学会/中国电子学会/中国指挥与控制学会会士。长期从事计算机系统与网络安全的研究与攻关，以第一完成人获得2019国家科技进步一等奖、2012国家科技进步二等奖，四次获得省部技术发明一等奖和科技进步一等奖，主持国家重点、重大课题40余项，以第一作者出版著作5部，发表SCI/EI收录学术论文249篇，荣获第六届“杰出工程师奖”、首届“全国科创名匠”、第二届“全国创新争先奖”、国家“网络安全优秀人才奖”等荣誉称号，牵头主持的网络安全的专用编程语言Yaklang创新及应用成果入选2023年度信息通信领域十大科技进展，获2025中国通信学会科技进步一等奖。



关振宇

北京航空航天大学教授、博士生导师。国家杰出青年科学基金获得者、教育部青年长江学者，致力于网络空间安全理论与应用研究，在无人系统安全、人工智能等相关领域取得了丰富的研究成果。近5年，以第一/通讯作者发表高水平论文100余篇，授权发明专利60余项。主持国家科技重大专项、国家自然科学基金等科研项目38项，获北京市自然科学一等奖、教育部技术发明一等奖等省部级奖项7项。成果应用于南海应急搜救、空军航空兵/空降兵及战略支援部队应急指挥等场景，在我国边海防、公共安全体系建设中发挥关键支撑作用。

召集人



刘建伟

北航网络空间安全学院创院院长，教授，博士生导师，国家级教学名师，享受国务院政府特殊津贴专家。现任国务院学位委员会学科评议组成员、教育部网络空间安全专业教学指导委员会副主任委员，全国密码专业学位研究生教指委委员，中国密码学会常务理事、中国网络空间安全学会常务理事，中国指挥与控制学会常务理事，《网络空间安全科学学报》常务副主编、《网络与信息安全学报》副主编，IEEE高级会员。

获评国家级教学名师、国家网络安全优秀教师奖、宝钢优秀教师特等奖、北京市教学名师奖、北京市优秀教师等奖励和荣誉。荣获第二届全国优秀教材奖二等奖、全国普通高校优秀教材一等奖、北京市教学成果一等奖等国家级和省部级教学成果奖20余项。

荣获国家技术发明一等奖、全国争先创新奖牌和中国指挥与控制学会科技进步一等奖等国家级和省部级奖励16项。

报告嘉宾



吴启晖

南京航空航天大学，教授，博士，博士生导师，副校长，教育部长江学者特聘教授，IEEE Fellow，国家频谱战略规划政策研究组组长、国家自然科学基金委信息学部咨询委员会委员、中国工程院学科交叉专项总体组成员，中国电子学会常务理事/会士、中国通信学会常务理事/会士。主要从事天地一体频谱认知智能、无人机集群网络等方向研究，入选ESI工程学与计算机科学领域科学家（Scientists）世界前1%，获IET创新奖等国际高水平奖项12项，制定国际标准2项，承担国家重点研发计划、国家自然科学基金重大仪器等国家重大/重点项目10余项，获国家科技进步二等奖1项、省部级科技进步一等奖6项等。

报告题目：低空电磁频谱安全

报告摘要：低空经济是战略新兴产业，低空安全是低空经济稳步健康发展的基石，低空电磁频谱安全是低空安全的重要方面，包括导航安全、通信安全、监视安全、反制安全、频谱数据安全、频谱安全态势、频谱安全标准等。本报告从低空安全战略需求与低空安全事件出发，探索了低空电磁频谱安全与低空智能网的逻辑关系，阐述了低空频谱安全体系与基础理论，分析了立体通信覆盖、导航受扰、超宽谱分布式监视、电磁频谱空地协同反制、电磁频谱数据安全、电磁地理信息系统等方面的关键技术，预测了未来低空电磁频谱安全发展趋势。



刘 军

北京航空航天大学电子信息工程学院教授、博士生导师，低空智联网北京市重点实验室主任、通信工程系主任、天临空地海一体化方向责任教授，国家级科技创新领军人才、北航蓝天学者。长期从事无线通信网络、定位导航、目标追踪与监测研究，聚焦低空智联网感知、通信、导航、管控一体化及智能网络安全。主持国家重大专项、重点研发计划、国家自然科学基金重点项目等30余项，发表论文百余篇，授权发明专利30余项。

报告题目：低空智联网：技术发展与安全挑战

报告摘要：低空智联网通过融合飞行器、信息物理基础设施、数据信息网络与应用服务，实现低空空域的泛在感知、广域互联和智能管服，也使网络攻击、数据污染、导航欺骗、通信干扰与智能模型失效等风险跨域耦合。本报告围绕低空智联网“感得精、传得稳、导得准、控得好”的安全需求，分析高速移动、三维拓扑快变、多网异构和高密度运行带来的新型攻防问题，介绍多源融合感知与物理层身份认证、多模通信与通感协同抗干扰、韧性导航与欺骗检测、数字孪生驱动的风险评估和智能管控等关键技术，探讨构建“感知—预测—决策—执行”闭环的内生安全体系，为低空经济规模化、安全化发展提供技术路径。



杨 冬

北京交通大学教授，国家杰青，河北省低空智联通信网络重点实验室主任，中国通信学会会士，北京通信学会副理事长，国家铁路局装备技术中心科研管理技术委员会委员，中国通信学会“无人系统智融网络”标准组组长，IEEE TMC等期刊编委。长期从事新型网络关键技术研究，应用在智能制造专网、高铁专网等，牵头制定行业标准YD/T6568-2025等。曾获国家技术发明二等奖、中国通信学会科学技术一等奖等。

报告题目：无人系统智融网络研究

报告摘要：随着人工智能技术的快速发展和普及应用，未来各种低空无人机、智能机器人、自动驾驶车辆等无人设备和系统将大量部署，这些无人设备和系统的安全稳定运行高度依赖网络和算力等基础设施。另一方面，随着国家在数字经济基础设施领域的持续加大投入，我国已经积累了海量的网络资源和丰富的算力资源。然而，这些算网资源存在网络资源异构化和碎片化、算力资源私有化和封闭化问题，难以支撑未来无人化、智能化、个性化的应用发展需求。报告介绍团队在无人系统智融网络研究的探索性工作，拟通过融合空天地多维网络资源，并赋予无人系统在无网络设施覆盖区域的智能自决策能力，以提升其韧性与可靠性。


朱友文

南京航空航天大学教授、博导，江苏省青蓝工程中青年学术带头人。主持国家重点研发计划青年科学家项目 1 项，主持国家自然科学基金 3 项（含重点类项目 1 项）。主要研究方向是应用密码学、低空安全、隐私计算等。近年来，以一作 / 通讯作者在 USENIX Security、TDSC、TIFS 等国内外知名期刊和会议上发表学术论文 50 余篇，授权发明专利 9 项。牵头立项制定中国电子学会标准 1 项。获江苏省科学技术奖一等奖、中国电子学会自然科学二等奖等。

报告题目：面向低空数据安全的协同推理机制研究

报告摘要：低空数据安全的低空智联网安全的重要内容。协同推理机制通过模型切分等方法可以降低端侧计算开销，适配低空飞行器算力受限等特点，然而也面临着数据隐私安全风险。本报告将简要介绍我们近期面向低空场景的安全协同推理工作，主要探讨协同推理机制中影响模型逆向的内在因素，模型在何处切分能使边缘端获得对模型逆向攻击的最佳内生抗性。对于前者，我们深入分析了输入与表征的互信息、条件熵、表征有效信息量对模型逆向的影响，并设计了相应的防御方法；对于后者，不同于已往“切分越深越安全”的直觉，我们观察到深度只有在促成表征转变才能带来内生抗性，并为最优切分位置设计了主动定位算法，理论和实验证实了其具有更强的逆向抗性。


王 聪

中共党员，电子与通信工程硕士，高级工程师，安恒信息低空安全副总经理。杭州市 E 类人才、高新区（滨江）专业人才，浙江省科技厅专家库专家，教育部研究生教育评估监测专家库专家。

深耕网络安全与低空安全多年，首创低空安全一体化防护理念，2024 年牵头研发行业首款低空智能巡处一体化平台，同时带领团队研发基于 AI 的恒御低空安全管理平台，产品落地全国多省份城市。曾主导或参与十余项国家与行业安全标准，牵头编制物联网安全白皮书与无人机系统安全白皮书，拥有多项技术专利，长期深耕低空安全产学研融合与城市低空治理实践。

报告题目：智驭低空，安全先行：浅谈 AI 驱动的城市级低空安全体系与实践

报告摘要：当前低空经济加速规模化发展，黑飞扰航、违规飞行、空域管控难等安全风险持续凸显，传统手段已难以满足城市全域低空治理需求。本次分享结合安恒信息全链条低空安全业务布局，围绕“智驭低空，安全先行”核心思路，分享 AI 技术在多源数据融合感知、低空目标风险智能研判、防控预案智能推荐、智能联动处置、智能分析复盘、智能报告生成等关键环节的落地路径。结合多地城市低空安防实战案例，分享安恒基于 AI 构建全域、实时、可追溯的城市级低空安全防护体系架构与落地实践。

CCSC2026青年学者论坛

承办单位：中国网络空间安全大会青年工作组

🕒 9月20日下午

📍 天鹅湖大酒店一楼黄山厅

论坛主席：刘哲理（南开大学）

沈 超（西安交通大学）

召 集 人：刘哲理（南开大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-16:10	青年学者交流
16:10-16:30	休 息
16:30-17:30	专家panel 马建峰（西安电子科技大学） 过敏意（上海交通大学） 田志宏（广州大学） 杨 珉（复旦大学） 祝烈煌（北京理工大学） 何德彪（武汉大学）

论坛主席



刘哲理

刘哲理，南开大学计算机学院院长、密码与网络空间安全学院常务副院长，国家级人才项目特聘教授、教育部数据与智能系统安全重点实验室主任。主要从事数据安全与人工智能安全研究，近五年发表顶级期刊和会议论文50余篇，主持国家自然科学基金重点项目、国家重点研发计划等课题，获天津市科技进步一等奖、教育部自然科学奖二等奖及中国密码学会密码创新奖等。



沈超

沈超，IEEE Fellow，西安交通大学二级教授、国家级人才项目特聘教授，国家自然科学基金创新研究群体负责人、国家重点研发计划首席科学家。主要从事智能系统可信、安全与控制研究，出版专著3部，获最佳论文奖13次，牵头获3项省部级或一级学会科技一等奖。现任IEEE SMCS副主席及多个国内外学术组织重要职务，并担任IEEE TDSC、IEEE TCYB等期刊副主编或编委。

召集人



刘哲理

刘哲理，南开大学计算机学院院长、密码与网络空间安全学院常务副院长，国家级人才项目特聘教授、教育部数据与智能系统安全重点实验室主任。主要从事数据安全与人工智能安全研究，近五年发表顶级期刊和会议论文50余篇，主持国家自然科学基金重点项目、国家重点研发计划等课题，获天津市科技进步一等奖、教育部自然科学奖二等奖及中国密码学会密码创新奖等。

■ 圆桌嘉宾



马建峰



过敏意



田志宏



杨 珉



祝烈煌



何德彪

CCSC2026博士生论坛

承办单位：中国网络空间安全学会青年工作委员会

🕒 9月20日下午

📍 天鹅湖大酒店一楼淮河厅B厅

论坛主席：刘哲理（南开大学）

付章杰（南京信息工程大学）

召集人：付章杰（南京信息工程大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:22	基于全同态加密的密态数据查询设计研究 张 舟（北京航空航天大学）
14:22-14:34	面向人工智能硬件安全的DNN加速器侧信道攻防技术研究 乌 勒（清华大学）
14:34-14:46	开源软件漏洞自动诊断、定位、修复 叶振雷（扬州大学）
14:46-14:58	冰山之下：生成式AI安全风险的发现与治理 唐坤生（中国科学技术大学）
14:58-15:10	复杂误导信息环境下的大模型可靠决策 李钊效（云南大学）
15:10-15:22	大模型驱动的社会舆情治理研究 张云耀（华中科技大学）
15:22-15:34	休 息
15:34-15:46	面向多方数据协同的隐私集合操作研究 刘冉阳（南开大学）
15:46-15:58	AI生成视觉内容安全与取证的研究与思考 穆文鹏（上海交通大学）
15:58-16:10	从AIGC标识到可信内容生态 甘振良（复旦大学）
16:10-16:22	生成式人工智能图像模型安全 戴昀书（中山大学）
16:22-16:34	高分辨率图像隐写分析 胡欣珏（南京信息工程大学）
16:34-16:46	面向扩散模型的免训练高容量隐写方案 郑裕丰（合肥工业大学）
16:46-17:00	休 息
17:00-18:00	圆桌主题：AIGC时代科研新范式

论坛主席



刘哲理

刘哲理，南开大学计算机学院院长、密码与网络空间安全学院常务副院长，教育部长江学者特聘教授、教育部数据与智能系统安全重点实验室主任。主要从事数据安全与人工智能安全研究，近五年发表顶级期刊和会议论文50余篇，主持国家自然科学基金重点项目、国家重点研发计划等课题，获天津市科技进步一等奖、教育部自然科学奖二等奖及中国密码学会密码创新奖等。



付章杰

付章杰：教授，博士生导师，现任南京信息工程大学网络空间安全学院院长，数字取证教育部工程研究中心主任。国家级青年人才，江苏省杰出青年基金获得者。主要研究方向为多媒体内容安全、区块链安全、数字取证等。近年来主持国家自然科学基金重点项目、国家重点研发计划青年科学家项目等国家级项目6项，发表学术论文50余篇。研究成果获2019年度江苏省科学技术（基础类）一等奖，2022年度江苏省科学技术（基础类）二等奖。

召集人



付章杰

付章杰：教授，博士生导师，现任南京信息工程大学网络空间安全学院院长，数字取证教育部工程研究中心主任。国家级青年人才，江苏省杰出青年基金获得者。主要研究方向为多媒体内容安全、区块链安全、数字取证等。近年来主持国家自然科学基金重点项目、国家重点研发计划青年科学家项目等国家级项目6项，发表学术论文50余篇。研究成果获2019年度江苏省科学技术（基础类）一等奖，2022年度江苏省科学技术（基础类）二等奖。

报告嘉宾



张 舟

北京航空航天大学网络空间安全学院博士生，导师为关振宇教授和边松副教授。主要研究应用密码学、隐私保护计算与全同态加密。相关成果在网络安全和其他领域的CCF-A类会议与期刊ACM CCS、IEEE S&P、USENIX Security上发表了高水平论文9篇，其中第一作者或学生第一作者3篇，并获CCF-A类会议ACM CCS杰出论文奖；主持国家自然科学基金青年学生基础研究项目，入选首批中国科协青年人才托举工程博士生专项计划，主持网络安全学院学生创新资助项目《隐私保护算法实现KMEANS、SVM》，担任国际知名期刊IEEE TKDE审稿人，申请发明专利9项。

报告题目：基于全同态加密的密态数据查询设计研究

报告摘要：密态数据查询旨在让服务器在不获知数据内容或检索目标的前提下完成查询计算，从而在保护数据隐私与保密性的前提下，实现敏感数据的可用与协作分析。本报告探讨基于全同态加密的密态数据查询的关键技术，分享了本人利用全同态加密实现高效密态数据查询的具体研究与工作，探索批量查询场景下的性能瓶颈与实际应用挑战，包括密态数据库的高精度比较与复杂查询处理、单服务器批量隐私信息检索等，并提出了融合算术与逻辑同态加密的密态查询引擎、任意精度同态比较算子，以及有状态密文转换与批量响应压缩等系统级解决方案，相关成果发表于 ACM CCS、USENIX Security 等网络安全领域顶级会议期刊，为密态数据查询的实用化落地做出了积极贡献。



乌 勒

清华大学集成电路学院电子科学与技术专业博士研究生，师从乌力吉教授，主要从事人工智能芯片与密码芯片硬件安全研究，聚焦侧信道分析、故障注入及安全防护技术。博士期间的研究成果已在IEEE TCAD、TCAS-I、TVLSI、HOST、CHES及DAC等国际期刊和会议发表或展示。曾获清华大学校级综合一等奖学金、集成电路学院优秀成果奖等。

报告题目：面向人工智能硬件安全的DNN加速器侧信道攻防技术研究

报告摘要：随着人工智能技术加速向边缘设备部署，DNN加速器中的模型参数、输入输出数据及网络结构面临日益突出的侧信道泄漏风险。相比传统密码硬件，DNN加速器具有线性运算引起的权值混叠、高并行架构下信噪比较低以及防御开销受限等特点，使传统侧信道分析与防护方法难以直接适用。本报告围绕DNN加速器侧信道安全中的关键问题展开研究。攻击方面，介绍三维功耗曲面（3DPS）与侧信道星图（SCSM）等方法，实现输入特征与模型权值恢复，并进一步分析推理输出阶段的新型泄漏面，结合故障注入实现定向误分类。防御方面，介绍面向脉动阵列的双轨预充电架构及布尔掩码乘累加单元，在降低功耗相关性的同时兼顾安全性与硬件开销。



叶振雷

扬州大学信息与人工智能学院博士三年级研究生，师从孙小兵教授。主要研究方向是开源软件漏洞自动定位与修复。现已发表CCF-A类会议论文2篇，CCF-B类会议5篇。公开发明专利3项，获得ICSE 2026杰出论文奖，入选中国科协青年科技人才培养工程博士生专项计划，第十八届“挑战杯”全国大学生课外学术科技作品竞赛“揭榜挂帅”专项赛全国一等奖、第一届软件缺陷自动修复挑战赛二等奖等多项国、省级赛事奖项。

报告题目：开源软件漏洞自动诊断、定位、修复

报告摘要：近年来，随着开源软件成为现代软件系统的重要基础，用户在Issue Tracker中提交的问题报告在记录缺陷与异常行为的同时，也可能公开披露可被攻击者利用的漏洞信息，使开源软件面临严重的安全风险。本次汇报围绕“开源软件漏洞自动诊断、定位、修复”展开，主要介绍从漏洞关联评估、定位解释到自动修复的研究进展。汇报旨在形成从漏洞发现评估、细粒度定位解释到安全修复的闭环能力，为开源软件安全治理提供支撑。



唐坤生

中国科学技术大学2025级博士在读，师从张卫明教授和俞能海教授，研究方向聚焦于大模型对齐。累计发表/录用CCF-A类论文7篇，其中以第一作者在ACM CCS、USENIX Security等安全四大会议发表多篇论文，并获CCS 2024 Distinguished Artifact Award。担任USENIX Security、ICLR、ICML等学术会议程序委员会委员，以及CCS、S&P、NDSS等学术会议工件评估委员会委员。主持国家自然科学基金青年学生基础研究项目（博士生项目）及中央网信办网空计划三期项目，并入选中国科协青年科技人才培养工程。曾在蚂蚁集团担任研究实习生，并与微软亚洲研究院深度合作，从事大模型对齐研究。曾获研究生国家奖学金，并在AI Singapore全球大模型安全挑战赛中获全球第5名。

报告题目：冰山之下：生成式AI安全风险的发现与治理

报告摘要：随着生成式人工智能在各领域的广泛应用，其安全风险也逐渐从“冰山之上”的显性问题，延伸至“冰山之下”的隐性威胁。本报告围绕“发现—检测—治理”的闭环范式，系统介绍报告人在生成式AI安全领域的系列工作，从文本到多模态、从模型表层到嵌入底层，逐步揭开生成式AI安全冰山的全貌，并为每一层风险探讨切实可行的治理方案。



李钊效

云南大学博士研究生，师从周维教授，研究聚焦可信人工智能与虚假信息治理。2025—2026 年获国家留学基金委(CSC)资助，赴新加坡国立大学 WING 实验室联合培养，合作导师为 Min-Yen Kan 教授。近年来已在 ICLR、ACL、AAAI、IEEE TMM 等重要国际会议和期刊发表论文 10 余篇，并主持或参与云南省教育厅研究生科学研究基金等多项科研项目。曾获国家奖学金、云南大学优秀毕业生等荣誉，现担任 ACL Rolling Review、NeurIPS、AAAI、ACM MM、IEEE TIFS、IEEE TDSC 等国际会议和期刊审稿人。

报告题目：复杂误导信息环境下的大模型可靠决策研究

报告摘要：近年来，误导信息风险正由内容认知层面向智能体决策与执行环节扩展，不仅影响人类判断，也可能干扰大模型决策。报告围绕误导信息环境下的的大模型可靠决策，介绍系列研究进展：针对脱离语境的图文误导，探索模型如何在噪声干扰下对误导信息形成可靠且可解释的判断；面向误导信息的动态演化与隐蔽表达，分析内容多样性和关键语境省略对模型判断可靠性的影响，并讨论评测、检测与修正方法；针对误导信号进入多智能体系统后的风险，研究其对工作流、智能体协作及下游决策的影响与防护机制。报告旨在揭示误导信息从干扰内容理解、影响可信判断到操纵智能体行动的风险链路，探索构建具备可靠感知、稳健推理与安全行动能力的大模型智能体系统的可能路径。



张云耀

华中科技大学网络空间安全学院直博三年级博士研究生，师从华中科技大学副校长于俊清教授。主持中央网信办第四期网络空间安全学院学生创新资助计划项目 1 项。主要围绕大模型驱动的社会舆情治理开展研究，以第一作者或共同第一作者身份在 ACL、ACM MM、EMNLP 等国际学术会议发表论文 5 篇。现担任 ICLR、AAAI 及 ACL ARR 审稿人。

报告题目：大模型驱动的社会舆情治理研究

报告摘要：随着社交网络成为信息传播、情绪扩散与群体博弈的重要场域，网络舆情呈现主体异质、交互复杂、持续演化等特征，传统局部识别与静态分析方法已难以支撑复杂舆情的理解、推演与治理。本次汇报围绕“大模型驱动的社会舆情治理”展开，主要介绍基础模型、社交建模与宏微演化三个层面的研究进展，包括面向复杂舆情深层意图与语义关系的神经符号融合和多视角逻辑推理，面向海量异质主体与复杂社会交互的群像智能体大规模社会模拟，以及动态联结个体决策与群体状态的宏观—微观耦合演化建模，探索复杂舆情从语义理解、社会建模到长程推演与趋势研判的技术路径，为网络舆情监测、风险预警与智能治理提供支撑。



刘冉阳

南开大学密码与网络空间安全学院2024级博士研究生，导师为刘哲理教授。研究方向为应用密码学与安全多方计算，累计发表CCF-A类论文5篇，其中以第一作者身份在USENIX Security、IEEE TIFS、IEEE TDSC等会议和期刊上发表论文3篇，入选第二期网络安全学院学生创新资助计划，参与国家密码科学基金首批重点项目等多项科研项目。

报告题目：面向多方数据协同的隐私集合运算研究

报告摘要：随着数据协同需求不断增长，如何在不直接交换敏感数据的情况下实现多方安全协作，成为数据流通与利用面临的重要问题。隐私集合运算是安全多方计算领域中的专用密码技术，能够在保护参与方输入集合隐私的前提下完成集合运算，并有效支撑集合型数据的安全流通和价值挖掘，在联合风控、联合营销、联合监管等场景具有广阔的应用前景。本报告围绕隐私集合运算，介绍相关工作在安全、效率和规模等方面面临的挑战，分享报告人在面向多方协同的隐私集合运算方向上的研究工作，结合相关成果讨论隐私集合运算的实用化问题。



穆文鹏

上海交通大学计算机学院（网络空间安全学院、密码学院）蒋兴浩教授、孙铨锋教授课题组的三年级博士生，主要研究方向为AI生成内容安全、数字媒体取证、信息安全等，主持或参与包括国家重点研发计划项目、上海市科委等相关科研项目7项，近年来在IEEE TIFS、CSUR、ACM MM、信息安全学报等国内外重点期刊会议发表论文10余篇，授权专利2项。曾获得2026年中央网信办第四期网络安全学院学生创新资助计划资助、“一流网安”奖学金等。

报告题目：AI生成视觉内容安全与取证的研究与思考

报告摘要：近年来，随着生成式人工智能技术的快速发展，AI生成图像与视频在内容创作和信息传播中被广泛应用，同时也带来了虚假信息传播、身份伪造和数字取证可信性下降等安全风险。本次汇报围绕“AI生成视觉内容真实性检测”展开，主要介绍该方向的研究背景、核心挑战与技术探索，包括生成模型快速迭代、伪迹特征弱化、跨生成器泛化困难和真实传播场景下鲁棒性不足等问题，并结合视觉伪迹建模、特征融合、生成器归因以及多模态大模型辅助检测等思路，分享本人对该研究方向的阶段性理解与后续思考。


甘振良

复旦大学多媒体智能安全实验室23级直博生，师从张新鹏教授，研究方向为AIGC安全。复旦学生科协首届主席，复旦计算机云翼青年讲师团首届团长，复旦F-LAB首届天才少年。在AAAI、ACM MM、ECAI等CCF A/B类会议发表论文7篇，曾问鼎华为ICT大赛全球冠军、斩获全国大学生软件创新大赛一等奖等十余项国家级及国际级学科竞赛奖项，获创客北京优胜奖、中关村人才创客大赛最具创新创业奖、北京·亦庄创新创业大赛优胜奖、复旦大学青年IT同学会创新创业一等奖等十余项创新创业奖项。

报告题目：从AIGC标识到可信内容生态

报告摘要：生成式人工智能正在重塑数字内容生产范式，也催生了新的安全挑战。一方面，AIGC生态日益丰富，生成模型、生成方式与应用场景持续扩展，海量生成内容不断涌现，传统单一模型或特定场景下的标识与溯源方法已难以适应复杂多样的生成环境；另一方面，内容生成、编辑与重构手段不断演进，针对AIGC标识的攻击由传统信号处理逐步升级为局部智能编辑、整体重生成和语义级修改等复杂形式，安全需求从“能否溯源”进一步拓展至“来源是否可信、内容是否完整、篡改能否定位甚至恢复”等。数字水印作为一种主动安全技术，正从标识溯源走向支撑可信内容生态的内生安全底座。本报告将梳理数字水印从传统后置、生成式内生到通用潜空间的发展历程，探讨通用生成式潜空间水印的核心问题与关键技术，并展望AIGC时代从内容标识走向可信内容生态的技术路径。


戴昀书

中山大学网络空间安全学院博士研究生，师从黄方军教授。研究方向为AI生成内容安全，累计发表IEEE TPAMI、CVPR、ICML、ICCV等国内外期刊与会议论文20余篇。入选中国科协青年科技人才培养工程博士生专项计划，主持国家自然科学基金青年学生基础研究项目（博士生项目）。

报告题目：生成式人工智能图像模型安全

报告摘要：生成式人工智能的快速发展使高逼真人脸与图像内容的生成门槛大幅降低，随之而来的伪造滥用、身份盗用与责任界定难题，已成为人工智能安全领域的核心挑战。本报告围绕生成式图像模型的安全性展开，汇报课题组近年来在相关方向上的探索与思考。报告将从生成内容的可辨识度出发，讨论在生成技术不断迭代、生成方式日益多样的背景下，如何提升安全机制对未知场景的适应能力；进而从被动应对转向前置保护，探讨在内容被恶意生成或篡改之前建立防护的可行路径，以及此类防护在真实环境下所面临的鲁棒性与可用性权衡；最后聚焦生成内容的来源确认与责任界定，分析如何在不损害正常使用的前提下，为生成结果与生成模型建立可信、可验证的关联。上述工作共同构成“检测—防御—溯源”相互衔接的整体思路，为生成式人工智能的可信可控应用提供支撑。



胡欣珏

南京信息工程大学计算机科学与技术专业博士生，师从付章杰教授，研究方向为图像隐写、图像隐写分析及伪造图像取证。近年来在 TDSC、TIFS 等期刊和会议上发表论文 4 篇，授权国家发明专利 3 项，国际发明专利 1 项。现担任 TCSVT、ACM MM 等期刊和会议审稿人。

报告题目：高分辨率图像隐写分析

报告摘要：随着社交媒体与移动互联网的发展，高分辨率图像逐渐成为日常信息传播的重要载体，也给隐蔽信息检测带来了新的挑战。在平方根定律约束的安全嵌入条件下，图像分辨率提升会进一步稀释隐写噪声，使具有显著判别作用的强决策区域更加稀少，同时大量弱决策区域容易对检测产生干扰。本次汇报围绕“高分辨率图像隐写分析”展开，主要介绍该场景下隐写噪声弱化、有效判别区域稀疏以及局部与全局信息协同不足等关键问题，并分享一种从决策区域“发现”到“利用”的两阶段检测思路。



郑裕丰

合肥工业大学计算机与信息学院博士研究生，导师为王垚飞副教授。研究方向为生成式隐写与隐蔽通信，目前聚焦于生成式图像隐写技术，同时参与课题组内AIGC防伪、检测与内容溯源的合作研究。部分成果已发表于ICML、ACM MM学术会议。

报告题目：面向扩散模型的免训练高容量隐写方案

报告摘要：扩散模型在图像生成领域的快速发展，为隐写技术提供了新的载体与实现空间。然而，在实际应用中，嵌入扩散模型生成过程中的信息会在多步去噪及 VAE 解码过程中逐渐衰减，从而影响接收端的信息恢复效果，使得隐写容量与提取准确率难以兼顾。针对这一问题，我们提出了一种无需额外训练且具备良好安全性的隐写框架，以有效缓解隐写容量与提取准确率之间的权衡。本报告将重点介绍该方案的核心观察、方法设计与实现机制，并进一步探讨相关技术的发展趋势。

■ 点评专家



黄欣沂

南京航空航天大学，教授/博导，长期从事数据与系统安全工作，主持国家自然科学基金青A、青B、重点等项目；在 CRYPTO、ASIACRYPT、ACM CCS、IEEE S&P、IEEE TDSC、IEEE TIFS 等重要学术期刊与会议上发表论文260余篇；获教育部自然科学奖一等奖、二等奖，中国密码学会密码创新奖一等奖；获包括欧洲计算机安全学术会议、Trust-Com等国际学术会议等最佳论文奖5项；担任80多个国际学术会议程序委员会委员，国际期刊IEEE TDSC、《中国科学：信息科学》（中英文版）等6个期刊编委等。



纪守领

纪守领，浙江大学计算机科学与技术学院求是特聘教授/长聘教授，国家级领军人才，可信人工智能研究中心主任，获佐治亚理工学院电子与计算机工程博士学位、佐治亚州立大学计算机科学博士学位。主要研究方向为人工智能安全、软件与系统安全。主持国家重点研发计划项目、国家自然科学基金联合重点项目等多项，发表CCF A类论文100余篇。获IEEE S&P 2025杰出论文奖、ACM CCS 2021最佳论文奖等10项最佳/优秀论文奖，全球黑客大会Pwntips Awards最具创新研究奖提名等。



李 琦

李琦，国家级人才，互联网体系结构全国重点实验室副主任，清华大学长聘副教授、博士生导师。长期从事互联网安全、物联网安全、人工智能安全等方向的研究，主持包括自然科学基金重点基金和重点研发等多项国家级项目和课题，获得包括浙江省科技进步一等奖、中国电子学会科技进步一等奖和中国计算机学会技术发明一等奖等多项科技奖励，以及USENIX互联网防御奖、USENIX Security杰出论文奖等多项顶会论文奖。



汪 定

汪定，南开大学二级教授，计算机学院、密码与网络空间安全学院副院长，教育部重大人才工程特聘教授，主要研究口令安全与抗量子密码协议。主持国家自然科学基金、科技部及华为、蚂蚁等科研项目20余项，在IEEE S&P、ACM CCS等发表论文100余篇，被引8000余次，H-index 40。研究成果推动美国身份认证标准NIST SP800-63修订。现任中国密码学会、中国网络空间安全学会理事及IEEE TIFS等期刊编委，获教育部自然科学奖一等奖、二等奖和中国密码学会密码创新奖一等奖。

抗量子密码迁移的“道法术器”论坛

承办单位：中国科学院软件研究所、中国科学院信息工程研究所

🕒 9月20日下午

📍 天鹅湖大酒店一楼贵宾厅

论坛主席： 张 阳（中国科学院软件研究所） 刘玉岭（中国科学院信息工程研究所）

召 集 人： 刘玉岭（中国科学院信息工程研究所）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	SHA-3算法的高速实现及其在后量子签名算法中的应用 何德彪（武汉大学）
14:40-15:10	基于格的高效抗量子密码算法设计 张 江（密码科学技术全国重点实验室）
15:10-15:40	量子计算的并行计算原理与实验进展 高 飞（北京邮电大学）
15:40-16:00	休 息
16:00-16:30	基于异构认证加密的敏捷抗量子密码迁移 路献辉（中国科学院信息工程研究所）
16:30-17:00	Diffie-Hellman 密钥交换算法抗量子迁移的挑战和思考 陈 隆（中国科学院软件研究所）
17:00-17:30	QKD与PQC融合的工程化探索与产业化实践 栗 栗（中移量子科技有限公司、中国移动研究院安全所）
17:30-18:00	自由交流讨论 主持人：咸鹤群（青岛大学）

论坛主席



张阳

中国科学院软件研究所正高级工程师、硕士生导师。中国科学院软件研究所、中国科学院大学双聘导师，中关村实验室双聘研究员。《Security and Safety》杂志副主编，某部专家委员会成员。曾任国家信息安全共性技术工程研究中心（主任负责制）主任助理。主要研究领域为软件与系统安全、信息物理融合系统安全。先后多次主持国家重点研发计划、国家自然科学基金、国家“863计划”重大项目子课题等国家级信息安全领域科研项目。在International Conference on Software Engineering(ICSE)、International Symposium on Research in Attacks, Intrusions and Defenses(RAID)、Computers & Security、《软件学报》、《计算机学报》等国内外一流学术期刊和会议上发表过40余篇学术论文，拥有专利2项，软件著作权5项，参与编写国家及行业标准近10项。获得中国电子学会科学技术奖一等奖1项、北京市科技进步一等奖2项、北京市科技进步三等奖1项。



刘玉岭

中国科学院信息工程研究所正高级工程师、博士生导师，中国科学院大学岗位教授，国家重点研发计划项目首席科学家，首都高端领军人才。长期从事网络安全态势感知、测评与认证技术研究，主持承担了装备预研、国家重点研发、公安部、海关总署、中科院项目20余项；相关技术成果已在多个国家部委机关以及多次国家重大活动网络安全保障中应用，多次获得网络安全主管机构书面感谢和表扬；发表学术论文50余篇，授权国家发明专利20项，制定安全相关的国家标准20余项，获得省部级一等奖3项，部委科研成果二等奖3项，科技冬奥先进个人。

召集人



刘玉岭

中国科学院信息工程研究所正高级工程师、博士生导师，中国科学院大学岗位教授，国家重点研发计划项目首席科学家，首都高端领军人才。长期从事网络安全态势感知、测评与认证技术研究，主持承担了装备预研、国家重点研发、公安部、海关总署、中科院项目20余项；相关技术成果已在多个国家部委机关以及多次国家重大活动网络安全保障中应用，多次获得网络安全主管机构书面感谢和表扬；发表学术论文50余篇，授权国家发明专利20项，制定安全相关的国家标准20余项，获得省部级一等奖3项，部委科研成果二等奖3项，科技冬奥先进个人。

报告嘉宾



何德彪

武汉大学计算机学院教授、院长、国家杰出青年科学基金获得者。主要从事密码学、人工智能安全、区块链技术等方面的研究工作。主持国家自然科学基金、科技部重点研发计划等国家级科研项目。近5年，在IEEE TIFS, IEEE TDSC, ASIACRYPT、PKC、CHES等国际知名期刊/会议上发表论文40余篇，谷歌学术引用总次数超过2万次。获教育部自然科学奖一等奖、湖北省自然科学奖二等奖等省部级奖励4项。担任SCIENCE CHINA Information Sciences、密码学报等多个期刊编委。

报告题目：SHA-3算法的高速实现及其在后量子签名算法中的应用

报告摘要：SHA-3函数广泛应用于后量子密码算法，然而其计算开销往往成为制约算法性能的主要瓶颈。许多后量子密码算法在设计时已充分考虑利用现代处理器的SIMD指令并行加速SHA-3运算，但在部分方案中，例如NIST标准签名方案ML-DSA，SHA-3运算仍可占据算法总运行时间的80%。本文提出了一种新的并行SHA-3算法优化思路，并提出了两种并行SHA-3计算模式：混合SHA-3模式（Hybrid SHA-3 mode）和同步吸收/挤压模式（Synchronous Squeezing and Absorbing mode）。这两种计算模式能够显著提高并行SHA-3运算调度的灵活性和资源利用效率。同时，我们提出了将上述计算模式高效集成到FSwA签名的方法。该方法可利用当前轮次的挑战值计算过程预计算后续轮次中多项式向量 y 的随机序列，从而进一步提升算法实现效率。我们将上述技术应用于ML-DSA和HAETA算法。与现有实现相比，ML-DSA算法的AVX2实现性能提升10.83%-19.92%，验签性能提升14.72%-18.05%。



张江

博士，现任密码科学技术全国重点实验室研究员/博导、副主任，国家青年科学基金项目（A类）、国家优秀青年科学基金项目获得者。主要从事密码学基础理论和密码算法协议设计研究，包括抗量子公钥密码、全同态加密等方向，主持国家网络安全国家科技重大专项、国家重点研发计划等项目，获省部级自然科学/科技进步奖一等奖3项、中国密码学会密码创新奖一等奖、全国密码算法设计竞赛一等奖2项等学术奖励。

报告题目：基于NTRU格的紧致抗量子公钥加密算法设计

报告摘要：NTRU公钥加密体制是最早的实用格公钥加密方案之一，也是IEEE 1363和ANSI X9.98的公钥加密标准。本报告将首先介绍原始NTRU加密体制及其优缺点。在此基础上，给出一种基于NTRU格的紧致抗量子公钥加密算法NEV，并详细介绍其主要设计考虑。NEV具有极高的通信和计算性能，在128比特安全强度下，NEV的公钥/密文仅约600字节，单核单线程软件实现加解密性能超过10万次/秒。与同等安全强度下的美国抗量子密码标准ML-KEM（Kyber）比较，NEV通信带宽远小于ML-KEM，同时计算性能超过ML-KEM的2.8倍。



高 飞

北京邮电大学教授，国家级高层次人才，Vebleo Fellow，中国密码学会理事会常务理事、党委委员，中国网络空间安全学会学术工作委员会委员，中国计算机学会量子计算专业委员会常务委员。主要从事量子密码、量子计算和网络安全等方向的研究，已在PRL/PRA、TKDE、TC、JSAC等重要期刊发表论文150余篇，SCI总引用7000余次，H因子52。近五年连续入选斯坦福大学“全球前2%科学家榜单”、爱思唯尔中国高被引学者。

报告题目：量子计算的并行计算原理与实验进展

报告摘要：量子状态的叠加性赋予微观世界很多新奇特性，同时也给量子信息处理带来了超凡脱俗的特殊能力。本报告将介绍量子态、量子操作和量子测量等量子信息基础知识，并在此基础上从求函数整体性质的角度介绍Deutsch、BV、Simon和Shor算法的基本思想，旨在帮助大家理解量子计算具有强大计算能力的基本原理。此外，本报告也将介绍量子计算的发展现状和实验进展。



路献辉

中科院信息工程研究所研究员，网络空间安全防御全国重点实验室副主任，密码理论与技术研究室主任，ISO/IEC国际标准化组织技术专家，中国密码学会理事，全国量子计算与测量标准化委员会委员。主要研究兴趣包括可证明安全理论、抗量子密码、全同态密码、物理层密码。带领团队设计了抗量子密码算法LAC参加美国抗量子密码标准征集，是进入第二轮的26个候选算法之一。提出了基于零因子多项式的双层编码框架，基于此设计了目前尺寸紧致的基于NTRU的公钥加密算法。针对全同态加密的计算瓶颈，设计了门自举、电路自举、AES转密、同态汇编指令集等同态加密核心算法和应用方案，研究成果发表于欧密2024、TCHES2024、CCS2025。担任国际标准化组织ISO/IEC密码工作组抗量子密码和全同态密码领域多项标准化项目负责人。

报告题目：基于异构认证加密的敏捷抗量子密码迁移

报告摘要：抗量子密码算法呈现出尺寸大、类型多、功能完备度不一等特点，给TLS、IPSEC、SSH等安全协议的抗量子密码迁移带来挑战。本报告介绍基于异构认证密钥交换的抗量子密码协议敏捷迁移技术，可实现持有不同密码原语证书的客户和服务端实现异步平滑迁移。除了敏捷迁移，异构认证机制还可以实现终端和服务端计算传输负载的差异化非均衡配置，提升系统整体性能。



陈 隆

博士，中国科学院软件研究所可信计算与信息保障实验室副研究员。他于2018年在中科院软件所获博士学位，2018年至2021年在新泽西理工学院计算机学院担任博士后研究员，2021年7月入职中国科学院软件研究所，入选中国科学院引才计划B类并获得择优支持。陈隆博士长期从事理论与应用密码相关领域研究，研究方向为抗量子密码和安全协议。陈隆博士正在担任中国密码学会算法专业委员会委员兼秘书，作为项目负责人承担国家重点研发计划青年科学家项目和国家密码科学基金重点项目，并承担国家自然科学基金重点项目课题，入选中科院基础研究青年团队。他在IEEE S&P、USENIX Security、CCS、CRYPTO、NDSS、ASIACRYPT、TCC、PKC、TIFS等顶级国际密码与信息安全会议与期刊上发表了十余篇论文，担任《密码学报》青年编委，并曾担任 USENIX Security、PKC 等国际会议程序委员会委员。

报告题目：Diffie-Hellman 密钥交换算法抗量子迁移的挑战和思考

报告摘要：Diffie-Hellman密钥交换算法是应用最为广泛的公钥密码算法之一。然而，现有基于格的密钥封装机制在功能和安全性方面与Diffie-Hellman协议并不完美匹配。在量子时代到来之前，如何实现Diffie-Hellman密钥交换算法的平滑无缝迁移，依然是一个亟需解决的难题。本报告从基础功能、证书兼容性以及性能三个维度，分析了Diffie-Hellman迁移可能遭遇的问题与挑战，并探讨相应的解决方案。



粟 粟

博士，中移量子科技有限公司副总工、中国移动研究院安全所所长，教授级高级工程师，国家重大专项专家，兼任移动互联网安全国家工程中心副主任。主要研究方向包括网络安全、人工智能安全。先后主持了国家重大专项、重点研发计划、国家242信息安全专项等重大项目4项，主导了移动通信网5G/6G内生安全、虚拟化安全等关键技术与国际标准化相关工作，实现了大规模应用。获得授权发明专利110余项、发表科技论文30余篇、主持撰写《5G安全技术标准》专著；曾获得工信部科技成果鉴定证书1项，省部级科技奖11项。

报告题目：QKD与PQC融合的工程化探索与产业化实践

报告摘要：报告从产业视角出发，聚焦量子密钥分发（QKD）从基础研发走向规模商用的核心命题，结合中国移动工程探索与产业实践成果，系统阐述量子通信基础设施建设、商业化推广与标准化体系建设的落地进展。立足未来产业布局，明确下一阶段PQC算法将要全面落地之际，QKD+PQC融合的技术攻关、工程短板补齐的重点方向，面向学术界与产业界提出产学研协同创新、联合攻克瓶颈、共建标准与人才体系的合作倡议，共同推动QKD+PQC从“可用”走向“好用”。

智能无人系统安全论坛

承办单位：西安电子科技大学

🕒 9月20日下午

📍 天鹅湖大酒店三楼11号会议室

论坛主席：马建峰（西安电子科技大学）

召 集 人：习 宁（西安电子科技大学）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:50	智能无人系统综合安全防护关键技术 习 宁（西安电子科技大学）
14:50-15:30	智能物联终端设备在新形势下的安全挑战与防御机制 王 滨（海康威视）
15:30-16:10	面向无人系统的跨域攻击和智能安防护技术 孙彦斌（广州大学）
16:10-16:30	休 息
16:30-17:10	面向大语言模型安全对齐的三角色自博弈强化学习框架 杨 全（北京大学）
17:10-17:50	智能无人系统安全标准研究进展 范科峰（中国电子技术标准化研究院）
17:50-18:10	自由交流讨论

论坛主席



马建峰

我国无线网络安全领域的学术带头人，提出了“云-网-端”协同安全一无线网络安全的新范式，以第一完成人获国家技术发明二等奖2项、省部级科技成果一等奖2项；发表高水平论文100余篇，连续入选科睿唯安全球高被引学者；出版英文专著（斯普林格）、国家网络安全优秀教材；授权发明专利72件，参与制定ITU安全标准3项、国家标准1项；获评中央网信办网络安全优秀人才、带领团队获首届国家网络安全先进集体；培养了以教育部长江学者、国家杰青项目获得者、国防卓青项目获得者为代表的网络安全优秀人才百余名，获国家教学成果奖二等奖2项。

召集人



习 宁

西安电子科技大学教授，博导，网络与信息安全教育部创新团队、陕西省网络与系统安全重点实验室技术骨干。先后主持及参与了国家自然科学基金重大研究计划，国家重点研发计划，装备预研专用技术项目、国家自然科学基金重点项目等项目的研究。在IEEE会刊，Science China等国内外期刊和会议上发表论文50余篇，授权技术发明专利30余项，研发了服务安全组合中间件、嵌入式可组合安全验证工具、嵌入式多级安全防护系统，应用于智慧医疗系统、国产首个机载操作系统等，获得年陕西省科学技术奖一等奖2项、中国电子学会科学技术奖一等奖1项、陕西高等学校科学技术研究优秀成果奖1项。

报告嘉宾



习 宁

博士，西安电子科技大学网络与信息安全学院教授、博导，国家自然科学基金重大项目负责人，陕西省网络与系统安全重点实验室无人系统安全方向技术负责人，西安市智能系统安全实验室副主任。主要从事数据安全、系统安全方面的研究。先后主持了国家自然科学基金重大项目，国家重点研发计划子课题，陕西省自然科学基金基础研究计划，航空基金项目等，发表论文70余篇，已授权中国技术发明专利40项，研发了智慧城市服务安全协同中间件，面向国产机载操作系统的多级安全防护系统，无人机综合安全防护系统等，获得中国电子学会技术发明一等奖1项，陕西省科学技术奖一等奖2项。相关理论与技术在智慧医疗、智能安防、智慧小区、工业互联网等领域的相关产品得到推广应用，并被纳入国家标准《GB/T 31167-2014，信息安全技术 云计算服务安全指南》，标准草案《智慧城市城市智能服务构建指南》中。

报告题目：智能无人系统综合安全防护关键技术

报告摘要：智能无人系统广泛用于社会、经济、民生诸多领域，多无人系统协同服务已成为战略性关键技术，智能化、异构化、多任务已成为无人系统应用趋势，形成了“平台-网络-任务”三级联动的典型应用框架。智能无人系统攻击种类繁多、攻击方式复杂，严重影响无人系统的安全性、可靠性，急需构建“平台-网络-任务”一体化的无人系统综合安全防护体系。本报告重点介绍西电项目组近期在基于可信计算的平台综合安全防护、轻量动态安全组网、智能异常检测与攻击识别等技术上的新进展，以及自研无人综合安全防护系统在支撑低空经济、海洋探测方面的应用情况。



王 滨

国家级高层次人才，享受政府特殊津贴，博士生导师，浙江全省智能物联网与数据安全重点实验室主任、视频感知国家新一代人工智能开放创新平台副主任、视觉感知技术研发与应用国家地方联合工程研究中心副主任、首届国家卓越工程师团队奖核心成员。长期从事物联网安全、密码学、数据安全等研究。近年来主持国家科技重大专项、国家自然科学基金重大项目重点支持项目、国家重点研发计划、国家工业互联网安全专项等10余个国家级项目，作为主要完成人获省部级科技奖励近20项。

报告题目：智能物联终端设备在新形势下的安全挑战与防御机制

报告摘要：物联网从“万物互联”走向“万物智联”，全面走向智能化，这使得边缘终端能够采集到更加精准的数据，具备了更加强大的功能，全球各国相继颁布了物联网终端安全标准和法规，本报告将从实践的角度探讨在新的形势下终端安全面临的新的安全挑战，以及如何有效的解决这些安全问题，采取何种对应的安全机制全面提升物联网的安全能力。


孙彦斌

博士，教授，广州大学网安学院副院长，国家级创新人才、广东省特支计划青年拔尖人才。广东省工业控制系统攻防对抗重点实验室常务副主任。深耕网络攻防、工控安全方向，在国内外重要学术期刊、会议上发表论文100余篇。主持国家科技重大专项课题、国家重点研发课题、国自然联合基金重点等项目20余项。相关成果获广东省哲学社会科学优秀成果奖二等奖、中国电子学会科技进步一等奖2项，多次参加国家重保任务并获致函感谢。

报告题目：面向无人系统的跨域攻击和智能防护技术

报告摘要：无人系统正加速由单机遥控向开放互联、自主智能及集群协同方向演进，攻击者所具备的网络-物理多域渗透、隐蔽触发与后果级联放大能力，构成了该领域的重大安全威胁，可能导致任务执行失败、装备失控乃至人员与环境受损。本报告首先分析无人系统面临的共性安全特征与关键防护难点，系统梳理典型攻击技术与现有防护方法，并审视跨域协同防御中的瓶颈问题。在此基础上，提出基于功能安全优先的跨域新型智能防御体系架构，研究对抗增强型智能防护技术，形成涵盖“事前多重抑制—事中跨域协同检测—事后全局动态处置”的闭环防御能力。本研究旨在为无人系统的安全可靠运行提供关键技术支撑，推动防御体系向智能化、体系化方向演进。


杨 全

北京大学PKU-DS实验室负责人，北京大学计算机学院研究员，北京大学人工智能研究院研究员，教育部青年长江学者，北大-小米大模型联合实验室主任，北大-360大模型联合实验室主任，北大-开物金如意人工智能联合实验室主任，北大-腾讯大模型联合实验室主任，北大-九章云极大模型联合实验室主任，北大临港大模型研发中心主任。主要研究方向为大语言模型关键技术与概率数据结构。实验室云集了众多优秀硕博学生100人，包括IMO、IOI、IBO、IPHO、NOI、CMO、CNBO、CChO 金牌。近年发表一作/通信作者CCF A类论文100余篇。提出首个复数大模型iFairy，实现了2比特的量化精度超过了量化前16比特实数大模型。提出了复域变换，首次实现实数-复数模型理论严格等价转换。研发的FairyR1模型以5%参数超越Deepseek满血版R1的数学和代码性能，已在北大校园网提供大模型服务。获得中国电子学会技术进步一等奖（排名第二）、6次获得华为火花奖（全国最多），1次奥林帕斯先锋奖。研究成果获得了国家自然科学基金委员会网站4次报道、教育部科技发展中心1次。

报告题目：面向大语言模型安全对齐的三角角色自博弈强化学习框架

报告摘要：近年来，大语言模型的安全风险日益突出，如何减少有毒、有害内容的生成已成为亟待解决的问题。当前主流的安全对齐方法通常包含攻击者、防御者和评估者三类角色，分别负责生成对抗性提示、安全防护及回复质量评估。本文提出三角角色自博弈强化学习框架 TriPlay-RL，通过构建统一的闭环学习机制，使三个角色在近乎零人工标注的条件下进行迭代训练与协同进化。实验结果表明，攻击者在保持较高输出多样性的同时，对抗攻击有效性提升了20%~50%；防御者的安全性能提升了10%~30%，且未损害模型的通用推理能力；评估者则在多轮迭代中持续增强细粒度判断能力，能够准确区分不安全回复、简单拒答和具有实际帮助的安全指导。总体而言，TriPlay-RL为大语言模型安全对齐提供了一种高效、可扩展且低标注成本的新范式，实现了攻击、防御与评估能力在统一强化学习闭环中的持续协同演化。

**范科峰**

正高级工程师，博士，中国电子技术标准化研究院副院长，兼任电子信息产品标准化国家工程研究中心主任、人工智能标准与应用工业和信息化部重点实验室主任、中国网络安全产业联盟秘书长、工业和信息化部电子信息科学技术委员会常委、工业和信息化部人工智能标准化工作组副主任委员等职务。享受国务院政府特殊津贴，研究方向为网络安全、人工智能等，主持完成国家级科研项目10余项，制定网络安全、人工智能等领域国家标准20余项，获省部级以上重要科技奖励10余项。

报告题目：智能无人系统安全标准研究进展

报告摘要：本报告聚焦智能无人系统的安全治理前沿，系统阐释了我国在该领域标准化建设的最新进展与战略思考。报告指出，随着人工智能与无人技术的深度融合，构建全生命周期的安全标准体系已成为护航产业高质量发展的必由之路。报告重点分享了两大核心维度的标准化突破：在物理空间，通过推进“一机一码”等强制性国家标准，为低空经济等无人系统筑牢安全闭环监管底座；在数字空间，依托《人工智能 智能体互联》系列国家标准与伦理安全指引，为智能体赋予“数字身份证”与通用交互规则，有效破解跨域协同中的“信息孤岛”与信任危机。报告强调，标准化不仅是技术的规范，更是产业重构的基石。未来将坚持“鼓励创新与规范发展并重”，以敏捷标准化安排与开源生态共建，提前划定安全合规边界，为智能无人系统产业的安全、可信、向善发展提供坚实支撑。

网安期刊创新发展论坛

承办单位：中国网络空间安全学会出版工作委员会、中国航天系统科学与工程研究院

🕒 9月20日下午

📍 天鹅湖大酒店三楼8号会议室

论坛主席：刘建伟（北京航空航天大学）

召集人：谢 雪（中国航天系统科学与工程研究院）

时 间	内 容
14:00-14:05	主持开场
14:05-14:10	致 辞
14:10-14:40	“借船出海”Vs“造船远航”——以中国激光杂志社为例 杨 蕾（中国激光杂志社）
14:40-15:10	科技活动助力科技期刊高质量发展 赵跃峰（长安大学学术期刊管理中心）
15:10-15:40	从编辑分工到流程重构，英文科技期刊的高效管理模式 尹 欢（南京农业大学英文期刊编辑部）
15:40-16:10	探索学术期刊影响力提升的策略——以《网络空间安全科学学报》为例 谢 雪（中国航天系统科学与工程研究院）
16:10-17:40	主题交流 1.科技期刊的创新发展路径探索； 2.AI赋能科技期刊高质量发展； 3.网络安全相关领域科技期刊集群建设。

论坛主席



刘建伟

刘建伟，教授，博士生导师，北京航空航天大学网络空间安全学院创院院长，享受国务院政府特殊津贴专家。现任国务院学位委员会学科评议组成员，教育部网络空间安全专业教指委副主任，全国密码专业学位研究生教指委委员，中国网络空间安全学会常务理事，中国密码学会常务理事，《网络空间安全科学学报》常务副主编，IEEE高级会员。获国家级教学名师，国家网络安全优秀教师奖，宝钢优秀教师特等奖，北京市教学名师奖，北京市优秀教师等国家级和省部级教学荣誉和奖励20余项，获国家技术发明一等奖，国防技术发明一等奖，中国指挥与控制学会一等奖等国家级和省部级奖励10余项。

召集人



谢 雪

现任中国航天系统科学与工程研究院第六研究所副所长、《网络空间安全科学学报》负责人，兼任中国网络空间安全学会副秘书长、中国期刊协会常务理事，长期深耕网络安全研究、科技期刊管理、科普教育等工作。

报告嘉宾



杨 蕾

杨蕾，中国激光杂志社总经理，二级编审。享受国务院政府特殊津贴，曾荣获中宣部“四个一批”人才，中国出版政府奖“优秀出版人物”、全国新闻出版行业领军人物，上海市出版人奖，上海杰出出版人物奖等奖项。

创建了中国激光杂志社有限公司，牵头成立了“中国光学期刊联盟”，创立了中国光学期刊网中文版，researching 英文平台，实现了科技出版全流程的数字化。

创办并运营了12种国际和国内知名期刊，其中多种被国内外知名数据库收录，并位居同行前列。担任中国科技期刊协会常务理事，上海科技期刊学会常务副理事长等社团工作。

报告题目：“借船出海”Vs“造船远航”——以中国激光杂志社为例

报告摘要：在当前全国上下高度重视科技期刊的态势下，国内各学科各行业的新生了很多的新刊。办刊理念、路径、方法的确立，人才队伍搭建，影响力传播等方面，都在进行积极探索。本报告以中国激光杂志社为例，重点汇报在“借船出海”Vs“造船远航”方面的思考与实践。


赵跃峰

赵跃峰，编审，现任长安大学学术期刊管理中心副主任，陕西高等学校学报研究会副理事长兼副秘书长，为陕西省出版高级职称评审专家、陕西省期刊审读专家。主要从事编辑出版学、出版实践等研究，所编辑出版的期刊被国内外重要数据库收录，荣获中国卓越行动计划一二期资助，多次荣获百种中国杰出学术期刊、中国百强报刊、中国高校百杰学术期刊等称号；本人担任中国科技期刊卓越行动计划项目负责人、主持和参与省部级项目6项，发表论文20余篇，荣获中国科协优秀编辑，全国高校科技期刊研究会优秀编辑，陕西省高校学报优秀编辑等奖项10余项。

报告题目：科技活动助力科技期刊高质量发展

报告摘要：立足长安大学的办刊实践，从汇聚学术资源、搭建高端学术平台、主办前沿论坛和学术年会等方面，阐述以学术活动提升科技期刊影响力的思路 and 措施，不断提高期刊影响力。


尹欢

尹欢，南京农业大学英文期刊编辑部主任、副编审、硕导，江苏省新闻出版优青。现任江苏科技期刊学会副理事长、中国期刊协会农业期刊分会理事、中国高校科技期刊研究会青委会副主任委员。研究方向：期刊创办与高效出版，编辑团队管理和协同发展。作为主要人员，参与中国科技期刊卓越行动计划领军期刊、第六届中国出版政府奖期刊奖 Horticulture Research，梯队期刊 Plant Phenomics 和高起点新刊 BioDesign Research 的创刊、运营与管理。主持和参与中国科技期刊卓越行动计划选育高水平办刊人才子项目 - 青年人才支持项目等课题 21 项；获评中国科技期刊卓越行动计划选育高水平办刊人才子项目 - 优秀编辑等

报告题目：从编辑分工到流程重构，英文科技期刊的高效管理模式

报告摘要：聚焦英文科技期刊在学术出版日益复杂化背景下的管理模式转型，提出从编辑分工向流程重构驱动的效能跃迁路径。通过重塑初审、送审、编校、传播等环节的职责边界，引入动态任务分配与数据化进度监控，实现人力与时间成本的双降；同时强调技术工具对流程的嵌入式支撑，最终形成一套兼具标准化与柔性的高效管理方案，为提升期刊竞争力、缩短出版周期并保障学术质量提供可操作的实践参考。


谢雪

现任中国航天系统科学与工程研究院第六研究所副所长、《网络空间安全科学学报》负责人，兼任中国网络空间安全学会副秘书长、中国期刊协会常务理事，长期深耕网络安全研究、科技期刊管理、科普教育等工作。

报告题目：探索学术期刊影响力提升的策略——以《网络空间安全科学学报》为例

报告摘要：以《网络空间安全科学学报》为例，探讨学术期刊的发展路径。首先提炼了办好期刊要立足两个行业、运用三支队伍、提升三个价值的ATQ模型；其次分享了学报如何进行选题策划、如何进行编委会运营、如何提升学术质量和学术影响的实践案例；最后探讨了AI编辑加工智能体在学报编加工作中的应用实践，展示了自主开发AI智能化工具的可复制性和应用价值。

海报展示

论 文

Stop Starving or Stuffing Me: Boosting Firmware Fuzzing Efficiency with On-demand Input Delivery

申珊珊, 周威, 赵柯名, Peng Liu, Chung Hwan Kim, Le Guan

HoSig-Align I: 边缘侧网络攻击归因的同源块发现方法

姚爱婷, Shantanu Pal, 东承祖, 邵迪, 冯文英, 李若南, 顾钊铨

VFCionX: Bridging Large and Small Models for Vulnerability-Fixing Commit Identification

崔星, 吴敬征, 欧文祥, 罗天悦, 李志远, 凌祥

Catching License Changes Before They Cost: Leveraging LLMs to Analyze License Evolution in Open Source Software

崔星, 吴敬征, 罗天悦, 欧文祥, 凌祥

Mining Long Tail Bugs: Identifying Rare and Overlooked Issues in Code

梁文韬, 武延军, 凌祥, 罗天悦, 刘丁豪, 张皓添, 吴敬征

LongFlowFormer: 鲁棒长序列加密流量分类

Mingshuai Pan, Ning An, Zelin Cui, Bo Jiang, Zhigang Lu, Baoxu Liu, Junrong Liu, Cong Dong

ADGFUZZ: Assignment Dependency-Guided Fuzzing for Robotic Vehicles

王允成, 郑尧文, 刘圃卓, 方栋梁, 程嘉兴, 石丁焱, 孙利民

面向开放世界通用的测绘探针组织溯源系统

贺鑫, 董恩焕, 韩骥远, 王之梁, 张辉, 刘亮, 孙连祯, 张苏沛, 薛鹏飞, 宋光磊, 李晗, 权小文, 杨家海

EchoFuzz: Empowering Smart Contract Fuzzing with Large Language Models

Juanen Li, Peng Qian, Guanyan Li, Rui Wang, Peixin Wang

终极加密流量特征工程: 基于还原后应用数据单元长度的HTTPS加密流量分类

陈子涵, 程光, 钮丹丹, 赵玉宇, 周余阳, 蒋山青

理解并形式化验证AWS Cognito中的跨服务漏洞

陈贞, 靳泽, 龚乐, 陈可欣, 曾祥翼, 王昊强, 刘奇旭

区块链预言机协议安全性分析

翟翟, 张家硕, 高健博, 刘天皓, 张涛, 王健, 刘吉强

论 文

Cutting the Fuse: Actionable APT Attack Blocking in Provenance-based IDS

Weiheng Wu, Wei Qiao , Teng Li , Yebo Feng , Zhuo Ma , Jianfeng Ma, Yang Liu

ARTO: Efficient Execution Integrity Attestation for Real-Time Operation of Cyber-Physical Systems

赵瑞喆, 孙聪, 李宗臻, 王天天, 王运柏

One Bad Token Spoils the Barrel: Assessment, Detection, and Remediation of Glitch Tokens in Large Language Models

唐坤生, 齐培贵, 宋轶德, 周文柏, 黄智聪, 郭青, 张天威, 张卫明, 俞能海, 张杰

Attesting Model Lineage by Consisted Knowledge Evolution with Fine-Tuning Trajectory

商卓逸, 李佳森, 陈鹏桢, 刘延伟, 古晓艳, 王伟平

Your Classifier Can Do More: Towards Balancing the Gaps in Classification, Robustness, and Generation

姜凯超, 王鹤, 郝孝帅, 杨秀隆, 刘阿建, 储琪, 刁云峰, 洪日昌

HENet: A Heterogeneous Encoding Network for General and Robust Adversarial Example Generation

Jiawei Zhang, Hao Wang, Hao Wu, Bin Li, Xiangyang Luo, Bin Ma, Jinwei Wang

RDTSM:基于可信影子模型的联邦学习投毒攻击鲁棒防御

徐辰晨, 陈付龙, 黄大荣, 李洪潮, 洪鑫, 王涛春

MPAS: Breaking Sequential Constraints of Multi-Agent Communication Topologies via Individual-Epistemic Message Propagation

Jingxuan Yu, Ju Jia, Simeng Qin, Xiaojun Jia, Siqi Ma, Yihao Huang, Yali Yuan, Guang Cheng

DRGW: Learning Disentangled Representations for Robust Graph Watermarking

李佳森, 刘延伟, 商卓逸, 古晓艳, 王伟平

Confused Deputy Attack Against Model Context Protocol

李志远, 吴敬征, 彭宇毫, 罗天悦, 崔星, 凌祥

PVMark: 面向大语言模型水印的公开可验证机制

段皓铎, 向立瑶, 张鑫, 李葆春, 李波

Sentient: Detecting APTs Via Capturing Indirect Dependencies and Behavioral Logic

燕文浩, 安宁, 乔伟, 吴伟恒, 卢志刚, 姜波, 刘宝旭, 刘俊荣

Efficient, Robust, and Anti-Collusion Fingerprinting of Image Diffusion Models

Jianwei Fei, Yunshu Dai, Zhihua Xia, Xiaochun Cao, Jiantao Zhou, Alessandro Piva, Benedetta Tondi

论 文

定制化文生图生态中的插件投毒与病毒式传播

陈佳豪, 贺兴, 杨勇, 李鑫锋, 周纯毅, 李俊豪, 马哲, 杜天宇, 纪守领

PlugMark: A Plug-in Zero-Watermarking Framework for Diffusion Models

陈鹏桢, 刘延伟, 古晓艳, 商卓逸, 季向阳, 刘武, Enci Liu

Rotation-Invariant Spherical Watermarking via Third-Order SO(3) Representation Coupling

陈鹏桢, 刘延伟, 古晓艳, Antonios Argyriou, 刘武, 王伟平

SafeGuider: Robust and Practical Content Safety Control for Text-to-Image Models

齐培贵, 唐坤生, 周文柏, 张卫明, 俞能海, 张天威, 郭青, 张杰

Before Agents Speak: Pre-hoc Failure Risk Inference in Multi-Agent Systems

Shi Lin, Chenpei Wang, Peng Qian, Dezhang Kong, Minghao Li, Yufeng Li, Xun Wang

Rethinking Smart Contract PoC Generation: Agentic Exploit Witness Construction from Bug Findings

Peng Qian, Chengyu Wu, Shi Lin, Jiachi Chen, Dinghao Liu, Renjie Sun, Jun Shao, Xun Wang

Are LLM-Enhanced Graph Neural Networks Robust against Poisoning Attacks?

马宇航, 王杰, 闫峥

From Extraction to Deduction: Resolving Functional Misalignment in RAG via a Collaborative Critic-Reasoner Framework

陈宇飞, 王尧, 张海宾, 周华林, 顾涛

Fine-Grained Privacy Extraction from Retrieval-Augmented Generation Systems by Exploiting Knowledge Asymmetry

陈宇飞, 王尧, 张海宾, 顾涛

Awakening the Hydra: Stabilizing Multi-Concept Backdoor Injection in Text-to-Image Diffusion Models

Kai Wang, Jiale Zhang, Chengcheng Zhu, Chuang Ma, Songze Li

Manhuff: Adaptive Inverted Huffman Coding for High-Capacity Imperceptible Generative Linguistic Steganography

Yu Liu, Jun Ma, Xun Liu, Yanmei Kang, JiaJia Li

Bridging the Knowledge Gap in Federated Prompt Learning via Text-Guided Feature Augmentation

王一斌, 周玉灿, 李洪成, 管增浩, 古晓艳, 王伟平

基于时空统计量聚合的高效联邦类增量学习

管增浩, 祝国隽, 周玉灿, 刘武, 王伟平, 罗杰波, 古晓艳

论 文

Accurate, Secure, and Efficient Semi-Constrained Navigation with Multiple Spatial Restrictions

Meng Li, Yao Hu, Yan Qiao, Zijian Zhang, Liehuang Zhu, Mauro Conti

An Anonymous Multi-Authority Key-Policy Attribute-Based Encryption Scheme with Expressive Keyword Search

赵陈斌, 林晓聪, 陈晶, 尤玮婧, 崔杰, 花忠云

ES-CSFL: Compressed Sensing-based Efficient Privacy-preserving Federated Learning with Enhanced Security

陈律君, 肖迪, 张懋兰, 任昱, 范新春

面向去中心化世界的可信数据治理

Meng Li, Yifei Chen, Yan Qiao, Guixin Ye, Zijian Zhang, Liehuang Zhu, Mauro Conti

广域网下大规模隐私且可摊销的决策树评估

许伟, 朱辉, 郑艳东, 李晖, 边松, 冯登国

代数结构驱动的高效伪随机相关性生成与安全多方计算

李哲, 刘弘庆, 邢朝平, 姚逸洲, 袁晨

Exploiting Body-Coupled Leakage for Keystroke Inference on Smartphones

王尧, 何岸, 李晓龙, 张海宾, 顾涛

Hallucination Mitigation for Large Vision-Language Models with Hypothesis-Guided Spatial Verification

汤化宇, 姚远志

科 技 项 目

具身智能物理安全评测与协同优化关键技术研究

何召锋, 任民, 李秋瑜, 齐翊深

MG-TCANet: 面向恶意流量分类的多粒度行为表征学习

王泽涵, 岳鹏飞

低轨卫星通信无线信道可定义安全认证技术研究

鲁信金

城市智能系统可信机理与关键技术

刘吉强

云计算下数据安全理论及方法

李兴华, 苗银宾, 马鑫迪, 王运帷

一流学院成果

筑基·砺刃·鼎新：网络强国战略下拔尖网络安全人才培养体系构建与实践

东南大学

北京航空航天大学一流网络空间安全学院建设成效

北京航空航天大学

湖南大学网络空间安全学院 创新引领·安全报国

湖南大学

天津大学网络空间安全学院建设成效

天津大学

武汉大学网络空间安全学院建设成效

武汉大学

中国科学技术大学网络空间安全学院建设成效

中国科学技术大学

西北工业大学网络空间安全学院建设成效

西北工业大学

西安交通大学网络空间安全学院建设成效

西安交通大学

教学筑基·平台赋能·实战育人·服务国家——一流网络安全学院建设成果

北京邮电大学

暨南大学网络空间安全学院研究成果与育人成效

暨南大学

大会合作单位



长三角安全人工智能安徽省实验室
Anhui Laboratory for Safe Artificial Intelligence in the Yangze River Delta

HYGON
中 科 海 光



数字化及AI解决方案领导者

NIO 蔚来



长三角安全人工智能安徽省实验室简介

Laboratory Introduction

长三角安全人工智能安徽省实验室是由安徽省人民政府批准，依托安徽星盾智能科技有限公司（科大讯飞全资子公司）建设的高能级科创平台，是以国家战略需求为导向、行业应用需求为牵引，集聚全国人工智能安全领域相关的重点高校、科研院所及科技企业联建。实验室致力于突破安全人工智能关键技术，构建人工智能安全治理和防护能力，护航人工智能技术和产业发展。

实验室聚焦人工智能安全前沿领域原始创新、关键技术研发和产业化应用，着力打造前沿科技研发引领平台、科产融合开放平台、人才引育汇聚平台、高能级产业培育平台，建立健全开放式协同创新机制，积极探索有效风险防护和治理方法路径，推动人工智能产业安全化、安全产业化，确保人工智能发展安全、可靠、可控。

产品介绍

Product Introduction

星界—大模型内容安全治理解决方案

方案覆盖数据构建、模型训练、应用上线。以内生安全工具夯实基础，结合内容安全护栏实现提示词注入和违法违规内容防护，以风控运营服务优化安全能力，全流程通过安全评测能力把关，稳居行业第一梯队。



星驭—智能体应用安全解决方案

方案覆盖上线前的Skill安全检测与靶场评测、运行中实时防护和事后日志审计，可充分暴露安全风险，对关键风险进行防护和分级处置，保障智能体应用安全稳定运行并兼顾业务可用性。方案支持场景共建与集成落地。



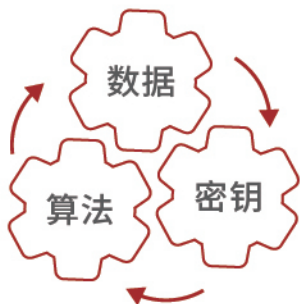
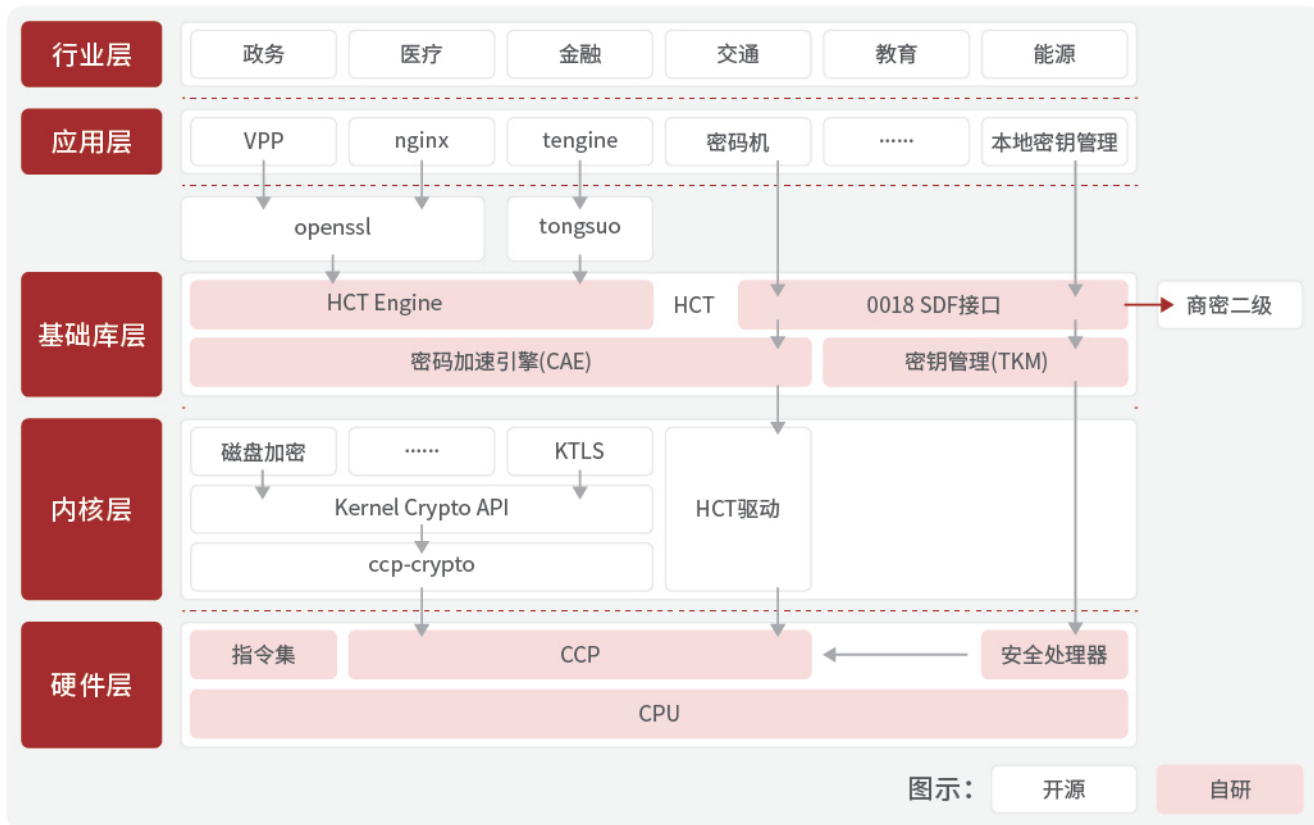
星鉴—AIGC 内容可信传播治理方案

方案提供显式水印和元数据水印能力保障内容合规，并进一步提供隐式数字水印；提供AI生成检测与有害内容识别能力，适用于票据审核、风控反诈、内容平台以及 AIGC 内容创作等多种场景，助力构建安全可信的内容传播环境。



海光芯片内生密码技术

商密引擎硬核守护，芯片级密码运算加速，密评合规无忧



密钥安全至关重要，密钥被窃取，数据安全无从谈起



密钥管理

- 合规强制要求：密评、商密产品认证等
- CPU内置密钥管理能力成为新的趋势
- 接口符合最新GM/T 0018-2023 《密码设备应用接口规范》

海光CPU，完整密钥生命周期管理，密钥“可用不可见”

生成

存储

使用

导入

导出

销毁



海光信息技术有限公司

国内唯一的同時拥有性能领先的CPU和GPGPU的IC设计公司

天融信科技集团

中国领先的网络安全与智算云解决方案提供商



AI安全解决方案引领者

提供从模型训练、部署到应用的全周期风险管控产品与解决方案，构建面向大模型的漏洞发现、安全评估、安全监测、安全防护的闭环体系。

全场景安全解决方案领军者



中国网络安全领军者，100余类、1000余款产品，覆盖基础网络、工业互联网、物联网和车联网等全业务场景，服务全行业客户数智化转型。



智算云新兴市场先行者

构建通算、智算、存储、网络、安全、智能一体化底层基础设施，提供从通算到智算的全栈云能力，助力AI应用加速落地。

地址：北京市海淀区西北旺东路10号院西区11号楼东侧天融信科技集团

电话：400-777-0777



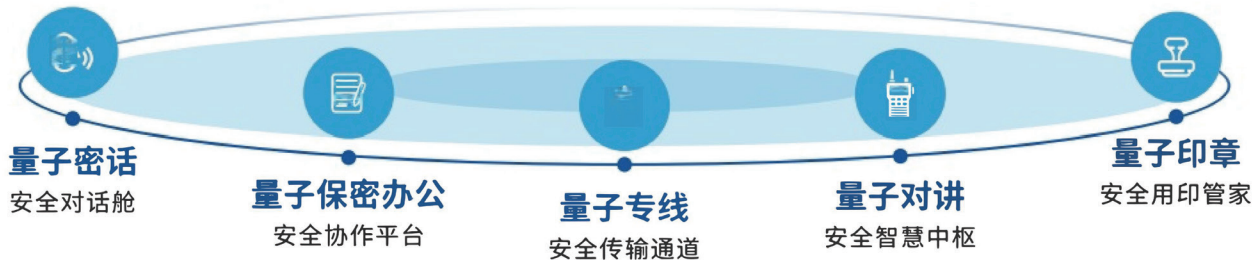
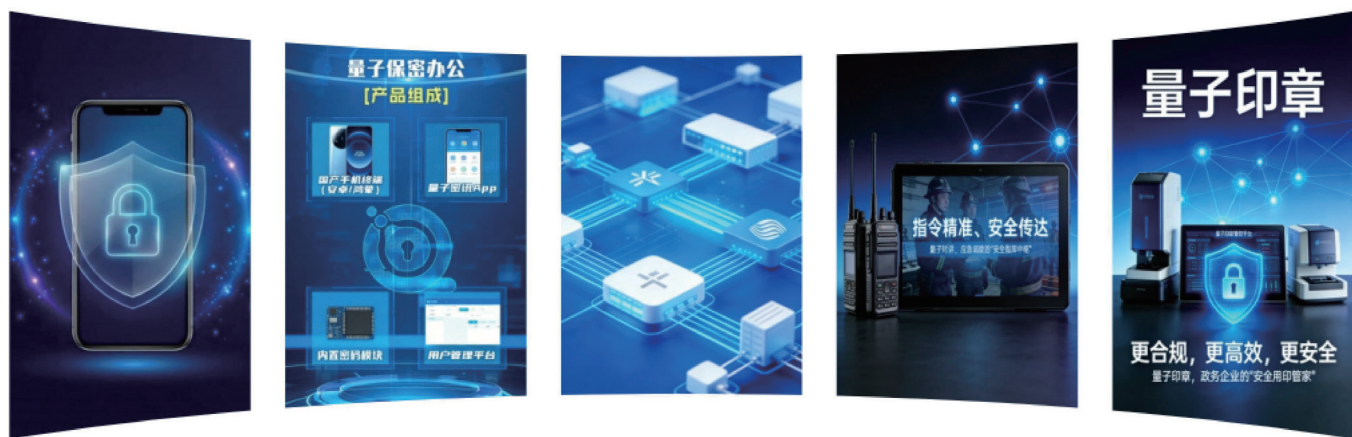
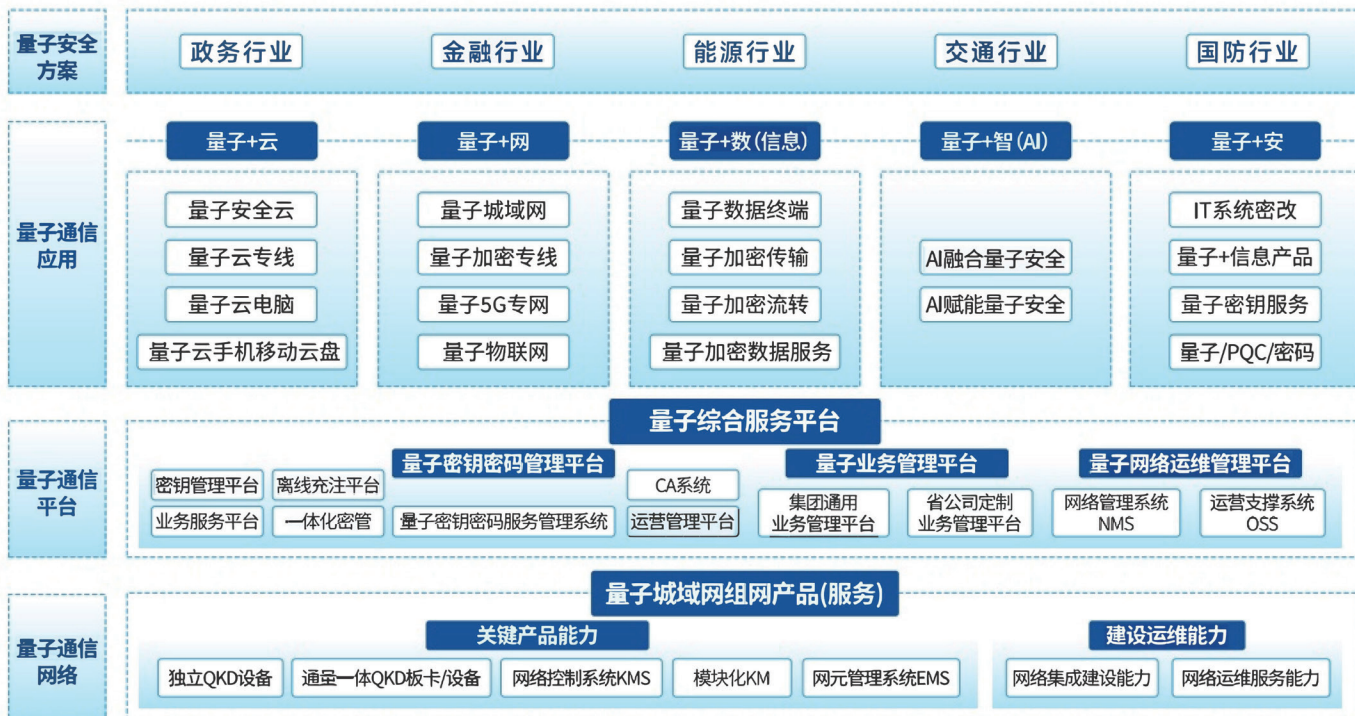
天融信官方微信



天融信官方网站

量子+产品体系

中移量子科技有限公司（简称“中移量子”）是中国移动控股的高科技企业，肩负国家量子科技战略落地使命。中移量子打造中国移动量子+“云、网、数、智、安”五维、一体（网、平台、应用三层一体化）量子安全能力矩阵，形成政务、金融、能源、交通、国防等行业解决方案，真正实现量子赋能千行百业。



中检集团天帷信息技术股份有限公司

中检天帷，是中国检验认证集团控股的网络数据安全服务机构。聚焦网络安全、数据安全、人工智能安全与检验检测数智化领域，构建面向数字经济时代的可信、安全、高质量发展体系。

网络安全

等保测评
商密评估
软件测试
风险评估
代码审计
攻防演练
渗透测试

数据安全

个人信息保护合规审计
数据分类分级
数据安全风险评估
数据跨境合规
智能网联安全

人工智能安全

高质量数据集检验检测
人工智能大模型评测
智能体检验检测
数字人安全评估
智能体管理能力成熟度评估

AI驱动网数安全合规治理 告别“人海战术”

大小模型协同驱动，全面覆盖境内外网数合规

中检天帷网络数据安全合规治理平台，内置100余种测评工具、Skill、Agent等，覆盖等保测评、商密评估、数据安全评估、个人信息保护合规审计、数据出境安全评估、软件测试、代码审计、ISO 9000、ISO/IEC 20000、ISO/IEC 27001、GDPR等主流合规要求，采用大小模型协同及知识图谱等人工智能技术，合规检测效率提升200%、7*24小时风险检测、穿透式监管降低IT供应链风险。助力企业实现网数安全合规治理自动化、智能化，实现资产全面精细化管理、风险精准预警及闭环管理，用AI对抗AI，打赢每一场网络空间战，守护企业网数安全，保障业务连续性！

7*24小时
风险检测

200%
效率提升

内置
100+
测评工具



☎ 电话:0551-62381688 ✉ 邮箱:swzc@tanovo.com 🌐 网址:tw.ccic.com
📍 公司总部(合肥公司):合肥市高新区望江西路900号中安创谷一期A1栋29层
中检天帷北京公司:北京市朝阳区西坝河东里18号中检大厦10层
中检天帷上海公司:上海市徐汇区虹梅路1801号A区凯科大厦1601号
中检天帷深圳公司:深圳市南山区科技南十二路中检大厦3层301

智护全域安全

AI-EMPOWERED CYBER · QUANTUM · LOW-ALTITUDE SECURITY

网络安全 × 量子安全 × 低空安全

央企担当 · 一站式安全运营服务



网络安全

皖安智盾

- **一体化安全运营托管**：7×24小时监测值守，威胁发现效率提升80%，分钟级响应
- 全省16地市城市安全运营中心，500+属地技术人员就近服务
- 安全大脑+等保助手，云边端服立体防护，攻击拦截率超98%



量子安全

量子+网络

- **OTN+量子融合专线**：全球首例，率先打造商用400G量子加密传输
- STN+量子：行业首个“SRv6+量子网络+国密算法”专线产品
- 量子组网网关轻量叠加，存量专线不改造即享加密；覆盖政务、金融、政法、卫健



低空安全

天翼星盾

- **5G-A多源融合侦测反制**：黑飞无人机秒级锁定，可侦测10平方厘米目标
- AI+ “1+1+4+N” 低空能力体系，全国31省落地1500+项目
- 护航春晚合肥分会场、十五运、数字中国峰会等重大活动低空安防保障

★ 中央网信办指定唯一省级运营商应急支撑单位

◆ 国家级攻防演练连续六年零失误零攻防

◆ 多年全国及省级安全技能竞赛一等奖

新华三主动安全

以 AI 筑盾，护智能未来

AI 赋能安全防御，威胁感知能力提升，风险处置能力增强



新华三主动安全理念持续演进，AI-Driven赋能安全未来



新华三 AI 安全战略以主动安全理念为底座，秉持以安全守护 AI 发展、以 AI 赋能安全防御双轮驱动核心理念，将安全嵌入云网算存端基础架构，融合流量、身份、应用、算力与行为等全域数据，依托AI分析与自动化编排，构建“建设一体、防护一体、运营一体”的内生安全体系，解决大模型、AI 智能体、算力集群带来的新型安全风险。

安全大模型底座

- 全网设备运维库：1.2 亿 + 政企网络设备故障、告警、配置样本；
- 攻防威胁知识库：二十余年累计百万级攻击样本、勒索病毒、APT、钓鱼、AI 生成恶意载荷；
- 合规政策库：等保 2.0、关基、数据安全法、个人信息保护法、行业专项合规条文；
- AI 新型风险库：Prompt 注入、数据投毒、模型窃取、智能体越权、向量数据库泄露案例。

AI in SEC

将AI能力深度内置到安全设备中，通过亿级云端情报、云端大模型与本地设备联动，实现云边协同、原生智能。

AI for SEC

依托于Agentic SOC自主式安全运营平台，实现自决策、自协同、自进化，让安全运营与业务防护同步升级。

SEC for AI

面向AI Agent 智能体全生命周期，从算力、模型、应用多个维度构建智能体安全防护体系，保障 AI 业务可信合规。



NIO es8

创 领 新 境

蔚来ES8

0首付5免2金融方案*

1万元选装基金限时享



*即日起至2026年9月30日(含), 锁单蔚来ES8的用户, 可申请蔚来指定合作机构提供的60期(5年)分期金融方案, 首付比例0%起, 前24期年化费率0%, 后36期年化费率3% (折合年化利率5.64%)。上述金融产品由官方合作银行农业银行等银行提供。具体贷款方案由合作机构根据用户信用状况等情况独立审批确定, 最终以金融机构实际审核结果和贷款合同约定为准。





CCSC
2026

Hefei, China
2026.09
中国·合肥



图片直播



学会公众号

