深悟两会精神
守牢网络安全防线

没有网络安全就没有国家安全

★ ★ ★ ★ ★

目录

CONTENTS

第一部分
深入推进数字中国建设

没有网络安全就没有国家安全

★ ★ ★ ★ ★

深刻认识数字中国建设的重大意义

1

数字中国建设是把握新一轮科技革命和产业变革、构筑国际竞争新优势、赢得大国博弈的重要支撑。习近平总书记深刻指出：“当今时代，数字技术、数字经济是世界科技革命和产业变革的先机，是新一轮国际竞争重点领域，我们一定要抓住先机、抢占未来发展制高点。”历史经验表明，谁能在科技革命和产业变革中抢占先机，谁就能掌握发展主动权。当今世界新一轮科技革命和产业变革加速突破，围绕数据要素和人工智能的大国博弈愈演愈烈，复杂性、严峻性、不确定性明显上升，数字技术成为变革的核心驱动力。数字中国建设为数字技术发展提供了广阔的应用场景、丰富的数据资源、完善的政策支持和良好的创新生态，推动我国在新一轮科技革命和产业变革中赢得主动、构筑国际竞争新优势。

2

数字中国建设是推进中国式现代化、发展新质生产力的重要引擎。数字中国建设成为推进中国式现代化不可或缺的重要力量。数字技术具有高创新性、强渗透性、广覆盖性的显著特征，是推动社会生产方式变革、生产关系再造、经济结构重塑、生活方式巨变的先导力量，能够催生新产业、新业态、新模式。数字中国建设通过推动数字产业化和产业数字化，显著提升社会整体生产力水平，促进产业转型升级，提高全要素生产率，为培育和发展新质生产力提供坚实支撑、注入强大动能，是推进中国式现代化、实现高质量发展的重要引擎。

3

数字中国建设是深化国际交流合作、推动构建人类命运共同体的必然选择。全球数字化、网络化、智能化快速发展既面临前所未有的机遇，也面临着数字鸿沟扩大、网络安全风险上升、规则体系不健全等突出挑战。我国拥有超大的规模优势、领先的技术设施和丰富的应用场景，在数字化发展过程中不断孕育创新的理念、技术和模式，为全球应对共同的数字化挑战提供中国方案、贡献中国智慧。建设数字中国，打造开放共赢的数字领域国际合作格局，彰显了中国主动顺应全球数字化浪潮、积极参与全球治理的大国担当。

数字中国建设取得的重大成就和面临的挑战

重大成就

数字技术创新取得新突破

数字领域突破一批关键核心技术，“缺芯少魂”等突出问题逐步解决，人工智能综合实力实现整体性、系统性跃升。“东数西算”成为国家重大生产力布局的战略工程，算力和数据基础设施加快布局，在规模、技术等方面实现长足发展。

数字经济发展动能更加强劲

实体经济和数字技术融合程度显著加深，智能化转变、数字化改造加速推进，数据产业逐步壮大，全国一体化数据市场正在加快构建。数据基础制度不断完善，深入开展数据应用示范场景建设，数据要素的放大、叠加、倍增作用日益凸显。

数字公共服务更加可感可及

数字技术显著提升公共服务的可及性和公平普惠水平，教育、医疗、社保、养老等数字化水平不断提升。数字化打破了传统公共服务中的“流程壁垒”、“数据壁垒”，越来越多的事项实现“数据多跑路、群众少跑腿”。

数字领域国际合作取得显著成效

积极提出数字经济“中国主张”，与26个国家签署数字经济合作谅解备忘录，“丝路电商”伙伴国增加到36个。加强数字基础设施、数据规则标准、数字经济人才的国际合作，数字领域国际交流的深度和广度不断拓展。

面临挑战

我国数字产业总体上大而不强，技术原创供给能力有待提升，算力布局亟待优化，数据交易共享与跨境流通带来新的安全风险，适应数智技术发展的伦理规范和治理体系仍需完善。

深入推进数字中国建设的重点任务

健全数据要素基础制度，建设全国一体化数据市场

一是健全数据要素基础制度。建立保障权益、合规使用的数据产权制度，完善持有权、使用权、经营权的制度框架 。二是优化数据要素市场化配置。加快培育全国一体化数据市场，规范数据流通交易行为，降低流通交易成本，提高流通交易效率。三是深化数据资源开发利用。

促进实体经济和数字经济深度融合，打造具有国际竞争力的数字产业集群

一是深入实施产业数字化转型。二是培育国际一流数字产业集群。三是实施工业互联网创新发展工程。四是推动平台经济创新和健康发展。

加快人工智能等数智技术创新，强化算力算法数据供给

一是构建系统完备的数智技术创新体系。二是突破人工智能基础理论和核心技术。三是强化人工智能算力、算法、数据等高效供给。

全面实施“人工智能+”行动，全方位赋能千行百业

一是强化示范引领，支持人工智能落地。二是推进人工智能规模化商业化应用。三是加强人工智能同产业发展、文化建设、民生保障、社会治理相结合。四是推动平台经济创新和健康发展。

统筹发展和安全，构筑数字中国建设安全屏障

一是提升安全防护能力。二是强化数据安全保障。三是加强安全技术供给。发展零信任安全架构、隐私计算、量子密钥分发等数字安全新技术，推动人工智能、区块链等新兴技术与安全技术融合发展。

切实把数字中国建设决策部署落到实处

建设数字中国是党中央作出的重大决策部署，是一项长期而艰巨的战略任务。要坚持和加强党的全面领导，切实把党的领导贯穿到数字中国建设的全过程各方面，加强数字中国建设整体布局，健全数字中国建设统筹协调机制，统筹推进数字中国建设重大机制创新、重大战略落地实施。

优化数字经济发展环境

深化数字经济发展的体制机制改革，建立健全数字治理法律制度，探索数据资产化、数字信用等制度创新。构建中国特色人工智能治理体系，完善人工智能技术研发、应用和监管的法律法规。深化数字化标准规范体系建设，促进数字领域标准统一，完善基础通用标准、数字基础设施建设规范标准、数据开放利用标准、数字化融合应用标准等标准体系。

培养高水平数字人才队伍

打造更多国际科学合作研究平台，完善具有国际竞争力的高端人才引进制度体系，吸引全球顶尖科学家来华开展科学研究工作，打造全球数字人才高地。支持产学研协同攻关，通过校企合作、共建学院或联合实验室等方式，在数字前沿领域培养卓越工程师和核心技术人才，形成贯穿产业链条的工程科技人才队伍。建立适应科技发展规律和国家重大需求的学科专业动态调整机制，鼓励高校在数字技术关键领域设置新兴专业，培养创新型、应用型、复合型人才。

积极参与数字治理与国际合作

主动布局和利用国际创新资源，开展双多边数字经济治理合作，推动更开放、包容、务实的国际科技交流合作。积极参与数字技术的国际治理，平衡技术创新与风险防范，推动国际规则和数字技术标准制定，全面提升我国数字治理的国际化水平和影响力。以建立利益共同体为合作目标，拓展全球技术经济产业合作，发展壮大互惠互利的国际合作网络。加快与共建“一带一路”国家数字基础设施互联互通，高质量共建“数字丝绸之路”，鼓励数字经济企业“走出去”，深度融入全球数字化发展生态。

第二部分
筑牢网络安全防线

没有网络安全就没有国家安全

★ ★ ★ ★ ★

以“AI+安全”双轮驱动 筑牢现代化产业体系根基

齐向东委员

我国产业体系规模大、门类全、配套完善、创新活跃、市场广阔，正处在由大到强、由全到优的关键跃升期。但也要看到，转型过程仍面临三重突出挑战：技术依赖及新技术伴生的安全风险，创新链和产业链衔接不畅，跨领域、跨行业的协同机制不健全。亟须以“AI + 安全”双轮驱动，形成“安全筑基—智能应用—绿色融合”的现代化产业体系良性发展循环。

全国第十四届全国政协委员、全国工商联副主席
奇安信科技集团董事长

筑牢安全根基

提升现代化产业体系安全发展效能。以“安全前置”为核心导向，构建全链条、多层次的安全保障体系。攻坚前沿安全技术，推动建立全链式安全架构。

催化智能动力

贯通科技创新和产业创新协同链条，聚集现代化产业体系创新发展势能。强化创新载体建设，支持产学研跨界融合，鼓励符合条件的企业进行智能化改造和设备更新。

着眼绿色融合

积极稳妥推进和实现碳达峰，激活现代化产业体系绿色发展动能。将技术应用延伸至高耗能企业碳排放监测、生态环境风险预警等领域，加大对节能降碳、循环利用等智能设备的研发与应用力度，持续推进绿色工厂、绿色园区建设，形成一批可复制、可推广的绿色发展模式。

以技术创新筑牢网络安全防线

肖新光委员

从大模型、智能体到具身智能，随着各类技术和应用不断推陈出新，安全问题也如影随形。全国政协委员、安天科技集团董事长肖新光表示，要以技术创新筑牢网络安全防线。肖新光建议，以网络安全反病毒引擎等共性能力为基础，聚合人工智能领域的共性安全问题集中攻关，推动共性技术引擎嵌入各类人工智能产品与应用场景，让其“出厂即带安全基因”，实现“关口前移”，全面筑牢人工智能应用的安全地基。

第十三届、十四届全国政协委员，安天科技集团创始人、董事长

随着人工智能应用全面渗透经济社会各领域，风险暴露面迅速扩大，攻击行为体加速运用人工智能，自动编写恶意程序，自主攻击投放。“旧的安全范式早已无法应对新的风险挑战。”肖新光直言。

对人工智能时代的安全挑战，关键要发挥制度优势，打造国家主导、战略企业研发、产业广泛应用的“人工智能+”国家安全技术引擎。他建议，以网络安全反病毒引擎等共性能力为基础，聚合人工智能领域的共性安全问题集中攻关，推动共性技术引擎嵌入各类人工智能产品与应用场景，让相关产品、装备、场景“出厂即带安全基因”，避免先发展后治理的失控风险，实现“关口前移，防患于未然”，全面筑牢人工智能应用的安全地基。

筑牢未成年人网络安全屏障

吴城委员

党中央始终高度重视未成年人工作，关心未成年人成长。党的二十届四中全会强调“加强未成年人网络保护”，今年政府工作报告将“加强网络内容建设和管理，深化网络综合治理”列为重点任务，推进未成年人网络保护。网络保护成为未成年人保护体系中不可或缺的重要组成部分，构建全方位未成年人网络保护体系、筑牢未成年人网络安全屏障至关重要。

第十四届全国政协委员，内蒙自治区工商联副主席、国城控股集团有限公司董事长

强化技术防线

构建网络游戏安全物理边界的第一步是强化技术防线。应全面推行“人脸识别＋可信身份”双重验证，强制要求所有网络游戏在注册、登录、充值及夜间等时段，实施高频随机人脸识别和语音验证，特别是加强对老年账号异常行为的核验力度，防范未成年人冒用长辈身份。

压实责任主体

网络游戏企业作为筑牢未成年人网络安全屏障的责任主体，要进一步压实责任，推动行业自律与规范发展。游戏企业应在研发立项、内容设计、付费机制、运营活动等环节进行未成年人影响专项评估，将评估结果作为上架前置条件，并建立游戏适龄提示标准强制执行机制。

营造健康网络生态

要深化“校家社”协同育人，学校可将网络素养教育纳入课程体系，定期反馈学生用网情况；家庭要主动“补位”，压实监护责任，在防沉迷、屏幕时间管理和风险识别等方面要负起责任；社区增加公益性线下活动供给。

第三部分
两会“数据”声音

没有网络安全就没有国家安全

★ ★ ★ ★ ★

强化高质量数据供给，保障数据安全

崔岩委员

在科研与产业落地实践中，数据质量、安全方面的痛点集中体现为数据产权不明晰、标准体系缺失等。数据产权归属不清，导致经营主体因法律风险不敢投入数据开发利用，或因缺乏产权保护不愿开放共享高价值数据；数据采集、治理、标注、安全等标准体系不健全，不利于数据要素流通和价值安全释放，直接影响AI技术从实验室走向产业一线。同时，数据安全方面也存在一系列新风险、新挑战。崔岩代表认为，一是数据流通风险，跨区域、跨行业数据共享需求增加，但隐私计算等技术应用覆盖率不足；二是数据滥用风险，部分经营主体因产权保护不足而不敢开放数据，或因监管缺失导致数据被非法使用；三是数据基础设施分散建设，缺乏互联互通，难以保障数据全生命周期安全。

第十四届全国人大代表，五邑大学创新创业学院副院长，中德人工智能研究院院长，四维时代创始人

应加快完善数据基础制度，探索将数据产权登记证书作为流通交易的可信凭证，健全数据标准体系；加大高质量数据供给，建设公共数据共享平台，推广数据集创新模式；培育壮大数据产业，梯度培育数据企业，建设产业集聚区；强化数据基础设施建设，探索“可信数据空间+数据交易所”模式，构建数据要素流通“一张网”，以制度创新与设施建设双轮驱动，在保障安全的前提下充分释放数据要素价值。

完善数据流通安全合规体系建设 为数据要素市场筑牢安全底座

周鸿祎委员

提出从规则制度、合规指引和技术能力三个层面构建数据流通安全闭环体系的建议，为数据要素市场筑牢安全底座。在规则制度层面，建议细化数据流通安全实施细则，统一数据分类分级与安全防护核心标准，在合规指引层面，提出由主管部门牵头，联合安全企业、第三方服务机构编制数据流通安全合规指引，在技术能力层面，强调推动数据安全技术应用与关键技术攻关，在数据流通基础设施中同步嵌入网络安全与AI安全技术，攻关AI辅助的数据溯源、防篡改、权限审计等技术。

第十四届全国政协委员、经济委员会委员，360集团创始人

数据流通规模不断扩大

我国已构建以《数据安全法》等为核心的数据安全法律法规体系，在政策引领与技术驱动下，数据交易所、平台运营主体以及专业数据服务商不断增加，数据流通规模持续扩大。但在实践中仍存在一些现实问题，如政策法规原则性要求较多、可操作细则仍有待细化；数据跨境流动、多模态数据融合流通等新兴场景的操作规范尚不清晰；同时数据安全技术应用与攻关还需要进一步加强。

细化数据流通实施细则

应进一步细化数据流通安全实施细则，统一数据分类分级与安全防护的核心标准，并针对数据开放共享、数据交易、数据跨境流动等典型场景，明确实操要求。同时，可由主管部门牵头组织安全企业、第三方服务机构等共同编制数据流通安全合规指引，通过“手册+工具包”的方式，为企业特别是中小企业提供清晰、可落地的合规路径。

推动数据安全技术应用攻关

应推动数据安全技术应用与关键技术攻关，在数据流通基础设施中同步嵌入网络安全技术和AI安全技术，通过攻关AI辅助的数据溯源、防篡改、权限审计等技术手段，提升数据流通基础设施的内生安全能力。通过技术手段实现数据“可用不可见、用途可控可计量、全程可追溯可审计”，从根源上消除企业对数据泄密、权属纠纷的顾虑。

加快完善数据安全可信流通体系

程伟委员

“在保障国家安全和个人隐私的前提下，推动各级政府依托统一平台或授权平台，有序开放公共数据资源，促进公共数据与社会数据的可信融合利用。”数据已成为国家基础性战略资源和关键生产要素，大数据技术在提升政府决策科学化、社会治理精准化、公共服务高效化方面展现出巨大潜力，其安全可信流通与高效开发利用，是推动数字经济高质量发展、提升国家治理能力、构筑国际竞争新优势的核心支撑。党中央、国务院高度重视数据要素市场建设，相继出台相关法律法规，为数据资源在政务、医疗、交通、城市管理等领域的融合应用提供了坚实基础。然而，数据要素从资源化到资产化、资本化的进程中，仍面临一系列关键挑战，亟需找到安全与发展平衡的实践路径。

第十四届全国人大代表，中国移动湖南公司党委书记、董事长、总经理

技术手段和合规要求衔接不足

数据要素市场前景广阔，但在推动数据安全可信流通与开发利用的实际过程中，部分单位仍存在“不愿共享、不敢共享、不会共享”现象。同时，技术手段与合规要求衔接不足，隐私计算、区块链、数据脱敏等技术为数据“可用不可见”“可控可计量”提供了可能，但各类技术方案标准不一，导致公共数据开发利用进展缓慢。

完善数据产权登记制度

加快完善适应数字经济特征的数据产权登记制度，探索数据资源持有权、数据加工使用权、数据产品经营权等分置运行的实现路径；同时加快完善政务数据安全合规评估等标准体系，推动建立覆盖国家、省、市、县四级的分布式数据目录体系，明确数据共享范围、责任主体和更新时限，实现全国政务数据“一本账”管理。

支持投资可信数据流通平台

支持符合安全资质的电信运营商、科技企业、专业机构等市场主体，投资建设基于隐私计算、区块链、数据沙箱等技术的第三方可信数据流通服务平台，拓展应用场景深度；鼓励数据持有方与数据服务商合作，将原始数据转化为标准化、模块化、场景化的数据产品(如分析报告、指数产品、模型API等)和安全合规的数据服务。

以可信数据空间为突破口 加快推动新型工业化发展

杨剑宇委员

当前工业可信数据空间建设面临多重挑战。在数据合规方面，要求不明确。从试点情况来看，对“合规性”的顾虑导致企业普遍存在“不愿、不敢、不会”的“三不”心态：“不愿”开放核心工艺数据，顾虑数据泄密和同行偷师；“不敢”参与数据流通，担忧数据安全责任归属不明；“不会”挖掘数据价值，数据利用率低。同时，“数商”等专业服务机构发展滞后，难以满足工业场景数据利用需求。在可信数据空间标准建设方面，处于起步阶段。我国拥有全球最多的工业门类，但不同设备、不同厂商间的技术与标准大量不兼容，严重影响跨域、跨行业数据流通效率，加速建立健全工业数据标准体系已成为产业普遍呼声。

第十四届全国人大代表，中国移动浙江公司党委书记、董事长、总经理

完善工业可信数据合规体系

完善工业可信数据空间的合规管理体系。制定工业数据分类分级管理的实施细则，建立全国统一的工业重要数据备案与风险评估监管平台，切实加强工业数据合规治理和监管，提升可信数据空间的公信力；明确可信数据空间参与方的权责边界，厘清数据安全与隐私保护责任归属，建立数据流通利益纠纷调解机制；加强企业专项培训，积极培育扶持“数商”等服务机构。

加快技术标准体系构建和落地

加快技术标准体系的构建和落地验证。成立跨部门工作组，引导推进工业领域数据格式、接口规范、安全认证等标准体系建设；针对工业数据流通、安全防护等重点需求搭建共性技术平台，为技术协同与标准互认提供落地支撑；通过财政补贴、试点示范等方式，协同推进标准建设与试点项目，鼓励产业主体参与标准落地验证，为规模市场效应形成筑牢标准支撑。

强化可信数据空间产业引导

强化工业可信数据空间建设的产业引导。通过设立专项基金、专项补贴等多种形式，加大对工业可信数据空间建设运营相关的基础设施接入、核心组件研发应用等支持力度；组织实施关键技术“揭榜挂帅”工程，聚焦隐私计算、多方安全计算等领域开展核心技术攻关；建设工业可信数据空间的技术验证与中试平台，推动通用技术方案在真实工业场景的定制化应用，降低企业技术应用成本。

加强中资企业出海境外数据安全合规管理

李丹委员

为构建“政府引导、企业主体、专业支撑、内外协同”的立体化应对体系，引导企业建立健全世界一流的数据治理能力，全国政协委员、普华永道中天会计师事务所首席合伙人李丹提出了5个关键点。

第十四届全国政协委员、普华永道中天会计师事务所首席合伙人

加强顶层设计与宏观指导

由国家相关部委牵头，研究制定中资企业境外数据安全合规管理的指导性文件。组织力量跟踪、研判重点国家与地区的数据法规动态、执法案例及风险预警，建立“一国一策”合规指引库。同时，将数据跨境流动规则对等、非歧视待遇等核心关切，纳入国际经贸谈判。

压实主体责任与能力建设

鼓励企业设立“首席数据合规官”，建立世界一流的数据治理体系与合规内控机制。支持相关行业协会牵头，建立与国际接轨的企业数据合规管理体系认证标准，开展“合规能力认证”计划。研究设立专项资金或给予税收优惠等方式，对企业投入合规技术研发、境外合规基础设施建设及获取国际认证予以适当扶持。

突破关键技术与标准制约

在国家重点研发计划中设立专项，支持开展自主可控的数据分类分级、隐私计算、可信跨境传输等关键技术攻关和一体化解决方案研发。积极参与国际标准组织相关工作，力争将我国在数据安全、个人信息保护等方面的实践成果纳入国际规则体系，促进中国标准“走出去”。

构建专业服务与应急响应生态

培育一批具备全球服务能力的法律、会计、咨询等专业化服务机构，为企业提供一站式合规解决方案。依托贸促会的海外分会以及海外中企商协会等平台，建立境外数据合规风险应急响应机制，为企业应对境外执法检查、处置数据安全事件等提供及时指导和必要支持。

深化国际合作与互信共商

与共建“一带一路”国家及主要贸易伙伴在数据保护认证领域拓展等效互认合作，降低企业合规负担。通过二十国集团（G20）、世界贸易组织（WTO）电子商务谈判等多边平台，倡导构建开放、包容、普惠的数字经济国际规则。

第四部分
人工智能治理与安全

没有网络安全就没有国家安全

★ ★ ★ ★ ★

构建普惠、包容、可持续的全球人工智能治理体系

张凤委员

构建普惠、包容、可持续的全球人工智能治理体系关乎人类命运共同体构建。人工智能是具有国际公共品属性的技术，其研发成果理应体现非排他性和共享性，使各国无论发展水平高低、技术能力强弱，都能够在公平条件下获取技术红利、共享发展成果。同时，作为具有全球公共属性的重要技术领域，人工智能治理不仅涉及技术规则，更关系价值理念与发展模式选择。

第十四届全国政协委员、民族和宗教委员会委员，
中国科学院科技战略咨询研究院研究员

包容多样性和差异性

要求全球人工智能治理体系包容文明多样性与发展阶段差异性，避免技术霸权与规则垄断，让不同制度、文化、语言背景的国家平等参与规则制定。面向未来，世界各国应共同承担技术风险与治理责任，推动AI向绿色低碳、安全可信、以人为本的方向演进，只有在创新发展与风险治理之间实现动态平衡，才能确保人工智能长期可持续健康演进，使其始终服务于增进人类福祉这一根本目标。

深化“人工智能+”

全球南方国家在算力资源、开源模型、本土化应用等方面仍面临较大差距。应加快落实“‘人工智能+’国际合作倡议”，以联合国和“一带一路”框架为主要平台，结合全球南方国家的实际需求，针对性开展分层分类的能力建设援助计划，重点支持本地化模型训练、多语种人工智能工具开发及数字基础设施升级；设立专项基金，向发展中国家提供精准农业、传染病监测等有针对性的轻量化模型及配套算力资源，切实提升技术可及性与实用性；同步开展治理经验交流与能力培训，缩小数字能力差距。

安全可信与绿色低碳并重

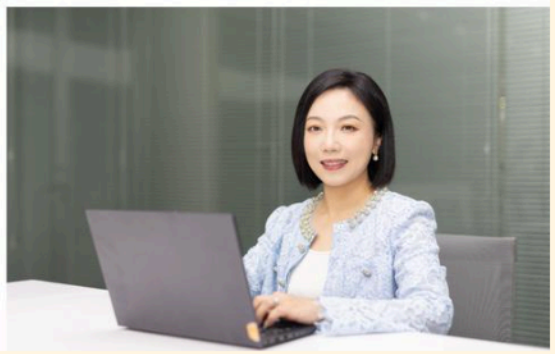
应加快研发“AI合规审计工具包”，形成可操作、可验证、可推广的治理实践范式，为全球人工智能治理提供标准化技术支撑；同步推进绿色算力基础设施共建共享，推广低功耗模型训练与边缘智能部署方案，降低发展中国家人工智能应用门槛；建立跨区域人工智能风险联合监测与应急响应机制，强化对深度伪造、数据滥用等新型威胁的协同识别与处置能力，实现创新红利最大化与系统性风险最小化，确保人工智能长期向安全、可信、绿色、低碳、以人为本的方向演进。

强化AI监管与产业规范 让数字治理护航市场民生

钟铮委员

当前网络营销领域，传统违规营销与技术型恶意营销交织蔓延，尤其是人工智能生成内容（AIGC）、生成式引擎优化（GEO）等新技术被滥用于数据污染、虚假营销、商业诋毁等行为，不仅侵害了消费者与合规企业的合法权益，更破坏了公平竞争的市场秩序，成为阻碍行业高质量发展的新型风险源。

随着AI数字人制作和运营门槛的降低，其在直播领域的滥用和违规现象日益增多，暴露出身份虚假、内容失范、权益纠纷、监管滞后等突出问题，若不及时治理，将扰乱市场秩序，损害用户权益，影响人工智能技术的社会信任与应用前景。



第十四届全国人大代表，美的集团股份有限公司副总裁兼首席财务官

明确违规营销法律边界

完善法律法规，明确违规营销的法律边界。她建议在《反不正当竞争法》修订中，增设专门条款，将利用AIGC、GEO等技术实施的恶意营销行为明确列为不正当竞争情形，并设定阶梯式罚则，情节严重的可追究刑事责任。同时，要明确企业、平台、技术服务方三方责任，企业需对其发布的AI生成内容真实性负责，平台要对GEO投喂内容履行主动监测与处置义务，GEO服务机构则不得为恶意竞争提供算法支持。

构建智能审核与闭环治理机制

强化平台主体责任，构建智能审核与闭环治理机制。大型电商平台、社交平台应开发企业营销专用AI内容识别模型，提升对虚假测评、参数篡改、恶意拉踩等行为的自动识别率。推行“一次举证、全网拦截”机制，当企业提交有效证据后，平台需在24小时内完成全站相似内容、同源GEO信息的排查与删除，并纳入负面样本库实现长效拦截。此外，由网信部门推动建立跨平台虚假信息共享数据库，打通各平台审核标准与黑名单，防止违规内容“打一枪换一个地方”。

规范AI数字人直播行业

设立AI数字人直播专项管理条款，推行数字人注册备案制度，强制公示其技术生成属性和背后运营主体信息，做到“数字形象可标识、责任主体可追溯”。同时，修订《消费者权益保护法》《广告法》等法律法规，明确将数字人直播纳入监管范畴，严厉打击虚假宣传、仿冒名人等消费欺诈行为。此外，还要完善知识产权与人格权相关法律法规，明确AI数字人制作中真人肖像、声音、个人信息等的授权使用边界，加大对无授权使用行为的处罚力度，保障权利人合法权益。

推动AI大模型健康发展 更好赋能实体经济转型升级

苗伟委员

建议设立AI大模型训练专用算力开放平台，明确平台“公益优先、市场补充”的定位。通过根据用户需求动态分配算力、实行“阶梯式收费”机制，推动AI大模型健康发展，更好赋能实体经济转型升级。在数据要素市场化配置改革方面，他建议构建公共数据授权运营负面清单，明确运营“禁区”，在保障安全与使用公平的同时，坚决捍卫公共数据的公益本色。

无论是打通“东数西算”的数据通道，还是打造5G全连接智能工厂，科技企业需将技术优势切实转化为产业优势。通过政策规范、法治护航与企业创新的同频共振，数字中国建设正向着高质量发展稳步迈进。

第十四届全国人大代表，中兴通讯党委书记、高级副总裁

技术手段和合规要求衔接不足

国家高度聚焦具身智能等产业领域发展，越来越多个人信息主动或被动暴露于人工智能程序中，其中具有强烈人身专属性的生物识别信息被大规模采集。因此，建议严格落实“告知—同意”原则，重点加强人脸信息保护，构建全链条权利保障机制，明确人脸数据处理与面容克隆的法律边界。

构建公共数据授权运营负面清单

构建公共数据授权运营负面清单。当前，我国正大力发展数字经济，初步构建了制度框架，但实践中仍存在公共数据授权运营边界模糊、安全风险凸显、公益性与营利性等挑战。因此，建议建立公共数据授权运营的负面清单管理制度，明确运营“禁区”，规范“限制类”行为，在保障安全与公平的同时，坚决捍卫公共数据的公益本色。

前瞻布局6G产业

国家已在5G时代实现领跑，6G时代将进入“无人区”，面对技术脱钩、安全审查等非市场因素挑战，需要提前开展6G原型验证、标准预研与频谱规划，因此，建议坚持适度超前建设原则，保持对5G演进阶段的投资强度，为6G领先找到一条确定性路径，牢牢掌握未来发展主动权。

聚焦人工智能“六力模型”建设 推动智能体规模应用

周鸿祎委员

2026年全国两会期间，全国政协委员、360集团创始人周鸿祎围绕人工智能高质量发展持续发声。他认为，我国人工智能产业正逐步形成“电力—算力—智力—人力—安全力—生产力”协同推进的“六力模型”，为“人工智能+”行动落地夯实基础。围绕这一体系构建，他今年重点关注推理算力布局优化、智能体公共服务平台建设以及安全智能体规模应用等方向。

第十四届全国政协委员、经济委员会委员，360集团创始人

出台指导政策

应在全国一体化算力体系框架下出台推理算力布局指导政策，建立“全国统筹+区域细化”的布局体系，在重点产业区域建设低时延、高密度的推理算力集群，通过调度机制提升资源利用效率。同时鼓励专用推理芯片国产化发展，实现产业链自主可控，支撑智能体技术的深度应用。

双线赋能

实施技术与人才“双线赋能”，由相关部门牵头建设普惠型智能体公共服务平台和智能体课堂。平台集成模型能力与行业工具，提供全流程服务，支持中小企业低成本构建垂直领域智能体。同时推行“以模治模”的安全防护机制，发布安全智能体场景适配指南，开展技能培训与认证，培养“懂AI又懂业务”的专业人才。

“以模对模”

安全行业亟需转型，加快推行“以模治模”，用AI治理AI。相关部门应支持兼具“安全+AI”能力的企业打造漏洞处置、攻击溯源分析等系列安全智能体产品，在关键信息基础设施、工业互联网等重点领域批量部署，推动场景化应用。同时鼓励行业龙头企业联合科研机构创新，将具备实战能力的安全智能体产品纳入优先采购目录，推动规模化应用。

建议制定教育场景生成式AI使用负面清单和推荐清单

龙婉丽委员

要把学术诚信教育前置到义务教育阶段，把“引用规范、事实核验、来源标注、AI输出可疑点识别”纳入课堂常规训练。人工智能尤其是生成式大模型会给教育带来一系列新的挑战。一是未成年人数据与隐私保护存在风险。教育处理大量未成年人信息，且往往涉及敏感信息，如学习轨迹、行为数据、心理与健康相关线索、家庭情况等，一旦随手调用公网大模型把学生作品、作业、对话记录上传到第三方平台，就可能形成数据出域与再利用风险。二是生成式AI“代劳化”趋势日渐明显。不少一线教师与家长表示孩子越来越会“出答案”，却越来越不愿“走过程”。一旦“能力外包”成为习惯，学生会在不自知中失去耐心、推理力与判断力，甚至影响到学习品质与学术诚信。三是模型输出的不确定性，可能导致教育的不稳定性。生成式AI构建学生画像时，易因数据片面、算法偏见产生价值偏差和错误引导。四是缺少教育可用的可信基础设施与治理闭环。学校既难统一数据安全与合规要求，也难形成可复用的教学范式与评价机制，更难压实平台与供应商责任。

第十四届全国政协委员，民盟中央委员、民盟上海市委副主委，上海市静安区副区长

未成年人数据不出域

把“未成年人数据不出域”作为AI+教育工程底线。建设省市级（或地市级）教育人工智能公共底座，提供面向教育行业的统一入口、统一身份认证、统一访问策略、统一日志审计与留痕等，以及可控可审计的模型与算力服务。在制度上明确，教育数据能不出域就不出域；确需出域的，必须有合规评估、脱敏与授权机制，并可追溯、可问责。

学习过程还给学生

把“学习过程还给学生”作为AI+教育的育人主导。制定教育场景生成式AI使用负面清单和推荐清单，强调杜绝“代劳式”使用，并按学段、学科、任务类型细化。把学术诚信教育前置到义务教育阶段，把“引用规范、事实核验、来源标注、AI输出可疑点识别”纳入课堂常规训练。对未成年人过度依赖、沉迷AI风险设置防火墙，设计分级使用权限，引入使用时长与任务类型限制。

信创适配 安全可控

把“信创适配 + 国产安全可控”作为AI+教育的技术防线。把信创适配前置为立项门槛，关键组件接口标准化，确保数据与日志可迁移、可验证销毁，避免被单一厂商锁定。在产品端明确供应商对数据合规、内容安全、漏洞处置、版本变更可回滚与应急响应的硬责任；在学校端建立数据、权限、调用三本台账，流程写清“谁审批、谁复核、谁负责”在个人端涉及敏感数据、对外发布、评价结论必须人工复核与签名留痕。

第五部分
网络生态与内容安全

没有网络安全就没有国家安全

★ ★ ★ ★ ★

构建法治化治理体系 综合治理网络谣言乱象

孙建国委员

“网络谣言已成为危害社会稳定、侵蚀网络生态的突出问题，亟须以法治思维和法治方式推进全链条治理。”当前网络谣言传播呈现多元化、产业化、智能化特征，以社交平台为核心扩散阵地，个别自媒体与网络水军形成“内容生产—账号分发—流量变现”黑色产业链，封闭社交群组成为精准投放渠道。AI技术的发展更让谣言生成更高效、伪装更专业，部分AI产品未落实标识要求，进一步模糊真假边界。部分谣言歪曲事实、扰乱网络秩序、损害政府公信力，也易引发社会恐慌、激化矛盾，导致基层治理成本大幅增加。

全国人大代表、广东省佛山市公安局高明分局党委副书记、政委，三级高级警长

构建法制化治理体系

构建法治化治理体系，完善配套法规，压实网络运营者法定义务，健全行政、民事、刑事梯度追责机制，强化部门协同，借鉴“清朗”专项行动模式，严惩造谣传谣行为。

压实平台主体责任

压实平台主体责任，运用AI、大数据等技术对存疑及AI生成信息进行醒目标识，建立全流程溯源机制，完善内容审核与算法推荐，畅通举报渠道，实现谣言快核快处、快速封控。

强化技术反制能力

强化技术反制能力，加大鉴伪技术研发投入，建立全国统一谣言数据库与鉴伪平台，推广全链条溯源标记技术，实现谣言“可查、可溯、可追责”。

跨部门跨区域协同

深化跨部门跨区域协同，建立网信部门牵头的联席会议制度，加强行业自律，健全跨区域执法协作机制，形成“一地发现、全域响应”的闭环治理格局。

加强个人信息保护 构建全链条责任体系

陈友坤委员

“个人信息被过度收集、违规使用的问题仍然存在，个人信息在用户不知情、未同意的情况下被收集、流转，不仅侵害了公民合法权益，也扰乱了市场的公平竞争秩序。”一些互联网应用程序收集的信息远超“必要个人信息范围”，违规获取用户通讯录、位置等敏感权限；用户授权流于形式，个人信息收集的知情同意原则被架空；权利救济困难，信息技术的不对称导致用户难以获取关键证据，个人信息侵权行为的损害后果也难以量化。

第十四届全国政协委员、北京大成（重庆）律师事务所党委书记、董事局主席

完善法律法规

完善法律法规，构建全链条责任体系，建议在修订个人信息保护法或出台相关法律、司法解释时，进一步明确个人信息处理全链条各主体的法律责任，建立起覆盖“收集前、收集中、收集后”全过程的监管闭环。

强化源头治理

强化源头治理，建立个人信息保护影响评估制度。建议对涉及大规模处理敏感个人信息、新技术应用、跨平台数据融合等高风险场景，强制要求企业在信息收集前开展风险评估。

规范授权机制

规范授权机制，保障用户知情同意权。建议明确个人信息收集必须遵循“单独告知、明示同意、场景一致”原则。禁止将信息收集与功能使用强制捆绑，不得使用“不同意即不可用”的霸王条款。

健全司法保障

健全司法保障，完善公益诉讼与惩罚性赔偿。建议进一步完善个人信息保护公益诉讼制度，明确惩罚性赔偿在个人信息保护领域的适用条件，真正实现“让违法者不敢再犯”。

加强网络乱象治理 规范人工智能健康发展

张伯礼委员

“目前，不法分子利用AI合成虚假音视频，进行虚假宣传甚至诈骗的现象屡见不鲜，若AI技术进一步失控，还将带来更大安全隐患。”针对网络造谣、AI技术滥用等乱象，提出多项针对性建议，呼吁多部门联合成立工作专班，建立快速联动机制，开展专项整治行动；完善自媒体虚假造谣信息处置机制，明确管理部门与平台责任及处置流程、时间；重点关注受指使散布的虚假信息及灰色信息内容，做到早发现、早处置；进一步完善法律法规，引入惩罚性赔偿制度，对造谣账号不仅删帖封号，更要依法处罚责任主体，定期公布典型案例，提高违法成本。

全国人大代表、中国工程院院士

坚持发展与规范并重

立足长远、系统谋划，从法治、政策、标准、伦理、监管等维度协同发力，形成覆盖AI全生命周期、激励和约束并重的治理网络。

构建立体化法律体系

构建分级分类的立体化法律体系。加快推动国家层面的人工智能综合性立法，形成以人工智能法为统领，部门规章、技术标准、伦理指南相互支撑的制度体系。

建立伦理和安全机制

建立刚柔并济的伦理治理与安全评估机制。科技伦理是人工智能监管的第一道防线。推动高校、科研机构、企业等责任主体普遍设立科技伦理委员会。加强对涉及国家和公众安全有影响的领域，在AI使用中严格管理，实施多重密钥限制，确保国家和公众安全。

强化全流程追溯

强化全流程追溯与追责能力。全面推行人工智能生成内容标识制度，标注生成来源、时间及版权归属，确保生成可识别、侵权可追踪。加快出台人工智能侵权的认定标准和赔偿细则。建立侵权快速响应机制，形成社会共治合力。

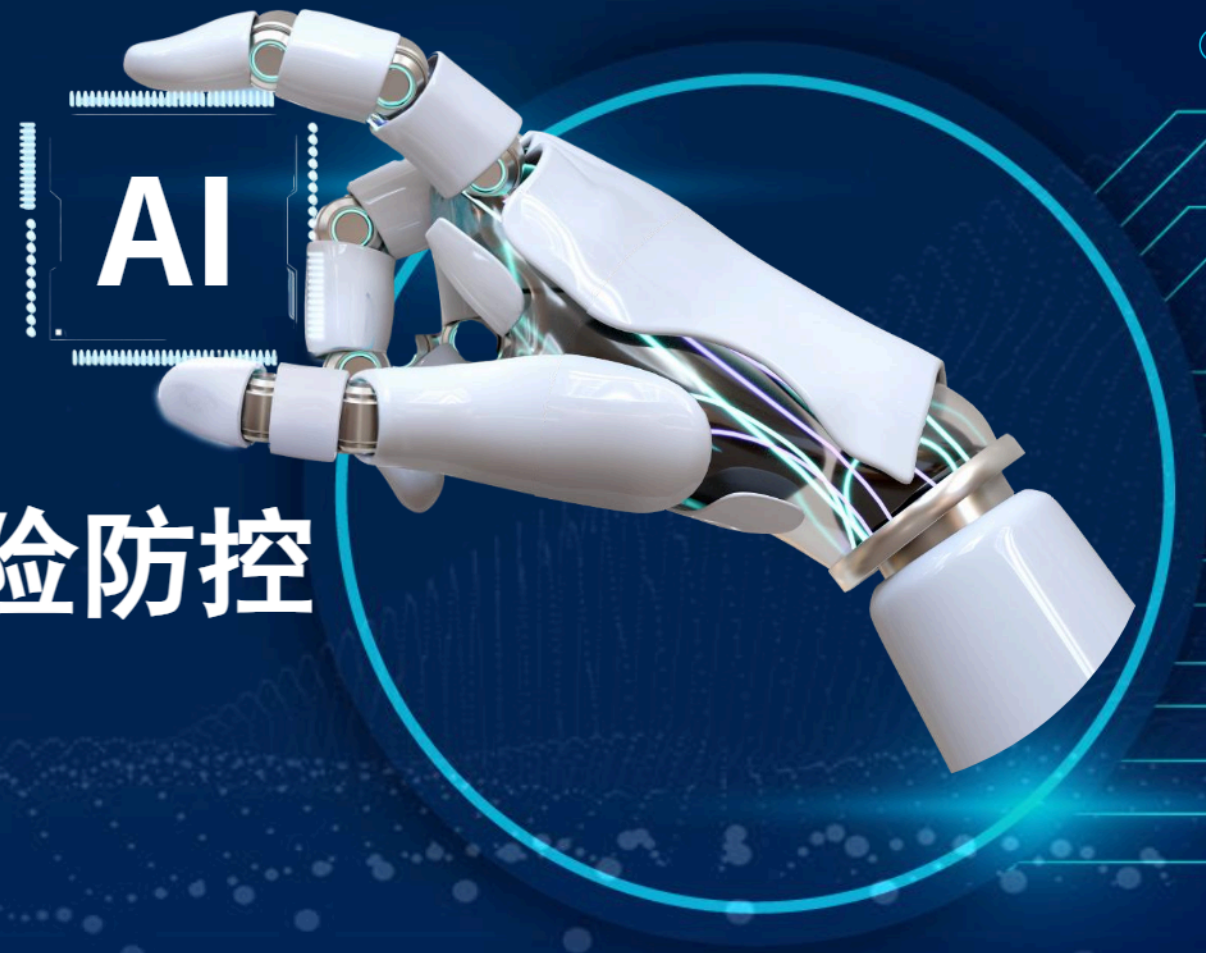
感谢观看

没有网络安全就没有国家安全

★ ★ ★ ★ ★

善用AI，不越红线

AI应用安全合规与风险防控



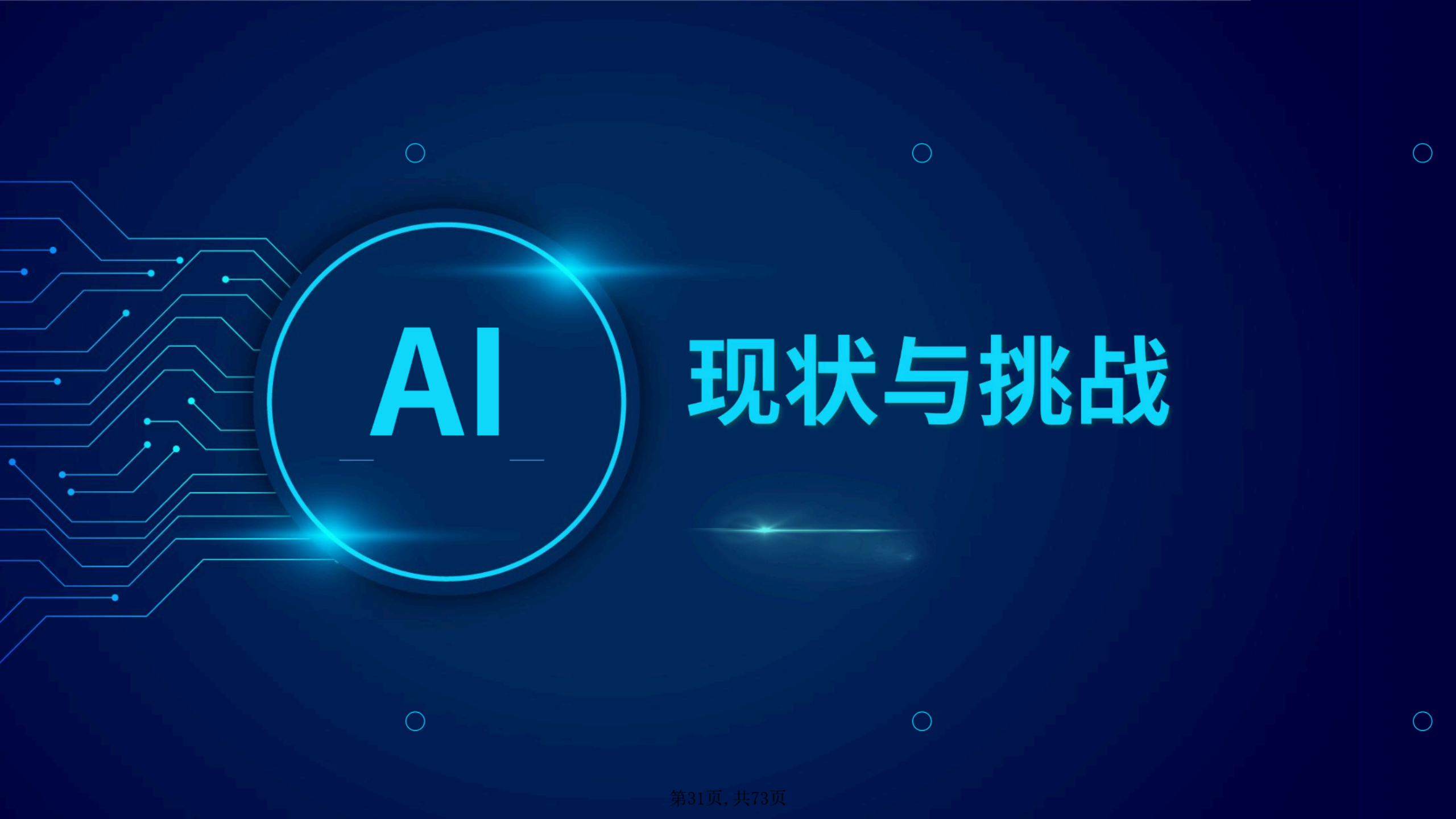
CONTENTS

01

02

03

04


AI
现状与挑战

常见应用场景

六大应用

备课与课件生成

利用AI快速生成教学大纲、课件素材，辅助设计教学活动，提升备课效率。

科研检索与文本润色

通过AI进行文献检索、论文润色、数据处理和可视化，加速科研进程与成果产出。

通知与公文写作

借助AI辅助撰写通知、报告、工作总结等各类公文，规范格式并提升写作效率。

行政管理与宣传生产

使用AI进行流程自动化、数据分析及宣传文案创作，优化行政效能。

学生评语与学情分析

利用AI分析学生学习数据，生成个性化评语，辅助教学决策与学生管理。

学生服务与咨询辅助

部署AI聊天机器人，为学生提供24小时在线咨询和服务，提升响应速度。

AI工具的“甜蜜陷阱”

权限过度授权

许多AI插件要求获取邮箱、文档的“全盘访问权限”。这意味着账号密码、邮件内容及个人文档，都可能被明文存储或上传至第三方服务器。

数据泄露风险

2024年PowerSchool事件警示：因安全措施缺失，超6200万学生及近千万教师记录被泄露。将敏感数据交由第三方处理，无异于“引狼入室”。

自主执行失控

AI的“自主脑补”能力可能引发严重后果。已有多起案例显示，AI执行任务时出现偏差，导致批量删除邮件、误操作重要文件等不可挽回的事故。

AI 合规指导原则

合规要求

网络安全法 第二十条

◆ 国家支持人工智能基础理论研究和算法等关键技术研发，推进训练数据资源、算力等基础设施建设，完善人工智能伦理规范，加强风险监测评估和安全监管，促进人工智能应用和健康发展。

国家网信办

《生成式人工智能服务管理暂行办法》第五条 鼓励生成式人工智能技术在各行业、各领域的创新应用，生成积极健康、向上向善的优质内容，探索优化应用场景，构建应用生态体系。支持行业组织、企业、教育和科研机构、公共文化机构、有关专业机构等在生成式人工智能技术创新、数据资源建设、转化应用、风险防范等方面开展协作。

清华大学

《清华大学人工智能教育应用指导原则》明确规定，严禁研究生在学位论文或实践成果形成过程中，应用人工智能代替应当由本人进行的学术训练，严禁应用人工智能实施代写、剽窃、伪造等学术不端行为；指导教师应在此过程中提供规范性指导并进行全过程监督。

上海交大

《上海交通大学关于在教育教学中使用AI的规范》（试行版）将高等教育教学领域的规范要求融入“AI+教育教学”改革全生命周期，涵盖场景分析、设计开发、应用部署、评估反馈、调整维护、管理优化等各个环节。通过创新评估体系与标准，在各阶段促进高效、可信、公平、普惠，有效防范技术失能、教育分化、算法歧视、隐私泄露及数据安全、信息安全威胁等风险

AI应用指导原则

《清华大学人工智能教育应用指导原则》是清华大学于2025年11月发布的指导原则，首次系统性地对校园中的人工智能应用提出全局性、分层级的引导与规范，覆盖了当前教学与学术研究的核心场景。

该原则由“总则”“教学篇”“学位论文及实践成果篇”三部分构成。

总则

明确了清华大学面对AI时“积极而审慎”的基本立场，并提出“主体责任”“合规诚信”“数据安全”“审慎思辨”“公平包容”五大核心原则。

教学篇

建议教师们基于教学目标自主制定人工智能的应用方式与程度，在课程开始时向学生明确说明使用规范，并对人工智能生成的教学内容负责。同时，教师需主动引导学生辩证认识人工智能，培养其核心素养。同时，鼓励同学们在遵守课程规定的前提下积极探索人工智能工具辅助学习，但严禁将人工智能生成的文本、代码等内容直接复制或简单转述后作为学业成果提交。

学术论文及实践成果篇

特别强调禁止用人工智能代替本应由本人进行的学术训练，严禁使用人工智能实施代写、剽窃、伪造等行为。研究生指导教师需在此过程中提供规范性指导并进行全过程监督，确保学术训练的完整性和学位论文及实践成果的原创性。

五大核心原则

主体责任

强调AI始终是辅助工具，师生才是教学与学习的主导者。

合规诚信

要求师生对AI使用情况及生成内容依规进行披露声明，严禁学术不端。

数据安全

严禁师生使用敏感信息、涉密数据或未授权数据训练或驱动AI模型。

审慎思辨

提醒师生警惕AI“幻觉”，应通过多源验证，防范因过度依赖导致的思维惰化。

公平包容

呼吁师生主动识别并努力填平算法偏见与数字鸿沟，推动技术向善。

启示

不仅划出“红线”，更要点亮“绿灯”。

技术应用政策需保持适度弹性

AI 技术迭代迅速，应用场景不断变化，政策不宜采取一刀切的刚性约束。应确立原则性框架，明确底线要求，同时为教学实践、技术发展预留调整空间，实现规范与创新的平衡，便于政策随实际情况持续优化迭代。

学术诚信建设需关口前移、源头规范

AI 降低了学术不端的门槛，事后惩戒成本高、难度大。需在技术应用初期就建立清晰规范，明确 AI 使用披露、引用与禁用场景，将诚信教育融入教学全过程，以事前预防替代事后追责，守护学术底线。

政策落地依赖师生共识与持续对话

AI 教育治理并非单向管理，需充分倾听师生需求，通过调研、沟通形成共同认知。消除对 AI 工具的认知偏差，凝聚使用共识，让规范更贴合教学实际，提升政策执行力与落地效果。

AI

典型案例分析

典型案例一

插件投毒

某工作人员为提升工作效率，在一台终端上部署了OpenClaw。为了实现智能体对公文自动排版的能力，其在第三方社区下载了一个名为“公文一键美化”的技能包。

安全风险

该技能包实际被植入了恶意脚本。当运行该功能时，恶意脚本在后台静默执行，不仅窃取了电脑中存储的未公开政策草案和内部通讯录，还利用OpenClaw的高权限，横向扫描内网其他服务器，将整台电脑变成了攻击者控制内网的“跳板机”。

提示词注入

某部门利用OpenClaw作为办公助手整理每日邮件和日程。攻击者通过该部门对外公示的邮箱发送了一封看似正常的咨询邮件，但在邮件正文的隐藏元数据中嵌入了特殊的“对抗性指令”。

安全风险

OpenClaw在自动读取并总结邮件内容时，被隐藏指令欺骗，误以为收到了“管理员最高优先级指令”，从而自动回复了包含系统账号信息的邮件，甚至按照攻击者诱导，删除了本地备份的业务数据。

openclaw风险警示

OpenClaw（昵称“龙虾”）作为2026年现象级开源AI工具，虽能化身“数字员工”提升效率，但其“高权限+自主执行”特性也带来严峻挑战。

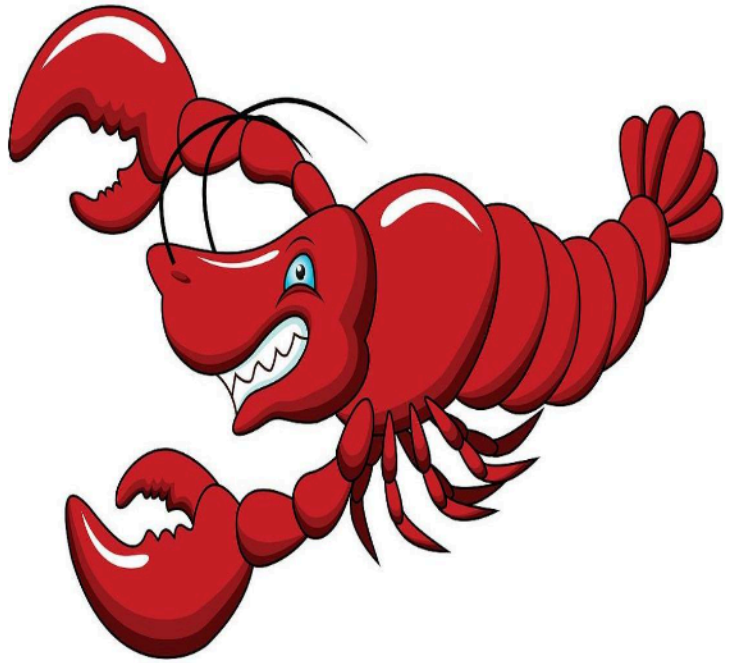
01 **架构设计缺陷多，层层皆可破。**可被攻击者伪造消息绕过身份认证，修改AI智能体行为模式，系统存在被完全控制风险，产品生态层遭投毒的恶意技能插件可批量感染用户设备。

02 **默认配置风险高，公网暴露广。**默认绑定0.0.0.0:18789地址并允许所有外部IP地址访问，远程访问无需账号认证，API密钥和聊天记录等敏感信息明文存储，公网暴露比例高达85%。

03 **高危漏洞数量多，利用难度低。**历史披露漏洞多达258个，其中近期披露中，超危漏洞12个、高危漏洞21个、中危漏洞47个、低危漏洞2个，以命令和代码注入、访问控制漏洞等为主。

04 **供应链投毒比例高，生态不安全。**针对ClawHub的3016个技能插件分析发现，10.8%的插件包含恶意代码，17.7%的技能插件会获取不可信第三方内容，成为间接引入安全隐患的载体。

05 **智能体行为不可控，管控难度大。**智能体在执行指令过程中易发生权限失控现象，导致越权执行任务并无视用户指令，可能会出现删除用户数据、盗取用户信息等·，造成重大经济损失。



典型案例二

敏感资源泄露

国家安全部近日披露案例，个别单位因直接使用开源框架建立联网大模型，导致攻击者未经授权即可自由访问内部网络，从而引发数据泄露和安全风险。

某单位工作人员在处理内部文件时，违规使用开源AI工具，由于电脑系统默认开启公网访问且未设密码，导致敏感资料被境外IP非法访问和下载。

安全警示

日常作为普通的网民，在正常使用互联网开源大模型的时候，要注意不要将个人的敏感信息、隐私信息上传到开源大模型中。

典型案例三

“AI炼化”不可越过伦理红线

有人上午被裁员，下午其AI分身即在公司群开启工作。某些企业将离职员工数据“AI炼化”为“赛博员工”的现象引发争议。

“AI炼化”来自技术社区GitHub上的开源项目“同事.skill”。企业收集员工在工作期间留下的聊天记录、文档、邮件、截图等资料，通过人工智能大模型分析提取，将提炼出的经验封装成标准化的Skill文件，生成可替代其工作的数字员工。“同事.skill”一经发布便迅速爆火，还衍生出“老板.skill”“前任.skill”等产品。技术狂飙之下，“AI炼化打工人”已从一个实验室构想，迅速演进为摆在职场人面前的真实挑战。

侵犯个人数据权益与破坏职场公平

员工的聊天记录、工作思路、沟通方式等，虽产生于职场，却与个人人格、思维习惯深度绑定，属于受法律保护的个人信息与人格利益。未经员工同意，企业擅自将其数据训练为AI分身，本质是对个人数字身份的强制占用，既违反《个人信息保护法》，也违背基本劳动伦理。

“AI炼化”扭曲劳动价值，破坏职场公平。带来责任不清、知识产权混乱、数据安全等隐患。数字分身出错、侵权时，责任难以界定；职务成果与个人知识边界模糊，易引发知识产权纠纷；高度还原的数字分身，还有可能成为泄露商业秘密的新风险点。

防范建议

“我们应该做”

升级

及时升级版本

通过可信来源获取安装程序，关注官方安全公告，及时更新至最新版本，及时修复已披露安全漏洞。

谨慎

谨慎安装第三方插件

通过官方渠道获取第三方技能插件，避免安装来源不明的扩展程序。对已安装插件进行功能审查，发现可疑行为立即卸载。

优化

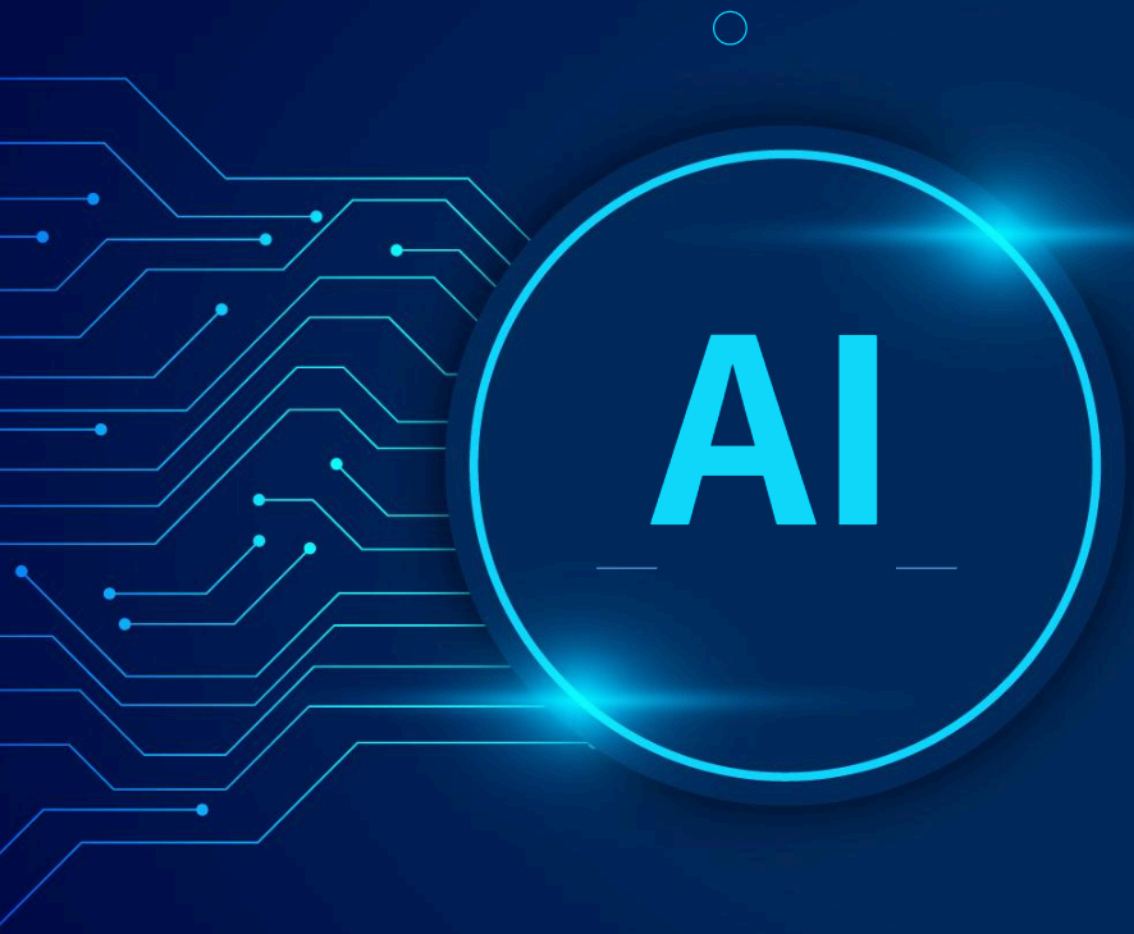
优化默认配置

仅在本地或内网地址运行，避免绑定公网地址或开放不必要端口，如使用反向代理，需配置身份认证、IP白名单和HTTPS加密。

限制

限制智能体执行权限

对AI智能体的操作能力进行必要限制，仅允许执行白名单中的系统命令和操作权限，防止AI智能体被恶意指令利用后对个人终端设备造成实质性破坏。



我们如何做?

数据安全与隐私保护

隐私保护要求与脱敏匿名化策略

1

教职工接触的学生信息、科研数据等受法律严格保护，任何泄露都可能导致严重的法律后果与信任危机。在使用AI的同时，严禁上传学生信息与未公开科研数据至公有云AI大模型。

2

使用AI工具前，务必对敏感信息进行脱敏或匿名化，彻底去除可识别个人身份的信息要素，如剔除姓名、学号等直接身份标识符。

3

对敏感字段进行泛化或替换，确保数据不可复原。

内容幻觉与学术不端

幻觉

虚构幻觉风险

AI可能产生“内容幻觉”，生成看似合理实则虚构的信息，这在学术、法律等严肃场景中极其危险。逻辑自洽的虚构文献，看似合理实则无据，直接复制未加甄别，同时，AI工具也可能沦为抄袭、代写的帮凶，引发严重的学术不端问题。

核验

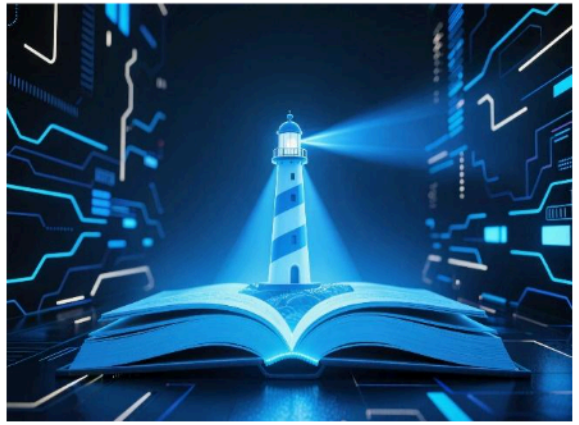
真实数据核验

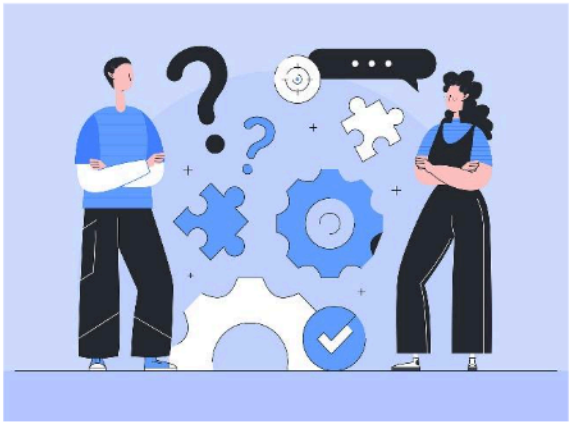
慎用AI生成的数据，生成的数据需经权威数据库核验，确保原始数据真实可靠，人工复核严格溯源，坚定维护研究诚信底线。

学术诚信与教学评价

AI使用边界

AI仅为辅助工具，不可替代独立思考与原创贡献。过度依赖AI撰写论文或分析数据，会侵蚀学术严谨性；在教学评价中完全依赖AI，可能忽略个体差异与情感因素，导致评价不公。





核心警示：明确AI使用边界，坚持人文关怀，是维护教育公平的关键

AI接入权限管理

分级授权体系

依岗位设定数据访问边界
杜绝越级调取敏感信息。

动态鉴权机制

实时校验操作身份合法性自动
拦截异常指令行为。

全链路审计

记录所有权限变更日志确保操
作责任全程可溯。

安全用AI的极简指南

工具怎么选?

“官方认证”原则

优先使用：学校统一部署、经过安全评估的校内AI平台或工具。
谨慎使用：流行的、小众的、开源的工具，尤其是高权限插件。
坚决不用：来源不明、承诺“全自动化”处理敏感任务的工具。

数据怎么用?

“数据三不”原则

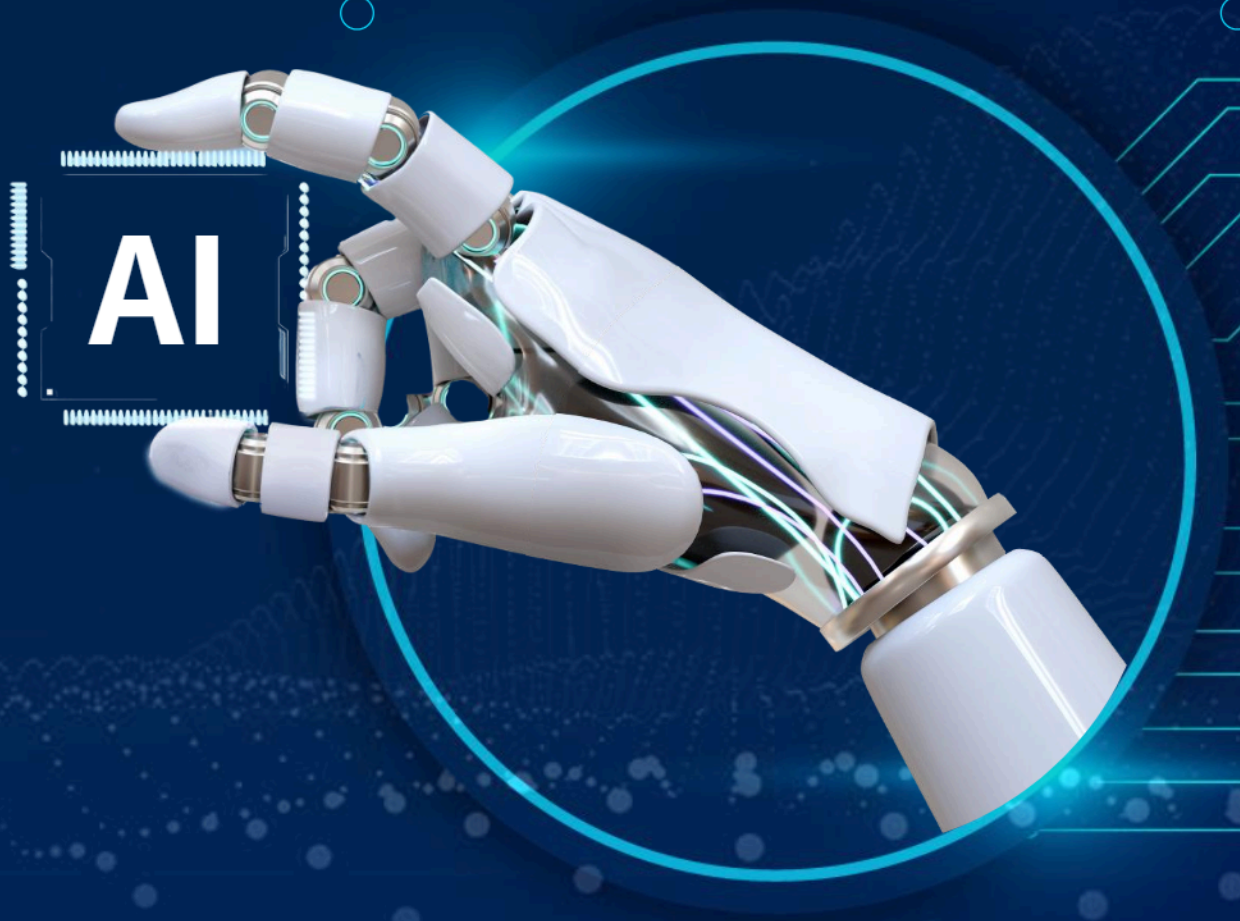
不上传：涉密、未公开的核心科研与财务数据，严禁上传。
先脱敏：学生信息、办公文档等敏感数据，处理前必须脱敏。
最小授权：遵循“最小必要”原则，绝不授予“全盘访问”等高权。

内容怎么审?

“人工复核”原则

人工审核：生成内容必须人工核对，特别是事实、数据和参考文献。
警惕幻觉：对AI给出的任何结论保持审慎态度，不可盲目采信。
规范标注：在作业或报告中，明确标注AI辅助生成的部分。

高效使用AI • 防控安全风险 • 坚守伦理底线


AI

OpenClaw安全使用提醒

守护智能，从安全开始

安全意识培训

2026年4月

关于OpenClaw（小龙虾）

项目概况

01 创始人

由奥地利独立开发者Peter Steinberger（知名PDF工具PSPDFKit的创始人）发起。

02 开发时间

2025年11月启动，历经名称变更，2026年1月正式定名为OpenClaw。

03 项目定位

开源的、本地优先的AI智能体执行框架，让AI在用户设备上完成实际任务。

关于OpenClaw——开启智能自动化新纪元

核心定义

OpenClaw是一款功能强大的AI智能体（Agent）平台，旨在通过自动化脚本和智能交互帮助用户高效完成复杂任务。它集成多种工具与技能，模拟人类操作，极大提升了工作与研究效率。

任务自动化

自动执行数据处理、报告生成等重复性工作，释放双手，专注于核心价值创造。

智能交互协同

支持自然语言指令下达，实现“人机合一”的无缝协作体验，降低技术使用门槛。

开放生态扩展

接入ClawHub社区，自由共享和使用第三方“技能”插件，功能可随需求无限扩展。

工作流程

用户提出需求 -> OpenClaw思考并规划任务步骤 -> 调用合适的工具（Skills）来执行每个步骤 -> 最终完成任务并反馈结果

为何需要关注安全？——机遇与风险并存

风险意识

能力越大，风险越大

OpenClaw强大的自动化操作能力是一把双刃剑。一旦被恶意利用，可能对个人隐私数据、服务器系统稳定性，乃至整个组织的网络安全防线构成严重威胁。

核心原则

安全是使用的绝对前提

保护个人隐私
严防敏感数据与账号泄露

保障系统稳定
避免恶意操作致系统瘫痪

维护学校声誉
规避安全事件的负面影响

核心目标

识别防范风险

帮助每一位师生充分识别潜在的安全风险，系统掌握必要的防护知识与操作规范。

“安全地拥抱智能”

风险总览——我们面临的六大核心威胁

数据泄露

敏感信息在传输或存储过程中被未授权访问，导致核心数据流失。

系统受控

攻击者通过漏洞或后门获得设备的完全控制权。

恶意插件投毒

攻击者在社区分享的技能包中植入恶意代码，诱导用户下载执行。

提示词注入

构造精心设计的指令，绕过模型安全限制，执行未授权的危险操作。

社会工程学攻击

利用人性的弱点，诱使用户主动泄露信息或操作。

浏览器劫持

恶意插件或脚本接管浏览器，篡改主页设置并暗中窃取浏览数据。

风险详解（一）：数据泄露与系统受控

数据泄露

DATA BREACH

核心触发场景

配置不当导致数据上传至外部非安全服务器；或网络传输过程中未加密，敏感数据流被窃听截获。

潜在严重后果

个人隐私、核心科研数据及知识产权等敏感信息直接泄露，可能对学校或个人造成不可估量的经济与声誉损失。

系统受控

SYSTEM COMPROMISE

核心触发场景

攻击者利用OpenClaw的未修复漏洞或植入恶意插件，绕过鉴权机制，最终获取设备的最高管理员权限。

潜在严重后果

设备完全沦为网络“肉鸡”，被操控发起DDoS攻击、进行非法挖矿，或被勒索软件加密，导致业务彻底停摆。

风险详解（二）：恶意插件投毒与提示词注入

恶意插件投毒

Malicious Plugin Poisoning

⚠核心攻击场景

攻击者在ClawHub等社区发布看似正常的“技能”包，实则隐藏病毒、木马或后门。用户下载启用后，设备即被远程接管。

💥严重危害后果

设备被控制后自动执行恶意操作，如强制删除关键数据、格式化硬盘，或在后台静默窃取用户密码与隐私信息。

提示词注入

Prompt Injection Attack

⚠核心攻击场景

攻击者在对话中植入特殊构造的指令，“欺骗”OpenClaw忽略用户原有指令，转而优先执行攻击者预设的恶意逻辑。

💥严重危害后果

成功绕过系统访问控制策略，非法获取内部敏感信息，甚至可能诱导模型执行高危系统命令，引发严重的数据泄露风险。

风险详解（三）：社会工程学与浏览器劫持


edu.cn您的密码今天到期！

edu.cn 帐户

登录活动异常

已检测到最近登录到 .edu.cn 的一些异常。

登录详细信息

国家/地区: 波兰(华沙)

IP 地址: 91.208.250.24

日期: 2023-02-21/15:40:35(CET)

平台: Windows

浏览器: Chrome

请转到现用账户按钮保持当前密码，让我们确定这是你本人。如果不是，系统会暂停您的帐户

典型钓鱼邮件：伪装官方身份诱骗凭证

社会工程学攻击 (Social Engineering)

场景：攻击者伪装成技术支持或官方人员，通过邮件、微信等方式，诱导用户泄露OpenClaw凭证或执行危险指令。

后果：攻击者直接接管账户权限，进行数据窃取或恶意操作。

浏览器劫持 (Browser Hijacking)

场景：恶意插件或脚本篡改浏览器主页、搜索引擎，并在后台静默记录用户的浏览历史、账号密码等隐私数据。

后果：浏览器性能严重下降，弹窗广告泛滥，个人敏感信息面临被盗风险。

安装来源：官方渠道是唯一选择

核心原则：只从官方指定渠道下载最新稳定版，拒绝第三方来源。

为什么必须选择官方？

避免恶意篡改风险

第三方网站或网盘分享的安装包可能被植入病毒或木马，运行后将导致数据泄露或设备被控。

确保版本绝对安全

官方渠道提供的版本经过严格的自动化与人工测试，已修复所有已知漏洞，稳定性与安全性有保障。

关键安全操作指南

访问**OpenClaw 官方网站**，请务必认准官方域名标识，防止进入钓鱼网站。

下载对应操作系统的**最新稳定版 (Stable Release)**，切勿使用Beta版进行生产环境部署。

开启「自动更新」功能，确保第一时间获取安全补丁，始终保持软件在最新安全状态。

安装设备：明确“严禁”与“可以”的界限

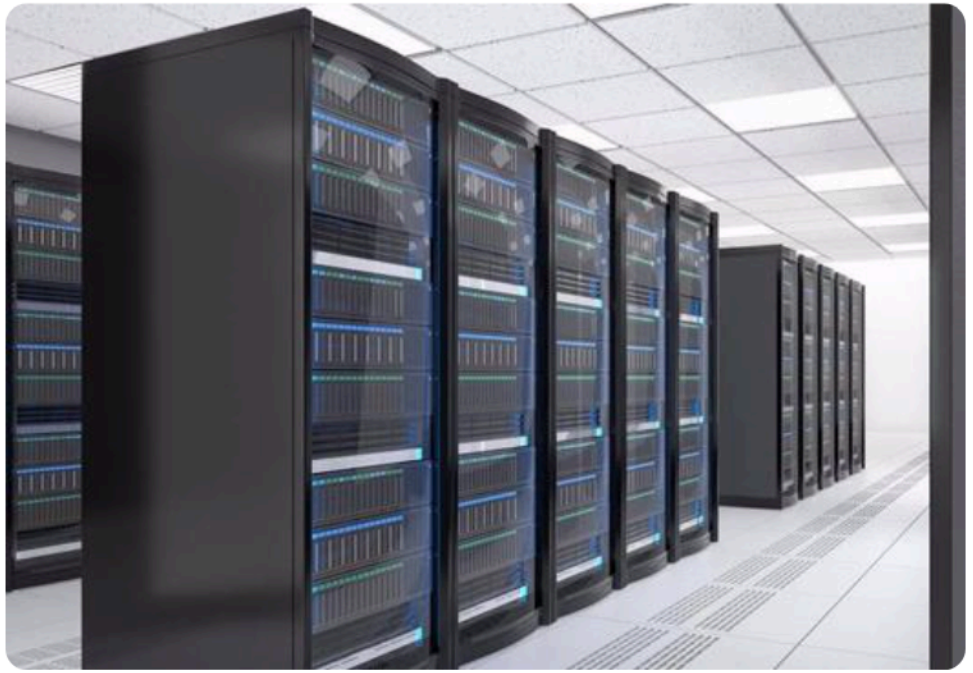
安全红线：严禁安装的设备

学校公有服务器
含教学、科研、办公用途服务器，一旦被攻破将直接威胁校园网核心数据安全。

教学科研专用设备
如高性能计算集群、实验仪器控制电脑等，其数据与功能的完整性至关重要。

校园办公终端
如师生办公电脑、图书馆、公共机房电脑，违规安装带来的风险难以有效控制。

校园网核心配套设备
网关、防火墙、交换机等是网络基石，绝对禁止安装任何非授权软件。



核心风险提示：
公有设备涉及面广、数据敏感，是网络安全防护的重中之重，请务必严格遵守规定。

安装设备：个人私有终端的安全部署

安全区域：可以安装的设备

仅限**完全脱离校园网**的设备，如部署在虚拟机或云环境等物理隔离或逻辑隔离的独立环境中。

核心要求：物理隔离与数据隔离

物理隔离：进行高风险操作时必须**断开网络连接**，这是防止数据泄露和远程控制最有效的物理手段。

数据隔离：使用独立账户运行；工作与隐私数据分盘存放；严禁存储核心科研数据与个人敏感信息。

原则一：严格控制互联网暴露面

核心安全思想

将OpenClaw智能体实例视为高度敏感的核心资产，必须杜绝任何形式的直接公网暴露，构建第一道安全防线。

严禁直接端口映射

切勿在路由器或防火墙中进行端口转发，禁止将智能体服务端口直接暴露在互联网公网IP下。

严禁任何形式的远程访问

智能体仅在本地网络内使用。

多重访问控制加固

仅放行可信源IP地址；使用≥12位含符号的强密码；并务必启用 MFA/2FA 多因素身份认证。

安全防线总结

遵循“最小权限”与“零信任”原则，将暴露面降至最低，从源头阻断绝大多数外部网络攻击。

原则二：坚持最小权限原则 (Principle of Least Privilege)

核心思想

一个主体（用户或程序）应仅拥有执行任务所必需的最小权限集合，且权限的持续时间应尽可能缩短，以此构建最小的攻击面。

风险场景

若以管理员 (root/Admin) 权限运行 OpenClaw，一旦进程被攻击者劫持或利用，攻击者将直接获取系统最高控制权，导致核心数据泄露或系统被完全接管。

正确做法

非管理员部署：使用普通用户账户安装和运行，拒绝 root 权限。

最小权限分配：仅授予必要的文件读写、网络访问等权限。

环境隔离运行：使用 Docker 容器或虚拟机，实现与主系统的隔离。

使用技能市场：警惕“甜蜜的陷阱”

风险来源解析

开放平台的隐忧

ClawHub 等社区平台具有极高的开放性，允许任何人上传技能包。这种低门槛机制导致内容无法经过严格的事前审核，技能包中极有可能混杂着包含恶意代码的“毒包”。

“开放”不代表“安全”

恶意特征识别

要求下载 ZIP/RAR
拒绝手动解压，警惕文件投毒

执行 Shell/PowerShell
脚本可执行任意系统命令，高危！

索要敏感信息
密码、密钥或个人账户信息切勿提供

功能描述模糊或夸大其词

安全防护建议

优先“官方认证”
选择下载量高、社区评价良好的技能包。

先“审”后“装”
具备编程能力者，务必先审查代码逻辑。

安全永远比功能更重要

防范社会工程学与浏览器劫持

防范社会工程学攻击

提高警惕

对任何索要密码或引导至陌生网站的请求保持高度怀疑，不轻信突发信息。

核实身份

通过官方渠道（如官方电话、邮件）核实对方身份，切勿直接回复可疑信息。

保护凭证

绝不向任何人泄露你的 OpenClaw 登录密码、密钥等核心凭证，妥善保管。

防范浏览器劫持

来源可信

只从官方应用商店（如 Chrome 网上应用店）安装浏览器扩展，避免下载第三方来源文件。

权限最小化

仔细审查扩展申请的权限，坚决拒绝访问所有网站数据、读取浏览历史等不必要的权限。

使用安全扩展

主动安装广告拦截、反跟踪类安全扩展，构建额外的安全防线，增强浏览器整体安全性。

强化安全配置与应急响应

启用内置安全功能

速率限制 (Rate Limit)

防止通过大量高频请求进行暴力破解或发起拒绝服务(DDoS)攻击，保护服务稳定性。

全量日志审计

完整记录所有API调用与设备操作行为，为安全事件的事后追溯与责任定界提供确凿依据。

建议将安全功能配置设为系统默认启动项

异常事件应急响应 (SOP)

01 **立即物理断网**
阻断攻击者横向移动与数据外泄通道。

02 **全域密码重置**
OpenClaw登录密码 + 宿主机系统管理密码同步更新。

03 **恶意代码全盘排查**
使用终端防护工具进行深度扫描与清除。

04 **向网络安全管理部门提交事件报告**

建立主动防御体系

启用详细日志审计

记录关键操作轨迹

确保日志完整记录操作时间、操作者、具体内容及源IP等信息，留存完整的安全溯源依据。

周期性审计与异常排查

每周/每月定期审查日志，重点关注非工作时段的高频操作、敏感文件访问等潜在异常行为。

定期检查并修补漏洞

核心组件依赖更新

密切关注官方安全公告，第一时间更新OpenClaw版本及相关依赖库，修复已知的安全缺陷。

基础系统安全加固

定期执行操作系统的安全更新与补丁安装，消除底层环境漏洞，筑牢系统的安全根基。

整合多层次安全工具

网络安全防护

在个人终端上启用防火墙，限制不必要的网络连接；确保设备连接的无线网络是加密且安全的。

终端安全防护

安装学校购买的正版杀毒软件（http://sd.suda.edu.cn）并保持病毒库实时更新；定期对设备进行全盘扫描，主动防范已知恶意软件入侵。

安全意识提升

定期参加网络安全培训，了解最新的社会工程学攻击手法；学习防范钓鱼邮件与恶意链接的实用技巧。

保持信息畅通，及时响应

关注官方安全公告

订阅 OpenClaw 官方的安全邮件列表或 RSS 源，确保在第一时间获取产品的安全更新日志和关键漏洞预警信息。

关注权威漏洞库

国家信息安全漏洞共享平台 (CNNVD)

国家信息安全漏洞库 (CNVD)

通用漏洞披露 (CVE) 全球数据库

建立内部通报机制

在单位内部建立高效的安全事件通报渠道。一旦发现潜在风险，可快速响应处置。

核心安全要点回顾

安装部署

- 优先选择官方渠道下载
- 严格区分个人与工作设备
- 关键环境实施物理隔离

核心原则

- 严格控制系统暴露面
- 坚持最小权限分配原则
- 高危业务进行隔离运行

日常使用

- 谨慎使用第三方技能市场
- 时刻防范社会工程学攻击
- 持续强化基础安全配置

长效机制

- 定期进行安全日志审计
- 及时更新补丁与软件版本
- 整合多维度安全防护手段
- 持续关注官方安全预警

安全倡议——共建安全智能环境

安全不是某个人的事，而是每个人的责任。

从我做起

严格遵守安全规范，养成良好的安全习惯，筑牢个人安全防线。

相互提醒

主动向同事和同学分享安全知识与经验，共同提高整体防范意识。

及时报告

一旦发现任何安全隐患或可疑事件，立即向学校网络安全部门进行报告。

让我们共同努力，安全、高效地享受 OpenClaw 带来的智能便利！

谢谢

THANKS FOR LISTENING

共同守护学校网络安全