头豹 LeadLeo

www.leadleo.com

中国AI在网络安全中的应用探析：AI攻防，新时代的安全利器

China Application of AI in Network Security Industry

中国ネットワークセキュリティにおけるAIの応用産業

概览标签：AI、网络安全、AI在网络安全中的应用

报告主要作者：林若薇

目录
CONTENTS

名词解释

- **机器学习：**机器学习（Machine Learning）是人工智能（Artificial Intelligence，AI）的一个子集，它专注于开发使计算机系统能够自动从数据中学习并改进其性能而无需明确编程的算法。通过从数据中寻找模式和洞察力，机器学习模型能够在没有显式编程的情况下做出决策或预测。
- **深度学习：**深度学习是机器学习的一个重要分支，是一种试图使用包含复杂结构或由多重非线性变换构成的多个处理层对数据进行高层抽象的算法。它基于人工神经网络（ANN）进行计算和训练，通过模仿人脑中的神经元连接和工作方式，实现对数据的有效表示和分析。
- **生成式AI：**生成式AI是一种能够生成新的、原创的数据、图像、音频、视频或其他形式的内容的技术。与传统的基于规则的程序或机器学习模型只能进行分类、识别或预测不同，生成式AI具有创造性，可以生成以前从未存在过的内容。
- 模型可检测性、可验证性：模型的可检测性是指通过某种方法或技术，能够系统地、全面地检测模型是否满足预定的要求或规范，包括模型的正确性、一致性、完整性和性能等方面。模型的可验证性是指通过特定的验证方法或流程，能够证明模型是否准确地描述了实际系统的行为或结构，并且其计算分析流程是否正确。
- **深度伪造：**深度伪造（Deepfake）是一种基于深度学习的人工智能技术，主要用于创建或修改音视频内容，使其看起来像是真实发生过的。
- **加密挖矿：**是一种处理加密货币交易并铸造新代币的过程。是加密货币（如比特币）网络中的一个关键环节，矿工使用计算能力来解决复杂的数学问题（或“难题”），以验证和记录交易，并获取加密货币奖励。
- **流量操纵：**流量操纵是指通过一定手段，在网络平台上进行人为干预，控制网站或应用程序的流量，使得特定的内容或服务得到更多的曝光和访问。
- **DDoS攻击**：DDoS攻击（Distributed Denial of Service Attack），即分布式拒绝服务攻击，是一种利用分布式网络来发起大量的请求，以占用目标服务器或网络资源的攻击行为。这种攻击方式的主要目的是通过瘫痪目标系统，使其无法正常提供服务。
- **Log分析：**Log分析，也称为日志分析，是指对系统、应用程序或服务生成的日志文件进行详细检查和解读的过程。这些日志文件通常记录了系统或应用程序在运行过程中产生的各种事件、状态变化、错误消息和其他相关信息。

第一部分：发展现状概述

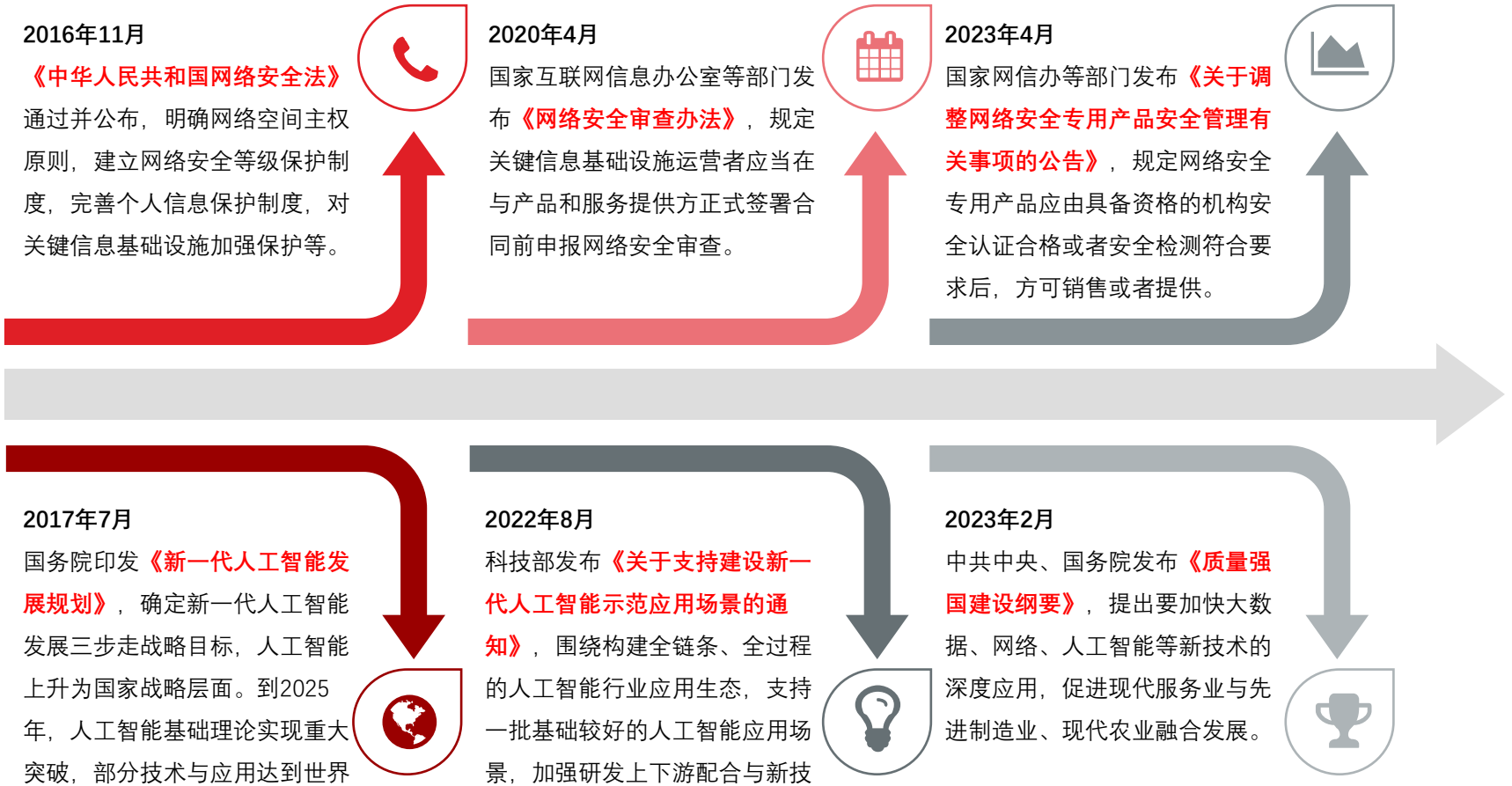
主要观点：

- 政策层面：AI技术通过自动化和智能化的方式，能够有效地提高威胁检测的速度和准确性，从而加快对安全事件的响应时间并提升处理效率
- 市场层面：网络安全事件频发，全球网络安全投入持续增加，行业规模稳定增长；生成式AI和多模态人工智能推动AI应用商业化加速，互联网、金融等行业广泛应用，中国AI行业规模持续增长
- AI网络安全：一方面，AI技术能够显著提高网络安全防御能力，通过自动化和智能化的手段增强对潜在威胁的识别和响应速度；另一方面，AI也可能被恶意利用，加剧网络安全威胁
- AI在网络安全中的应用：机器学习、生成式AI和大模型在网络安全领域广泛应用，AI技术通过高级数据分析增强检测系统，降低误报率和人力成本，提高威胁识别与预测效率，在网络安全领域的应用潜力大

AI在网络安全中的应用——政策层面

AI技术通过自动化和智能化的方式，能够有效地提高威胁检测的速度和准确性，从而加快对安全事件的响应时间并提升处理效率

网络安全和人工智能相关政策



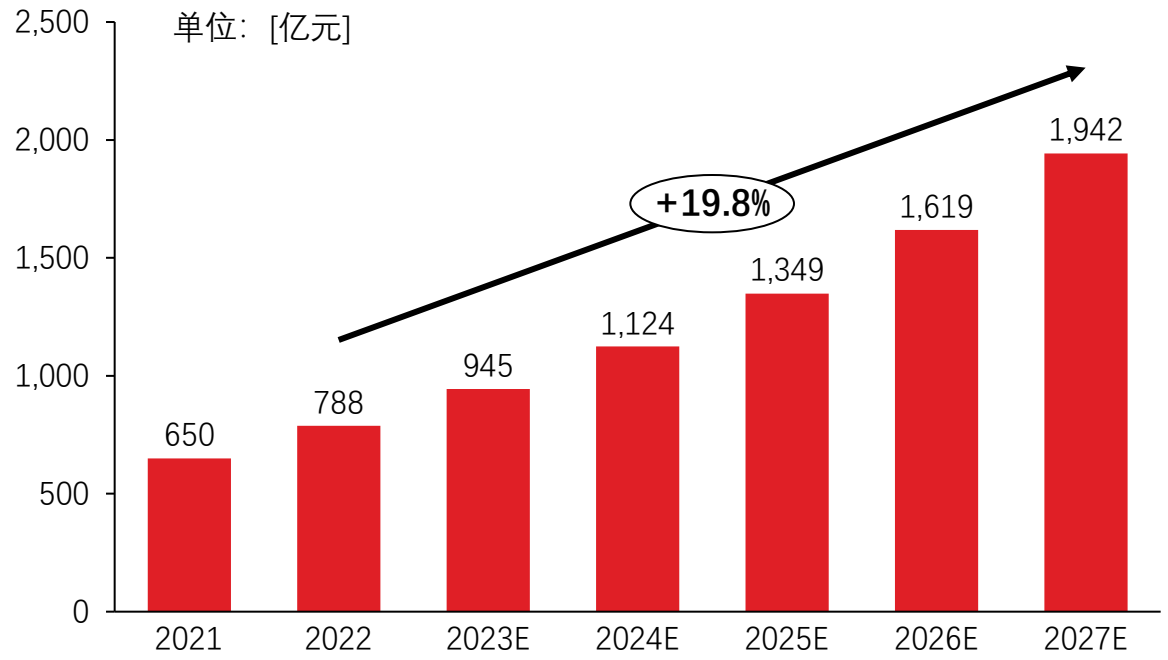
- 2024年**全国两会《政府工作报告》**中**“人工智能+”**行动引发热议，这是“人工智能+”首次被写入政府工作报告中，“AI+安全”也屡被提及。
- 在**《关于鼓励兼具“安全和AI”能力的企业解决通用大模型安全问题的提案》**中，阐述了解决通用大模型安全问题的必要性和紧迫性。
- **《关于创新发展“AI+安全”护航中国式现代化的提案》**则建议要大力探索“AI+安全”创新应用，抢占国家安全的人工智能战略制高。

来源：政府官网，头豹研究院

AI在网络安全中的应用——市场层面

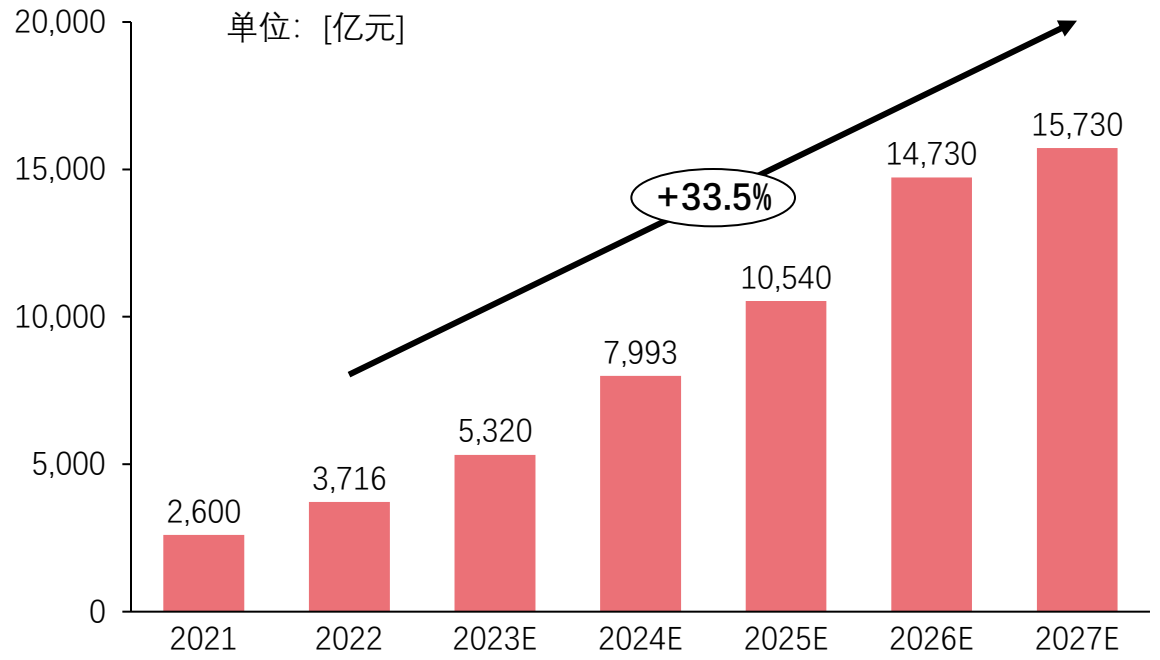
网络安全事件频发，全球网络安全投入持续增加，行业规模稳定增长；生成式AI和多模态人工智能推动AI应用商业化加速，互联网、金融等行业广泛应用，中国AI行业规模持续增长

中国网络安全市场规模，2021-2027预测



在0 day攻击、ATP攻击、DDoS攻击等网络安全事件频发催化下，网络安全已上升至国家战略高度，企业与个人信息安全备受各界重视，并着力部署网络安全防御策略，伴随数字经济的发展、物联网建设的逐步推进，网络安全作为数字经济发展的必要保障，其投入将持续增加。2022 年全球网络安全技术总支出711 亿美元，同比增长 15.8%。其中网络安全是收入最多的安全产品，增长了8.9%，达到了 54 亿美元。2022年，中国网络安全行业规模达到788亿元，预计2027年，行业规模将增长至1,942亿元。

中国人工智能市场规模，2021-2027预测



随着生成式AI的兴起和多模态人工智能的发展，人工智能的应用商业化落地进程不断加速，行业应用场景不断拓宽。目前，互联网、金融、政府、电信和制造业是人工智能应用最为广泛的行业。此外，中国政府在全球范围内都在积极推动AI技术的发展和应用。例如，中国政府提出了加快推动人工智能发展的政策措施，并在国际合作中发布了《全球人工智能治理倡议》，推动行业发展。2022年，中国人工智能行业规模达到3,716亿元，预计2027年，行业规模将增长至15,730亿元。

来源：头豹研究院

AI在网络安全中的应用——AI+网络安全

一方面，AI技术能够显著提高网络安全防御能力，通过自动化和智能化的手段增强对潜在威胁的识别和响应速度；另一方面，AI也可能被恶意利用，加剧网络安全威胁

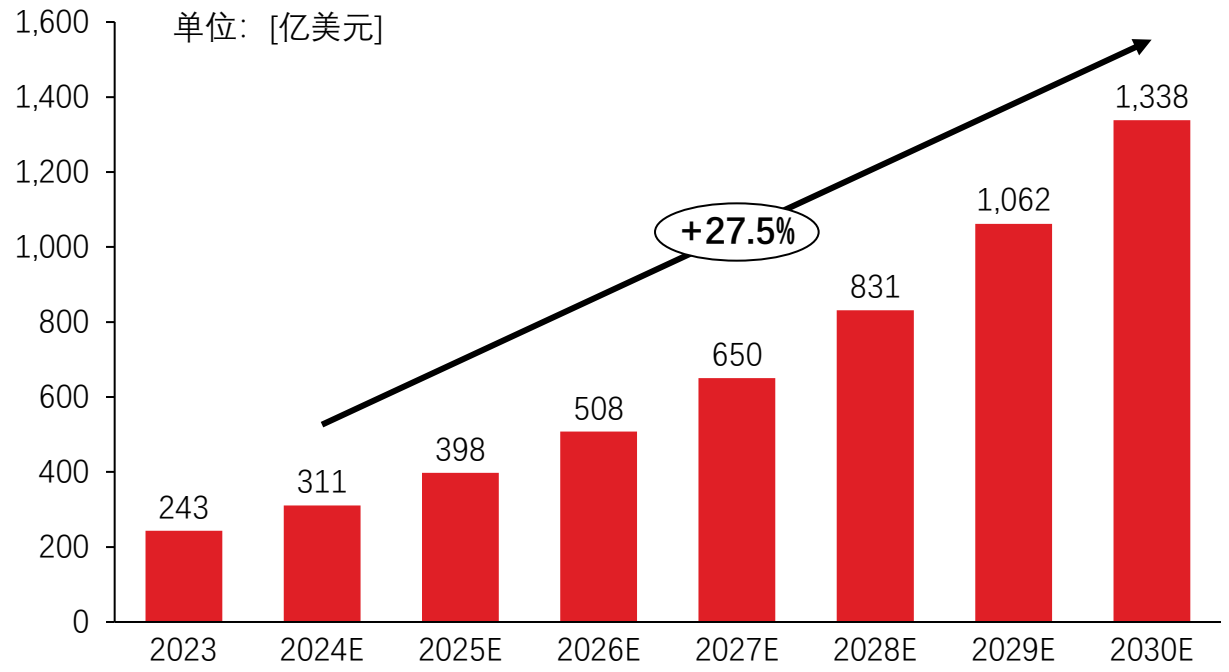
AI在网络安全中呈现的“两面性”

来源：CSDN，《中国工程科学》，头豹研究院

AI在网络安全中的应用——AI在网络安全的应用

机器学习、生成式AI和大模型在网络安全领域广泛应用，AI技术通过高级数据分析增强检测系统，降低误报率和人力成本，提高威胁识别与预测效率，在网络安全领域的应用潜力大

全球AI网络安全市场规模，2023-2030预测



随着机器学习技术、生成式AI和大模型的发展，其在网络安全领域的应用越来越广泛。AI技术能够通过高级数据分析功能，增强早期检测系统的能力，从而更有效地识别和预测网络威胁。2022年一项调查显示，使用AI驱动式入侵检测系统的企业的误报率降低了43%，AI驱动式的电子邮件安全解决方案则最多可将误报率降低70%。2023年，全球AI网络安全市场规模为243亿美元，预计2030年规模将达1,338亿美元。未来，AI将在威胁发现、调整防御和确保可靠的数据备份等方面发挥重要作用。

全球公司数据泄露平均成本、识别时间， 2023年

未基于AI的网络安全		基于AI的网络安全
泄露成本：445万美元	↓ 40%	泄露成本：265万美元
威胁识别时间：277天	↓ 36%	威胁识别时间：177天

69%的企业表示，由于人类分析师无法应对数量庞大的网络威胁，因此网络安全中应用人工智能是必要的。IBM Global报告的受访高管则表示，其组织2023年的AI网络安全预算相比2021年增加了**51%**。而且，他们预计到2025年，这一预算将再增加**43%**。

来源：Techopedia，IBM Global，Ponemon Institute，头豹研究院

第二部分：AI在网络安全中的应用

主要观点：

- AI技术应用总览：AI技术具备的速度、规模、范围优势，在数据处理和自适应防护、自动化和实时监控、优化资源分配、提升安全运营效率等方面赋能于网络安全，使得防御策略更加高效、自动化、智能化
- AI赋能检测与响应：AI技术通过自动化和智能化的方式，能够有效地提高威胁检测的速度和准确性，从而加快对安全事件的响应时间并提升处理效率
- AI赋能安全运营：安全运营领域自动化程度低，自动化和智能化成为提升效能关键，如大模型技术能增强知识语义、改善研判，提高人机协作效率，降低时间和人力成本，推动安全运营更高效应对威胁
- AI安全大模型：AI大模型在网络安全中应用日益广泛，目前处于第二层级，协助处理特定业务；未来有望达第三层级，实现全面网络防御

AI在网络安全中的应用——AI技术应用总览

AI技术具备的速度、规模、范围优势，在数据处理和自适应防护、自动化和实时监控、优化资源分配、提升安全运营效率等方面赋能于网络安全，使得防御策略更加高效、自动化、智能化

AI在网络安全中的应用总览

自动化渗透测试

通过自动化测试工具和机器学习算法，模拟黑客攻击行为,快速发现安全漏洞。

安全运营

AI可以辅助安全分析人员分析安全数据，自动完成安全事件的分析和响应，提高安全运营效率。

安全知识库

整合安全漏洞信息、攻击手法和防御策略，利用AI技术进行知识管理和智能检索。

流量检测

AI可以协助分析网络流量数据来识别异常流量和恶意行为，及时发现并阻止潜在的安全威胁。

异常行为分析

分析用户和设备行为，学习用户正常行为模式，识别出异常行为，进行风险评估和预警。

安全数据接入

利用AI技术实现安全数据的自动化采集、处理和存储，为安全分析和决策提供可靠数据支持。

可视化图表分析

利用AI技术对安全数据进行可视化分析，将复杂的安全信息以直观图表形式展现出来。

数据分类分级

利用AI对数据进行分类和分级，根据数据的重要性和敏感性制定相应的安全策略和控制措施。

产品资料编写

AI可自动生成安全产品资料和文档，包括安全漏洞报告、产品白皮书等，提高安全团队工作效率。

网络靶场

模拟网络攻击和防御场景，提供个性化训练，帮助安全人员进行实战演练和技能培训。

来源：51CTO，《计算机与自主智能研究进展》，头豹研究院

AI在网络安全中的应用——检测与响应

AI技术通过自动化和智能化的方式，能够有效地提高威胁检测的速度和准确性，从而加快对安全事件的响应时间并提升处理效率

AI在检测与响应中的应用

1. 自动化威胁分析
2. 深度学习应用
3. 多模态数据处理
4. 对抗性样本的防御
5. 动态调整安全策略
6. 自然语言处理
7. 生成式AI应用
8. 知识图谱
9. ……

AI在检测与响应中的应用

恶意域名系统检测

AI应用：N-gram 模型、LSTM 模型，随机森林、KNN、马尔科夫链等算法
应用效果：可对随机生成的DGA 域名进行有效把控和识别，取得更高的检测准确率和更低的误报率。

异常检测

AI应用： UEBA 技术、安全知识图谱、序列异常挖掘算法等
应用效果：处理和分析海量数据，通过建立用户画像和资产画像，与历史数据进行比较，发现异常行为。

骚扰诈骗电话检测

AI应用： 人工智能助手、DNN模型、RNN模型等
应用效果：实时调整模型并进行对抗识别，可自学习参数，实现营销电话柔性拒绝，提高欺诈电话的识别率与覆盖率。

手机恶意软件检测

AI应用：增压模型、深度神经网络等算法
应用效果：提取有效特征区分手机恶意软件和正常软件，提高手机恶意软件检测的自动化、智能化水平。

威胁情报

AI应用： 知识图谱、 NLP 中的命名实体识别任务、BERT 预训练模型等
应用效果：从海量信息提取出高价值威胁特征，提升对 APT 组织的分析追踪能力。

态势感知

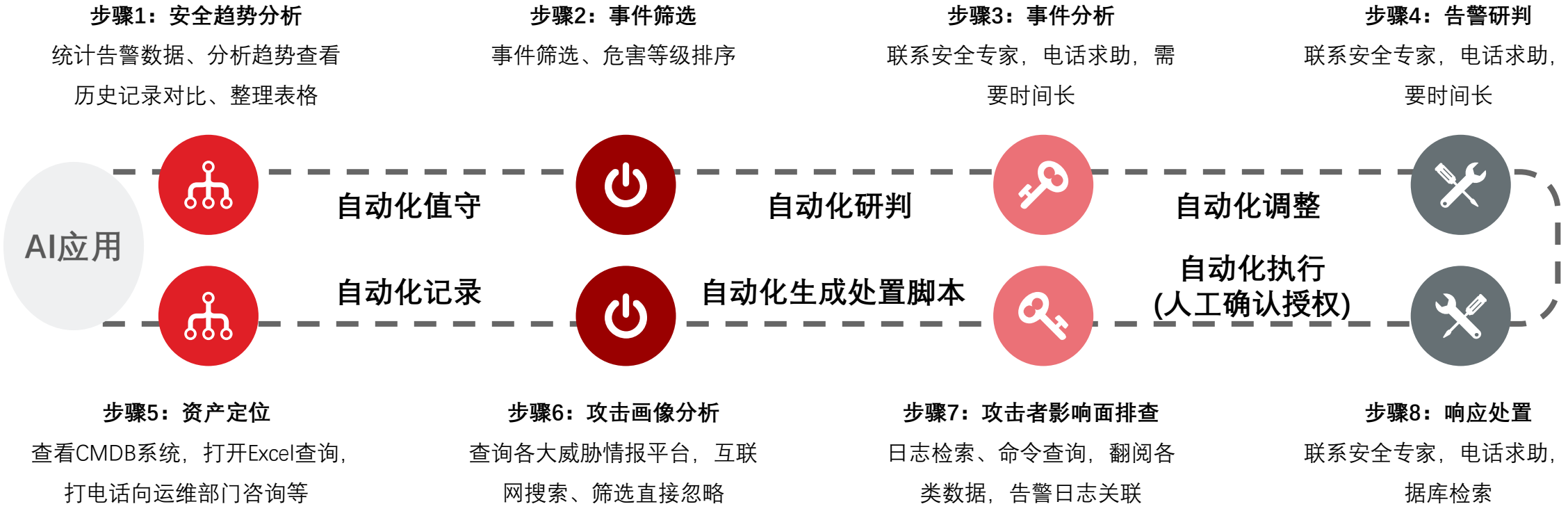
AI应用：随机森林、增压模型、二分类模型等
应用效果：综合提取样本的统计特征、概率特征及 NLP 特征并进行样本描述，自动化降低规则生成周期，提高训练出的判定模型准确度。

来源：GSMA，头豹研究院

AI在网络安全中的应用——安全运营

安全运营领域自动化程度低，自动化和智能化成为提升效能关键，如大模型技术能增强知识语义、改善研判，提高人机协作效率，降低时间和人力成本，推动安全运营更高效应对威胁

安全事件分析处置全流程



安全运营领域是自动化程度较低的领域。安全事件的频繁发生和种类多样性要求安全运营能够快速响应并处理各种威胁，然而，现有的自动化技术在处理复杂的安全事件时仍存在局限性，如告警无效、误报率高、缺乏研判上下文等问题。此外，安全运营中心面临的核心问题包括大规模告警、24小时运营周期以及高要求的运营指标，这些都直接影响了安全运营的效率和状态。因此，自动化和智能化被视为提升安全运营效能的重要方向。例如，基于大模型的重建可以增强知识语义、改善任务分析研判，并提高人机协作效率，可以降低安全运营的时间和人力成本。

来源：国脉电子政务网，头豹研究院

AI在网络安全中的应用——AI安全大模型

AI大模型在网络安全中应用日益广泛，目前处于第二层级，协助处理特定业务；未来有望达第三层级，实现全面网络防御

网络安全大模型的发展与现状

A1 1.0/小模型

技术与特点

- **核心技术：**行为分析、小模型、机器学习
- **特点：**能够快速预测和检测未知威胁，但小模型只能用于单一的检测场景

主要产品

态势感知	NDR	TIP
SIEM	UEBA	SOAR
XDR	EDR	AIOPs

AI 2.0/大模型

技术与特点

- **核心技术：**生成式AI、大模型
- **特点：**能够协助检测和分析，达到较为高效的场景化安全应用

主要产品

- 安全大脑
- 安全Copilot
- AI安全平台

网络安全大模型的发展

市面上的安全大模型

- 奇安信：Q-GPT安全机器人
- 安恒信息：恒脑·安全垂域大模型
- 360：360安全大模型
- 绿盟科技：风云卫大模型
- 深信服：安全GPT
- 启明星辰：盘小古安全大模型
- 亚信安全：信立方安全大模型
- 长亭科技：问津安全大模型
- 微步在线：情报智脑XGPT
- ……

在过去十年，人工智能的机器学习技术已被厂商广泛应用于网络安全产品中，反病毒、垃圾邮件过滤和钓鱼检测工具等产品的自动化和智能化水平大幅提升，而“智能”网络防御成为网络安全行业的基本营销要素。2023年，生成式AI、大模型技术大爆发，其技术价值正在多产业持续释放。国内众多网络安全厂商也分别聚焦于安全运营、威胁情报、应用程序安全、数据分类分级等应用场景，在2023年陆续推出了各具特色的网络安全垂直领域大模型。专家表示，若将AI大模型在网络安全中的应用分为三个层级，目前的应用情况属于第二层级，即是面向工业应用的，这类模型可以自动处理特定的业务，协助安全专家处理行业任务。而未来第三层级则是直接替代传统的网络安全产品，具备全面网络防御功能。

来源：AI科技评论，企业官网，头豹研究院

第三部分：AI网络安全未来展望

主要观点：

- 应用趋势：AI在网络安全中将广泛应用，通过实时威胁分析、增强网络防御、生成式AI应用、自适应防御、智能对抗AI及零信任微隔离，有效应对复杂威胁，实现高效安全防护与持续验证权限管理
- 网安大模型发展痛点：网安大模型发展面临数据安全与隐私、基础模型脆弱性、内容合规性、模型记忆风险、算法不透明、新兴威胁及外部和内部安全威胁等多重挑战

AI在网络安全中的应用——应用趋势

AI在网络安全中将广泛应用，通过实时威胁分析、增强网络防御、生成式AI应用、自适应防御、智能对抗AI及零信任微隔离，有效应对复杂威胁，实现高效安全防护与持续验证权限管理

发展趋势

01

实时威胁分析与自动化响应

AI技术能够处理和分析海量数据，从而更准确地识别潜在的威胁，并通过自动化响应系统迅速采取措施。

02

自适应网络安全防御系统

基于深度学习、监督学习和强化学习等算法的自适应网络安全防御系统能通过实时分析网络流量和行为模式来自动调整安全策略。

03

增强网络防御

AI的自主学习和数据分析能力使得它在发现未知威胁及高级持续性威胁（APT）方面具有显著优势，还能提升对新型威胁的适应能力。

04

AI对抗AI

生成式AI技术的出现，使得网络攻击者可以更快地开发出新的攻击方法。因此，为了应对日益复杂的网络威胁，通过模拟攻击者行为来测试和改进防御系统变得极为重要。

05

生成式AI的应用

通过对历史威胁数据的分析，生成式AI可以预测未来的威胁。在漏洞发现上，通过学习和分析大量代码和应用程序，可以自动发现潜在的漏洞。

06

零信任微隔离

零信任模型强调每次访问都要重新验证用户的身份和权限，以及限制用户的访问权限到最低必要水平。AI可以通过自动化的方式实现持续验证和权限管理。

来源：深圳市人工智能产业协会，头豹研究院

AI在网络安全中的应用——网安大模型发展痛点

网安大模型发展面临数据安全与隐私、基础模型脆弱性、内容合规性、模型记忆风险、算法不透明、新兴威胁及外部和内部安全威胁等多重挑战

网安大模型发展痛点

数据安全与隐私

大模型在训练和部署过程中需要处理大量敏感数据，如何保护这些数据不被未授权访问或泄露是一个重大挑战。

内容合规

AIGC的内容合规性也是一个重要问题，需要确保生成的内容符合法律法规和社会伦理标准。

算法透明度和可解释性

由于AI模型通常是“黑箱”式的，其决策过程缺乏透明度，这可能导致用户对AI系统的信任度下降。

外部攻击安全问题

包括数据投毒、模型后门、对抗样本、数据泄露、模型窃取、软件漏洞等安全隐患。

基础模型的安全性影响

基础模型的脆弱性可能被下游模型继承，如果基础模型对部分训练数据进行了“记忆”，则下游模型也面临相同的风险。

模型记忆风险

经过模型的训练和推理后，模型可能会形成对特定输入的“记忆”，这可能导致模型输出偏见或错误。

新兴威胁场景

随着AI技术的发展，新的威胁场景不断出现，如智能生成虚假新闻等，这些都需要在安全策略中加以考虑。

模型流转/部署过程中的安全问题

在大模型的非私有化预训练、精调和推理服务过程中，数据需要在不同的主体或部门之间传输，这增加了传输截获的风险。

来源：经济观察网，新京报，头豹研究院

方法论

- 头豹研究院布局中国市场，深入研究19大行业，532个垂直行业的市场变化，已经积累了近100万行业研究样本，完成近10,000多个独立的研究咨询项目。
- 研究院依托中国活跃的经济环境，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

法律声明

头豹业务合作

会员账号

可阅读全部原创报告和百万数据，提供PC及移动端，方便触达平台内容

定制报告/词条

行企研究多模态搜索引擎及数据库，募投可研、尽调、IRPR等研究咨询

定制白皮书

对产业及细分行业进行现状梳理和趋势洞察，输出全局观深度研究报告

联系我们

业务热线

袁先生：15999806788

李先生：13080197867

招股书引用

研究覆盖国民经济19+核心产业，内容可授权引用至上市文件、年报

市场地位确认

对客户竞争优势进行评估和证明，助力企业价值提升及品牌影响力传播

云实习课程

依托完善行业研究体系，帮助学生掌握行业研究能力，丰富简历履历