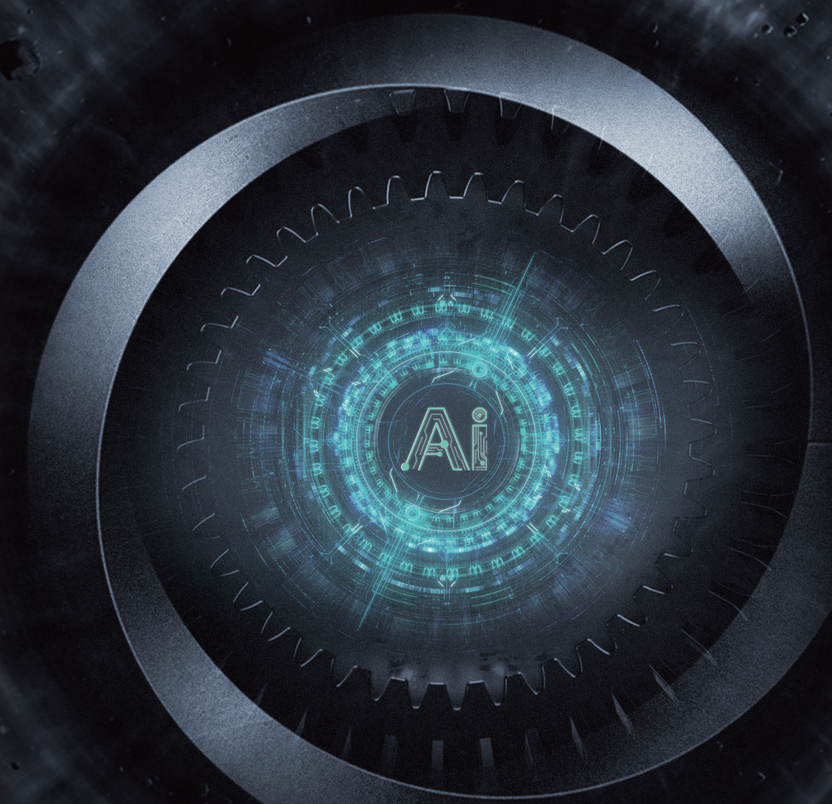
2024 全球AI网络安全产品洞察报告

2024 Global AI Cybersecurity Product Insights Report



FreeBuf 咨询　天融信 TOPSEC　蚂蚁集团 ANT GROUP　瑞数信息 RIVER SECURITY

前 言

人工智能（以下简称“AI”）是未来发展的重要技术。随着机器算法、算力地不断突破，AI技术进入飞速发展阶段。同时，云、管、端、边等基础设施持续深入融合，在以5G、物联网、数字化等新兴技术为核心的“新基建”的加持下，万物互联和数据积累持续加速，全面推动AI技术广泛应用于千行百业，并且产生了具有革命性的深远影响。

网络安全正是其中之一，且天然适合AI技术应用的行业。目前AI技术在网络安全行业应用广泛，包括网络行为与威胁分析、安全事件管理、入侵检测、安全漏洞治理、端点保护、安全知识问答、数据与文件分类等多个方面。

更重要的是，AI对于传统网络安全产品领域也产生了重大影响，对于未来安全产品的形态和发展趋势有着不可替代的作用，以“AI”为核心的安全产品时代正在到来。

目前，全球网络安全厂商正在持续发力AI安全产品的研究与落地，将AI的能力深度融合于安全产品体系之中，推动网络安全行业智能化发展。

2024年3月，Microsoft Security Copilot正式发布，是全球首款真正意义上的AI安全产品，借助微软庞大的全球威胁情报和每天数十万亿个信源提供的信息，以快速检测、响应来帮助企业更好地应对当下日益严峻的网络安全形势。

例如用户可以向 Security Copilot 询问特定时间段内的可疑用户登录情况；甚至可以使用它来创建概述事件及其攻击链的 PowerPoint 演示文稿；还可以接受文件、URL 和代码片段进行分析。同时，Security Copilot还提供了可视化的工具，安全人员能够让机器人一键生成演示文件，展示安全威胁的路径。

Microsoft Security Copilot的发布正式拉开了全球AI安全产品的浪潮。国外知名网络安全公司CrowdStrike、Palo Alto Networks、Fortinet等纷纷跟进，国内网络安全厂商也不断发布对应的AI大模型，时至今日，我们已经看到不少安全产品已经具有“AI”的能力。

FreeBuf咨询的《全球AI网络安全产品洞察报告》重点研究网安行业AI安全产品的市场规模、商业模式、发展挑战等，对国内外AI安全产品的竞争力进行全面评估，展望当前国内外AI安全产品的发展阶段。**本报告中的AI安全产品，主要有三种呈现形式，分别是AI安全助手产品、AI驱动的安全平台、AI赋能安全产品，其中包括针对AI大模型自身的安全产品和解决方案。**本报告将为市场提供网安行业AI安全产品格局的全面、客观的见解。

AI安全任重道远。作为网络安全行业发展的观察者，FreeBuf咨询对于AI安全的发展将保持常态化观察，也欢迎业界同仁共同参与进来，发现AI安全的价值。

关于报告

AI对于传统网络安全产品领域也产生了重大影响，对于未来安全产品的形态和发展趋势有着不可替代的作用，以“AI”为核心的安全产品时代正在到来。全球网络安全厂商正在持续发力AI安全产品的研究与落地，以及将AI的能力深度融合于安全产品体系之中，推动网络安全行业智能化发展。

FreeBuf咨询的《全球AI网络安全产品洞察报告》重点研究网安行业AI安全产品的市场规模、商业模式、发展挑战等，对国内外AI安全产品的竞争力进行全面评估，展望当前国内外AI安全产品的发展阶段。本报告中的AI安全产品，主要有三种呈现形式，分别是AI安全助手产品、AI驱动的安全平台、AI赋能安全产品，其中包括针对AI大模型自身的安全产品和解决方案。本报告将为市场提供网安行业AI安全产品格局的全面、客观的见解。

关于FreeBuf咨询

FreeBuf咨询集结安全行业经验丰富的安全专家和分析师，常年对网络安全技术、行业动态保持追踪，洞悉安全行业现状和趋势，呈现最专业的研究与咨询服务，主要输出四个种类的咨询报告：行业研究报告、能力评估报告、产品研究报告以及甲方定制化报告。

FreeBuf咨询自成立以来，已积累了1500+ 甲方安全智库资源，为行业研究报告、企业咨询服务提供指导。访谈上百位行业大咖，为业界输出真实、丰富的安全管理价值与实践经验，具备超过80万+ 精准用户，直接触达CSO、企业安全专家、投资人等专业人群。

编者阵容

FreeBuf咨询： 杨雨翔 龚子程 宋丹丹 赵家钰

天融信团队： 雷晓锋 余小军 于国强

蚂蚁集团团队： 翁海琴 刘焱 仲震宇 王宇 韦韬 王珉然 朱泽韬

瑞数信息团队： 吴剑刚 邓巧华 俞加鸣

CONTENTS 目录

AI推动网络安全产业变革

伴随着AI及相关大模型产品的增多以及AI在攻击方的快速运用，主打AI的网络安全产品及服务将成为行业的主要竞争领域，能够深度整合AI技术，不断创新和优化安全解决方案的企业，将在未来的网络安全市场中占据领先地位。

网络安全领域的头部厂商正在快速推出AI安全产品，以应对日益增长的网络安全挑战。尤其是在数据安全、入侵检测、智能防护等方面，网安巨头企业通过AI技术提高了安全防护的效率和智能化水平。

随着AI技术的不断进步与成熟，AI在网络安全中的应用场景越来越广泛，涵盖了从异常检测到身份认证，再到自动化响应等多个方面，在网安领域的应用实例越来越普遍。

以下列举部分典型场景

在数据处理方面

AI大模型通过高效的算法和强大的计算力，能够在海量安全数据中快速挖掘出有价值的情报信息，筛选、分析和解读数据，从中识别出潜在的安全威胁和异常行为，极大地提升了数据处理的速度和准确性，为安全分析和响应提供坚实的支撑。

在自动化渗透测试与漏洞管理方面

AI大模型可以模拟各种复杂的攻击场景，快速发现系统中的漏洞和弱点，能够帮助安全团队优先处理高危漏洞，提供修复建议，并持续跟踪漏洞的修复进度，确保系统的安全性得到及时提升。

在实时安全预警方面

相较于传统安全防御难以应对快速变化的威胁环境，过度依赖规则匹配和签名检测，AI通过自主学习往往可以实现精准、实时的威胁检测。这不仅提高了检测效率，还显著降低了误报率和漏报率，为构建更加稳固的安全防线提供了有力支持。

在关联分析方面

AI大模型能够自动化地整合来自不同源头的异构数据，并进行必要的预处理，为后续分析奠定统一的数据基础。通过关联分析和特征提取，AI能够识别出不同数据源间的相关性和潜在模式，从而有助于识别异常网络行为和潜在的安全威胁。

在用户行为分析方面

AI大模型能够学习并理解用户和设备的正常行为模式，一旦检测到登录失败次数过多、访问敏感数据等异常行为，便会立即触发警报，并可能自动采取隔离措施，防止潜在的安全威胁扩散。这种能力对于防范内部威胁和外部攻击都至关重要。

在智能安全决策方面

AI大模型能够整合来自多个数据源的信息，进行综合分析，为安全团队提供全面的安全态势感知和决策支持。通过预测潜在的安全风险、评估不同安全策略的效果，AI大模型能够辅助安全团队做出更加明智的决策，提高整体的安全防护水平。

在知识库与培训体系优化方面

AI大模型可以构建和维护一个全面的安全知识库，包括最新的漏洞信息、攻击手法和防御策略等。这不仅可以为安全分析人员提供及时、准确的参考信息，还可以作为智能培训系统的数据源，通过模拟实战场景、提供个性化学习路径等方式，帮助安全人员不断提升技能水平。

AI大模型在网安行业的应用推动技术创新和产业升级

通过不断引入新的技术和方法，AI大模型为网安行业带来了更多的可能性和解决方案，促进整个行业的快速发展和进步。

AI安全发展的风险与威胁

AI对于网络安全的发展产生了重大影响，不仅是指AI推动了技术创新和产业升级，被用于构建恶意检测、攻击识别系统，提升安全防御的智能化水平，降低响应时间，提高整体的安全防御效果。

更是因为AI在与网络安全深度结合的过程中，引入了大量全新、复杂的风险与威胁，极大地降低了黑灰产与网络犯罪的门槛，例如攻击者可一键自动化生成高质量钓鱼邮件。因此AI安全产品研究不能只关注AI和安全产品本身，也需要关注在不同业务场景下AI带来的新威胁。

具体来说，AI安全风险主要分为AI大模型自身安全风险和黑客利用AI大模型提升入侵效率和成功率两部分。

（一）AI大模型存在安全风险

AI大模型存在的安全风险主要可以分为三类，分别是训练数据的安全风险、大模型本身的安全风险和外部攻击风险。

训练数据的安全风险

AI大模型训练数据的质量直接决定了其实际能力，因此训练数据在提升AI大模型的性能和应用效果中扮演着重要角色。训练数据的形式通常包括文本、图像、语音、视频等结构化与非结构化数据。在实际训练过程中，AI大模型训练数据中往往包含敏感信息，且屡屡发生数据泄露事件，涉及的数据安全问题也随之凸显，包括个人隐私泄露、商业机密泄露等。

以下是AI大模型训练数据泄露的安全风险：

1）数据泄露

AI大模型在训练过程中可能会泄露训练数据中的敏感信息。攻击者可以通过模型的输出推断出训练数据的内容，或者利用模型的漏洞获取未经授权的数据访问。例如，模型可能会在输出中无意间泄露训练数据中的个人信息或其他敏感数据。

其次，攻击者可以在不接触隐私数据的情况下，利用模型输出结果等信息来反向推导出用户的隐私数据。这种攻击可以通过模型的输出向量来恢复训练数据或者预测数据中的图像，导致用户的肖像信息或其他敏感信息被泄露。

2）数据投毒

数据投毒是AI大模型面临的一种严重的安全威胁，它涉及在训练数据集中故意引入错误或虚假的数据，以误导模型并产生错误的预测结果，从而影响模型的训练效果和最终输出，会对AI系统的可靠性、稳定性和安全性造成严重影响。

典型数据投毒的形式如下：

恶意数据注入：攻击者故意向训练数据集中注入错误或虚假的数据，以误导模型。这些数据可能包括具有误导性标签或特征的数据，导致模型学习到错误的知识。恶意数据注入是一种直接且有效的攻击手段，能够显著提高模型的错误率，影响其在实际应用中的表现。

数据篡改：攻击者篡改现有数据集中的某些数据，例如修改图像、文本或音频文件。这种攻击通常针对标签数据，通过更改标签来误导模型。数据篡改能够直接影响模型的输出结果，导致模型在各种应用场景中表现不佳。

数据污染：攻击者通过控制数据源或数据传输过程，使数据集受到污染。这种污染可能包括注入恶意数据、数据损坏等。数据污染是一种隐蔽的攻击手段，能够在模型训练过程中不被察觉，但其影响可能在模型部署和运行时显现。

3）数据版权

根据《生成式人工智能服务安全基本要求》，人工智能的训练数据主要来源于开源语料、自采语料、商业语料、使用者输入信息语料，其可能存在大量安全风险。

开源语料：虽然开源语料的数据集公开可用，但可能包含受版权、商标、专利或其他知识产权法律保护的元素，该语料无法保证知识产权清洁性。

商业语料：是指开发者与数据提供商签订协议后获得的内容，但这部分内容通常难以追根溯源，确保每项内容都有准确授权。

自采语料：是指开发者利用爬虫技术获取的网络信息，其爬取的内容受相关网络平台政策的影响，若出于商业产品开发使用目的，则依然可能存在不正当竞争的风险。

用户输入信息语料：是指用户在使用AI大模型过程中输入的信息，其中包括个人敏感信息和公司敏感数据，被AI大模型调用后可引发严重的安全风险。

外部网络攻击与风险

与AI大模型快速发展、应用不同的是，AI如今仍然面临着大量的外部攻击风险。由于大模型自身存在脆弱性，对安全风险的防范能力不足，容易受到指令攻击、提示注入和后门攻击等恶意攻击。尤其是在政治、军事、金融、医疗等关键的涉密应用领域，任何形式的恶意攻击都可能造成严重的后果。

以下列举几种较为典型的外部攻击和安全风险：

海绵攻击：是指攻击者通过针对海绵结构的弱点来破坏其安全性。攻击的目标是从给定的输入中生成具有特定性质的输出。攻击者通过选择恰当的输入数据和进行特定的修改来实现这一目标，以及通过观察输出数据的反馈信息来指导修改进程。CalypsoAI的首席执行官Neil Serebryany称，该攻击试图让神经网络使用更多的计算，以达到可能超过系统可用计算资源的程度，并导致系统崩溃。类似的攻击形式还有tokens攻击。

武器化模型：是指将人工智能模型或机器学习模型应用于恶意目的或攻击性行为的过程。通常情况下，这种行为是非法且具有破坏性的，旨在侵犯隐私、破坏系统、欺骗用户或其他恶意活动。武器化模型可以被用于各种攻击场景，例如在钓鱼攻击中，可使用自然语言处理模型生成欺骗性的电子邮件或社交媒体消息，以诱骗用户提供个人信息或执行恶意操作。

伦理与歧视风险：基于商业秘密保护等原因，人工智能存在“算法黑箱”，使得算法可解释性不强，输出结果的公正性难以被证成，这有违基本的程序正义要求。同时训练过程中也会吸收并放大训练数据中的偏见导致在处理特定问题时对某些群体产生不公平对待，并且随着时间的推移，歧视风险可能会不断升级、叠加。

安全合规风险：全球各国政府陆续出台相关法规，对AI大模型提出新的合规要求。例如中国发布《生成式人工智能服务管理暂行办法》，旨在促进生成式人工智能健康发展和规范应用。新的法规要求AI大模型在开发、部署和使用过程中必须遵守更严格的安全和隐私标准，增加了AI大模型的合规风险。同时，AI技术的快速发展可能使得现有的法规难以跟上技术变革的步伐，导致合规风险的不确定性增加。

（二）AI大模型成为攻击者的利器

通过AI大模型，攻击工具能够自动学习、适应和进化，攻击者无需手动编写复杂的代码或策略，即可发起高效、精准的攻击，这使得网络攻击相较于以往更加容易实施且难以防范。

以下列举部分使用场景：

高效生成恶意内容：AI大模型能够快速地生成看似合理且难以区分的欺诈性文本、电子邮件和网站。例如，FraudGPT就是黑客借助开源模式模仿ChatGPT等合法模型，基于有害语料库训练而生的、专门用于网络犯罪、诈骗等非法行为的一类非法大模型。

据国际安全研究团队Perception Point与Osterman Research此前联合发布的《人工智能在电子邮件安全中的作用》报告显示，网络罪犯正在迅速采用AI工具以推动其利益，而有91.1%的组织称他们已遭受到了被AI增强的电子邮件攻击。

自动化漏洞利用：大模型对编程语言和软件漏洞的深入理解，使得攻击者能够更快速地识别并利用这些漏洞。通过自动化扫描和测试，攻击者可以更加高效地发动攻击，对目标系统造成破坏。此外，基于大模型的攻击手段还可能更加隐蔽和难以检测，增加了防御难度。

深度伪造与身份冒充：AI大模型在图像和声音处理方面的能力也为攻击者提供了新的伪造手段。利用这些技术，攻击者可以创建逼真的虚假图像、视频和音频文件，进行深度伪造，冒充他人身份进行诈骗或传播虚假信息。

根据德勤的最新报告，与深度伪造相关的网络攻击损失预计将从2023年的123亿美元飙升至2027年的400亿美元，复合年增长率达到惊人的32%，其中，银行和金融服务业将成为主要目标。报告显示，作为对抗性AI攻击的典型代表，深度伪造攻击仅去年一年就暴增了3000%。

复杂攻击策略的制定：大模型所具备的逻辑推理和策略规划能力，使得攻击者能够设计出更加复杂和难以预测的攻击策略。这些策略可能涉及多个步骤、多个目标和多种攻击手段的组合，增加了防御的复杂性和难度。攻击者借助AI大模型的泛化能力，还可以将针对特定目标的攻击策略迅速扩展到更广泛的范围。通过自动化工具和脚本，攻击者可以同时对多个目标发起攻击，迅速扩大攻击范围，增强攻击的影响力。

定制化攻击：定制化攻击策略，也是AI大模型为攻击者提供的“超能力”，攻击者能够借此利用目标网络的特点、用户习惯等信息，动态生成定制化的攻击策略，使得攻击更加难以被检测和防御。这种个性化的攻击方式极大地提高了攻击的成功率。

恶意AI大模型：AI大模型由于存在安全控制措施，并不能直接被用于恶意活动，但在2023年却诞生了专门被用户黑客攻击/入侵的恶意AI大模型，包括WormGPT和FraudGPT在内的最受欢迎的黑暗聊天机器人的销售在黑暗的网络上激增。据ThreatLabz研究团队发布报告，这些恶意AI大模型使用门槛低，功能十分齐全，可以用于创建，测试一下或优化任何种类的恶意代码，包括恶意软件和勒索软件等。

全球AI安全市场概况与宏观环境分析

随着人工智能相关技术的快速发展，AI已经逐渐成为网络安全领域的重要驱动力。AI技术在网络安全产品中的应用，不仅提升了产品的防护能力，还为网络安全带来了全新的视角和解决方案。例如赋能传统威胁检测，使其能够自动分析网络流量，识别异常行为；强化自动化响应与处置，AI技术可以实现对事件的自动分类和分级，并根据预设的策略自动采取相应的处置措施等。

IDC最新发布的《全球未来信任十大预测--中国启示》报告也指出，到2025年中国40%的2000强企业将在其安全运营中心中基于第一方数据部署GenAI，以便辅助高级分析师进行检测和响应工作，并同时解决幻觉、偏见、隐私和强化学习等问题。

2024年全国两会期间，“人工智能+”首次被写入《政府工作报告》。报告提出，深化大数据、人工智能等研发应用，开展“人工智能+”行动，打造具有国际竞争力的数字产业集群。此次“人工智能+”行动的提出，预计也将为人工智能技术在各行业的广泛应用开启新篇章，人工智能的重要性正在逐步提升。

AI技术在网络安全领域的应用为行业带来了新的商业机会。许多企业和机构开始寻求基于AI的安全解决方案，以提升自身的网络安全防护能力。未来随着技术的不断进步和应用需求的持续增长，全球AI安全产品市场将迎来更加广阔的发展前景。

（一）全球市场总体规模分析

目前全球AI安全产品市场还处于初级阶段，相关安全大模型的应用主要聚焦在安全检测、安全运营、数据安全等领域，在金融、大型制造业、互联网等多个行业落地，其发展势头较为迅猛，市场规模不断扩大，竞争格局日趋激烈。

而随着网络攻击手段日益复杂和频繁，企业对网络安全的需求不断增长，推动了对AI驱动的安全解决方案的需求增加。根据Marketsandmarkets报告，全球网络安全人工智能市场规模在2023年达到224亿美元，并预计到2028年将达到606亿美元，复合年增长率达到21.9%。根据ResearchAndMarkets报告预测，2030年全球网络安全中AI产值将突破千亿美元，年复合增长率达到25.7%。这表明AI网络安全产品正受到越来越多企业和组织的青睐，市场需求强劲。

2025年之前，AI自动执行日常事件响应任务将成为主流，缩短响应时间，最大限度地减少手动错误，同时整合可解释的AI，以提高透明度，促进更好地了解威胁检测机制。2028年之前，网络安全企业将能够更多地使用联合机器学习模型进行协作威胁情报。2030年之前，AI将与数据安全能力紧密结合，以确保敏感数据的保护和数据交易的安全。

从行业应用来看，目前AI安全已经在通信、能源、金融、教育、医疗、关键基础设施等行业具有落地方案，并在某些方面开始发挥出人工无法替代的重要作用，例如自动处理告警信息、自动化识别网络威胁、自动化威胁响应、敏感信息保护等。

从地域分布来看，北美地区占据全球AI安全产品市场的最大份额，其次是欧洲和亚太地区。中国、日本和印度等国家的AI安全产品市场需求增长迅速，成为推动亚太市场增长的主要动力。

北美市场规模领先

北美地区预计将继续保持全球最大的AI安全产品市场地位，北美AI安全产品市场规模巨大，这主要得益于其在技术创新、资金投入和政策支持方面的优势。随着全球对AI安全需求的不断增长，北美市场预计将继续保持领先地位。

北美AI安全产品市场的主要参与者包括微软、思科、博通、Palo Alto Networks、Fortinet、CrowdStrike、Cloudflare、Zscaler、Netskope等。主要参与者通过收购、合作和技术创新等手段，不断提升自身的市场竞争力和市场份额。例如，思科收购Splunk，微软与IBM达成广泛合作，CrowdStrike收购数据安全态势管理公司Flow Security等。竞争格局的激烈使得市场中的主要参与者不断寻求新的增长点和竞争优势，有助于推动整个行业的发展和创新。

欧洲AI安全产品市场需求旺盛

近年来，欧洲AI安全产品市场需求旺盛，主要驱动因素包括欧盟《人工智能法案》的实施、企业对网络安全投资的增加、生成式AI技术的应用，以及数据安全和隐私保护的需求提升。

网安法规驱动市场需求：2024年5月1日，欧盟《人工智能法案》正式生效，对AI系统的安全性提出了严格要求，从而推动了AI安全产品市场的发展。

数据安全和隐私保护的需求提升：随着企业和个人对数据安全和隐私保护意识的增强，对AI安全产品的需求也随之增加。

2023年12月，“数字欧洲计划”正式被提出，欧盟委员会为包括网络安全和人工智能在内的数字解决方案提供了7.627 亿欧元的资金。随着数字欧洲计划加速实施，对于AI安全产品的需求也将更加旺盛，例如包括部署数字化安全运营中心网络，将使用AI安全技术或以提升对网络威胁的监测和响应能力；进一步完善人工智能实验及测试设施，支持人工智能技术的研发等，将进一步刺激AI安全市场的需求。

中国AI安全产品市场发展迅速

Check Point 与 Vanson Bourne联合发布的市场调研报告显示，AI在网络安全领域应用效果明显，其中接近50%的亚太地区网安管理层表示，AI 可简化安全运维和资源分配，这一比例在全球三大地区中位居首位。

虽然AI大模型在国内网络安全领域的应用还处于初阶，但是已经展现巨大发展潜力，众多厂商推出了集成AI技术的网络安全产品，已经在各行各业中落地，AI安全产品市场发展迅速，呈现出“多强竞争”的格局。

众多老牌厂商包括天融信、360、奇安信、安恒信息、迪普科技、绿盟科技、深信服科技、微步在线、火山引擎、华为等都已发布了相应的网络安全AI大模型或AI安全产品。例如将AI大模型能力深度融合于防火墙、入侵检测与防御、Web应用防火墙、DDoS缓解、零信任网络访问等此前已经成熟的网络安全产品之中，进一步提升相关产品在自动化、准确度、效率等方面的效果。

国内知名安全厂商天融信是国内AI安全产品的主要玩家之一，一直积极探索“AI+安全”的融合创新发展，基于“平台、产品、场景”构建企业级AI安全能力体系，将大模型、小模型、机器学习等AI技术能力与网络安全产品深度融合，实现AI 对威胁检测、安全运营、知识问答和算力管理等应用场景的全面赋能。

通用AI大模型在网络安全行业也有广泛的应用。例如蚂蚁集团AI大模型可应用于威胁识别、态势感知、风险评分、恶意检测等方向，提升网络网络空间智能安全整体防护水平。

政策法规层面，目前各国政府对AI安全的重视程度不断提高，出台了一系列法规和标准，促进了AI在网络安全领域的应用。中国2023年10月发布《生成式人工智能服务管理暂行办法》，这也是中国针对生成式AI服务的首个法规，要求生成式人工智能提供商在提供面向公众的服务时必须获得许可。该办法旨在规范生成式AI服务，保护用户隐私和数据安全。

美国在2023年1月发布了《人工智能风险管理框架》，旨在对人工智能系统全生命周期实行有效的风险管理。该框架将可信度考量纳入设计、开发、使用和评估AI产品、服务和系统中，并基于其他机构的AI风险管理工作，确保制定过程的公开、透明。

2024年8月，欧盟《人工智能法案》将正式生效。这是全球首个全面监管人工智能的法案，旨在实现与《通用数据保护条例》（GDPR）相同的“布鲁塞尔效应”。该法案对人工智能系统进行了分类，并根据风险等级制定了相应的监管要求。

（二）AI安全市场融资情况

2023年，全球网络安全行业投融资情况不容乐观。根据网络招聘公司Pinpoint Search Group的报告，2023年全球网络安全行业共进行了346轮融资和91笔并购（M&A）交易，投资总额为87亿美元，比2022年的145亿美元减少了40%，下降幅度较大。

2023年第四季度的跌幅尤为明显，初创公司仅筹得16亿美元，创下自2018年第三季度以来的最低水平，当时网络安全公司仅筹得13亿美元。2023年第四季度只有三家网络安全初创公司的单轮融资超过1亿美元。

这里面既有市场在经历了2021年融资激增后回调的原因，也反映出投资者开始趋于谨慎，对项目的选择和评估更加严格。其中，人工智能、零信任、数据安全、智能汽车等方向受到市场青睐，随着数字化转型的推进，网络安全需求在不断扩展，特别是在人工智能等新兴技术中的应用。这意味着AI是网络安全领域中最热门且最有想象空间的投资方向，并在2023年获得了新的增长。

FreeBuf咨询根据公开数据统计发现，2022年，全球AI安全的融资额约为2000万美元左右；2023年，AI安全融资迎来了爆发期，融资总额约为2亿美元，同比增幅高达1000%。

2023年人工智能安全的融资额显著增长，主要得益于AI技术的突破性发展，在促进网络安全产品联动、自动化方面有着巨大的优势，可实现安全威胁实时检测和自动化处理，有效提升了网络安全产品的能力。

其次是出于对AI大模型自身安全性的担忧，随着ChatGPT等AI工具的广泛流行，以及AI大模型数据泄露事件、AI大模型越狱事件等屡屡出现，导致市场对大模型自身的安全性的需求激增。

Gartner的研究预测指出，由于生成式人工智能（GenAI）引入了新的攻击面，为提供相关防护，企业机构必须对应用和数据安全实践以及用户监控进行变革。到2025年，GenAI的采用将导致企业机构所需的网络安全资源激增，从而使应用和数据安全支出增加15%以上，进一步拉动了网络安全相关需求。

2024年，AI+网络安全（AI安全）依旧是行业融资最热门的领域。FreeBuf对公开数据进行整理分析后发现，截止7月31日，全球AI安全共进行了102轮融资交易，投资总额为超过10亿美元，同比2023年增幅达到500%，预计在2024年下半年还有大量AI安全企业将活动融资，2024年总融资金额或接近20亿美元。

AI网络安全企业融资表

国家地区	企业名称	业务方向	轮次	融资金额
中国	中科睿鉴	AI反欺诈	未披露	近亿元
	云起无垠	安全智能体	天使+轮	数千万元
	瑞莱智慧	AI安全	战略融资	未披露
	路为科技	云原生	股权融资	未披露
	木卫四	车联网安全	Pre-A轮	数千万元
	中科睿鉴	AI伪造检测	股权融资	未披露
	深可信云	安全解决方案	未披露	未披露
	掌数科技	金融行业大数据平台	未披露	未披露
	锘崴科技	隐私计算	未披露	未披露
美国	Cyera	数据安全	C轮	3亿美元
	Corelight	NDR	E轮	1.5亿美元
	Huntress	安全运营	D轮	1.5亿美元
	Bugcrowd	安全服务		1.02亿美元
	Cyberhaven	数据安全	C轮	8800万美元
	Cowbell	网络保险	C轮	6000万美元
	Protect AI	AI安全	B轮	6000万美元
	Transcend	隐私安全	B轮	4000万美元
	Seven AI	安全运营	种子轮	3600万美元
	SafeBase	AI GRC自动化	B轮	3300万美元
	WitnessAI	AI安全	A轮	2750万美元
	SOCRadar	XTI&DRP	B轮	2520万美元
	CyberSaint Security	网络风险管理	种子轮	2100万美元
	Credo AI	AI安全	B轮	2100万美元
	Kindo	AI SSPM	股权融资	2060万美元
	Alethea	舆情监测	B轮	2000万美元
	XBOW	AI渗透测试	种子轮	2000万美元
	Lakera	LLM漏洞防护	A轮	2000万美元
	Aim Security	AI安全	A轮	1800万美元
	LightBeam.ai	零信任	A轮	1780万美元
	Axion Ray	AI可观测	A轮	1750万美元
	Patronus AI	LLM安全	A轮	1700万美元
	Dropzone AI	AI安全运营	A轮	1685万美元
	Andesite AI	安全分析平台	A轮	1520万美元
	Apptega	安全合规平台	未披露	1500万美元
	Bolster AI	反钓鱼	B轮	1400万美元
	Aurascape AI	AI安全	种子轮	1280万美元
	StrikeReady	AI+SOC	A轮	1200万美元
	Prophet Security	AI安全运营	种子轮	1100万美元
	Reken	AI安全	种子轮	1000万美元
	Bedrock Security	数据安全	种子轮	1000万美元
	Allure Security Technology	DRP	A轮	1000万美元

国家地区	企业名称	业务方向	轮次	融资金额
美国	Simbian	AI安全	种子轮	1000万美元
	Nova Microsystems	数据安全	未披露	1000万美元
	CultureAI	风险管理平台	A轮	1000万美元
	Abstract Security	AI驱动SIEM	种子轮	850万美元
	Averlon	AI安全	种子轮	800万美元
	Dymium	数据访问控制	种子轮	700万美元
	Goodfire	AI可观测性	种子轮	700万美元
	P0	开发安全	种子轮	650万美元
	Liminal	AI治理	种子轮	过500万美元
	Prompt Security	AI安全	C轮	500万美元
	AirMDR	MDR	种子轮	500万美元
	ZEST Security	AI驱动云安全	种子轮	500万美元
	Promptfoo	AI应用漏洞修复	种子轮	500万美元
	RagaAI	AI安全	未披露	470万美元
	Redcoat AI	AI安全	种子轮	424万美元
	Trustwise AI	AI安全	种子轮	400万美元
	HydroX AI	AI安全	天使轮	400万美元
	Coris	风险评估	种子轮	370万美元
	Knostic	AI访问控制	前种子轮	330万美元
	Amplifier Security	AI安全自动化	Pre种子轮	330万美元
	HoundDog.ai	数据安全	种子轮	310万美元
	SydeLabs	AI安全	种子轮	250万美元
	Bricklayer AI	安全运营	预孵化资金	250万美元
	Enkrypt	AI治理	种子轮	235万美元
	Resonance Security	Web2/Web3安全评估	Pre种子轮	150万美元
	Guard Dog Solutions	安全运营	未披露	100万美元
	PromptArmor	AI安全	Pre种子轮	50万美元
	Holistic AI	AI治理	风险轮	未披露
	Traceable	API 安全	未披露	未披露
	Harmonic Security	AI安全	种子轮	未披露
	Velotix	DSPs	未披露	未披露
	Vidoc Security Lab	AI代码安全	种子轮	未披露
	Almanax	区块链安全	前种子轮	未披露
	Blackbird.AI	虚假信息对抗	战略投资	未披露
	Haize Labs	AI安全	未披露	未披露
	Vorlon	API安全	A轮	未披露
加拿大	TrojAI	AI治理	种子轮	580万美元
	Armilla AI	AI安全	未披露	450万美元
	AllHeart Web	数字风险防护DRP	种子轮	10万美元
以色列	Pindrop Security	AI鉴伪	债务融资	1亿美元
	Clarity	AI安全	未披露	1600万美元
	Aim Security	AI安全	A轮	1000万美元
以色列	DeepKeep	AI安全	种子轮	1000万美元
	PVML	数据访问平台	种子轮	800万美元
	Seal Security	软件供应链安全	未披露	740万美元
	Apex	AI安全	种子轮	700万美元
	Upstream Security	车联网安全	战略融资	未披露
瑞典	Xensam	安全运营	未披露	4000万美元
立陶宛	Cyber Upgrade	安全运营	未披露	65万欧元
爱沙尼亚	BotGuard OÜ	Anti-Bot	A 轮	1200万欧元
法国	Mindflow	SOAR	种子轮	500万欧元
澳大利亚	Redactive AI	AI DLP	种子轮	750万美元
	Nullify	开发安全	种子轮	520万澳元
新加坡	StealthMole	暗网情报	A 轮	700万美元
西班牙	Onum	可观测性	A 轮	2800万美元
	Nymiz	隐私数据	种子轮	280万欧元
英国	Cynomi	安全运营	A轮	2000万美元
	qomodo	工控安全	Pre种子轮	130万英镑
爱尔兰	Tines	安全流程自动化	B轮	5000万美元
印度	Treacle Technologies	欺骗防御	种子轮	4亿卢比
科特迪瓦	DANAYA	LLM身份验证	Pre种子轮	75万欧元

表：2024年全球AI安全企业融资情况

从2024年全球AI安全企业融资情况表可以发现，AI安全创业公司十分活跃，资本市场对该领域未来的发展持续看好，并呈现出以下特点：

融资金额高

从融资金额来看，大额融资金额高，小额融资数量分散。从上表中可以看到，Cyera以3亿美元的C轮融资金额位居首位，其次是Corelight和Huntress，均获得了1.5亿美元的融资。这三家企业均来自美国，且业务方向主要集中在AI+数据安全和AI+安全运营领域。中国中科睿鉴融资金额接近1亿美元关口，融资轮次未披露，其主营业务是AI反欺诈。

除了大额融资外，还有大量企业获得了数百万至数千万美元的融资。这些企业涵盖了AI+隐私计算、AI+车联网安全等多个领域，显示出AI安全领域的投融资活跃度。这些初创企业在早期阶段就能获得相对较高的融资金额，显示出投资者对其未来发展的信心。

从投融资轮次分析，种子轮和天使轮融资事件占比较高。这表明AI网络安全领域仍处于快速发展阶段，大量初创企业涌入该领域寻求发展机会。

但随着AI网络安全市场的不断成熟，部分企业开始进入A轮、B轮等中后期融资阶段，这些企业通常已经在市场中取得了一定的份额和影响力，正在寻求更多的资金支持以加速发展。这反映出市场对企业发展速度和规模的要求逐渐提高。

除了传统的股权融资外，部分企业还通过债务融资等方式筹集资金，显示出AI网络安全企业在融资方式上的多样性和灵活性。

融资活动全球化

从地区分布来看，硅谷、北京、上海等国际大都市已经成为AI网络安全领域的重要集聚地，这些地区汇聚了大量的创新资源和人才优势，为企业的发展提供了良好的环境和条件。

从国家分布来看，AI网络安全融资呈现出一超多强格局。美国以绝对的企业融资数量成为AI网络安全领域投融资最为活跃的国家。这主要得益于美国科技创新体系、金融市场成熟度、研发资源丰富度及开放创新文化。顶级高校与研究机构提供人才储备与知识产出，硅谷等创新高地汇集全球活跃风投，形成良好投融资生态。因此，强大的科技创新能力、完善的市场环境和配套的政策支持是美国AI网络安全投融资活跃的主要原因。

中国在AI网络安全领域的发展势头强劲，包括中科睿鉴、云起无垠、木卫四等多家公司获得了大额融资，显示出资本市场对于AI网络安全的看好。其中中科睿鉴融资金额近亿元，达晨财智资本领投。中科睿鉴致力于运用AI技术赋能数字内容安全。围绕全类型伪造检测、多模态数据生成、内容合规审核等核心技术，布局了“模型-数据-算力”的AI基础设施，面向国家、行业、个人安全场景，提供音视图文全栈全类型鉴伪技术和产品服务。

资本市场的活跃与中国对科技创新的大力支持和庞大的市场需求密不可分。政策方面，中国陆续出台了《新一代人工智能发展规划》《促进新一代人工智能产业发展三年行动计划（2018-2020年）》《关于加快推进人工智能创新应用的指导意见》等多部重要意见，设立科技创新2030–“新一代人工智能”重大项目，各地方政府设立相关产业投资基金，引导社会资本投入人工智能领域。

需求方面，随着中国经济的快速发展，各行各业都在寻求通过技术创新来提升效率和竞争力。人工智能作为一项颠覆性技术，能够帮助企业实现自动化、智能化，推动产业升级和经济增长。从智能制造、智慧城市到医疗健康、教育、金融等领域，人工智能技术都有着巨大的应用潜力。

以色列紧随其后。作为网络安全老牌强国之一，在AI网络安全领域也有着不俗的表现。多家以色列企业凭借其独特的技术优势和创新思维获得了国际资本市场的青睐，融资金额自700万美元至1亿美元不等。其中，Pindrop Security债务融资1亿美元，Hercules Capital领投。Pindrop Security使用基于神经网络的生物识别引擎优化语音安全，确保无缝被动身份验证和高效的欺诈检测，还结合了网络级数据、用户设备和行为分析以及风险情报，以在IVR和座席级别对呼叫者进行全面身份验证。

随着全球化的深入发展，AI网络安全领域的投融资活动呈现出全球化趋势，加拿大、法国、澳大利亚、英国、新加坡、西班牙、瑞典、立陶宛、爱沙尼亚、爱尔兰、印度、科特迪瓦等国家皆有相应的企业获得投资，这意味着AI网络安全正在被全球大多数资本看好。

细分市场多样化

从细分市场来看，AI网络安全的方向呈现出多样化的特点。

其中，AI安全（AI自身的安全性）占比高达23%，是最受资本和创业者青睐的细分方向。

随着深度学习、计算机视觉、自然语言处理等技术的不断发展和成熟，AI的应用领域和能力得到了极大的拓展。各行各业都在寻求通过AI技术提升运营效率、创新服务模式，满足市场和消费者的新需求。

大量的AI应用需求带来了全新的安全风险，数据显示，AI既放大现有网络安全威胁，又引入了新型威胁，引发网络安全事件指数级增长。例如2023年基于AI的深度伪造欺诈暴增3000%，基于AI的钓鱼邮件数量增长了1000%。

由于企业现有的安全体系难以有效应对此类威胁，AI安全存在巨大的想象空间，从而吸引大量初创公司和资本投身于该领域。

安全运营是AI安全投融资热度第二的细分方向，占总投融资事件比例达到10%。AI在安全运营中的应用主要体现在以下几个方面：

智能防御与实时监测：AI技术通过机器学习和深度学习算法，能够自动识别网络流量和用户行为的异常模式，实现实时监测和预警。

自动化响应与处置：AI系统可以自动对检测到的威胁进行响应，如自动加入黑名单或通知管理员。这种自动化的响应机制提高了安全事件的处置速度，并减轻了安全团队的工作压力

风险预测与策略优化：AI技术通过分析历史攻击数据和网络安全日志，预测未来可能面临的安全风险，帮助安全运营团队提前做好准备，调整防御策略

智能化的安全保障：AI技术使得安全保障系统能够动态地学习和适应不断变化的威胁环境，通过深度分析网络流量和用户行为，更准确地识别出潜在的威胁

数据安全是AI安全投融资热度第三的细分方向，占总投融资事件比例达到8%。一方面，AI的发展带来数据安全问题，包括大规模数据采集与个人隐私风险，AI对数据的依赖性及其衍生问题；网络攻击与数据泄露，数据篡改与AI模型的安全性等。

另一方面，AI可进一步赋能数据安全，包括提升数据分类分级的效率，对结构化和非结构化数据进行处理；通过对进出口流量的识别来强化数据安全管控；借助数据匿名化与差分隐私技术，强化网络安全措施与加密技术等。

AI正在渗透至安全各个细分领域，包括反欺诈、车联网安全、云原生、漏洞检测、NDR等，获得资本市场的认可。从2022年––2024年7月，AI安全投融资总金额持续攀升，表明市场与资本对AI安全强烈看好，在当下全球网络安全行业投融资不足的背景下，AI安全的火热融资情形更加凸显出未来市场具有庞大的想象空间。

AI技术持续推动网络安全领域的技术创新。作为全球网络安全行业新兴技术的风向标，RSA Conference创新沙盒大赛一直备受关注。2023年和2024年RSA Conference的创新沙盒大赛冠军都是AI安全初创企业，分别是HiddenLayer和Reality Defender，表面网络安全行业对于AI安全技术发展趋势的认可。

HiddenLayer 是机器学习算法和模型安全解决方案的提供商，成立于 2022 年。该公司基于轻量化的软件平台方案和AI大模型的能力，提供针对机器学习系统的威胁建模、机器学习的风险评估、人工智能 / 机器学习的模型扫描等服务，实现定制化的攻击面识别、攻击防护、攻击模拟等功能。

Reality Defender是一家专注于利用人工智能技术检测深度伪造和AI生成的媒体内容的初创企业。该公司的创新技术主要是深度伪造检测工具和多模型方法，前者可实时识别AI合成与深度伪造的欺诈、虚假信息和有害内容，包括文本、图像和音视频等；后者对深度伪造检测平台和API对人工智能生成媒体的平台具有强大的抵御能力，使团队能够实时识别欺诈、虚假信息和有害内容。

AI安全企业并购

在并购方面，过去一年内已发生多起AI网络安全公司收购事件，其中具有代表性的并购事件如下：

2023年9月，思科宣布以280亿美元收购Splunk，以实现在AI安全方面的强强联合，推动实现 AI 驱动的下一代安全性和可观测性。此次被收购方Splunk是近年来安全创业公司领域的明星公司，主要帮助企业监控和分析数据，以最大限度地降低被黑客攻击的风险，并更快速地解决技术问题。

2024年3月，CrowdStrike宣布斥资收购Flow Security，购完成后，CrowdStrike计划在Falcon Cloud Security中全面提供原生Flow Security DSPM功能，作为Falcon XDR平台的一部分，使客户能够整合云端解决方案并保护整个云资产，以进一步加强其 AI 网络安全解决方案。此外，CrowdStrike宣布将与戴尔（DELL.N）联手，帮助企业利用人工智能抵御网络攻击，以防范GenAI、隐形社工攻击和端点攻击。

2024年4月，私募股权巨头 Thoma Bravo 宣布以 53.2 亿美元的价格收购英国网络安全AI公司DarktraceDarktrace。Darktrace公司成立于2013年，主要利用人工智能检测IT网络内的攻击和漏洞。2022年2月，Darktrace以 5370 万美元的价格收购了攻击面管理公司 Cybersprint，加速 Darktrace 进入主动 AI 网络安全等新领域。

2024年7月，谷歌母公司Alphabet拟以230亿美元收购以色列知名AI安全初创公司Wiz，但被后者拒绝。值得注意的是，该公司2023年的收入约为3.5亿美元，在最近的一轮私人融资中筹集了10亿美元，估值达到120亿美元。

随着网络威胁的持续增长，投资者将积极投入AI防御性安全解决方案，预计未来将会有更多更多大型科技公司、网络安全巨头以及具有创新潜力的初创企业之间的战略合作与并购交易。

全球AI安全产品主要产品形态

根据海外市场的表现，网安行业AI安全产品的商业化效果较好，目前已有厂商实现不错的业绩，包括Palo Alto Networks基于AI驱动的新产品XSIAM；CrowdStrike公司的AI安全平台Falcon等。

AI安全产品的成功进一步刺激了厂商对于其应用方向的探索，给传统的网络安全产品带来了新的变革。AI在产品自动化与智能化响应、实时监控与预警、优化资源配置、增强用户体验、跨平台整合、持续学习与进化、辅助决策等方面有着明显的优势。

因此，AI安全产品的主要应用方向如下：

网络行为与威胁分析：AI在网络行为与威胁分析中扮演着重要角色。通过AI、深度学习或机器学习技术可以实现监测、理解和响应网络中的各种活动和潜在威胁。其中包括用户和实体行为分析（UEBA）、流量分析、协议分析、应用行为分析等，涉及沙箱、威胁情报、攻击模拟与预测、异常检测等技术的联动。

入侵检测：基于AI的入侵检测是一种利用AI技术对网络或系统中的潜在威胁进行识别和检测的方法。通过分析网络流量、系统日志和其他数据源，利用机器学习、深度学习和自然语言处理等技术来识别异常行为，从而及时发现并应对网络攻击。AI可区分正常网络流量和恶意入侵行为，有效发现网络中的入侵行为，对恶意入侵行为进行有效的检测和分类。

安全策略管理：基于AI的安全策略管理是指利用AI技术来制定、实施、监控和优化安全策略的过程。这种方法旨在提高安全防御的效率和准确性，以应对日益复杂的网络威胁，具有高效性、实时性、智能性、可持续性等有点。例如自动化风险评估系统，对访问请求进行实时的风险分析和评级，根据行为分析和风险评估的结果优化安全决策，或根据行为趋势和风险预测自动采取防护措施。

端点保护：基于AI的端点保护是一种利用AI技术来保护端点设备免受网络攻击的方法。通过在端点设备上部署AI模型，实时监控和分析设备行为，以识别和预防潜在的威胁。AI技术的加持可简化数据收集和分析，提升系统可视性和检测异常活动，并自动执行大部分分析、监视、检测和响应活动，使信息安全团队成员能够专注于更高优先级的操作。

漏洞管理：基于AI技术的漏洞管理是指利用AI技术来识别、评估、优先排序和修复软件或系统中的安全漏洞的过程。这种方法通过自动化和智能化的手段，对软件代码、系统配置、网络流量等进行分析，可明显提高了漏洞管理的效率、准确性和优先级。例如AI可生成基于场景的漏洞风险评分，确定优先级排序，帮助安全团队优先处理最关键的漏洞。

垃圾邮件处理与反钓鱼：基于AI的垃圾邮件处理与反钓鱼技术是一种利用AI的强大数据处理和智能分析能力，提高电子邮件安全性和反钓鱼技术的一种方法。通过分析大量邮件数据，学习识别垃圾邮件的特征，AI可提取出更加复杂和抽象的特征，提高过滤的准确率。和传统垃圾邮件处理方式相比，AI可更好地适应不断变化的垃圾邮件发送方式，具有更强的鲁棒性。另外AI在反钓鱼方面已经展现出新的潜力，其核心优势包括钓鱼邮件类型的覆盖、意图提取与分析能力以及多语言能力等，可有效应对和检测各类复杂的钓鱼邮件。

数据与文件分类：基于AI技术的数据与文件分类是一种利用AI算法自动识别、分类和管理数据及文件的方法。这种方法根据文件内容、元数据或其他特征将文件分配到预定义的类别中，可以帮助组织更有效地保护信息资产，提高工作效率和安全性，具有自动化、准确性高、可拓展性强、实时更新等优势。

安全知识问答：基于AI技术的安全知识问答是一种利用人工智能技术来提供安全相关知识和信息的服务。这种方法通过自然语言处理（NLP）技术，理解用户的问题，并提供准确的安全建议或解决方案。AI安全知识问答具有自然语言理解的能力，根据对话管理判断应该采取的策略，生成相应的回答，可有效提升服务效率，增强用户体验。

（一）国外AI安全产品分析

目前国外的AI安全产品商业化推进较快，已经有不少成熟的AI安全产品进入市场，并得到了不分企业用户的肯定。全球网络安全的头部厂商纷纷加快AI安全产品的动作，既是快速抢占尚未填满的空白市场，也是在网络安全智能化大战略布局上的先手。AI安全产品的落地提高了安全防护的效率和智能化水平，通过智能分类分级、安全策略配置、数据脱敏等技术，为企业提供全方位的数据安全解决方案。

总的来看，AI安全产品主要分为安全助手产品（AI Copilot）和AI驱动的安全平台。

安全助手产品

安全助手是AI与网络安全相结合最早的产品形态，也是目前大部分生成式AI产品聚焦的方向。

AI安全助手产品是基于AI技术，专门为网络安全领域设计的智能助手，通过对话的方式为用户解决安全相关的问题，可与安全工具、安全产品进行联动，以便更有效应对网络安全威胁。

AI安全助手可以快速理解并挑选合适的插件来回答用户的需求，支持文字交互，文件、图片的方式，以及插件扩展，从而提高工作效率。这是其核心价值所在，不仅提高了安全专家的工作效率，也降低了非安全工程师完成安全相关任务的难度。

随着AI技术的不断进步，AI安全助手有望在未来成为网络安全领域的重要辅助工具。

AI安全助手应具备自然语言理解能力、上下文关联能力、网安领域专业化能力、输出能力、插件扩展等多种能力。

自然语言理解能力：能够理解用户的自然语言输入，无需特定格式的命令，支持多种语言类型。

上下文关联能力：能够在对话中理解上下文，执行多个任务并综合回答。

网安领域专业化能力：能够准确理解安全行业专业术语，并输出相应的专业内容。

输出能力：可根据用户的问题输出各种内容，执行用户下达的各项指令，以及其他预设安全策略等。

插件扩展：支持灵活的插件扩展，以适应不同的安全需求。

AI安全助手具备安全咨询、自动化任务执行、安全事件响应、风险评估、知识库和情报集成、可视化与报告生成等核心功能。

安全咨询：提供有关网络安全、数据保护、合规性等方面的建议和信息，快速回答网络安全问题，辅助安全分析和运维工作。

自动化任务执行：根据预设的安全策略自动化执行各项安全操作，例如自动化执行渗透测试、更新安全规则、修复已知漏洞等，辅助用户完成网络安全相关的任务。

安全事件响应：在检测到威胁后，能够自动采取预定义的措施进行初步应对，例如隔离受感染的系统、阻断恶意IP地址或执行紧急补丁安装等。

风险评估：监控和分析网络流量，识别系统潜在的安全威胁，包括异常行为、恶意软件等，评估组织的网络安全风险状况，并提供降低风险的建议。

知识库和情报集成：整合来自多个来源的安全知识库和最新威胁情报，提供防护策略建议，帮助用户了解当前威胁环境和最佳实践。

可视化与报告生成：将复杂的安全数据和分析结果以直观的图表和报告形式展示，帮助安全分析师快速将安全状况转化成报告。

个性化学习与适应能力：根据用户的行为和组织的网络环境进行自我学习和优化，逐渐了解特定组织的业务流程和安全需求，提供定制化的安全建议和服务。

安全策略管理：根据用户需求制定、实时调整安全策略，例如动态管理用户身份和访问权限等，自动调整安全配置，以适应不断变化的威胁环境。

国外安全厂商已经陆续推出多款AI安全助手产品，例如Microsoft Copilot for Security/Fortinet Advisor/Crowdstrike Charlotte。

1）Microsoft Copilot for Security

Microsoft Copilot for Security于2024年4月1日正式上线，是一款基于OpenAI LLM大模型的安全助手，可帮助安全人员以更快地速度，更高的准确度发现、识别并处理网络威胁。

借助生成式AI技术以及融合微软情报体系，Microsoft Copilot for Security可快速处理大量数据，发现潜在的安全风险并进行深度威胁分析，确定风险处理优先级排序，为安全人员提供安全分析和决策支持，提升安全团队响应速度，从而提升整体安全能力。

Microsoft Copilot for Security扮演着“安全专家”的角色，安全人员只需以自然语言描述即可获得专业化能力支撑，通过自动化执行安全任务可有效减少低级/高重复度工作，简化了安全操作过程，提高安全部门的工作效率。

Microsoft Copilot for Security能够理解并回应8种语言，产品界面支持25 种不同语言，具有个人化提示手册（Custom promptbooks）、知识库整合（Knowledgebase integrations）的预览版，可与Microsoft 的多个安全产品集成，可通过插件与第三方服务集成。

Microsoft Copilot for Security主要应用场景包括威胁检测与响应、安全合规性管理、安全数据分析、安全培训与模拟、安全自动化工作流以及安全咨询等。

2）Fortinet Advisor

Fortinet Advisor 是 Fortinet 最新推出的生成式 AI 安全助手，旨在帮助安全运营（SecOps）团队加速威胁调查和修复，提高安全团队的技能和效率，从而增强企业整体的网络安全防护能力。

Fortinet Advisor 能够帮助安全分析师构建复杂调查查询，通过自然语言输入提问即可生成问询结果，能够快速分析安全告警，生成易于理解的事件摘要，提供威胁修复计划建议，帮助安全团队快速响应威胁，并可根据分析师的实时反馈进一步完善所建议的响应计划。

安全架构师可以通过咨询 Fortinet Advisor 来生成 Playbook 模板，快速将流程转化为可执行的计划。目前Fortinet Advisor 已无缝集成至 Fortinet 安全信息和事件管理解决方案 FortiSIEM 以及 Fortinet 安全编排、自动化和响应解决方案 FortiSOAR 中，帮助用户将威胁识别和遏制所需时间从 20 多天锐减至 1 小时内，并将威胁调查和修复时间从 18 小时以上缩减至 15 分钟以内。

3）CrowdStrike Charlotte

CrowdStrike Charlotte是CrowdStrike重磅推出的生成式AI安全助手，通过自然语言处理技术，理解并回答用户的问题，提供快速、准确的信息，旨在提高网络安全团队的效率。

CrowdStrike Charlotte可以自动执行数据收集、提取和基本威胁搜索等任务，简化高级安全操作，也可以执行实时响应脚本，帮助用户快速消除威胁，例如杀死进程、隔离可疑文件等。具有独特的人类验证内容数据集，包括来自 Falcon OverWatch 管理的威胁狩猎、Falcon Complete 管理的检测和响应、CrowdStrike 服务和 CrowdStrike 情报的反馈。

CrowdStrike Charlotte的核心优势之一，通过提高安全操作的效率和自动化响应能力，将检测到响应的时间从小时甚至天数缩短到了几秒钟，显著提升了客户的安全防护能力，并减少了人工操作的需求，释放了更多的人力资源用于战略性任务。所提供的答案都可以追溯到原始数据，确保透明度和可审计性，用户可以检查底层源数据，确保AI提供的见解和推荐可信。

AI驱动的安全平台

以AI驱动的安全平台是AI安全另外一种重要的应用场景，集成了包括AI技术在内的安全解决方案，通过机器学习、深度学习、自动化处理、数据分析等技术来提升平台的整体安全能力。和AI安全助手产品相比，AI安全平台具有更加全面的功能和更加广泛的范围。

AI安全平台旨在提供一个更加智能和主动的安全防护机制，以应对日益复杂和隐蔽的网络攻击，通常具备高级威胁检测、自动化安全任务、智能安全运维、工具集成和拓展、高级分析和响应、统一管理控制台等特点。

高级威胁检测：实时监控网络、系统和设备，快速检测威胁，包括未知威胁如零日攻击，分析历史数据预测潜在威胁，采取更主动防御措施。

自动化安全任务：自动化执行重复性任务，如日志分析、漏洞扫描和补丁管理，减轻安全团队负担。

智能安全运维：可实现从安全分析到响应处置全流程闭环，将AI融入事件分析研判、可视化编排等技术之中，通过与专家经验进行有机整合。

工具集成和拓展：AI安全平台一般会集成多种安全工具和流程，实现自动化的威胁检测、响应和修复，同时支持扩展，以适应不同规模的组织和不断变化的安全需求。

高级分析和响应：平台可以执行高级的安全分析，包括用户和实体行为分析（UEBA）、安全信息和事件管理（SIEM）等，并能够自动响应安全事件。

统一管理控制台：平台提供统一的管理控制台，使得安全团队能够从一个中心位置监控和管理整个安全态势。

目前海外已经有头部安全厂商陆续推出多款AI安全平台产品，例如Palo Alto Networks公司的Cortex® XSIAM（扩展安全智能和自动化管理）、Vectra 公司的 Cognito平台和CrowdStrike公司的 Falcon平台。

1）Cortex® XSIAM平台

Cortex® XSIAM平台是一个云交付的集成SOC平台，将EDR、XDR、SOAR、ASM、UEBA、TIP和SIEM等多个产品和功能整合到一个集成平台中，底层完全由AI驱动，统一了一系列的安全能力。该平台通过扩展型安全智能和自动化管理，将各种基础架构遥测数据转化成智能数据基础，加快威胁响应速度。

XSIAM利用机器学习模型来识别和响应安全威胁，提供快速准确的威胁检测和响应能力，能够整合来自不同来源的安全数据，并自动化处理大部分安全事件，减少人工干预的需求。它还提供了对云资产、事件、覆盖范围和漏洞的可见性，并通过与Prisma Cloud集成以增强事件分组与导航功能。

XSIAM 2.0增加了全新的自定义机器学习（BYOML）框架，允许用户集成定制的机器学习模型，实现第三方EDR数据的出色整合，并且还可以充分利用云检测和响应功能。

2）Vectra Cognito平台

Vectra是AI驱动网络安全的头部企业，旗下Cognito平台以AI为驱动，结合机器学习、深度学习等技术，可以自动执行威胁响应任务，减少人工干预的需求，提升检测和响应网络攻击灵敏度的安全解决方案。

通过自动化威胁检测和响应，Cognito平台有效提高了SOC（安全运营中心）的效率和响应速度，并且可以根据威胁的潜在影响和严重程度自动对检测到的威胁进行优先排序，极大地降低了安全团队的负担，帮助安全团队将精力集中在最关键的事件上。

Cognito 平台的 AI 引擎还可分析大量网络元数据，包括流量模式、用户行为和云交互，能够检测已知威胁以及可能逃避传统安全措施的潜在攻击的细微指标。

3）CrowdStrike Falcon平台

CrowdStrike Falcon平台是一款基于云原生架构和集成的XDR解决方案提供全面的终端安全服务，是CrowdStrike公司的明显产品，引入AI技术驱动增强了该平台的能力，显著提升了安全运营的效率和响应速度。

通过AI和机器学习算法，Falcon平台能够实时监测和分析网络流量，自动检测可疑活动，帮助组织及时响应安全事件。此外它还能够实时检测高级威胁，提供比传统微软解决方案更高的检测准确性和响应速度。

AI技术的引入使得Falcon平台可以更顺畅地自动执行威胁响应任务，减少人工干预的需求，提高安全运营的效率，进一步将从检测到响应的时间从小时甚至天数缩短到了几秒钟。

（二）国内AI安全产品分析

国内AI安全产品发展较欧美市场晚，尽管2023年为萌芽之年，当年也有十余家安全企业发布AI安全产品，2024年AI安全产品的竞争将全面铺开。头部网络安全厂商的战略部署中，AI安全产品是不可或缺的一环。

在大模型落地方式上，以私有化本地化部署为主，同时考虑算力成本以及大模型本地实施需求，以及国内政府、金融等头部客户对数据安全要求较高，训练/推理一体机的模式可能成为主流，但混合部署模式易带来管理、维护、运营效率低下等问题，未来云安全一体化仍然是重要趋势。

目前来看，国内AI安全产品的模式主要分为三类，分别是AI安全助手、AI驱动的安全平台和AI赋能安全产品，后者是通过将AI技术与安全产品深度融合，全面提升安全产品的能力。例如以AI技术赋能威胁检测与响应、身份认证和行为分析、反欺诈与风险控制、恶意行为检测、安全托管服务（MSS）、SecOps、数据安全等。

AI安全助手产品

作为AI安全典型安全产品形态，目前国内安全厂商已经陆续推出多款AI安全助手产品，例如天融信打造的AI安全助手小天、启明星辰研发的AI安全助手盘小古等。

1）天融信AI安全助手小天

天融信AI安全助手小天基于天问大模型系统，具备自然语言理解、命名实体识别等能力，能够智能分析客户行为、快速响应客户需求。同时，小天具有良好的兼容性，能够与不同厂商的AI系统和平台无缝对接，降低了客户的迁移成本。

小天产品形态分为两种：一是云上小天，提供天融信产品、技术、解决方案、行业知识等服务；二是产品小天，以组件的形式嵌入到各类网络安全产品中，以智能问答的形式，提供告警分析、处置建议、报告生成等功能。

小天的核心功能如下：

安全Copilot：通过交互式对话提供各种安全问答与能力订阅，覆盖网络安全解决方案、产品功能配置、故障排查等各类场景。产品小天与天融信脆弱性扫描与管理系统、数据库审计与防护系统等产品深度融合，可通过对话的方式下发检测任务、获取评估结果、查看异常事件、了解事件类型分布，掌握网络中的整体资产和风险情况，简化人工操作，提升运维管理效率。

安全运营：与天问大数据分析系统融合后，可实现基于AI的海量数据分析和智能AI交互能力，自动输出图表或文本式结果，解读告警信息等能力。

知识管理：小天依托天问大模型能力，通过对网络安全垂直领域专业知识的不断学习和积累，持续提升问答质量。

2）启明星辰AI安全助手盘小古

盘小古是一款基于AI的网络安全运营智能助手，旨在通过自然语言处理、大数据分析等技术，提升网络安全运营的效率和效果。

基于安全行业的特点，盘小古发展了独有的基于AI安全智能体的安全垂直领域“大小模型自主协同”技术体系，把启明星辰深厚的安全产品能力和AI专项安全小模型积累，与大模型强大的意图理解和推理能力进行融合，构建了由安全大模型驱动的全场景安全智能自动化运营中心。

盘小古适用于以下场景：

知识问答场景：提供网络安全基础知识、防护措施等问题的答案和建议；

威胁分析场景：通过接收告警数据识别威胁，帮助用户快速了解安全告警、风险资产等；

安全生成场景：生成威胁态势概览、安全事件详细分析报告等；处置响应场景，发现安全威胁时立即触发预设的响应动作。

AI驱动的安全平台

AI驱动的安全平台也是国内发力的重要方向，国内安全厂商中绿盟科技、云起无垠等已分别发布了风云卫AI安全能力平台、无极 AI 安全智能体平台。

1）风云卫AI安全能力平台

绿盟科技“风云卫AI安全能力平台”是一个集成了绿盟科技多年AI与机器学习研究经验、攻防知识与威胁情报积累、实战化专家能力于一体的AI安全平台。其内置了多种大小模型、知识库、情报库，支持本地安全知识应用以及基于AI Agent的模型能力拓展。

平台将AI能力原子化融入到安全运营体系，有效促进安全数据、分析能力与运营决策三者有机结合，能够统一调度各类AI安全应用，执行实时监控网络环境、自动分析日志数据及识别异常行为等功能，迅速定位潜在的安全威胁。

此外平台还可实现安全风险自动化识别、告警、研判、拟定处置建议、生成对应事件处置脚本等，并在事后进行复盘，深入挖掘安全事件的根源，揭示隐藏的攻击路径与关联事件，为后续加固提供支持，并自动生成全面、专业的网络安全事件研判报告。

平台适用于以下场景：

安全运营智能化场景：提升安全运营的响应时间和工作效率，适用于面对日益增长的安全威胁和数据的传统安全运营方式。

威胁情报智能化场景：整合、精炼高价值内容，并根据应用场景对情报进行深度的挖掘、演绎、推理和应用。

攻防对抗智能化场景：辅助智能渗透测试，攻防对抗博弈，集成各类攻防工具，提升对抗效率。

2）无极 AI 安全智能体平台

无极 AI 安全智能体平台是云起无垠基于“云起 AI 安全大脑”自主研发与训练的AI安全智能体平台，涵盖安全知识问答、智能模糊测试、代码分析与生成、漏洞威胁情报等功能，以帮助企业自动化完成各类安全任务。

平台具备内容追溯和记忆库功能，支持13种主流编程语言，具有多模态分析能力，能够处理文本、图像、源代码、二进制文件等多种数据格式，还能在函数级、文件级和仓库级等不同层面进行安全分析。

平台提供安全知识问答、智能模糊测试、代码分析与生成、漏洞威胁情报、智能攻击面探测、攻防技战法辅助、钓鱼邮件生成、智能文档分析、渗透测试等功能。

平台适用于以下场景：

开发安全场景：平台通过高效精准的智能闭环，实现多工具的统一接入与调度，提升检测的准确度和效率，并支持缺陷的统一分析与智能修复，提供完整的修复方案。

攻防渗透场景：平台可以高效处理情报数据、进行精确分析，并辅助进行渗透测试，从而替代初级安服工程师，提高整体效率和检测准确度。

知识检索场景：平台通过自动化数据采集，覆盖多个漏洞披露源，利用Text-to-SQL技术进行自然语言分析，并对接业务资产库，实现漏洞自主清查、验证和通告。

AI赋能安全产品

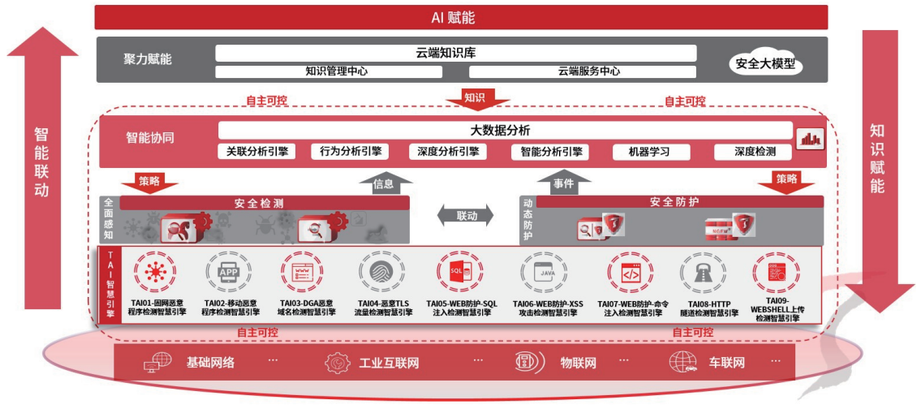
AI赋能安全产品是指将AI技术深度融合于安全产品之中，再加上机器学习、深度学习等技术的使用，从而实现更高精度的安全风险识别，或者提高安全监测、分析与响应等工作的自动化程度，大大增强了原安全产品的能力，以应对日益复杂的网络安全威胁。

目前，国内大部分安全厂商都在发力AI赋能安全产品赛道，应用场景十分广泛，涉及数据安全、内容安全、业务安全、终端安全等，包括威胁检测与响应、恶意行为检测、智能安全运营等。

例如天融信基于多年在AI领域的研发和积累，2024年，在下一代可信网络安全架构（NGTNA2.0）下，将 AI 技术与网络安全能力深度融合，发布天问系列产品，实现 AI 全面赋能,主要应用于威胁检测、安全运营、知识问答等场景。

1）AI和网络安全融合框架

天融信持续探索“AI+安全”融合发展之道，将AI技术与天融信下一代可信网络安全架构（NGTNA2.0）深度融合，通过将加密流量检测引擎、隐蔽隧道检测引擎、关联分析引擎、行为分析引擎等AI引擎应用到网络安全及大数据产品中，大幅提升产品安全检测与分析能力。同时，将自然语言理解、命名实体识别等技术应用到安全知识库、威胁情报等业务中，结合云端知识库，全面提升安全防护能力。

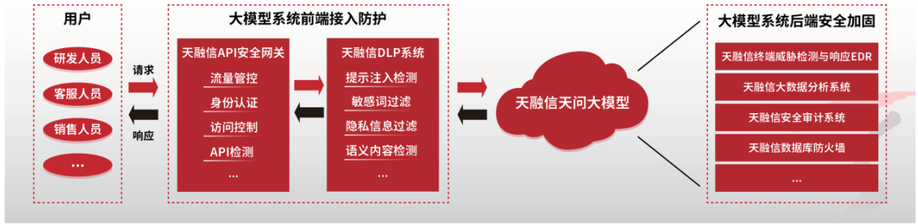


在威胁检测方面，天融信具有九大TAI检测智慧引擎，包括DGA恶意域名检测、隐蔽隧道检测、恶意TLS流量检测、移动恶意程序快速检测、命令注入检测、固网恶意程序检测、WEBSHELL检测、XSS攻击检测、SQL注入检测等检测引擎，可集成在天融信防火墙、入侵检测、入侵防御、僵木蠕监测、APT监测等网络安全产品中，增强未知威胁检测能力。

在安全运营方面，天融信天问大模型系统具备自动化研判能力，大幅提升安全日志研判效率，实现更加精准和便捷化的安全运营；内置产品小天，让安全人员用自然语言方式获取事件告警、资产、漏洞、日志等信息；面向天融信其他网络安全产品提供知识问答服务，快速获取威胁情报、漏洞解读等安全知识；与天融信防火墙、入侵检测、入侵防御、僵木蠕监测、WAF、漏扫等设备联动，可获取系统所需的数据，提升全面感知、动态防护的安全能力。

2）天问大模型安全防护方案

天问大模型是天融信基于大模型技术及安全大数据语料构建的安全领域大模型，主要功能包括：安全知识问答、威胁情报查询、威胁分析等。天问大模型作为软件系统部署在云环境中，并以云服务方式提供访问，通过云管理平台实施运行维护。为了防止模型被滥用或恶意利用，天融信提出了前后端联合的安全防护方案，整体防护架构如下：



天问大模型安全防护架构

天问大模型防护方案由大模型系统前端接入防护和大模型系统后端安全加固两部分构成。前端接入防护主要实现大模型安全接入和数据泄露防护；后端安全加固实现对大模型系统核心软硬件威胁检测与分析、系统操作安全审计及容灾备份功能。天融信API安全网关可以提供流量管控、身份认证、访问控制、API检测等功能，有效拒绝非法访问用户和流量。天融信数据防泄漏能够提供prompt注入检测、敏感词过滤、隐私信息过滤、基于语义的内容检测等功能，实现大模型输入输出内容的敏感数据识别、拦截。除上述安全防护措施外，天问大模型系统还引入人工审查机制，保障大模型输入输出的内容安全性，降低不良信息输入和输出概率。

3）AI应用安全防御产品：动态安全Botgate

动态安全Botgate是由瑞数信息自主研发，以“AI智能威胁检测+动态安全”为核心技术能力的AI应用安全防御解决方案。该方案涵盖了机器学习、智能人机识别、智能威胁检测、全息设备指纹、智能响应等多项AI技术，变革传统安全依赖攻击特征、行为规则特征及阈值的被动式防御技术，通过对服务器网页底层代码的持续动态变换，增加服务器行为的“不可预测性”，为企业提供提供面向应用和业务层面的全面高效防御能力。

瑞数动态安全Botgate旨在保护企业免受恶意机器人和自动化攻击的侵害，主要功能包括：

网站漏洞隐藏：使漏洞扫描攻击无法获取漏洞信息，使漏洞利用工具无法完成有效攻击，在网站未打补丁和补丁空窗期提供有效安全保护；

网站代码隐藏：可以针对网站底层的代码进行封装，使攻击者无法分析网站应用的源代码；

应用安全威胁防护：有效防止SQL注入、越权访问、跨站攻击、网页后门等攻击行为；

自动化攻击防护：有效防护攻击者利用自动化攻击脚本或工具对应用发起的漏洞探测与利用，以及针对新兴LLM应用的提示词注入、越狱及欺诈滥用等攻击行为；

模拟合法操作防护：可有效应对攻击者利用工具、合法身份，模拟正常人工访问的攻击行为；

数据防泄漏：可以有效防护恶意人员使用工具或脚本程序通过前台应用批量获取数据的攻击行为。

国内AI安全产品典型应用案例

（一）天融信--AI大模型在金融行业安全运营场景的应用实践

案例背景

某基金管理有限公司是国内业务资质全面、产品种类丰富、经营业绩优秀、资产管理规模领先、业务发展均衡的基金管理公司之一，为近1亿境内外个人和机构投资者提供财富管理服务。截至2024年6月30日，该公司旗下管理数百只公募基金和多个年金、私募资产管理计划，资产管理总规模近2万亿元。

为满足数字化、智能化时代的网络安全运营工作需要，该基金管理公司规划建设智能化安全运营中心。天融信基于客户需求与业务逻辑，以依法合规为基本原则，以具备智能化分析引擎的运营平台为底座，融合天问大模型帮助客户实现高效、智能安全运营。

问题与需求

当前，客户面临新业务不断拓展与新型技术频繁应用的双重挑战，导致传统网络边界不断泛化。在数字化转型的浪潮中，数据资产与信息系统资产呈现出爆炸式增长，而与此同时，新型攻击手段、僵尸木马、勒索挖矿病毒等安全风险也层出不穷，且愈发多变、隐蔽和仿真，给客户的安全运营工作带来了巨大冲击。持续产生的海量运维日志和安全告警日志极大加重了网络安全工作负担，同时海量资产的检查、监控、安全事件处置等管理工作也使运营团队应接不暇。此外，智能化工具催生的新型攻击武器和手法等难以发现和溯源，更是对安全运营团队的能力和传统工具构成了严峻挑战。

为了应对挑战，客户亟待建设一套可以全面掌控全网安全动态、智能化识别新型安全风险，并对风险、漏洞进行自动流程化闭环管理的智能化安全运营工具。因此，依托于大数据分析系统和天问大模型系统，天融信帮助客户设计并建立了智能安全运营平台，以实现对安全风险的高效智能化管理。

实施内容及效果

本案例基于安全运营工具智能化的需求，对客户的安全运营业务流程、IT架构和安全体系进行了全面的调研，采用天融信大数据分析系统+天问大模型系统搭建智能安全运营平台的技术方案。该方案主要实现以下三个建设目标：

1)打造智能化安全运营中心

在传统安全运营平台的基础上融合各类人工智能模型，实现了对资产安全管理、安全检测和监测以及风险处置决策等功能的全面智能化改造。从不同维度对资产进行统一管理，依靠智能模型进行自动化分类，并与安全基础设施、安全基线、数据分类分级库等多维度实体关联。将安全策略与业务流程紧密结合，使客户更全面地了解自身的安全风险状况，并能够根据业务需求和风险变化动态调整安全策略，实现安全与资产的深度融合。通过大模型+小模型的综合应用，实时分析海量安全数据和运营数据，精准识别和预警潜在的安全威胁，并结合资产信息高效响应处置，有效降低了企业因安全事件导致的业务中断和数据泄露风险，保障了业务的连续性和稳定性，将安全转化为企业数字业务的重要组成部分。基于大模型智能决策模块，综合历史经验数据、威胁情报数据和实时信息，自动生成处置建议和应急预案，为安全专家提供科学决策依据。利用融合安全规范、安全最佳实践、前沿技术等信息的安全知识库，由大模型根据专家意图进行信息抽取生成，为管理层提供基于安全现状的运营分析，更加精准地把握安全风险，辅助制定更加有效的安全策略，实现安全运营与管理决策的高效同频。

整体上智能安全运营中心实现了对安全威胁的实时监测、预警、处置响应和运营管理，极大提升了安全运营的效率，有效缓解了数字业务飞速扩展与安全运营团队资源有限之间的不平衡矛盾。

2)可支撑海量数据管理和智能化分析

智能化的数据管理与分析检测是重塑安全运营模式的重要基石。将全网资产信息、安全威胁信息、业务运行日志和安全威胁情报、企业/行业舆情等数据汇聚到智能安全运营平台，确保了数据的全面性和时效性。在数据治理方面运用先进的数据分类分级技术和全生命周期管理策略，通过清洗、富化等处理，构建了高质量的元数据体系，为后续智能模型数据分析奠定了坚实基础。依托天问大模型与多种安全专用智能分析引擎，显著增强了威胁分析能力。大模型基于语义理解，能够精准识别复杂攻击手段，如钓鱼邮件、流量威胁识别、恶意软件等，无需验证样本即可实现高效检测与定位。同时，多种专用模型，如安全知识图谱、用户异常行为分析等，为深度挖掘潜在威胁提供了有力支持。

通过告警降噪技术自动识别和过滤大量重复、低效的告警信息，有效减少了安全运营人员的负担，并提高了对真正安全事件的响应速度和准确性。智能引擎调度管理模块根据任务风险等级智能分配资源，实现对重点风险的优先分析与处置，在面对大规模攻击时有效集中力量防御重点威胁。

相较于传统方式，人工智能技术大幅提升了海量数据的管理效率与威胁检测能力，为企业安全运营提供了强有力的技术支撑和安全保障。

3)自动化流程管理提高运营效率

智能安全运营平台具备松耦合接入模块，支持多种协议和第三方系统接入，确保系统的高度灵活性和可扩展性。通过该模块，平台能够无缝集成各类业务管理系统和安全设施，实现多种安全运营工作的自动化。例如：面向威胁事件处置，大模型+SOAR可以显著增强响应能力，大模型基于广泛的安全知识库，迅速生成针对性的威胁处置预案，经过与安全专家交互调优后，驱动SOAR模块能够自动触发预设的响应流程，对事件进行快速、准确的处置，有效缩短安全响应时间，降低安全风险。

面向安全巡检，平台能够定期自动化执行资产检查、漏洞扫描等任务，由大模型生成详细的检查报告，并自动推送漏洞信息至相关系统，触发漏洞修复操作，同时跟踪修复情况，形成自动化漏洞管理闭环等。与业务系统融合的自动化流程，使安全运营工作效率大幅提升，同时降低人力资源依赖和人为操作失误损失。

关键成功因素

在实践智能安全运营平台的过程中，金融行业的优质数据是保障AI模型高质量运作的基础。同时业务中对业务目标和判断风险标准也至关重要。在实施过程中天融信与客户安全部门紧密合作，梳理出一系列的风险分析模型、自动处置脚本等工具成果，并成功应用于日常安全业务实践。

在建设层面，因为金融行业涉及大量个人敏感信息，而AI技术应用存在黑箱、偏见、信息泄露等安全风险，大模型自身也存在运营成本过高、技术集成困难等风险。因此在项目建设过程中，详细梳理各种AI应用风险并建立保障措施，从使用安全、成本控制、技术能力保障等多方面为客户建立信心，将新型技术用于实践。

实施收益与反响

1）提升安全管理水平

智能安全运营平台提供了安全管理和防护能力，覆盖了安全态势感知、安全风险监测与安全运营管理等多个方面，同时也辅助安全专家对技术、流程等进行完善，提升了安全管理水平。

2）增强检测防护能力

智能安全运营平台具备预测、分析、生成、推理等核心能力，可对恶意邮件、恶意网页、恶意样本、代码逻辑漏洞等进行智能识别和预测，使企业可以更快速地发现潜在的威胁，并采取措施进行防范。

3）提升安全运营效率

智能安全运营平台具备安全分析、风险预测和决策行动等核心能力，以智能助手形式赋能于安全产品中，安全人员通过对话方式即可获得更精准的预测、研判结果和决策执行指令，使得安全运营更加高效。

案例总结与推广价值

1）探索新技术应用可起到良好的行业示范作用

将网络安全技术与人工智能技术相结合是安全创新的重要途径，通过不断地探索新技术、新产品和新模式，迭代网络安全能力体系，推动科技创新的持续发展。新技术的成功应用，可以激发更多的创新灵感和动力，推动金融科技创新水平不断提升。

2）经过不断实践可使人工智能技术更好服务网络安全

技术的成熟和广泛应用是相辅相成的。一方面，技术的成熟需要不断的实践和改进；另一方面，广泛的应用又为技术的进一步成熟提供了更多的数据和反馈。通过“抛砖引玉”的形式，让更多企业在已有的方案和成果上实践，可让这项技术更加的成熟，从而推动技术的广泛普及和深入应用。

（二）蚂蚁集团--切面融合智能在威胁检测的应用

案例简介

现有的安全威胁检测技术普遍存在误报多、未知威胁发现能力弱、可解释性差等问题。究其原因一方面是由于安全内视能力不足，无法深入系统内部获取足够的数据，另一方面是由于智能化水平不足，传统基于规则的检测系统通常无法兼顾检出率和误报率。

针对上述威胁检测中的问题，蚂蚁集团通过切面技术将系统、网络、应用等多维度的数据引入到威胁检测领域，通过安全大模型将专家智能和机器智能快速融合，并依托安全大模型自身的行业知识，高效检测出网络威胁并对之进行研判。具体方式如下，首先通过切面技术获得系统、网络、应用等多维度数据，并提取多维度数据中网络主客体凭证进行关联分析构建出完整攻击链路。然后，安全大模型融合专家智能与机器智能，取长补短，发挥算法和专家各自的优势，提升未知威胁的发现能力；同时构建以应用为主体的安全基线，安全大模型基于基线和专家经验识别正常的业务行为，降低告警误报。最后，使用系统、网络、应用及情报数据构建安全对抗知识图谱，提升模型的可解释性，实现攻击模式的发现与溯源。

代表性成果介绍

“切面融合智能与威胁检测技术”方案分为4个模块，功能如下：

1）基于攻击链路的安全切面数据关联分析能力

该模块有2个主要功能，涉及网络主客体凭证关联方式，将网络安全领域中多种类型的数据有效关联。主客体凭证关联分析和链路构建，是将一次网络访问行为的主体、客体、凭证所关联的数据串联起来。例如，一次web服务的访问，主客体链路关联分析需要将网络访问的发起IP、入网凭证、网络访问web站点这三者关联起来，形成一个有效链路。

2）基于安全大模型与融合智能的未知威胁发现能力

利用安全大模型对安全原语义强大理解能力和问题处理能力，融合专家经验和机器智能特征，更准确检测出网络风险。专家经验一般属于强特征，它可以快速检测出网络风险，但泛化能力弱；机器智能属于弱特征，单个机器智能特征无法有效刻画并精确检测出风险，多个机器智能特征的组合可以检测出风险并且相对专家有较高的泛化能力。将专家经验与机器智能结合属于网络安全领域的一项大胆的创新，两者优势互补后有望在网络安全领域的威胁检测取得突破性进展。

3）基于业务自学习的告警降误报能力

该模块将历史网络、流量、主机数据与应用数据进行关联，得到各应用的行为数据。再通过行为基线生成算法，形成诸如命令基线、进程基线、搜索关键字基线等各类正常业务行为的基线。然后，在网络威胁告警运营的过程中，依托大模型的能力将告警数据按照特定的模糊匹配算法同业务行为基线进行关联匹配，能够实现匹配的即为正常业务行为，可以直接过滤，达到告警降误报能力。

4）用于提升模型可解释性的安全对抗知识图谱

将ATT&CK战术技术、威胁情报信息融入知识图谱的图节点属性中。基于构建的知识图谱，利用图谱推理和关联分析技术，检测潜在的威胁模式和关联。这使威胁检测系统能够发现不易察觉的威胁，有助于提前预防和干预威胁事件。这种技术路线技术路线在威胁检测具备高度的先进性和可解释性。

应用背景及实施效果

随着全球数字化转型的浪潮持续涌动并不断加快步伐，网络空间和数字化业务已经日益成为经济发展的核心载体与驱动力，其承载的价值量级呈指数级增长。然而，伴随着的是网络威胁风险态势的日趋严峻复杂，恶意攻击手段日新月异，网络安全事件频发，对国家的信息安全构成了前所未有的挑战。

蚂蚁集团作为领先的金融科技企业，在其广泛的业务范围内，涉及的信息化系统数量庞大且结构复杂，涵盖了生产环境网络、研发测试环境网络、内部办公网络等多个维度，连接着数以百万计的各类IT资产，如服务器、终端设备等。每日产生的网络安全数据异常庞大，仅入侵威胁检测实时探针日志每秒钟就高达千万条之多，全天候累积的数据总量更是达到惊人的万亿级别。在如此庞大的数据洪流中，蚂蚁集团面临着极其艰巨的安全防护任务，既要确保能够从海量数据中实时精准地发现潜在的安全威胁，满足高效、敏捷的威胁对抗实时性要求；又要能够在面临突发安全事件时快速响应、有效应变，构建起一套既能灵活应对又能稳固支撑业务连续性的安全体系。尤为关键的是，鉴于蚂蚁集团所处的金融领域对安全性的极高要求，其网络安全工作必须具备金融级别的稳定性和可靠性，以维护用户信任和社会责任，从而在数字化转型的大潮中扮演好保驾护航的角色。

在此背景下，为了保护企业的核心业务和敏感数据，蚂蚁集团建立了以“安全平行切面融合智能”为核心的多层次网络安全纵深防护体系，该体系通过人工智能技术与安全平行切面技术的深度融合，利用基于攻击链路的安全切面数据关联分析能力，实时捕捉潜在的威胁信息。同时，借助于基于融合智能的未知威胁发现能力，该体系能够有效应对新型、未知的安全挑战，提升整体安全防护的前瞻性和有效性。在此基础上，该体系还具备业务自学习的告警降误报功能，通过自我学习和降噪，大幅减少无效的安全告警，显著提升风险运营的效率。

在历次演练和实践中，该体系对异常行为自动化研判率达到95%以上，能够有效应对新型、未知的安全挑战，提升整体安全防护的前瞻性和有效性。相关成果曾获评2024 WIC Find智能科技创新应用典型案例、2022年度上海市网络安全产业创新攻关成果，入选《2024人工智能先锋案例集》及大模型应用落地“样板间工程”优秀案例等。

应用实施难度与复杂性

蚂蚁集团切面融合智能系统经过多次迭代和演进。

1）通过安全平行切面将安全能力系统化地融入技术基础设施与应用服务内部；

2）基于ATT&CK攻击知识框架体系，对攻击行为、攻击阶段等维度提供统一刻画，围绕行为画像、威胁情报等多源异常联合风险研判分析定性；

3）通过离线数据分析平台/知识图谱平台/模型服务平台以及大模型服务平台，将智能威胁检测和研判能力融入到入侵检测体系。

蚂蚁切面融合智能系统需要对蚂蚁集团内所有业务提供入侵检测服务，因此系统处理蚂蚁集团内部生产网、研发测试网、办公网、终端等百万级IT资产的日志数据，每秒处理千万条入侵威胁检测实时探针日志，每日高达万亿条。与此同时，还要满足威胁对抗的实时性要求、快速应变要求，以及金融级别的稳定性和可靠性。总结来说，蚂蚁集团入侵检测系统面对的挑战主要包括万亿级超大规模数据的实时检测与离线分析、安全攻防强对抗的快速应变、金融级别的稳定性和可靠性、复杂安全环境的入侵检测水位量化。

市场影响力与推广性

基于人工智能的智能安全切面技术在市场上的影响力体现在以下几个方面：

首先，有助于减少网络犯罪和数据泄露事件的发生，企业可以更好地保护客户数据和业务机密，从而降低潜在的经济损失。

其次，能够促进安全技术和人工智能领域的产业发展。这将有助于减少安全事件的损害范围，降低恶意攻击造成的社会成本。

此外，对于安全专业人士而言，智能安全切面技术的发展为他们提供了更强大的工具和资源。

智能安全切面技术的推广性在于其应对网络威胁的全方位能力和跨行业的适用性。现在各个行业都迫切需要强大且智能化的安全防护系统。智能安全切面不仅能够为企业提供个性化、高效的安全解决方案，还能够适应各种规模和需求的变化，使其成为市场上的通用防护产品。其整合人工智能的特性，使得它能够不断学习和适应新的威胁模式，增强了其在市场上的竞争力和推广潜力。随着安全意识的提高和技术的不断进步，智能安全切面的应用和需求预计将持续增长。

（三）瑞数信息——LLM平台滥用下的AI应用安全防御

案例背景

随着大型语言模型（LLM）在企业应用中的迅速普及，相关的安全威胁也日益增多，呈现出多样化和复杂化的趋势。这些威胁不仅对企业的经济利益和技术基础设施构成威胁，还可能导致用户体验受损和合规风险增加。

瑞数动态安全Botgate作为国内最早以AI为核心能力的应用安全防御产品之一，旨在保护企业免受恶意机器人和自动化攻击的侵害，并特别关注防御那些针对大型语言模型（LLM）平台的滥用行为，例如通过非法反向代理和未经授权的平台副本进行的攻击。帮助企业维护其服务的完整性和安全性，同时保护客户数据不受侵害，确保业务的连续性和品牌声誉。

应用场景与防护策略

瑞数动态安全Botgate适用于多种应用场景，并能提供针对性的AI防护策略。

1）LLM平台滥用威胁

攻击者通过构建未经授权的平台镜像并利用反向代理技术绕过地域限制，实施匿名化攻击。这种复杂的攻击手法不仅涉及知识产权侵犯，更可能演变为更具规模的网络犯罪行为。

具体而言，攻击者通过部署平台克隆站点、突破地理访问限制并借助代理服务隐匿身份，以盗用平台技术资源、构建仿冒服务并规避安全防护。此类非法行为显著增加了LLM应用的算力消耗和运营成本，同时可能被滥用于自动化生成钓鱼内容、制作深度伪造媒体等欺诈活动，不仅严重损害品牌声誉，更将对企业和个人网络安全构成重大威胁。

瑞数动态安全Botgate防御策略： 通过部署多层检测机制和实时风险评估引擎，及时识别并拦截未授权的平台接入行为，有效甄别和阻断可疑代理服务器流量，全面保护客户数据安全，切实维护平台知识产权与品牌声誉。

2）提示词注入攻击威胁

攻击者通过部署恶意机器人，实施系统化的提示词注入和数据抓取，能够大规模窃取机密信息、企业核心知识以及个人身份信息（PII）等敏感数据。此类攻击行为不仅直接导致数据泄露风险，还可能衍生出难以预见的安全隐患，严重损害用户对平台的信任，并对企业和个人的安全防护体系造成重大冲击。

瑞数动态安全Botgate防御策略：通过部署先进的机器人行为识别技术，实时监测并拦截异常提示词请求，结合多维度的数据防泄露机制，全方位保护企业与个人核心数据安全。

3）短信欺诈攻击

攻击者利用自动化机器人，批量发起大规模短信欺诈攻击，耗费平台短信资源、造成经济损失的同时，也会增加运营成本，严重损害企业品牌声誉。

瑞数动态安全Botgate防御策略：通过部署先进的机器人行为识别技术，实时监测并拦截异常提示词请求，结合多维度的数据防泄露机制，全方位保护企业与个人核心数据安全。

4）账号盗用攻击

恶意机器人利用撞库攻击或系统漏洞实施账号盗用，从而执行未经授权的非法操作并窃取敏感信息，不仅可能导致数据泄露，还可能引发财产损失，对企业构成重大安全威胁。

瑞数动态安全Botgate防御策略：基于访问行为异常检测技术，构建实时账号安全监测体系。

5）新账号欺诈

攻击者利用自动化工具批量注册虚假账号实施资源盗用，不仅大量消耗系统资源，还可能导致服务性能显著下降。

瑞数动态安全Botgate防御策略：基于智能风险评估模型和异常客户端识别技术，构建严格的账号注册验证机制，有效防控批量虚假账号注册。

核心技术能力

1）AI智能威胁检测

瑞数信息通过自主研发的AI智能威胁检测技术，提升了应对复杂自动化攻击的响应能力。传统的安全检测往往依赖于固定规则和特征库，而此类方法面对不断变化的新型威胁时却变得力不从心。

瑞数信息的AI智能检测体系则打破了这一困境，采用机器学习和大数据分析技术，能够从海量网络流量行为和模式中快速识别异常。通过标记的威胁检测模型，系统不仅能够发现已知的威胁，还能捕捉到那些暂未出现过的未知威胁。

AI智能威胁检测系统能够在攻击发生行为之前，通过对系统日志、用户行为、网络流量的多维度分析，识别潜在的安全隐患。其收敛性机器学习引擎能够自我学习与进化，使得能够检测模型在面对变化多端的攻击手段时，也能保持高度的姿态和准确性。

此外，系统通过全息指纹设备技术，实现对用户终端、网络设备的深度识别，确保精准定位攻击源，快速阻止恶意行为。

2） 动态安全引擎

瑞数信息“动态安全”采用了创新的动态防护技术，包括动态封装动态、动态交互、动态验证和动态令牌。通过这些动态防护机制，系统能够在应用层面增加不可预测性，从而有效阻止攻击者对漏洞的利用。每次访问请求和响应都会被重新包装和加密，使得攻击者无法轻松读取或利用网页底层代码和数据。

动态安全引擎能够根据威胁态势对各类网站应用及业务交易的全过程进行动态感知、分析与预测，精准识别并阻止自动化攻击工具和调试行为，及时追溯与阻断恶意攻击来源，打击伪装正常交易的业务作弊、利用合法账号窃取敏感数据，以及假冒合法终端应用的各类网络欺诈与攻击行为，从而最大限度地主动透视风险，实现安全风险防护。通过持续学习和分析攻击模式，动态安全引擎还可以根据实际威胁对防护措施进行优化调整。这种能力让瑞数信息的动态安全防护系统能够始终保持对新型威胁的有效抵御，特别是在AI自动化面对攻击的持续进化时，动态安全引擎提供了持续升级的防护能力。

3） 全面的安全防护体系

通过整合AI智能威胁检测和安全动态引擎，瑞数信息构建了一套全面的安全防护体系。无论是Web、APP、H5、API接口以及各类混合业务，该体系都对应提供智能化的防护。

瑞数信息的自动化攻击防护体系通过结合AI智能技术，实现了从自动化攻击的精准识别到处理的闭环防护。AI驱动的智能威胁检测和人机识别能力，能够自动化地识别和阻断各种自动化工具和群控行为，提供了针对性强、响应迅速的防护措施。在业务层面，瑞数信息的防护类型包括盗刷、虚假注册、霸占库存、验证码绕过等问题。动态应用保护系统和API动态安全防护系统的结合，能够有效抵御大规模自动化工具的攻击，包括Bots攻击、爬虫攻击等。

这个安全防护体系不仅能够实时响应复杂的自动化攻击，还能够通过威胁情报与攻击者画像引擎对攻击者进行追踪溯源，帮助企业预防未来的潜在威胁。瑞数信息高效的全面防护体系还具备快速的恢复能力，在遭遇攻击后，能够快速恢复系统运行，确保企业业务的连续性。

产品优势

面对传统应用安全威胁的持续演进以及LLM应用带来的新型安全挑战，企业亟需构建全方位的安全防护体系。瑞数动态安全 Botgate通过先进的多层防护机制，为企业提供了强有力的安全保障：

- 全方位的威胁检测能力
- 实时的风险响应机制
- 精准的防护策略执行
- 持续的安全态势感知

实施收益与推广价值

在数字化转型加速的背景下，部署专业的安全解决方案已成为企业保护数字资产、维护业务持续性的关键举措。

瑞数动态安全Botgate定位于为企业提供一个全面高效的防护体系，有效应对传统应用安全威胁的持续演进及LLM应用带来的新型安全挑战。它通过先进的多层次的AI智能威胁检测和动态安全引擎，提供了全面的安全防护策略，适用于多种应用场景，为企业提供了强有力的安全保障，具有强大的核心能力和市场竞争力。

随着AI技术的不断发展和应用，瑞数动态安全 Botgate不仅能够有效应对当前的安全挑战，还能为企业未来的安全建设提供坚实基础。未来Botgate将继续升级和优化，以适应不断变化的安全环境，保护企业的数字资产，维护业务的持续性和稳定性。

全球AI安全发展趋势

AI对于安全产品具有革命性意义

AI在提高网络安全产品能力方面具有革命性意义，当传统安全产品找到与AI的契合点，将会迸发出1+1大于2的效果。以AI防火墙和AI零信任为例。

1）下一代AI防火墙

下一代AI防火墙作为市场中的刚需产品，预计将成为产品创新力最强的安全技术解决方案。融合了AI大模型能力的防火墙可在实际工作中自适应学习，根据实时流量数据持续训练、优化模型，提升威胁检测能力。

AI还将赋予防火墙自动化处置能力，可自动调测策略、分析威胁流量，监控设备状态，结合先进的对话式AI技术，能够通过自然语言与用户进行交互，实现智能运维。

因此，AI大模型可大幅提升防火墙功能，赋予其更强大的生命力，预计将成为网络安全硬件产品中市场规模最大的产品品类。根据市场研究报告，下一代AI防火墙的市场规模正在快速增长，并且预计未来几年将继续保持高速增长。

目前，下一代AI防火墙竞争日趋激烈，主要玩家包括华为、深信服、Palo Alto Networks等，未来将更加注重全流量分析与加密流量检测，以及语音/自然语言处理应用，进一步提高网络安全防护的效率和用户体验。

2）AI零信任解决方案

零信任解决方案在实际落地中存在用户体验不高、手动定义和管理访问策略、资产收集不全等问题，并导致该方案无法被甲方采纳。而AI大模型的加持可有效解决零信任体系在落地时遇到的各类问题，实时分析网络流量和用户行为，自动识别潜在的安全威胁，动态调整安全策略，简化混合云环境的集成，提高了零信任网络的自适应性和智能性。

微软在去年11月举办的Microsoft Ignite 2023大会上提出全面推进零信任战略，明确提出零信任全面渗透到微软的安全策略中。在微软的零信任战略中，人工智能扮演者至关重要的作用，不但用于帮助客户部署成熟的零信任框架，还将广泛赋能零信任技术创新，覆盖了从持续监控、自适应威胁响应到新兴威胁防御的几乎所有零信任技术堆栈，强化信息源、决策引擎和策略执行三个关键组成部分，显著提升了零信任战略的实施效果。

AI将成为网络安全市场的新增长引擎

目前，AI技术在网络安全领域的应用已经逐渐展现出巨大价值和发展潜力，赋能网络安全产品更强的功能和更高的效率。与此同时，越来越多的企业和机构对于AI的高接受度，并逐渐开始部署、使用AI安全相关的产品和解决方案，促使网络安全厂商在AI安全领域投入更多资源开发、迭代AI安全产品，使其更具生命力。

AI网络安全市场的高速增长和AI安全相关企业融资额在网安产业的高额占比，凸显出AI对于网安产业发展的重要性和推动力。未来，随着数字化转型和AI大模型技术的进一步发展，AI安全的产品力将更加凸显，包括帮助终端用户降低安全运营成本；利用多模态模型和自动化操作增强网络安全威胁检测和响应，为企业减少暴露面和攻击面等，这些都将为网络安全市场发展提供新的引擎。

以AI对抗AI的场景将愈发普遍

现阶段，众多企业已经或准备开始运用各类AI工具进行网络安全维护或治理工作。Splunk在《2024年安全现状》报告中指出，93%的受访企业安全负责人表示已在企业或组织中使用生成式AI，但有34%的受访者表示未制定相关的AI政策，

65%的受访者还未充分认识到AI带来的影响。值得注意的是，44%的受访者已将AI列为2024年的首要计划，超过了云安全部署，成为企业安全的首要事务。

在黑客及其他攻击组织视角，AI凭借着分布式、智能化、自动化等特性，也被利用来强化攻击手段，包括自动化攻击流程、高效率发掘并利用漏洞、智能规避防御措施等。出于更加直接的经济动机或政治因素，攻击者往往会展现出更加迅速的AI攻击利用策略，包括开发出多种恶意的AI大模型，进一步增加了传统安全产品体系的压力，并促使防守方利用AI进行反击。

由于在攻守双方的运用，未来的网络攻防将逐渐AI化，但当前仍处于AI赋能网络攻防的发端之际。未来攻守双方将会进一步在技术、数据、模型等层面的找到最佳的结合点，抢占网络攻防领域的“技术差”“应用差”，从而获取对抗博弈优势。

AI自身的安全性将成为行业关注焦点

AI的价值已经成为一种共识。随着AI技术的快速发展及其在各个领域的广泛应用，给大量的行业带来了全新的变革。但安全是AI赋能千行百业的前提，AI想要大规模应用，首先要解决的问题就是如何保障安全。

由于AI还在部分关键基础设施领域部署、应用，因此AI系统的安全性不仅直接关系到个人隐私、企业机密，甚至还事关国家安全。目前AI在系统、算法、数据、应用等层面都还存在脆弱性，无法满足机密性、完整性和可用性的需求。

基于此，AI自身安全的需求十分迫切。从2024年全球AI安全融资表中也可以发现，AI自身的安全性占比高达23%，是最受资本和创业者青睐的细分方向。可以预测，在不久的未来，AI安全将长期占据网安行业的焦点。

AI存在合规挑战

在世界多国越发重视数据安全、确保数据合法合规处理与流通的背景下，对数据高度依赖的AI正面临严苛的合规要求，AI网络安全产品的商业化运用也越发以合规为首要应对场景。在全球市场，欧盟的GDPR（通用数据保护条例）对AI关于数据收集与处理、数据保护与安全、算法透明度与可解释性以及责任与问责等多个方面做了多项规则，而最新的《人工智能法案》要求具有系统性风险的AI模型提供方履行进行风险评估、减轻风险、报告严重事件、进行尖端测试和模型评估、确保网络安全等一系列合规义务。

中国于2023年10月发布《生成式人工智能服务安全基本要求（征求意见稿）》对于生成式人工智能服务的基本安全要求予以规定，包括语料（即训练数据）安全、模型安全等方面的具体要求、人工智能服务提供者应遵循的相应安全措施以及安全评估的程序和内容，其中明确要求生成式人工智能服务提供者在进行算法备案申请前，应先行进行安全性评估，并在算法备案时提交内部评估结果以及证明材料。

免责声明

本报告行业数据及相关市场预测为公司研究院采用桌面研究、行业访谈、市场调查及其他研究方法，部分数据文字采集于公开信息；FreeBuf咨询对该信息准确性、完整性或可靠性做最大的努力追求，但不做任何保证。在任何情况下，本报告中的信息或所表达的观点不构成任何建议。

FreeBuf

中国最具影响力的网络安全行业门户，创立于2010年，提供专业的安全咨询与网安行业研究服务，汇集全球海量安全资讯和深度研究报告。

天融信（002212.SZ）

创始于1995年，是上市公司中成立最早的网络安全企业，围绕网络安全、大数据与云服务三大业务持续创新，以安全护航各行业客户数字化转型。多年来，集团为政府、运营商、金融、能源、医疗卫生、教育、交通、制造等各行业客户提供网络安全产品与服务。未来，天融信将始终以捍卫国家网络空间安全为己任，创新超越，致力于成为民族安全产业的领导者、领先安全技术的创造者和数字时代安全的赋能者。

蚂蚁集团

起步于2004年诞生的支付宝，经过二十年的发展，已成为世界领先的互联网开放平台。聚焦发展，蚂蚁集团形成五大业务增长曲线，即数字支付、数字互联、数字金融、全球化、数字科技。全国工商联发布的2024年中国民营企业500强榜单，蚂蚁集团约1785亿营收位列全国第44位、服务业第12位，研发投入约212亿全国第8位。中国企业联合会发布的2024中国企业500强榜单（同时包含央国企和民营企业），蚂蚁集团位列全国第147位。在安全可信、人工智能、隐私计算等领域获相关科学技术奖6项，包括国家科学进步奖、吴文俊人工智能科技进步奖、浙江省科技进步奖等。

瑞数信息

中国动态安全技术的创新者和Bots自动化攻击防护领域的专业厂商。聚焦新一代应用安全与数据安全建设，提供全面覆盖Web、APP、API的全渠道应用安全、业务安全、数据安全及云安全等领域的产品及服务，持续为用户创造安全的价值，高效应对各种未知威胁和挑战。目前业务覆盖三大运营商、金融、政府、制造、能源、交通、医疗、教育、电商互联网等众多行业千余家头部客户。

内容咨询：

FreeBuf咨询 杨先生　15757171611　yuxiang.yang@tophant.com

合作咨询：

FreeBuf咨询 赵女士　13127723186　jiayu.zhao@tophant.com