Workshop Report

Securing Critical Infrastructure in the Age of AI

Authors

Kyle Crichton*
Jessica Ji*
Kyle Miller*
John Bansemer*
Zachary Arnold
David Batz
Minwoo Choi
Marisa Decillis
Patricia Eke
Daniel M. Gerstein
Alex Leblang
Monty McGee
Greg Rattray
Luke Richards
Alana Scott

***Workshop Organizers**

CSET CENTER *for* SECURITY *and* EMERGING TECHNOLOGY

October 2024

微软公司的慷慨捐助使这次研讨会和最终报告的制作成为可能。本文件中的观点严格属于作者的观点，并不一定代表美国政府的观点。政府、微软公司或作者可能隶属的任何机构、组织或实体

通过商品名称、商标、制造商或其他方式引用任何特定的商业产品、流程或服务，并不构成或暗示美国政府的认可、推荐或支持，包括美国财政部、美国国土安全部、网络安全和基础设施安全局，或作者可能隶属的任何其他机构、组织或实体。

执行摘要

随着人工智能能力的不断提高，关键基础设施（CI）运营商和提供商寻求在其企业中集成新的人工智能系统;然而，这些能力伴随着风险和好处。人工智能的采用可能会带来更强大的系统、业务运营的改进以及更好的工具来检测和应对网络威胁。与此同时，人工智能系统还将引入CI提供商必须应对的新网络威胁 去年的人工智能行政命令指示各部门风险管理机构（SRMA）“评估并提供.....评估与关键基础设施部门使用人工智能相关的潜在风险，包括部署人工智能可能使关键基础设施系统更容易受到严重故障，物理攻击和网络攻击的方式。

尽管行政命令最近的方向，人工智能在关键基础设施中的使用并不新鲜。 擅长预测和异常检测的人工智能工具多年来一直用于网络防御和其他商业活动。例如，提供商长期以来一直依赖由人工智能驱动的商业信息技术解决方案来检测恶意活动。发生变化的是，新的生成式人工智能技术变得更加强大，并为人工智能运营商提供了新的机会。 潜在的用途包括用于客户交互的更强大的聊天机器人，增强的威胁情报合成和优先级排序，更快的代码生成流程，以及最近可以根据用户提示执行操作的AI代理。

CI运营商和行业正试图驾驭这一快速变化和不确定的环境。幸运的是，我们可以借鉴网络安全的类似物。
几年前，网络连接的创新为CI运营商提供了一种远程监控和操作许多系统的方法然而，这也为恶意行为者创造了新的攻击媒介。过去的经验教训可以帮助企业了解如何整合人工智能系统。如今，风险可能以两种方式出现：人工智能漏洞或CI内部署的系统故障，以及恶意使用人工智能系统攻击CI部门。

本研讨会报告提供了在关键基础设施中管理人工智能使用的技术缓解措施和政策建议。这次讨论得出了若干结论和建议。

- 行业内和行业间CI提供商之间的资源差距对AI采用和AI相关风险管理的前景产生了重大影响。**需要更多的项目来支持资源不足的供应商**

与人工智能相关的援助，包括财政资源，培训模型的数据，必要的人才和工作人员，交流论坛，以及更广泛的人工智能话语中的声音。**扩大正式和非正式的互助手段**有助于缩小差距。这些计划在组织间共享它们包括正式的计划，如共享人员以应对事件或紧急情况，以及非正式的努力，如开发最佳实践或审查产品和服务。

- 人们认识到有必要将人工智能风险管理整合到现有的企业风险管理实践中;然而，在当前的企业结构中，人工智能风险的所有权可能是模糊的。一位参与者将这种风险称为AI“烫手山芋”，被抛到了高管层。**需要在公司结构内明确指定人工智能风险的责任。**

- 人工智能安全和人工智能安全之间的模糊性也对人工智能风险管理的运营提出了重大挑战。组织通常不确定如何应用美国国家标准与技术研究所最近发布的人工智能风险管理框架以及网络安全框架的指导。**需要进一步指导如何实施统一的人工智能风险方法**。对这一指南进行调整并确定其优先次序，将有助于使资源不足的提供者和那些有具体需求（往往是定制需求）的人更容易获得这一指南。

- 虽然有完善的网络安全信息共享渠道，但在人工智能方面没有类似的渠道。**SRMA应该利用现有的场所，如信息共享和分析中心，进行人工智能安全信息共享**。分享人工智能安全问题、缓解措施和最佳实践也至关重要，但这样做的渠道尚不清楚。明确什么是人工智能事件，应该报告哪些事件，报告的阈值以及现有的网络事件报告渠道是否足够将是有价值的。为了促进跨部门的可见性和分析，包括人工智能安全性和安全性，各**部门应考虑建立一个人工智能安全性和安全性的集中分析中心。**

- 管理网络和人工智能风险的技能相似但不完全相同。人工智能系统的实施将需要许多CI提供商目前不具备的专业知识。因此，**供应商和运营商应积极提高其现有劳动力**的技能，并寻求机会对员工进行交叉培训，

相关的网络安全技能，以有效地解决人工智能和网络相关的风险范围。

- 生成式人工智能引入了新的问题，这些问题可能更难以管理，需要仔细研究。人工智能提供商**在采用新的人工智能技术之前应保持谨慎和知情，特别是对于敏感或关键任务**。评估一个组织是否准备好采用这些系统是关键的第一步。

目录