# Experts Warn of Security Gaps as AI Expands into UK Critical Infrastructure



The UK’s push to embed artificial intelligence (AI) into critical national infrastructure (CNI)—including power grids, water systems, and transport networks—is raising red flags among cybersecurity and academic experts who warn the technology remains too immature to manage its associated risks.

Francesca Boem of UCL cautions that AI’s integration into infrastructure systems introduces complex cyber-physical vulnerabilities. She warns adversaries could exploit AI decision-making to destabilise energy or water systems, particularly as AI is used in sensing, forecasting, and automation.

AI’s dual dependence on software and hardware increases its exposure to threats like data poisoning, prompt injection, and model manipulation. Experts like Richard Allmendinger (Manchester Business School) stress that even small data tweaks can cause disproportionate impacts, from outages to contamination.

Operational complexity further compounds the risk. Noel Chinokwetu of Orange Cyberdefense highlights that many sectors are still aligning IT and OT systems. He cautions against rushing AI adoption, especially with known issues such as hallucinations in LLMs like ChatGPT.

Despite initiatives such as AI sandboxes by the Office for Nuclear Regulation and the FCA-Nvidia partnership for financial AI testing, experts say coverage remains insufficient. Without rigorous real-world testing, AI failures in CNI could propagate rapidly and dangerously.

A recent study shows nearly 75% of CNI organisations fear AI-enabled threats, including phishing, automated hacking, and adaptive cyberattacks. To mitigate risks, experts call for strict human oversight, clear role delineation between AI and operators, robust security frameworks, and transparent regulatory standards.

While AI offers transformative potential for UK infrastructure, safe deployment will depend on security-first strategies, strong regulation, and responsible design to avoid systemic vulnerabilities.

Created by [Amplify](https://www.hbmadvisory.com/amplify): AI-augmented, human-curated content.

## Bibliography

1. <https://www.newcivilengineer.com/latest/integrating-ai-into-critical-national-infrastructure-presents-severe-risks-experts-warn-01-10-2025/> - Please view link - unable to able to access data
2. <https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper/safety-and-security-risks-of-generative-artificial-intelligence-to-2025-annex-b> - The UK government's discussion paper outlines the safety and security risks associated with generative artificial intelligence (AI) up to 2025. It identifies three primary risk domains: digital risks, political and societal risks, and physical risks. The paper highlights concerns such as cyber-attacks, increased digital vulnerabilities, erosion of trust in information, political and societal influence, insecure use and misuse, and weapon instruction. It emphasizes the need for proactive measures to address these risks and ensure the safe integration of AI technologies into critical infrastructure.
3. <https://professionalsecurity.co.uk/news/interviews/ai-cyber-and-critical-infrastructure/> - An interview in Professional Security Magazine discusses the impact of artificial intelligence (AI) on critical infrastructure. It highlights the significant risks associated with AI integration, including increased complexity, potential over-reliance on AI systems, and exposure to sophisticated cyber-attacks. The article emphasizes the importance of ensuring high reliability and safety, seamless integration with existing legacy systems, and robust cybersecurity measures. Ethical considerations such as transparency, accountability, and bias are also addressed, underscoring the need for careful implementation of AI in critical infrastructure sectors.
4. <https://www.hitpaw.com/ai-act-tips/impact-of-ai-on-critical-infrastructure-management.html> - This article examines the impact of artificial intelligence (AI) on critical infrastructure management, focusing on the risks under the EU AI Act. It discusses compliance and regulatory risks, data privacy and security concerns, algorithmic bias and transparency issues, and systemic reliability and safety risks. The piece emphasizes the need for high-quality, representative datasets, detailed technical documentation, and third-party conformity assessments. It also highlights the importance of addressing data privacy, ensuring transparency, and maintaining human oversight to mitigate potential unintended consequences and enhance the resilience of critical infrastructure systems.
5. <https://www.bridewell.com/insights/news/detail/three-quarters-of-uk-critical-national-infrastructure-organisations-concerned-about-ai-driven-cyber-threats> - A study by Bridewell reveals that 76% of UK critical national infrastructure (CNI) organisations are concerned about AI-driven cyber threats. The research highlights worries about AI-powered phishing attacks, adaptive AI cyber-attacks, AI-driven exploit development, and automated hacking using AI. The findings underscore the growing apprehension among CNI organisations regarding the sophistication and scale of cyber threats facilitated by AI technologies, emphasizing the need for enhanced cybersecurity measures and proactive strategies to address these emerging risks.
6. <https://www.sourcesecurity.com/news/quarters-uk-critical-national-infrastructure-organisations-co-1616132660-ga.1715682798.html> - An article on SourceSecurity.com reports that 76% of UK critical national infrastructure (CNI) organisations have identified AI-driven cyber threats as a current security concern. The study highlights specific worries about AI-powered phishing attacks, adaptive AI cyber-attacks, AI-driven exploit development, and automated hacking using AI. These findings reflect the increasing apprehension among CNI organisations about the evolving nature of cyber threats enabled by AI, emphasizing the necessity for robust cybersecurity strategies and vigilance in the face of these challenges.
7. <https://www.reuters.com/business/finance/uk-financial-regulator-partners-with-nvidia-ai-sandbox-2025-06-09/> - Reuters reports on the UK's Financial Conduct Authority (FCA) partnering with Nvidia to launch a 'Supercharged Sandbox' aimed at enabling financial services firms to test AI technologies in a controlled environment. Set to begin testing in October 2025, the initiative is part of a broader government strategy to foster innovation and boost economic growth. The sandbox provides participants with access to Nvidia's advanced computing platform and AI software, along with technical expertise, better datasets, and regulatory guidance, especially for firms at the early stages of AI exploration.