Summary: Generative artificial intelligence and cyber security in central banking

Overview

This summary provides an overview of the document titled 'Generative artificial intelligence and cyber security in central banking' published by the Bank for International Settlements (BIS) in May 2024. The paper, identified as BIS Papers No 145, delves into the implications of gen AI for central banks' cyber security, examining the technology's potential to both strengthen cyber defences and introduce new vulnerabilities. It draws on survey data from cyber security experts at major central banks to assess the current adoption of gen AI tools, their benefits and risks, and the need for enhanced human capital. The document also contemplates the future role of gen AI in cyber security and the necessity for regulatory frameworks and international cooperation.

Section 1: Gen AI and cyber risk

Generative artificial intelligence (gen AI) represents an evolution in machine learning, with neural networks and transformers being key techniques. Neural networks, which mimic the human brain's neurons, are structured in layers and use parameters to improve during training. They are foundational to technologies like face recognition and voice assistants. Transformers, introduced in 2017, have significantly advanced natural language processing by understanding the context within text sequences, leading to large language models (LLMs) such as ChatGPT. These models, trained on extensive internet text data, can generate content that closely resembles human language.

Gen AI's potential as a general-purpose technology is significant, with the ability to generate new content across various formats and perform tasks beyond language recognition. Central banks have already implemented traditional AI tools for data analysis and cyber security, and gen AI offers further opportunities and challenges, particularly in cyber security management. Gen AI can be used offensively by cyber threat actors to conduct sophisticated social engineering, create malware, and run disinformation campaigns. Conversely, it can also be used defensively to enhance cyber security measures.

However, the adoption of gen AI introduces new risks, such as data/model poisoning, data leakage during inference, and vulnerability discovery. These risks necessitate careful consideration of data security and privacy, including the handling and quality of training data. Despite these challenges, gen AI can strengthen cyber security through proactive fraud prevention strategies, which depend on the AI programming ability of IT staff and data availability.

Section 2: Gen AI in central banking

The survey among members of the GCRG in January 2024 reveals that 71% of central banks are currently using generative AI (gen AI), with an additional 26% planning to do so within the next two years, indicating a potential near-term adoption rate of nearly 100%. Despite this, only 19% of central banks have a concrete strategy for gen AI integration, with 55% reporting their strategy is still in development, reflecting a cautious approach due to uncertainties about the technology's proper use. Central banks generally see more benefits than risks in using gen AI, with 19% fully agreeing and 56% partially agreeing that the benefits outweigh the risks.

Central banks identify cyber threat detection as the primary benefit of gen AI, with 57% of respondents selecting it. Other areas where gen AI is expected to be beneficial include code creation and debugging,

fraud detection, cyber threat response, summarizing documents and meeting notes, and drafting emails and documents. These findings underscore the importance of cyber security for central banks and the potential role of gen AI in enhancing their defenses.

Section 3: Opportunities, risks and challenges for cyber security

Central banks face the challenge of developing IT infrastructures that harness the advantages of generative artificial intelligence (gen AI) while mitigating associated cyber risks. The section under review delves into the opportunities presented by gen AI, such as improved cyber threat detection and response capabilities. It also examines the risks and challenges, including the potential for social engineering attacks and unauthorized data disclosure. Surveyed leaders from central bank IT cyber security units contribute their perspectives on these issues. The analysis underscores the need for significant investment in human capital, particularly in professionals skilled in both cyber security and AI programming, to effectively manage the integration of gen AI into cyber security practices.

Section 3.1: Opportunities

Central banks have been using traditional machine learning tools for cyber risk management, which include systems for threat detection, transaction security, and payment system integrity. The introduction of generative artificial intelligence (gen AI) could potentially enhance these cyber security capabilities. A survey revealed that 44% of respondents view gen AI as very effective or effective in cyber security, while 41% see it as moderately effective. Only a small fraction consider it not very effective, and some did not provide an evaluation. When asked to rate the impact of gen AI on cyber security enhancement on a scale from 1 to 10, a quarter of respondents gave a score of 5 to 6, over 45% scored it between 7 and 8, and 11% rated it 9 to 10, indicating a generally positive view of gen AI's potential to improve cyber security, although there is still some uncertainty about its overall impact.

The survey also assessed the potential benefits of AI in cyber security, with "Automation of routine tasks" receiving the highest average score, suggesting that gen AI's ability to automate labor-intensive tasks is a key advantage. Other highly rated benefits include improved response times, deep learning insights for analyzing complex data patterns, and enhanced threat detection. These findings suggest that gen AI is expected to positively impact the detection and response to cyber threats, and could lead to an expansion of cyber security units at central banks through automation and the development of new soft skills.

Section 3.2: Risks and challenges

Central banks are increasingly concerned about the sophisticated cyber threats posed by generative artificial intelligence (gen AI), with social engineering and unauthorized data disclosure being the top vulnerabilities. These AI-enabled attacks could lead to unauthorized access to internal networks and sensitive data exposure, potentially undermining trust and financial stability. A comprehensive policy is needed to educate all central bank employees on the risks associated with gen AI tools.

The survey revealed that central banks also face significant challenges when integrating gen AI into cyber security systems, with the skill gap being the most pressing issue. Other challenges include ensuring the security of AI systems, understanding AI decisions, addressing ethical and privacy concerns, and obtaining high-quality data for training AI models. Nearly all respondents rated the risk associated with implementing gen AI for cyber security as moderate to high.

Graphs in the section illustrate the average scores assigned by respondents to various vulnerabilities and challenges, with social engineering, unauthorized data disclosure, and skill gaps receiving the highest

concern. The next section will delve deeper into the skill gap challenge, highlighting the need for expertise in both AI programming and cyber security.

Section 4: IT investments and human capital

Central banks have significantly increased their investment in cybersecurity and are now focusing on integrating generative artificial intelligence (gen AI) into their operations. This integration necessitates an urgent update in skills, which can be addressed through training existing staff and hiring new employees with the necessary expertise. A particular security concern is "model poisoning," where attackers manipulate AI systems to compromise their integrity.

The integration of gen AI in central banks involves two main dimensions: the adoption of gen AI tools by all employees and the specific application within IT divisions for cybersecurity. A survey revealed that most central banks allow staff access to cloud-based gen AI applications, such as ChatGPT, with certain restrictions, while a minority have no plans to enable such access. Concerns about staff preparedness to effectively use AI systems are prevalent, with 40% of respondents expressing high or extreme concern.

In terms of IT investments, gen AI is seen as increasing efficiency and reducing workload in cybersecurity units, although some banks reported no significant change in human resource allocation. The need for skillset upgrades and strategic task reallocation was also noted. The scarcity of AI-qualified personnel is a major concern, with limited AI expertise and dependency on external vendors being the top-rated issues.

Graphs in the section illustrate the benefits of AI for human resources, issues for AI-qualified personnel, and the nature of AI-human interaction. AI is primarily viewed as a decision support tool requiring human oversight, with a complementary relationship between human and AI capabilities in cybersecurity. The survey indicates that central banks recognize the need for continuous learning and adaptation in the evolving collaboration between AI systems and human experts.

Section 5: Future perspectives and regulatory insights

Central banks have been focusing on enhancing cyber resilience, shifting from compliance-based to risk management approaches. The integration of generative artificial intelligence (gen AI) is expected to continue this trend, with a move towards proactive threat management and customized network defenses. However, the core strategies of cyber security are not anticipated to change drastically with the adoption of gen AI. Ethical and regulatory concerns rated highest by central banks include autonomous decision-making and data protection, with accountability for AI actions also being significant. Roles that will gain importance as gen AI is more widely used include data scientists, AI supervisors, security analysts, and developers, emphasizing the continued need for human expertise in managing AI systems. Graphs from the survey illustrate these concerns and the anticipated importance of various roles, with autonomous decision-making and data scientists scoring highest in their respective categories.

Section 6: Conclusion

Cyber attacks are on the rise, becoming more complex and sophisticated, paralleling significant technological advancements, particularly in generative artificial intelligence (gen AI). A survey conducted with central bank cyber security leaders through the CRCC-administered GCRG forum in January 2024 assessed the adoption of gen AI tools in cyber security. The survey revealed a consensus on the need for common AI usage rules and new forms of cooperation to establish data protection standards and address the skills gap in human personnel.

The Bank for International Settlements (BIS) facilitates central banks' cyber security efforts and global cooperation via the CRCC, which has been crucial since its establishment in 2019 for integrating gen AI into cyber security. The GCRG forum, comprising CISOs from BIS member banks, is key in tackling AI technology challenges. The CRCC also offers a global cyber resilience collaboration platform with over 300 professionals, which serves as a knowledge-sharing and collaboration hub on AI challenges and adoption.

Additionally, the CRCC leads the Cyber Resilience Assessments project, providing a framework for central banks to assess and enhance their cyber resilience, and has delivered a global benchmark for comparison and informed decision-making on cyber security investments. The CRCC's community events, including annual seminars and cyber range exercises, maintain engagement with emerging cyber security issues and threats, ensuring operational readiness.

The document underscores the growing importance of cooperation and information-sharing in mitigating cyber risks and managing incidents, especially with the development of gen AI systems. International work on cyber security continues under the guidance of standard-setting bodies like the Financial Stability Board and G7.