SUMMARY

Generative artificial intelligence and cyber security in central banking

Bank for International Settlements (BIS) (May 2024)

Disclaimer: This summary was created using generative AI. Mistakes are possible.

Overview

This summary provides an overview of the BIS Paper No 145 titled 'Generative artificial intelligence and cyber security in central banking', authored by Iñaki Aldasoro, Sebastian Doerr, Leonardo Gambacorta, Sukhvir Notra, Tommaso Oliviero, and David Whyte. The paper explores the dual role of generative artificial intelligence (gen AI) in central banking, highlighting its potential to enhance cyber security while also presenting new risks. The authors utilize data from a unique survey conducted among cyber security experts at major central banks to examine the adoption of gen AI tools, their perceived benefits and risks, and the necessary investments in human capital. The paper reveals that most central banks have adopted or plan to adopt gen AI for cyber security due to its benefits in threat detection and response time reduction. However, gen AI also increases the risks of social engineering attacks and unauthorized data disclosure. To mitigate these risks, central banks anticipate substantial investments in human capital, particularly in staff with expertise in both cyber security and AI programming. The paper also discusses the future landscape of AI in cyber security, including the roles that will become increasingly critical for human workers and the ethical and regulatory concerns that arise with the integration of gen AI.

Section 1: Introduction

The financial sector, and central banks in particular, face increasing cyber threats, with generative artificial intelligence (gen AI) emerging as both a potential asset and a new vector for attacks. The US Department of the Treasury's report in March 2024 underscores the critical role of gen AI in financial sector cybersecurity. Central banks are leveraging gen AI for enhanced data processing and proactive security measures, but they also confront challenges such as AI-generated social engineering and zero-day attacks. An ad hoc survey among Global Cyber Resilience Group members reveals that most central banks are adopting or planning to adopt gen AI tools, recognizing their benefits in cyber threat detection and operational efficiency. However, this adoption is accompanied by significant challenges, particularly in

human capital investment, as central banks grapple with the need for staff skilled in both AI and cybersecurity. The survey also indicates a shift towards more strategic cybersecurity initiatives and the necessity for human oversight of AI systems. The paper contributes to the discourse on cyber risk regulation and the need for common guidelines and practices among central banks to address the benefits and challenges of gen AI. It also highlights the importance of preparing IT staff for rapid technological advancements and the potential for gen AI to increase productivity in cybersecurity roles without replacing human experts. The paper concludes by discussing the future landscape of AI in cybersecurity, emphasizing regulatory insights and the critical roles of data scientists, AI security analysts, and AI supervisors in integrating gen AI with existing security tools.

Section 2: Gen AI and cyber risk

Generative artificial intelligence (gen AI) represents a significant evolution in machine learning, with neural networks and transformers being key techniques that have led to advancements such as large language models (LLMs) exemplified by ChatGPT. These models have the ability to generate new content across various formats and have shown human-like language processing capabilities. Gen AI's potential as a general-purpose technology is recognized for its transformative impact on industries, including central banking, where it has been used for data analysis, payment systems oversight, and cyber security.

However, gen AI also introduces new cyber risks by enhancing the capabilities of cyber threat actors. It can be used for sophisticated social engineering attacks, creating advanced malware, and conducting disinformation campaigns. Additionally, gen AI's reliance on data raises concerns about data security and privacy, including the handling and quality of training data.

Central banks face the challenge of defending against gen AI-powered cyber threats. Risks include data or model poisoning, data leakage during inference, and the use of AI tools by threat actors to discover vulnerabilities. To counter these risks, central banks must develop proactive fraud prevention strategies and strengthen their cyber security defenses using AI.

The paper underscores the importance of having IT staff with adequate AI programming skills and access to quality data for training and testing AI systems. This is crucial for leveraging AI's benefits in cyber security and mitigating the heightened risks associated with gen AI.

Section 3: Gen AI in central banking

A survey among members of the GCRG reveals a significant uptake of generative artificial intelligence (gen AI) in central banks, with 71% already using it and another 26% planning to do so within two years, indicating a potential near-term adoption rate close to 100%. Despite this trend, only 19% of central banks have a concrete strategy for gen AI integration, while a majority are still developing their approach. This cautious stance is due to uncertainties about the optimal use of gen AI, balancing perceived opportunities against risks and challenges. Central bank experts generally acknowledge the net advantages of gen AI, with 75% agreeing to varying extents that AI offers more benefits than risks.

The survey identifies cyber threat detection as the primary benefit of gen AI, with 57% of respondents highlighting it. Other significant areas where gen AI is expected to contribute include code creation and debugging, fraud detection, and cyber threat response. Additionally, gen AI is anticipated to aid in routine tasks such as summarizing documents and

drafting communications. These insights underscore the central banks' focus on enhancing cyber security and streamlining daily operations through gen AI, while also navigating the complexities of its integration into their systems.

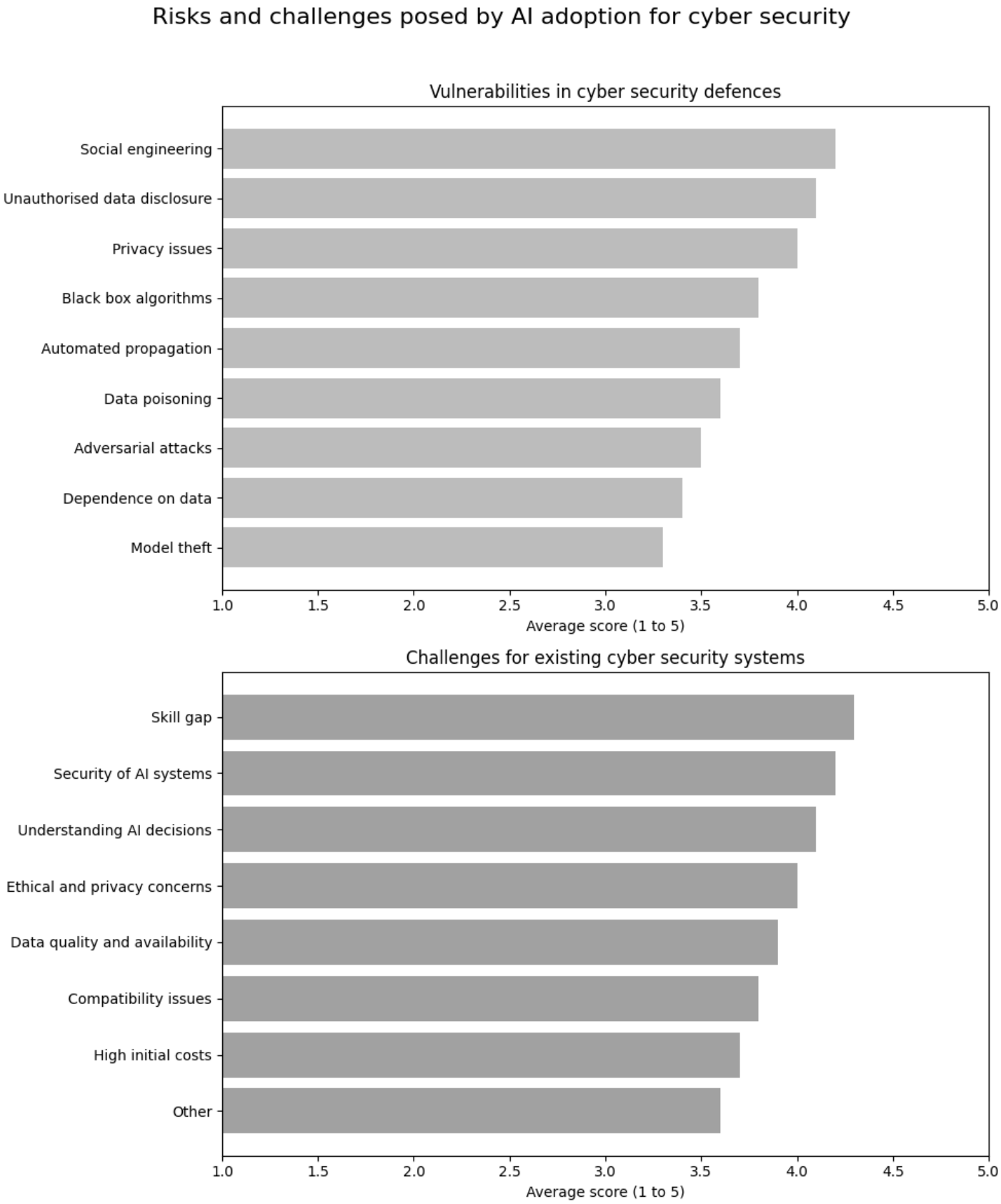
Section 4: Opportunities, risks and challenges for cyber security

Central banks are exploring the use of generative artificial intelligence (gen AI) to bolster cyber security, with a survey of IT cyber security leaders revealing a generally positive perception of gen AI's effectiveness. Approximately 44% of respondents consider gen AI to be very effective or effective in identifying and responding to cyber threats, while 41% view it as moderately effective. The potential benefits of gen AI in cyber security are recognized, particularly in automating routine tasks, improving response times, and enhancing threat detection through deep learning insights.

However, the adoption of gen AI also introduces new risks and challenges. Respondents express significant concern over social engineering and unauthorized data disclosure, which are exacerbated by gen AI's capabilities in creating sophisticated attacks like deepfakes. Privacy issues, the opacity of "black box" algorithms, and automated propagation of threats are also major concerns, reflecting a fear of losing control over cyber security systems.

Central banks face challenges in integrating gen AI into existing cyber security frameworks, with the skill gap being the most prominent. This gap refers to the shortage of professionals skilled in both AI programming and cyber security. 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.

In response to these challenges, central banks are expected to make significant investments in human capital, particularly in staff with dual expertise in cyber security and AI. The integration of gen AI into cyber security is anticipated to transform the scope and scale of cyber security units, necessitating a strategic approach to manage the ethical and regulatory implications of this technology.



Section 5: IT investments and human capital

Central banks have significantly increased their investment in cybersecurity and are now focusing on enhancing their human capital to effectively integrate generative artificial intelligence (gen AI) into their operations. Since 2020, there has been a notable rise in

cybersecurity budgets, driven by the emergence of gen AI and the need to update skills through training and recruitment. A particular security concern is "model poisoning," where attackers manipulate AI systems to compromise their integrity, necessitating heightened vigilance and expertise.

The integration of gen AI tools, such as cloud-based applications, is being cautiously adopted by central banks, with most allowing restricted access and a minority planning or not planning to enable it. There is a high level of concern among cyber security experts about the current staff's ability to onboard and operationalize AI systems, suggesting a need for regulatory policies and internal practices for safe adoption.

The benefits of gen AI in cybersecurity units are primarily seen in increased efficiency and reduced workload, with some central banks reporting no significant change in human resource allocation. However, there is a lower emphasis on skillset upgradation and strategic task reallocation. The scarcity of AI-qualified personnel is a major concern, with limited AI expertise and dependency on external vendors being the top-rated issues. This underscores the importance of talent retention and recruitment.

Central banks view the collaboration between AI systems and human experts as complementary, with AI serving as a support tool requiring human oversight and continuous learning. This suggests a future where AI enhances productivity but does not replace the need for human expertise in cybersecurity.

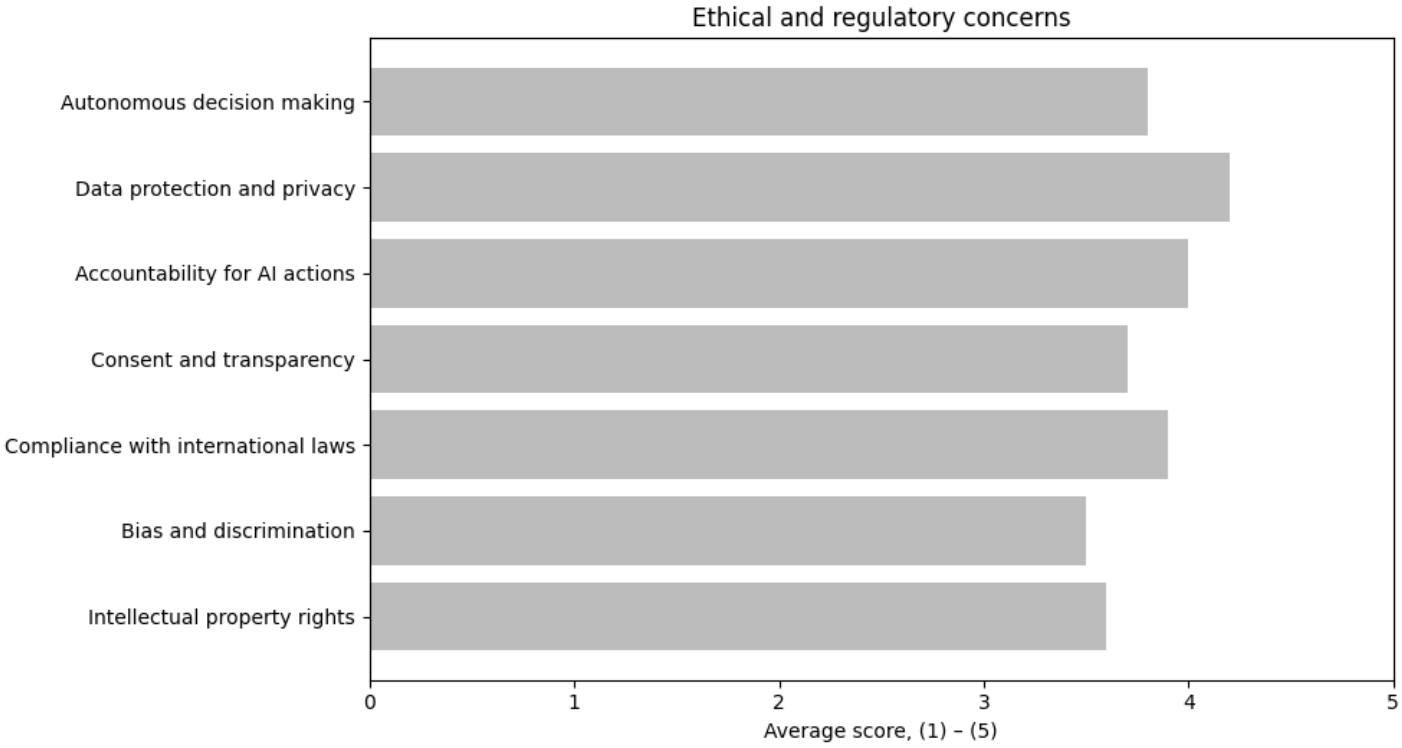
Section 6: Future perspectives and regulatory insights

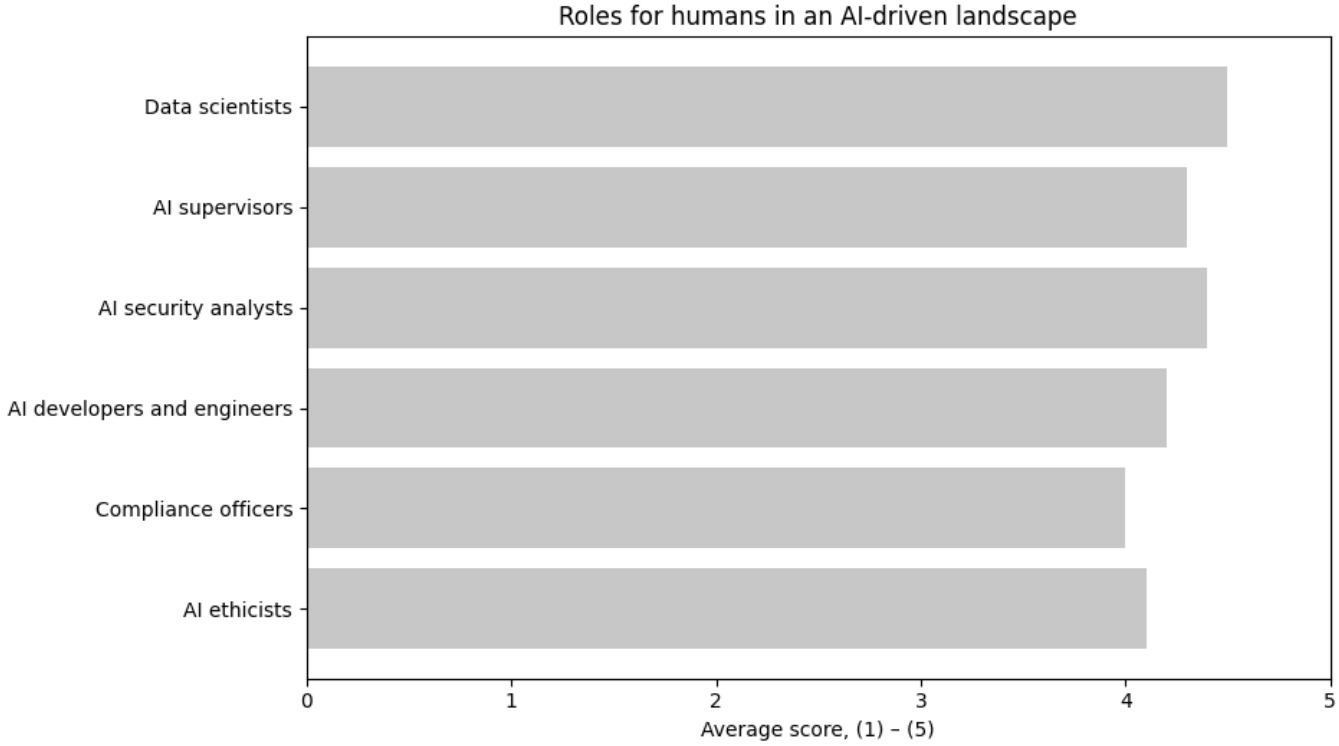
Central banks have been investing in cyber security, shifting from compliance to risk management and resilience, and this trend is expected to continue with the integration of generative artificial intelligence (gen AI). Survey results from central banks suggest that gen AI will enhance proactivity in cyber risk management, allowing for more customized defenses and dynamic risk assessments. However, the core strategies of central banks regarding cyber security are not anticipated to undergo major changes with the adoption of gen AI, as the focus remains on improving cyber resilience.

Ethical and regulatory concerns associated with gen AI in cyber security are primarily focused on autonomous decision-making, data protection and privacy, and accountability for AI actions. These concerns underscore the need for clear guidelines and standards to manage AI's decision-making capabilities and protect sensitive data. Compliance with international laws and issues of consent and transparency are seen as less urgent, likely due to ongoing regulatory adaptations.

The survey also highlights the increasing importance of human roles in an AI-driven cyber security landscape. Data scientists, AI supervisors, AI security analysts, and AI developers and engineers are identified as key professionals who will support the correct adoption and use of gen AI tools. Their expertise will be crucial in interpreting data, refining AI learning processes, ensuring ethical standards, and maintaining compliance with legal and regulatory frameworks. These insights emphasize the continued need for human expertise alongside advanced AI technologies in central banking cyber security.

The future landscape of AI and cyber security





Source: Authors' calculations.

Section 7: Conclusion

The increasing frequency and sophistication of cyber attacks, coupled with rapid advancements in generative artificial intelligence (gen AI), have prompted central banks to evaluate the adoption of gen AI tools for cyber security. A survey conducted among central

bank cyber security leaders through the CRCC-administered GCRG forum has shed light on the current adoption status, benefits, risks, and challenges of gen AI in this context. There is a consensus on the need for common AI usage rules and enhanced cooperation among central banks to address data protection and the skill gap in human personnel.

The Bank for International Settlements (BIS) plays a supportive role in these efforts through the CRCC, which has been pivotal since its establishment in 2019. The CRCC's initiatives include the GCRG forum, a global cyber resilience collaboration platform, and the Cyber Resilience Assessments project, which provides a framework for central banks to assess and improve their cyber resilience. These tools and events, such as annual seminars and cyber range exercises, help central banks stay engaged with emerging cyber security issues and maintain operational readiness.

The CRCC's work, including community events and collaborative platforms, is crucial for knowledge-sharing and training in the central bank community. This collective approach is essential for reducing cyber risk and effectively responding to incidents, especially as the role of gen AI systems in cyber security continues to grow. International cooperation on cyber security, guided by standard-setting bodies, is ongoing and increasingly important with the integration of gen AI technologies.

References

1. Aldasoro, I, S Doerr, L Gambacorta and D Rees (2024a): "The impact of artificial intelligence on output and inflation", BIS Working Papers, no 1179, April.
2. Aldasoro, I, J Frost, L Gambacorta, T Leach and D Whyte (2020): "Cyber risk in the financial sector", SUERF Policy Notes, no 206, November.
3. Aldasoro, I, L Gambacorta, P Giudici and T Leach (2022): "The drivers of cyber risk", Journal of Financial Stability, vol 60, June.
4. Aldasoro, I, L Gambacorta, A Korinek, V Shreeti and M Stein (2024b): "Intelligent financial system: how AI is transforming finance", BIS Working Papers, forthcoming.
5. Araujo, D, G Bruno, J Marcucci, R Schmidt and B Tissot (2022): "Machine learning in central banking", IFC Bulletins, no 57, November.
6. Araujo, D, S Doerr, L Gambacorta and B Tissot (2024): "Artificial intelligence in central banking", BIS Bulletins, no 84, January.
7. Basel Committee on Banking Supervision (2018): Cyber-resilience: range of practices, December.
8. --- (2021): "Newsletter on cyber security", September.
9. Boissay, F, G Cornelli, S Doerr and J Frost (2022): "Blockchain scalability and the fragmentation of crypto", BIS Bulletins, no 56, June.
10. Brynjolfsson, E, D Li and L Raymond (2023): "Generative AI at work", NBER Working Papers, no 31161.
11. Doerr, S, L Gambacorta and J Serena Garralda (2021): "Big data and machine learning in central banking", BIS Working Papers, no 930, March.
12. Doerr, S, L Gambacorta, T Leach, B Legros and D Whyte (2022): "Cyber risk in central banking", BIS Working Papers, no 1039, September.
13. Enterprise Strategy Group and Information Systems Security Association (2020): The life and times of cybersecurity professionals 2020, July.
14. Falade, P (2023) "Decoding the threat landscape: ChatGPT, FraudGPT, and WormGPT in social engineering attacks", arXiv:2310.05595.
15. Felten, E, M Raj and R Seamans (2021): "Occupational, industry, and geographic exposure to artificial intelligence: a novel dataset and its potential uses", Strategic Management Journal, vol 42, no 12, pp 2195-217.

16. Financial Stability Board (2020): Effective practices for cyber incident response and recovery: final report, October.
17. G7 (2016): G7 fundamental elements of cybersecurity for the financial sector, October.
18. Hitaj, D, G Pagnotta, F De Gaspari, D Ruko, B Hitaj, L Mancini and F Perez-Cruz (2024): "Do you trust your model? Emerging malware threats in the deep learning ecosystem", arXiv:2403.03593v1.
19. Hitaj, D, G Pagnotta, B Hitaj, L Mancini and F Perez-Cruz (2022): "MaleficNet: hiding malware into deep neural networks using spread-spectrum channel coding," in V Atluri, R Di Pietro, C Jensen and W Meng (eds), Computer Security - ESORICS 2022, Lecture Notes in Computer Science, vol 13556, Springer, Cham.
20. Improta, C (2024): "Poisoning programs by un-repairing code: security concerns of AI-generated code", arXiv:2403.06675v1.
21. Kashyap, A and A Wetherilt (2019): "Some principles for regulating cyber risk", AEA Papers and Proceedings, vol 109, May, pp 482-7.
22. McKinsey (2023): The economic potential of generative AI: the next productivity frontier, McKinsey Digital report, June.
23. Neupane, S, I Fernandez, S Mittal and S Rahimi (2023): "Impacts and risk of generative AI technology on cyber defense", 10.48550/arXiv.2306.13033.
24. Noy, S and W Zhang (2023): "Experimental evidence on the productivity effects of generative artificial intelligence", Science,