BALANCING INNOVATION AND RISK: OSFI'S PRINCIPLES FOR RESPONSIBLE AI IN FINANCE

Categories: Insights, Publications

The Office of the Superintendent of Financial Institutions ("**OSFI**"), in collaboration with the Global Risk Institute ("**GRI**"), recently convened the second Financial Industry Forum on Artificial Intelligence ("**FIFAI II**"). This forum brought together leaders from the financial sector, government, and academia to further explore the risks, opportunities, and oversight considerations posed by artificial intelligence ("**AI**") in Canada's financial system. The event builds on previous discussions and research, including OSFI's AI and Quantum Questionnaire conducted in December 2023 and the joint AI Risk Report with the Financial Consumer Agency of Canada ("**FCAC**") released in September 2024.

This latest development follows earlier regulatory and policy milestones we have covered in detail, including OSFI and FCAC's initial recommendations for sound AI risk management and the first joint OSFI–GRI risk report, both of which signaled the groundwork for more formalized regulatory guidance. (See our earlier bulletins: "*AI Use by Financial Institutions: OSFI and FCAC Recommendations for Sound Risk Management*" and "*AI in Financial Services: Joint OSFI and GRI Report Highlights Need for Safeguards and Risk Management as a Prelude to Enhanced OSFI Guidance*".)

FIFAI II comprises four thematic workshops—security and cybersecurity; financial crime; consumer protection; and financial stability—each culminating in interim reports, with a consolidated final report to follow.

The first workshop, held on May 28, 2025, brought together 56 Canadian and international AI experts, including representatives from banks, insurers, asset managers, fintechs, academia, policymakers, and regulators. It was co sponsored by OSFI, the Department of Finance Canada, and GRI.

AI adoption continues to grow rapidly among federally regulated financial institutions ("**FRFIs**"). In 2019, approximately 30 percent of FRFIs reported using AI in their operations. That number had risen to about 50 percent by 2023, and projections indicate that over 70 percent will adopt AI technologies by 2026. Institutions are increasingly integrating AI into areas such as fraud detection, customer service, document automation, underwriting, trading, and claims management.

At the core of the forum's findings is a proposed framework for responsible AI use, anchored in four guiding principles referred to as "EDGE": Explainability, Data, Governance, and Ethics. Explainability emphasizes the importance of making AI decisions understandable to both internal and external stakeholders. Reliable and well-governed data is essential to building trustworthy models. Strong governance frameworks are necessary to oversee AI implementation across the

enterprise. Finally, ethical considerations must be central, including transparency, privacy, consent, and the mitigation of algorithmic bias.

The forum identified several internal risks associated with AI. These include poor data governance, opaque or overly complex models, legal and reputational vulnerabilities, excessive reliance on third-party AI vendors, cybersecurity threats, and exposure to market and credit risks through automated decision-making. External risks include the rising sophistication of cyber threats enabled by generative AI, such as deepfakes and phishing, as well as the competitive pressure that could incentivize rapid AI adoption without adequate safeguards.

Speakers underscored the urgent nature of AI-enabled cyber threats, noting that such incidents may cause losses equivalent to 1–10 percent of global GDP, and that deepfake attacks have increased twenty-fold in the past three years. Security was defined more broadly than cyber-risk alone—encompassing protection of physical infrastructure, personnel, technology, and data, with implications for national security. A participant survey identified the top internal hurdles to managing AI security risk: 60 percent flagged the rapid pace of AI innovation, 56 percent raised concerns over third-party vendor vetting, and 49 percent highlighted governance uncertainty.

To address these challenges, participants stressed the importance of robust, adaptable governance structures. Existing and proposed risk management frameworks, such as OSFI's draft Guideline E-23, should be extended to include AI-specific considerations. Institutions should adopt lifecycle-based approaches to managing AI risk, including thorough model validation, human oversight, and transparent communication with consumers about how their data is used and how decisions are made. Monitoring of third-party vendors and investing in employee training on AI ethics and data literacy were also highlighted as critical steps.

Looking ahead, OSFI is advancing toward a more formal regulatory approach to AI. The revised draft Guideline E-23 now explicitly addresses AI and machine learning risks with publication expected by September 11, 2025. The regulator continues to work closely with other agencies, including FCAC and Innovation, Science and Economic Development Canada, to align financial-sector oversight with broader federal AI legislation, such as the *Artificial Intelligence and Data Act* (AIDA).

In conclusion, the forum underscores that as AI becomes more deeply embedded in the financial sector, institutions must approach its deployment with a high degree of caution and accountability. The EDGE principles offer a structured way to manage the risks while unlocking the transformative potential of AI. OSFI encourages all regulated entities to act proactively, ensuring their AI strategies are both innovative and responsibly governed in anticipation of Canada's evolving regulatory landscape. Financial institutions should consider this forum as not only a continuation but also an acceleration of supervisory focus on AI.

We will continue to monitor and report on these developments as the regulatory environment matures ahead of key milestones in Guideline E-23.

By [Darcy Ammerman](), [Jerry Z. Huang](), and [Mathurhaen Siri]()

A Cautionary Note

The foregoing provides only an overview and does not constitute legal advice. Readers are cautioned against making any decisions based on this material alone. Rather, specific legal advice should be obtained.