First in Two-Part Series

Managing AI Fraud Risks:

Fiduciary Duties for Pension and Benefits Leaders—Part 1

by | **Jerrard Gaertner, Paul Catenacci** and **Donald Broggi**

AI is already emerging in the fraud landscape, affecting pension and benefits plans, with its role expected to expand in the near future. In this first article, the authors examine how these evolving fraud risks intersect with Canadian fiduciary obligations and risk-management expectations.

plans & trusts
education | research | information

For most pension and benefits plans, fraud used to mean familiar things: a forged cheque, a stolen password, a fake invoice that somehow slipped past the internal controls. Today, it looks increasingly different. A member receives a call from the "plan office" that sounds exactly like your staff. A "vendor" email lands in an inbox with perfect grammar, the right logos and just enough urgency to rush a payment. Your "IT department" asks you to log in to a slightly different web address due to technical issues. In these and many more cases, the human employee is dealing with something quite new: fraudsters quietly using artificial intelligence (AI) to impersonate the people and organizations your plan relies on.

Regulators are starting to connect the dots. In Canada, OSFI's risk outlooks have elevated technology and operational risk for federally regulated pension plans. CAPSA's Guideline No. 4 now weaves cyber, data and third-party risk into what it means to run a prudent plan. South of the border, securities and consumer regulators are warning about AI-enabled scams and products that are "too good to be true" and can reach plan members directly. None of these scenarios are abstract. They show up as real incidents, difficult and time-consuming investigations, regulator inquiries and uncomfortable questions from stakeholders.

When questions do arise, they are not aimed at your IT vendor. They land on trustees, plan administrators, union leaders and employer/employee representatives. An AI-generated phishing campaign that captures member credentials, or a convincing deepfake that bypasses security controls, becomes more than a one-off incident. It raises core issues about governance; vendor oversight; incident readiness; and, ultimately, whether the plan can show that it took reasonable, forward-looking steps to protect members' money and information.

We wrote this article with fiduciaries, managers and their staff in mind. We'll describe how AI is already showing up in the fraud landscape around pension and benefits plans today, as well as how it is likely to evolve in the near term. We'll look at how these new fraud risks interact with Canadian fiduciary duties and risk-management expectations, province by province. Our U.S. co-authors will add a comparative view of ERISA and other U.S. regimes, along with case examples illustrating how courts and regulators are responding to these new exploits. Finally, we'll provide a playbook and a short checklist (not legal advice) for boards and committees to stress-test their own arrangements, along with a few simple rules of thumb you can bring to your next meeting.

Takeaways

- AI-driven fraud is already affecting pension and benefits systems, and regulators now view technology and third-party risk as core governance issues, requiring trustees, administrators, unions and employers to demonstrate structured, prudent responses.
- Many, but not all, AI-related fraud risks today are enhanced versions of familiar schemes: impersonation, phishing, account takeover and payment diversion. However, some are brand new. AI makes these attacks more scalable and convincing, reducing the obvious "red flags" staff used to rely on.
- Pension and benefits plans are vulnerable to AI-enabled fraud directly, indirectly through members and vendors, and through the AI tools embedded in their own operations.
- Across Canada, regulators are steadily moving technology and cyber risk, including AI-enabled fraud, into the heart of pension plan governance and risk management.
- Federally regulated plans face OSFI expectations on technology and cyber incidents, alongside broader prudential guidance that increasingly touches on AI-related risks. Ontario's FSRA and CAPSA guidelines together establish clear, Canada-wide expectations for managing IT, cyber and third-party risks in pension plans.

AI Fraud in and Around Pension and Benefits Plans: What It Actually Looks Like

In most plans, AI does not first arrive as a shiny new system with a line item in the IT budget. It shows up on the other side of the table: in the tools fraudsters are using. That doesn't mean basic fraud patterns are necessarily changing. It does mean that AI fraudsters can teach an old dog some new tricks.

1. Impersonation Is Getting an Upgrade

The simplest fraud pattern has always been someone pretending to be someone you trust.

Regulators and investor-protection agencies have been warning for some time about fraudsters impersonating registered firms, advisors and even regulators themselves, using very professional-looking websites, social media accounts and documents. AI makes doing this almost trivial. Logos, boilerplate text and even "About Us" pages can be cloned and customized in minutes. Email and text messages can be generated at scale, precisely tuned to sound like a specific firm

or individual. Voice-cloning tools can reproduce a familiar voice from a few seconds of audio, making a fake call very hard to detect. Some firms have already switched to video calls only for critical interactions.

Unfortunately, video fakes are rapidly becoming almost as easy to generate as audio fakes. Fake calls from members, vendors, custodians or even regulators are especially hard to spot when they hit busy staff who are using phone, email and messaging apps all day. The infrastructure around your plan (recordkeepers, administrators, custodians, insurers, consultants) is also a tempting target. If a fraudster can convincingly pose as any trusted partner, they may be able to trick staff into disclosing information or even approving a transaction.

2. Account Takeover and Payment Fraud Are More Convincing

The most powerful tools available today for phishing and account takeover have changed.

Fraudsters now use large language models (including tools like Claude and ChatGPT) to create emails and messages that are free of the spelling and grammar mistakes that used to be easy red flags. Recent surveys of corporate finance and treasury teams, including the 2025 *AFP Payments Fraud and Control Survey*, report significant year-over-year increases in cyber-fraud incidents, with deepfakes, voice cloning and increasingly sophisticated email schemes singled out as key drivers.

From a plan's perspective, the scenarios are familiar, but success rates can be very much higher with AI-enabled fraud. Between improved attack quality and greater attack volumes, the financial consequences of fraud are predicted to be significantly higher in the short to medium term. Batch runs and high-value payments are particularly vulnerable and can result in substantial, immediate losses; regulatory reporting obligations; and long, difficult remediation work.

3. AI Inside Your Own Ecosystem: Models, Vendors and Operational Risk

OSFI and other Canadian regulators have already highlighted how AI models can create new combinations of cyber risk, third-party risk, financial crime and operational breakdowns when used for core financial functions.

A second pattern is quieter but just as important: AI embedded in the systems used by your plan and its service providers. Today, many pension and benefits operations already rely on vendor systems that use machine learning to flag unusual transactions, detect possible fraud or automate portions of member communication. In the near term, many functions are likely to be "AI-assisted," such as chatbots that answer member questions, tools that help staff draft emails and models that prioritize which claims to review—sometimes powered by agentic AI.

Each of these uses comes with its own failure modes. A model that fails to detect an AI-enhanced fraud attempt, or a chatbot that provides misleading guidance on security steps, may not look like "fraud" at first glance, but it can be part of the chain of events leading to loss.

For fiduciaries, this means thinking about AI not only as something used by outsiders, but as a feature of the vendor and internal systems they already oversee. Even vendors that don't incorporate AI today will likely do so tomorrow. As a result, questions about model governance, data quality, access controls and incident response start to overlap directly with traditional duties to protect plan assets and member information. More and more, adoption of recognized AI governance frameworks, such as ISO/IEC 42001 or the NIST AI Risk Management Framework, is becoming part of the default due diligence standard for larger organizations.

4. Fake "AI" Investment and Retirement Products at the Edges of Your Ecosystem

A fourth cluster of risks, although less direct, sits at the edges of the plan: products and services marketed to members or sponsors under the banner of "AI."

Securities regulators in Canada and the U.S. have issued alerts about scams that promise exceptional returns from proprietary "AI trading systems" or "quantum AI" platforms, as well as about firms that overstated their use of AI to mislead investors. Often, these schemes target members directly through social media and online advertising, thereby sitting outside the formal, legal plan. But they can still create problems for fiduciaries and sponsors, such as when a legitimate pension or retirement plan transfers assets (at the request of the member or a deepfake of the member) to an illegitimate but seemingly real AI-branded scheme. In such cases, sponsors and unions may then find themselves fielding very legitimate questions about due diligence on providers, member education and how easily funds can be moved.

Fiduciary Duties in a Changing Fraud Landscape: What "Reasonable" Looks Like

With all of this talk about AI, it is fair for trustees and plan managers to ask a simple question: What does the law actually expect of us? Canadian courts and regulators do not demand perfection. They know that fraudsters innovate and that no system is bulletproof. What they do expect is a reasonable process in light of the risks that are known, or should be known, at the time. That is where AI-enabled fraud starts to matter for fiduciaries.

In a pension or large benefits plan, there are usually several people and entities "on the hook" to get that process right. The statutory administrator under pension standards legislation, trustees of a pension trust, boards and committees that make investment or administrative decisions, and sometimes employers or unions when they step into a decision-making role can all owe duties with a fiduciary flavour. Those duties do not disappear just because the plan hires a recordkeeper, an insurer or a technology vendor. Delegation of day-to-day work is permitted. Delegation of ultimate responsibility is not.

Neither are the core duties that intersect with AI-enabled fraud new. The duty of loyalty requires decision makers to act in the best interests of plan members and beneficiaries, rather than in their own or the sponsoring employer's interests. The duty of prudence or care requires them to act with the care, diligence and skill that a reasonable person in their position would use, taking into account the size and complexity of the plan and the roles they have accepted. That usually includes a duty to select, instruct and monitor agents and delegates carefully as well as to maintain appropriate records and controls over plan operations and assets.

For many plans, the best practical expression of those duties is found in guidance rather than in case law. CAPSA's Guideline No. 4 (December 2016) on pension plan governance highlights the need for structured governance frameworks, clear allocation of responsibilities and regular reviews

of key risks (including technology, cyber and third-party risks). CAPSA's Guideline No. 10 (September 2024) complements, reinforces and clarifies these requirements. Even though these guidelines are not legislation, regulators and courts increasingly look to them, along with similar guidance from OSFI and provincial regulators, as benchmarks for what a prudent administrator or trustee should be doing in practice.

Seen through that lens, AI does not necessarily create a new category of fiduciary duty. Rather, it changes the context in which existing duties are applied. If regulators across Canada and the U.S. are warning that impersonation, deepfakes and AI-enhanced phishing are now part of the mainstream fraud landscape, it becomes harder to argue that a plan can safely ignore those developments in its governance, vendor oversight or incident response planning. At the same time, no one expects every plan to build its own AI lab or to chase every headline about the latest exploit.

In practical terms, what "reasonable" looks like is this: Fiduciaries ask the right questions about how AI-enabled fraud could affect their plan, demand clear answers from their vendors and internal teams, refresh key policies and controls periodically, and document their decisions. Prudent process also includes considering risk transfer mechanisms—cyber insurance, fidelity bonds and D&O coverage that responds to technology-related incidents. Insurance is not a substitute for controls, but it is part of a layered defense. A plan that can show that it identified AI-related fraud risks, considered them alongside other operational risks and took proportionate steps to manage them will be in a very different position than one that treated AI as someone else's problem.

Provinces at a Glance: Same Duties, Different Emphases

Canadian pension and benefits law is famously jurisdiction-heavy. The good news is that, when it comes to AI-enabled fraud, the themes are more consistent than they are different. Administrators across the country are being pushed toward the same basic expectations: understand your risks, manage your vendors, and be ready to deal with technology and cyber incidents in a structured way.

1. Federally Regulated Plans: OSFI's Tech and Cyber Lens

For federally regulated private pension plans, OSFI has been clear that technology and cyber-risk are now part of mainstream prudential supervision, not side issues. Its recent advisory on technology and cybersecurity incident reporting sets out expectations for how FRPP administrators should report significant incidents that affect their plans. While OSFI's B-13 Technology and Cyber Risk Management Guideline is formally aimed at financial institutions, not pension plans, it reinforces the same message: Boards and senior management are expected to oversee technology and cyber risk in a disciplined, risk-based way.

For AI-enabled fraud, this translates into a simple practical point: If a deepfake-enabled fraud or AI-boosted attack causes a material incident at a federally regulated plan, OSFI will likely ask not only "What happened?" but also "What was your process for managing this category of risk before it happened?" OSFI's pending statement E-23 (effective May 1, 2027) strongly reinforces this orientation toward AI risk identification and management.

On the legislative front, the Artificial Intelligence and Data Act (AIDA), which had been proposed as part of Bill C-27, died on the order paper when Parliament was prorogued in January 2025. Evan Solomon, Canada's Minister of Artificial Intelligence and Digital Innovation, is currently leading the development of a refreshed national (Canada) AI strategy, with new comprehensive legislation and policy expected to be released sometime in 2026. Early indications suggest a principles-based approach with strong privacy orientation—closer to guidance than to prescriptive rules. For pension fiduciaries, that trajectory carries an important implication: If federal AI regulation remains light-touch, the onus of governing and controlling AI systems will fall more heavily on trustees and administrators themselves. The driver will be responsible for their duties to members and beneficiaries, not detailed regulatory checklists. Plans that wait for Ottawa to tell them exactly what to do may find themselves behind those that treat AI governance as an extension of existing fiduciary obligations.

2. Ontario: IT Risk and Cyber Guidance With Teeth

In Ontario, the Financial Services Regulatory Authority (FSRA) has moved beyond general expectations and into fairly detailed guidance on IT and cyber risk. Its Information Technology (IT) Risk Management Guidance applies across FSRA-regulated sectors, including pension plan administrators, and sets out "practices for effective IT risk management" and an IT risk incident notification protocol.

The guidance also sets out several expectations for administrators, including the following.

- Identify and assess key IT and cyber risks (which now clearly include AI-enabled threats).
- Implement controls and governance processes appropriate to the size and complexity of the plan.
- Report material IT risk incidents to FSRA using a prescribed form.

From an AI-fraud perspective, this means Ontario administrators will find it harder to argue that a significant AI-related incident was unforeseeable or outside their remit. The regulator has already said that IT and cyber risk belong squarely on the governance and risk-management agenda.

3. Québec: A Distinct Landscape

Québec combines strong governance expectations from Retraite Québec with Canada's most stringent provincial privacy regime. The result is a higher bar for AI-related incident response.

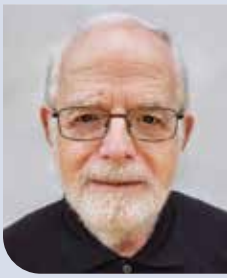
4. Other Provinces and Territories: CAPSA as the Common Denominator

Outside the federal, Ontario and Québec regimes, most jurisdictions look to CAPSA guidelines as the baseline for good practice. CAPSA's Governance Guideline (Guideline No. 4) and Risk Management Guideline (Guideline No. 10) set out expectations for structured risk management, clear allocation of responsibilities and ongoing monitoring. They also explicitly address cybersecurity, third-party risk and other operational risks as areas where administrators are expected to be active, not passive.

Provincial regulators frequently endorse these documents and encourage—or strongly expect—plans in their jurisdiction to align with them. For many single-employer and multi-employer plans, especially those that operate in multiple provinces, the practical path is to treat CAPSA's governance and risk-management guidelines as the "floor" and then layer on any province-specific rules or guidance.

Part 2 of this article will look at real incidents and examples of fraud, U.S. and international regulatory approaches in contrast to our own, the unique legal and regulatory regime regarding AI in Québec, and finally how management can begin to address this new risk landscape in a manageable and realistic way.

BIOS



Jerrard Gaertner is principal at Bizcom, an AI governance, monitoring and risk management firm. He is also a part-time professor of computer science. He earned his undergraduate degree from MIT and studied medicine at University of Pennsylvania and business at McGill University. His experience spans over 30 years at major multinational accounting firms, where he was responsible for complex statutory systems auditing, technology risk and computer security engagements. Gaertner is an expert in artificial intelligence, data analytics, computer security, forensic/fraud investigation, privacy law and regulatory compliance.

Paul Catenacci focuses his practice on all areas of employee benefits law and the operation of fringe benefit plans under the Employee Retirement Income Security Act (ERISA) and the Internal Revenue Code (IRC). He regularly advises single- and multi-employer fringe benefit plans across the country on the full range of issues that affect these plans.

Donald Broggi is a senior partner in Scott+Scott Attorneys at Law LLP's New York office. For the past 25 years, he has represented institutional investors, including public pension funds and Taft-Hartley funds in a wide variety of complex litigation, including securities fraud, derivative/corporate governance, antitrust/price-fixing and consumer cases.

pdf/326