

A1: Individual Assignment — AI Proof of Concept

Podcast script and AI readiness assessment: a Canadian bank's Microsoft Copilot pilot

Presented to

Mr. Renzo Casapia

Agentic AI for Business Transformation

DAT-1007

By

Alexandre Ly

47053031

28/06/26

PART 1 — INTRODUCTION

AL : Happy Sunday everyone and welcome back to the Agentic AI Investigators Podcast. I am your host, Alexandre Ly, and we're back on another Sunday to hear the most recent news on AI implementation across all of the industries you can think of, and to do our best to entertain you through this AI revolution. Today's topic revolves around the finance and banking industry in North America, specifically in Canada, and the role that chatbots and other forms of conversational agents are starting to play there.

In an attempt to streamline processes, many banks are now using chatbots to help employees find and research past agreements and contracts, and to draft investment proposals, all to solve efficiency gaps in their daily workload.

One of Canada’s major banks is no exception: it is planning to test-run a pilot project with Microsoft Copilot. The goal of this agent is to scan huge internal databases and retrieve valuable information quickly, in order to offload the repetitive, time-consuming research work that currently sits on employees' shoulders. Strictly speaking, this first pilot is an assistant rather than a fully autonomous agent: it finds and drafts, and a person decides. But assistants like this are usually the first step toward agents that act on their own, which is why some of today's questions look further ahead.

Banking is a useful lens for this conversation, because it's one of the most heavily regulated industries there is. A bank can't simply experiment freely. Every decision touches client money, client trust, and a wall of compliance requirements. So when a bank starts letting AI act autonomously, it's doing so under far more pressure and scrutiny than, say, a tech startup would face.

To make sense of what we hear, I'll use a lens from our class: the AI readiness framework of Tehrani and colleagues (2024), which looks at eight dimensions of readiness: informational, technological, environmental, infrastructural, data, participants, customers and process. At the end of the episode, I'll score the bank on each one, name the adoption strategy that fits, and give five recommendations.

PART 2 — THE INTERVIEW

Today, I am joined by our guest, a lead data analyst in commercial banking at one of Canada’s major banks, who has asked to remain anonymous for this episode.

She has kindly given us the privilege of hearing about her role at the bank and her view on the rise of AI.

For this interview, our guest recorded her answers in advance, which allowed her to give us fuller, more considered answers.

So, first question: what's really driving the bank to look into agentic AI, the kind that doesn't just respond, but can actually make decisions and act on its own?

Guest : You know, if I had to boil it down to one word, I'd say: exhaustion. Not in a dramatic way, but when you work in a large organization like ours, you see very talented people spending a huge chunk of their day on tasks that, frankly, don't require their talent. Processing the same type of request for the hundredth time, copy pasting information between systems, answering questions that have the same answer every single time. It wears people down, and it's also just not a good use of anyone's potential.

So when agentic AI came into the conversation, systems that don't just answer a question but actually follow through on a task from start to finish, there was a genuine sense of "okay, this could actually help." Not in a hype driven way, but in a very practical, ground level way.

And then you add the client side of it. People expect their bank to respond instantly, at any hour, without making them repeat themselves three times. That's a real pressure. The bank knows that if it doesn't find smarter ways to operate, the gap between what

clients expect and what can actually be delivered is only going to grow. Agentic AI is one of the most credible answers to that gap right now.

I'd add one thing that often gets overlooked: compliance. Counterintuitive as it sounds, a well governed AI agent can actually make compliance more consistent than a team of tired humans working under pressure at the end of a quarter. That matters a lot in banking.

AL : The bank already uses smart technologies, but how do you see the arrival of these new autonomous AI agents? Is there a specific department, like customer service or fraud detection, where you'd want to deploy them first?

Guest : Right, so I want to make sure we're clear on something first, because this trips a lot of people up. The bank has been using AI for years, but mostly in a "hey, look at this, a human should probably deal with it" kind of way. What's different now is that these new agents can actually close the loop themselves. They don't just flag the issue, they handle it.

So where do you start with that? You look for the places where the volume is high, the task is fairly structured, and a mistake, while unpleasant, isn't catastrophic. Customer service is the obvious first answer. Not because it's simple, but because the payoff is immediate. If an AI agent can resolve a straightforward inquiry at 2am without waking up a call centre representative, that's a win for the client and a win for the employee who gets to work on something more interesting during business hours.

Fraud detection is another one that gets everyone excited, and for good reason. The speed advantage is almost unfair, AI can analyze patterns across millions of transactions in the time it takes a human analyst to open their spreadsheet. When you're dealing with fraud, every second matters.

And then, and this one is less glamorous but honestly just as important, back office operations. The unglamorous stuff. Document verification, onboarding checks, internal IT tickets. Enormous amounts of employee time go into that, and it's exactly the kind of structured, rules based work that agents handle really well.

The guiding principle I always come back to is: start where you can show results quickly and where the stakes of getting it wrong are manageable. You need early wins to build internal confidence, because without that, you lose momentum fast.

AL : Letting an AI take initiative requires solid technical preparation. Does the bank feel truly ready to integrate this into daily operations?

Guest : Honestly? "Ready" is a word I use carefully. I think the bank is ready to start, but I wouldn't say it's ready to go full speed, and I don't think any serious institution should be, at this point.

The technical infrastructure is solid. The data is there, the cloud capabilities are there, there's already a history of AI investment. So you're not building from zero. But technical readiness and operational readiness are two very different things, and sometimes organizations confuse them.

What I mean is, do you have your processes documented clearly enough that an AI can follow them? Do you have the governance in place to catch problems early? Do you have people who understand both the business context and the technology well enough to oversee these systems? Those are harder questions, and they take longer to answer than "do we have the infrastructure."

What I've seen work well is a very deliberate, almost cautious approach: run a pilot in one well defined area, observe closely, document what goes wrong and why, fix it, and then and only then think about scaling. It's not the most exciting story to tell, but it's the one that actually leads somewhere.

The bank knows this is a multi year journey. And I think that's the right mindset. You don't transform an institution of that size overnight, and you probably shouldn't want to.

AL : Banking is a very regulated space, and letting a machine operate autonomously can raise concerns. How do you ensure these AI systems make sound decisions, remain secure, and align with the bank's values?

Guest : This one I feel strongly about, so bear with me.

The thing that worries me most, not just at our bank but across the industry, is when organizations treat ethics and compliance as a checklist. "Did we tick the fairness box? Great, we're done." That's not how it works, and in banking, the consequences of getting it wrong are serious.

The foundation, for me, is accountability. Autonomy does not mean no one is responsible. If an AI agent makes a decision that harms a client, someone owns that. Someone designed the system, someone approved it, someone is supposed to be watching it. That chain cannot have gaps.

In practice, what that looks like is: every agent has clearly defined limits, things it's allowed to do, things it absolutely cannot do, and situations where it has to stop and escalate to a human. Every action it takes is logged. Not in a vague "we have records" way, in a way where, if a regulator walks through the door tomorrow and asks why a specific decision was made on a specific account, you can pull up a clear, readable explanation. That's not optional in banking.

And then there's the bias question, which people sometimes dismiss as abstract but really isn't. If you train an AI on historical credit decisions and those historical decisions had discriminatory patterns baked into them, the AI learns those patterns. It doesn't know they're wrong. That's on us to catch, through regular audits, diverse teams

reviewing outputs, and a genuine willingness to slow down and fix things when problems surface.

I think of a well governed AI agent kind of like a very capable new employee. They might be technically brilliant, but you still give them clear guidelines, you still supervise them, and you still do performance reviews. The difference is the AI doesn't get offended when you look over its shoulder.

AL : Introducing systems that work on their own can stress employees out. Have you noticed any resistance or anxiety within your teams, and how are you managing that change?

Guest : Oh, absolutely. And look, I think the anxiety is completely understandable. I'd be more worried if people weren't reacting at all.

What I've observed is that the fear comes in layers. There's the obvious one, "is this going to take my job?", which people sometimes ask directly but more often just carry around quietly. And then there's a second layer that's almost harder to address, which is the fear of becoming irrelevant. Even if your job technically survives, what happens when the interesting parts of it get automated? What's left?

The questions that keep coming up in team meetings, in hallway conversations, are always some version of "what does this mean for me specifically?" General reassurances from leadership don't really land at that level. What lands is when a manager sits down with their team and walks through, concretely, how their day to day is going to evolve. That's where trust is either built or broken.

I'll be honest with you, I think organizations sometimes make it worse by being overly optimistic in their messaging. "AI will only make your job better, don't worry!", people can see through that, and when they do, they stop trusting the rest of what they're told.

A more honest message, something like "some things will change, we're going to help you navigate it, and here's how", that actually tends to land better.

The real goal of change management isn't to eliminate anxiety. It's to make people feel like they're moving through this with support, rather than having it happen to them.

AL : Do the leaders see AI as a powerful tool to make employees more effective, or more as a way to reduce headcount? And how will you know if this AI initiative is actually a worthwhile investment for the bank?

Guest : I want to answer this one honestly, because I think you deserve a straight answer rather than the corporate version.

The official position, and I do believe the people saying it mean it, is that this is about making employees more effective, not replacing them. The vision is genuinely about freeing up human capacity for higher value work. And at this stage of the journey, that's largely true.

But, and this is important, any senior leader evaluating a major technology investment is also thinking about cost. That's just reality. Automation does have long term implications for how organizations are staffed, and pretending otherwise wouldn't be honest. The question is what you do with the efficiency gains. Do you invest them back into your people, through training and more meaningful roles? Or do you take the short route and just reduce headcount and call it a success?

The organizations I respect are the ones that explicitly plan for the first option. They don't just say "we'll upskill people", they have actual programs, actual timelines, actual commitments.

On measuring success, the trap I see constantly is measuring only what's easy to quantify. Yes, you track cost reduction and processing speed. But if your client

satisfaction scores are dropping because the AI agent keeps mishandling edge cases, or if your employees are so demoralized they're leaving, those numbers should terrify you more than a slow efficiency gain. Real success is multidimensional, and the metrics you choose before you launch tell you a lot about whether an organization actually understands that.

AL : Drawing on your experience with these technology projects, what's the biggest lesson you've taken away? If you had one piece of advice for another bank thinking about launching autonomous AI, what would it be?

Guest : The biggest lesson, and I've had to learn this more than once, honestly, is that the technology is almost never the real problem.

Every time I've seen an AI initiative struggle or stall, if you dig deep enough, it's not because the model wasn't good enough or the infrastructure wasn't there. It's because nobody clearly defined what problem they were actually solving. Or they tried to automate a process that was already broken and expected the AI to fix the brokenness. Or they didn't bring their employees along and then wondered why adoption was so low.

Those are human and organizational problems. And they don't get solved by better algorithms.

So if I had one piece of advice, just one, I'd say: before you talk to a single vendor, before you run a single demo, write one clear sentence about the specific problem you're trying to solve. Not "we want to leverage AI capabilities", that means nothing. Something like: "Our fraud analysts spend 60% of their time chasing false positives that never lead anywhere, and it's burning out our best people." That kind of clarity is worth more than any technology decision you'll make.

Start there. Start small. Learn from what actually happens, not from what you hoped would happen. And let genuine results build the case for doing more. That's the only path I've seen actually work.

PART 3 — AI READINESS ASSESSMENT

AL : Let's step back and assess what we just heard. The core motivation was efficiency, but in a deeper sense than just “doing things faster.” Our guest described very talented people spending a huge chunk of their day on repetitive work that doesn't require their talent. In the Copilot pilot, that work is digging through old contracts, hunting for precedents and assembling the building blocks of a proposal by hand. The tool is meant to take over that research so people can focus on the analysis and judgment that actually require a human. There is also a client angle: people expect fast, polished service at any hour, and the bank knows the gap between what clients expect and what it can deliver will only grow. And our guest made a counterintuitive point: a well-governed agent can make compliance more consistent than a tired team working under end-of-quarter pressure.

To judge whether the bank is ready, I'm using the framework from Tehrani and colleagues, published in *Technovation* in 2024. They describe AI readiness as how willing and able an organization's people are, and how well suited its environment, processes, data and resources are, to adopt and operate AI. They split it into eight dimensions and insist that readiness is a matter of degree, not a yes-or-no answer. That fits our guest's own verdict perfectly: the bank is ready to *start*, not to go full speed. So here is my scorecard, and you'll find the full table in Appendix A.

Where the bank is strong. Informational readiness is the clearest strength. Our guest knows the use cases, she has a rule for choosing them (high volume, structured work, and a mistake that is unpleasant but not catastrophic), and her biggest lesson is about problem recognition: define the problem before you look at the technology. The one piece missing is a comparison of Copilot against other options, including simpler

automation. Technological maturity is also strong: the bank has used AI for years, mostly to flag issues for humans, so it isn't building from zero.

Where the bank is partly ready. On the environment, our guest is clearly alert to regulation and accountability, but the specific rules still need to be mapped to this pilot. In Canada, that means OSFI's Guideline E-23 on model risk, which explicitly covers AI models and takes effect in May 2027; OSFI's Guideline B-10, under which the bank remains accountable for whatever a vendor like Microsoft does on its behalf; and privacy law, including Quebec's Law 25, which already requires organizations to tell people when a decision about them is made by automated processing alone, a rule that matters as soon as agents start deciding things about clients. On infrastructure, the cloud and IT resources are there, but our guest flagged the harder question herself: does the bank have people who understand both the business and the technology well enough to oversee these systems? The budget never came up, and the paper stresses that AI funding has to be flexible and continuous, not a one-time purchase.

On participants, staff readiness is the most visible gap. The anxiety is real, both the fear of losing a job and the quieter fear of becoming irrelevant, and our guest was clear that trust is built or broken in day-to-day conversations with managers. The partner side, meaning how aligned Microsoft is with the bank's rules, didn't come up at all. On process, our guest's own question still hangs in the air: are processes documented clearly enough for an AI to follow them? The pilot loop she described (observe, document what goes wrong, fix it) is a good start on a feedback mechanism, but it isn't a formal one yet. And on customers, the bank understands what clients need, but the interview didn't cover client privacy and security, which the paper flags as especially sensitive in banking.

Where readiness is unverified. For a retrieval tool, data readiness decides everything, and it's the dimension we know least about. “The data is there” tells us about availability and volume, not quality. Copilot will only be as good as the contracts it searches: if those repositories hold outdated versions, duplicates, or files that too many people can

open, the assistant will surface the wrong precedent with total confidence. Our Philippine Airlines case called this being data rich but information poor, and as we saw in class, data preparation usually takes about 80% of an AI project's effort.

Put together, my verdict matches our guest's: the bank is ready for a controlled pilot with a human reviewing every output, but not yet ready to scale it or give it more autonomy.

That also tells us which adoption strategy fits. The paper maps strategies on two axes: the managerial goal, cost reduction or differentiation, and the level of risk. Here the goal is mainly productivity, freeing up analyst hours, with a differentiation upside in faster, sharper client proposals. The risk is high: confidential client data, regulators, and employees who might reject the tool. That places the bank in the high-risk half of the matrix, where the paper points to *crawling*, meaning small, contained use cases with the least impact on the whole operation, and *partnership*, meaning you lean on a vendor's capability while building your own. That is exactly our guest's advice: start small, show results, then scale. Her examples were customer service, fraud detection and back-office work, but information retrieval and proposal drafting fit her rule even better, because the assistant helps rather than decides, as long as a human reviews everything before it goes out.

PART 4 — RECOMMENDATIONS

Based on that assessment, here are my five recommendations, each aimed at one of the gaps.

One: write the problem down, and measure it before you start. Before talking to any vendor, write one clear sentence describing the specific problem you're solving. Not “we want to use AI,” but something like: “our commercial banking analysts spend hours every week digging through old contracts to find precedents for new proposals.” Then put a number on it by timing that search-and-draft work for a few weeks before the pilot, so there's a baseline to measure against. As our session on retail forecasting put it, value lives in the decision, not in the model: here, the value is in hours given back and

better proposals, not in how impressive Copilot's answers look. This closes the informational gap and covers the first two steps of our implementation process, the value and the target.

Two: make data readiness the first workstream, not an afterthought. Inventory the repositories Copilot will search, retire or clearly flag superseded contracts, fix access rights so the assistant can't surface documents a user shouldn't see, and name an owner for every source. In class terms: identify, store, process, provision and govern the data, and answer the question “who owns the data?” before launch, not after.

Three: build governance before you build autonomy. Document what the assistant can do, what it must never do, and when it has to hand off to a human. Every answer should cite its source so an analyst can check it in seconds, and every action should be logged so it's auditable and explainable. Run a privacy review before any client files are indexed, bring Microsoft through the bank's third-party risk process, and assess the tool under the model risk framework that E-23 formalizes. Treat bias testing as an ongoing process, not a one-time checkbox. The goal is an AI you can supervise like a capable new employee: clear guidelines, real oversight, regular review.

Four: define success on three levels, and agree on the bar for scaling before launch. Our implementation framework separates model success, user success and business success. For the model: does it find the right, current document, and are its answers accurate, relevant, complete and free of bias? For users: are analysts actually using it, and do they trust it? For the business: hours saved against the baseline, proposal turnaround time, error and compliance incidents, and client satisfaction. In banking especially, one major failure can wipe out years of efficiency gains, so set go/no-go thresholds in advance and test in real work, not just in demos, because backtesting flatters reality. And give users a simple way to flag bad answers, the way Philippine Airlines used pilot feedback to keep improving its fuel program.

Five: treat adoption as the project. In our retail forecasting session, the biggest blocker turned out to be fear, not technology, and our guest described the same thing. So don't promise that nothing will change; people see through it, and it erodes trust. Be honest that some roles will evolve, then back it up with real training, including how to prompt and how to verify an answer, and with a clear plan for reinvesting the hours saved into people. Let direct managers lead the conversation, and involve employees in choosing which tasks to automate. That turns them from anxious bystanders into active participants.

CLOSING

AL : So, is the bank ready? Ready to start, yes. Ready to scale, not yet, and that's fine. Readiness is built one dimension at a time, and a well-run pilot is how you build it.

Thank you everyone for listening to this episode of the Agentic AI Investigators Podcast, produced in podcast format for the course Agentic AI for Business Transformation and submitted to Professor Renzo Casapia, whom I would like to thank for the opportunity and the great research. And of course, thank you to our guest for her contribution and for sharing her knowledge in this endeavour.

APPENDIX A — AI READINESS SCORECARD

Framework: Tehrani et al. (2024), eight dimensions of AI readiness. Evidence comes from the interview; "not evidenced" means the topic did not come up and would need to be verified with the bank. Readiness is scored on a continuum: Strong, Partial, Not evidenced.

Table A1 — AI readiness of the bank for the Copilot pilot

Dimension	Evidence from the interview	Rating	Gap to close
Informational use cases · problem recognition · AI solutions	Clear criteria for choosing use cases (volume, structure, cost of error); candidate areas named; central lesson is to define the problem first.	**Strong**	Quantify the Copilot problem statement; compare Copilot with alternatives, including simpler search or automation.
Technological maturity · AI techniques · IT support	Years of AI use, mostly flag-for-human; "not building from zero." Techniques and IT support not discussed.	**Strong (maturity only)**	Confirm source-citation set-up, evaluation method and a support channel for pilot users.
Environmental culture · leadership · regulatory environment	Strong accountability mindset (limits, escalation, logging, explainability to regulators); augmentation-first leadership position; cautious, pilot-first culture.	**Partial**	Map obligations: OSFI E-23 (model risk incl. AI, effective May 1, 2027), OSFI B-10 (third-party risk), PIPEDA and Quebec Law 25; name an executive sponsor and a reinvestment commitment.
Infrastructural financial · human · IT resources	Cloud and IT infrastructure described as solid; people who understand both business and technology flagged as the harder question; budget not discussed.	**Partial**	Fund the pilot, data clean-up and training as a multi-year line; build hybrid talent through training and hiring.
Data availability · quality · volume	"The data is there" (availability, volume); quality, versioning and access rights of contract repositories not evidenced.	**Not evidenced (critical)**	Inventory sources; retire or flag superseded versions; remove duplicates; fix permissions; assign data owners; test retrieval accuracy before launch.
Participants managers · staff · partners	Real anxiety (job loss, irrelevance); concrete, manager-led conversations identified as what works; vendor alignment not discussed.	**Partial**	Manager-led change plan; employee co-design of automated tasks; pilot champions; vendor governance.
Customers needs · privacy & security · acceptance	Clear view of client expectations (speed, any hour, no repetition); client privacy and acceptance of AI-assisted proposals not discussed.	**Partial**	Privacy impact assessment before indexing client files; human sign-off rule for any AI-assisted client document.
Process operational integration · feedback · integrated communication	Process documentation raised as an open question; pilot learning loop (observe, document, fix); honest, manager-led communication advocated.	**Partial**	Map the real proposal workflow; in-tool feedback button; formal go/no-go gates.
Overall: ready to start a controlled, human-reviewed pilot; not ready to scale or increase autonomy. **Fitting adoption strategy** (Tehrani et al., 2024, Fig. 3): crawling combined with partnership, the high-risk quadrant.			

Table A2 — The Copilot pilot through the six implementation steps (course framework)

Step	Application to the Copilot pilot
1. Value deployment goal	Give analysts back the hours spent searching past agreements; faster, more consistent proposals.
2. Target prediction goal	For each query, return the correct, current source with a citation; produce a first-draft proposal section for human review.
3. Performance evaluation metrics	Model: retrieval accuracy, citation rate, error rate. User: adoption, satisfaction, trust. Business: hours saved vs. baseline, turnaround time, compliance incidents, client satisfaction.
4. Fuel data	Contract and proposal repositories, cleaned, versioned, permissioned and owned (identify, store, process, provision, govern).
5. Algorithm model	Pre-trained vendor language model grounded in the bank's own documents; no custom training. Favour traceable, citable answers.
6. Launch deployment	One team, fixed duration, go/no-go gates, then phased rollout; monitor and iterate.

APPENDIX B — REFERENCES

Primary source: recorded interview responses from a lead data analyst in commercial banking at a major Canadian bank (personal communication, 2026; name withheld at the interviewee's request).

Agentic AI for Business Transformation (DaT-1007). (2026). *AI fundamentals* and *AI implementation* [Lecture slides]. Hult International Business School.

AI & Data in Practice. (2026). *Demand forecasting* [Guest session slides]. Agentic AI for Business Transformation, Hult International Business School.

Bernabe, L. E. R. (2022). Philippine Airlines steps up to fuel efficiency with an intelligent solution. *Aircraft IT Operations*, Winter 2022/23, 48–57.

Office of the Superintendent of Financial Institutions. (2023). *Guideline B-10: Third-party risk management* (effective May 1, 2024).

Office of the Superintendent of Financial Institutions. (2025). *Guideline E-23: Model risk management (2027)* (effective May 1, 2027).

Québec. *Act respecting the protection of personal information in the private sector*, CQLR c P-39.1, s 12.1 (as amended by Law 25).

Tehrani, A. N., Ray, S., Roy, S. K., Gruner, R. L., & Appio, F. P. (2024). Decoding AI readiness: An in-depth analysis of key dimensions in multinational corporations. *Technovation, 131*, 102948. https://doi.org/10.1016/j.technovation.2023.102948