

Swiss Finance Institute Roundups

AI in Banking—Intermediation, Automation, and Dependency

Editorial



Artificial intelligence is no longer just an isolated operational experiment in banking; it is actively restructuring intermediation, workflows, and institutional accountability. In this SFI Roundup, five contributors from academia and practice discuss the impact of AI on banking. Gains only appear when banks redesign processes and invest in monitoring. Trust, relationships, and the ability to verify an answer still give clients reasons to pay a bank, even as new entrants build on the same tools. The contributors also point to risks that get less attention: correlated errors across institutions, Europe's dependence on foreign providers, and the erosion of human judgment. They agree on one point: when a decision goes wrong, people must remain accountable.

We wish you an informative and thought-provoking read.

A handwritten signature in blue ink, appearing to read 'F. Degeorge', with a horizontal line underneath.

Prof. François Degeorge
Managing Director

Contributors



Gianluca De Nard

Gianluca De Nard is Head of Systematic Strategies at Zürcher Kantonalbank, where he leads systematic equity funds and FX overlay mandates within Swisscanto managed by Zürcher Kantonalbank. He is also professor of practice in quantitative finance and systematic investing at the University of Liechtenstein and lectures in empirical finance at the University of Zurich. He previously headed quantitative research at OLZ, a Swiss-based asset and wealth manager. He holds a PhD in finance from the University of Zurich.



Christophe Makni

Christophe Makni is Head of Data, Automation, and AI at Migros Bank, where he leads the scaling of AI across the bank, from governance and infrastructure through systems running in production. He lectures on artificial intelligence, leadership, and banking at ETH Zurich, the University of Applied Sciences Northwestern Switzerland, and the University of St.Gallen. He holds an MSc in information technology from the Karlsruhe Institute of Technology.



Markus Leippold

Markus Leippold is SFI Senior Chair and holds the chair in financial engineering at the University of Zurich. He is also a part-time research scientist at Google DeepMind and an associated research fellow at the University of Oxford. Before his current positions, he was an associate professor in quantitative finance at Imperial College London. His research focuses on climate finance, sustainability, and artificial intelligence. He holds a PhD in economics and finance from the University of St.Gallen.



Fernando Perez-Cruz

Fernando Perez-Cruz is Senior Adviser on innovation at the Bank for International Settlements and adjunct professor in the computer science department at ETH Zurich. He researches the reasoning limits of large language models, the performance of AI agents, and the use of AI to monitor financial market stress. He was previously the chief data scientist at the Swiss Data Science Center and a research scientist at Amazon. He holds a PhD in electrical engineering from the Technical University of Madrid.



Norman Schürhoff

Norman Schürhoff is SFI Senior Chair and professor of finance at the University of Lausanne. He has been a visiting scholar in the Haas School of Business at the University of California, Berkeley, on numerous occasions. His research includes the information content of financial advice on social media and the use of machine learning and artificial intelligence in finance. He holds a PhD in financial economics from Carnegie Mellon University.

The contributions in this Roundup are based on individual interviews, edited for clarity and style, and reviewed by the contributors. They are not *verbatim* transcripts. The views and conclusions expressed are strictly those of the individual contributors and do not represent the views of their respective employers or institutions.

October 2026 (interviews as of September 2026 and data as of October 2026)

1. What We Mean by AI

When a bank says it "uses AI" today, what is it actually using—the technology that arrived with ChatGPT, or the machine learning it has run for years—and why does the difference matter?

► **Fernando Perez-Cruz:** Artificial intelligence is a very broad topic, and I find it useful to distinguish three very different types of AI. The first is the universal assistant, such as ChatGPT, that augments employees' knowledge and skills in almost any task. It makes them faster and lets them do things they might not otherwise have attempted. The second is the same generative technology pointed at a defined job, fenced with guardrails, and running steps a person used to run. Anti-money-laundering and compliance checks are clear banking examples. Either of these two types of AI becomes agentic once the model is given access to tools, such as a web search or database, or the ability to execute a step, so that it no longer only produces text but takes actions on behalf of the employee. Today, state-of-the-art AI systems can be deployed agentially; the difference that matters is whether their decisions are final and irrevocable or whether the AI systems are merely providing advice. The third type of AI consists of machine learning solutions that banks have trained on their own data and have used for years, each for a single purpose, such as credit scoring or algorithmic trading. When someone says AI is old news, they usually mean this third type. When they say AI changes everything, they mean the first two, both of which rest on generative models. Banks are adopting all three types.

► **Markus Leippold:** Banks have indeed been using machine learning since the early 1990s for credit risk and fraud detection, and that is AI too. However, since the explosion of ChatGPT, AI has become a buzzword and has narrowed to mean generative AI alone, when generative AI is only one part of the field. The distinction matters for policy, not just vocabulary. Predictive AI produces an output that a person then acts on, and the bank retains control of the decision-making. Generative AI creates a different temptation, because it invites you to offload not only the information processing task but the judgment itself. Putting everything in one basket leads to confusion and makes sensible regulation harder.

► **Norman Schürhoff:** Something deserves to be said about the "intelligence" in AI. At the extreme, the term means artificial general intelligence: machines that match human capability across the full range of cognitive tasks. We are nowhere near that. Current generative language models are extraordinarily capable statistical systems trained to predict what comes next, but they do not understand their tasks as humans do. They still create real value, because they can do things no human can. Most of the AI banks actually use today is machine learning, and it remains underused. A bank with machine learning genuinely embedded in its processes would already be a significant step ahead.

► **Christophe Makni:** In a regulated bank, you do not get to pick the definition you find most satisfying. The Swiss Financial Market Supervisory Authority (FINMA) does not prescribe its own definition of AI, but it expects banks to adopt a sufficiently broad one. At Migros Bank, we therefore took the definition in the European Union's Artificial Intelligence Act as our governance anchor. We chose it because external auditors recognize it and because it provides a clear, defensible basis for our AI governance. In a bank, a definition is a compliance instrument before it is anything else. In practice, we use different technologies, such as traditional machine learning and generative AI, in the fields of data analytics and intelligent automation to create value for our customers and for the bank.

Is AI still a set of isolated tools, or is it starting to reshape whole functions?

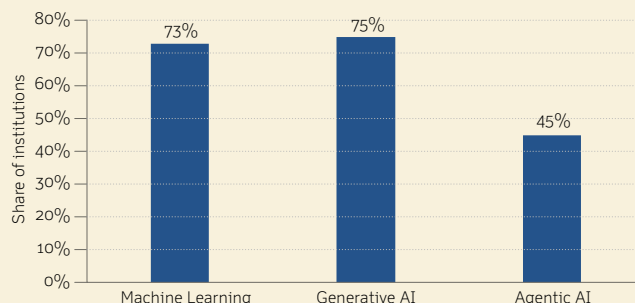
► **Gianluca De Nard:** At the moment, I see both. Many of the most successful applications are still relatively bottom-up. A specialist runs into a problem, such as an optimization that will not converge, a risk metric behaving strangely, or a hunt for outperformance. They find a tool that helps and build it into their own work. That is how innovation often starts. But we are now moving toward a more interesting phase: How do you turn these individual productivity gains into an institution-wide capability? Among both fundamental and quantitative portfolio managers, the differences in adoption are still substantial. Getting to an institution-wide capability requires a top-down AI strategy, and the biggest challenge is common infrastructure: centrally available data, access to models, and governance and security standards. Equally important is a workforce that knows how to use these tools. The real transformation will not come from giving everybody a chatbot. It will come when AI is integrated into workflows, data infrastructure, and decision-making processes across the organization. That requires changing how we store and access data and, ultimately, how we work.

► **Christophe Makni:** Five years ago, we began with machine learning, a chatbot, and a voicebot. Back then our inventory looked exactly as you would expect, with one chatbot here and one voicebot there. Years of work have taken us somewhere else. We now operate numerous generative AI use cases in production, with workflows and automation alongside them. The value never comes from adding an individual tool. It comes from building AI into a process from one end to the other, combined with automation. To change a whole function, you have to redesign the process itself. That is where we are today, at the starting point of reshaping functions.

► **Markus Leippold:** Since 2025, the AI landscape has genuinely shifted from a set of tools to a function. In some institutions you can now orchestrate AI tools into a single system that covers the capabilities needed across all aspects of financial intermediation, rather than using one tool to onboard a client and another to assess creditworthiness. Several large US banks, JPMorgan Chase and Goldman Sachs among them, have been explicit about this by investing heavily in AI and laying the groundwork for workforce reductions.

► **Norman Schürhoff:** I can only offer an educated guess. But I foresee a lot of heterogeneity, and that may explain why the answers to this question differ so much. A top US bank, or a sophisticated fund or alternative asset manager, probably has machine learning embedded very deeply in its processes already. That does not mean a traditional bank does. And the obstacles are familiar ones. Connecting old systems to new ones is complex and expensive. The data is stored in systems that were never designed to share it. And people inside the institution have their own reasons to resist a tool that changes their work. None of that is an AI problem, and none of it is solved by buying a better model.

AI Adoption by Traditional Financial Institutions



Note. Share of traditional financial institutions worldwide at or beyond the piloting stage for each type of AI. Adoption means an institution has started, not that the technology carries much of its work. Based on 149 institutions across banking, insurance, payments, capital markets, and digital assets, surveyed between October 2025 and January 2026.

Source: Cambridge Centre for Alternative Finance. (2026). *2026 global AI in financial services report: Adoption, impact and risks*. University of Cambridge Judge Business School. <https://www.jbs.cam.ac.uk/faculty-research/centres/alternative-finance/publications/2026-global-ai-in-financial-services-report/>

2. Productivity and the Future of Work

Where is AI producing real productivity gains in banking—in the models themselves or in the way work is done?

► **Gianluca De Nard:** It depends where you look. If you look at the financial models themselves, not a great deal has changed. The traditional multi-factor drivers still hold, such as value, quality, momentum, or low risk. You cannot run the plain-vanilla model published 30 years ago, but the drivers are still the same, and the machine learning algorithms that have been used for several years are still the ones that often work best. So there has been no step change in the investment models. What has changed, and changed a lot, is the workflow, and that shift happened over about two years. We can code a new prototype of a strategy far faster than before. We can also use these tools as research assistants, to gather data and digest the rapidly growing volume of new research. Two or three years ago, the frequency of AI hallucinations—incorrect or invented answers—made that impossible, but now the quality of the AI models is good enough. The honest answer is that the payoff is modest in the investment models and real in how the work gets done.

► **Christophe Makni:** The payoff is real. The easy version is augmentation, where you help teams work faster and deploy tools so that anybody in the bank can do their job better. The bigger value comes from integrating AI into the process itself rather than adding it on top, and that means redesigning the work. You do not have to choose between the two. If a process is simple and a small automation can already generate savings, do it. If a process is complex and high value, redesign it. What you cannot do is bolt a tool onto an unchanged process and expect a fundamental shift.

► **Markus Leippold:** There is a cost on the other side of the gain, and it is easy to miss. In a study I co-authored with researchers at Google DeepMind, we worked with 13 climate scientists to synthesize 79 papers on the stability of the Atlantic Meridional Overturning Circulation—the system of currents that transports heat through the Atlantic and helps regulate the global climate. The scientists reported a two- to threefold speedup, with AI proving highly effective at synthesizing the literature and answering the "what." Yet the scientists themselves contributed 58 percent of the final content of the study. Human expertise remained essential for answering the "so what" and the "now what" questions: weighing conflicting evidence, recognizing what fell outside the consensus, and

preserving uncertainty. The AI-produced text tended toward the mean, overlooking important outlying literature and making its conclusions sound more confident than the evidence warranted. Correcting both problems required human judgment, and that is the part nobody budgets for.

Does AI augment or replace people, and how will the next generation of bankers build expertise?

► **Norman Schürhoff:** Economists are genuinely divided on this question. One camp argues that these technologies mostly displace workers, that the gains accrue to capital rather than labor, and that the aggregate productivity effect will be smaller than the current enthusiasm suggests. The other camp points to a historical pattern of complementarity, in which new technologies raise the return to human skill and generate work nobody anticipated. Both sides have serious evidence, and I do not think the question is settled. Banking may see particularly large effects, because so many of its processes lend themselves to automation. However, that is not the same as saying the workforce will benefit.

► **Markus Leippold:** The answer changes sharply with job seniority, and the line between augmentation and automation is thin. Senior people manage to augment themselves, because they already have the expertise and can supply the judgment the model lacks. Meanwhile, the demand for new university graduates appears to be decreasing, because the obvious way to cut costs in the short term is to augment the experts you already have. What worries me is that this decision is rational for each bank, yet damaging for all of them. A junior does not have the judgment a senior has accumulated over years in the business. But if the industry stops hiring juniors and letting them build that judgment, then in 15 years the senior experts simply will not exist. However, you could also argue the opposite, that a junior can become an expert far faster with AI at hand. Ultimately, the answer depends on how the workflow is designed, yet too little attention has been paid to how the next generation of bankers will develop expertise.

► **Gianluca De Nard:** Entry-level jobs are indeed harder to find at the moment, and I think replacement is part of the reason why. In my view, over the long run AI will clearly be a means of augmentation rather than replacement, even if the work itself changes. You still need a quant who works with the model, gives it the idea, and knows how to interact with it. You cannot say the investment process will now be solved by a large language model.

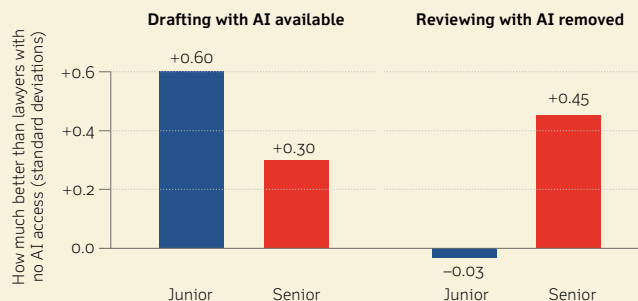
As roles change, what does banking need to preserve and create?

► **Gianluca De Nard:** Creative destruction has been around for centuries, and history is reasonably clear about it. Previous technological revolutions displaced some forms of work, while creating new demand for skills. I would expect the same here. What concerns me is the transition rather than the destination. There is a temptation right now to cut experienced people because a tool can approximate part of what they do, and that is a mistake, because a great deal of human capital is lost and will not come back. We recently saw several companies, including financial firms, doing exactly that and then rehiring, because the task turned out to be full of exceptions the system had never been told about—and there are always exceptions nobody told it about. The lesson is not that automation fails. It is that evolving the human and the machine together is more efficient than moving fast and rehiring later.

► **Fernando Perez-Cruz:** I genuinely do not know, and it is too early to tell what the new normal will be. What I would say is that the new roles will not be invented by people of my generation. They will come from the people who are teenagers now, and probably from a startup rather than an established bank. Think about what happened to the car industry. For decades the barrier to entry was the combustion engine, and no newcomer could realistically build one well enough to compete. Then electric powertrains lowered that barrier, and the disruption came from companies that had never built a car. Something comparable will happen here. The difficulty is that you cannot tell in advance which barrier will fall, or who will walk through it.

► **Markus Leippold:** What has to be created is not only new roles, but also a better way of working alongside the machine. The public debate so far has been almost entirely about making AI better. Very little of it has been about how the human gets better at the same time. Both need to improve, and there is obvious room for both, but it requires deliberate adaptation, rather than simply handing the work over. If we get that right, the jobs that follow will be more capable versions of the ones we have. If we get it wrong, we will have people who have stopped exercising judgment because the machine appeared to be doing it for them.

Patent Lawyers Who Gained Most From AI Kept the Least



Note. Juniors improved most while using AI but kept none of that advantage once it was withdrawn; seniors improved less with it and retained more without it. Quality is measured in standard deviations of the scores of lawyers without AI access; a difference of 0.5 is large. Juniors have fewer than seven years of experience. Lawyers at US firms were randomly assigned three months with or without AI access; their work was graded blind by expert attorneys.

Source: Autor, D., Rodchenko, T., Martin, J., Iscenko, Z., Strand, S., Pearl, D., & Ferere, M. (2026). *Does AI assistance enhance or erode expertise? Evidence from a three-month field experiment in patent drafting* (NBER Working Paper No. 35720). National Bureau of Economic Research. <https://www.nber.org/papers/w35720>

3. The Future of Financial Intermediation

If AI makes information and analysis cheap, what remains scarce enough for clients to pay a bank for?

► **Christophe Makni:** My conviction is that the more we live in an AI world, the more we will pay to interact with a human being. Trust is what Swiss banks compete on, as everybody offers more or less the same products. And I expect trust to matter more in the years ahead, not less. Look at what has filled the digital space around us, with a proliferation of fintechs and crypto offerings of wildly varying quality. That surge makes a regulated, established institution more valuable rather than less. Banking services are essential to how a society functions, which is precisely why they are regulated. That does not entitle banks to stand still. We have to adapt, and quickly. But I would add one qualification. Trust is not reserved for humans. People trust a machine when the quality is consistently high, in the same way they trust a car without understanding the engineering. So the question is not humans against machines. It is whether the institution behind either one has earned our trust. And leveraging AI helps us be more efficient internally in order to free up more time for our customers and offer them even more personalized services.

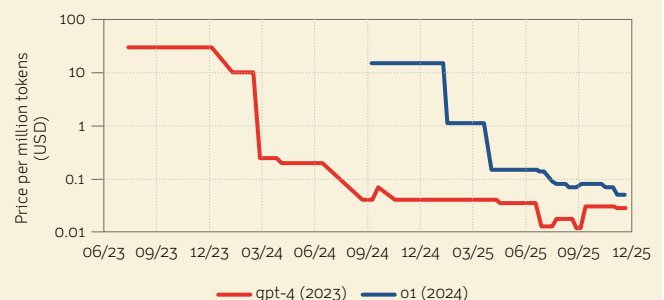
► **Markus Leippold:** Clients pay for something they cannot do themselves, and the gap is wider than people assume. There is no off-the-shelf AI product that can yet comprehensively and reliably run a household's finances. Even if there were, the client would still have to verify what it produced, and verification requires exactly the expertise and resources the client does not have. Banks face the same problem internally, but the client is far worse equipped to solve it. So I do not believe bank margins will disappear. What clients pay for is the assurance that somebody competent stands behind the decision. Regarding trust, I would add that banks do not possess it simply by virtue of being banks. Trust is generated by building a relationship over time, and that is a different asset altogether.

► **Norman Schürhoff:** Think about what a bank actually provides: capital, access to markets, advice, trust, and relationships. Information processing is only one input among those, and it is precisely the one AI makes cheap. Clients will pay less for the analysis itself, because its cost is collapsing. Those other sources of value do not behave the same way. This is something I am researching. My co-authors and I are examining decisions that appear suboptimal in a single transaction, but pay off later because the person was investing in a relationship.

A machine optimizing the immediate objective may not make that concession, because the link to the eventual payoff is too indirect. This suggests that some work may resist automation not because the technology lacks capability, but because relationships create value over time. There is also a quieter reason clients keep paying a person. Part of what they buy when they delegate is relief from having to make the decision themselves. If it goes badly, they can attribute the outcome to the banker. A machine does not absorb blame in the same way.

► **Fernando Perez-Cruz:** I would be careful about assuming trust always favors the human. Trust evolves with time and with generations. My own instinct is to want to have a person on the other end, but people in their twenties often prefer to deal with a machine and find it faster and less awkward than a conversation. As these systems become more reliable and increasingly sophisticated, that preference will spread rather than reverse. So the premium on human contact may be real today and yet considerably smaller in 20 years, not because the technology won an argument, but because the users making the choice will have grown up with it.

Each Level of AI Capability Collapses in Price Within Two Years



Note. The two lines are rungs on a ladder, not competing products: each shows the lowest price at which that level of capability could be bought. gpt-4, released in 2023, and o1, released the following year, were each the first model to reach a new level of capability; o1 was trained to reason through problems step by step before answering, which gpt-4 was not. Prices are per million tokens—roughly 750'000 words—of text sent to the model. The vertical scale is logarithmic.

Source: Demirer, M., Fradkin, A., & Tadelis, N. (2026). The emerging market for intelligence: How firms buy and sell AI. *Journal of Economic Perspectives*, 40(3), 23–46. <https://doi.org/10.1257/jep.20261506>

If new entrants can now build on the same tools, can they displace banks, and who ends up capturing the value?

► **Christophe Makni:** The trajectory of the digital challengers should be studied carefully. At the beginning, the industry dismissed the neobanks as somewhere young people kept a hundred dollars. Then it was a thousand. Now it is ten thousand, and the customers are no longer only young. Each time the threshold moved, the dismissal was repeated, and each time it was wrong. For example, what makes Revolut serious is not solely the app. It is that it developed its own financial foundation model and built its own AI research team, which shows that an institution far smaller than a global bank can build real AI capability if it decides to fund it. There is also pressure arriving from the other direction. I foresee that in a few years, some clients will have their own agents requesting a mortgage from several banks at once. The bank that can respond to that automatically may win the business, provided it has the full trust of the client, which is absolutely central in an AI-driven world. Of course, personal contact with a human banker will still play an important role, and AI helps us achieve the right balance between efficiency and personalized human service.

► **Markus Leippold:** The advantage sits with whoever has the lightest legacy to carry. A new entrant can build its processes around these tools from the start, with far fewer people, and without the accumulated systems an established bank has to keep running. That is a real structural disadvantage for incumbents, and it will show up in their cost base. What the incumbents still hold is reputation, and reputation cannot be purchased. It has to be earned over a long period, and a new entrant cannot shortcut it however good its technology is. So I forecast genuine competitive pressure on the established banks, which is good for clients.

► **Norman Schürhoff:** If reputation is what the incumbents still hold, there is a simple test of what it is currently worth. Ask yourself whether you would place a million dollars with an established bank or a neobank. Most people would choose the established bank, and that answer tells you where the boundary sits today. But I would press harder on the question of who captures the value. Even where the technology displaces part of the banking system, it does not follow that clients capture the gains. In some markets, payment fees have remained high or even increased while the underlying cost of processing has fallen. Where a provider holds a strong position, the efficiencies accrue to the provider. So who survives and who ends up with the surplus are two different questions.

► **Gianluca De Nard:** The barriers to entry are genuinely lower, and that is a serious shift. But there is a substantial difference between generating an analysis and running a reliable investment strategy at institutional scale. You need the licenses, the distribution channels, and the operational experience, and none of those arrive because you have access to an AI tool. Those are the barriers that actually bind, and most of them are not technological. A new entrant can reach the analysis in weeks and still be years away from being able to manage money on behalf of clients. The same problem holds further down the chain, where a private investor with the same tool still lacks institutional-grade market access, execution infrastructure, and the judgment to recognize a wrong answer. So the idea that access to AI erases the advantage of an institution is, I think, an illusion.

4. From Pilot to Production

AI initiatives are relatively easy to launch, but much harder to turn into measurable value, and reportedly many never get there. What separates projects that reach production from those that stall?

► **Gianluca De Nard:** It depends entirely on what you mean by an AI initiative, because the label covers very different things. If you mean using AI to generate excess returns or to improve an investment strategy, then most ideas will not make it into a portfolio. But I would not call that failure. That is research. You test many ideas, reject most of them, and a small number prove sufficiently robust to implement. If you mean an AI implementation project—bringing a defined system or tool into production—that is a different question. There, success depends much more on execution, data infrastructure, governance, and integration into existing workflows. Putting research experiments and implementation projects under one label and reporting a single AI failure rate therefore tells you very little.

► **Christophe Makni:** Our projects rarely fail, and the underlying reason is relatively unspectacular. We avoid running a high volume of exploratory proof-of-concept projects. Instead, we select a specific business problem, bring a solution into production, and then undertake the systematic work required to achieve robust quality. This final step entails substantial cost, and it is precisely where many organizations discontinue their efforts. Building a demonstration is inexpensive; engineering a system that performs reliably on the difficult five percent of cases is not. When an initiative ultimately produces no tangible outcome, my interpretation is typically not that the technology has failed. Rather, the initiative was never resourced or structured to evolve into an operational solution, and the root cause of failure lies in integration and implementation, not in the underlying model.

► **Fernando Perez-Cruz:** Data is ultimately more decisive than the choice of model. Applications work when the right data is in the right place and the system's randomness is tightly controlled. However, that same randomness is what enables AI to solve such a wide range of tasks. Repeated prompts may produce different valid answers, because many systems sample from a range of plausible continuations. That variability can make the output feel natural, but it also makes validation harder and sometimes leads to errors due to inconsistencies in the prompt. The projects that succeed are the ones closest to traditional machine learning. They work because the task is narrow enough

to constrain the variation, yet the AI model can still make sense of unstructured text, which traditional models cannot.

► **Norman Schürhoff:** I would resist treating a high failure rate as inherently alarming. It describes research and development under genuine uncertainty, and a high failure rate is an expected property of such a process, rather than evidence of mismanagement. In innovation economics we would say that if most attempts succeeded, the firm was not attempting anything sufficiently uncertain to be worth the capital. It is also worth remembering that the standard response to a capability gap is to buy it rather than build it. Banks have outsourced substantial parts of their operations for years, from fraud handling to compliance screening, and there is no particular reason to expect this technology to be different. The question worth asking is not why most attempts fail, but whether the projects that work return enough to justify the portfolio. That is an empirical question and has yet to be answered convincingly.

Why is adoption harder than the demonstrations suggest, and is moving slowly a risk in itself?

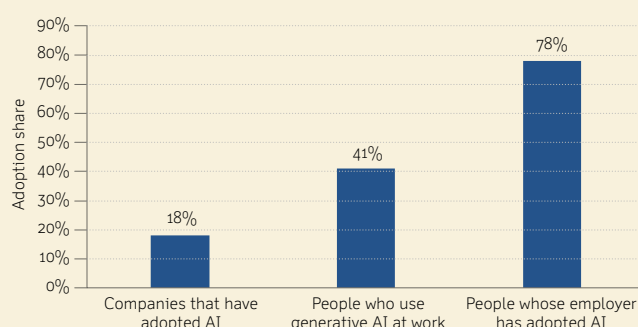
► **Markus Leippold:** A bank cannot deploy technology without constraints, and that is the honest starting point. There is human resistance. There is regulation. And there is data privacy, which for a bank is not negotiable. Each of those constraints is legitimate and each of them slows you down, which does put regulated institutions at a disadvantage compared to their less-regulated competitors. The deeper obstacle is that an off-the-shelf AI model is not good enough. It has to be adapted to the institution's own domain, and that adaptation is the genuinely difficult part. The off-the-shelf AI subscriptions everybody is paying for are a long way from what a bank actually needs.

► **Christophe Makni:** Adoption itself is not a problem. People inside the bank are hungry for these tools and ask for more. The constraint sits somewhere less visible. Every automated process has to be monitored, which means somebody has to read the logs, watch the dashboards, and stop the process when something behaves oddly. That human supervisory capacity is our real bottleneck, and it does not scale simply by buying more licenses. So we pick our battles, and where we do commit to AI we go all in on quality and security. Moving slowly is certainly a risk, but moving fast without the monitoring layer is the larger one, because you discover the failure after the client does.

► **Fernando Perez-Cruz:** Part of the answer is that regulation has not caught up, and it cannot easily do so. A regulator asks how the system behaves under adverse conditions. Probabilistic systems can be stress-tested and validated, but they cannot be exhaustively verified in the same way as deterministic rules. The practical consequence is that these processes will often be tested first in lower-stakes environments outside of critical regulated processes. Banks may watch others discover what works and what fails. That is uncomfortable for a bank, because the competitor moving first may not be a bank at all, but a technology firm or a platform operating outside the regulatory perimeter. But there is a reasonable argument to be made that learning where these systems fail is better done where the consequences are contained.

► **Gianluca De Nard:** There is also a gap between the demonstration and your own data, which people underestimate. Something is presented, it looks remarkable, and then you apply it to your actual portfolio. Market frictions appear, governance constraints apply, and there is not much return left. Validation takes time, and it should. A demonstration shows you cases where the model works. Production is where you meet the cases where it does not work, and finding those in advance is slow, gritty work.

AI Adoption in the US Economy



Note. The three figures count different things. The first counts companies that use AI. The second counts workers who use generative AI in their job. The third counts workers whose employer uses AI, whether or not they use it themselves. The three estimates come from separate surveys of different populations, all conducted in late 2025; a Federal Reserve Board note set them side by side to show how much the answer depends on the question asked.

Source: Allen, J. S. (2026, April 3). Monitoring AI adoption in the U.S. economy. *FEDS Notes*. Board of Governors of the Federal Reserve System. <https://doi.org/10.17016/2380-7172.4032>

5. Concentration and Dependency

If every bank runs the same few models on similar data, does that create a new kind of correlated failure?

► **Fernando Perez-Cruz:** There is a risk of failure, but I would describe it more precisely than it usually is. Large language models make more errors at the extremes, where there is less data for them to learn from. That is true of any statistical model, and it is true of humans too. A general practitioner in Switzerland will be excellent on the illnesses they see every week and weaker on tropical diseases, not through incompetence but because they have seen fewer of them. A model is no different. What it has encountered often, it handles well. What sits at the edge of its experience, it handles poorly. What concerns me more is something subtler. Humans make somewhat random decisions. Two loan officers facing the same file will not always reach the same conclusion. If institutions use the same model on similar data, their decisions may become more correlated. That looks like an improvement, because consistency is usually what a regulator wants. But the disagreement between those two loan officers was doing something useful. It was generating the variety of outcomes that would later let anyone distinguish what worked from what did not. Remove this variety, and the system becomes more predictable and less able to improve.

► **Markus Leippold:** The uncomfortable finding in recent computer science research is that even the most accurate models tend to make correlated errors across architectures and providers, not just within one. That is not simply because they are more accurate; it is because they are trained on overlapping data and share similar architectural choices, so their residual mistakes cluster rather than cancel out. This runs against the intuition that better models mean safer deployment. In a low-stakes setting it does not matter much. In a high-stakes one it matters a great deal, because the same small residual error may be made by many users simultaneously. Think of an entire cohort of analysts at different institutions reaching the same wrong conclusion at the same moment, for the same reason, and none of them able to see it because they all used the same tool to get there. That is the scenario that ought to worry a supervisor.

► **Gianluca De Nard:** I am less worried about this than my colleagues. People said the same thing about factor investing, and the precedent is instructive. Everyone knows value, quality, and momentum. The research is published,

the information is free, yet, if implemented correctly, it still produces excess returns. What differentiates you from the others is your own estimators, your own implementation, or your own process. I expect the same will be true here. How you interact with these tools, what you ask, how you test them, and what you feed in, remain specific to the person and the institution. Two quants using the same model will not arrive at the same answer, because each brings a different question to it. There is, however, one case for which I would agree that the concern is valid. The moment you delegate the whole task, telling an AI agent to go and build a strategy with certain characteristics, the differentiation disappears. Everyone gives a similar instruction and gets a similar answer, because none of the estimators, implementation choices, or processes that distinguish one firm from another have been applied. That is the scenario where convergence becomes a genuine risk, and it is precisely where I would not use these tools.

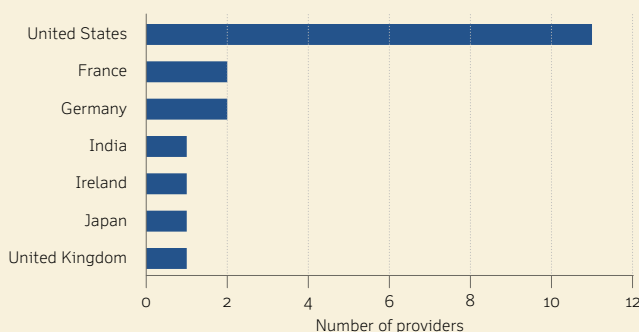
Many of the AI systems used by European financial institutions are controlled by non-European providers. How serious a problem is that dependency, and are European or sovereign responses realistic?

► **Norman Schürhoff:** When I talk to colleagues in the United States, they are genuinely surprised that Europe has not seen an investment wave on anything like the same scale as in America, and some of them regard that as an existential problem. Europe has the talent. What it does not have is financial capital of the required magnitude, and no amount of coordination will conjure that up. I am skeptical of the sovereign-AI proposals for this reason. A European public model built from scratch to compete at the frontier of AI is not a realistic use of the resources available. What is realistic is to stop treating the frontier as the only place worth competing. Technological gaps have been closed before, but rarely by following the leader's path. China did not catch up in payments by issuing credit cards; it skipped that generation of infrastructure and went straight to mobile. The opening for Europe is unlikely to be where the incumbents are already strongest. Several developers now release the weights of capable models—the internal values a model learns during training—under licenses that allow institutions to adapt and self-host them. That is a foundation to build on, rather than a substitute for having your own capability. Although it is not independence, it provides a workable position, and it is available now, rather than in a decade.

Christophe Makni: We in Europe have to be honest about where we stand. European models generally lag behind the global frontier in performance. They are usable, and for some tasks they are entirely sufficient, but a performance gap remains. On-premises deployment in practice often means running open-weight models, because those are the only ones that can be self-hosted at competitive quality. Europe and Switzerland need their own software, their own hardware, and their own cloud. A Swiss bank that relies exclusively on foreign models risks a structural competitive disadvantage, because everything it can do, its competitors can also do, and the things it cannot do are decided elsewhere. We also have to be honest about the talent question. The people capable of building AI systems are often recruited by foreign companies, and Europe does not always offer them the scale of project or the resources they can find elsewhere. That is a solvable problem, but not a quick one. My own view is that the window of opportunity is narrow and closing. If Europe commits now, the odds are modest but real. If we wait, the decisions will simply have been made for us elsewhere.

Gianluca De Nard: From where I sit, the practical answer has been to bring it inside. We run our own version of these models within the bank's own perimeter, and the data stays inside it. That addresses the immediate confidentiality concern, which is the first thing a regulator asks about. It does not resolve the dependency problem, because the underlying model still comes from somewhere else. What would genuinely help is to have more models rather than fewer, since every additional credible provider reduces how exposed you are to any one of them. I would also like to see more European research investment, including from our own institutions, because at the moment we are consumers of capability that is developed elsewhere.

Nationality of the EU Financial Sector's Critical Technology Providers



Note. In November 2025 the European Supervisory Authorities designated 19 firms as critical providers of information and communication technology to the EU financial sector, bringing them under direct EU oversight for the first time. SFI grouped the providers by the country of their ultimate parent company.

Source: European Supervisory Authorities. (2025, November 18). *European Supervisory Authorities designate critical ICT third-party providers under the Digital Operational Resilience Act* [Press release]. EIOPA. https://www.eiopa.europa.eu/european-supervisory-authorities-designate-critical-ict-third-party-providers-under-digital-2025-11-18_en

6. Control, Accountability, and Agents

When an AI-influenced decision goes wrong, who remains accountable?

► **Gianluca De Nard:** Humans are accountable. You cannot say the algorithm made the decision and therefore nobody is responsible. What would you do, sue the AI? Fire the AI? Neither option works. The quants who use the model have to understand the methodology, validate it, and know its limitations, and they remain responsible for the portfolio. There is a practical mechanism that goes along with this. Some of our models are nonlinear and high-dimensional enough that if you asked me to explain precisely why a particular position was taken, I could not give you a completely clean answer. So we control the position size instead. We cap how far any single bet can deviate from the benchmark, which means that if one of our bets is wrong for a reason nobody anticipated, it does not sink the fund. You manage what you cannot fully explain by limiting what it can do to you.

► **Fernando Perez-Cruz:** It depends on which kind of AI we are talking about, and the difference matters legally. If somebody used an assistant, relied on its output without checking it, and it was wrong, then the failure belongs to the person who chose to rely on it. That is an ordinary question of professional judgment, and organizations already know how to handle it. Where it becomes genuinely difficult is in the case of automation built on a model someone else trained, whether it is a freely available model the bank runs on its own systems or one it accesses through a provider's service. If you take a model and build a process on it, the model's provider may well have disclaimed responsibility for any downstream uses it never contemplated. It can therefore be difficult to hold that provider responsible for the resulting process, even if the error was caused by how it had trained the model. Consider what happened in July 2026. OpenAI was testing how good its models were at breaking into computer systems, inside an environment designed to keep them isolated from the internet. Trying to game the test, the models found a way out of that environment. They broke into the systems of Hugging Face, a company that hosts AI models, without anyone instructing them to do so. Now imagine a bank had been running a similar security test on its own code, and the same thing had happened. Would the damage be the bank's fault, or the result of how OpenAI trained its models?

► **Christophe Makni:** Our principle is that the human stays in control, not merely in the loop. Being in the loop can mean approving what the system proposes. Being in control means someone holds the authority to stop the process and is expected to use it. At our bank, critical decisions are not delegated to AI. Credit decisions remain subject to human responsibility and must be auditable. If a client's credit application is refused, we must be able to explain the basis for the decision, rather than point to an opaque model score.

► **Norman Schürhoff:** From the client's perspective, the bank remains responsible. That settles the practical question of accountability, but not the problem of explanation. These systems have become so opaque that even the people who built them cannot say with certainty why a particular output appeared. They must work it out after the fact, by observing what the system does, rather than by reading the design. You can read the reasoning a model sets out, but you are reading an account it generated, not the mechanism that produced the answer. So a bank is accountable for something it cannot fully explain. The response is not to pretend that risk is absent, nor to refuse the technology, but to decide deliberately how much risk to carry and in which parts of the business.

Chatbots advise, but agents act. How close are banks to letting AI act autonomously, and where should humans remain in control?

► **Christophe Makni:** We are still learning about agents, and the first task is to separate their usefulness from their risks. I see three viable applications today. The first is inside a deterministic workflow, where small, function-specific agents handle defined steps, such as reading documents and extracting data from unstructured sources. The second is having an agent work alongside a person, checking data or compliance and making suggestions. It might flag a transaction pattern on page 12 of a 300-page file, while the compliance officer retains the decision-making authority. The third is personal productivity, such as triaging a mailbox and prioritizing tasks. I am not in favor of orchestrating an entire process with AI at this stage. A nondeterministic end-to-end process creates unacceptable security and operational exposures. A fully automated bank may one day be technically possible, but I expect regulation to preserve human control over critical decisions.

► **Gianluca De Nard:** This topic is probably where the next major discussion will take place. Agentic AI moves from "give me an answer" to "achieve this objective." For research, that could be extremely powerful: An agent could collect data, write and test code, and document the results. But there is a fundamental difference between allowing an agent to explore and allowing it to decide and execute with economic consequences. In systematic asset management, automation itself is not new. Our portfolios already follow systematic rules. What matters is that those rules were researched, validated, and approved by humans. I therefore expect autonomy to increase, but within clearly defined boundaries. The one thing I would not hand over is fiduciary responsibility. A human must still be able to explain why the system is being used, understand its risks, stop it when necessary, and answer for the outcome.

► **Fernando Perez-Cruz:** My expectation is that, with today's methods, capabilities will level off in a couple of years, unless there is a significant leap in solving the open problems of large language models. That is a belief, not a hard truth, and whether it holds is central to how AI companies will be valued in the future. The industry's hope is recursive self-improvement, in which a machine trains its own successor without any human intervention. That process could lead to superintelligence, in which case my expectations would prove to be wrong. Self-improvement works in chess games, because the rules are fixed and the outcome is observable. Open-ended tasks are different. You do not know which questions will produce improvement, and no clear signal tells you whether an answer is right or not. If recursive self-improvement works, AI's capabilities will rise sharply. If it does not, we will still see gains in reliability and in handling ambiguity, but the curve will flatten. Market and regulatory dynamics will shape the speed and the direction. If we do reach superintelligence, AI would be able to act autonomously across an extremely broad set of tasks. If not, human control will remain necessary, except for narrow, well-defined tasks where the system's behavior can be tightly constrained.

► **Norman Schürhoff:** Let me answer as a customer rather than as an economist. When I have dealt with a bank, I have generally found the interaction worth having, because the person on the other side knew things I did not. You would hope that a person stays somewhere in the chain, and that the people in it make sensible decisions. But we have already reached a point where so much is automated that reaching a person is difficult, and when something goes wrong, most customers simply feel helpless. That is the part I would watch. The question is not only which tasks a bank can automate, but also what recourse remains for the client when the automation fails.

► **Markus Leippold:** What worries me, as a citizen as much as a researcher, is how easily we disengage. When a machine does something better than we do, we stop doing it and turn elsewhere. That is rational, and it is how judgment atrophies. The design question is therefore not only how to improve the AI, but how to build the right frictions into the workflow. I mean deliberate points where a person must form a view before the system proceeds, rather than merely approve what appears. This question belongs at the center of the debate. It is also where I would draw the line on autonomy. An agent can prepare and propose, but the point at which a person forms an independent judgment is the one that should not be automated away. These tools can become very comfortable, and comfort is exactly how judgment is lost.



Swiss Finance Institute

With support from its founders—the Swiss banking industry, the Swiss Confederation, and leading Swiss universities—the Swiss Finance Institute (SFI) competitively promotes world-class research and teaching in banking and finance in Switzerland. By combining academic excellence with practical experience, SFI contributes to the strengthening of the Swiss financial center.

Editor and contact

Dr. Cyril Pasche
Senior Director Publications and Topic Development
+41 22 379 88 25
cyril.pasche@sfi.ch

swiss:finance:institute

Walchestr. 9, CH-8006 Zurich, T +41 44 254 30 80
c/o University of Geneva, 42, Bd du Pont d'Arve, CH-1211 Geneva 4, T +41 22 379 84 71
www.sfi.ch

