ANALYSIS



Double-entry AI:
what banks earn from artificial intelligence and what it costs to answer for it

Article for AECA journal
November 2026

The falling cost of artificial intelligence is accelerating its adoption in banking, but there is a second, less visible bill: the cost of controlling it and being accountable for its decisions.

Drawing on the accounting principle of double-entry bookkeeping, this article examines both costs and proposes four metrics for measuring the returns on AI once the costs of oversight, validation and resilience are considered.

Artificial intelligence is paid for twice in a bank: once to produce an answer, once to be able to answer for it. The first cost is collapsing and already buys measurable gains. The second is the cost of control over models the bank does not own; it is rising and has no line in the accounts. This article starts from the Spanish origins of compulsory double-entry bookkeeping to explain why that second cost is missing from business cases and proposes four measures that let a board read the value of AI net of its control.

On 4 December 1549, at Cigales, near Valladolid, a royal pragmatic ordered the banks, public exchanges and merchants of the kingdom to keep their books in Castilian and “por debe y ha de haber”, by debit and credit. With that text and a second issued in Madrid in 1552, Spain became the first country in Europe to make double entry compulsory by law (Hernández Esteve, 2013). The Crown wanted to stop coin and precious metals leaving the kingdom, and a transaction written twice was one an inspector could reconstruct.

Banking has paid twice ever since: once to carry out a transaction, once to be able to prove it. Artificial intelligence falls under the same rule. Every answer a model gives a banker, or a customer has a cost of production and a second cost, which comes from the bank having to answer for it. Everyone watches the first, which is collapsing. The second has no line in the accounts, and it will decide the return.

1. The credit side: gains that already carry numbers

Spanish banks are further along than most. A study published by Funcas, on 12 November 2026, finds their adoption of AI above the European average and puts the technology investment of the ten largest at EUR 5.6 billion in 2025, equal to 8.7% of their operating expenses.

Such figures need care before they enter a business case. Hours released become a financial benefit only when they generate additional output, avoid an expense or reduce a loss.

Volume is no measure of productivity either. The best-known randomized trial, run by METR in 2025, found experienced developers 19% slower with AI tools while believing themselves faster; in February 2026, its authors wrote that the tools probably helped developers more than a year earlier and that their experiment no longer gave a reliable signal.

A stage that runs faster can also pass its errors and its workload to the next, so the only unit that holds is the end-to-end process, with its exceptions and its controls. The ECB supervisors reached the same point after workshops with thirteen banks: “the financial quantification of realized benefits remains a challenge” (ECB Banking Supervision, 2025).

One gain shows up in capital. In a simulation on an anonymized loan portfolio from a large Spanish bank, the Banco de España researchers found that a machine-learning model could lower regulatory capital requirements by 12.4% to 17% under the internal ratings-based approach, compared with a simpler statistical model (Alonso-Robisco and Carbó, 2022). The saving exists only if the model passes supervisory validation, which is second-entry work, and the Basel output floor would cap it.

2. A falling price, a rising bill

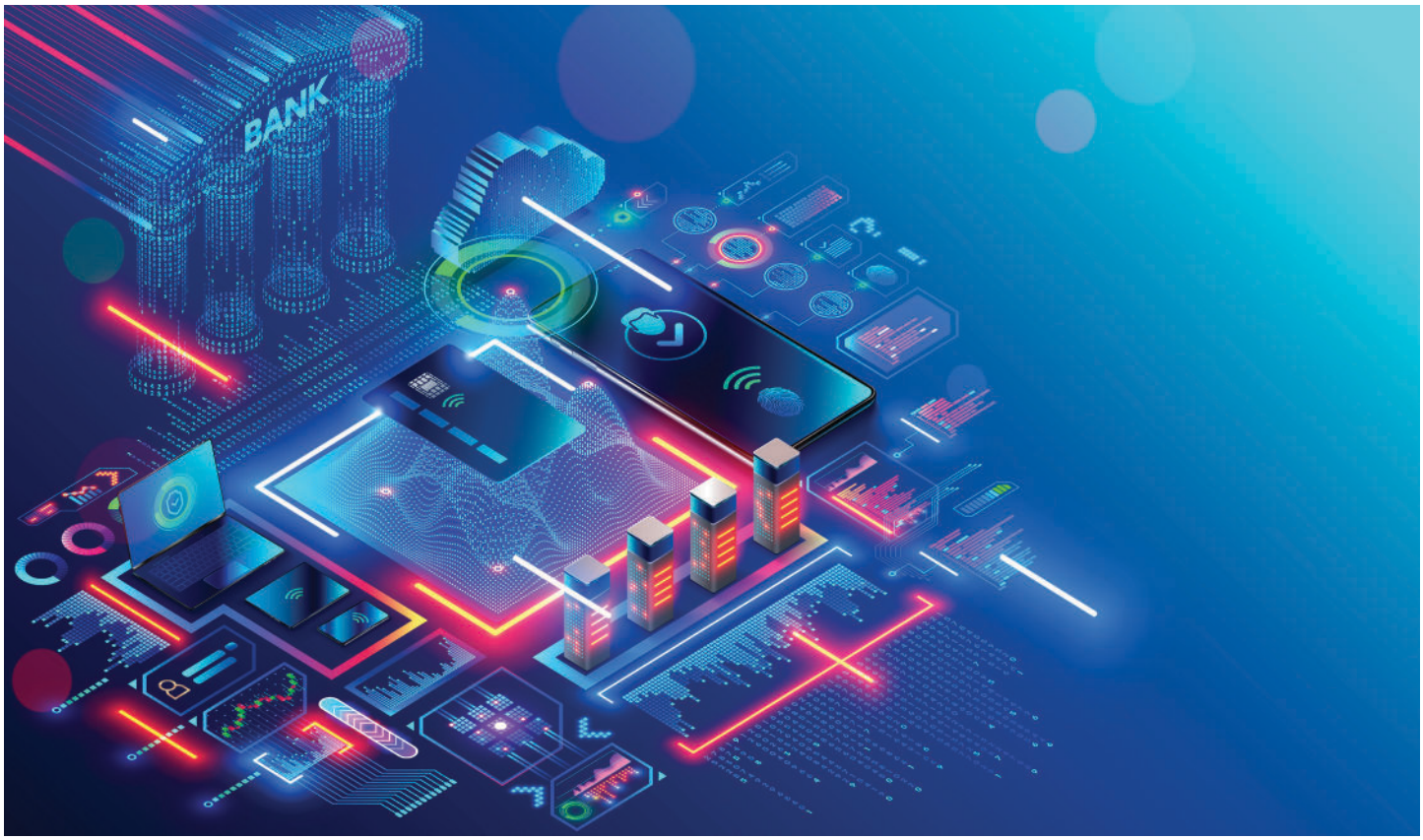
Few inputs a bank buys have ever fallen in price this fast. The Stanford AI Index (2025) puts the cost of GPT-3.5-level performance at USD 20 per million tokens in November 2022 and USD 0.07 in October 2024.

The bill rises all the same. An agent that plans, calls tools, checks its own work and starts again consumes many times the tokens of a single question, and a cheaper resource is one that gets used more. So, the unit that matters for a budget is the completed task, with its retries and its checks.

The choice between buying inference and building capacity follows the same logic. Hosting a model internally is an option to be justified against its full cost, which includes hardware, energy, engineers and validation. European hosting does not end the dependence: Mistral trained Large 4, the model it unveiled in October 2026, on 3,800 Nvidia GPUs.

For the finance function, the choice leaves an accounting signature. Inference bought as a service is an operating expense, recognized as it is consumed. Since the IFRS Interpretations Committee's 2021 agenda decision on cloud arrangements, the cost of configuring such a service is often expensed too and only code the bank itself controls can be capitalized (IAS 38.13–17).

What the bank builds can become an asset, whether its own capacity or the software written around a rented model. IAS 38 amortizes that software over its useful life, and EU prudential rules, which deducted it in full of capital until 2020, now recognize it over the shorter of that life and three years (Commission Delegated Regulation (EU) 2020/2176; EBA Question ID 2020_5171; and Article 36(1)(b) of Regulation (EU) 2024/1623, CRR3, which neither amended nor repealed the prudential amortization method for software).



3. The debit side: what it costs to answer for a model

Answering for a model starts with knowing what it does. Neither the place where a model was trained nor the place where it runs settles the matter. A European server changes where the data sits and leaves the model's decisions as they were; Chinese, American or European origin proves nothing about fairness in each portfolio. At a bank's scale a small deviation becomes material, and when several institutions rent the same model, the error is shared.

The bank then must prove what it knows. From 2 December 2027, following the Digital Omnibus adopted in June 2026, the high-risk requirements of the AI Act apply to systems that assess the creditworthiness of natural persons or establish their credit score (Annex III, point 5(b)).

The provider answers for the design of the system, including data quality, documentation, accuracy and cybersecurity. The bank that deploys it keeps obligations of its own: competent human oversight, monitoring, retention of logs and, for credit scoring, a fundamental rights impact assessment before deployment. Buying technology does not transfer accountability.

A bank also must be able to do without the model. On 12 June 2026, three days after their release, a US export-control directive barred any foreign national from Anthropic's two newest models; the company disabled them for all customers, and general access to the first resumed only on 1 July. Three days is too short for a bank process to depend on a model. The episode mainly showed how long access to a frontier model can be suspended by a decision taken outside Europe.

DORA already requires exit strategies for critical functions, and on 7 July 2026 the ECB asked the chief executives of significant institutions for action plans against AI-enabled cyber threats by 31 October. The first list of critical ICT providers designated under DORA, in November 2025, includes the large cloud providers; OpenAI, Anthropic and Mistral are not on it.

Agents add a question of authority. The ability to execute an action does not confer the right to decide it.

4. Why the second entry is missing from the books

These costs are already in the accounts, scattered. Validation sits in Risk, monitoring in Technology, human review in the business lines, evidence in Compliance, the second supplier in Procurement and assurance in Internal Audit. No chart of accounts has a line called control of AI, so business cases are written gross, with the value on one side and the price of tokens on the other.

Markets are not waiting for the net figure. The ECB analysis published on 12 November 2026 finds that AI-exposed firms in the STOXX Europe 600 earn an equity return premium of about 1.3 percentage points a quarter and pay about 5 basis points less on their bonds (ECB, 2026). The narrative is priced before the cash flows can be audited.

A ledger kept use case by use case closes the gap, and four measures are enough to keep it. The first is the cost per completed task: inference, retries, human review and rework, divided by the tasks that reached their end. It replaces the cost per token. The second is the control ratio, the euros spent on testing, monitoring, review, evidence and redundancy for each euro of inference. It rises as inference gets cheaper, a sign that the cost has moved from the model to its supervision.

The third measure is switching time, the number of days needed to move a critical use to another model, tested once a year. It prices the option whose value the June 2026 episode made visible. The fourth is retained value: verified benefits net of both costs, measured against a baseline fixed before deployment, with the criteria for continuing or withdrawing set in advance.

Financial discipline does the rest. A benefit counted in the cash flows is not counted again in the discount rate, and control is recorded as what it is: an expense today whose return takes the form of lower expected losses and continuity of service.

5. Scale without merger

The cost of control is largely fixed. A test suite, a monitoring platform or a second supplier costs a mid-sized bank much the same as a large one. That favours scale, and it has revived the question of consolidation. Accessing scale and owning it are, however, different decisions. An acquisition is justified only by a demonstrable improvement in risk-adjusted profitability after integration costs.

Spanish banking has its own precedent for the first route. Bizum, built in common by the sector, has 32 million users ten years after its launch.

What the sector did for payments can be done for the testing of models that its members rent from the same suppliers. Test sets for bias could be shared, and so could the evidence prepared for supervisors and the negotiation of exit terms. The route has a limit, since a shared platform is itself a common dependency.

6. Where to start measuring

Technology development is the obvious place to start, because its output is already counted: time to production, defects, maintainability, with the review effort included. For asset management, ESMA accepts AI in advice and portfolio management while keeping management bodies accountable and MiFID II fully applicable. Operational efficiency and the suitability of recommendations are the items to record there.

7. The entry opposite

In 1549, what made a banker's books worth reading was the entry opposite, the one that let a third party reconstruct the transaction. A board that reads both entries knows what the bank retains once its AI has been governed. For now, the first entry is nearly free. The second is still to be accounted for.

REFERENCES

1. Alonso-Robisco, A., & Carbó, J. M. (2022). Can machine learning models save capital for banks? Evidence from a Spanish credit portfolio. International Review of Financial Analysis, 84, 102372.
2. Bowen, D. E., Price, S. M., Stein, L. C. D., & Yang, K. (2024). Measuring and mitigating racial disparities in large language model mortgage underwriting. SSRN Working Paper 4812158.
3. Commission Delegated Regulation (EU) 2020/2176 of 12 November 2020 as regards the deduction of software assets from Common Equity Tier 1 items.
4. Regulation (EU) 2024/1623 of the European Parliament and of the Council of 31 May 2024 amending Regulation (EU) No 575/2013 as regards requirements for credit risk, credit valuation adjustment risk, operational risk, market risk and the output floor.
5. ECB Banking Supervision (2025). AI's impact on banking: use cases for credit scoring and fraud detection. Supervision Newsletter, 20 November.
6. European Central Bank (2026). Artificial intelligence and European firms: the cost of market-based financing. Economic Bulletin, Issue 7/2026, box.
7. Funcas (2026). Spanish banks could exceed €8 billion in technology investment by 2031 as part of their commitment to artificial intelligence. Press release, 12 November.
8. Hernández Esteve, E. (2013). An approach to the study of Spanish accounting thought. Madrid: AECA.
9. Maslej, N., et al. (2025). The AI Index 2025 Annual Report. Institute for Human- Centered AI, Stanford University.
10. Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonized rules on artificial intelligence (AI Act), as amended by Regulation (EU) 2026/1744.

HIGHLIGHTED QUOTES

- "Artificial intelligence is paid for twice in a bank: once to produce an answer, once to be able to answer for it."
- "An European server changes where the data sits and leaves the model's decisions as they were."
- "No chart of accounts has a line called control of AI, so business cases are written gross."
- "Accessing scale and owning it are different decisions."
- "The first accounting entry is nearly free. The second is yet to be accounted for."

Contacts



Jean Barrère
Partner
jean.barrere@accuracy.com
+33 1 58 75 70 48



Philippe Wüst
Partner
philippe.wuest@accuracy.com
+33 1 58 75 72 34



Carla Azorí
Senior Regulatory Affairs Manager
carla.azori@accuracy.com
+34 91 406 7326

Accuracy is a wholly independent international consulting firm providing advice to company management and shareholders for their strategic or critical decisions, notably in transactions, disputes and crises.

Our strength is to connect strategy, facts and figures. Accuracy's teams are international and multicultural, combining various skills to provide bespoke services to our clients. We recruit consultants from the best.

Accuracy is present in 14 countries in Europe, North America, the Middle East and Africa and leads engagements all over the world.



Accuracy partners and professionals are available to discuss your needs and design an appropriate solution with the relevant experts.