



AI Is an Engine

Why collections AI only works when a
behavioural philosophy steers it



WHITE PAPER



Executive Summary

The collections industry is adopting artificial intelligence faster than almost any other back-office function: 93% of collections companies now use or plan to use AI and machine learning, up from 49% only two years earlier.¹ The results have not followed the spending. Across industries, only 37% of organizations can attribute any earnings impact to AI, a figure flat year over year; 42% of companies abandoned most of their AI initiatives in 2025, up from 17% the year before.^{2,3}

This paper argues the gap has a specific cause. AI is an engine. It amplifies whatever operating philosophy it is pointed at, and most deployments are being pointed at a collections playbook that was already incomplete — built on lagging indicators, score-band segmentation, and a pressure model that misreads why people fall behind (the argument of Paper 1 in this series). An engine given a bad map arrives at the wrong destination faster.

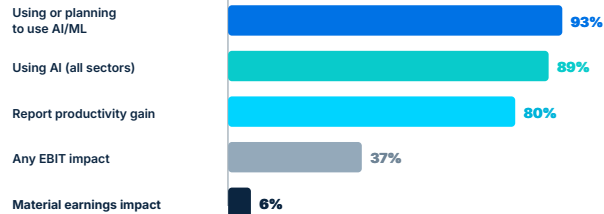
Getting it right requires three commitments. First, respect the line between deterministic and stochastic systems: credit and treatment decisions belong to explainable, auditable machine-learning models; large language models (LLMs) belong in communication — heavily tested — and in accelerating analysis and configuration. Second, buy outcomes, never deflection: the key performance indicator is cure and roll rate, never cost per contact. Third, welcome the rising barriers around data. They are rising for good reason, and they favour operators who build governance in rather than bolt it on. Done this way, AI in collections is effective, cost-efficient, and compliant at the same time.

AI is an engine. It amplifies whatever operating philosophy it is given.

01 The stampede, and the 95% problem

The adoption numbers describe a stampede. Worldwide AI spending is forecast to reach \$2.6 trillion in 2026, up 47% in a year.⁴ In collections specifically, TransUnion's industry survey found 93% of companies using or planning AI and machine learning — with voice communication at 44% and negotiation support at 42% — while 98% of firms now offer at least one self-service channel.¹ Every top-ten US bank ran a customer-facing chatbot at the CFPB's last count.⁵

Adoption Is Not the Constraint



TransUnion, 7th annual Debt Collection Industry Report, February 2026; McKinsey & Company, *The State of AI*, 2026 wave (n=1,719); S&P Global Market Intelligence, 451 Research, March 2025.

The outcome numbers describe something else. MIT researchers reported in 2025 that 95% of enterprise generative-AI pilots produced no measurable profit-and-loss impact — a finding whose methodology has been fairly contested, but whose direction matches independent data.⁶ S&P Global found the share of companies abandoning most of their AI initiatives leapt from 17% to 42% in a single year.³ Gartner predicts more than 40% of agentic-AI projects will be cancelled by the end of 2027, and its research on the current generation of self-service tools found only 14% of customer-service issues are fully resolved in self-service.^{7, 8} McKinsey's 2026 global survey completes the picture: 89% of organizations use AI, 80% of people report personal productivity gains, and still only 6% of companies achieve material earnings impact.²

Productivity without profit is the signature of a tool being used on the wrong problem. The collections industry should study that signature closely, because it is currently writing the largest technology cheques in its history.

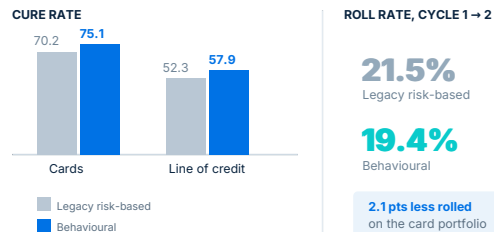
What this means for issuers: adoption rate is not a strategy. Before asking "which AI?", ask "pointed at what?" — the question the next section answers.

02 Automating the broken playbook

The most common failure mode in collections AI is precise: institutions are automating the playbook they already had. The same score-band waterfalls, the same escalating pressure sequences, the same contact-rate metrics — now executed by machines, at scale, around the clock. To the extent existing practice rests on incomplete assumptions about why customers fall behind, AI applied to it inherits the incompleteness and multiplies its output.

Paper 1 laid out the evidence: 94% of mortgage defaults trace to cash-flow shocks and life events, and 70% of all defaults to cash flow alone; behavioural distress signals appear one to two years before delinquency; identical credit scores conceal opposite treatment responses.⁹ A pressure-based playbook misreads a customer whose problem is cash flow, not willingness. An AI-accelerated pressure playbook misreads them thousands of times per hour, while reporting excellent efficiency metrics to its owners.

Behavioural vs. Legacy Segmentation



Symend program data, a major North American bank, May 2025 placements. Control (~10% of population) ran legacy risk-based segmentation; treatment ran behavioural segmentation from a supervised machine-learning model — not a generative system. Results validated as statistically significant. Subsequent-month figures excluded as non-mature at time of analysis.

The KPI regime does the damage. When conversational AI is bought as a labour-cost story — and the marquee forecast is precisely that: Gartner's 2022 call of \$80 billion in contact-centre labour savings by 2026, still quoted, never trued up — the system optimizes for deflection: conversations contained, calls avoided, headcount removed.¹⁰ Klarna became the emblem. Its AI assistant, Klarna boasted, did "the work of 700 agents"; within eighteen months the company conceded the cost focus had produced lower-quality service and began rehiring humans — while keeping AI on two-thirds of its inquiries.¹¹ The lesson is a rebalancing, and it is instructive: the failure was never the engine. The failure was the objective function handed to it.

What makes this genuinely painful is that collections is one of the few domains where the consumer evidence favours well-designed AI. Consumers say they dislike AI in customer service generally: 64% would rather companies not use it, and half say they would consider switching over it.⁸ Debt appears to be the exception. A 2026 study across eleven European countries — a preprint, and so far the only large test — found people report roughly half the stigma when a debt conversation is with an AI agent rather than a human, 11% versus 19%, with trust essentially equal; humans still rate higher on empathy and fairness.¹² For a customer who is behind on a bill and ashamed of it, a capable, judgment-free digital counterpart is often the preferred one. The dignity dividend is real and measurable. It is squandered when the AI on the other end is running a dunning script.

What this means for issuers: audit what your AI is optimizing for. If the answer is deflection, containment, or cost per contact, you have automated the broken playbook. The fix is to hand the engine outcome objectives — cure rates, roll rates, customer lifetime value — and a behavioural model of the customer worth optimizing toward.

03 The deterministic line

The architecture question is where enthusiasm most needs discipline. Our position, stated plainly: credit rules and credit decisions must remain deterministic — explainable machine-learning models with audit trails. Automated communication, in voice or text, can use LLMs, provided it is heavily tested inside guardrails. LLMs are also superb accelerants for data analysis and platform configuration. What they must not do, yet, is decide who gets credit, forbearance, or a settlement.

Where the Line Sits

DETERMINISTIC

Explainable · audited · fixed

- Eligibility
- Treatment boundaries
- Payment terms
- Disclosure content
- Anything moving money
- Anything creating a record

STOCHASTIC

Tested · sandboxed · bounded

- Reading the situation
- Choosing phrasing
- Conducting conversation
- Comprehension at scale
- Data analysis
- Platform configuration

Magesh et al., *Journal of Empirical Legal Studies* (Stanford RegLab), 2025; Equal Credit Opportunity Act and Regulation B, 12 CFR 1002.9; Zaharia et al., Berkeley AI Research, 2024.

The reliability evidence sets the line. Stanford's peer-reviewed audit of purpose-built, retrieval-grounded legal AI tools — systems engineered for accuracy in a regulated domain — found they still hallucinated 17% to 33% of the time; general-purpose LLMs erred on 58% to 82% of legal queries.¹³ Language models are stochastic instruments. In conversation, inside tested bounds, that flexibility is the feature: comprehension of a customer's actual words, negotiation within pre-approved limits, warmth at scale. In a credit decision, it is a liability no institution should accept — and one the law does not accommodate. The Equal Credit Opportunity Act and Regulation B require creditors to give specific, accurate reasons for adverse action, whatever the sophistication of the model; a decision that cannot be explained cannot be lawfully defended.¹⁴

Note carefully what regulation is and is not doing here, because the ground moved in the past eighteen months. The US window is deregulatory: the Consumer Financial Protection Bureau withdrew its AI adverse-action circulars in May 2025; the banking agencies rescinded the fifteen-year-old model-risk guidance SR 11-7 in April 2026 and explicitly carved generative and agentic AI out of its replacement; enforcement of Colorado's AI act was stayed before it ever took effect; Europe deferred its high-risk credit obligations to late 2027.¹⁵ No AI-specific rule currently forces discipline on a collections program; the durable constraints are the old statutes — fair-lending, debt-collection, and telephone-consent law — and the state attorneys general and plaintiffs' bar that enforce them. The AI-specific discipline, in other words, is now a choice. The institutions that choose it are building compliance into architecture — deterministic rails around stochastic capability — because that is sound engineering under statutes that never expire, not because a circular told them to.

The practical division of labour looks like this. Deterministic systems own eligibility, treatment boundaries, payment terms, disclosure content, and every action that moves money or creates a legal record. Stochastic systems own understanding and expression: reading a customer's situation from her words, choosing phrasing, conducting a conversation — always inside the sandbox the deterministic layer defines, and always tested against thousands of adversarial cases before a single customer hears it. The AI research community has converged on the same conclusion from the engineering side: state-of-the-art reliability comes from compound systems that wrap models in retrieval, verification, and deterministic control logic, and never from a monolithic model handed the keys.¹⁶

What this means for issuers: ask any AI vendor two questions. Where, exactly, is the deterministic line in your architecture? And how many adversarial test cases did this conversation flow survive before production? Vendors who answer with model names have answered both questions.

04 The barriers are rising. Good.

The final trend runs opposite to the industry's instincts: access to consumer data is getting harder, and this is healthy.

Banks are gating data more tightly precisely because AI raises the stakes of misuse. Bank regulators' third-party risk guidance puts vendor AI squarely inside full-lifecycle risk management; the US Treasury has flagged third-party AI dependence and data provenance as system-level concerns; the OCC's 2026 risk perspective names data privacy and explainability among the central vulnerabilities of advanced AI in banking; and in surveys, data leakage now tops bank executives' list of AI worries.^{17, 18, 19} Every one of these pressures raises the cost of the move-fast, data-hungry integration model.

Three consequences follow, each favourable to institutions and vendors that take governance seriously.

The trust premium. When data is the exposure, trust becomes the procurement criterion. Advantage shifts from the flashiest demo to the vendor with governance credentials, purpose-built models, and a track record inside regulated environments — a moat that compounds, because trust is earned on a clock that cannot be accelerated by funding rounds.

Compliance by architecture. The only sustainable way to run stochastic systems in a regulated industry is architectural: deterministic guardrails, complete audit trails, explainable decision paths, tested boundaries. Barriers force good engineering.

Data gravity. As raw data becomes harder to move, value shifts to whoever can bring intelligence to the data — embedded, governed platforms operating inside the institution's perimeter and its policies, rather than point tools that demand the data come out to them. MIT's adoption research — the same contested survey — found externally partnered deployments succeeding at roughly twice the rate of internal builds, though McKinsey records a countercurrent: 32% of organizations now build in-house because agentic tools let them.^{6, 2} The embedded-platform case has to be earned deployment by deployment.

Together with the architecture discipline of Section 3, this is how the three goals that are usually traded off against each other — effectiveness, cost-efficiency, and compliance — stop being trade-offs. A behaviourally steered engine cures more customers (effective), automates the communication layer where automation belongs (cost-efficient), and runs on deterministic rails with audit trails (compliant). Institutions that hit all three will set the benchmarks the rest of the industry spends the next decade chasing.

The macro backdrop makes the timing urgent rather than optional. Paper 1 documented a concentrated, structural stress building beneath calm aggregates — a cohort of customers whose distress is visible in behaviour long before it reaches the loss line. The institutions that meet that cohort with instrumented early warning, behavioural segmentation, and a well-steered engine will resolve it. The institutions that meet it with an automated dunning script will discover what the engine does to a bad map.

Endnotes

1. TransUnion, 7th annual Debt Collection Industry Report, published February 2026, via ACA International. "Using or planning to use" AI/ML: 93% (2025) vs. 49% (2023); voice 44%; negotiation support 42%; self-service 98%.

2. McKinsey & Company, The State of AI (2026 wave; surveyed May–June 2026, n=1,719). 89% using AI; 37% attributing any EBIT impact (flat vs. 2025); 6% high performers; 80% reporting individual productivity gains. Same survey: 32% of organizations declined to buy software because agentic tools enabled in-house builds.

3. S&P Global Market Intelligence (451 Research), March 2025 survey, n>1,000, North America and Europe.

4. Gartner, worldwide AI spending forecast, May 2026. Over 45% is infrastructure rather than applications.

5. Consumer Financial Protection Bureau, Chatbots in Consumer Finance, June 2023.

6. MIT NANDA, The GenAI Divide: State of AI in Business 2025, August 2025. Methodology contested (failure defined as absence of rapid P&L impact); cited here with that caveat. Partnership-vs-internal success rates from the same report.

7. Gartner, June 2025 prediction on agentic-AI project cancellations by end-2027.

8. Gartner, August 2024: 14% of customer-service issues fully resolved in self-service; July 2024: 64% of customers prefer companies not use AI for customer service, and 53% would consider switching to a competitor over it. General customer service, not collections-specific.

9. Ganong & Noel, Quarterly Journal of Economics 138(2), 2023 (94% cash-flow defaults); Bank of Canada Staff Analytical Paper 2026-3 (distress behaviour ~2 years before delinquency); Federal Reserve FEDS Notes, January 2024 (score migration). Full sourcing in Paper 1.

10. Gartner, August 2022 forecast: conversational AI to reduce contact-centre agent labour costs by \$80 billion in 2026.

11. Customer Experience Dive and Forbes coverage of Klarna, 2024–May 2025. Klarna retained AI on roughly two-thirds of inquiries.

12. Academic preprint (arXiv 2602.00050), published February 2026; eleven European countries, ages 18–70. Stigma probability 11% (AI contact) vs. 19% (human); trust 84% vs. 85%. Preprint — peer-review status pending.

13. Magesh, Surani, Dahl, Suzgun, Manning & Ho, Journal of Empirical Legal Studies (Stanford RegLab), 2025. Legal domain; best available proxy for LLM reliability on regulated text.

14. Equal Credit Opportunity Act; Regulation B, 12 CFR 1002.9. The CFPB's interpretive circulars on AI adverse action (2022-03, 2023-03) were withdrawn May 2025; the statutory and regulatory requirements stand.

15. CFPB withdrawn-guidance list and Federal Register, May 2025; OCC/Federal Reserve/FDIC, Revised Guidance on Model Risk Management, April 17, 2026 (generative and agentic AI excluded from scope); Colorado SB 24-205 enforcement stayed April 2026; a narrower replacement (SB 26-189, ADMT transparency framework) signed 2026-05-14, effective 2027-01-01; EU Digital Omnibus (Regulation (EU) 2026/1744, in force July 2026) deferring Annex III high-risk obligations (including credit scoring) to 2027-12-02. Status as of 2026-08-31. Regulation F (12 CFR Part 1006) remains in effect and unchanged.

16. Zaharia et al., "The Shift from Models to Compound AI Systems," Berkeley Artificial Intelligence Research, February 2024.

17. Interagency Guidance on Third-Party Relationships: Risk Management (Federal Reserve/OCC/FDIC), June 2023.

18. US Department of the Treasury, Artificial Intelligence in Financial Services (post-RFI report), December 2024.

19. OCC, Semiannual Risk Perspective, May 2026; Wolf & Company 2026 banking AI survey (small sample; directional).



About Symend

Symend is a **collection strategy platform** that has cured over 250 million delinquencies and recovered over \$50 billion in payments.

+10%

Recovery rate
improvement

−50%

Operational cost
reduction

10x

Return on
investment

Meet with our team for a tailored demo

symend.com

THE SYMEND PLATFORM

SymendCure

The debt recovery platform that combines behavioral science with AI to deliver deeply personalized engagement journeys — improving recovery rates while strengthening customer relationships.

SymendConverse

The same collection strategy extended to voice. Behavioral science-powered conversations that read each customer's situation and adapt in real time, so routine resolutions complete without a human and agents reach the accounts where they change the outcome.

SymendPrevent

Bill protection that keeps cured customers cured. Covers a customer's bill for up to six months in the event of job loss, disability, critical illness or death — protecting the relationship, not just the balance.