FIELD BRIEFING

AI value traceability, from pilot to P&L

How a board traces a claimed AI benefit from demonstration to a defensible line in the accounts, and what it authorises at each step.

3 AUGUST 2026 · V1.0.0 · PUBLIC

Decision mandate

This briefing serves one decision. An AI deployment has been demonstrated somewhere in the group, a benefit has been claimed for it, and the board is asked to authorise the money and to carry the claimed benefit into a plan. The accountable executive, normally the chief financial officer with the sponsoring business head, must decide what the claim is allowed to become: a capability recorded and nothing more, a funded measurement programme, a deployment with a benefit recognised in a named profit-and-loss line, or a refusal.

The decision horizon is the budget cycle in which the claim would land, typically the next twelve months, with a three-year run-rate attached to it whether or not anyone has costed one. The value at stake is rarely the licence. It is the plan line built on top of the licence: a headcount assumption that will be used to justify not hiring, a margin promise made to a lender or a shareholder, a cost-reduction target cascaded to managers who cannot deliver it. A licence written off is an expense. A benefit booked into a plan and never realised is a credibility event that surfaces two budget cycles later, when the people who authorised it have moved on and the baseline it was measured against no longer exists.

The reader should leave able to decide differently in one specific way: to refuse to carry any AI benefit into a plan unless it can be traced, link by link, from the claim to a named line with a named owner and a named baseline, and to treat a break in that chain as a stop condition rather than a documentation gap to be tidied later.

The alternatives, stated plainly, with their reversibility. Recording the demonstration and doing nothing else is fully reversible and costs a quarter. Funding a measurement programme without a deployment commitment is nearly reversible: the instrumentation survives the tool and is reusable against the next claim. Deploying with a recognised benefit is the least reversible of the three, not because the software is hard to remove but because the recognised benefit has already been spent, in a headcount plan, a covenant conversation or a target. Refusal is reversible at the cost of a renegotiation. Reversibility, not conviction, is why the sequence in this briefing puts measurement before authorisation and authorisation before recognition.

Two companion notes state the firm's positions that stand behind this instrument: that a productivity claim is only real once it survives translation into a profit-and-loss line, and that the pilot-to-production gap is a governance artifact because pilots are scoped to succeed. Those notes argue. This briefing instruments. It does not restate them.

This is a field briefing, not a technology assessment. It assumes a reader who already operates in the corridor, who has a vendor quotation in front of them, and who needs a chain, thresholds and owners rather than a survey of what AI can do.

EVIDENCE

Official statistics count AI adoption and do not measure realised value. Eurostat records 20.0 per cent of EU enterprises with ten or more employees using AI technologies in 2025, up 6.5 percentage points from 13.5 per cent in 2024, with national shares from 42.0 per cent in Denmark, 37.8 per cent in Finland and 35.0 per cent in Sweden down to 5.2 per cent in Romania and 8.4 per cent in Poland. In the United States, Census Bureau estimates held between 17 and 20 per cent of businesses from December 2025 to May 2026, standing at 19.8 per cent nationally on 3 May 2026, with 39.7 per cent in Information and 33.9 per cent in Finance and Insurance.

Sources: Eurostat, ICT usage and e-commerce in enterprises, 2025 release published 11 December 2025; US Census Bureau, Business Trends and Outlook Survey, releases to May 2026. Geography: European Union; United States. Method: official enterprise surveys, self-reported use. Grade SEG-5. Caveat: adoption counts only, with no outcome or financial measure in either series; Eurostat excludes enterprises below ten employees; the US survey uses a two-week reference period and is an experimental product; the two series are not directly comparable and are cited side by side rather than combined.

The conventional assumption

The conventional assumption, held in most operating companies and reinforced by most vendor engagement models, is that the benefit case is a procurement artifact. On this view the work is to obtain a credible-looking business case before signature, approve it, deploy, and then look for the savings. Benefit tracking is a reporting exercise scheduled after go-live, staffed by whoever has capacity, and measured against whatever numbers can be assembled at the time. The claim is treated as an input to a purchase decision, and once the purchase is made the claim's job is done.

This briefing takes the opposite position, and states it as the house position, an interpretation from operating practice rather than a measured finding (grade SEG-1, marked as such deliberately): the benefit case is not an input to the decision, it is the decision, and it is a chain rather than a number. A claimed benefit becomes a defensible profit-and-loss line only by passing six sequential links, each with a required artifact, a test, a numeric decision rule, a named owner, a fixed timing and a stated consequence of failure. Claim registration. Measured effect in a defined population. Attribution to the intervention. Transfer to this firm's conditions. Recognition in a named line against a named baseline with a named owner. Verification after the fact. A break in any link stops the chain at that link. The chain does not average: a strong measurement does not compensate for an absent counterfactual, and an excellent counterfactual does not survive a population that does not resemble yours.

The consequence for the board is procedural and immediate. The question changes from "is this a good business case" to "which link are we at, and what is the artifact". A board that asks the second question cannot be handed a plausible number; it has to be handed a document. Most claims fail at links two and three, before anyone reaches the arithmetic, and they fail visibly.

The corridor makes this discipline harder to skip and more expensive to skip, for reasons set out in the next movement.

The corridor structural difference

Five structural facts distinguish the traceability problem in the WAEMU-euro corridor from the same problem in the markets where the vendor's business case was assembled. Each one attacks a different link in the chain.

First, the licence is a hard currency claim on a soft currency benefit, and the peg does not fix this. The CFA franc has traded at 655.957 to the euro since 1999, so a euro-priced licence carries no exchange-rate volatility for a corridor buyer. That is the beginning of the analysis and not the end of it. The licence is a fixed nominal claim denominated in the currency of the vendor's cost base, while the benefit is a variable saving denominated in the buyer's payroll. The two are not the same size relative to their local economies, and no exchange-rate movement will bring them closer, because there is no exchange-rate movement. Where the licence is priced in dollars, volatility returns in full and the fixed claim becomes a floating one. The instrument response is to convert every licence and integration line into hours of the firm's own labour at the firm's own loaded rate before the decision, and to state the licence as a headcount equivalent. A tool that costs five analysts to release three is legible in that unit and illegible in any other.

Second, connectivity and power are inputs to the benefit, not background conditions. Internet use in Africa averaged about 36 per cent of the population in 2025 against a global average of 74 per cent, falling to about 21 per cent in rural areas; fixed broadband in the least developed countries stood at about 2 subscriptions per 100 inhabitants; 5G coverage reached about 84 per cent of people in high-income countries against about 4 per cent in low-income countries. Access to electricity was 74.2 per cent of the population in Senegal in 2023 and about 71 per cent in Cote d'Ivoire, against a sub-Saharan average of about 50 per cent, with an urban and rural split in Cote d'Ivoire of about 92 and 38 per cent. Those are national and regional aggregates and no firm should plan on them. What they establish is the prior: the vendor's throughput assumption was measured where a round trip to a hosted model is uninterrupted and cheap, and the corridor deployment will meter latency, retries, partial batches and power interruptions into the cost line. The instrument response is that the measurement window must include the bad days, and that a measurement window which excludes them has measured the vendor's conditions, not yours.

EVIDENCE

Corridor operating conditions differ structurally from those in which vendor business cases are built: internet use in Africa averaged about 36 per cent of the population in 2025 against a global average of 74 per cent, about 21 per cent in rural areas; fixed broadband in the least developed countries stood at about 2 subscriptions per 100 inhabitants; 5G coverage reached about 84 per cent in high-income countries against about 4 per cent in low-income countries. Access to electricity stood at 74.2 per cent of the population in Senegal in 2023 and about 71 per cent in Cote d'Ivoire, urban 92 and rural 38, against a sub-Saharan average of about 50 per cent.

Sources: ITU, Measuring digital development: Facts and Figures 2025, published 17 November 2025; World Bank, World Development Indicators, access to electricity (EG.ELC.ACCS.ZS), 2023 reference year, with World Bank Group country reporting for Cote d'Ivoire. Geography: Africa region, least developed countries, income groups, Senegal, Cote d'Ivoire. Method: ITU statistical compilation from national administrations and operators; household-survey and administrative compilation under the global electricity access tracking framework. Grade SEG-4. Caveat: regional and national aggregates, not site conditions; access to electricity measures connection, not continuity or voltage quality, which is what a batch process depends on; the Cote d'Ivoire figures combine a World Bank Group country statement with the indicator series, hence SEG-4 rather than SEG-5.

Third, data availability and quality change the denominator, not the margin. The vendor's accuracy claim was measured on documents that exist in machine-readable form, in one language, in one layout family, from counterparties already present in master data. A corridor process receives photographs of documents sent through a messaging application, part-payments, transit and customs steps between order and receipt, suppliers without tax identifiers, and one currency mixed with another on the same file. This does not degrade the benefit by a few points. It changes which files the tool touches at all, and every file it does not touch is a file where the tool adds a step rather than removing one. The instrument response is that the exception rate is measured on live traffic before authorisation, never inferred from the pilot sample, and that the exception rate is a threshold in the chain rather than a caveat in the appendix.

Fourth, oversight and exception handling require staffing ratios that vendor business cases do not carry, because their home markets assume them as ambient. Where a system evaluates the creditworthiness of natural persons or establishes their credit score, human oversight is a legal duty and not a discretionary line: Article 6(2) with Annex III point 5(b) of Regulation (EU) 2024/1689 classifies such systems as high-risk, Article 14 requires that they be capable of effective oversight by natural persons throughout use, and Article 14(4) requires that oversight be assigned to a person with the necessary competence, training, authority and adequate resources. A corridor firm outside the Regulation's territorial scope is not bound by it; it is still the clearest published statement of what competent oversight actually requires, and the cost of that oversight is a board decision either way. The instrument response is that the oversight ratio is authorised as a number, that falling below it suspends the automated path, and that the oversight cost sits inside the benefit calculation rather than beside it.

Fifth, the adaptation window is shorter and the people who would use it are the people running the business. Implementation is mutual adaptation: the technology is changed and the organisation is changed, and misalignments are resolved by adjusting both. Adaptation is also discontinuous, with a relatively brief window after initial implementation during which users explore and modify a new process technology, after

which routinisation congeals the technology and its context and embeds unresolved problems into practice. In an operating company with thin supervisory layers, the staff who would exercise that window are the same staff clearing the daily queue. The instrument response is that the window is scheduled and staffed explicitly, with named people released from production work, or it is not used at all, and whatever was unresolved when it closes becomes the operating model.

The executive consequence of the five differences together: in this corridor the benefit case is destroyed most often not by the technology failing but by four cost lines the vendor's case did not carry, oversight labour, exception rework, connectivity resilience and the adaptation window, arriving against a benefit denominated in a payroll several multiples smaller than the one the evidence was measured in. Traceability is what makes those four lines visible before signature rather than at the first close after go-live.

Economics and uncertainty

The economics of AI benefit claims are governed by a multiplication, and the multiplication is deflationary at every step. Firm-level gain equals task-level gain, times the share of work that task represents, times the share of cost that work represents, times the share of released capacity that actually leaves the cost base. Every term is a fraction below one. This is the core quantitative content of the briefing and it is arithmetic, not opinion.

The first term is the only one that has been measured well, and it has been measured on a narrow object. Peer-reviewed field and laboratory experiments report large task-level effects: issues resolved per hour rose 14 per cent on average among 5,179 customer-support agents, with a 34 per cent gain for novice and low-skilled workers and minimal effect for experienced ones; among 453 college-educated professionals on incentivised writing tasks, average time fell 40 per cent while graded output quality rose 18 per cent. That is real evidence, and its narrowness is what makes it real. Neither study measures payroll, end-to-end cycle time, error cost, or what the released minutes became.

> EVIDENCE
>
> Peer-reviewed experiments measure large task-level generative-AI effects: 14 per cent more issues resolved per hour among 5,179 customer-support agents, 34 per cent for novices and minimal for the most experienced; 40 per cent less time and 18 per cent higher graded quality on incentivised writing tasks among 453 college-educated professionals.
>
> Sources: Brynjolfsson, Li and Raymond, Quarterly Journal of Economics 140(2), 2025, pp. 889-942; Noy and Zhang, Science 381(6654), 2023. Geography: United States, one firm and one remote study population. Method: staggered field rollout of a conversational assistant; randomised controlled experiment. Grade SEG-4. Caveat: task-level, population-specific, short-horizon; instrumented, high-volume, text-centred work; neither measures a firm's profit and loss, and neither concerns senior judgment work or exception-rich processes. A widely cited controlled study of an AI coding assistant is deliberately not used here: see the limits movement.

The second and third terms deflate the first. A task-based macroeconomic model calibrated on published experimental estimates puts the total factor productivity gain at no more than roughly 0.5 to 0.7 per cent in total over ten years, and lower still if hard-to-learn tasks yield less than the easy-to-learn tasks studied. The inputs are contested and other economists estimate materially larger effects; the structure is not contested, and the

structure is what a board needs. A 40 per cent saving on a task that occupies 5 per cent of the work of 20 per cent of staff is 0.4 per cent of that staff cost, which is not a programme.

EVIDENCE

Aggregating measured task-level savings across the share of work actually exposed yields modest economy-level effects: a task-based model calibrated on experimental estimates puts total factor productivity gains at no more than roughly 0.5 to 0.7 per cent in total over ten years.

Source: Acemoglu, "The simple macroeconomics of AI", Economic Policy 40(121), 2025 (NBER Working Paper 32487, 2024). Geography: United States, economy-level model. Method: task-based model calibrated on published experimental estimates. Grade SEG-4. Caveat: a projection with contested inputs; other estimates run materially larger; this briefing uses the multiplicative structure, not the point estimate.

The fourth term is the one boards consistently forget, and it is the firm's position rather than a finding (grade SEG-2, interpretation): capacity released is not cost removed. Hours freed inside a shift are not money until a cost leaves the cost base or a revenue line is invoiced. The evidence of removal is a cancelled requisition, a terminated contract, a withdrawn overtime authorisation or an invoiced incremental volume. Everything else belongs in a capacity register, which is a legitimate and useful artifact, and not in a benefit line. This rule is deliberately conservative and it will understate benefit where released capacity is genuinely absorbed by growth without hiring; the counter-case is stated below.

Between the third term and the fourth sits the depth problem, and it has been observed once by an institution with a defensible sampling frame. In the European Central Bank's SAFE ad hoc module for the fourth quarter of 2025, covering more than 5,000 firms, around 70 per cent reported some level of AI use while only 7 per cent described their use as significant or intensive, and more than 84 per cent of intensive users had invested in the technology against 33 per cent of moderate users. That is the pilot-to-P&L distribution seen from outside the firm: a ten-to-one gap between having a tool and running on it, and a strong association between depth of use and willingness to spend capital rather than subscription. It is not a failure rate and must not be read as one.

EVIDENCE

Breadth and depth of AI use are separated by roughly an order of magnitude in the euro area: around 70 per cent of firms reported some level of AI use in the fourth quarter of 2025, while only 7 per cent described their use as significant or intensive; more than 84 per cent of intensive users had invested in the technology, against 33 per cent of moderate users.

Source: European Central Bank, Survey on the Access to Finance of Enterprises, ad hoc questions for the fourth quarter of 2025, reported in the Economic Bulletin and in the ECB blog of 24 June 2026, covering more than 5,000 firms. Geography: euro area. Method: firm survey with an ad hoc module, self-reported intensity, comparisons within industry, country, size and age cells. Grade SEG-4. Caveat: self-reported against the survey's own scale; no financial outcome measured; a snapshot of a fast-moving distribution and explicitly not a project failure rate.

The appraisal correction. Official appraisal guidance does not treat benefit optimism as an occasional lapse. It treats it as systematic, and it requires an explicit, quantified correction: appraisers must increase cost estimates and both reduce and delay estimated benefits, using empirically based adjustments drawn from past or comparable projects, which may be relaxed only as project-specific risk is identified and managed. Independently, a supreme audit institution examining government digital change programmes found that such programmes often experience significant increases in cost and decreases in expected benefits. The published uplift factors are calibrated on public capital programmes and do not transfer to a corridor AI deployment; the requirement transfers completely. A benefit case with no optimism correction, no delay assumption and no source for either is below the standard a public body would apply to a road.

EVIDENCE

Official appraisal guidance treats over-optimism in benefit and cost estimates as systematic and requires an explicit correction: costs increased, benefits both reduced and delayed, on empirically based adjustments from past or comparable projects, relaxed only as project-specific risk is identified and managed. Separately, an audit of government digital change programmes found that such programmes often experience significant increases in cost and decreases in expected benefits.

Sources: HM Treasury, The Green Book (2026 edition) and Supplementary Green Book Guidance: Optimism Bias; National Audit Office, The challenges in implementing digital change, HC 575, 21 July 2021. Geography: United Kingdom central government; method of general application. Method: government appraisal guidance; supreme audit institution value-for-money examination. Grade SEG-5. Caveat: published uplift factors are calibrated on public capital programmes and are not transferable to a private AI deployment; the audit finding is qualitative, concerns programmes examined to 2021, and supplies no rate, from which none should be inferred.

Counter-case. The house position would be wrong, or materially weakened, in four identifiable conditions. One: where the deployment is cheap relative to the cost base and the instrumentation would cost more than the tool, the chain is disproportionate and should be replaced by a spending cap and a review date; the test is always proportionate to the money at stake. Two: where the firm is growing fast enough that released capacity is absorbed by volume without hiring, the recognition quantum understates real benefit, and the correct treatment is a hiring-avoidance register signed by the executive who would otherwise have raised the requisition, not a relaxation of the rule. Three: where the benefit is a risk reduction rather than a cost saving, for example a lower rate of undetected fraud or fewer regulatory findings, the profit-and-loss line is a loss line and the baseline is a loss history, which is legitimate but requires a longer measurement window and a wider confidence interval than a throughput claim. Four: where the deployment is defensive, undertaken because a counterparty or a regulator requires it, in which case the traceability chain should be run against the compliance obligation and not against a benefit that was never the reason.

Highest-leverage unknown. The exception rate on live traffic. Every other term in the multiplication can be estimated from published evidence or from the firm's own ledger. The exception rate cannot: no official series measures it, no vendor case study will report it honestly for your document mix, and the pilot was designed not to encounter it. It is also cheap to measure, it is measurable before any licence is signed, and it moves the answer more than any other input. If measured exception rates on a full cycle of live traffic come in below about 10 per cent for the firm's actual document mix, the arithmetic in this briefing loosens considerably and the

deployment case may well stand on its own; if they come in above about 40 per cent, no plausible licence price rescues it. That is the falsification condition and it is testable in one operating cycle by any process owner with a sampling frame and a stopwatch.

What counts as evidence inside the chain. The same grading discipline applied to this briefing applies inside the reader's own file. Benefit claims rank as follows: instrumented system logs with a defined population and a pre-period, strongest; the firm's own general ledger and payroll records; time-and-motion measurement on a documented sample; process owners' structured estimates recorded with a range; and, weakest, vendor-supplied benchmarks and user satisfaction surveys, which are graded as interpretation and never carry a recognition decision alone. Measurement method must be declared before measurement begins, in line with the discipline the OECD manual imposes on productivity measurement generally: choose gross output or value added and say which, keep the output measure independent of the input measure, and adjust for quality change rather than counting faster output of worse work as a gain.

What this briefing is not valid for is set out in the failure-modes movement below and must travel with any internal circulation of the document.

Instruments: the benefit traceability chain

The house instrument for this decision is the benefit traceability chain: six sequential links, each producing an artifact that can be handed to a board, tested against a numeric rule, owned by a named person and dated. It is presented here as a method. No governed public framework label is bound to it in this version, and any such label requires resolution in the firm's signature-IP register before publication.

The chain is not a maturity model and it is not scored. It is passed or it is stopped, link by link, in order. The order matters: attribution cannot be assessed before there is a measurement, and transfer cannot be assessed before there is an attribution.

Link 1. Claim registration. Artifact: a one-page claim record stating the claim in a single measurable sentence, the claimed effect size, the source of the claim, the population it was observed in, the profit-and-loss line it would eventually land in, and the executive who would own that line. Test: can the claim be falsified, and does it name a line before any money is spent. Rule: a claim that cannot name its line and its owner in one sentence does not enter the pipeline. Owner: the sponsoring executive. Timing: before any pilot funding is released. On failure: **stop**. The proposal is returned, not deferred.

Link 2. Measured effect in a defined population. Artifact: a measurement protocol, written before measurement begins, plus the baseline dataset it produces. Test: is there a baseline that predates the intervention, and is the population defined by an inclusion rule rather than by who volunteered. Rule: a baseline of at least three months or one full seasonal cycle, whichever is longer; at least 80 per cent of live volume in scope, with exclusions listed and justified; measurement error no greater than 10 per cent of the claimed effect, or the claimed effect is unmeasurable with the available instrumentation and must be re-specified. Owner: the chief financial officer with the process owner. Timing: begins at least 90 days before any go-live. On failure: **re-baseline**. The decision is deferred until the baseline exists; this is the most common stop in practice and the cheapest.

Link 3. Attribution to the intervention. Artifact: an attribution note setting out the counterfactual construction and the confounders considered. Test: would the observed change have happened anyway. Rule: an explicit counterfactual is mandatory, built as a randomised or staggered rollout, a matched comparison group, or a

documented pre-period comparator adjusted for trend and seasonality; the observed effect must exceed the pre-period trend by at least twice the residual standard deviation of that trend; every simultaneous process change in the window must be listed and either held constant or separately estimated; a population composed only of volunteers disqualifies the measurement for attribution purposes and it reverts to Link 2. Owner: the chief financial officer. Timing: at the close of the measurement window. On failure: **re-scope**. The claim reverts to capability evidence and may be recorded as such; it may not be carried forward.

Link 4. Transfer to this firm's conditions. Artifact: a transfer test sheet scoring the measured population against this firm on six dimensions: task mix, skill distribution of the affected staff, input data quality, live-traffic exception rate, loaded wage rate, and oversight regime. Test: does the population in which the effect was measured resemble ours on the dimensions that drive the effect. Rule: state a transfer coefficient explicitly and justify every dimension scored below one; any single dimension scored below 0.5 requires local measurement before recognition is permitted; the live-traffic exception rate is measured over one full cycle and gates the link directly, with 25 per cent or below permitting authorisation, 25 to 40 per cent requiring re-scope to the subset of traffic that clears, and above 40 per cent stopping the deployment. Owner: the process owner with the chief financial officer. Timing: before deployment authorisation, never after. On failure: **stop or re-scope**, according to which dimension failed.

Link 5. Recognition in a named line. Artifact: a recognition memo carrying the named line, the dated baseline value, the named owner, the quantum rule, and the full three-year run-rate netted against the gross benefit. Test: has a cost actually left the cost base, or has a revenue line actually been invoiced. Rule: recognition only against evidence of removal, being a cancelled requisition, a terminated contract, a withdrawn overtime authorisation or an invoiced incremental volume; fractional full-time equivalents and redeployed hours are entered in the capacity register and are not benefits; benefit is always stated net of licence, integration, oversight, monitoring, retraining, connectivity and version-change cost. Owner: the chief financial officer, countersigned by the line owner. Timing: at the first close after go-live and at every close thereafter. On failure: **re-classify**. The gain is capacity, not benefit, and the plan line is removed.

Link 6. Verification after the fact. Artifact: a post-implementation review written against the Link 1 claim record, not against a revised claim. Test: did the named line move by the named amount by the named date. Rule: verification at 6, 12 and 24 months on a fixed calendar; movement below 50 per cent of the claimed amount at 12 months triggers re-scope; movement below 25 per cent at 12 months, or a negative net position at 24 months, triggers the decommission protocol. Owner: the audit committee, reporting to the board. Timing: fixed at authorisation and not movable by the sponsor. On failure: **kill**, with the decommission reserve released.

LINK	ARTIFACT	TEST	RULE / THRESHOLD	OWNER	TIMING	ON FAILURE
1 Claim	Claim record, one page	Falsifiable and names a line	No line and owner in one sentence, no entry	Sponsoring executive	Before pilot funding	Stop

LINK	ARTIFACT	TEST	RULE / THRESHOLD	OWNER	TIMING	ON FAILURE
2 Measurement	Protocol plus baseline dataset	Baseline predates intervention	3 months or one seasonal cycle; 80 per cent of live volume; error under 10 per cent of claimed effect	CFO with process owner	Starts 90 days before go-live	Re-baseline
3 Attribution	Attribution note	Would it have happened anyway	Effect exceeds pre-period trend by 2 residual standard deviations; simultaneous changes listed; volunteers disqualify	CFO	At close of measurement window	Re-scope
4 Transfer	Transfer test sheet, six dimensions	Does their population resemble ours	Any dimension under 0.5 requires local measurement; exception rate 25 per cent authorise, 25 to 40 re-scope, over 40 stop	Process owner with CFO	Before authorisation	Stop or re-scope
5 Recognition	Recognition memo	Has a cost actually left	Removal evidence only; fractional FTE to capacity register; net of full run-rate	CFO, countersigned by line owner	First close and every close	Re-classify

LINK	ARTIFACT	TEST	RULE / THRESHOLD	OWNER	TIMING	ON FAILURE
6 Verification	Post-implementation review	Did the line move as claimed	Under 50 per cent at month 12 re-scope; under 25 per cent at month 12 or negative at month 24 decommission	Audit committee	6, 12, 24 months, fixed	Kill

Benefit attribution, the hard part, handled honestly. Link 3 is where most benefit cases quietly fail, and it fails for five reasons that are all foreseeable. Official evaluation guidance is unambiguous that no impact evaluation is possible without an explicit counterfactual, and that attribution is a separate question from outcome. The five that matter in an operating company:

The counterfactual. The comparison is never "before and after". It is "with and without", and the without has to be constructed. In descending order of strength: randomised assignment of users or files, which is usually available and usually refused for reasons of convenience; staggered rollout across branches or teams, which is nearly always available and costs a quarter of calendar; a matched comparison group in an untreated unit; a documented pre-period comparator with trend and seasonality adjustment, which is the weakest defensible option and requires the adjustments to be shown. Theory-based methods can support attribution without precise effect sizes and should be used where the process is too small or too entangled for a comparison group, provided the causal chain is written down and each step is separately evidenced.

Baseline drift. Processes improve without intervention, because managers manage. If throughput was already improving at 3 per cent a quarter before the tool arrived, the tool must beat 3 per cent a quarter to have done anything. The baseline is therefore a trend line, not a point, and the pre-period must be long enough to estimate its slope. A single pre-month is not a baseline; it is a data point that will be quoted for three years.

Selection of volunteer users. Pilot participants volunteered, and volunteers are the population most disposed to make a tool work. They are also frequently the most capable staff, which cuts the other way, since the strongest measured effects concentrate among novices. Either direction of selection destroys attribution. The rule is an inclusion rule: all files of a defined type, or all staff in a defined grade and unit, with opt-out recorded and counted.

Seasonality. Corridor processes have pronounced annual and monthly shapes, driven by fiscal deadlines, import cycles, harvest and holiday periods. A measurement window that spans a quiet month and compares to a busy one will produce a large effect made entirely of the calendar. The measurement window must either cover a full cycle or be matched to the same period in the prior year, and the choice must be declared in the protocol before the data are seen.

Simultaneous process changes. Tools rarely arrive alone. A new tool usually arrives with a redesigned form, a new checklist, additional training, management attention and, in the corridor, often a new supervisor. Every one of those is an intervention. They must be listed in the protocol, and either frozen for the duration of the

measurement window or estimated separately, most cheaply by staggering them. A benefit attributed to a tool that in fact belongs to a checklist will not survive the checklist being copied elsewhere for free.

Maintenance run-rate. A deployment is an asset with a whole life, and the international asset-management standard states the general principle plainly: realisation of value normally involves a balancing of costs, risks, opportunities and performance benefits. Six run-rate lines are costed for three years at authorisation, not one: licence or subscription, including contractual escalation; monitoring of model performance and data drift; retraining or re-fitting, with a stated frequency; vendor version changes, costed as regression testing and re-acceptance days for both business and technical staff; human review and exception-handling labour at the authorised oversight ratio; and connectivity and power resilience attributable to the deployment. The cost side of the ledger is also an accounting question, and the reporting framework is relevant: configuration or customisation activities in a software-as-a-service arrangement do not generally create a resource controlled by the customer separate from the software, so those costs are generally recognised as an expense. That treatment is developed in a separate estate publication and is referenced here only to locate the cost side of the chain; local statutory treatment must be confirmed with the statutory auditor.

Kill criteria and the decommission trigger. These are thresholds, not sentiments, and they are set at authorisation:

TRIGGER	THRESHOLD	CONSEQUENCE
K1 Exception rate	Live-traffic exception rate above 40 per cent for two consecutive months	Deployment stopped; automated path withdrawn; re-scope or decommission
K2 Net time released	Net analyst time released at or below zero over a full quarter	Immediate decommission; the tool is consuming labour
K3 Recognised movement	Named line moved less than 25 per cent of the claimed amount at month 12	Decommission unless the board re-authorises against a re-baselined claim with a new date
K4 Run-rate revision	Three-year run-rate revised upward by more than 30 per cent against the authorised figure	Re-authorisation required; absent re-authorisation, non-renewal at term
K5 Oversight floor	Authorised oversight ratio unmet for more than 30 days	Automated decisioning suspended; recorded as a control breach, not a cost saving
K6 Version integrity	A vendor version change altering model behaviour without an accepted regression pack	Suspension until re-tested; two suspensions in 12 months triggers decommission review

A decommission reserve is budgeted at authorisation, at a minimum of 15 per cent of one year's run-rate, covering contract exit, data extraction and return, re-training of staff on the manual path, and the temporary throughput loss during reversion. A deployment with no exit budget is not reversible, whatever the contract says.

Decision sequence

Five actions, each with owner, threshold, timing and escalation. They are sequential, and the worked scenario shows why the order is not negotiable.

#	ACTION	OWNER	TRIGGER / THRESHOLD	TIMING	EVIDENCE PRODUCED	ESCALATION
1	Register every live AI claim in the group on a single claim record and refuse entry to any claim without a named line and owner	CFO	Any claim already inside a plan line without a claim record	Within 30 days of adopting this briefing	Claim register, CFO-signed	Unregistered claim already in a plan goes to the audit committee as a control finding
2	Measure the live-traffic exception rate over one full operating cycle for every candidate process, before any licence commitment	Process owner	Exception rate at or below 25 per cent authorises; 25 to 40 re-scopes; above 40 stops	One full cycle, starting immediately, ahead of procurement	Exception count by type with handling cost	Above 40 per cent, the procurement is closed by the CFO, not paused by the sponsor
3	Build the pre-intervention baseline and freeze the measurement protocol in writing before any data are seen	CFO with process owner	Baseline of 3 months or one seasonal cycle, whichever longer; 80 per cent of live volume in scope	Starts at least 90 days before any go-live	Protocol document and baseline dataset, dated	Protocol changed after data are seen: measurement void, restart

#	ACTION	OWNER	TRIGGER / THRESHOLD	TIMING	EVIDENCE PRODUCED	ESCALATION
4	Authorise deployment only against a full three-year run-rate and a stated oversight ratio, with kill thresholds and a decommission reserve set in the same paper	Board, on CFO recommendation	Recognised benefit net of run-rate positive over three years at the measured exception rate	At the authorisation decision	Authorisation paper with thresholds and reserve	Negative net at the measured exception rate: authorisation refused or volume threshold set as a condition precedent
5	Verify against the original claim record on a fixed calendar and apply the kill criteria without renegotiation	Audit committee	Under 50 per cent of claim at month 12 re-scopes; under 25 per cent, or negative net at month 24, decommissions	6, 12 and 24 months	Post-implementation review	Sponsor requests a revised claim as the comparison basis: refused, and the request is logged

The sequence exists because action 4 is impossible to do honestly without actions 2 and 3, and because action 5 is meaningless unless action 1 preserved the original claim. A firm that starts at action 4, an authorisation paper without a measured exception rate, produces a business case whose largest single variable is an assumption made by the party selling the software.

Worked scenario: an illustrative Abidjan lender

All figures in this scenario are ILLUSTRATIVE and synthetic. They describe no client and no real company. They are chosen to be representative of a mid-market corridor lender and to make the arithmetic reproducible from the stated inputs. Basis: synthetic; every input is a scenario assumption. Currency conversion uses the fixed parity of 655.957 XOF per EUR.

Situation. Lender L is a mid-market credit institution based in Abidjan, lending to small and medium enterprises across Cote d'Ivoire, with a Dakar branch. A vendor has demonstrated a document-processing tool that extracts and spreads financial statements, bank statements, invoices and tax attestations at the front of the credit-decisioning workflow. The pilot ran for six weeks on 400 pre-selected files and showed a 40 per cent reduction in analyst handling time on the extraction and spreading segment. The sponsor asks the board to authorise deployment and to book the resulting saving in the plan.

Inputs. Every value below is a scenario assumption.

INPUT	VALUE (ILLUSTRATIVE)
Credit files processed	24,000 per year, about 2,000 per month
Credit analysts	18 full-time equivalents
Fully loaded cost per analyst	XOF 7,200,000 per year, about EUR 10,977
Total analyst payroll	XOF 129.6 million per year
Productive analyst hours	1,700 per year, giving XOF 4,235 per analyst hour
Baseline handling time per file	95 minutes end to end
Treated segment, extraction and spreading	57 minutes of the 95, the remaining 38 being judgment, calls and committee
Pilot-measured effect on the treated segment	40 per cent reduction, saving 22.8 minutes per clean file
Exception rework multiplier	1.25 times the treated segment, so an exception file consumes 71.25 minutes against 57 baseline
Live-traffic exception rate, base case	22 per cent
Licence	EUR 54,000 per year, XOF 35.4 million, escalating 5 per cent a year
Integration, one-off	EUR 45,000, XOF 29.5 million
Baseline build and instrumentation, one-off	XOF 3.2 million
Oversight	1.0 senior credit officer at XOF 12.0 million per year
Model monitoring and retraining	0.5 technical full-time equivalent at XOF 9.0 million, so XOF 4.5 million
Connectivity and power resilience attributable	XOF 3.6 million per year
Vendor version changes	two per year, 20 combined business and technical days at XOF 45,000, XOF 1.8 million
Payroll escalation	5 per cent a year

Calculation 1, net time released per file. Net minutes per file equal 22.8 times (1 minus e) minus 14.25 times e, where e is the live-traffic exception rate. The 14.25 is the rework penalty: an exception file does not merely fail to save, it costs a quarter of the treated segment more than the manual path, because the analyst reviews the machine output and then redoes the work. At the base case of e equal to 0.22, net minutes per file are 14.65. Over 24,000 files that is 351,576 minutes, or 5,860 hours, or 3.45 full-time equivalents, worth about XOF 24.8 million of capacity a year. The rate at which net released time reaches zero is e equal to 61.5 per cent: above that exception rate the tool consumes more analyst time than it saves, at any volume and at any licence price.

Calculation 2, capacity released against cost removed. Capacity released is 3.45 full-time equivalents. Cost removed, under the recognition quantum, is 3 cancelled requisitions, XOF 21.6 million a year. The residual 0.45 full-time equivalent is entered in the capacity register and is not a benefit. This gap of about XOF 3.2 million a year is not an accounting nicety; it is the difference between a plan line that can be defended at the first close and one that cannot.

Calculation 3, the run-rate in analyst units. The annual run-rate is XOF 35.4 million licence, XOF 12.0 million oversight, XOF 4.5 million monitoring, XOF 3.6 million connectivity and XOF 1.8 million version changes: XOF 57.3 million. At XOF 7.2 million per analyst that is the equivalent of 7.96 analysts. The licence alone, XOF 35.4 million, is 8,359 analyst hours, or 4.9 analysts working full time purely to pay for the tool. The deployment costs the equivalent of about 8 analysts in order to release the equivalent of about 3.4.

Calculation 4, the three-year position.

YEAR	COSTS (XOF MILLION)	RECOGNISED BENEFIT (XOF MILLION)	NET
1	90.0 (run-rate 57.3 plus integration 29.5 plus baseline build 3.2)	7.2 (one avoided hire, benefit from month 7)	minus 82.8
2	59.9	22.7 (three avoided hires at escalated payroll)	minus 37.2
3	62.6	23.8	minus 38.8
Total	212.5	53.7	minus 158.8

Three-year net is minus XOF 158.8 million, about EUR 242,000 at the fixed parity, or roughly 1.2 times the entire annual analyst payroll of the function the tool was meant to make cheaper.

Calculation 5, break-even volume and the sensitivity that matters. Break-even requires recognised benefit to cover the run-rate, which needs 7.96 avoided full-time equivalents. The table below gives the annual file volume at which that is reached, by live-traffic exception rate and by oversight staffing. Current volume is 24,000 files.

LIVE-TRAFFIC EXCEPTION RATE	OVERSIGHT 0.5 FTE	OVERSIGHT 1.0 FTE	OVERSIGHT 2.0 FTE
10 per cent	38,000 files	42,500 files	51,500 files
22 per cent (base case)	49,500 files	55,500 files	67,000 files
35 per cent	74,000 files	82,500 files	100,000 files
45 per cent	118,500 files	132,500 files	160,000 files

No cell is within reach of current volume. The base case requires 55,500 files a year, 2.3 times today's throughput. A realistic deterioration to a 35 per cent exception rate, entirely plausible for a document mix that

includes photographed statements and part-payments, pushes it to 82,500. The benefit case does not degrade gracefully under exception rates; it collapses, because exceptions do not merely fail to save time, they consume it.

Calculation 6, the licence question, answered. At current volume and one oversight full-time equivalent, the non-licence run-rate alone, XOF 21.9 million, already exceeds the recognisable benefit of XOF 21.6 million. No licence price closes that gap: the tool would have to be free and the case would still be negative. At half an oversight full-time equivalent, the non-licence run-rate is XOF 15.9 million, leaving XOF 5.7 million a year of headroom for the licence, about EUR 8,700, against a quotation of EUR 54,000, a factor of 6.2. Expressed per file, the maximum defensible price is about XOF 238 per file, roughly EUR 0.36, against a quoted equivalent of about XOF 1,475 per file, roughly EUR 2.25.

Result. The board's decision on this claim is: do not authorise deployment; do authorise measurement. Specifically, fund the exception-rate count and the baseline build, XOF 3.2 million, as a standalone measurement programme with a 90-day term; register the claim at Link 1 with the named line, credit operations payroll, and the named owner, the head of credit operations; return to the vendor with a per-file price ceiling of XOF 238 and a volume-triggered structure rather than a flat annual licence; and set a volume condition precedent of 50,000 files a year, whether reached organically or by consolidating the Dakar branch's files onto the same path. The claim is recorded as capability evidence, which it is, and not as a benefit, which it is not.

Interpretation. The scenario is not a prediction and no number in it is a market observation. It demonstrates three things a board can use. First, the currency mismatch is not about volatility: the peg is fixed and the licence is still 4.9 analysts, because the mismatch is between a hard-currency price and a corridor payroll, not between two moving exchange rates. Second, oversight and exception handling, the two lines vendor cases carry lightest, together outweigh the licence in their effect on the answer. Third, volume is the variable that decides, and it is the one nobody negotiates, because everybody negotiates price. Readers must re-run every number on their own ledger before acting; the sensitivity table is the template, and the exception rate is the input to measure first.

Failure modes and invalid uses

Five recurring failure modes, each with its trigger and its owner. The trigger, not the mood of the quarter, forces the escalation.

FAILURE MODE	ESCALATION TRIGGER	OWNER	RESPONSE
Claim drift: the claim verified at month 12 is not the claim registered at Link 1	Any variance between the claim record and the review's comparison basis	Audit committee	Review re-run against the original record; the substitution is logged as an exception event
Baseline laundering: the baseline is re-cut after the data are seen	Any protocol change dated after first data extract	CFO	Measurement void; restart with a fresh protocol and a fresh window
Gross-benefit reporting: savings quoted before run-rate	Any benefit figure circulated without the netted run-rate on the same page	CFO	Figure withdrawn from circulation; corrected version reissued to the same list

FAILURE MODE	ESCALATION TRIGGER	OWNER	RESPONSE
Oversight erosion: the authorised ratio quietly falls as staff move	Authorised oversight ratio unmet for more than 30 days	Head of the affected function	Automated path suspended; recorded as a control breach, never as a saving
Zombie renewal: a deployment that failed verification renews because nobody owns the exit	Any renewal proposed for a deployment below its month-12 threshold	Board	Renewal refused by default; continuation requires a fresh authorisation with a re-baselined claim

NOT VALID FOR

This briefing must not be used as: a technology selection guide or a comparison of tools or vendors; an accounting policy, for which the applicable framework and the statutory auditor govern; a legal compliance opinion on Regulation (EU) 2024/1689 or on any national personal-data regime, for which counsel is required; a basis for asserting any project failure rate, since this briefing publishes none and refuses the circulating figures; a model-risk or model-validation methodology, which is a separate discipline with its own standards; a staff-reduction plan, since the recognition quantum is a measurement rule and not an instruction to remove people; a template for deployments outside the corridor without reworking the currency, connectivity and oversight reasoning; or any use of the illustrative scenario figures as market data, benchmark data or pricing guidance.

Stakeholder and institutional map

The traceability decision sits inside an institutional lattice the executive should map by name. The table is the field version; the authorisation paper behind it holds the detail.

INSTITUTION	ROLE IN THE TRACEABILITY DECISION	WHAT TO MONITOR
The board and its audit committee	Sole authority to authorise recognition and to apply the kill criteria; the only body that cannot be lobbied by the sponsor	That the verification calendar is fixed at authorisation and not moved
Statutory auditor	Governs whether implementation and configuration cost is expensed or capitalised, and will test any benefit asserted in the accounts	Accounting framework applicable locally; treatment of software-as-a-service configuration cost
Banking supervisor for the union	Where the deploying firm is a credit institution, supervises outsourcing, model use and operational resilience in the credit process	Outsourcing and operational-resilience expectations; notification duties before deployment
BCEAO	Monetary anchor: the fixed parity that converts a euro licence into a CFA franc cost line without volatility	The parity framework, which is the conversion basis for every euro-denominated cost in the chain

INSTITUTION	ROLE IN THE TRACEABILITY DECISION	WHAT TO MONITOR
National personal-data authority in each country of operation	Governs processing of the personal data inside credit files and documents, and the lawful basis for automated processing	Registration, notification and transfer requirements before any data leaves the country
Group or parent entity, where EU-established	May bring the deployment within the scope of Regulation (EU) 2024/1689, converting oversight from a board choice into a legal duty	Territorial scope, the group's own classification of the system, and the current application timetable on EUR-Lex
Vendor	A counterparty with a version roadmap that changes your system without your decision; also the source of the claim being tested	Version-change notice periods, regression obligations, exit and data-return terms
Connectivity and power providers	Supply an input to the benefit, not a background service	Contracted availability, redundancy, and the cost of the resilience the deployment requires
Works council or staff representation, where constituted	Workflow redesign changes job content, and job content is negotiated	Consultation obligations triggered by role change, ahead of the adaptation window

The French edition of an authorisation paper, and of this briefing, is not a courtesy: French is the language of legal record across the OHADA states, and a benefit case that exists only in English will be tested in a language it was not written in.

Operating constraints and dependencies

Four external dependencies sit outside the firm's control and are monitored, not managed:

DEPENDENCY	WHAT WOULD CHANGE THE BRIEFING	WHERE TO MONITOR
Vendor pricing structure and version policy	A per-file or volume-triggered price, or a longer version-change notice period, changes the break-even volume directly	The contract, at each renewal window
Regulatory scope of AI oversight duties	Bringing the deployment within a high-risk regime converts the oversight ratio from a board decision into a floor	EUR-Lex for the Regulation as amended; the group's own scoping determination
Connectivity and power availability at the deployment site	Site-level degradation raises the exception rate, which is the single most sensitive input in the chain	The firm's own incident log, not national statistics

DEPENDENCY	WHAT WOULD CHANGE THE BRIEFING	WHERE TO MONITOR
Local accounting treatment of configuration and implementation cost	A different capitalisation outcome changes the shape of the three-year position, though not its sign	Statutory auditor, at each reporting date

Five constraints bound implementation. **Data constraint:** the exception count requires document-level logging that many mid-market systems capture but do not report; budget one analyst-month of extraction and a manual tally sheet, not a new system. **People constraint:** the oversight role must be senior enough to override the machine and junior enough to be available, and firms consistently under-price this; an oversight officer who cannot in practice reverse a decision is a control on paper and a cost in fact. **Timing constraint:** the adaptation window opens once, briefly, and closes into routine; if the people who would use it are not released from production work in that window, whatever is unresolved becomes the operating model permanently. **Governance constraint:** the verification calendar must be immovable by the sponsor, which means it belongs to the audit committee and not to the programme. **Proportionality constraint:** the full chain is disproportionate below a materiality threshold each board should set explicitly, below which a spending cap and a review date replace it; running a six-link chain on a small subscription is a governance cost with no governance benefit.

Implementation cadence and field checks

The cadence is fixed, not aspirational; each window has one owner.

WINDOW	ACTIVITY	OWNER
Days 1 to 30	Claim register built; every live claim in the group given a claim record or refused (action 1)	CFO
Days 1 to 60	Live-traffic exception count over one full cycle, by type, with handling cost (action 2)	Process owner
Days 30 to 120	Measurement protocol frozen and baseline built; falsification condition tested (action 3)	CFO with process owner
Before authorisation	Transfer test sheet, three-year run-rate, oversight ratio, kill thresholds and decommission reserve in one paper (action 4)	CFO to board
Adaptation window, first 90 days after go-live	Named staff released from production to modify the process; unresolved items logged before the window closes	Process owner
Monthly	Exception rate and oversight ratio reported against their thresholds	Head of the affected function
Quarterly	Run-rate actuals against the authorised run-rate; K4 tested	CFO
Months 6, 12 and 24	Verification against the original claim record; kill criteria applied (action 5)	Audit committee

WINDOW	ACTIVITY	OWNER
Annually	Full re-run of the scenario arithmetic on actual ledger numbers, including a fresh break-even volume	CFO

Field checks, asked in the room by the chair before any authorisation proceeds:

1. Show me the claim record. What line does this land in, and who owns that line?
2. What was the exception rate on live traffic, measured over what period, by whom?
3. What is the counterfactual, and what would have happened anyway?
4. What is the three-year run-rate, and what is it in units of our own staff?
5. What has actually left the cost base, and what is only capacity?
6. What are the kill thresholds, and what is the decommission reserve?
7. What date is the verification, and who cannot move it?

A no, or a hesitation, on any of these postpones the authorisation. That postponement rule is itself the discipline, and it costs a month against a three-year commitment.

Decision log

Every material AI benefit decision is recorded at the moment it is made, in this format, and the log is reviewed quarterly by the audit committee:

FIELD	ENTRY
Date and deployment	
Link reached (1 to 6) and link outcome	
Decision (record only / fund measurement / authorise deployment / recognise benefit / refuse / decommission)	
Claim record reference and original claimed amount	
P&L line named; baseline value and baseline date	
Measured exception rate on live traffic; measurement window	
Counterfactual construction used	
Three-year run-rate authorised; oversight ratio authorised	
Recognised benefit, XOF per year, net of run-rate	
Capacity register entry, where capacity was released but not removed	
Kill thresholds set; decommission reserve set	
Owner and countersignature	
Verification dates (6, 12, 24 months)	

The log converts AI benefit management from anecdote into an auditable series. After four quarters it becomes the firm's own evidence base for the next claim, and it is superior to any external benchmark for the processes it covers, because it is measured on this firm's documents, this firm's staff and this firm's exception profile.

Exception and escalation protocol

Exceptions are permitted; unlogged exceptions are not. A benefit may be recognised outside the quantum rule, or a threshold waived, only by the board, in writing, with the value at risk quantified, a stated reason and an expiry date. Standing exceptions are re-authorised or lapse at each quarterly review.

RUNG	TRIGGER	DECIDED BY	CLOCK
1	Any link threshold missed by a candidate deployment	Process owner with CFO	Within 5 working days
2	A request to recognise benefit outside the removal-evidence rule, or to book capacity as benefit	CFO	Before the close in question
3	A request to move a verification date, revise a claim record, or waive a kill threshold	Audit committee	Before the date in question, never retrospectively
4	Cumulative authorised AI run-rate above a board-set share of the cost base, or any single deployment above the materiality threshold	Board	Next board session, with the claim register attached

Any attempt to route around the ladder, a sponsor seeking authorisation directly from an executive who does not own the line, is itself logged as an exception event and reported to the audit committee in the quarterly review.

Transferability, limits and sources

What transfers. The six-link chain, the distinction between capacity released and cost removed, the attribution discipline, the kill thresholds and the decision log transfer to any firm evaluating any technology benefit claim, in any geography and in any sector. None of that is corridor-specific, and none of it is specific to AI. It is the appraisal discipline that public bodies apply to capital programmes, applied to a subscription.

What does not transfer. The numeric thresholds are calibrated to the illustrative scenario and must be re-set by each board on its own economics: an exception-rate ceiling of 25 per cent, a break-even volume of 55,500 files, a rework multiplier of 1.25 and an oversight ratio of one full-time equivalent are scenario values, not standards. The currency reasoning is peg-specific in one direction only: the absence of volatility is corridor-specific, while the mismatch between a hard-currency licence and a local payroll is common to most emerging markets and is worse where the currency floats. The connectivity and electricity anchors are regional and national aggregates and describe a prior, not a site. The corridor description assumes a document-heavy, exception-rich, transaction-processing context; a firm deploying AI into research, design or senior judgment work faces a different measurement problem, and the task-level evidence cited here is not about that work.

Adjacent contexts, for contrast. A reader who also operates in Ghana or Nigeria should not carry the currency arithmetic across the border: there, the licence is a floating claim, the break-even volume moves with the exchange rate, and a three-year run-rate must be computed in a depreciating unit, which changes the

meaning of a multi-year commitment entirely. A reader deploying in a group whose parent is EU-established should assume the oversight ratio is a legal floor rather than a board choice, and should confirm the current application timetable directly.

Anchors dropped, and why. Three figures were considered and are not used, and the drops are stated here rather than buried. First, the circulating claim that between 70 and 95 per cent of AI projects fail is refused explicitly. There is no published sampling frame; the failure criterion, generally an absence of measurable profit-and-loss impact, is not defined in a way anyone could reproduce; interview and survey counts differ between secondary accounts of the same source; the most-cited version is an unrefereed report amplified by consultancy and vendor commentary, and the seventy per cent variant is older transformation lore with no traceable measurement behind it at all. None of it is official, multilateral, standards-body, national-statistical or peer-reviewed, and citing it to sell governance would mean adopting the vendor's evidentiary method in order to criticise the vendor's product. This briefing therefore publishes no failure rate anywhere, and the adoption-versus-intensity distribution above is explicitly not one. Second, a widely cited controlled study reporting that developers with an AI coding assistant completed a defined task 55.8 per cent faster is dropped here, although the paper exists and was located. It is an unrefereed working paper and its authors are affiliated with the vendor whose product is under test; under the rule that no vendor statistic is cited anywhere, vendor-affiliated authorship on the vendor's own product disqualifies it regardless of plausibility. The companion note cites it with an explicit unrefereed flag; this briefing applies the stricter rule, and the divergence is deliberate. Third, an intermediate euro-area figure for firms at an advanced stage of adoption did not surface against the publishing institution on this verification pass and is not asserted; the verified pair, around 70 per cent reporting some use against 7 per cent reporting significant or intensive use, carries the same argument more sharply. Separately, the staged application dates for high-risk obligations under Regulation (EU) 2024/1689 are not asserted, because only secondary commentary surfaced for the amended timetable; readers must confirm the current dates against EUR-Lex before relying on them.

Data gaps that no source closes. No official series measures value realised from AI deployment at firm level, in any geography. No official series measures live-traffic exception rates for document-processing or credit workflows. No official corridor equivalent of the Eurostat or Census adoption series exists, so the adoption figures used here describe European and United States populations and are cited for what they establish about the limits of official measurement, not as corridor conditions. The oversight staffing ratios in this briefing are firm judgment and are not measured anywhere. Each of these gaps is the reason the chain requires local measurement rather than published benchmarks.

Refresh conditions. Re-validate this briefing on any of the following, and in any case every two quarters:

1. A new Eurostat ICT usage release or a material revision to the Census Bureau AI use series.
2. A new ECB SAFE ad hoc module on AI adoption intensity.
3. Any amendment to Regulation (EU) 2024/1689 affecting classification, oversight duties or application dates.
4. Any revision to HM Treasury appraisal or evaluation guidance affecting the optimism correction or the counterfactual standard.
5. Any change to the CFA franc parity or its guarantee framework, which would alter every currency conversion in the chain.
6. Publication of any peer-reviewed field evidence measuring firm-level financial outcomes rather than task-level effects, which would materially strengthen links 3 and 4.

Claim ledger. The briefing's material claims, as graded in the governed evidence register:

CLAIM	TYPE	GRADE	SOURCES	VALID AS OF
Official statistics count adoption, not value; EU 20.0 per cent 2025, US 19.8 per cent May 2026	Observed fact	SEG-5	Eurostat; US Census Bureau	2026-05-03
Breadth versus depth: about 70 per cent any use, 7 per cent significant or intensive	Observed fact	SEG-4	ECB SAFE ad hoc module, Q4 2025	2026-06-24
Task-level effects: 14 and 34 per cent among 5,179 agents; 40 per cent time and 18 per cent quality among 453 professionals	Observed fact	SEG-4	QJE 140(2) 2025; Science 381(6654) 2023	2026-08-03
Aggregation deflates: no more than about 0.5 to 0.7 per cent TFP over ten years	Observed fact	SEG-4	Economic Policy 40(121), 2025	2026-08-03
Appraisal optimism is systematic and requires explicit correction; audited benefit erosion in digital programmes	Observed fact	SEG-5	HM Treasury Green Book and optimism-bias guidance; NAO HC 575	2026-08-03
Attribution requires an explicit counterfactual	Observed fact	SEG-5	HM Treasury Magenta Book	2026-08-03
Measurement method is governed; official productivity statistics are vintaged	Observed fact	SEG-5	OECD Measuring Productivity manual; BLS programme documentation	2026-08-03
Value realisation balances cost, risk, opportunity and performance	Observed fact	SEG-5	ISO 55000:2024	2024-07
Corridor infrastructure conditions	Observed fact	SEG-4	ITU Facts and Figures 2025; World Bank WDI	2025-11-17
Oversight is a legal duty for credit-scoring systems in scope	Observed fact	SEG-6	Regulation (EU) 2024/1689, Art. 6(2), Annex III 5(b), Art. 14	2026-08-03
Implementation is mutual adaptation with a brief window	Observed fact	SEG-3	Research Policy 17(5) 1988; Organization Science 5(1) 1994	2026-08-03
Fixed parity 655.957 XOF per EUR since 1999	Observed fact	SEG-6	BCEAO	2026-08-02

CLAIM	TYPE	GRADE	SOURCES	VALID AS OF
SaaS configuration cost generally expensed (cross-reference only)	Observed fact	SEG-6	IFRIC agenda decision, March 2021 (IAS 38)	2026-08-03
Scenario arithmetic: net minutes, break-even volume, three-year position	Calculated inference	SEG-5	Stated synthetic inputs; fixed parity	2026-08-03
The six-link traceability chain (house position)	Interpretation	SEG-1	None: firm position, so labelled	2026-08-03
Capacity released is not cost removed (house position)	Interpretation	SEG-2	None: firm position, so labelled	2026-08-03
Kill criteria must be numeric and set at authorisation (house position)	Interpretation	SEG-1	None: firm position, so labelled	2026-08-03

Sources. Eurostat, ICT usage and e-commerce in enterprises, 2025 release published 11 December 2025, and Statistics Explained, use of artificial intelligence in enterprises; United States Census Bureau, Business Trends and Outlook Survey, AI use supplement, releases to May 2026; European Central Bank, Survey on the Access to Finance of Enterprises, ad hoc questions for the fourth quarter of 2025, reported in the Economic Bulletin and in the ECB blog of 24 June 2026; Brynjolfsson, Li and Raymond, "Generative AI at Work", Quarterly Journal of Economics 140(2), 2025, pp. 889-942; Noy and Zhang, "Experimental evidence on the productivity effects of generative artificial intelligence", Science 381(6654), 2023; Acemoglu, "The simple macroeconomics of AI", Economic Policy 40(121), 2025; HM Treasury, The Green Book (2026 edition) and Supplementary Green Book Guidance: Optimism Bias; HM Treasury, The Magenta Book, including Annex A analytical methods; National Audit Office, The challenges in implementing digital change, HC 575, 21 July 2021; OECD, Measuring Productivity: OECD Manual, 2001; United States Bureau of Labor Statistics, productivity programme release and revision documentation; ISO 55000:2024, Asset management: vocabulary, overview and principles; International Telecommunication Union, Measuring digital development: Facts and Figures 2025, 17 November 2025; World Bank, World Development Indicators, access to electricity (EG.ELC.ACCS.ZS); Regulation (EU) 2024/1689 (Artificial Intelligence Act), Official Journal L, 2024/1689, 12 July 2024; IFRS Interpretations Committee agenda decision, Configuration or Customisation Costs in a Cloud Computing Arrangement (IAS 38), March 2021, cited as cross-reference only; Leonard-Barton, Research Policy 17(5), 1988, pp. 251-267; Tyre and Orlikowski, Organization Science 5(1), 1994, pp. 98-118; BCEAO, history of the CFA franc, reused from the estate-verified source pack of STG-PUB-BRIEFING-PRICING-DEFENSIBILITY. All retrieved and verified 3 August 2026 unless otherwise dated. No consultancy or vendor statistic is cited anywhere in this briefing. The worked scenario is synthetic and sourced to no external data. House positions are graded SEG-1 or SEG-2 and labelled as interpretation wherever they appear.