Programme for Handover

An applied-AI research programme has three governance documents. The model write-up describes what was built; the invoice schedule describes how it was funded; the KPI contract describes what had to be measured for the money to move. The first two are published often. The third is the one that decides whether an open-ended research bet stays honest, and it is rarely read. We reconstruct the KPI contract behind a three-phase, roughly twenty-month subsurface computer-vision programme built for a major operator in Oman. Every phase carries a lead KPI and a lag KPI: Phase 1 pairs Time-to-Deploy within the phase window and a Standardization measure (a complete, ready dataset for the selected wells) as leads against reduced Time-to-Market as the lag; Phase 2 is gated on a contractual minimum of at least one fully functioning algorithm and at least one fully functioning model, versioned, with accuracy sound enough to carry into Phase 3; Phase 3 names self-service inference, tested, as its lead indicator, with model validation, operations and performance on distributed and HPC compute as the lags. Between Phase 2 and Phase 3 sits an intermediary handover report that is itself a contractual gate: it had to carry an infrastructure-cost advisory and an API-performance budget covering call duration and latency, so the running cost of the next phase was priced in writing before that phase was entered. A continuous bi-weekly progress cadence with validation, verification and preventive controls ran underneath all of it. The five payment tranches were wired to the gates, Part-1 due at signing, and the production and commercial phases (four and five) were held explicitly out of scope until a final recommendation. We argue that this KPI-to-gate-to-tranche wiring is the actual value-governance instrument of the programme: money follows proven capability rather than effort, and later commercial phases stay out of scope by design until the research proves out. Three interactive instruments walk the argument, and the effort-obligation clause work and the cash-flow ledger of this same programme are pointers here, not re-derivations.

Tarry Singh
Founder & CEO

Tannistha Maiti
Senior AI Researcher

LEAD-KPI / LAG-KPI / KPI-CONTRACT / PHASE-GATE / MILESTONE-TRANCHES / HANDOVER-REPORT

CONTENTS

A board approving an AI research programme is asked to spend against a promise it cannot verify. The model does not exist yet, the accuracy is unknown, and the science may not converge. The usual reflex is to reach for the invoice schedule and the effort-versus-result clause, and those matter. But they answer a narrower question than the one a director actually holds: not *how is this paid* and not *who is on the hook for a result*, but *what will we measure, phase by phase, to know the money should keep moving*. That third document is the KPI contract, and on the programme we read here it was written with unusual care.

This is a governance paper, not a model paper. It reconstructs the KPI contract behind a three-phase, roughly twenty-month subsurface computer-vision engagement built with a major operator in Oman, on a fractured carbonate reservoir, using borehole image logs from FMI and CMI tools. The programme carried a budget on the order of 165,000 OMR across those three phases, and its Phase-3 correlation work ranged across two to eighty wells. The technical core of that programme, the fracture and bedding detectors and the vug work, is written up elsewhere and is not the subject here. The subject is the measurement layer: the lead and lag KPIs attached to each phase, the contractual gate each phase had to clear, the handover report that had to be priced before the next phase could start, and the way all of it wired into the payment tranches so that money followed proven capability instead of effort spent.

It matters that the programme also carried an Omanization commitment, young Omani engineers and researchers trained through the work with the local academic partner, because that is the version of a subsurface-AI programme that a board in the region actually approves: not a black-box capability rented from abroad, but a capability the client's own people can eventually operate. That commitment is not a side note to the KPI contract; it is the reason self-service inference sits where it does in the measurement framework, as we will get to. A capability that leaves with the vendor trains nobody. The KPIs were written to steer toward the opposite outcome.

Two prior chapters of this same programme are worth naming so we do not re-walk them. The cash-flow architecture of the engagement, the mobilisation invoice, the five tranches read as a treasurer would read them, is covered in the money-and-milestones chapter. The clause-level design, the effort obligation rather than a results warranty, the phase-capped liability, the IP split, is covered in the contracting-for-uncertainty and effort-obligations

chapters. We lean on both as settled ground. What follows is the layer above the clauses and beneath the deliverables: the KPIs that told the parties, every two weeks, whether the programme was earning its next tranche.

Two documents govern the money; a third governs the measurement

Read an applied-AI engagement from the outside and you will find two documents doing the visible work. The scope of work names the phases and the deliverables. The payment schedule names the amounts and the dates. Between them they answer what gets built and how it gets paid. Neither of them, on its own, tells you the thing a director most needs to know at each review: has this phase actually produced the capability that justifies releasing the next payment.

On this programme that question had its own contract table, a per-phase KPI and business-impact schedule that sat alongside the scope and the payment terms. It did something the other two documents could not. It named, for each phase, a lead indicator and a lag indicator, and it tied the lead indicators to concrete artefacts that either existed or did not. A lag KPI tells you where you have been. A lead KPI tells you whether the next thing is on track. Governing a research programme on lag KPIs alone is governing by rear-view mirror; by the time reduced Time-to-Market shows up in the numbers, every decision that produced it is already sunk. The point of pairing a lead with a lag at every phase was to give the board something to steer on before the lag arrived.

The literature-review deliverable of Phase 1 was required, in writing, to detail the path to automated deployment with *both* lead and lag KPIs named. That is a small clause with a large consequence. It meant the measurement framework was itself a Phase-1 deliverable, agreed before any model was trained, rather than a reporting convenience negotiated later when the numbers were already in hand and easier to flatter.

Phase 1: a lead you can check today, a lag you cannot

Phase 1 was Dataset Development: literature review, exploratory analysis, the ingestion pipeline, and the assembly of a clean, model-ready dataset for the selected wells. Its KPIs were split cleanly across the lead-lag line.

The lead KPIs were Time-to-Deploy within the Phase-1 window and a Standardization measure, defined as a complete, ready dataset for the selected wells by the end of the phase. Both are checkable on the day. Either the dataset for the selected wells is standardised and complete or it is not; either the phase deployed inside its window or it slipped. Neither requires you to wait for a downstream outcome. The lag KPI was reduced Time-to-Market, and that one you genuinely cannot check in Phase 1, because a market outcome is months and phases away. Pairing them was the design: the lead indicators let the board confirm Phase 1 had produced the substrate that a shorter Time-to-Market would eventually depend on, without pretending the lag itself was already visible.

There is a reason a dataset counts as a governance KPI at all, rather than as invisible plumbing. On this engagement the data arrived late and in pieces, and the contract had anticipated that: delay costs for late data sat with the party that owed the data. Making a standardised dataset an explicit, gate-bearing lead KPI meant the programme could not paper over a data shortfall by moving on to modelling. The dataset gate had to clear on its own terms first, which is exactly where the risk on a subsurface-AI programme actually concentrates.

Phase 2: the gate is a working algorithm and a working model, not a metric

Phase 2 was ML Operations on the image logs: setting up the pipeline, developing algorithms and models, evaluating performance. Its gate is the sharpest clause in the whole KPI contract, and it is worth quoting the shape of it precisely, because it is doing something a headline accuracy target could not.

The Phase-2 minimum was written as *at least one fully functioning algorithm* and *at least one fully functioning model*, with appropriate versioning, and with accuracy sound enough to carry into Phase 3. Notice what that gate refuses to be. It is not "reach F1 of 0.7." A numeric accuracy target written at the start of a research phase is a promise about an unknown, and on genuinely uncertain science it pays perverse incentives, rewarding the safe unfalsifiable claim over the honest experiment. That failure mode is the subject of the contracting-for-uncertainty chapter and we take it as read. The Phase-2 gate sidesteps it by asking for *existence and function* rather than a *level*: a working algorithm, a working

model, versioned, and accuracy the parties jointly judge sound enough for the next phase to build on. The gate is falsifiable without being a hostage to a number nobody could commit to at signing.

That is the difference between a KPI that governs and one that performs. "Working, versioned, sound enough to proceed" is checkable at the review and is honest about the uncertainty; "hit 0.7 by month nine" is a number that either freezes the science into conservatism or gets quietly renegotiated. The gate the parties chose is the one that stays signable and stays honest.

The versioning requirement inside that gate is easy to skip past and worth stopping on, because it is where the gate connects to handover. "Appropriate versioning" of the algorithm and the model is not a hygiene footnote; it is what makes the Phase-2 output reproducible and therefore inheritable. A model that a client cannot rebuild from a pinned version is a model the client does not really own, and the whole programme was pointed at ownership. Requiring versioning at the gate means the thing that clears Phase 2 is not a one-off artefact that ran once on the vendor's machine, but a reproducible build the next phase, and eventually the client, can pick up and re-run. The gate asks for a working algorithm and a working model, and it asks for them in a form that survives being handed over. Existence, function, and reproducibility, judged together, are what release the tranche.

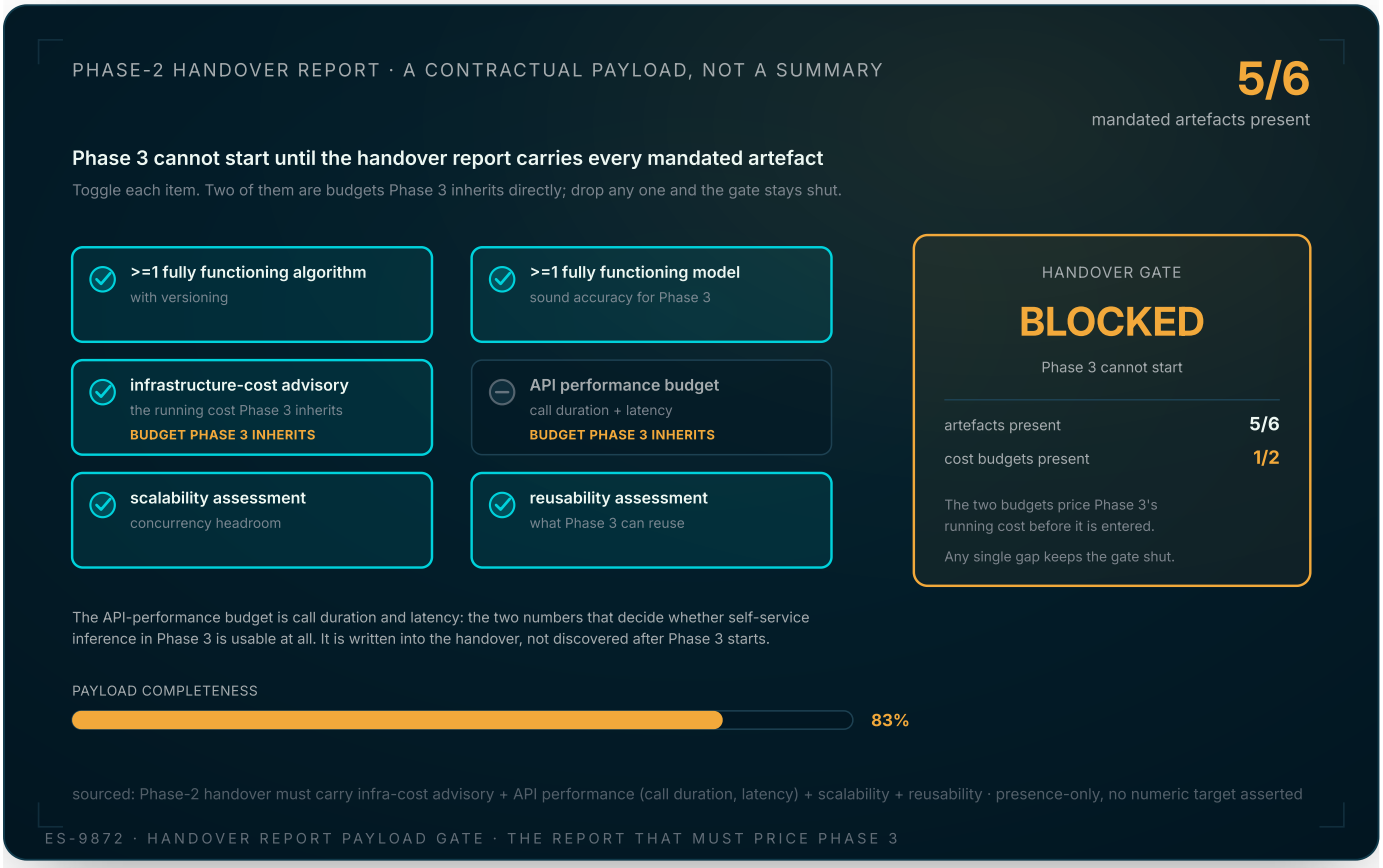
The handover report is a gate, and it prices the next phase before you enter it

Between Phase 2 and Phase 3 the contract placed an intermediary handover report, and this is the clause a board should read twice, because it is the one that most changes how the next phase is governed. The handover report was not a courtesy summary of Phase 2. It was a contractual payload with a required content list, and two items on that list are budgets the next phase inherits directly.

The report had to carry an infrastructure-cost advisory for Phase 3, a scalability assessment, a reusability assessment, and an API-performance budget covering call duration and latency. Read those two budget items together. The infrastructure-cost advisory says what Phase 3 will cost to run. The API-performance budget says how fast the served model responds and how long a call takes. On a programme whose Phase-3 lead

indicator was self-service inference, those two numbers are not reporting decoration; they decide whether self-service inference is usable at all. A self-service inference capability that costs too much to run or answers too slowly is not a capability, it is a demo.

Putting those budgets *inside the handover*, as a precondition for starting Phase 3, meant the running cost and the latency envelope of the next phase were priced in writing before the next phase was entered, rather than discovered halfway through it when the compute bill and the response times arrived. The handover report converts the boundary between two research phases into a contract renewal with a costed prospectus. The board does not approve Phase 3 on faith; it approves it against a written cost-and-latency budget the previous phase had to produce to get paid.



The Phase-2 intermediary handover report treated as a contractual payload rather than a courtesy summary. Six mandated artefacts have to be in the report before Phase 3 can start: at least one fully functioning algorithm and one fully functioning model, an infrastructure-cost advisory, an API-performance budget covering call duration and latency, a scalability assessment, and a reusability assessment. Toggle each artefact present or missing; the orange-or-teal verdict clears only when all six are present, and drop any single one and the gate stays shut. Two of the artefacts are budgets Phase 3 inherits directly, which is the point of the exhibit: the running cost and the latency envelope of self-service inference are priced in writing at handover, before Phase 3 is entered, not discovered after it starts. The mandated payload is sourced from the engagement KPI memo; the exhibit reasons over the presence of each artefact only, and asserts no numeric latency target because none is sourced.

The exhibit makes the mechanism concrete. Six mandated artefacts have to be present before the handover gate clears, and dropping any single one keeps it shut. Two of them, the infrastructure-cost advisory and the API-performance budget, are the ones a treasurer cares about, because they are the ones that price the next phase. The gate does not reason over whether the latency is *good*; no numeric latency target is asserted here, and none is sourced. It reasons over whether the number is *present and in writing at the boundary*. That is the governance property: the cost of the next phase is a deliverable of the previous one.

Phase 3: self-service inference as a lead indicator

Phase 3 was Well-to-Well Correlation and Integration, a correlational task across a range of two to eighty wells, plus AI model integration, rollback and serving. Its KPIs invert the usual instinct about what counts as a leading signal.

The lead indicator for Phase 3 was self-service inferencing, tested. Not model accuracy, not correlation quality, but whether an operator's own people could run inference themselves, self-service, without the research team in the loop. That is named explicitly as a lead indicator for the client's operational and cost footprint in production, and the choice is deliberate. Self-service inference is the thing that predicts whether the capability survives handover. A model that only the vendor can run is a consulting dependency; a model the client can serve itself is an owned capability. Making self-service the *lead* KPI means the programme was steering, from early in Phase 3, toward the client operating the thing without help, which is the only version of "done" that reduces the client's IT-operations overhead and infrastructure cost in the way the business case promised.

The lag KPIs for Phase 3 were model validation, model operations, and model performance on distributed and HPC compute, the trailing confirmations that the served system holds up at scale. And the phase closed on a final recommendation report summarising the project and the way forward. The lead told the programme whether it was heading toward a client-operable capability; the lags confirmed, after the fact, that the capability ran at scale.

The scale itself is worth naming, because it is why the distributed-and-HPC lag KPIs are not boilerplate. Phase 3 was a correlational task across two to eighty wells. Well-to-well correlation is combinatorial in the number of wells, and a method that runs comfortably on a handful of wells can fall over on eighty. Measuring model performance on distributed and HPC compute as a lag KPI meant the programme could not declare the correlation

capability done on a small pilot and quietly assume it would scale; the trailing metric had to show it holding up across the full well range on the compute that Phase 3 would actually use in operation. This is the second place the KPI framework refuses a convenient shortcut. The Phase-1 leads refuse to let modelling start on an incomplete dataset; the Phase-3 lags refuse to let the correlation capability be signed off on a scale it has not been tested at.

There is a quiet ordering here that a director should notice. The lead indicator, self-service inference, is checked first and predicts ownership; the lag indicators, validation and scale performance, are checked later and confirm robustness. A programme that only measured the lags would learn whether the model worked at scale but not whether anyone but the vendor could run it. A programme that only measured the lead would learn whether the client could operate the model but not whether it held up on eighty wells. Pairing them is what makes Phase 3's sign-off mean both things at once: operable by the client, and robust at the scale the client will actually use.

The bi-weekly cadence is the verification layer under all of it

None of these gates verifies itself. A gate written as "at least one working model, sound enough to proceed" or "self-service inference, tested" is a judgement made at a review, and a judgement made at a review is only as good as the evidence trail underneath it. The contract supplied that trail as a standing governance term: continuous bi-weekly project progress reporting, with validation, verification and preventive controls.

Bi-weekly is a specific choice. It is frequent enough that a gate judgement at a phase boundary is the summary of ten or twelve prior checkpoints rather than a single high-stakes examination, and infrequent enough not to become its own overhead. Validation, verification and preventive controls, running every two weeks, mean that by the time a phase gate comes up for sign-off, the parties have already watched the capability accrete in the open. A gate is not a surprise at the end of a phase; it is the moment a trend that both sides have been tracking crosses a line they agreed on in advance. That is what lets an effort-obligation contract, which by design does not promise a result, still give a board something firm to steer on. The result is not guaranteed, but the *evidence about the result* arrives on a fixed cadence, and the gate is the point where accumulated evidence releases money.

The three verbs in that governance term are not synonyms, and reading them apart is worth the moment. Validation asks whether the thing being built is the right thing, whether the dataset, the algorithm, the model actually answers the geological question. Verification asks whether it was built correctly, whether the pipeline does what it claims, reproducibly. Preventive controls ask what could go wrong next and what is in place to stop it, the forward-looking half that keeps a programme from walking into a known failure. A cadence that ran only validation would confirm relevance but miss defects; one that ran only verification would confirm correctness but miss whether the correct thing was even the useful thing; one without preventive controls would keep discovering the same avoidable problem. Requiring all three, every two weeks, is what turns the bi-weekly report from a status update into a governance instrument. On this programme it also gave the parties a paper trail dense enough that a phase gate could be argued from the record rather than from memory, which is precisely the kind of evidence a director wants before releasing the next tranche.

A useful way to see the cadence is as the sampling rate on an uncertain signal. A research programme's true state, is the capability converging or not, is a noisy thing you cannot read continuously. Sample it too rarely and you learn the phase failed only at the gate, too late to correct. Sample it every two weeks, with validation, verification and preventive controls, and you get enough resolution to catch a drift while it is still cheap to fix, without paying for the overhead of watching constantly. The gate is then not a test the programme either passes or fails cold; it is the reconciliation of a trend the sampling has already made visible.

Wiring KPIs to gates to tranches: money follows proven capability

Here is where the three layers, the KPIs, the gates and the payments, resolve into a single instrument. The programme's five payment tranches were structured to mirror the phase gates: Phase 1 in two parts with Part-1 due at signing, Phase 2 as a single tranche, and Phase 3 in two parts. The count and the ordering of those tranches are not incidental. They are the phase gates, expressed as money.

The consequence is the thesis of this paper. A tranche is released when its gate clears, and a gate clears when its phase's KPIs are met. So the money moves when, and only when, the capability the KPIs measure has been proven at a review that the bi-weekly cadence has been building toward for weeks. Part-1 of Phase 1 is the exception, and a deliberate one: it

is due at signing, funding the un-glamorous data and infrastructure work that makes the rest possible before any capability exists to gate against. Every tranche after it follows a proven gate. Effort spent between gates earns nothing on its own; only a cleared gate releases the next payment.

How value is governed in an AI research-and-development programme: three contractual phases, each carrying a lead KPI and a lag KPI that wire into a single deliverable gate, with money released only when the gate clears. Drag the capability-proven lever along the R&D arc and the gates clear left to right: the Phase 1 dataset gate, the Phase 2 working-algorithm-plus-working-model gate, and the Phase 3 self-service-inference gate. The teal step line is the cumulative payout the contract permits at each point; the orange marker is the only element that argues, stepping up the ladder only as a gate is actually met, so money follows proven capability rather than effort. Part-1 of Phase 1 is due at signing by contract; Phases 4 and 5 (production and commercial) stay out of scope until the research proves out and a final recommendation is delivered, with continuous bi-weekly progress reporting verifying each step. The five-tranche structure (Phase 1 in two parts, Phase 2 as one, Phase 3 in two, Part-1 at signing), the gate wording, the lead-and-lag KPI split, and the out-of-scope rule are sourced from the engagement KPI memo and the draft service agreement; the tranche VALUES are shown only as normalised illustrative shares of the R&D fee that sum to 100, never as absolute client fees, which are commercially sensitive.

The ledger above is the argument in one picture. Drag capability along the research arc and the gates clear left to right; the payout steps up only as each gate is actually met, never in between. That step function is the whole governance philosophy rendered as cash flow. Money does not accrue smoothly with time or effort. It sits flat through a phase and then

jumps at the gate, because the contract pays for proven capability at discrete, verifiable moments and for nothing else. The two out-of-scope boxes at the bottom, the production and commercial phases, are the other half of the same discipline, and they deserve their own section.

We can state the wiring compactly. If a phase's lead KPI L_i and lag KPI G_i define whether its gate γ_i has cleared, and each gate carries a tranche of normalised share s_i, then the fraction of the fee released at any point in the programme is a step function of proven capability:

RELEASED SHARE AS A STEP FUNCTION OF CLEARED GATES

Formula | LaTeX

Copy LaTeX

$$\text{Released}(\text{capability}) = \sum_i s_i \cdot \mathbf{1}[\gamma_i(L_i, G_i) \text{ cleared}]$$

The indicator function is the point. There is no partial credit for a gate half-met; the term is either in the sum or it is not. Money is a sum over cleared gates, not an integral over time or effort.

Phases 4 and 5 are out of scope on purpose

The KPI contract named five phases and governed three. Phase 4, production, and Phase 5, commercial rollout, were explicitly out of scope, deferred pending the final recommendation of Phase 3. This is not an omission or a scoping accident. It is the same governance principle applied to the largest decision in the programme.

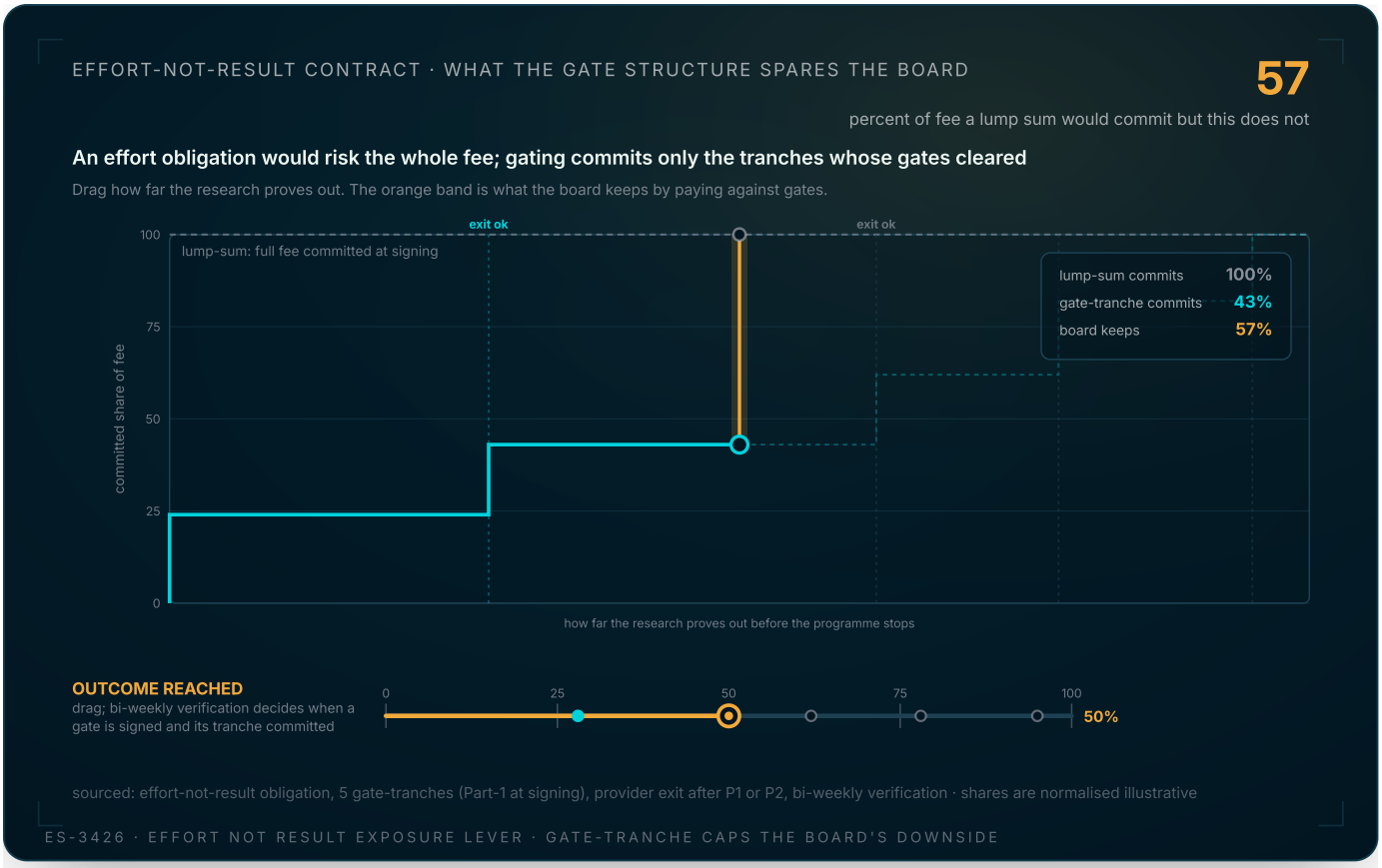
The logic is identical to the tranche logic, one level up. You do not commit to production and commercialisation before the research has proven the capability exists, any more than you release a Phase-2 tranche before the Phase-2 gate clears. Holding Phases 4 and 5 out of scope keeps the board from pre-committing to a build-out and a go-to-market on a capability that Phase 3 has not yet shown to be real, operable, and affordable to serve. The final recommendation report is the instrument that reopens that question, armed with the actual self-service inference result, the actual infrastructure cost, and the actual latency budget from the handover. The commercial phases stay behind a gate that only the completed research can clear.

An engineering read and a board read agree here. The engineering read is that you cannot productionise what you have not validated. The board read is that you do not want optionality you have paid to foreclose; keeping the commercial phases out of scope preserves the choice to stop, pivot, or scale, made against evidence rather than against a plan written before the evidence existed. Out-of-scope, in this contract, is not a gap. It is a preserved option.

The effort obligation is what makes measurement, not promises, the governor

There is one clause without which none of this KPI structure would hold its shape, and it belongs to a chapter we have pointed to rather than re-derived: the engagement was written as an effort obligation, not a result obligation, an explicit research project where the provider is not responsible for achieving a specific result. We take that clause as settled ground. What matters here is what it does to the KPIs.

Because the contract does not promise a result, the KPIs cannot be a covert result guarantee smuggled back in through the measurement layer. That is precisely why the Phase-2 gate is written as existence and function rather than a numeric level, why Phase 1's checkable leads are dataset-ready and in-window rather than an accuracy figure, and why self-service inference, an operational fact, is the Phase-3 lead. Every gate in the contract is checkable without asserting an outcome nobody could commit to. The effort obligation and the KPI contract are two halves of one design: the first says the vendor is not on the hook for a number, and the second says the money still moves only against proven, verifiable capability. Remove the effort obligation and the KPIs would harden into the very result guarantee the science cannot bear.



What the gate-and-tranche structure spares the board when the contract is written as an effort obligation rather than a result guarantee. The agreement is an explicit research project, so the provider is not responsible for achieving a specific result; on its own that wording would put the whole fee at risk on an uncertain outcome. Drag how far the research proves out before the programme stops and compare two contracts: a lump sum commits the full fee at signing regardless (the faint reference line), while the gate-tranche contract commits only the tranches whose gates were actually cleared, plus Part-1 at signing. The teal step is the fee the board has actually committed; the orange band is the spread it keeps, the fee a lump sum would have committed but this structure did not, because payment followed proven capability. The provider's right to exit cleanly after Phase 1 or Phase 2 caps the exposure at a phase boundary, and continuous bi-weekly verification decides when a gate is signed and its tranche committed. The effort obligation, the five gate-tranches with Part-1 due at signing, the exit-after-Phase-1-or-2 right, and the bi-weekly cadence are sourced from the draft service agreement and the engagement KPI memo; the per-tranche values are normalised illustrative shares that sum to 100, never absolute fees.

The exhibit reads that pairing from the board's side. Under a lump-sum result contract the full fee is committed at signing regardless of how far the research proves out. Under the gate-tranche structure, only the tranches whose gates cleared are committed, and the provider's clean exit right after Phase 1 or Phase 2 caps the exposure at a phase boundary. The orange band is what the board keeps by paying against gates instead of against a promise. Drag how far the research proves out and the band is largest exactly where it matters most, in the region where the science might not converge, which is the region a

lump-sum contract would have paid for in full and this structure does not. Bounded downside, verified every two weeks and released only at gates, is what makes an uncertain research programme signable at all.

What a board should take from this KPI contract

The transferable lesson is not the specific KPIs; it is the wiring. Three moves carry it, and a fourth guards it.

Pair a lead KPI with a lag KPI at every phase, and make the lead something you can check at the review, a dataset that exists, a model that runs, an inference an operator can perform, rather than a downstream outcome you can only confirm long after the decision. Write the gate as existence and function, not as a numeric level, so it stays honest on uncertain science and stays falsifiable at the review. Make the intermediary handover carry the next phase's cost and latency budgets, so each phase boundary is a costed renewal rather than an act of faith. Then wire the tranches to the gates, so money is a sum over cleared gates and effort between gates earns nothing on its own. The fourth move guards all three: run a fixed verification cadence, bi-weekly here, so a gate judgement is the summary of many prior checkpoints rather than a single examination, and hold the commercial phases out of scope until the research has cleared its own gates.

A board approving an AI research programme is still spending against a promise it cannot verify at signing. This KPI contract does not remove that uncertainty. It makes the uncertainty pay out in the right direction: every two weeks the parties learn more, at each gate the accumulated evidence releases the next tranche, and the largest commitments, production and commercialisation, wait behind the last gate the research has to clear. Value follows proven capability because the KPIs are the gate, the gate is the tranche, and nothing in the structure pays for effort that has not yet turned into something you can measure.

Limitations

This is a reading of one programme's KPI contract, not a template. The specific lead-lag pairings, the wording of the Phase-2 gate, and the handover payload were fitted to a subsurface computer-vision engagement on a fractured carbonate reservoir with a data-supply risk that sat with the client; a programme with different risk concentration would

place its gates differently. The tranche values shown in the interactive ledger and exposure exhibits are normalised illustrative shares that sum to one hundred, never absolute client fees, which are commercially sensitive and are not surfaced anywhere in this paper; the tranche count and ordering, the gate wording, the lead-and-lag split, the handover payload, the bi-weekly cadence, and the out-of-scope rule are sourced from the engagement's KPI and business-impact schedule and its draft service agreement. No numeric latency or accuracy target is asserted, because none is sourced; the handover exhibit reasons over the presence of a written budget, not its value. The effort-obligation clause, the phase-capped liability, the IP split, and the detailed cash-flow ledger of this same programme are treated as settled ground and pointed to rather than re-derived; readers who want those layers should read the contracting-for-uncertainty, effort-obligations and money-and-milestones chapters. Finally, this paper governs the measurement layer only; whether the underlying detectors met their scientific bar is a separate question answered in the technical write-ups.

EARTHSCAN

Lead KPIs, Lag KPIs, and a Bi-Weekly Verification Cadence: Governing an AI R&D Programme for Handover

Published April 2026. Generated 2026-07-22.
https://www.earthscan.io/whitepapers/lead-lag-kpis-biweekly-verification-cadence-ai-rd-handover