

Month Applied-AI Engagement: How the Money and Milestones Were Structured

A multi-phase applied-AI research contract is usually remembered for its deliverables. The document that decides whether the programme is fundable is the invoice schedule, and it is rarely read. This whitepaper reconstructs the cash-flow architecture of one such contract: a three-phase, roughly twenty-month engagement with a major operator in Oman, approved at a total of 533,000 US dollars. We trace the money along two axes. Along time, the total is paid as five milestone-gated tranches, 130,000, 104,000, 104,000, 104,000 and 91,000 US dollars, the first of them a 28 percent mobilisation invoice paid at kickoff to clean and prepare the data infrastructure and cloud before the first model runs, with each subsequent tranche released under a 96-hour client approval SLA and a phase-0 runway sitting ahead of phase 1. Along category, the same total resolves into seven budget lines, a 193,485 US dollar collaboration fee and 169,865 in salaries, scholarships and overhead together carrying more than two thirds of the contract, against 157,300 for all cloud and capital combined. We argue that these two structures are the actual risk-management instrument of the engagement: the mobilisation invoice funds the un-glamorous data plumbing that decides whether the science is even possible, the milestone gates convert an open-ended research bet into a sequence of small, cancellable ones, the SLA caps the schedule cost of the client's own review latency, and the people-heavy ledger names where the value and the risk really sit. The prior chapters of this programme, the manday-level pricing of a single phase and the year-two technology-transfer agreement, are pointers here rather than re-derivations; the subject is how the whole twenty months was funded.

Tarry Singh

Founder & CEO

Narendra Patwardhan

Research Collaborator

ENGAGEMENT-ECONOMICS / CASH-FLOW / MILESTONE-GATED-PAYMENTS /
MOBILISATION-INVOICE / APPROVAL-SLA / PHASED-DELIVERY

CONTENTS

01	The 533,000 total is the most inert number in the contract
02	Along time: five tranches, and the one that leads the work
03	Why gating five tranches turns one bet into five cancellable ones
04	Along the SLA: capping the latency the buyer controls
05	Along category: reading the ledger for where the risk really sits
06	The formula the whole structure is protecting against
07	What this structure is not, and where it can fail
08	Reading a research contract as a buyer
09	Limitations
10	References

A research contract is two documents wearing one cover. The first is the scope: the phases, the deliverables, the science that will be attempted. The second is the payment schedule, and it is the one almost nobody reads and the one that actually decides whether the programme happens. Scope tells you what a vendor promises to try. The schedule tells you when the money moves, what has to be true before each transfer, and who is carrying the risk in the gap between a payment and the result it was meant to buy.

We are writing as the people who ran the technical side of a multi-phase applied-AI engagement with a major operator in Oman and sat close to how it was funded. The published record of this programme already covers the science at length, the set-prediction fracture and bedding detector, the vug-quantification pipeline, the well-to-well correlation, and it has begun to cover the commercials too. One earlier whitepaper reconstructed how a single research phase was priced from a manday ladder and a rate card. Another read the year-two services agreement that handed the capability over. This paper sits between and beneath both. It is about the founding contract for the whole twenty months, and specifically about its cash-flow architecture: the total, how it was cut into milestone-gated tranches over time, and what those tranches were spent on. For a treasurer, a procurement lead, or an investor deciding whether an applied-AI programme is a fundable shape, this is the anatomy that matters.

The 533,000 total is the most inert number in the contract

The approved total was 533,000 US dollars, across three phases and roughly twenty months. Stated alone, that figure tells you almost nothing useful. It is a size, not a structure. A buyer cannot tell from 533,000 whether the programme is front-loaded or back-loaded, whether the vendor is exposed or the client is, or whether a bad first phase can be stopped before it drains the budget. All of that lives in how the number is cut, and it is cut two ways at once: along time, into a payment schedule, and along category, into a budget ledger.

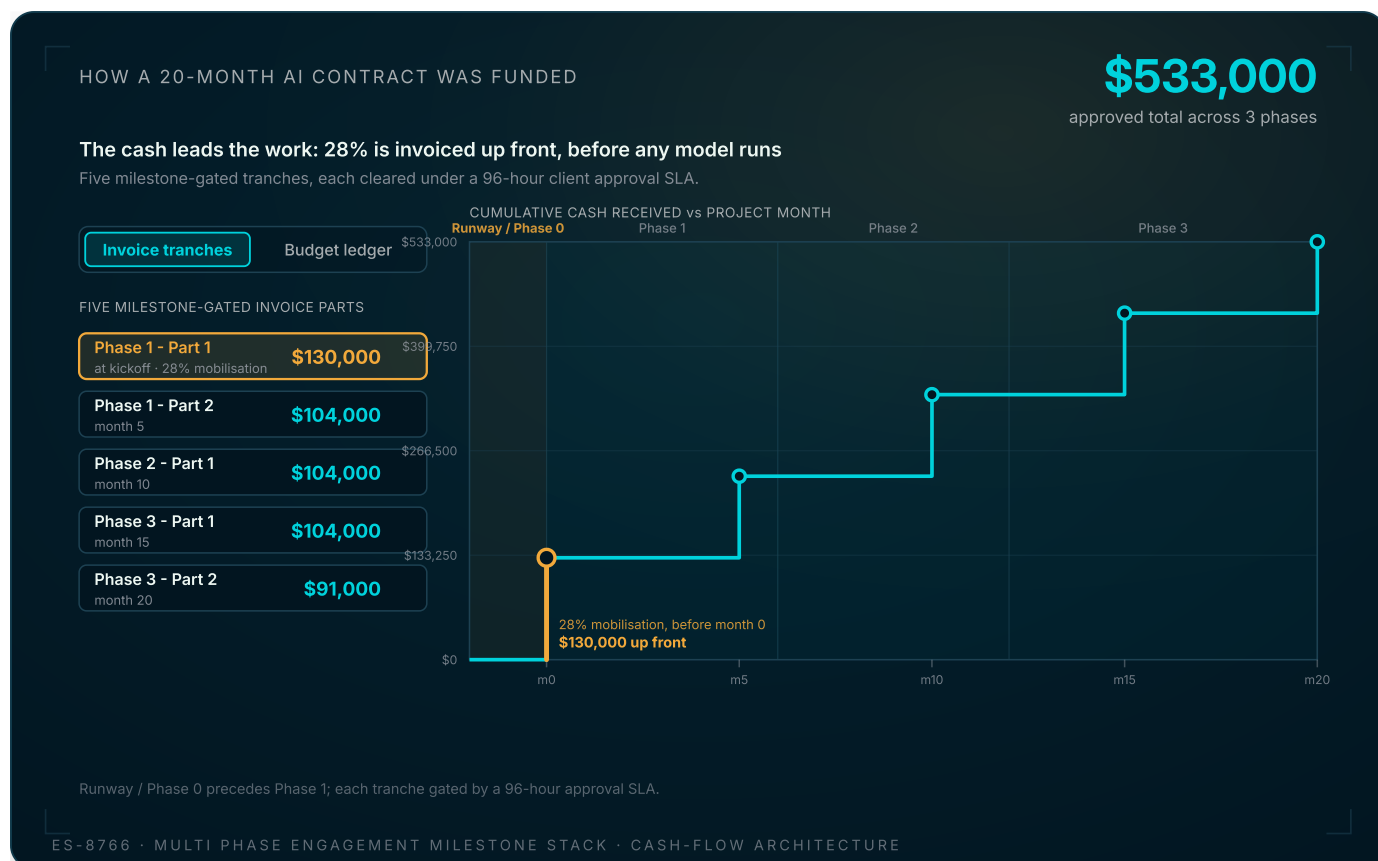
The temptation with any headline contract value is to reason about it as a lump. That instinct is exactly backwards. The total is the one quantity in the contract that is deliberately inert. It does not move, it does not gate anything, and it carries no information about who is at risk when. Everything that makes the contract a risk-management instrument rather than a wager is in the two decompositions, and the rest of this paper is about reading them.

One clarification before we go further, because the archive holds two currencies. An internal planning deck framed the programme against a 165,000 Omani rial budget line, and the earlier pricing whitepaper worked that rial structure in detail. This paper works the US dollar contract figure from the founding proposal, 533,000, and its own seven-line ledger. The two are different denominations of the same engagement seen through different documents; we do not reconcile them here, and where the rial-denominated manday build is the right lens, that earlier piece already owns it.

Along time: five tranches, and the one that leads the work

The 533,000 total was scheduled as five milestone-gated invoice parts. Phase 1 was split in two: a 130,000 US dollar part invoiced up front at kickoff, and a 104,000 part due around month 5. Phase 2 opened with a 104,000 part due around month 10. Phase 3 carried the last two, a 104,000 part around month 15 and a final 91,000 part at close, around month 20. Five payments, front-loaded, each tied to a milestone rather than to a calendar tick.

The word that does the work in that schedule is *mobilisation*. The first 130,000, framed in the proposal as a 28 percent mobilisation payment, was not an advance against future science. It was earmarked, explicitly, to clean and prepare the data infrastructure and the cloud before any model work began. That distinction is the whole argument of the payment structure. In an applied-AI programme on real subsurface data, the work that decides whether the science is even possible is not the modelling. It is the un-glamorous plumbing: pulling binary image logs out of the operator's systems, normalising imaging tools whose value ranges differ by orders of magnitude, standing up the compute, and building the ingestion path that turns a 1.5 GB raw log into trainable patches. That work has to be paid for before it can be done, and it has to be done before anyone can honestly promise a model metric. A mobilisation invoice funds the data plumbing that everything else stands on.



The cash-flow architecture of a three-phase, roughly twenty-month applied-AI research contract with a major operator in Oman. The staircase plots cumulative cash received against project month: five milestone-gated invoice parts of 130,000, 104,000, 104,000, 104,000 and 91,000 US dollars, summing to the approved total of 533,000. The first riser, drawn in orange, is the load-bearing fact: a 130,000 US dollar mobilisation invoice, framed in the proposal as a 28 percent mobilisation payment, paid at kickoff before any model work begins, to clean and prepare the data infrastructure and cloud. A phase-0 runway precedes phase 1, and every tranche clears under a 96-hour client approval SLA. Toggle the left column to read the same total as its seven-line budget ledger, where the collaboration fee and salaries dominate and compute is a smaller line than either. All figures are sourced from the founding engagement proposal; only the continuous month axis between milestones is drawn for legibility rather than asserted as a payment date.

Read the staircase above and the shape of the deal is legible in one glance. Cash climbs to nearly a quarter of the total before month zero, then steps up at each gate. This is the opposite of a milestone-only contract where the vendor funds the entire first phase from its own balance sheet and gets paid only on delivery. Here the buyer accepts more early exposure in exchange for something concrete: a vendor that is capitalised to do the data preparation properly rather than one cutting corners on plumbing to protect its own cash. The mobilisation invoice is a choice about *who* carries the setup risk, and the operator chose to carry it, deliberately, because the alternative is a first phase built on a rushed foundation.

There is a second, quieter piece of structure in the schedule: a phase-0 runway. Before phase 1 formally starts, the contract allows a setup and governance window, the period in which the mobilisation money is actually spent. It does not appear as a payment of its own, but it appears as time, and it matters because it is the interval where the programme's fate is largely decided. If the data does not arrive, if the imaging tools defeat normalisation, if the compute is not standing, the whole twenty months is compromised regardless of how good the eventual model is. The runway is the contract acknowledging that the first risk to retire is not a modelling risk at all.

Why gating five tranches turns one bet into five cancellable ones

Splitting a total into five tranches is not, by itself, de-risking. Anyone can chop a lump sum into five. What makes the structure a risk-management instrument is that each tranche is *gated*: it is released against a milestone, and the client signs off before the cash moves. That converts one large, open-ended research bet into a sequence of smaller, cancellable ones.

The mechanism deserves to be stated plainly, because it is easy to wave at and hard to design. At each gate, the vendor submits a deliverable, the client reviews it, and the next tranche is contingent on acceptance. If phase 1 fails to convince, phase 2 does not automatically fund. The buyer's downside on a programme that goes wrong is not the whole 533,000; it is the tranches already released plus the one in flight. For a research engagement, where the honest expected outcome is a distribution and not a point, this is the difference between a fundable structure and an un-fundable one. A board can approve a bet it can stop. It struggles to approve a bet it cannot.

Milestone gates are common enough that the interesting question is not whether to have them but what they cost. A gate is a stop point, and a stop point where cash is contingent on sign-off puts the client's review latency directly onto the critical path. The vendor can finish a deliverable on a Friday, but if approval takes three weeks, three weeks of schedule are gone, and on a fixed twenty-month clock that latency is expensive. Left uncapped, the client's own slowness silently lengthens the programme the vendor is being judged on.

Along the SLA: capping the latency the buyer controls

The contract's answer to that is a 96-hour client approval SLA. Every milestone sign-off is contracted to happen within 96 hours. Read as a legal clause, it is unremarkable boilerplate. Read as a schedule instrument, it is doing real work: it caps the schedule cost of the one input the vendor cannot control, the buyer's own review speed, and it does so at a known, budgetable four days per gate.



The 96-hour client approval SLA read as a schedule-risk instrument. Because cash moves only on sign-off, each of the five milestone gates puts client review latency directly on the critical path. Under the contracted cap, review consumes a fixed four days per gate and twenty days across the whole engagement, drawn as the teal bar. Drag the lever to lift the uncapped baseline, an open-ended review round that could drift to twenty days per gate, and the orange bar fans out; the dashed bracket between the two bars is the schedule risk the SLA retires. The 96-hour cap and the count of five milestone gates are sourced from the founding engagement proposal; the uncapped-latency baseline the reader sweeps is an illustrative comparison, not a contracted number, and the day counts are calendar days, not adjusted for working-day calendars.

The comparison above is the point. Under the 96-hour cap, review across all five gates consumes a fixed twenty days, a number a project plan can absorb without flinching. Without the cap, that same review is an open-ended hole, and the bar fans out as latency drifts. The gap between the two bars is not a nicety. It is schedule risk that the SLA retires,

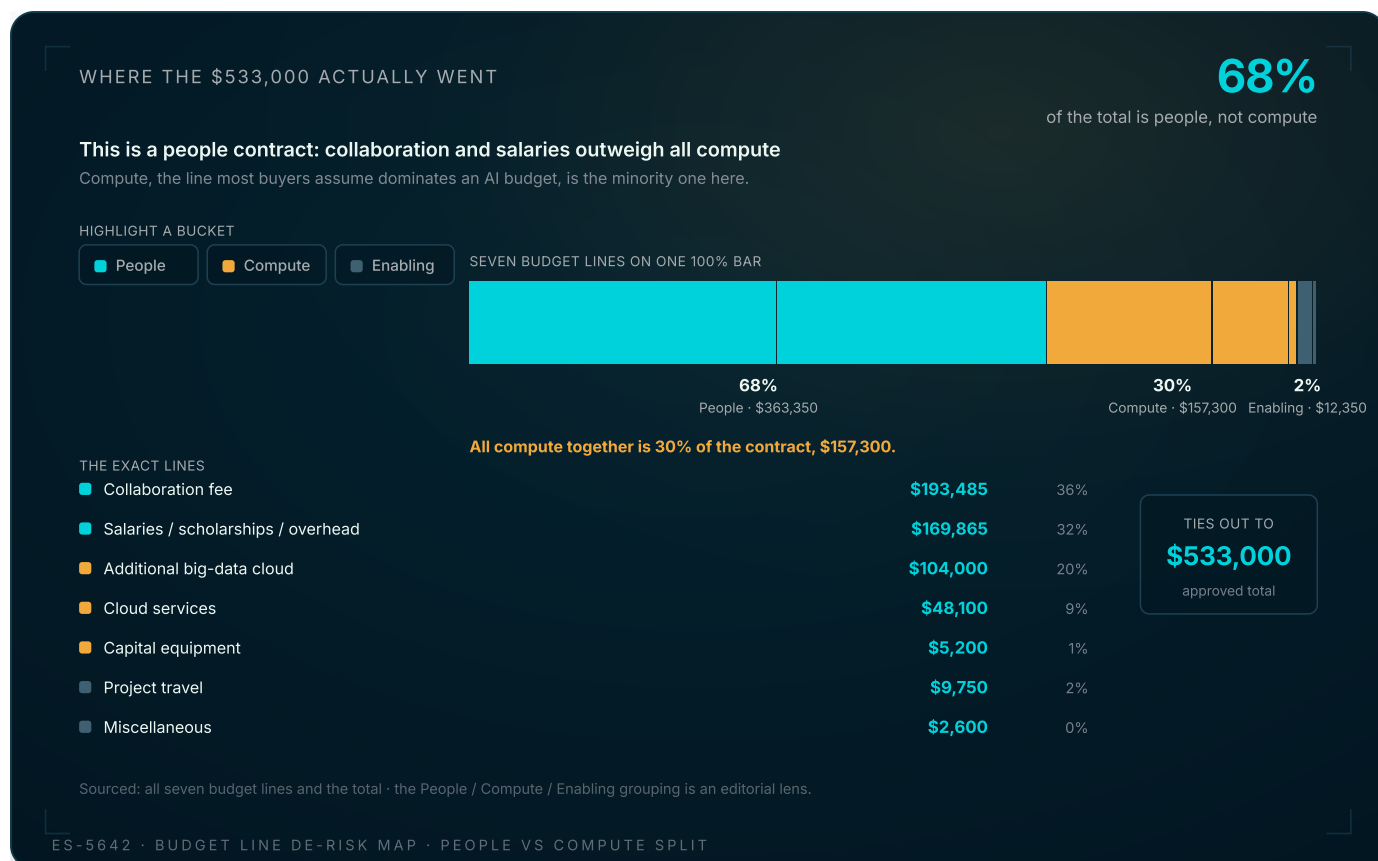
and it is retired symmetrically: the SLA protects the vendor from a client who sits on deliverables, and it protects the client by forcing the vendor to package each milestone into something reviewable in four days rather than a sprawling artefact that cannot honestly be judged that fast. A 96-hour SLA is a two-sided discipline. It says the vendor will produce gate-shaped deliverables and the client will act on them promptly, and it puts a bound on the cost of the handshake between them.

It is worth being precise about what the SLA does not do. It does not guarantee the deliverable is good, and it does not guarantee acceptance. A client can review inside 96 hours and reject. What the SLA guarantees is that the *decision* is fast, which is the part that sits on the critical path. The quality of the deliverable is governed elsewhere, by the scope and the acceptance criteria; the SLA governs only the latency of the verdict. Conflating the two is a common misreading of approval clauses, and it leads buyers to over-value an SLA as a quality lever when it is really a schedule lever.

Along category: reading the ledger for where the risk really sits

The payment schedule tells you when the money moves. The budget ledger tells you what it buys, and it is where the character of the engagement is most exposed. The same 533,000 resolves into seven lines: a collaboration fee of 193,485 US dollars, salaries, scholarships and overhead of 169,865, additional big-data cloud services of 104,000, cloud services of 48,100, project travel of 9,750, capital equipment of 5,200, and miscellaneous of 2,600.

Add those up by what they de-risk and the shape is unambiguous. People, the collaboration fee plus salaries and scholarships and overhead, come to 363,350, more than two thirds of the contract. All compute together, the big-data cloud plus cloud services plus capital equipment, comes to 157,300, under a third. Travel and miscellaneous account for the rest. This is a people contract, not a compute contract, and the ledger says so in the plainest possible terms.



Where the 533,000 US dollar contract total actually went, grouped by what each line de-risks. The seven sourced budget lines sit on a single 100 percent bar and roll up into three buckets: People, which is the collaboration fee plus salaries, scholarships and overhead and comes to more than two thirds of the total; Compute, which is the big-data cloud, cloud services and capital equipment together; and Enabling, which is travel and miscellaneous. The load-bearing fact, called out in orange, is that all compute combined is a minority of the contract, well under a third, against the common assumption that compute dominates an applied-AI budget. Highlight any bucket to isolate it. Every dollar figure and the total are sourced from the founding engagement proposal; the three-bucket grouping is an editorial lens laid over the sourced lines, and the percentages are rounded to whole numbers.

The instrument above lays those seven lines on a single bar, and the counter-intuitive fact is the one worth sitting with. The line most buyers assume dominates an applied-AI budget, compute, is the minority line here. The expensive, scarce, risk-bearing input is human: the research staff who build and validate the models, and the collaboration fee that funds the partnership through which they are engaged. Compute is real and it is not trivial, but it is a smaller line than either of the two people lines on its own.

Two consequences follow from that ledger, and both are decision-relevant for a buyer. First, the salaries-and-scholarships line is doing double duty. It is not only paying to get the work done; it is funding local capability, the training of local researchers and students through the university partnership, an Omanization dividend that sits inside the labour line rather

than in a separate programme. The money that de-risks the science is the same money that builds durable local skill, which is a large part of why a national operator funds an engagement this way at all.

Second, the compute proportion tells the buyer something about scalability that the headline total cannot. If most of a contract is people and the model, once trained, serves every subsequent log at a low marginal compute cost, then the expensive part of the programme is the one-time build, not the ongoing run. That is the economic argument the earlier operating-model and unit-cost pieces make in detail, and this ledger is its upstream evidence: the money went where the one-time risk was, into people, and only a minority went into the compute that the eventual served model amortises across every scan. We treat that downstream unit-cost story as a pointer rather than re-derive it; the point here is only that the ledger predicts it.

The formula the whole structure is protecting against

There is a way to make the logic of the schedule exact, because a research programme has a governing inequality and the payment structure exists to keep the buyer on the right side of it. The value a buyer should be willing to commit at any gate is bounded by the expected value of what the programme delivers, discounted by the probability that it delivers at all, minus what has already been spent. Written compactly, the decision at each gate is whether the forward expectation still clears the forward cost:

THE GATE DECISION, AT EACH MILESTONE

Formula

LaTeX

Copy LaTeX

$$\mathbb{E}[V \mid \text{phase } k \text{ passed}] \cdot P(\text{success} \mid k) > \sum_{j>k} \text{tranche}_j$$

The milestone structure is what makes this inequality evaluable rather than notional. Every gate updates the left-hand side: passing phase 1 raises the estimated probability of success and sharpens the estimate of value, and the buyer re-checks whether the remaining tranches on the right are still worth committing. A lump-sum contract has no such update; the buyer commits the whole right-hand side once, at the start, when the probability estimate is at its most uncertain. The gated schedule lets the buyer spend in step with what

has been learned, which is the entire point of de-risking a research bet by phase. The mobilisation invoice is the one deliberate exception, money committed before the first update, and it is defensible precisely because the thing it funds, the data plumbing, is a prerequisite for generating any update at all.

What this structure is not, and where it can fail

It would be dishonest to present this as a template that de-risks by itself. The structure has real failure modes, and naming them is part of reading it properly.

The mobilisation invoice front-loads the buyer's exposure. Nearly a quarter of the contract is committed before a single milestone is passed. If the vendor is the wrong vendor, that money is the least protected in the whole schedule, because it is spent before the first gate can catch anything. The mobilisation invoice trades the vendor's setup-cash risk for the buyer's early-exposure risk, and that is only a good trade if the buyer has done the diligence to be confident in the vendor before signing. It is a structure that rewards a well-chosen partner and punishes a poorly-chosen one more than a milestone-only contract would.

The gates are only as good as the acceptance criteria behind them. A milestone gate with vague criteria is theatre: the client cannot honestly reject, so the tranche releases regardless, and the cancellability that justified the structure evaporates. The gate de-risks only if failing it is a real possibility, which means the deliverable for each phase has to be specified sharply enough that a 96-hour review can reach a genuine verdict. A soft gate is worse than no gate, because it looks like protection and is not.

The SLA cuts both ways on quality. Forcing a 96-hour decision is a schedule discipline, but it can pressure a client into rubber-stamping a deliverable they have not truly absorbed, especially a technical one. The cap protects the schedule at some risk to the depth of review, and a buyer who leans on the SLA to move fast can find they have approved something they did not understand. The mitigation is not to weaken the SLA but to invest in making deliverables reviewable, which is a cost that lands on the vendor and belongs in the scope, not the schedule.

And the people-heavy ledger concentrates key-person risk. When two thirds of a contract is human, the programme's fate rides on a small number of researchers. That is efficient and it builds local capability, but it means the loss of a key person is a larger shock to a contract like this than to a compute-heavy one. The ledger that makes the engagement a good capability-building instrument also makes it fragile to attrition, which is one more reason the year-two agreement that follows is written to transfer the capability out of individuals and into the client's own team.

Reading a research contract as a buyer

Pull the pieces together and the engagement's commercial architecture is a coherent argument, not a collection of clauses. The mobilisation invoice funds the data plumbing that decides whether the science is possible. The milestone gates turn one open-ended bet into five cancellable ones and let the buyer spend in step with what each phase reveals. The 96-hour SLA caps the schedule cost of the buyer's own review latency at a budgetable four days a gate. The phase-0 runway names the setup risk as the first to retire. And the seven-line ledger, two thirds people and under a third compute, says where the value and the risk actually sit and, not incidentally, funds the local capability the operator wanted to build.

For a buyer looking at a proposal for applied AI, the decision criteria fall out of that architecture directly:

- Is there a mobilisation payment, and is it earmarked for the data and infrastructure work that gates whether the science is even possible, rather than being a generic advance?
- Are the payments milestone-gated with acceptance criteria sharp enough that a gate can genuinely be failed, so the structure is cancellable and not theatre?
- Is there a cap on the buyer's own approval latency, so the client's review speed does not silently lengthen the schedule the vendor is judged on?
- Does the budget ledger match the risk profile of the work: people-heavy where the risk is the one-time build, compute-light where a served model amortises the run?
- Does the labour line do double duty, funding delivery and durable local capability at once, or is capability-building bolted on as a separate cost?

This engagement was structured to answer yes to all five, and the structure is transferable even though the numbers are specific. The point is not that every applied-AI contract should total 533,000 or invoice 28 percent up front. It is that the payment schedule and the

budget ledger are the real risk-management instrument of a research engagement, and reading them as such, rather than reading only the scope and trusting the total, is how a buyer tells a fundable programme from a hopeful one.

Limitations

This is a reconstruction from the founding engagement proposal and the surrounding archive, anonymised, and it carries the limits of that provenance. The 533,000 US dollar total and its seven-line ledger, the five-tranche invoice schedule, the 28 percent mobilisation framing, the 96-hour approval SLA, and the phase-0 runway are all sourced from the proposal; the People, Compute and Enabling grouping of the ledger is an editorial lens we laid over the sourced lines, and the month positions of the tranches are the proposal's planned milestones, not a record of when each invoice actually cleared. The proposal also predates the programme's real start, which slipped later than planned and ran a different technical core than the one first scoped, so the schedule described is the intended architecture rather than a retrospective audit of realised cash flow. The 165,000 Omani rial planning figure that appears elsewhere in the archive is a separate denomination of the same engagement and is not reconciled to the dollar figure in this paper. The uncapped-latency baseline in the SLA instrument is an illustrative comparison, not a contracted number. None of the arguments here should be read as disclosing the client, the field, or the commercial terms of any specific named counterparty; they are offered as a transferable reading of how applied-AI research can be priced and de-risked, grounded in one real contract.

References

Proposal V3.0, founding engagement proposal, March 2021. The source document for the total, the seven-line budget ledger, the five-tranche invoice schedule, the 28 percent mobilisation payment, the 96-hour approval SLA, and the phase-0 runway. Held in the engagement archive; anonymised here.

Anatomy of a Multi-Phase, 20-Month Applied-AI Engagement: How the Money and Milestones Were Structured

Published March 2026. Generated 2026-07-22.

<https://www.earthscan.io/whitepapers/anatomy-of-a-multi-phase-20-month-applied-ai-engagement-money-and-milestones>

Copyright EarthScan. All rights reserved.