

EARTHSCAN

WHITEPAPER 18 PAGES / MAY 2026

Own Your AI Stack: A Whitepaper for Energy Operators

AI in oil and gas will hit \$17B by 2030, yet half of enterprise pilots never ship. The fix is engineering execution, not better models.

Tannistha Maiti

Senior AI Researcher

AI-STRATEGY / SUBSURFACE-AI / FOUNDATION-MODELS / HULDE / FORMATION-EVALUATION / MLOPS

CONTENTS

01	Executive summary
02	The opportunity
03	The technical landscape
04	Our approach
05	Case examples
06	Implementation roadmap
07	Risk and mitigation
08	Conclusion and next steps
09	References

The AI in oil and gas market is projected to grow from \$7.2B in 2025 to \$17B by 2030 — yet up to half of all enterprise AI projects never reach production. The cause is not model quality. It is the missing engineering layer between a notebook that works and a workflow that ships.

Executive summary

The energy industry has spent five years running AI pilots with a predictable outcome: proofs-of-concept that succeed in controlled conditions, then stall before production. This is not a technology failure. It is a structural one. Operators have acquired point tools, depended on vendors with conflicting incentives, and handed proprietary subsurface intelligence to closed cloud platforms. The result is a portfolio of decks, a few licensed dashboards, and very little production AI on the asset.

Three forces make 2025 decisive. The AI in oil and gas market stands at \$7.2 billion and is forecast to reach \$17 billion by 2030 at an 18.5% CAGR — operators who cannot deploy are ceding ground now. National oil companies are making sovereign AI a board-level priority: ADNOC's ENERGYai programme cut seismic model-build times by 75% in 2024. And the IEA's Energy and AI Special Report (2025) confirmed that missing internal expertise — not data, not compute — is the dominant adoption barrier across the sector.

EarthScan's thesis is direct: own your AI stack or fall behind. We are an AI engineering company purpose-built to close the gap between pilot and production — six capability layers from HPC to deployment, five formation evaluation modules already running on operator data, and a three-phase engagement model (Discover, Pilot, Scale) that takes a single asset from readiness assessment to production AI in under twelve weeks, with the operator retaining full data sovereignty and platform ownership throughout.

WHAT THIS WHITEPAPER ARGUES

- 1 AI in oil and gas grows from \$7.2B (2025) to \$17B (2030) — a 2.4x expansion in five years.
- 2 Up to half of enterprise AI projects never reach production; the bottleneck is engineering execution, not model quality.
- 3 Vendor-platform models that ingest your subsurface data convert operator IP into vendor leverage.
- 4 EarthScan's full-stack model — Hulde foundation model, five formation-evaluation modules, agentic workflows — gets a single asset to production in 12 weeks.

The opportunity

The market signal is unambiguous. AI spending in oil and gas is compounding faster than almost any other category of operational technology investment in the sector. Upstream applications — exploration, reservoir characterisation, drilling, production optimisation — account for the dominant share, reflecting where the highest-value, hardest-to-replicate workflows sit. This is not a horizontal IT story. It is a story about who owns the intelligence layer on top of the world's most expensive industrial data.

MARKET CONTEXT

\$7.2B

AI in oil & gas market, 2025

\$17B

Projected market, 2030

18.5%

CAGR through 2030

61%

Share of market held by upstream

The pull is no longer experimental. Devon Energy reported 25% productivity gains from AI drilling optimisation at CERAWeek 2025. Industry-wide, predictive maintenance is cutting unplanned downtime by up to 30%. ADNOC's ENERGYai programme — built in collaboration with hyperscalers and national partners — compressed seismic model-build times by 75% in its first year of deployment. These are not pilots. They are budget items competing with capital projects on the same return hurdles.

And yet, the IEA's Energy and AI Special Report (2025) found that the single biggest barrier to scaled adoption is not data availability and not compute access — it is the missing internal capability to take a working model and run it as an operational service. McKinsey's Global Energy Perspective (2025) reaches a similar conclusion from a different angle: digital transformation in the sector is advancing at roughly half the pace required to meet operators' own stated targets. The constraint is execution, not ambition or capital.

THE CONSTRAINT IS NOT CAPITAL

Operators are not under-investing in AI. They are under-investing in the engineering layer that converts AI into production workflows. That is the gap this whitepaper addresses.

The technical landscape

To understand why pilots stall, it helps to look at what production AI in an asset team actually requires. A working model is roughly 15% of the journey. The other 85% is the engineering stack underneath it: high-performance compute that can run seismic-scale inference; data engineering that can ingest fragmented LAS files, raster well-logs, and incompatible SEG-Y variants; a unification layer that joins subsurface, operational, and sensor data on a common ontology; an AI/ML layer that hosts trained domain models with versioning and drift monitoring; an agent layer that orchestrates upload, run, visualise, and report; and a deployment platform that delivers all of it inside the operator's security perimeter.

VI

inside the operator's security perimeter; full ownership transfer at scale.



SPEED ·

ACCURACY ·

V

Upload → run → visualise → report. Agentic orchestration of multi-step subsurface workflows.



SPEED ·

ACCURACY ·

IV

Include foundation model + five domain modules. Versioning, drift, MLOps.



SPEED ·

ACCURACY ·

III

Subsurface, operational, and sensor data on a single ontology.



SPEED ·

ACCURACY ·

II

Ingestion + provenance for LAS, SEG-Y, raster logs, image logs.



SPEED ·

ACCURACY ·

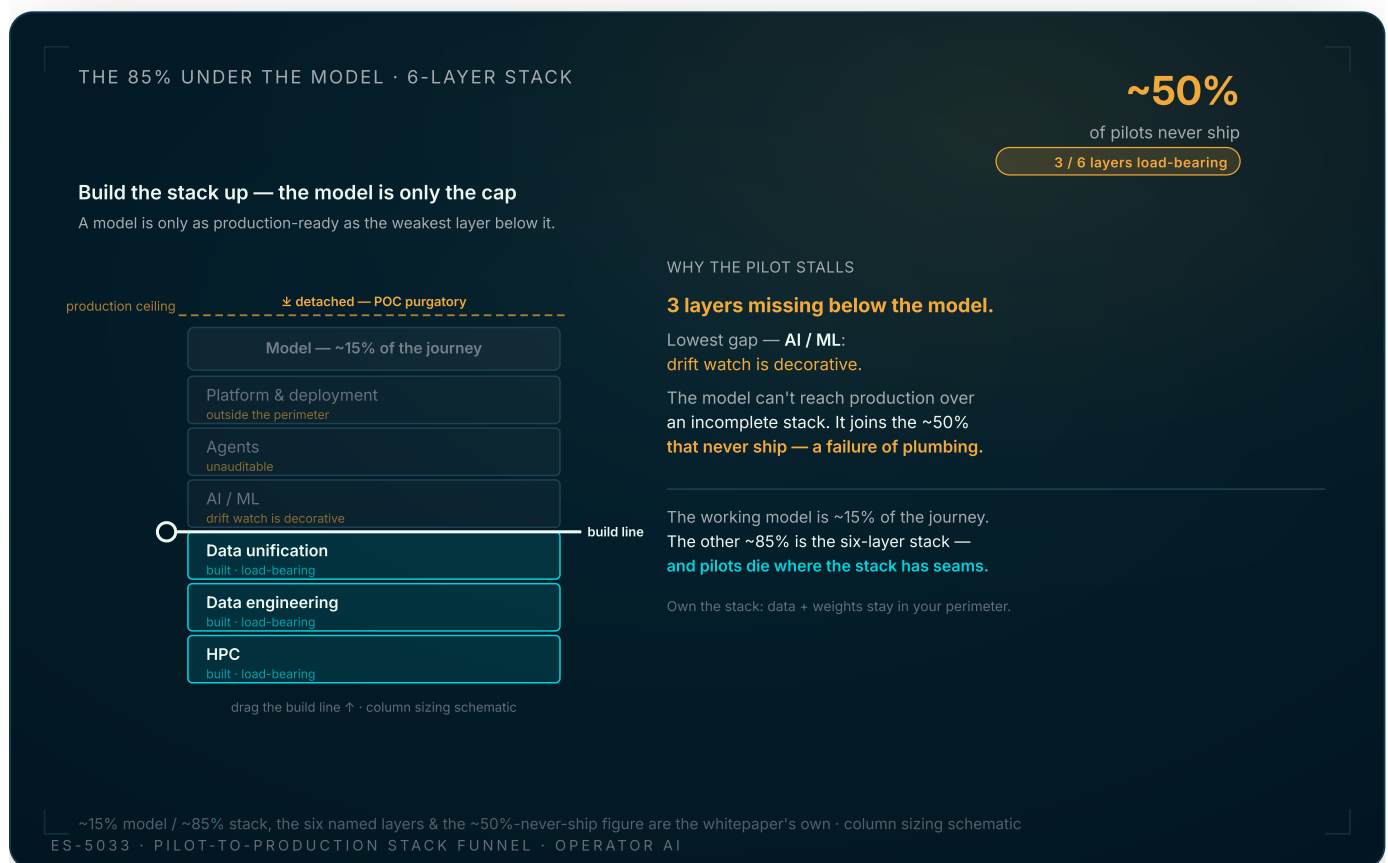
I

Seismic-scale compute provisioned for inference and retraining.





Most operators have none of these layers built to production grade. They have a data lake, a notebook environment, and a vendor or two. A model that performs beautifully on a curated 200-well training set fails on the operator's actual 4,700-well portfolio because the LAS headers are inconsistent, the raster logs were never digitised, and the SEG-Y was acquired across three decades by four different contractors. This is not a failure of machine learning. It is a failure of plumbing.



Pilots don't stall because the model is weak. The working model is only ~15% of the journey; the other ~85% is a six-layer engineering stack (HPC → Data engineering → Data unification → AI/ML → Agents → Platform/deployment), and a project ships only when every layer below the model is built to production grade. Drag the build line up the load-bearing column: with all six built the model reaches the production ceiling; with any gap below it the model detaches into POC purgatory — the ~50% that never ship. The ~15%/~85% split, the six layers and the ~50% figure are the whitepaper's own; the equal-sixths column sizing is schematic.

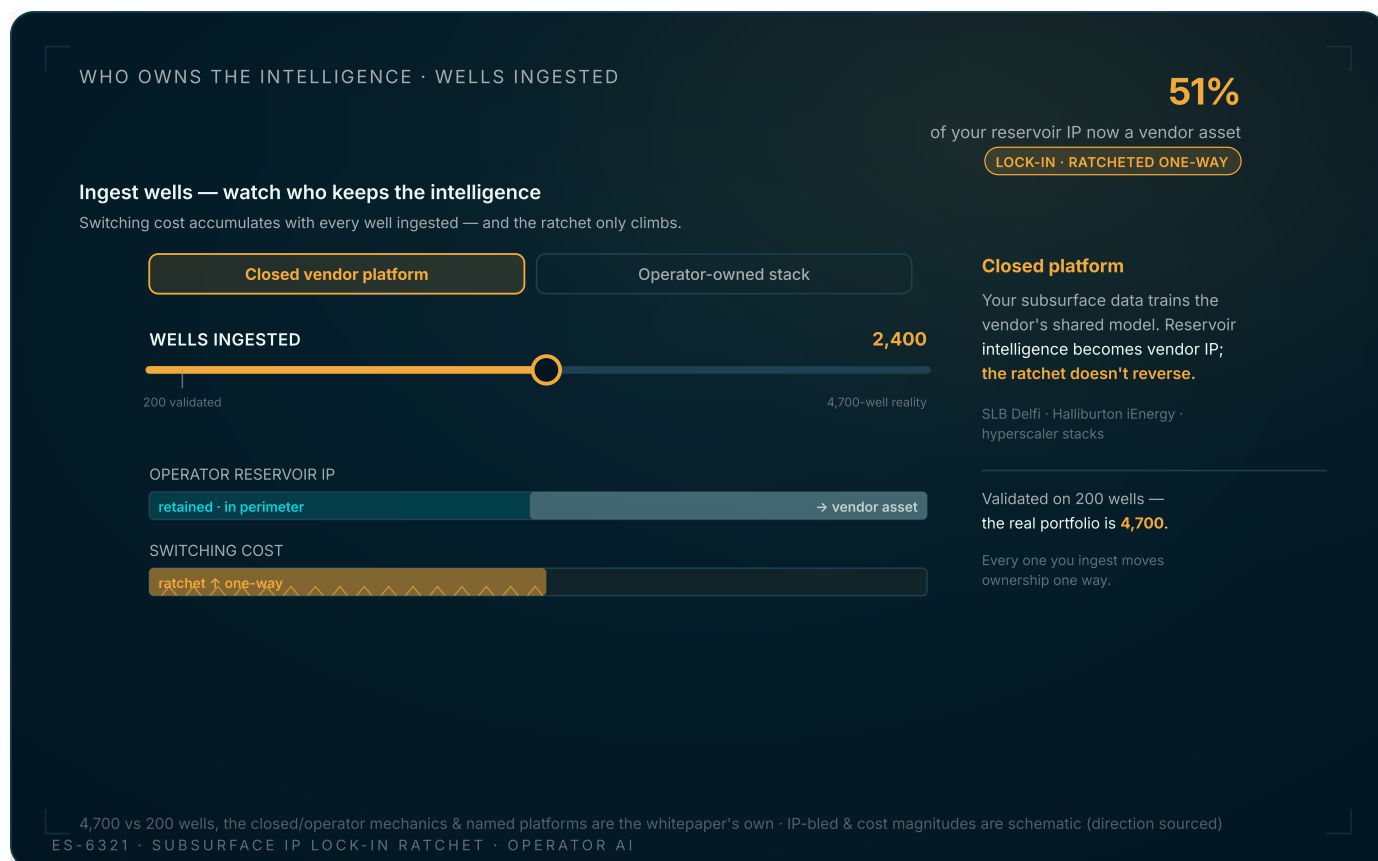
The vendor response — the one that has dominated the last five years — is to package the plumbing into a closed platform: SLB's Delfi, Halliburton's iEnergy, and the major hyperscaler stacks. These offerings solve a real problem. They also solve it in a way that ingests operator subsurface data to train shared platform models, converting the operator's hard-won reservoir intelligence into a vendor asset. McKinsey's State of AI 2025 identifies workflow redesign as the single strongest EBIT correlate of AI value. Outsourcing that redesign to a vendor whose commercial incentive is to maximise billable services is a misalignment that no contract clause can fully correct.

CLOSED VENDOR PLATFORM

- Operator subsurface data trains shared models
- Reservoir intelligence becomes vendor IP
- Workflow redesign owned by vendor services arm
- Switching costs accumulate with every well ingested

OPERATOR-OWNED FULL STACK

- Open-weights foundation model (Hulde, RAIL licence)
- Data and models remain inside operator perimeter
- Workflow redesign owned by the asset team
- Capability transfer is the explicit end state



A closed vendor platform and an operator-owned stack diverge on one axis: who owns the intelligence as wells accumulate. Every well ingested into a shared-model platform converts more operator reservoir IP into vendor asset and ratchets switching cost up — one-way: ingested data can't be un-ingested (drag the slider back and the lock holds). Toggle to the operator-owned stack and IP stays inside the perimeter, switching cost stays flat. The 4,700-well portfolio vs 200-well training set, the closed/operator mechanics and the named platforms are the whitepaper's own; the IP-bled fraction and switching-cost magnitude are schematic (the direction is sourced).

Our approach

EarthScan is an AI engineering company. We are not a model lab, not a services boutique, and not a platform reseller. We build and operate the full stack — six capability layers from HPC up to deployment — and we hand the keys back to the operator at scale. The product surface today is five formation evaluation modules that run on operator-uploaded LAS, SEG-Y, and image log data: Vug Detection, Fracture and Bedding Detection, VSP Analysis, Well Log Detection, and Geomechanical Analysis covering 1D mechanical earth models and wellbore stability. Each module is in production with at least one operator.

The foundation underneath those modules is Hulde, our procedural foundation model from the Hominis family, released with open weights under a RAIL licence. Hulde is purpose-built for a property that matters more in subsurface than in any consumer AI domain:

hallucination resistance. A general-purpose LLM optimised to sound confident is the wrong tool for a million-dollar well decision. Hulde prioritises correct execution of a procedural workflow over fluent narration of one. That design choice — explicit, auditable, conservative — is what makes it deployable inside an asset team's decision loop.

"A general-purpose LLM optimised to sound confident is the wrong tool for a million-dollar well decision. For subsurface workflows, hallucination resistance is a prerequisite, not a feature."

The full-stack posture matters because every layer of the stack has its own failure mode, and a model is only as production-ready as the weakest layer below it. A drift-monitoring layer without a clean data ontology is decorative. An agent layer without provenance tracking is unauditable. By owning all six layers and integrating them as one product, we eliminate the seams where pilots historically die — and we ship the result inside the operator's security perimeter, not on a shared platform.

Five production modules

- Vug Detection
- Fracture & Bedding Detection
- VSP Analysis
- Well Log Detection
- Geomechanical Analysis (1D MEM, wellbore stability)

Hulde foundation model

- Hominis family, open weights
- RAIL licence — no data lock-in
- Procedural execution, not narrative generation
- Designed for hallucination resistance on subsurface workflows

Six capability layers

- HPC
- Data engineering
- Data unification
- AI / ML
- Agents
- Platform deployment

Case examples

Two engagements illustrate the difference between a stalled pilot and a deployed workflow. Both are anonymised; both reflect the dynamics we see repeatedly across mid-market operators, independents, and NOCs.

Operator A — North Africa independent — ran a six-month POC with a tier-one vendor on an unconventional play. The technical results were validated. Recovery factor predictions correlated with offset wells, and the model's uncertainty bounds were defensible. Eighteen months later, the model is not in production. The reason is mundane and instructive: no one owned the integration layer between the vendor's model environment and the asset team's

existing G&G workflow. Every monthly run required manual data preparation by a single geoscientist who left the company. The model did not fail. The engineering around it did not exist.

THE POC TRAP, IN ONE SENTENCE

The model did not fail. The engineering around it did not exist. This pattern accounts for the majority of stalled pilots we are asked to revive.

Operator B — Gulf region independent — engaged EarthScan to deploy Well Log Detection across a 47-well evaluation portfolio. The previous workflow took a geoscientist approximately 14 hours per well to evaluate. After eight weeks of engagement, the agentic workflow reduced evaluation time to under 90 minutes per well, with geoscientist review explicitly retained for anomaly cases and final sign-off. The model did not replace the geoscientist. The engineering layer freed roughly 600 hours of expert time across the portfolio for higher-leverage work. The operator retained the platform, the weights, and the data.

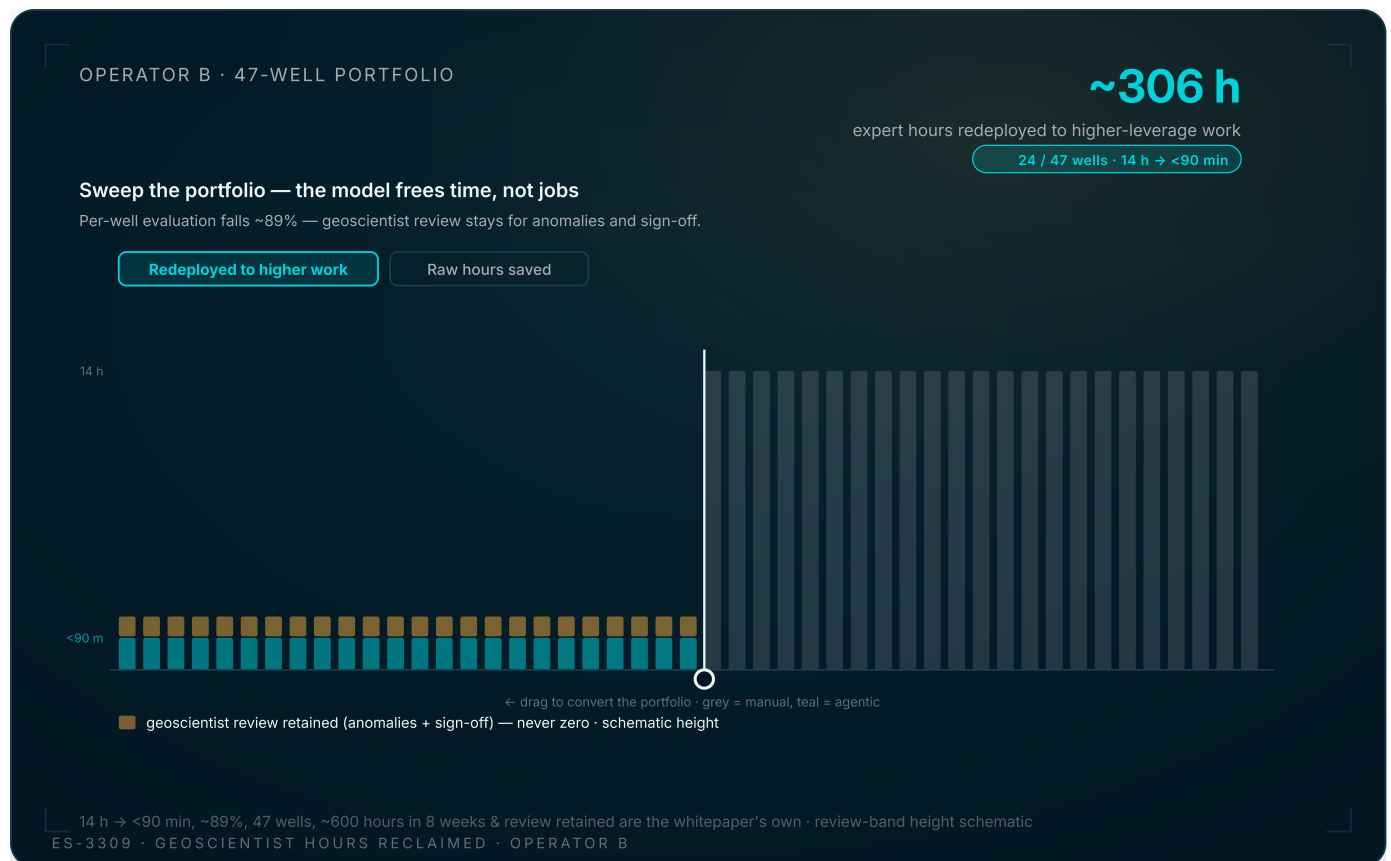
OPERATOR B — WELL LOG DETECTION (47 WELLS)

BEFORE

14 hours

AFTER

<90 minutes



The engineering layer didn't replace the geoscientist — it moved expert time off rote evaluation onto judgment. On Operator B's 47-well portfolio, per-well Well-Log-Detection evaluation fell from 14 hours to under 90 minutes (~89%), reclaiming ~600 expert hours in 8 weeks — while geoscientist review was explicitly retained for anomaly cases and final sign-off. Sweep across the portfolio: each 14-hour bar collapses to a <90-minute sliver, the reclaimed-hours total fills toward ~600, and an orange 'review retained' band rides every converted well and never shrinks to zero. All headline numbers are the whitepaper's own; the review-band height is schematic (review is retained but its hours aren't quantified).

Implementation roadmap

We deliver in three phases: Discover, Pilot, Scale. The cadence is intentional. Discover is short enough to fit inside a single budget cycle and produces an engineering artefact, not a strategy deck. Pilot is short enough to keep stakeholder attention and long enough to expose every layer of the stack to real operator data. Scale is paced to the operator's risk appetite, not the vendor's revenue model.

1

Discover (Weeks 1–4)

Data inventory, readiness assessment, infrastructure fit-gap. Output: a signed-off engineering brief — not a strategy deck.

2

Pilot (Weeks 5–12)

One production module integrated into the asset workflow. Ingestion + provenance, domain model on Hulde, agentic workflow, MLOps scaffolding.

3

Scale (Months 4–12)

Data unification, 2–3 additional modules, agent-to-agent orchestration, HPC provisioning, full capability transfer.

Discover runs four weeks. We conduct a data inventory and readiness assessment on the target asset, identify the single highest-ROI workflow to attack first, and produce an infrastructure fit-gap analysis. The deliverable is a signed-off engineering brief: which module, which data, which integration points, which success metrics, which acceptance criteria. This is the document that decides whether Pilot proceeds. If the data is not ready, we say so — and we recommend remediation before module deployment, not in parallel with it.

Pilot runs eight weeks (weeks five through twelve). We deploy one production AI module integrated directly into the asset team's existing workflow — not as a parallel sandbox. Deliverables include an ingestion pipeline with full provenance tracking, the domain model running on Hulde, an agentic upload-run-visualise-report workflow, baseline performance metrics measured against the geoscientist-validated ground truth, and MLOps scaffolding for versioning and drift monitoring. By the end of week twelve, the asset team is using the module on live data.

Scale runs months four through twelve. We unify data across subsurface, operational, and sensor sources; deploy two to three additional modules; introduce agent-to-agent orchestration for multi-step workflows; provision HPC at production scale; and execute full capability transfer. By month twelve, the operator owns the platform — the weights, the pipelines, the agents, the runbooks, and the team know-how to run and extend it independently. EarthScan transitions to a steady-state support role, or out entirely.

Risk and mitigation

No engagement is risk-free, and an honest whitepaper says so. The dominant risks we encounter, in roughly the order they appear, are data readiness, integration debt, change management, and over-scoping.

Data readiness is the most common reason a Pilot slips. Operators routinely underestimate the heterogeneity of their own LAS, SEG-Y, and image log archives. Our mitigation is the Discover phase itself: we surface the readiness gap before signing a Pilot statement of work, and we recommend remediation as a discrete workstream rather than burying it inside model development. If the data is not ready, we say so out loud.

Integration debt is the second pattern. Pilots designed as parallel sandboxes rarely convert; pilots designed as workflow integrations almost always do. We insist on integration into the asset team's existing tools as a non-negotiable Pilot requirement, even when it is the slower path. Change management is the third — a working module that no geoscientist trusts is shelfware. We mitigate by retaining explicit geoscientist review for anomaly cases, by surfacing model confidence and provenance on every output, and by training the asset team on the agentic workflow rather than handing them a black box.

Over-scoping is the fourth and most preventable risk. A Pilot that tries to deploy three modules in twelve weeks delivers none. We scope to one module, one asset, one workflow — and we expand only after that workflow is in production and trusted. The operator retains the option to slow down, pause, or terminate at any phase boundary. Capability transfer is built into the contract structure, not added as a courtesy at the end.

PHASE-BOUNDARY EXITS ARE EXPLICIT

The Discover, Pilot, and Scale phases each end at a defined decision point. The operator can pause, re-scope, or terminate at any boundary without forfeiting deliverables or accumulating switching costs. This is the inverse of the closed-platform model.

Conclusion and next steps

The decision in front of operators in 2025 is not whether to adopt AI. The market is moving from \$7.2 billion to \$17 billion in five years; the adoption question is settled. The decision is whether the AI deployed on your asset is owned by you or by a platform vendor whose interest is maximising billable services on your subsurface data.

Mid-market operators, NOCs, and independents who have been underserved by seven-figure oilfield service contracts have a third option now. A specialist AI engineering partner — Hulde foundation model, five formation evaluation modules, six capability layers, twelve weeks to production, full capability transfer at scale. Conflict-free by construction, because EarthScan does not sell drilling services, completion services, or anything else that competes with the operator's own value chain.

The next step is a four-week Discover engagement on a single asset. The output is a signed engineering brief — what is ready, what is not, what the highest-ROI first module is, and what the twelve-week Pilot looks like. If the readiness is not there, we will say so. If it is, the path from POC purgatory to a production workflow on the asset starts there.

BY THE NUMBERS

\$7.2B → \$17B

AI in oil & gas market, 2025 → 2030

18.5%

CAGR through 2030

~50%

Of enterprise AI projects that never reach production

75%

Seismic model-build time cut by ADNOC's ENERGYai (2024)

12 weeks

Discover + Pilot to production on a single asset

References

Knowledge Sourcing Intelligence, 2025 Knowledge Sourcing Intelligence. AI in Oil and Gas Market (2025). Market sized at \$7.2B (2025), forecast \$17B by 2030, 18.5% CAGR.

<https://www.knowledge-sourcing.com/>

Mordor Intelligence, 2025 Mordor Intelligence. AI in Oil and Gas Market (2025). Upstream accounts for 61% of market; ADNOC ENERGYai cut seismic model-build times by 75% in 2024.

<https://www.mordorintelligence.com/>

IEA, 2025 International Energy Agency. Energy and AI Special Report (2025). Missing internal expertise identified as the dominant adoption barrier.

<https://www.iea.org/reports/energy-and-ai>

IDC, 2025 IDC (2025). Up to 50% of enterprise AI projects collapse before production.

<https://www.idc.com/>

Devon Energy, CERAWeek 2025 Devon Energy, CERAWeek 2025. Reported 25% productivity gains from AI drilling optimisation.

Business Research Insights, 2025 Business Research Insights. AI in Oil and Gas (2025). Predictive maintenance cutting unplanned downtime by up to 30%.

<https://www.businessresearchinsights.com/>

McKinsey State of AI, 2025 McKinsey & Company. The State of AI (2025). Workflow redesign identified as the single strongest EBIT correlate of AI value capture.

<https://www.mckinsey.com/>

McKinsey Global Energy Perspective, 2025 McKinsey & Company. Global Energy Perspective (2025). Digital transformation in energy advancing at roughly half the pace required. <https://www.mckinsey.com/>

MDPI JRFM, 2026 MDPI Journal of Risk and Financial Management. Financial Performance Outcomes of AI Adoption in Oil and Gas (2026). <https://www.mdpi.com/1911-8074/19/1/44>

Own Your AI Stack: A Whitepaper for Energy Operators

Published May 2026. Generated 2026-07-22.

<https://www.earthscan.io/whitepapers/own-your-ai-stack-upstream-2025>

Copyright EarthScan. All rights reserved.