From Manual Picks to a Continuous Interpretation Workflow: Operationalizing AutoFrac and AutoVug for Reservoir Engineering

Borehole-image interpretation is run almost everywhere as a batch craft: a senior geoscientist takes a single high-resolution borehole image log, picks bedding planes, fractures and vugs sinusoid by sinusoid over a multi-week pass, hands a static result to the asset team, and moves to the next well while the queue grows. This whitepaper argues for replacing that batch craft with a continuous interpretation workflow, drawn from a roughly twenty-month, three-phase formation-evaluation engagement with a mid-sized Middle East carbonate operator we partnered with. We describe the two production capabilities at its core — AutoFrac for bedding-and-fracture detection and AutoVug for vug analysis, each running roughly five times faster than manual picking — not as models but as stages in a standing pipeline a well moves through: ingest, normalise, patch, infer, validate, and emit a structured eight-column output of predicted bedding planes, sinusoids, fractures and vugs that becomes the interface to reservoir engineering. We treat blind-zone validation on a held-out well and depth interval as the pipeline's standing acceptance gate rather than a one-time benchmark, and we show how the interpreter capacity reclaimed by the speed-up is redeployed — onto the operator's 80-plus-well backlog, and into continuous vug-percent and fracture-density reservoir maps that a manual team could never have produced at scale. The reframing is the point: interpretation stops being a three-to-four-week-per-well event and becomes a running service that reservoir engineering can query.

Tannistha Maiti
Senior AI Researcher

Tarry Singh
Founder & CEO

BOREHOLE-IMAGE-INTERPRETATION / FORMATION-EVALUATION / CONTINUOUS-WORKFLOW / AUTOFRAC / AUTOVUG / RESERVOIR-ENGINEERING

CONTENTS

A borehole image log is read the way it has been read for thirty years: one well, one interpreter, one pass that runs for weeks. The geoscientist scrolls the unrolled image of the borehole wall, fits a sinusoid to every bedding plane and fracture by hand, traces the dark spots that might be vugs, records dip and azimuth, and produces a static interpretation that the asset team receives, files, and rarely revisits. It is careful, expert work. It is also a batch process with a single server, and the queue behind it never empties — every well an operator acquires adds another multi-week pass to a backlog the team cannot clear. The bottleneck is not the geoscientist's skill. It is the architecture: interpretation is run as a sequence of discrete, manual events when it should be run as a continuous service.

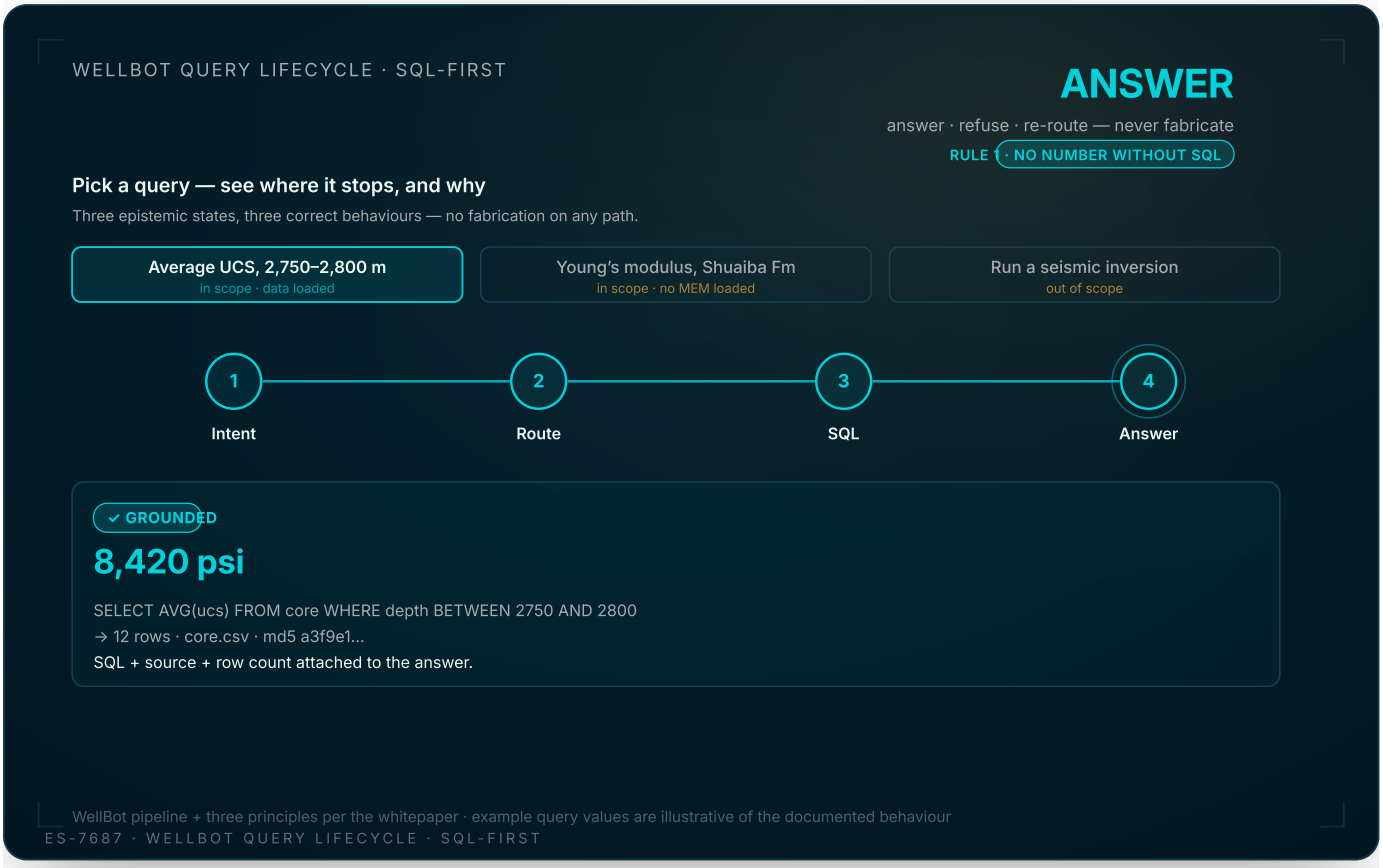
This whitepaper is about that architectural change. Over roughly twenty months, across three phases, working with a mid-sized Middle East carbonate operator, we built and handed over two production interpretation capabilities — AutoFrac, which detects bedding planes and fractures and regresses their dip and azimuth, and AutoVug, which quantifies vugs — and the well-to-well correlation tooling that consumes their output. The deep-learning work behind them is real and is documented elsewhere. Here the subject is different and, for a reservoir engineer or a workflow owner, more consequential: how those capabilities stop being models that produce a result and become **stages in a standing pipeline a well flows through**, and what reservoir engineering can do once interpretation is a feed rather than a deliverable.

The unit of work changes: from a well-event to a well-in-flight

The single most important shift is conceptual, and it is worth stating plainly before any number. In the manual world, the unit of work is a *finished interpretation*: a well arrives, weeks pass, a static result is handed over. In the continuous world, the unit of work is a *well in flight* — an object moving through a fixed sequence of stages, observable at each one, emitting a structured result at the end that downstream systems can consume without a human in the loop for the mechanical part.

The stages are not abstract. A well enters as raw inputs — the binary wireline log file, apparent dip and azimuth picks, well radius and the interpreter's reference PDF — and is carried through normalisation (reconciling the static and dynamic image channels, correcting for tool and well angle), patch generation (cutting the unrolled image into overlapping fixed-height tiles the model can consume), inference by AutoFrac and AutoVug,

geoscientist-validated evaluation, and emission of a structured output. Each stage is inspectable; a well does not vanish into a black box for three weeks and reappear as a PDF. It has a position in the pipeline, and an engineer can see exactly where it is.



WellBot's SQL-first, refusal-aware pipeline. Pick a query and watch it flow through Intent → Route → SQL → Answer and stop at the correct place: an in-scope query with data executes SQL and returns a grounded answer with a full audit trail; an in-scope query with no data returns 0 rows and a RULE 1 refusal (no interpolation); an out-of-scope query is refused at the router and re-routed. Every path is the right behaviour — answer, refuse, or re-route, never fabricate. Pipeline and principles per the whitepaper; example values illustrative.

The lifecycle above is the architecture made literal: a request enters, is routed and scoped, runs against loaded data, and returns an answer that carries its own provenance — or a structured refusal when the data to answer it is not present. That discipline is exactly what separates a continuous interpretation pipeline from a faster version of the manual one. A well that lacks the inputs a stage needs does not get a fabricated interpretation; it gets a recorded, inspectable stop, and the asset team knows precisely what is missing before anything reaches a reservoir model. Speed without that gate is just a faster way to ship a wrong pick.

The output is a contract, not a picture

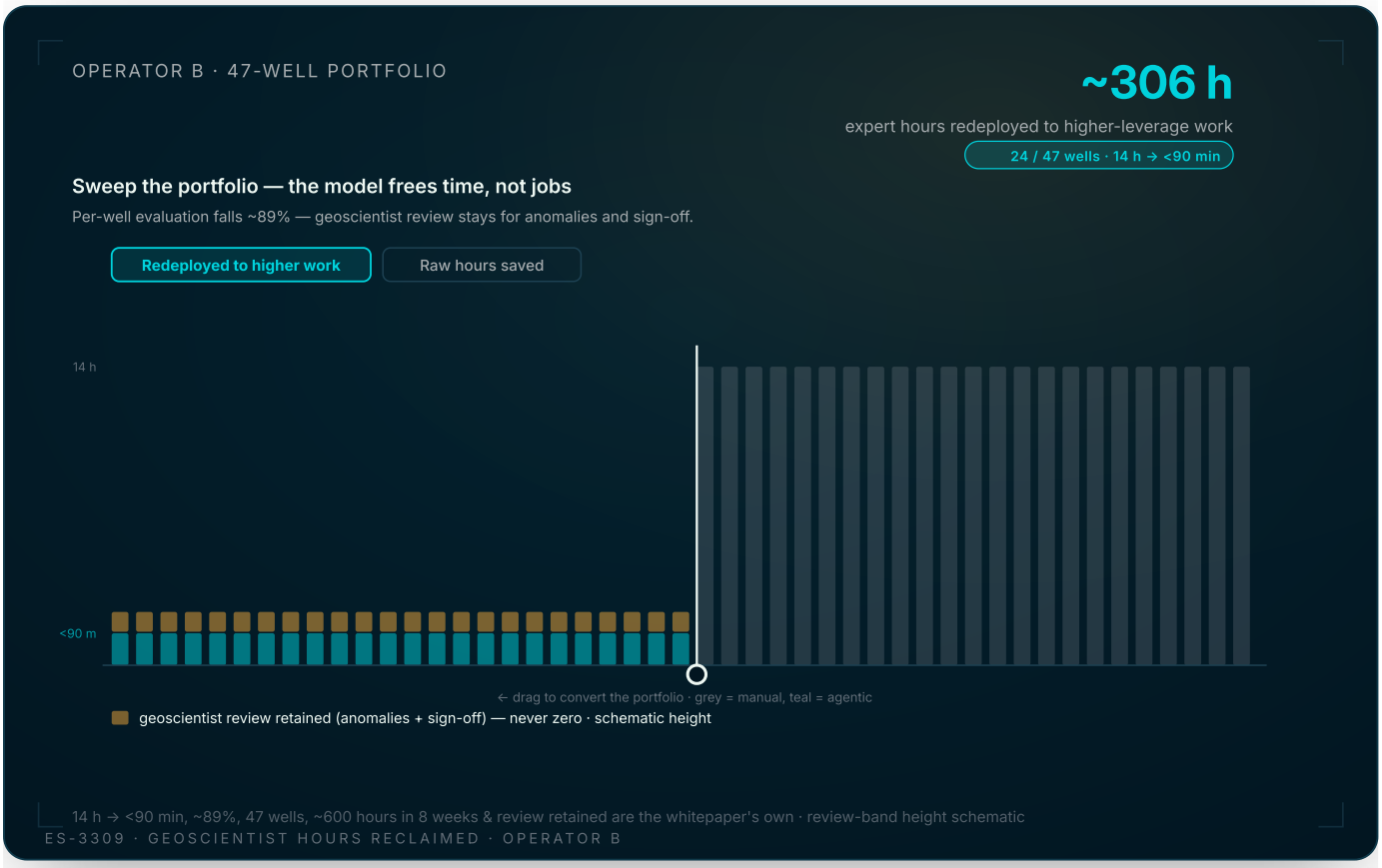
Manual interpretation hands the asset team a marked-up image and a set of notes — a picture a human made for other humans. A continuous pipeline has to hand downstream systems something a machine can consume deterministically. In this engagement that something was an eight-column structured output: for each detected feature, the pipeline emits predicted **bedding planes, sinusoids, fractures and vugs**, with their geometric parameters, as columns a reservoir-engineering workflow can read directly.

That eight-column emission is the interface between interpretation and reservoir engineering, and treating it as a stable contract is what makes the rest of the workflow possible. A reservoir engineer building a fracture-density map does not want to re-read a borehole image; they want a column of fracture locations and orientations they can window and aggregate. A completions engineer assessing a zone does not want a marked-up log; they want vug counts and a porosity proxy per depth interval. When interpretation emits a typed, columnar output with a fixed schema, those downstream consumers can be built once and run forever, against every well the pipeline processes, rather than being hand-assembled per well from a geoscientist's notes. The picture is for the geoscientist who validates; the contract is for everyone downstream.

> **WHY THE SCHEMA IS THE PRODUCT**
>
> The eight-column output — predicted bedding planes, sinusoids, fractures and vugs with their parameters — is the load-bearing deliverable of the whole workflow. A model that produces an accurate interpretation only a human can read keeps reservoir engineering dependent on the interpreter for every downstream use. A model that emits a stable, typed schema lets reservoir engineering build standing maps and aggregations that run unattended across the full well stock. The architectural win is not a better pick; it is a pick that downstream systems can consume without a person in the loop.

Speed is the enabler, not the headline

The two interpretation stages run roughly **five times faster** than manual picking. That number matters, but not for the reason it is usually quoted. A 5x speed-up on a single well is a convenience. A 5x speed-up applied continuously to a standing backlog is a change in what the asset team can attempt at all.



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).

The instrument above shows where the reclaimed time comes from: the mechanical labour — fitting sinusoids, tracing vug boundaries, transcribing dip and azimuth — collapses, and what remains for the geoscientist is validation, anomaly review, and sign-off. This is the correct division of labour. The pipeline does the repetitive geometry at machine speed; the expert is retained for the judgement calls the model should not make alone. Crucially, the expert stays *in* the loop as the acceptance authority, not *on* the critical path for every

sinusoid. Well-to-well correlation built on top of these stages lifted interpretation productivity by about 60% and accuracy by about 75% once the capabilities were productized — gains that come not from any single faster pick but from the workflow no longer stalling on the interpreter for the mechanical pass.

A reservoir engineer should read the 5x speed-up as a capacity figure, not a latency figure. The question it answers is not "how fast is one well now" but "how many wells can the same team move through interpretation per quarter" — and the answer reframes the backlog from an immovable queue into something a standing pipeline can actually work down.

Validation is a standing gate, not a one-time benchmark

The failure mode of every "we automated interpretation" claim is the same: a model validated once, on a curated test set, declared accurate, and then trusted indefinitely as the formation it sees drifts away from the formation it learned on. A continuous workflow cannot be operated that way. It needs a validation gate that runs as a standing part of the pipeline, on data the model has never seen, in geological conditions that resemble production.

We built that gate as **blind-zone validation on a held-out well**. A specific well was withheld from training entirely, and within it a continuous depth interval — on the order of 25 metres in the reservoir section — was reserved as a blind zone: the model produced predictions across it, and those predictions were scored against the geoscientist's independent picks for the same interval. Because the zone was continuous and non-overlapping with anything the model trained on, it is an honest proxy for the next real well, not a leak-contaminated split. A held-out *interval* inside a held-out *well* tests two things at once: that the model generalises across depth, and that it generalises across wells.

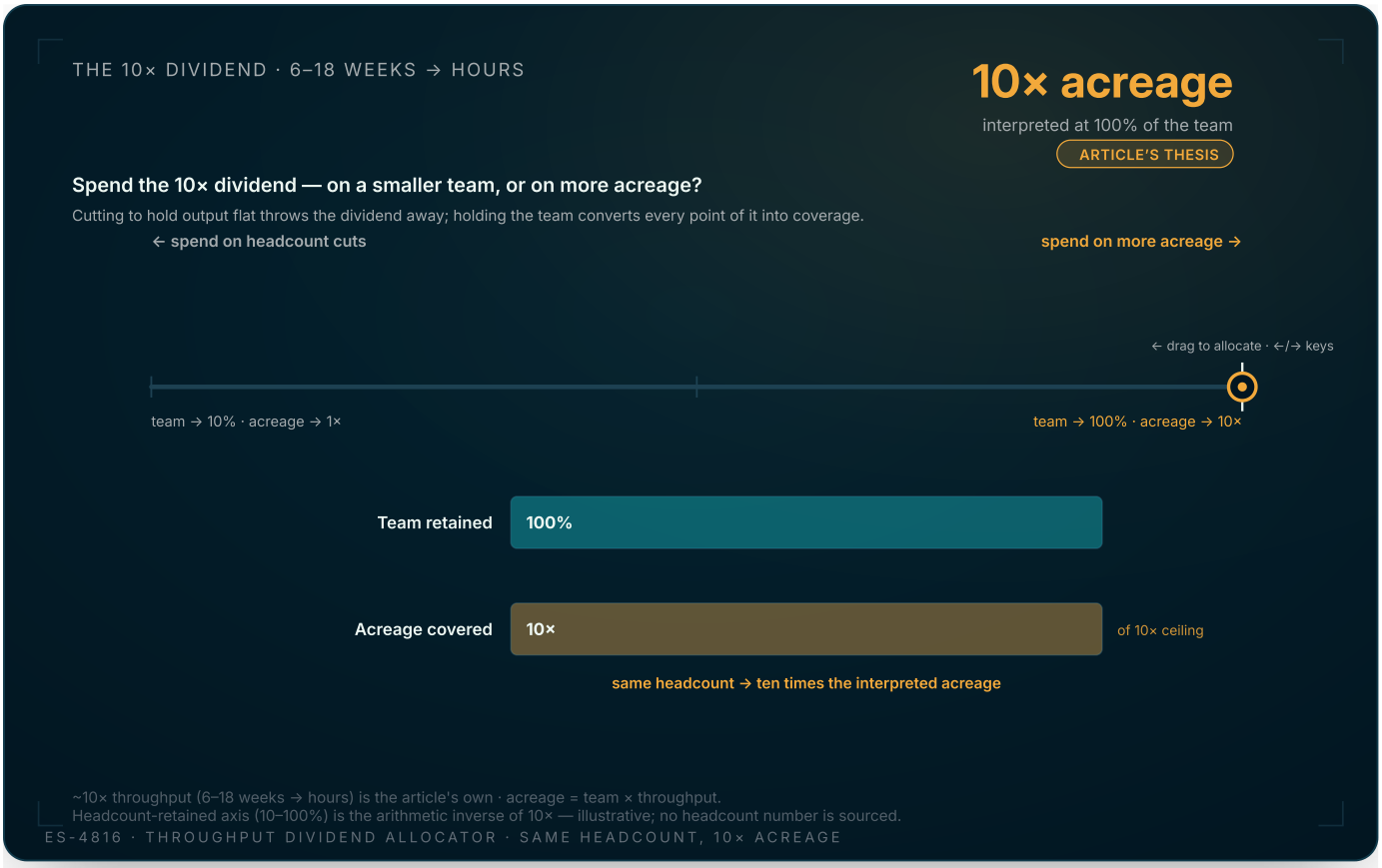
> **READ THE METRIC AGAINST THE INSTRUMENT FLOOR**
>
> Blind-zone scores have to be read against a physical limit, not in spite of it. At the image resolution in play, a single image-log pixel corresponds to about 3 centimetres of depth, so a ±3 cm uncertainty is present in the input before the model does anything. A predicted feature sitting 3 cm off a ground-truth pick is the instrument, not the network. The localisation tolerance that defines a correct detection has to be

> chosen and recorded against that floor — report a blind-zone number without naming the depth tolerance and the number is meaningless. A standing validation gate that ignores the instrument floor will either reject good models or pass bad ones.

The operational discipline is to keep this gate running. Every time the pipeline is retrained on new wells — and in this engagement the training set grew from 8 to 11 to a 14-well dataset and beyond as data arrived — the blind zone is re-scored before the refreshed model is promoted. Validation stops being an event that happened once at the end of the build and becomes a checkpoint the pipeline passes through on every update. That is the difference between a model you trust because it once tested well and a workflow you trust because it keeps proving itself.

What the reclaimed capacity is actually for

A productivity gain that is simply banked — the same number of wells interpreted, with idle geoscientists — is a gain no asset manager will fund twice. The case for a continuous workflow rests on what the reclaimed interpreter capacity is *redeployed* onto. There are two destinations, and both are things the manual workflow could never reach.



The naive read of a 10× productivity jump is 'cut 90% of the team.' The article argues the opposite: at the same headcount, the dividend buys ten times the interpreted acreage. Drag the allocator — spend the dividend on headcount cuts and acreage stays at 1×; hold the team and acreage climbs toward 10×. The single orange marker is the article's own position. The ~10× throughput multiple (6–18 weeks → hours) is sourced; the headcount-retained axis is the arithmetic inverse of 10× and is labelled illustrative, since the article states no specific headcount figure.

The first destination is the **backlog**. The engagement was scoped against more than 80 processed and interpreted image logs — a stock of wells far larger than any manual team could keep current. When interpretation runs roughly five times faster as a standing pipeline, the reclaimed capacity is not a bonus; it is the only mechanism by which that backlog gets worked down at all. The allocator above is the explicit decision an asset team now gets to make: how much of the throughput dividend goes to clearing the queue of un-interpreted wells, and how much goes to deepening the analysis on wells already processed. That choice did not exist when every well consumed a multi-week manual pass — there was no dividend to allocate.

The second destination is **reservoir mapping that only continuous interpretation makes feasible**. Once the eight-column output exists for many wells, reservoir engineering can compute things across the field that were impractical to assemble by hand:

- **Fracture-density volumes.** Fracture locations from the pipeline's output were aggregated into density grids over rolling depth windows — computed at 2-, 5- and 10-metre windows, settling on a fine per-interval grid — turning a column of individual fracture picks into a continuous fracture-intensity profile a reservoir model can ingest. Across many wells, those profiles become a field-scale picture of where the natural-fracture network concentrates.
- **Vug-percent and pore-statistics logs.** AutoVug's output supports per-interval statistics — counts, total and mean vug area, an area-and-porosity spectrum, a circularity spectrum and an azimuth spectrum computed per short depth window — producing a continuous secondary-porosity log rather than a one-off vug count. That is a direct input to porosity and completion decisions, available for every well the pipeline touches.
- **Well-to-well structural correlation.** With consistent, typed interpretations available across the well stock, the correlation tooling can tie bedding and fracture trends between wells over the field's spacing — work that depends entirely on having comparable, machine-readable interpretations for many wells at once, which only a continuous pipeline supplies.

None of these are exotic analyses. They are the reservoir-engineering products an asset team would obviously want — and could never produce when interpretation was a per-well bottleneck consuming the entire interpretation budget. The continuous workflow does not just speed up the old deliverable. It makes a class of field-scale deliverables possible for the first time.

The economics of the reframe

It is worth being explicit about the commercial logic, because the workflow change has to pay for the engineering behind it. The cost of manual interpretation scales linearly with the well count: every well is another multi-week expert pass, and the queue grows faster than the team clears it. The cost of the continuous workflow is a one-time build of the pipeline plus a small, roughly five-times-cheaper review pass per well thereafter.

INTERPRETATION COST: MANUAL BATCH VS CONTINUOUS PIPELINE

Formula | LaTeX

Copy LaTeX

$$C_{\text{manual}} = N \cdot c_{\text{week}} \cdot w \quad \longrightarrow \quad C_{\text{pipeline}} = C_{\text{build}} + N \cdot c_{\text{review}}$$

The left expression is the manual regime — well count times the cost of a week of expert time times the number of weeks per well — and it is the reason the backlog is economically immovable: there is no term that does not grow with the number of wells. The right expression is the continuous regime: a fixed build cost, then a per-well review cost that the 5x speed-up makes a fraction of the manual pass. The build cost is real and the engagement was scoped accordingly, but it is paid once. Past the crossover, every additional well is interpreted at the review cost, and the reclaimed expert weeks are the dividend that funds the backlog and the reservoir maps. The reframe is not "the model is faster." It is "interpretation is no longer a cost that scales with the well count" — and that is a reservoir-engineering economics statement, not a machine-learning one.

How to operate it: the workflow owner's checklist

For the workflow or process owner who has to run this — not build it — the continuous interpretation pipeline is a system with operating requirements, and most of them are unglamorous. The questions that decide whether it stays a running service rather than decaying into a one-off are not about model architecture:

- **Is the eight-column output schema frozen and documented as a contract?** Downstream maps and aggregations break silently when the interpretation schema drifts. The schema is an interface; treat it like an API.
- **Does the blind-zone gate run on every retrain, against a recorded depth tolerance?** A validation gate that runs once is a benchmark. A validation gate that runs on every promotion is an operating control. The instrument floor (±3 cm) has to be written into the acceptance criteria, not assumed.
- **Is the geoscientist positioned as the acceptance authority, not the bottleneck?** The pipeline does the mechanical geometry; the expert validates, reviews anomalies and signs off. If the workflow routes every sinusoid back through a human, the 5x speed-up evaporates and you have rebuilt the manual process with extra steps.
- **Is the reclaimed capacity explicitly allocated, not implicitly banked?** Decide — and revisit — how the throughput dividend splits between clearing the backlog and deepening reservoir maps. Capacity that is not allocated is capacity that quietly disappears into the old way of working.
- **Can the operator retrain and re-validate without the build team in the room?** A continuous workflow the vendor has to run is not continuous; it is a subscription. The retrain-and-revalidate loop has to be owned in-house, which is why the capability was handed over as a

complete unit — dataset, model, schema, runbooks and the validation gate — alongside a trained local cohort to operate it.

A continuous interpretation workflow answers yes to all five. A faster manual process answers no to most of them and is back to a growing queue within a year of the consultants leaving.

Conclusion: interpretation becomes a service

The manual interpretation pass is not slow because geoscientists are slow. It is slow because it is the wrong architecture for the volume of wells a modern asset acquires — a batch craft with a single human server, run as a sequence of discrete events, handing static pictures to systems that need running feeds. Replacing it with AutoFrac and AutoVug as stages in a standing pipeline — roughly five times faster, validated continuously against a blind zone, emitting a stable eight-column contract into reservoir engineering — does more than compress a per-well cycle from weeks to a reviewed pass. It changes interpretation from a deliverable an asset team waits for into a service an asset team queries, and it turns the reclaimed expert capacity into field-scale fracture-density and vug-percent maps that the old workflow could never have produced. The model is the means. The continuous interpretation throughput, feeding reservoir-engineering decisions across the entire well stock, is the asset.

WHAT THIS WHITEPAPER ARGUES

1. The unit of work changes from a finished per-well interpretation (a multi-week batch event) to a well in flight through a standing pipeline: ingest, normalise, patch, infer, validate, emit.
2. AutoFrac and AutoVug run ~5x faster than manual picking — read that as a capacity figure (more wells per quarter), not a latency figure for one well.
3. The eight-column structured output (predicted bedding planes, sinusoids, fractures, vugs) is the real product: a typed contract reservoir engineering consumes without a human in the loop, enabling standing maps across the full well stock.
4. Blind-zone validation on a held-out well and depth interval is a standing acceptance gate re-run on every retrain — scored against the ±3 cm image-log instrument floor — not a one-time benchmark.
5. Reclaimed interpreter capacity is redeployed, not banked: onto the 80-plus-well backlog and into continuous fracture-density (2/5/10 m windows) and vug-percent reservoir maps that manual interpretation could never produce at scale.

References

International Energy Agency, 2025 International Energy Agency. Energy and AI Special Report (2025). Missing internal expertise and unscalable expert workflows identified as dominant barriers to AI value capture across the energy sector. https://www.iea.org/reports/energy-and-ai

McKinsey & Company, 2025 McKinsey & Company. The State of AI (2025). Workflow redesign — not model adoption alone — identified as the strongest correlate of AI value capture. https://www.mckinsey.com/

Carion et al., 2020 N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko. End-to-End Object Detection with Transformers (DETR). ECCV 2020. Architectural basis for the set-prediction detection approach underlying AutoFrac. https://arxiv.org/abs/2005.12872

Sculley et al., 2015 D. Sculley et al. Hidden Technical Debt in Machine Learning Systems. NeurIPS 2015. The canonical argument that the model is a small fraction of a production ML system — and that the surrounding pipeline is the engineering. https://papers.nips.cc/paper/2015/hash/86df7dcfd896fcaf2674f757a2463eba-Abstract.html

EARTHSCAN

From Manual Picks to a Continuous Interpretation Workflow: Operationalizing AutoFrac and AutoVug for Reservoir Engineering

Published September 2025. Generated 2026-07-22.
https://www.earthscan.io/whitepapers/continuous-interpretation-workflow-autofrac-autovug-reservoir