

AI for Long-Term Energy Project Engagement and Stakeholder Management

Long-lived subsurface projects, and carbon capture and storage most of all, carry two machine-learning problems that people keep collapsing into one. The first is site selection, a bounded vision task: read the operator's legacy raster archive, of the order of 136,771 scanned log images, digitise the curves that constrain the reservoir, and choose where to store. A pure convolutional digitiser does this and is graded on the reconstructed curve, where our own work peaks at R-squared 0.9891. The second problem begins the day the site is chosen and runs for the life of the project: the community relationship, the regulatory posture, and the stakeholder engagement that decide whether a storage site that is geologically sound is also socially durable. That second problem is not a segmentation task with a different head; it is a different model family. This whitepaper sets out the case for a hybrid fuzzy-logic-plus-deep-net architecture as that family, anchored on the published result that such a model forecasts CO₂-storage project engagement at 90.476% accuracy, and argues three things with three instruments: that the fuzzy-DNN engagement forecast reaches usable accuracy as a distinct family from the digitiser, that the two families diverge on every axis of the problem that matters, and that the value of the forecast is not that it freezes engagement but that it buys the lead time to defend a decaying-trust asset before it crosses the acceptance floor. The deliverable is a way for an operator to organise its subsurface AI portfolio so the model that reads the rock and the model that reads the room are built, graded, and governed as the two different things they are.

Tannistha Maiti

Senior AI Researcher

Quamer Nasim

ML Research Engineer

Tarry Singh

Founder & CEO

CCS / CARBON-STORAGE / STAKEHOLDER-ENGAGEMENT / FUZZY-LOGIC / DEEP-NEURAL-NETWORK / HYBRID-MODELS

CONTENTS

01	What people collapse into one problem
02	The published anchor for the second family
03	Why fuzzy logic sits under the net
04	What usable accuracy means and does not mean
05	Same project, opposite problem shape
06	Why the divergence has governance consequences
07	What the forecast is actually for
08	Lead time is the product
09	How the two families share a project without merging
10	What an operator should take from this
11	Limitations
12	References
13	Get the full whitepaper

”

A storage site outlives the model that chose it. The vision network that read the archive is finished the day the site is picked. The model that reads the room is only getting started, and it is a different animal.

”

THE TWO PROBLEMS

A storage project carries one vision task and one relationship task

What people collapse into one problem

We build subsurface AI for operators, and the largest single confusion we meet is the assumption that a subsurface project is one machine-learning problem with a long tail. It is not. A long-lived storage project, and carbon capture and storage is the sharpest example, carries two problems whose only shared feature is the site they both point at.

The first problem is bounded and technical. Before an operator can store anything, it has to choose where, and that choice is constrained by the rock, which is described in the operator's own legacy archive of scanned well logs. That archive is large and unstructured: the engagement we draw this method from held on the order of 136,771 scanned log images, most of them raster with no vector curve behind them. Turning that pile into the curves that constrain a reservoir is a vision task, and it is the one we have written about at length elsewhere. Our digitiser reads the raster, segments each curve, and reconstructs it, and it is graded on how well the reconstructed curve matches ground truth. On that grading it peaks at an R-squared of 0.9891. The problem has a clean end: once the archive is digitised and the site is chosen, the vision task is complete and does not recur.

The second problem starts on the day the site is chosen and does not end for the life of the project, which for storage is measured in decades. It is the relationship: the community around the site, the regulators who license it, the institutions that fund and insure it, and the shifting social acceptance that decides whether a geologically sound store is also a durable one. This problem is not bounded, not technical in the same sense, and above all not the same shape as the first. Treating it as a second output head on the vision network, which is the reflex we keep seeing, misreads what it is.

This whitepaper is about that second problem, and about the specific claim that it deserves its own model family. We are not re-telling the digitiser here. We are placing a different architecture next to it, one built for the relationship rather than the rock.

The published anchor for the second family

The reason we can be concrete rather than aspirational about the second family is that the result already exists in the literature. Buah and colleagues asked directly whether AI can assist project developers in the long-term management of energy projects, and used carbon capture and storage as their case [1]. Their model is a hybrid: a fuzzy-logic layer that grades the soft, ambiguous signals a community relationship produces, feeding a deep neural network that learns to forecast engagement over the project. On the social-acceptance forecast, that hybrid reaches 90.476% accuracy.

That number is the spine of this document. It is not a curve R-squared and it is not comparable to one; it is the accuracy of a forecast about people, produced by an architecture built for people-shaped data. What it establishes is that the second problem is tractable with a real, distinct model family, and that the family has a recognisable shape: fuzzy membership under a deep net, graded on a decision rather than a reconstruction. The rest of this paper takes that anchor and works out what it means to run the two families side by side inside one operator's portfolio.

THE FORECAST

A hybrid model forecasts engagement at usable accuracy

Why fuzzy logic sits under the net

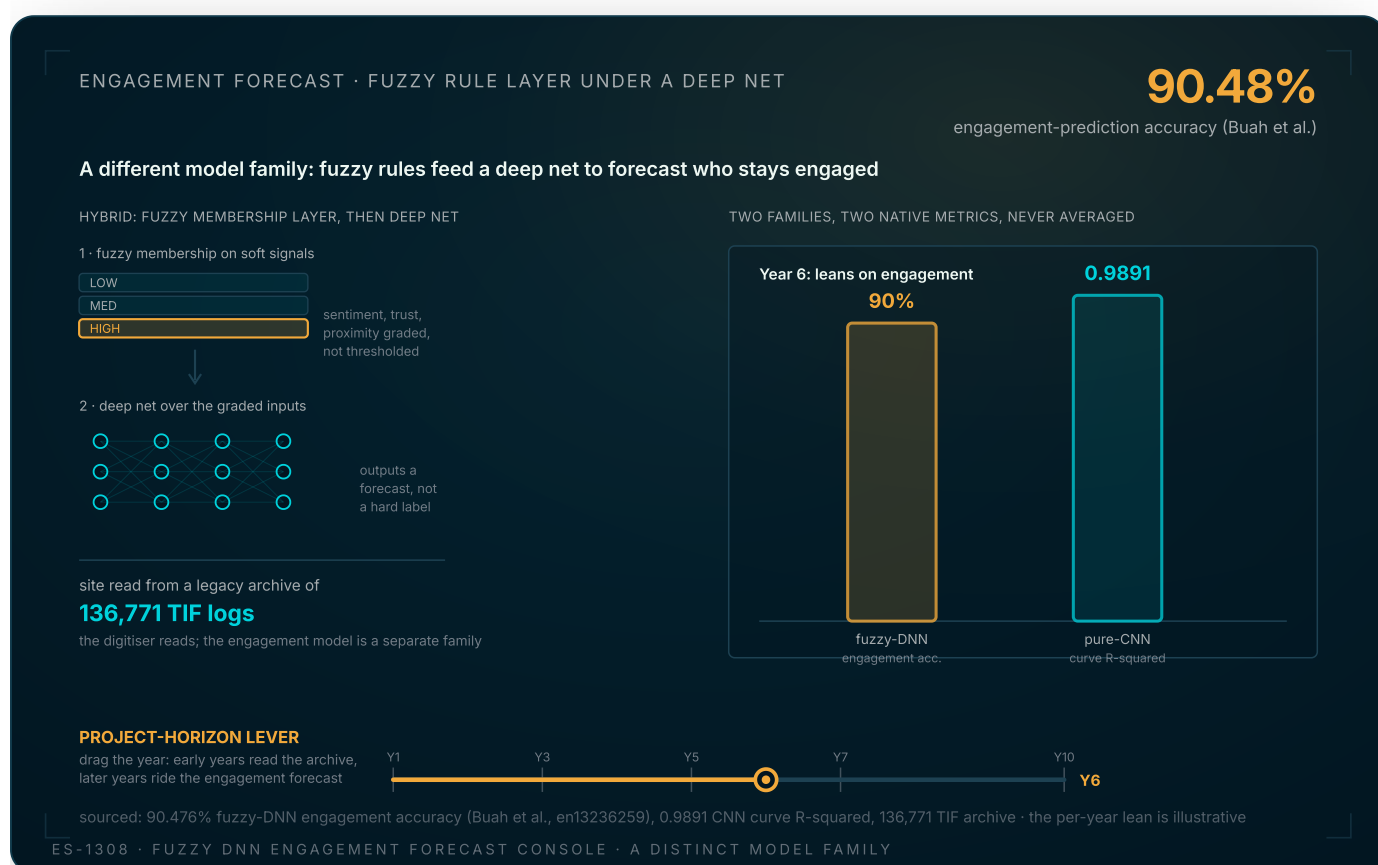
The instinct of a team that has just built a strong vision model is to reach for the same tools on the next problem: pixels in, a label out, a hard threshold on the label. That instinct fails on engagement data, because the inputs are not crisp. Whether a community trusts an operator is not a one or a zero; it is a degree, and it moves. Proximity to a site matters, but not as a binary inside-or-outside a radius; it matters on a gradient. Sentiment in a public consultation is a spectrum, not a class. Feed those signals to a network as hard categories and you throw away exactly the information that carries the forecast.

The fuzzy layer exists to keep that information. It grades each soft signal into a degree of membership, so a community that is moderately trusting enters the model as moderately trusting rather than being forced into a bucket. The deep net then learns over those graded inputs and returns a forecast, again as a degree rather than a hard verdict. This is the architecture that reaches the 90.476% accuracy in the published case, and the reason the fuzzy layer is not decoration is that the grading is what matches the model's representation to the shape of the data it has to predict.

There is a second, quieter reason the fuzzy layer earns its place, and it is about who has to trust the model. An engagement forecast is not a private engineering artefact the way a segmentation mask is; it is a claim about people that other people, project directors, regulators, and sometimes the communities themselves, will want to interrogate. A pure deep net that maps raw signals to a hard label is opaque in a way that is hard to defend in that room. A fuzzy layer, by contrast, exposes its intermediate grading: you can read that the model considered this community's trust to be high and its proximity concern to be moderate, and that legibility is worth real accuracy in a setting where the forecast has to survive a conversation, not just a test set. We would not claim the hybrid is fully

interpretable, because the deep net over the graded inputs is still a deep net, but the front of the model speaks a language a non-specialist stakeholder can follow, and on a decade-long relationship that is not a small thing.

The first instrument holds that whole picture in one frame: the hybrid architecture stated as two stacked layers on the left, and the accuracy it reaches placed next to the accuracy the digitiser reaches, each on its own native scale so nothing is conflated.



Engagement forecasting for a long-lived CO₂-storage project as a distinct model family. On the left the hybrid architecture is stated as two stacked layers: a fuzzy membership layer that grades soft signals (sentiment, trust, proximity) rather than thresholding them, feeding a deep net that returns a forecast rather than a hard label. On the right two accuracy rails stand side by side, each drawn against its OWN native scale and never averaged: the fuzzy-DNN engagement predictor at its sourced 90.476% accuracy (Buah et al., *Energies* 2020, 13, 6259) in orange, and the pure-vision raster-log digitiser at its peak R-squared of 0.9891 in teal. The orange rail is the only element that argues, because the engagement predictor is the model family this whitepaper is about, adjacent to but separate from the digitiser that reads the 136,771-file legacy archive the storage site is selected against. Drag the project-horizon lever to sweep a ten-year engagement: early years lean on the digitiser for site selection, later years ride the engagement forecast as the live community relationship becomes the binding risk. The two accuracies and the archive size are sourced; the per-year lean weighting is an illustrative schematic of where each family carries the project, not a measured allocation.

The console makes the central claim legible without ever averaging the two numbers. The orange rail is the engagement forecast at its sourced 90.476% accuracy, and it is the only element that argues, because that forecast is the model family this paper is about. The teal rail is the pure-vision digitiser at its peak R-squared of 0.9891, drawn against its own ceiling so that the reader can see two strong models without being invited to add them together. Dragging the project horizon shows the second half of the claim: early in a storage project the operator leans on the digitiser, because site selection is reading the archive, and later in the project it leans on the engagement forecast, because the live risk has moved from the rock to the relationship. The two families are not competing for the same job at the same time; they are carrying different years of the same project.

What usable accuracy means and does not mean

We are deliberate about the word usable. A 90.476% engagement forecast is not a promise that nine out of ten community outcomes are locked in; it is a signal strong enough to change what an operator does with its attention and its budget. On a decade-long project, a forecast that is right most of the time about where engagement is heading is the difference between acting on a trend and reacting to a crisis. That is the standard the number has to clear, and it clears it. It does not have to clear the standard of a physical measurement, and pretending it does would be the wrong way to sell it.

The accuracy also does not transfer blindly. It is the reported result for the published CO₂-storage case, with that study's data and that study's definition of the engagement outcome. An operator adopting the family should expect to re-establish the number on its own projects and its own definition of engagement, not inherit it. What transfers is the architecture and the problem shape, not the specific percentage.

THE POSTURE WE ADOPTED ON THE FORECAST

Grade the engagement model on a decision, not on a reconstruction. Keep the fuzzy layer, because the inputs are soft and thresholding them early throws away the signal. Treat 90.476% as evidence the family works, then re-earn your own number on your own projects rather than importing theirs.

THE DIVERGENCE

The two families diverge on every axis that matters

Same project, opposite problem shape

The strongest objection to treating engagement forecasting as its own family is the tidy one: it is all deep learning, so surely it is one system. We think that objection dissolves the moment you lay the two problems side by side on the axes that actually define a machine-learning problem, because they diverge on all of them.

The input diverges. The digitiser consumes raster pixels; the engagement model consumes graded soft signals that have been fuzzified precisely because they are not pixel-crisp. The output diverges. The digitiser returns a reconstructed, depth-indexed curve; the engagement model returns a forward-looking engagement probability. The grading metric diverges, and this is the one people miss most often: the digitiser is scored with R-squared against a ground-truth curve, a regression metric, while the engagement model is scored with accuracy against a realised decision, a classification metric. You cannot put those two numbers on one axis honestly, which is why the first instrument refused to. And the horizon diverges: the digitiser runs a one-off pass and is done, while the engagement model runs for the multi-year life of the relationship.

The second instrument walks those four axes one at a time.

The engagement predictor is not a second head on the digitiser; it is a separate family

AXIS	DIGITISER · pure CNN · R-squared 0.9891 ENGAGEMENT · fuzzy-DNN · 90% acc.	
INPUT	● raster pixels of a scanned log	vs ● graded soft signals, fuzzified
OUTPUT	● a reconstructed depth-indexed curve	vs ● a forward engagement probability
METRIC	● R-squared against a ground-truth curve	vs ● accuracy against a realised decision
HORIZON	● a one-off digitisation pass	vs ● a multi-year community relationship

ON THE INPUT AXIS

the digitiser works in raster pixels of a scanned log; the engagement model in graded soft signals, fuzzified

Click or arrow through the axes: on every one, the two families sit apart.

sourced: CNN peak R-squared 0.9891, fuzzy-DNN accuracy 90.476% (Buah et al., en13236259) · the axis contrasts are qualitative problem-shape statements

ES-1417 · ENGAGEMENT MODEL FAMILY DIVERGENCE MAP · TWO FAMILIES ONE PROJECT

Why the engagement predictor and the raster-log digitiser are two model families rather than one pipeline with two heads. Four axes are laid out as a comparison ledger: INPUT (raster pixels versus fuzzified soft signals), OUTPUT (a reconstructed curve versus an engagement probability), METRIC (R-squared against a ground-truth curve versus accuracy against a realised decision), and HORIZON (a one-off digitisation pass versus a multi-year community relationship). Click or arrow through the axes and each contrast is lifted one at a time, with the footer stating the verdict in a single line. Teal carries the digitiser column, anchored at its sourced peak R-squared of 0.9891; the single orange column carries the engagement family, anchored at its sourced 90.476% accuracy (Buah et al., Energies 2020, 13, 6259), because that family is the subject of this whitepaper. The two accuracies are sourced; the axis descriptions are qualitative statements of the two problem shapes, not numeric claims.

The map is a ledger, not a chart, and that is deliberate, because the argument here is categorical rather than quantitative. Click through the axes and on each one the two families sit apart: pixels against fuzzified signals on input, a curve against a probability on output, R-squared against accuracy on metric, a one-off pass against a multi-year relationship on horizon. The single orange column is the engagement family, carried at its sourced 90.476% accuracy, and the teal column is the digitiser at its sourced 0.9891 R-squared. The point of walking the axes rather than asserting the conclusion is that once a reader has seen the divergence four times in four ways, the idea that these are one pipeline with two heads stops being plausible.

Why the divergence has governance consequences

This is not a taxonomy for its own sake. If the two families are genuinely different, then the operational choices that follow are different too, and collapsing them costs real money and real risk.

Build. A vision digitiser and an engagement forecaster do not share a training pipeline, a labelling process, or a validation harness. The digitiser is validated against digitised ground-truth curves; the engagement model is validated against realised engagement outcomes that take years to accumulate. Trying to run them through one MLOps track produces a pipeline that fits neither well.

Grade. Because the metrics diverge, the review boards that sign off on the two models have to ask different questions. A curve at R-squared 0.9891 is a strong deliverable and the review is about geometric fidelity. A forecast at 90.476% accuracy is a strong deliverable and the review is about calibration, about whether the model is right when it is confident, and about the cost of the errors it does make. Handing an engagement forecast to a review board that only knows how to read a regression metric is how a good model gets rejected for the wrong reason, or a fragile one waved through.

Govern. The engagement model touches people, which pulls in a set of obligations the digitiser never triggers: fairness across communities, transparency of the forecast to the stakeholders it is about, and the question of what an operator is allowed to do with a prediction about a community's future support. None of that applies to a model that reads a scanned log. Treating the two as one system means either over-governing the digitiser or under-governing the forecast, and both are failures.

The failure directions are worth naming, because they are not symmetric. Over-governing the digitiser is the cheaper mistake: it slows down a vision project with review that its risk profile does not warrant, wastes reviewer hours, and frustrates a team whose model is graded on geometry. That is a cost, but a bounded one. Under-governing the forecast is the dangerous mistake, because a model that quietly shapes how an operator engages a community, without anyone checking whether it is fair across communities or honest about its own uncertainty, can steer real decisions about real people on the strength of a number nobody stress-tested. When an organisation runs both families through one governance track calibrated to the vision problem, it defaults toward exactly this second, dangerous

under-governance of the forecast, because the track was built for a model that never had these obligations. That asymmetry is the practical reason we insist the two families be governed apart rather than a philosophical one.

"The model that reads the rock and the model that reads the room need different training data, different reviewers, and different rules. The only thing they share is the site."

— FROM OUR ENGAGEMENT NOTES



THE DECAY

Engagement is a decaying asset, and the forecast buys lead time

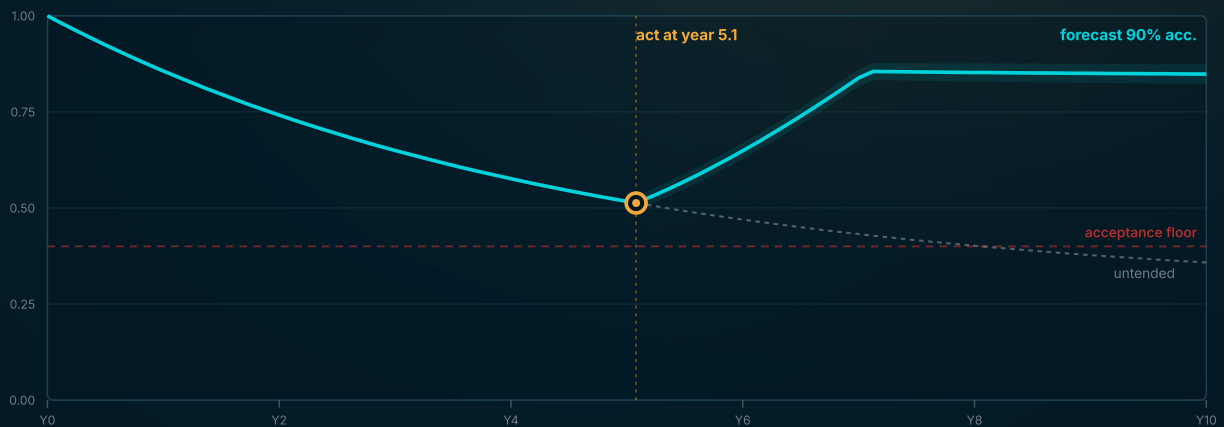
What the forecast is actually for

There is a seductive misreading of an engagement forecast, which is that its job is to keep engagement high. It cannot do that, because it is a forecast and not an intervention. What it can do is see the decay coming early enough that the operator has time to act, and that reframing is the whole practical value of the second family.

Engagement in a long project behaves like a decaying asset. Left untended, a community relationship erodes: attention drifts, early goodwill spends down, and the relationship trends toward an acceptance floor below which the project carries social risk regardless of how sound the geology is. The forecast does not arrest that erosion. It tells you, with 90.476% accuracy in the published case, roughly when the untended path crosses the floor, and that timing is the thing an operator can convert into action.

The third instrument makes the decay and the intervention concrete.

The forecast does not stop the decay; it lets you act before it crosses the floor



LEAD-TIME LEVER

drag how early you act on the forecast:
earlier action lifts more of the curve



sourced: 90.476% forecast accuracy (Buah et al., en13236259) caps the defensible lift · the decay curve, floor, and recovery are illustrative decaying-asset dynamics

ES-2691 · SUBSURFACE ENGAGEMENT TRUST DECAY LADDER · FORECAST BUYS LEAD TIME

Stakeholder engagement in a long-lived storage project read as a decaying asset. The dashed grey curve is an untended engagement index that erodes over a ten-year horizon and drops through the red acceptance floor, below which the project carries social risk. The teal curve is the defended path: the operator acts on the forecast at the orange intervention marker and lifts engagement back up, but only by the fraction the forecast can be relied on to deliver, which is why the recovered lift is capped by the sourced 90.476% forecast accuracy (Buah et al., *Energies* 2020, 13, 6259) and carries a teal confidence band for the remaining uncertainty. Drag the lead-time lever to act earlier or later: acting earlier catches more of the decay and keeps the year-ten index above the floor, while acting late leaves the crossing on the board. The claim is that a forecast does not stop the decay, it buys the lead time to intervene before the floor. The forecast accuracy is sourced; the decay curve, the acceptance floor, and the recovery dynamics are illustrative properties of a decaying-trust asset, not measured project telemetry.

The dashed grey path is engagement left alone: it decays across the ten-year horizon and drops through the red acceptance floor. The teal path is the defended one, and the orange marker is where the operator acts on the forecast. Two things about the defended path are worth reading carefully. First, the operator cannot lift engagement all the way back to full trust; it can only recover the fraction the forecast can be relied on to deliver, which is why the recovered lift is capped by the 90.476% accuracy and carries a confidence band for the remaining uncertainty. Second, when you act matters more than almost anything else: drag the lead-time lever and acting earlier catches more of the decay and keeps the year-ten

index above the floor, while acting late leaves the crossing on the board. The forecast's contribution is the lead time. Everything downstream of that is the operator's own engagement work.

Lead time is the product

We want to be blunt about what an operator is buying, because it is easy to oversell a forecast and easy to undersell it. The operator is not buying a guarantee that a community will stay supportive. It is buying lead time: the number of years between now and the projected crossing of the acceptance floor, delivered with enough accuracy to be worth acting on. Lead time is a real, ownable product. It is the difference between a scheduled engagement programme with room to work and an emergency response after opposition has already organised.

This is also why the second family is worth the cost of building and governing separately. The digitiser saves an operator the manual labour of reading its archive, and that saving is real; the broader upstream context puts the cost saving from machine-learning-optimised operations at roughly 20% [2], which is the kind of number that funds a portfolio. But the engagement forecast protects something the digitiser cannot touch, which is the social licence the whole storage project depends on. A geologically perfect site with a collapsed community relationship is not a store; it is a stranded asset. Lead time on the relationship is what keeps the asset from stranding, and no vision model, however accurate on the curve, produces it.

THE PORTFOLIO

Running two families inside one operator

How the two families share a project without merging

The argument of this paper is that the digitiser and the engagement forecaster are different families, but different does not mean disconnected. They share a project and they hand off to each other across its life, and the practical work is designing that handoff without pretending the models are the same.

In the early years the digitiser dominates. Site selection is reading the archive, and the archive is a vision problem, so the operator's AI attention and its review capacity go to the curve. The engagement model is running in the background at this stage, establishing a baseline of the community relationship, but the binding decision is geological. In the middle and later years the balance inverts. The site is chosen, the rock question is settled, and the live risk is whether the relationship holds. The engagement forecast moves to the front, and the operator's attention follows it. The first instrument's horizon lever is a picture of exactly this handoff: the same project, two families, each carrying the years where its problem is the binding one.

The organisational mistake to avoid is a single team owning both models with a single mental model of what a good result looks like. The vision team's instinct, quite correctly for its own problem, is geometric fidelity, and that instinct is wrong for engagement, where the question is calibration and the cost of confident errors. The engagement team's instinct is about people and process, and that instinct would over-complicate the digitiser, which needs none of it. Two families, two ways of being right, and a deliberate handoff between them is the structure that works.

What an operator should take from this

The transferable claim is not a percentage; it is a way of organising. An operator with a long-lived storage project should hold two model families in its subsurface AI portfolio and resist every pull to merge them. The first family reads the rock: a vision digitiser that turns a raster archive into curves and is graded on the curve, where strong work lands near R-squared 0.9891. The second family reads the room: a fuzzy-DNN hybrid that forecasts stakeholder engagement and is graded on the decision, where the published result is 90.476% accuracy. Build them on different pipelines, grade them with different metrics, govern them under different rules, and design the handoff between them across the project life. The prize for getting the structure right is that neither the rock question nor the relationship question is answered with the wrong model.

WHAT TO CARRY OUT OF THIS

- 1 A long-lived storage project carries **two** machine-learning problems, not one: a bounded vision task that reads the roughly **136,771**-image legacy archive to select a site, and an unbounded relationship task that runs for the life of the project.
- 2 The relationship task is its own model family. A fuzzy-logic layer under a deep net forecasts stakeholder engagement at **90.476%** accuracy in the published CO2-storage case, a distinct architecture from the pure-vision digitiser that peaks at R-squared **0.9891** on the curve.
- 3 The two families diverge on every axis that defines a machine-learning problem: input, output, grading metric, and horizon. That divergence has build, grade, and govern consequences, so collapsing them into one pipeline costs money and risk.
- 4 An engagement forecast does not freeze engagement; it buys lead time. Engagement decays like an asset toward an acceptance floor, and the forecast's value is seeing the crossing early enough to act, with the defensible recovery capped by the forecast accuracy.
- 5 Run both families inside one portfolio with a deliberate handoff: the digitiser carries the early site-selection years, the engagement forecast carries the later relationship years, and neither is graded by the other's metric.

Limitations

The 90.476% engagement-forecast accuracy is the reported result for the published CO₂-storage case in Buah and colleagues, with that study's data and its own definition of the engagement outcome; it is evidence that the fuzzy-DNN family works, not a portable guarantee, and an operator should re-establish the number on its own projects and its own engagement definition rather than inheriting it. The R-squared of 0.9891 is our digitiser's peak on the reconstructed curve and belongs to the vision family; it is placed beside the engagement accuracy only to show two strong models on two native scales, and the two numbers are never combined, because a regression metric and a classification metric are not comparable. The 136,771-image archive size is the sourced scale of the legacy raster corpus in the engagement it is drawn from and describes the vision problem's input, not the engagement model's. In the third instrument, only the forecast accuracy is sourced; the engagement decay curve, the acceptance floor, the recovered-engagement ceiling, and the year-by-year lift from acting early are illustrative dynamics of a decaying-trust asset, drawn to argue the shape of the problem rather than to predict any specific project's telemetry, and the per-year lean weighting in the first instrument is likewise an illustrative schematic of where each family carries the project. The roughly 20% cost saving from machine-learning-optimised upstream operations is the general upstream figure from Koroteev and Tekic and is cited as portfolio context, not as a return specific to either model here. Finally, the governance points about fairness, transparency, and permitted use of engagement predictions are the obligations we would apply, not a legal standard, and any operator deploying a model that forecasts a community's future support should treat those obligations as the floor rather than the ceiling of its own review.

References

- 1 Buah, E., Linnanen, L., Wu, H., Kesse, M. A. (2020). *Can Artificial Intelligence Assist Project Developers in Long-Term Management of Energy Projects? The Case of CO₂ Capture and Storage*. *Energies*, 13(23), 6259. The hybrid fuzzy-logic-plus-deep-net engagement predictor that reaches 90.476% accuracy forecasting social acceptance of a CO₂-storage project. <https://doi.org/10.3390/en13236259>
- 2 Koroteev, D., Tekic, Z. (2021). *Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future*. *Energy and AI*, 3, 100041. The upstream machine-learning cost and speed context, including the roughly 20% cost saving from ML-optimised

operations that funds the portfolio these models live in.

<https://www.sciencedirect.com/science/article/pii/S2666546820300033>

Get the full whitepaper

This page is the long-form summary. The complete whitepaper adds the fuzzy-membership design for the engagement inputs, the calibration protocol we use to re-earn the accuracy number on a new operator's projects, the handoff schedule between the two families across a storage project's life, and the review-board checklist that grades a forecast on calibration rather than on a regression metric.

AI for Long-Term Energy Project Engagement and Stakeholder Management

Published June 2024. Generated 2026-07-22.

<https://www.earthscan.io/whitepapers/ai-long-term-energy-project-engagement-stakeholder-management>

Copyright EarthScan. All rights reserved.