EON AI VENTURES

EON Delta

Why, what, and how.

From the beginning — including every wrong turn.

No jargon until Part 3. If a slide before that needs a technical background, it is badly made.

Internal · For decision · 1 August 2026

EXECUTIVE SUMMARY — READ THIS ONE SLIDE

Where we land

1	**The gap**	Nobody records how many accidents were CATCHABLE. Regulators record what broke. That number does not exist anywhere in the world.
2	**The product**	A small AI that reads a plant's own readings and says, at each shift handover: here is what is most likely to hurt you in the next 48 hours, and here is the check that would confirm it.
3	**The obstacle**	To teach it, we need the boring week BEFORE an accident. Three independent public sources were measured. None of them publish it. Nobody does.
4	**The answer**	A simulator writes that boring week on purpose — validated against the real cases we do have. Published research says simulation alone fails, but simulation plus a little real matches real.
5	**The asset**	Anyone can rent the computers. Almost nobody owns a physics simulator of industrial work with 25 years in it. EON does.

NEW IN THIS VERSION — READ THIS SECOND

The frame changed. Nothing built is withdrawn.

Prediction is three jobs, not one	Material failure (crowded), the person at the moment of work (open), latent organisational conditions (partial). And the conjunction of all three, which is where people actually die.
The open one needs no sensor	A human omission has no trajectory to extrapolate and no physical signature. It needs a device in the hand that knows both the procedure and the equipment. Field IQ, Genesis and Universal are that device.
The fastest product is a parser	Score any procedure against fifty years of human-factors research before anyone performs it. No customer, no sensor, no incident. Demonstrable in a week.
We already hold the best dataset for it	Every failed step in a Genesis simulation is a labelled human-error instance. We have been filing it as trainee assessment scores for years.

Part Five is the new material. None of it replaces the Delta — it tells us what the Delta is for.

THE ASK

What we are asking for

The decision	Run BOTH tracks. Build the procedure linter first — it needs nobody's permission — and proceed with the model narrow: liquids pipelines and slow degradation, 64.6% of incidents in the business our customers actually run.
The money	About $1,200 for the first model. The linter is a week of work on material we already hold. It was never the money.
The timeline	One week to a demonstrable linter. Four weeks to a first trained model, two more to a measured score.
The risk	Public data alone gives a thin corpus. We would rather find that out on $1,200 than on a customer commitment.

One clear number on a small claim beats a vague number on a big one.

0

What you actually get

Before any explanation of how — here is John's Tuesday, before and after.

THE WHOLE THING IN ONE PARAGRAPH

If you only remember one slide

EON has built a prediction engine.

It has studied every industrial accident it can legally read. It knows how equipment wears out. And it reads the live data from YOUR plant.

From those three things together it predicts what could go wrong — live, in the situation the crew is actually standing in — and tells them, with a stated confidence, the one thing to go and check.

Not the procedure done correctly. Not the permit done correctly. A level deeper than either — and deeper again every month.

Already analysed: **21,386** real incidents · **19,057** first-hand accounts · **20.7 million** characters · **40** years

That is roughly 83 full-length books of real industrial failure, already read and counted. Everything else in this deck is the evidence, and the honest account of what is not built yet.

WHAT HAS ALREADY BEEN READ

Twenty million characters of real failure

Before a single line of the model was trained, this much of the world's industrial failure record has already been pulled apart and counted.

21,386	**real industrial incidents**	Every reported US pipeline incident since 1986 — hazardous liquid, gas transmission, gathering, distribution and LNG.
19,057	**first-hand accounts**	Written by the operator who was there, on a regulatory form, describing what happened in their own words.
20.7 M	**characters read and counted**	About 83 full-length books. Every one machine-read, cause-coded and measured — not skimmed.
40	**years of record**	1986 to today. Long enough that the reporting rules themselves changed twice underneath it.

None of this is a plan. It is done, it is counted, and every figure can be reproduced from the same public files by anyone who wants to check.

USE CASE — BEFORE

John's Tuesday today

06:00	Handover is five minutes and spoken. He gets whatever the night shift happens to remember.
06:15	A work order list. Twelve jobs. Ordered by how old the ticket is — not by what might hurt anybody.
09:40	He reaches pump P-114B for a routine check. Does it by the book. Everything within limits. Signs it off.
	Nobody told him vibration has been climbing for nine days. That is on a screen in another room.
	Nobody told him the lube sample is three weeks overdue. That is in a different system entirely.
+11 days	The bearing seizes. Unplanned shutdown. Roughly £400,000, and it could have been a person.

USE CASE — AFTER

John's Tuesday with EON Delta

05:40	Overnight the model read the plant's own memory — sensors, inspections, work orders, permits, last night's log. 312 assets reviewed. Nothing left the site.
06:00	The 3D unit is on the wall. Two things are glowing amber. He can SEE which pump — no tag number, no drawing.
06:02	He touches it. The bearing housing lifts out and turns. The nine-day vibration trend draws itself beside it. Underneath: likely lubricant starvation, moderate confidence, and a real documented case that looked like this.
06:05	ONE thing to do. Not a project. Take a lube sample — and here is where the drain valve is, highlighted on the pump, before he walks out.
09:40	He takes it. It is cloudy. The photograph attaches to the pump, not to an email.
Saturday	Water confirmed in the oil. Bearing changed during planned downtime. About £8,000, no lost production.

Nothing happened. Which is the entire product, and the reason it is hard to sell.

THE DIFFERENCE, NAMED

Five things changed

Before	After

None of these are about the AI being clever. They are about the right thing being in front of the right person at the moment they can still act on it.

£400,000 and a possible injury, against £8,000 and a Saturday. That is the whole business case on one line.

USE CASE — HOW THE MODEL GETS MADE

Six steps, no jargon

1	**We collect accident reports**	Public ones, where somebody has read the licence and signed for it. 21,386 pipeline incidents so far.
2	**We turn each one into a practice question**	Take the moment before it went wrong. Keep what a person could have seen. Hide the ending.
3	**Checkers make sure the question is fair**	Including one kept deliberately ignorant. If IT can guess the answer, the answer was hiding in the question — so we bin it.
4	**The model studies thousands of them**	About $5 to $20 per attempt. The whole first model is roughly $1,200.
5	**We set an exam it has never seen**	Built in a simulator, so the situations never happened and nobody could have memorised them. We publish the results, including our losses.
6	**Then it reads YOUR plant**	And it keeps learning from what it gets right and wrong there — which is the part that compounds.

Steps 1 to 4 are built and tested. Step 5 is designed and costed. Step 6 needs a customer.

1

Why

A thousand people a day, and a number that nobody in the world writes down.

THE PROBLEM

A thousand a day

About a thousand people are killed at work every day, worldwide. The cost runs into trillions a year. Those are numbers that are easy to skim past.

So here is the part that matters instead:

Almost every one of them gave a warning first.

Somebody did not see it — or saw it and did not act.

A pump gets noisier over three weeks. A pipe gets a fraction thinner each year. A checklist step gets skipped on a Tuesday because the shift was short-handed — and by Friday nobody remembers it was a step.

The knowledge already exists

Experienced engineers can read these signs. They do it every day.

But it never ADDS UP.

Money adds up: 100 plus 100 is 200. This knowledge does not. Two plants each with 30 years of hard-won experience have 30 years — not 60.

Nothing collects it. Nothing carries it next door. And when the engineer retires, their thirty years leaves the building with them. That is what "does not add up" means.

WHY IT DOES NOT ADD UP

Two companies, same pump, twenty miles apart

One learns something the hard way. The other will learn the same thing, the same hard way, two years later.

They cannot simply tell each other, and both reasons are real:

It gives away how they operate

Detailed failure data reveals what you run, how hard you run it, and where you are weak. That is commercially sensitive and no operator will hand it over.

In many places it is illegal

Competitors systematically exchanging operating information is exactly what competition law exists to prevent. The lawyers are not being difficult. They are being correct.

So the same accident gets discovered independently, over and over, at full price each time. That is the real problem — and it is structural, not technical.

You cannot buy your way out of a problem shaped like this.

THE GAP

A question with no answer anywhere

Of all the industrial accidents last year — how many could have been seen coming?

Every regulator records what broke. None records whether anyone could have known in advance.

We checked 21,386 pipeline incident reports going back to 1986. The cause is filled in on every single one. The answer to the question above is filled in on none of them.

That number does not exist.
It is precisely the number this programme would produce.

We are not proposing to do better at something everyone measures. We are proposing to measure something nobody measures.

TWO NUMBERS THAT LOOK LIKE THEY FIGHT

“70% is human” and “8.5% is human”

Our own material says more than 70% of oil and gas accidents involve a human or organisational cause. The PHMSA data says incorrect operation is 8.5%. Both are correct. They answer different questions.

PHMSA asks:

WHAT BROKE?

One answer per incident. A physical mechanism. The pipe corroded through. The weld cracked. Somebody opened the wrong valve.

“Incorrect operation, 8.5%” means human error was the ONLY and IMMEDIATE cause — nothing was wrong with the metal at all.

HFACS asks:

WHY DID IT GET THERE?

Many answers per incident. Who set the inspection interval? Who deferred the work order? Who read the report and filed it?

“More than 70%” counts every incident with a human or organisational factor ANYWHERE in the chain — including most of the corrosion.

The corrosion that ate through the pipe was, physically, corrosion. Somebody still set the inspection interval, somebody deferred the work order, and somebody filed the report.

Same incident. 100% corrosion in one column, and three human factors in the other. Neither number is wrong.

WHY THIS MATTERS MORE THAN IT LOOKS

The 70% IS the catchability argument

If accidents were purely physical, there would be nothing to catch — metal fails when it fails.

The reason more than 70% have a human or organisational cause is that somebody, somewhere, had a chance to act — and the chance was missed.

A missed chance is, by definition, a chance that existed.

So the 70% figure is not a competing statistic. It is independent evidence that the number we want to produce — how many were catchable — is large. Two different research traditions, pointing the same way.

And it is why we now build TWO tracks: one that reads drifting numbers, and one that reads missing steps.

Source for >70%: HFACS-OGI framework analysis of oil and gas accidents, Taylor & Francis 2021. Accidents are typically multi-causal.

2

What

Three pieces, and the idea that lets two customers take part without sharing anything with each other.

THE THREE PIECES

A student, a school, and an exam board

THE STUDENT

EON Delta

A small AI trained on industrial accidents instead of on the whole internet.

THE SCHOOL

The Distillery

The machine that turns accident records into lessons the student can learn from.

THE EXAM BOARD

The Shift Test

A public exam that measures whether any AI can spot trouble coming.

We build the exam BEFORE we finish the student. Build the student first and you will design an exam it passes — without noticing yourself doing it.

None of the three works alone.

THE OBVIOUS QUESTION

Why not just use ChatGPT?

1

It learned from everything

Which means it learned from mostly not-this. A tiny fraction of the internet is industrial safety; a smaller fraction still is reading a drifting measurement. A brilliant generalist in a job that rewards a specialist.

2

It lives somewhere else

Using it means sending plant data to somebody else's computer. Many facilities legally cannot. Many more simply will not. Ours is small enough to carry into the building.

3

Nobody can check either claim

Every vendor says their AI understands industrial operations. Not one can show it, because there is nothing to show it on.

Three separate problems. The third is the one that stops every deal.

DISCIPLINE

What we are NOT claiming

We are not claiming to be smarter than the best model in the world.

We are claiming that a model which studied industrial accidents beats one that studied the whole internet — at understanding industrial accidents.

That is a small claim on purpose. Small claims are much harder to attack — and this one happens to be true.

It is also falsifiable in a single afternoon by anybody who wants to try. That is what makes it worth stating at all.

If a frontier model wins on general reasoning but cannot cite a real case, we have proved exactly the point we need to prove.

THE TWO-LAYER IDEA

How a customer's data stays theirs

The model comes in two pieces that click together when it runs, and pull apart at any time.

Layer	Learns from	Lives	Belongs to
Tenant Delta	Only that customer's own material	Inside their building. Never leaves.	Them. Deletable on demand — and we can prove it is gone.
Base Delta	Only patterns confirmed at several different companies	Shared	EON. Licensed back to every participant, free, forever.

Picture a locked room.

Their data stays in the room. We send in a student. The student reads everything, learns, and stays in the room forever. What comes back out is only the general lesson — and only after we have seen it at three or four unrelated companies, so it points at nobody.

This is what lets two customers who compete with each other both take part — without either seeing the other's data.

THE SAME IDEA, CONCRETELY

What leaves the room, and what never does

“A pump starved of oil runs hot before it seizes.” ← this comes out

“The pump at their Baytown site ran hot last March.” ← never leaves

A dataset is a pipe — whatever goes in comes out. A model is a teacher — it carries the lesson without carrying the source.

Which is how we satisfy two of our own rules that otherwise contradict each other.

THE PART THE LAST VERSION LEFT OUT

So how do we actually CATCH anything?

A model on its own catches nothing. It has to sit somewhere, be fed something, and produce something a person acts on. Here is that chain.

WHERE IT SITS

Field IQ Edge

On the tablet or box already in the facility. No cloud. Nothing leaves the site.

WHAT FEEDS IT

The plant's own readings

Historian trends, inspection results, work orders, shift logs, permit status. Data the operator already has and already owns.

WHAT IT PRODUCES

A shift handover briefing

Not an alarm. A short ranked note: what is most likely developing, why, and the one check that would confirm or kill it.

WHO ACTS

The crew

A human decides. The model never touches the plant — it cannot open a valve or stop a pump, by construction.

The Distillery BUILDS the model. Field IQ is where it goes to work. Two different things, and the deck previously only described the first.

A DAY IN THE LIFE

What the crew actually sees

06:00. Night shift hands over to day shift. The tablet has already read the last ninety days.

```
SHIFT BRIEFING — Unit 4, 06:00

1. PUMP P-114B   confidence: moderate
   Vibration 4.1 -> 7.8 mm/s over 9 days.
   Lube sample overdue by 3 weeks.
   Similar pattern preceded a bearing seizure
   in a documented case (ref: CSB-2019-04).
   CHECK: take a lube sample this shift.

2. LINE 22 PERMIT   confidence: high
   Excavation permit open. No locate
   verification recorded in 6 days.
   CHECK: confirm markings before work starts.

3. NOTHING ELSE FLAGGED
   41 other assets reviewed, no pattern found.
```

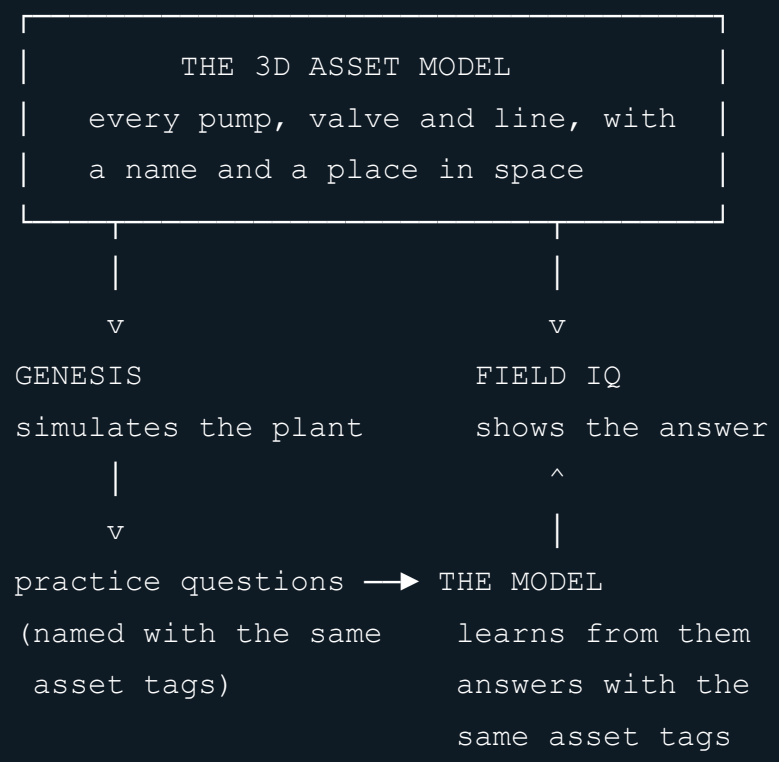
Note what it does

- Ranks — it does not alarm on everything
- Says how sure it is, honestly
- Cites a real case you can go and read
- Names ONE check, not a project
- Reports the 41 assets it cleared
- Covers BOTH kinds: a drifting number and a missing step

The 41-assets line matters: a system that only ever speaks up gets ignored. Saying “I looked and found nothing” is what earns the right to be believed on the day it does speak.

THE THING NOBODY ELSE CAN BUILD

The loop closes on itself



One vocabulary

The simulator, the training questions and the display all use the SAME asset names and the SAME geometry.

So the model's answer is not a sentence about a pump.

It is a pointer AT the pump.

Nobody else owns both ends of that loop.

And a bonus we did not design for: if the model names a pump that does not exist, the pointer fails to resolve — so a made-up answer is caught by a lookup, not by a human.

That last line is now Mechanism L in the patent filing — the most enforceable claim in it, because a customer can see it working.

WHAT THE CUSTOMER IS ACTUALLY BUYING

It shows you. It guides you. Now it warns you.

Three capabilities. Each one is rarer than the last, and the third has never existed.

IT SHOWS YOU

Genesis

A 3D plant you can stand inside and practise in.

Simulators exist. Most companies stop here, and use it only for training.

IT GUIDES YOU

Field IQ

Walks you through the actual job, on the actual equipment, in place.

Already rare. Almost nobody guides a worker visually at the job itself.

IT WARNS YOU

EON Delta

Predicts what is likely to go wrong before it does — and names the one check that would settle it.

New. Nobody predicts AND points at it in the same world.

No general AI can give the warning — it needs a model that studied industrial accidents. No monitoring vendor can give it — they have no 3D world to point into. And nobody can copy it at scale, because it needs both ends of the loop.

And every customer's own model is theirs alone — which is precisely the reason the next one wants to build one too.

3

How

The one dangerous fact, the two locks, and where the lessons actually come from.

THE DANGEROUS FACT

Every accident report is written backwards

It is written by somebody who already knows how the story ended. That sounds obvious, and it is the single most dangerous fact in this whole project.

> *“The bearing was degrading in March.”*

Nobody knew that in March. A scientist established it in June, by cutting the bearing open. But the report states it in the same flat tone as the temperature reading that genuinely was on a screen at the time.

Train on that and the AI learns to be certain about things that were genuinely unclear.

And here is the nasty part

It makes the test scores go UP.

If the answer is hiding in the exam question, the model scores 95% — because every question tells it the answer.

You celebrate. You ship it.

Then in a real plant nothing has happened yet — so the question contains no answer, and the model is lost.

The test lied to you. And the test is the only thing you had.

You cannot test your way out of it, because the test is the thing that is broken. You have to build it out.

THE OBVIOUS OBJECTION — AND IT IS A GOOD ONE

“But the scientist was RIGHT”

The June finding is not a guess. Crack propagation is real science. If a metallurgist says the bearing was degrading in March, it WAS degrading in March. So why throw it away?

We are not throwing it away because it is wrong. We are moving it because it was not AVAILABLE.

Is it TRUE?

YES

The science is sound. The bearing really was degrading in March.

Could the CREW have known?

NO

Nobody cut the bearing open in March. There was no metallurgist on the night shift.

Training a model on it would be like teaching a doctor using the autopsy as the symptoms. “Patient presented with: aortic dissection, confirmed post-mortem.” Perfectly true. Useless for diagnosis — because a living patient does not arrive with an autopsy report.

The word “discard” was wrong. Nothing is discarded. The next slide shows where it goes.

WHERE THE SCIENTIST'S FINDING ACTUALLY GOES

It becomes the answer, not the question

The June finding is not deleted from the system. It is the only reason we know what the right answer was.

THE EXAM PAPER

What the crew could see on Tuesday

Vibration 4.1 rising to 7.8. Lube sample overdue. A whine on start-up.

The metallurgist's finding is NOT here.

THE ANSWER KEY

What actually happened, and why

The bearing seized. The metallurgy showed contaminated lubricant caused spalling.

The metallurgist's finding IS here — it is what makes the example teachable at all.

And the general science — "bearings with this vibration signature fail in about six weeks" — is exactly the lesson we WANT the model to learn. What must never appear in the question is the finding about THIS bearing, which only existed afterwards.

Keep the physics. Remove the spoiler.

THE FIX

Every fact carries a date stamp

Not the date it was written down. The date a person standing in the plant could FIRST have known it.

Stamp	What it means	Allowed in a question?
Instrument	Came off a gauge or a log with a real timestamp	Yes
Stated	Somebody wrote it down at a stated time	Yes
Structural	It was in Tuesday's log, so it was known on Tuesday	Yes
Investigation	Established afterwards, by the inquiry	NEVER
Unknown	The document does not say when	NEVER

Why “Unknown” is thrown away rather than guessed.
It would be easy to estimate a date and note the uncertainty. That is exactly how contamination gets in — the note is not machine-checkable and nobody reads it.

Expect most of every report to be discarded. A machine that found a usable question in every paragraph would be doing something wrong.

THE TWO LOCKS

Permission, and no peeking

LOCK ONE

Permission

Before touching any source, the machine checks a list. Not on the list — refuses. On the list but nobody signed it — refuses. It will not even open the website.

A signature needs a real licence AND a named person who read the terms. A plausible licence with nobody's name is refused, because that is what somebody writes when they assumed rather than checked.

LOCK TWO

No peeking at the answer

Every fact carries its date stamp. Anything the investigation established afterwards is barred from every question, at every setting.

Not barred by arithmetic — barred by where it came from. So even a badly chosen setting cannot let investigation findings slip in.

These are not promises. They are code, and they are tested — 266 automated tests, and about half the machine is tests rather than features.

One test puts unlicensed documents on the disk and confirms the machine still refuses them. Having the files is not the same as being allowed to use them.

THE PIPELINE

How a lesson gets made

```
1  PERMISSION      is this source signed off?  no -> stop
2  READING         pull facts out, date-stamp each one
3  QUESTION        cut the story 48 hours before the accident,
                   keep what was knowable, hide the ending
4  SEVEN JUDGES    check the question does not give itself away
5  BALANCE         no single company may dominate the lessons
6  TRAINING        the student studies
7  RECEIPT         a signed record of exactly what went in
8  UNDO            any contributor can pull material back out
9  FINAL CHECKS    two tests before anything ships
```

Nine steps

Every one can be run on its own.

Every one writes a file a person can open and read.

That is what makes it auditable at all — and it is worth more than the speed a cleverer design would buy.

All nine are built and tested. Two plugs remain: something to read a web page, and something to do the actual training.

The next three slides walk through all nine, because the whole argument lives in the detail.

Making the question

PERMISSION

1

We check a signed list. Not on it — we stop. On it but unsigned — we stop. We do not even open the page.

Because if we train on something we had no right to, everything built on it has to be thrown away later.

READING

2

We pull individual facts out of the report and stamp each one with WHEN a person in the plant could first have known it.

This is the slow, unglamorous step that everything else depends on. Most facts get thrown away here.

QUESTION

3

We pick a moment 48 hours before the accident. Everything knowable by then goes into the question. Everything after is hidden — including the ending.

Now we have a fair exam paper: exactly what a real engineer could have seen, and nothing they could not.

At the end of step 3 we have a question with a known answer that the question itself does not contain — in theory. Step 4 checks whether that is actually true.

STEP 4 — WHY A QUESTION NEEDS CHECKING AT ALL

The problem step 4 solves

Step 3 hid the ending. But the ending can still LEAK into the question without anyone noticing, in ways that look completely innocent:

A word gives it away

A shift note says “bearing noise investigated”. Nobody said it failed — but the word “bearing” tells you where to look.

A date gives it away

Three readings in ninety days, then five in the final week. The sudden attention is itself the clue.

A missing thing gives it away

The one asset with no maintenance record is the one that broke. Absence is information too.

A human reviewer would catch some of these. Not all of them, and not across three thousand questions. So we built seven checkers instead — and one of them is deliberately kept ignorant.

This is the slide that was missing. Step 4 came out of nowhere before.

STEPS 5 TO 9 — FROM LESSON TO SHIPPED MODEL

After the question is approved

5

BALANCE

We count how much of the material came from each company, for each type of lesson. If one dominates, the build stops or trims.

Otherwise the model quietly becomes a portrait of whoever contributed most.

6

TRAINING

The student studies. About $5 to $20 per run.

Small model, small adapter, one rented graphics card.

7

RECEIPT

A signed record: every question used, every source, every setting, and a fingerprint of the model itself.

A trained model cannot be inspected. The receipt is the only auditable thing, so it is bound to the model by fingerprint.

8

UNDO

Any contributor can withdraw their material. Every model built on it is marked invalid and rebuilt without it.

You cannot un-teach a model. So we rebuild instead — and the new receipt proves the material is gone.

9

FINAL CHECKS

Two tests. Did anything leak? Can the model recite its training material back?

Fail either and it does not ship. A build with no recorded checks is treated as NOT releasable — silence is not a pass.

Nine steps, every one auditable, every one able to refuse.

STEP 4

The judge we keep ignorant on purpose

Among the seven judges is one that sees ONLY the question — never the source document. It is asked what happened next.

If the ignorant judge gets it right, the question is thrown away.

Because that means the answer was hiding in the question all along.

It is the inverse of how a model is normally used. Here, the machine being GOOD at its job is the alarm bell — not the goal.

One more rule. The judge that settles disputes must be a different AI from the one that wrote the question. Same family, same blind spots — you would be marking your own homework. Enforced in code, not written in a policy.

This is the piece we have applied to patent.

We need a boring diary

Monday the reading was 4.1. Wednesday 5.2. Saturday 7.8. Then it broke. Nothing dramatic — the dull week where a number drifted and nobody looked.

We went looking for that diary in three places. All three were measured, not guessed.

Source	What we hoped	What we found
CSB chemical investigations	Detailed accident stories	Verdicts, not diaries. For every sentence about what was visible beforehand, four about what investigators concluded afterwards.
PHMSA 21,386 pipeline incidents	Operating data	Forms filled in afterwards. 19,057 written accounts; only 5.9% carry a dated observation — and those start when somebody found the leak.
Situation Room our own 211 courses	Rich case material	ZERO dated readings in 1.6 million characters. Courses teach the lesson. They do not record an instance of it.

Three completely independent sources. Same wall.

THE FINDING THAT CHANGED THE PLAN

Nobody publishes the run-up

Regulators publish verdicts. Operators file forms afterwards. Courses teach the lesson.

Nobody on earth publishes the run-up — because the run-up is boring. Nothing has happened yet.

The boring diary exists in exactly two places:

Inside the plant

The operator's own machine memory. It exists. It is theirs.

Inside a simulator

Where we can make one on purpose.

This is better news than it sounds: it explains why nobody else has built this — a much stronger position than claiming to be first.

And it points straight at the two things EON already has: customers with plants, and twenty-five years of simulators.

NVIDIA Isaac: training in a simulator, working in reality

This is not a new or exotic idea. It is how modern robotics is built, and the reason to believe it is that the results are in hardware, not in slides.

What Isaac is	A physics simulator from NVIDIA. Robots, factories, warehouses, vehicles — simulated with real physics: friction, contact, gravity, momentum.
What they do with it	Train a robot entirely inside the simulation, thousands of copies running in parallel, then load the result onto a real robot that has never moved before.
What they achieved	Zero-shot transfer: a policy trained only in simulation performing gear assembly on a real UR10e arm, with no retraining on the real machine.
The trick that makes it work	DOMAIN RANDOMISATION. They do not build one perfect simulation. They build thousands of slightly WRONG ones — randomising friction, damping, controller gains, sensor noise, part sizes and positions.

A model trained on one perfect simulation learns the simulation. A model trained on a thousand imperfect ones learns the physics. That is the whole idea.

THE SAME IDEA, IN ONE PICTURE

Why thousands of wrong simulations beat one right one

ONE PERFECT SIM

The model learns exactly that world. Move one bolt a millimetre and it fails.

Brittle. Looks brilliant in the lab.

A THOUSAND IMPERFECT SIMS

Friction varies. Weight varies. The sensor is noisy. Nothing is ever the same twice.

The model can only survive by learning what stays TRUE across all of them — which is the physics.

For us this means: do not simulate ONE pump perfectly. Simulate a thousand pumps with slightly different wear rates, oil qualities, load patterns and sensor drift — so the model learns what degradation looks like, not what our simulator looks like.

NVIDIA randomise friction, damping, controller gains and observation noise. The industrial equivalent is wear rate, lubricant condition, duty cycle and instrument error.

If nobody keeps the diary, we write one

Tell a simulator: start a pump, make the oil slightly dirty, run it for two weeks. It runs. Vibration climbs. On day eleven the bearing fails.

Cut the recording at day nine. Throw away the ending. That is a question — and we know the true answer, because we caused it.

Unlimited questions. Every one with a guaranteed answer. And no AI on earth has read them — because they never happened.

The catch is real: a made-up diary is only worth something if experts cannot tell it from a real one.

How we check

We mix simulated situations in with real ones and ask working engineers to say which is which.

If they can tell, we throw the set away and rewrite the generator.

No AI anywhere in that check. Engineers, a list, and a coin-flip threshold.

Exactly how airlines validate a flight simulator.

THE EVIDENCE

Published numbers, not intuition

From a peer-reviewed study using NVIDIA's Isaac simulator. This is the constraint that shapes our whole plan.

Trained on	Score
100 real examples only	0.978
1,000 simulated only	0.626
5,000 simulated only	0.656
1,000 simulated PLUS 100 real	0.983

NVIDIA's own documentation says the same thing from the other direction: they train on simulated data, then test on REAL data — because they do not trust the simulator to grade its own work.

Simulation alone FAILS

0.626 against 0.978. Anyone selling a purely synthetic solution is overselling.

Simulation + a little real MATCHES real

A thousand fake plus a hundred real beat a hundred real alone. That is the whole strategy in one line.

More simulation does not help

1,000 to 5,000 barely moved it. Realism beats volume.

Citing the result that constrains us is what makes the rest of it credible.

WHY US

Anyone can rent the computers

Almost nobody has a physics simulator of industrial work with twenty-five years in it.

NVIDIA's entire argument is that the simulator is the hard part.

EON already owns one.

And the Situation Room turned out to be the second asset: 211 curated cards, every aerospace one anchored to a named official report with tail number and date. Not fuel — an INDEX of which investigations are worth reading.

Somebody already did the expensive work, and did it better than any automated filter will.

4

Everything we got wrong

Seven ideas that did not survive contact with the evidence, in the order they died.

Seven ideas that died

We thought	Why it died	What replaced it
Test AI on recent accidents it cannot have seen	The big labs retrain constantly on material they do not publish. We can never verify a cutoff.	Simulated situations, which provably never happened
Human experts review the training material	Three weeks and too expensive. Most of the job was mechanical. Challenged four times before it changed.	Seven AI judges, including the ignorant one
A training run costs $300–900	Misreading. That was the whole programme.	$5–20 per run. ~$1,200 for the first model.
The customer owns the model	Contradicted our own data rights schedule.	The two-layer split
CSB reports are the source	Measured: four sentences of hindsight for every one of foresight. A verdict is written backwards.	Still used, for the fraction that carries a run-up
Situation Room courses are training data	ZERO dated readings in 1.6M characters. Our own scoring tool rated them highest of everything — it was counting the vocabulary of hindsight.	The index, not the fuel. Tool rebuilt.
Equipment failure is an equipment problem	The filter silently discarded every case where the equipment was fine and somebody skipped a step.	A second track for procedural warnings

The reasoning behind the current plan is only convincing if you can see what it replaced.

THE CHALLENGE THAT WAS RIGHT

“Isn't it only 30% equipment?”

Put to us as a challenge. We checked it against 21,386 real pipeline incidents rather than arguing.

Equipment failing on its own: 31.4%

Cause	Share
Material / weld / equipment failure	31.4%
Corrosion	18.1%
Excavation damage — somebody digging	17.6%
Other / unclassified	11.8%
Incorrect operation	8.5%
Vehicles, weather, ground movement	12.6%

The data agreed with the challenge

Our scoring tool had a rule: if a document does not mention pumps or corrosion, skip it.

It silently discarded every case where the equipment was FINE and somebody skipped a step.

Roughly 39% of the problem, thrown away by a filter built for the easy half.

For LIQUIDS pipelines specifically it is 41.7% equipment and only 8.8% excavation — the combined figure understates their exposure by a third.

THE MOST USEFUL IDEA WE HAVE

Two kinds of warning sign

A DRIFTING NUMBER

The metal is slowly giving up. Vibration climbing. Wall thickness dropping.

Like a tyre slowly going flat.

You can watch it.

64.6% of liquids incidents

A SKIPPED STEP

Nothing is wrong with the equipment at all. No locate ticket. No verification. A handover that never happened.

Like a front door left unlocked.

No reading anywhere shows it.

~39% across all types

Both cause the loss. Only one of them can be watched. They need different systems — and we had built only one.

Your Situation Room cards already cover the second kind. That is why they scored zero on our first test — and why they matter.

5

The engine is bigger than the model

What we thought was one job turns out to be three — and only one of them is open ground.

THE REFRAME

We set out to build a model. We are building an engine.

WHAT WE THOUGHT THE JOB WAS

- Read the accident corpus
- Train a model that knows how things break
- Test it against a benchmark
- Sell the model

WHAT THE JOB ACTUALLY IS

- Predict three different kinds of event
- Two of which no sensor can ever see
- Using instruments we already own and never counted
- And deliver it to a person about to do a job

The model is a component. It was never the product.

Everything in Parts 1 to 4 stands — the corpus, the distillery, the temporal discipline, the two-layer Delta. None of it is withdrawn. But the frame around it was too small. A model that knows how equipment breaks addresses one of the three ways an incident is actually produced, and it is the one every competitor already sells.

THE CORE PROBLEM

“Prediction” is three different sciences wearing one word

A Material

Metal degrades along a physical trajectory. Bearing wear, wall thinning, fouling.

METHOD

Prognosis — extrapolate the curve

WARNING TIME

Weeks to months

CROWDED

B The person, at the moment of work

Slips, lapses, omissions. Forgot to restore the bypass. Opened the wrong valve.

METHOD

Anticipation of omission

WARNING TIME

Seconds to hours

OPEN

C Latent conditions

Barrier out of service. Temporary change never closed. Procedure no longer matches the asset.

METHOD

State assessment — score today

WARNING TIME

Weeks to months

PARTIAL

They differ in what generates the event, in the method that works, and in how far ahead it can be seen. Mixing them produces an engine that is mediocre at all three.

From the Cowork architecture brief. The measured evidence for the split is three slides on.

CLASS A · MATERIAL FAILURE

The physics guarantees a signal. Everyone sells this.

- Degradation is monotonic and governed by physics
- It always leaves a trace — vibration, temperature, current draw, acoustic emission
- If the signal exists, a sensor can find it
- An entire industry has spent twenty years building those tools

Aspen Mtell · GE APM · Bently Nevada · Augury · Senseye · Samotics · Uptake

The honest position: EON cannot beat a vibration specialist at vibration analysis, and should not try. Class A belongs in the engine as an input — not as the product.

One nuance we should hold on to. Corrosion is Class A, is 18.1% of all recorded incidents, and has the LOWEST control-room detection rate of any cause at 13.0%. So Class A is not uniformly well served — it is well served on rotating equipment and badly served on slow metal loss. Do not fight vibration specialists at vibration. Corrosion prognosis is open ground.

Everyone is predicting the machine.

Nobody is predicting the omission.

Forgot to restore the bypass. Opened the valve they opened last time, on a different unit. Skipped the gas test because the last four were clean. Resumed after an interruption at the wrong step.

No trajectory exists

There is nothing to extrapolate. The method is anticipation of omission, not prognosis.

No sensor can see it

The thing being predicted is an intention — and intentions have no physical signature until they are already consequences.

Why nobody covers it

It needs a device in the hand that knows both the procedure and the equipment. Field IQ + Genesis + Universal is exactly that device.

This is the class that no predictive-maintenance product on the market can reach, for a structural reason rather than a technical one.

CLASS C · LATENT CONDITIONS

Registers exist. Prediction does not.

- Barriers degraded or out of service
- A temporary change installed for a turnaround and never closed out
- A procedure that no longer matches the asset it describes
- Safety-critical maintenance overdue
- The experienced operators all rotated out of the crew
- Fatigue accumulated across a run of night shifts
- Four permits live in the same congested area at once

These are not forecasts. They are the current state of the plant's defences — readable today, if anyone bothers to read them.

Barrier and bowtie tools already hold this — as a register. A register is paperwork. None of them connect the degraded barrier to the specific task a named person begins at six o'clock tomorrow morning.

CLASS D · THE CONJUNCTION

Where people actually die

A **Equipment degrades quietly for months**

C **The barrier that would have caught it is out of service**

B **A person under time pressure misses the check that remained**

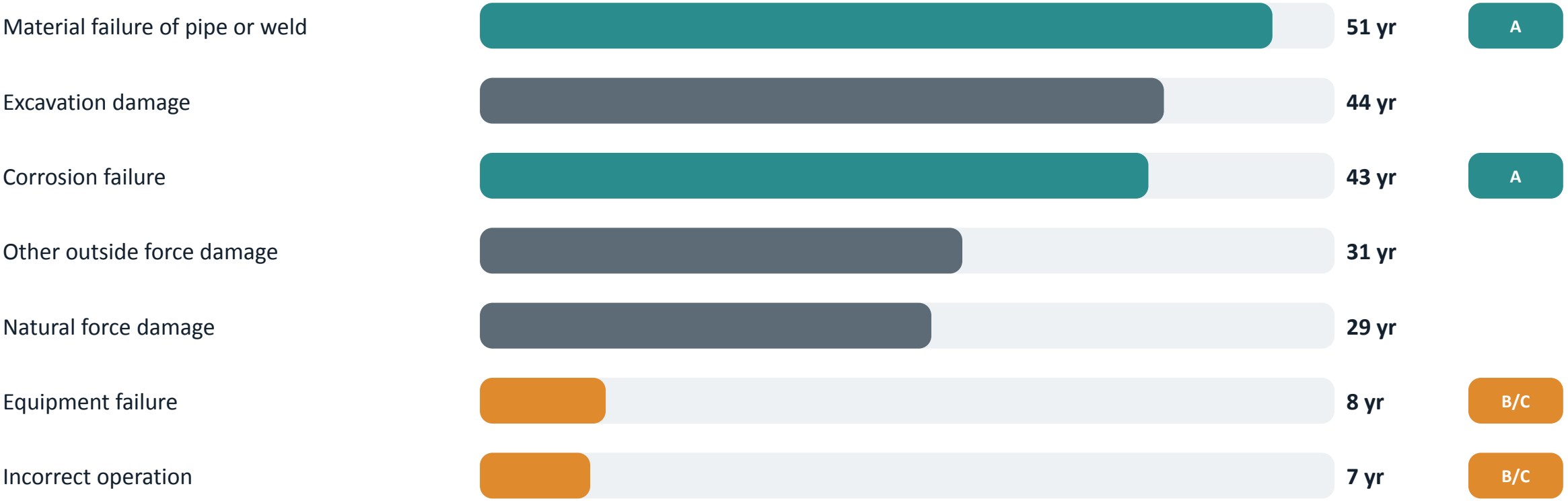
The event is the conjunction. Very few serious incidents are caused by one class alone.

Uncovered for a structural reason: nobody else holds all three inputs. Analytics vendors hold A. EHS platforms hold a lagging version of C. Nobody holds B at all — because B only exists where the work is being done.

AND IT IS NOT JUST A NICE FRAME

The split shows up in the federal data on its own

Median years from installation to failure. The taxonomy was argued from first principles — then it appeared in the numbers unprompted.



Class A mechanisms kill assets in decade five. Class B and C mechanisms kill them in decade one. Two regimes, two horizons — and one model with one horizon serves neither.

n = 7,187 incidents recording both installation year and incident year. data/engine.py.

MEASURED, NOT ASSERTED

We can count Class B against Class C

PHMSA asks contributing-factor questions that map onto the taxonomy almost exactly. Within incorrect-operation incidents on forms carrying those questions (n = 1,095):

51.6% — CLASS B — **A person failed to follow the procedure**

24.9% — CLASS C — **The procedure itself was inadequate, or none existed**

62.6% — CLASS C — **Operator-qualification programme implicated — 610 of 975, a real yes/no field**

Caveat on the record: most of these are tick-only fields, so the denominator is the population whose form carried the block — not all 21,386 incidents. The qualification field is the one with genuine yes/no values, which is why it is called out separately.

AND HOW THE CURRENT STACK PERFORMS

Every layer fires. It breaks anyway.

24.3% **SCADA was the detector**

where installed and running. A human found the other three in four

63.4% **failed within 2 years of inspection**

92.6% within five. Lower quartile: zero years

63.3% **of dig strikes followed a proper ticket**

the excavator called, the line was marked, it was hit regardless

13.0% **SCADA detection for corrosion**

the slowest and most predictable mechanism there is

Not a story about missing sensors. Four correct observations that were never put in the same room and asked one question: taken together, what does this mean, and when does it matter?

Independently corroborated: a 2026 peer-reviewed CO_2-pipeline study, same federal source, different subset, reports SCADA detection at 19.7%.

All four reproducible from the public federal files by data/engine.py, in about four minutes.

THE CLEANEST CLASS B MEASUREMENT WE HAVE

They called before they dug. They hit it anyway.

Excavation damage is 3,766 of 21,386 incidents — 17.6% overall, and 38.9% of everything in gas distribution.

63.3% of excavation strikes happened AFTER a one-call ticket was properly raised

The system worked exactly as designed. The excavator called. The locate was requested. A ticket exists — 614 of these incidents carry the number on the federal form. The line was marked. And it was struck anyway: a mismark, a wrong depth, a tolerance-zone violation, a machine over a painted line.

That is Class B, measured, in a cause category nobody thinks of as human error. No sensor sees it. No condition-monitoring vendor addresses it. 17.6% of all incidents, squarely inside the reach of a guided technician with a visual instrument.

CLASS B IN DEPTH

Which step gets forgotten is predictable

Fifty years of human-factors research. The structural features that make a step error-prone are known and enumerable — and none of this requires customer data.

Post-completion steps	Replacing the guard, closing the drain, restoring the bypass. The mind marks the task done at the goal and the tail of the sequence evaporates. The single most studied error pattern that exists.
Steps with no feedback	Nothing visible, audible or tactile confirms it happened — so nothing signals that it did not.
Steps after an interruption	Resumption reliably restarts at the wrong point, usually one step late. Field IQ can detect the interruption directly.
Invisible restorations	The abnormal state was created deliberately and is remembered. Its removal is not.
Differs from the next unit	People perform the sequence they perform most often, not the one in front of them.
Survivable most of the time	The check that has been clean forty times stops being performed on the forty-first.
At a handover boundary	Ownership of a partially completed sequence is where sequences are lost outright.

Every one of these can be detected in a written procedure by a parser. That is the next slide, and it is the most immediately buildable thing in this deck.

CLASS B IN DEPTH

Published, calibrated, and never once used live

Human reliability analysis supplies factors that multiply error probability — developed for nuclear regulatory use and validated over decades. Every one is already known to EON at the moment of work.

Available time	In the permit
Experience and training	In the competence record
Fitness for duty / fatigue	Derivable from the roster
Ergonomics and instrumentation	In the Universal equipment model
Task complexity	In the procedure
Procedure quality	Scored statically — see next slide
Stress and time pressure	Task context, concurrent work
Work process quality	Permit IQ and handover records

The legitimacy argument, and it is worth more than the accuracy.

When a major asks why the system said what it said, the answer is a named, published, regulator-accepted method applied to observed conditions — not an opaque model output. That is what makes a prediction defensible inside a safety case, and it is something no pure machine-learning product can offer.

THE THING WE CAN DEMONSTRATE IN A WEEK

A linter for procedures

Every SOP can be parsed before anyone performs it, scored against the error-prone features. No incident. No telemetry. No customer data — only the procedure we already receive to build a Genesis simulation.

```
$ eon-lint  SOP-4021-pump-changeout.docx

  step 14  POST-COMPLETION      replace coupling guard      risk 0.81  HIGH
           -> after the stated goal; no feedback; commonly dropped
  step 09  INVISIBLE RESTORE    close 2-inch drain valve    risk 0.74  HIGH
           -> abnormal state created at step 6, removal not cued
  step 22  NO FEEDBACK          confirm CP lead reattached  risk 0.58  MED
  step 17  UNIT DIVERGENCE      differs from B-train order  risk 0.44  MED

  4 high-risk steps  ·  procedure score 62/100  ·  3 fixes suggested
```

It predicts risk before a single person is deployed. And it turns every procedure a customer sends us for simulation into an immediate diagnostic finding — which makes it a sales instrument as much as a product feature. Nothing comparable appears to be on the market.

Illustrative output format. The scoring function is a work package, not a result — no procedure has been scored yet.

AND THE PART THAT SHOULD STING A LITTLE

Three sources we already generate and never counted

1

Genesis trainee failures

Every failed step in a simulation is a labelled human-error instance: the exact step, the exact equipment, the exact experience level, the exact preceding sequence. We have been generating the best Class B training set in the industry for years and filing it as assessment scores.

2

Field IQ behavioural telemetry

Time on step, back-navigation, repeated re-reads, help requests, hesitation before an action, abandoned and resumed sequences. The live signal that a person is about to make an error — as opposed to the record that they made one.

3

Procedure static analysis

Lint every SOP before anyone performs it. No incident, no telemetry, no customer data — only the document we already receive to build the simulation.

None of the three requires a single conversation with a customer's IT department.

The three deltas are how we measure, not what we predict

These are frequently confused with the three event classes. They are not the same thing — the classes are what is being predicted, the deltas are the instruments doing the measuring.

Delta	What it measures	Senses	Source	Latency
Procedure	What the person did, against what the procedure specified	B, drift into C	Field IQ at the point of work	Seconds
Physics	What the plant did, against what the EON Live twin says it should have	A, early edge of D	EON Live residual vs. observed	Minutes–weeks
Competence	What this person can demonstrably do, against what the task demands	C, pre-condition of B	Genesis + Assess IQ + Competence IQ	Weeks

- They fail in different directions. A procedure delta with no physics delta is an unexplained deviation. A physics delta with no procedure delta is an unattributed anomaly. Together they are a diagnosis.
- The physics delta is a second sensor-independent signal — the twin's residual works on sensors that already exist.
- The competence delta is the multiplier. The same deviation means something different from a fifteen-year operator than from someone three weeks in.

OUTPUT DESIGN

What the six a.m. screen actually says

NOT THIS

“This compressor will fail in 14 days.”

A date of failure is a prophecy. It is either right, or it is a wolf — and after two wolves nobody looks at the screen again.

THIS

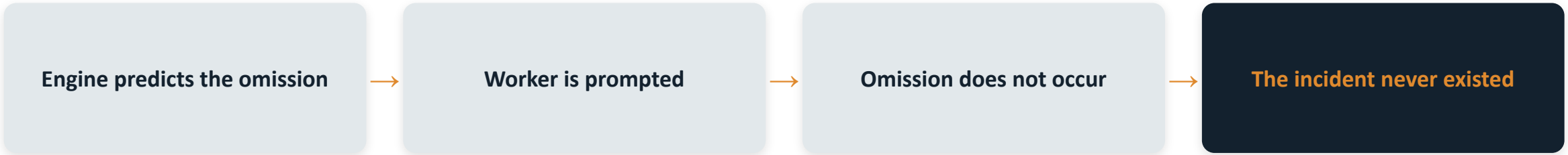
“For the task you are starting: three credible failure paths for this equipment in this state, two barriers currently degraded, the step most commonly missed on this procedure — and the one check that reduces your risk most today.”

A conditional statement cannot cry wolf, because it is not an alarm. It is an explanation offered to somebody who is about to do something.

This is a better answer to the false-alarm arithmetic than any improvement in accuracy could be. We were trying to raise precision. The real move is to change the output type so precision stops being the binding constraint.

THE COMMERCIAL LANDMINE

A working engine erases its own evidence



Measure prompted behaviour change — not incidents avoided.

- The predicted step was flagged, the worker acted on the flag, and the delta closed
- Calibration of every prediction against the outcome ledger
- The historical miss rate for that step — how often it would otherwise have been dropped
- Near-miss and stop-work frequency as the leading indicator

Design this before the first pilot. It is the most likely way for a technically successful engine to fail commercially.

ARCHITECTURE

Two tiers — and the second is never a precondition

TIER 1 — WORKS ON DAY ONE

- The customer's procedures
- Equipment class and the Universal model
- Task and permit context
- The EON accident corpus
- EON Live physics
- Genesis trainee failure data
- Procedure static analysis
- Field IQ behavioural telemetry

TIER 2 — SHARPENS OVER TIME

- Historian and vibration
- CMMS work-order history
- Alarm and override logs
- Management-of-change records
- Barrier and bowtie registers
- Inspection and integrity data
- Rosters and organisational state
- Near-miss and stop-work cards

Pilots that wait for a historian connection die in procurement. No Tier 1 capability may depend on Tier 2. And commercially: customer data makes that customer's instance better — it is the cross-customer pattern layer that makes EON's asset compound.

METHOD

Five rules for the build

1	**Negatives are half the training set**	Every uneventful Field IQ episode is a negative label. Positives alone teach what an accident looks like, not what distinguishes it from an ordinary Tuesday — which is exactly how false alarms are manufactured.
2	**Base rates before models**	Standard failure-rate data on the Universal taxonomy gives priors per equipment class. Without priors, rare-event models are miscalibrated by construction.
3	**The simulator sets coverage, never probability**	EON Live establishes what can happen and by which mechanism. Real data establishes how often. The simulator must never set a threshold.
4	**Keep an outcome ledger**	Every prediction stored with its conditions, its recommendation, and what actually happened. Without it the engine can never be shown to work.
5	**Explanation, not prophecy**	A conditional tied to the task cannot cry wolf, because it is not an alarm.

Rule 1 is the one we were already getting wrong, and it is the same finding as the intervention confounder in the prediction-engine paper.

THE BUILD

Work packages, in order

	Work package	Depends on	Output
P0	Define the delta formally — the event schema for procedure, physics and competence	Nothing	Schema document
P0	Procedure static analysis — score every SOP against the error-prone features	Procedures held today	Working linter + demo
P0	Mine Genesis trainee failures into a step-level error-frequency table	Assessment records	Error-rate table
P1	Live performance-shaping-factor scoring at the point of work	Permit IQ, Competence IQ	Scoring module
P1	Physics-delta residual — the twin's predicted state against observed	EON Live gas physics gap	Residual detector
P1	Base-rate library aligned to the Universal equipment taxonomy	Universal library	Prior probability tables
P2	Outcome ledger and calibration reporting	Delta schema	Ledger + reliability curves
P2	Tier 2 connectors — historian, CMMS, alarm, MOC — additive only	Customer agreement	Connector specification
P2	Conjunction scoring across A, B and C — the Class D model	All of the above	Risk conjunction model

The three P0 items need no customer, no sensor and no agreement. Two of them run on material we already hold.

ANSWER THESE — DO NOT ASSUME THEM

Six open questions

1. How thin can Tier 1 be before the six a.m. output stops being credible to an operator who knows the plant better than the system does?
2. Where is the threshold between useful prompting and alarm fatigue — and who controls it: the engine, the supervisor, or the worker?
3. Does the error-frequency table generalise across equipment classes, or must it be rebuilt per class and per sector?
4. At how many deployed Field IQ seats does the field delta start contributing more than the static analysis does?
5. **How is a prediction about a named individual handled, contractually and ethically, under our worker-protection commitments? Answer before the first pilot, not after.**
6. Which of the three deltas is genuinely necessary for release one — and which are being carried because they are conceptually elegant?

Question five is the one that can stop a pilot dead, and it is a legal question rather than a technical one. It needs an answer from counsel, not from us.

PART FIVE, IN ONE PARAGRAPH

What actually changed

The model is a component	not the product. Everything built so far stands — the frame around it was too small
Three classes, not one problem	A is crowded, C is partial, B is open, and D is where people die
The fastest thing we own is a parser	the SOP linter needs no customer, no sensor and no incident — and demos in a week
We already hold the best Class B dataset	and have been filing it as trainee assessment scores for years
Stop promising dates	a conditional tied to the task cannot cry wolf, which dissolves the false-alarm problem instead of fighting it
Design the metric before the pilot	because an engine that works produces nothing to count

None of this replaces the Delta. It tells us what the Delta is for.

6

The decision

Where we are, what it costs, and the three honest options.

THE STRATEGY

In one paragraph

Public records prove the method.

The simulator manufactures the volume.

The operator's own data is the prize — and we only earn the right to ask for it once the first two have produced a number.

Real first

Trustworthy and scarce. About 1,300 anchor cases available.

Simulated second

Unlimited, and needs the real material to validate against.

Customer data last

Asking before we have a number is asking them to take a chance on us.

Asking afterwards is offering them a measured thing. That is the whole difference, and it is why the exam comes before the sales call.

Real proves it. Simulator scales it. Never the reverse.

STATUS

Where we actually are

Piece	State
The Distillery — all nine stages	Built and tested. 266 automated tests. Runs end to end.
Permission register	Working. CSB, PHMSA and the Situation Room signed off. Nine sources still pending.
CSB crawler	Built, and run against the live site.
PHMSA data	21,386 incidents downloaded and analysed.
Base rates	Computed. The first thing we have that nobody else in this market has bothered to work out.
The document reader	NOT BUILT. Needs the crawl work.
The trainer	NOT BUILT. A thin wrapper around standard tools.
A trained model	Does not exist yet.
Any performance number	Does not exist yet — and none appears in any document.

Two plugs and a decision away from a first result.

COST AND NEXT SOURCES

It was never the money

Item	Cost
Everything up to the AI judges	$0
Seven judges over ~3,000 questions	~$400
Renting a machine to train on, ~20 runs	~$400
Final checks	~$150
Contingency	~$250
TOTAL for the first model	~$1,200

It was never the money. It is the permissions and the data.

Still worth chasing

NASA ASRS

Tens of thousands of NEAR-MISS narratives. The only public source that is precursor without an accident. Ranked first.

National Response Center

~30,000 spill reports a year. The largest structured feed nobody uses this way.

EPA RMP

Mandatory 5-year accident history for every US chemical facility.

Published degradation physics

Crack growth and corrosion equations are textbook. Physically correct series TODAY, without simulator integration.

NASA ASRS is the one I would chase first. Near-misses are precursors where nothing happened — exactly our shape.

THE DECISION

Three honest options

Option	What it means	Risk
Proceed as planned	Build the reader, train on public data, run the Shift Test, then approach an operator.	Four weeks and ~$1,200 before there is anything to show.
Narrow first	Pick one equipment class — pumps, or corrosion — and go deep before widening.	Slower to a headline, but a far more defensible first number.
Stop and reconsider	The evidence says public data alone gives a thin corpus. If that is disqualifying, better to know now.	We leave the base rates and a built machine on the shelf.

Recommendation: proceed, but narrow.

Liquids pipelines and slow degradation — 64.6% of incidents in the business EON's customers actually run, and the half with a watchable warning sign.

One clear number on a small claim beats a vague number on a big one. It is also the version that survives being attacked — which it will be.

Everything above is either built, measured, or costed. What remains is a direction.

TO CLOSE

Nobody records how many accidents were catchable.

That number does not exist. It is precisely the number this would produce.

No performance figure, score or comparison appears anywhere in this deck. Nothing has been trained and nothing has been measured. Every number here is either a count from a public file or a published result from someone else's research — and both can be checked.

Real proves it. The simulator scales it. The customer's data is the prize.

EON AI VENTURES · INTERNAL

The Prediction Engine

What it is made of. What each part contributes. What we missed. And why the layer nobody captures is the one that matters most.

Every layer of the existing stack already works. The equipment fails anyway.

Version 1.0 · Companion to the white paper of the same name

WHY THIS DOCUMENT EXISTS

Three questions that kept coming back

1

“You keep saying we need their data. What exactly do we need?”

Section 7 answers it item by item, with the reason each one is answerable.

2

“How much does each of our five sources actually contribute?”

Section 4. And the honest answer is that four of the five cannot predict anything on their own.

3

“How does prediction work today, and are we better?”

Section 3. Measured from the federal record, not from vendor marketing.

All three are answered from measured evidence. Where a number is a judgement rather than a count, the slide says so.

THE CONCLUSION, BEFORE THE ARGUMENT

Every layer fired. It broke anyway.

24.3% **of incidents were detected by SCADA**

where SCADA was installed and running. A human found the other three in four.

63.4% **of failures came within 2 years of inspection**

The tool ran. The report was read. 92.6% within five years.

63.3% **of dig strikes had a one-call ticket**

The excavator called. The line was marked. It was hit regardless.

13.0% **SCADA detection for corrosion**

The slowest, most predictable mechanism there is — and the least visible.

Not a story about missing sensors. A story about four correct observations that were never put in the same room.

All four computed from operator-completed fields on federal incident forms. Reproducible: data/engine.py.

THE CONCLUSION, PART TWO

We have five sources. Only one predicts.

Four of them tell the model what failure looks like after the fact. That makes it knowledgeable. It does not make it predictive.

Source	What it supplies	Est.	Do we have it?
Plant signals — vibration, DCS, history	THE INPUTS. The run-up itself	35%	No — customer owns it
Simulation (Isaac-class)	Negatives. The cry-wolf problem	20%	Designed, not built
Field IQ / the Delta flywheel	Human intervention + outcome	20% ↑	Exists nowhere. Ours to make
The accident corpus (21,386)	Labels, taxonomy, base rates	15%	Yes — public, processed
The Situation Room	Causal chains; the answer key	10%	Yes — EON owns outright

Percentages are engineering judgement, not ablation results. They say what each source structurally can and cannot supply. The experiment that would replace them with measurements is in section 9 — until it runs, treat this table as a hypothesis.

The two sources at the top of the table are the two we do not have. That is the whole strategic problem in one row-order.

What a prediction engine needs

Five ingredients. They are not substitutes for each other.

STATE THE PROBLEM ONCE

What we are actually estimating

```
P( failure of asset a in window [t, t+h]  |  everything knowable at t )

# fitting it requires five different things:

   Y    labels          what failed, how, when
   X    inputs          observations from BEFORE  <- the run-up
   N    negatives       same run-up, nothing happened
   Pr   priors          physics that constrains the extrapolation
   I    interventions   what a human did in between
```

The fifth term is the one everyone forgets. If a human acted between t and t+h and nothing failed, that is not a negative. It is a positive that was interrupted.

More labels do not compensate for missing inputs. A million perfect failure records will not predict anything if none of them contains what the asset looked like the week before.

THE ARITHMETIC THAT GOVERNS EVERYTHING

Why a 99% model is unusable here

```
prior of failure in window   pi = 0.0001
sensitivity                     = 0.99
specificity                     = 0.99

precision = (pi x .99) / (pi x .99 + (1-pi) x .01)
          = 0.0098
```

~100

false alarms
for every real catch

This is why honest vendors are cautious and the rest are not. You cannot fix it with better sensitivity.

So negatives matter more than positives here.

You escape the false-alarm trap by having seen enough near-misses that resolved to tell the difference between a trajectory that ends badly and one that does not. Real operations produce far too few of those to learn from in any reasonable time. That is what the simulator is for — and it is the whole of its 20%.

Alarm fatigue is not a UX problem. It is this equation, showing up in the control room.

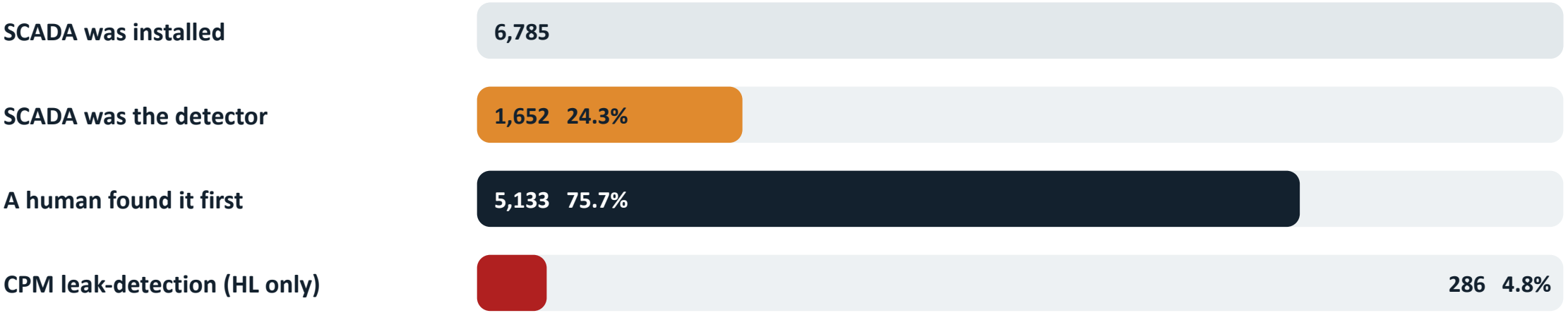
How prediction works today

Measured from 21,386 federal incident reports. Not asserted.

THE CONTROL ROOM

Three in four incidents are found by a person

PHMSA asks, on every report, whether SCADA-based information was how the incident was identified.



Independently corroborated. A 2026 peer-reviewed study of CO_2 pipeline failures, same federal source, different subset, reports SCADA detection at 19.7%. Our 24.3% across the full corpus sits alongside it.

And note: this is DETECTION — noticing it as or after it happened. Not prediction. The industry does not record predictions, so that question cannot even be asked of this data.

AND THE PART THAT SHOULD WORRY THE INCUMBENTS

The most predictable mechanism is the least visible

Corrosion is slow, physical, well understood and inspectable — the longest, most legible run-up in the industry. And it is the one the real-time layer sees least, because SCADA watches flow and pressure, and corrosion moves neither until the wall is gone.

THE INSPECTION LAYER

It was inspected. Recently. It failed anyway.

For 781 incidents we can compute the gap between the last dated in-line inspection and the failure.

63.4%

failed within 2 years of being inspected

495 of 781

92.6%

failed within 5 years

723 of 781

2 yrs

median gap between inspection and failure

lower quartile: 0 years

A quarter of them failed in the same calendar year they were inspected.

Read it correctly. This is conditioned on failure. It describes assets that failed and says recent inspection did not protect them. It is NOT a measure of inspection efficacy in general — this data cannot show the far larger population that was inspected and was fine. Nobody publishes that.

Stated precisely: among assets that failed, almost all had recently been measured by the best tool available, by a competent contractor, and the result was read and filed. The information was in the building. What was missing was the step from a measurement to a date.

THE PROCEDURAL LAYER

They called before they dug. They hit it anyway.

Excavation damage is 3,766 of 21,386 incidents — 17.6% overall, and 38.9% of everything in gas distribution.

63.3%

A one-call ticket WAS raised

611 of 965 where the answer is recorded. 614 incidents carry the ticket number on the federal form.

36.7%

No notification was made

354 of 965. This is the share the industry's safety campaigns are aimed at.

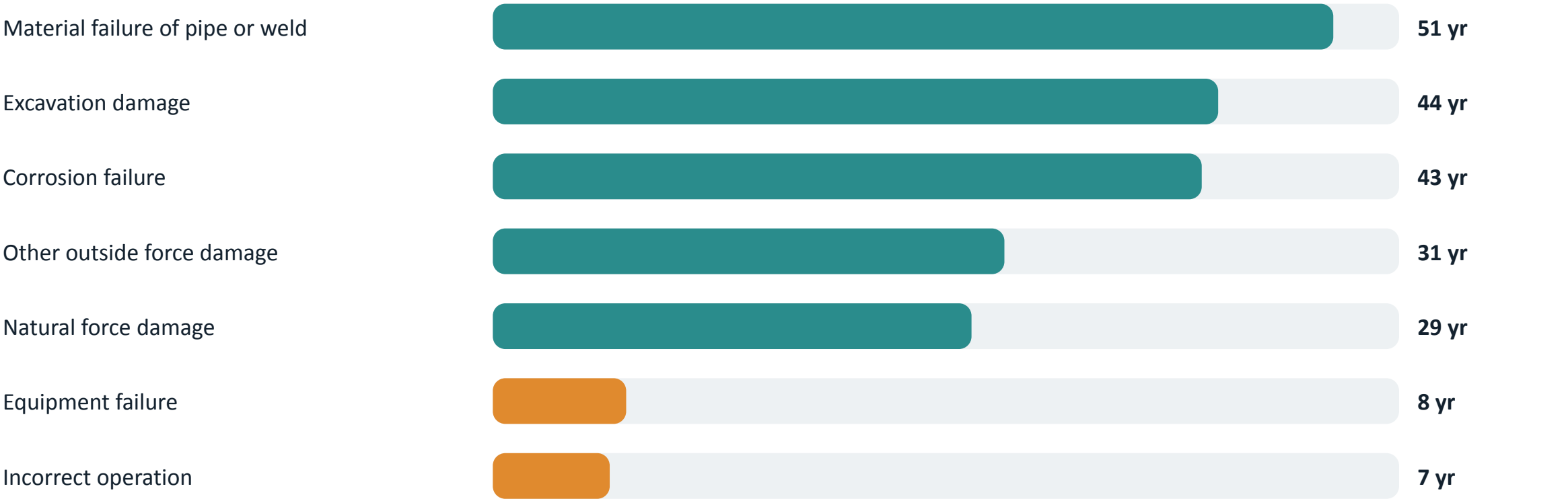
In two thirds of strikes the system worked exactly as designed.

The excavator called. The locate was requested. The line was marked. And it was struck regardless — a mismark, a wrong depth, a tolerance-zone violation, a machine over a painted line. No sensor on earth sees that. No condition-monitoring vendor addresses it. It is 17.6% of all recorded incidents, entirely outside the technology the industry is currently buying, and squarely inside the reach of a guided technician with a visual instrument.

TWO DIFFERENT PROBLEMS WEARING ONE NAME

Age at failure splits the population in half

Median years from installation to failure, for the 7,187 incidents recording both dates.



SLOW · metal loss and material defects kill in decade five and six. Inspectable. Forecastable years ahead.

FAST · equipment and human error kill in decade one. Days to weeks of warning, if any.

One model with one horizon serves neither. A vendor selling one horizon is selling to half the problem.

The five sources, audited

What each one supplies — and what it structurally cannot.

SOURCE 1 OF 5 · ESTIMATED CONTRIBUTION 15%

The accident corpus

21,386 incidents · 19,057 first-hand accounts · 20.7M characters · 40 years

SUPPLIES

- Labels (Y) at high quality and completeness
- The mechanism taxonomy — how things really break
- Base rates by pipeline type, cause and age
- Forty years of drift in how failure is classified

DOES NOT SUPPLY

- Inputs (X). There is no run-up in a post-hoc report
- Negatives. Only failures get reported
- Interventions. The form never asks
- Anything at all about when

High floor, low ceiling. Without it nothing is calibrated and the output has no units — you cannot say “12% chance” unless you know the base rate. But it makes the model knowledgeable, not predictive.

Expected yield after temporal-cutoff and knowability filtering: 3,000–5,000 usable examples, all on the label side. Any competitor could obtain the same files in a week — the moat is not the data, it is the extraction discipline.

SOURCE 2 OF 5 · ESTIMATED CONTRIBUTION 10%

The Situation Room is the answer key, not the exam paper

1,133 anchor cases. The aerospace set is 31 of 31 traceable to a named NTSB or AAIB report.

It is curated	A human expert already reconstructed the causal chain from initiating condition to outcome. That is exactly the structure a raw narrative lacks.
So it evaluates	It is how we check whether corpus extraction found the right chain — because here we already know the answer.
It carries the pedagogy	How to explain a finding to someone who must act in the next ten minutes. A different skill from predicting.
And we own it outright	No licence question. The only substantial corpus here we can use without a rights argument.

We misjudged this early on — treated it as fuel. It is smaller than the corpus by an order of magnitude and worth more per case.

SOURCE 3 OF 5 · ESTIMATED CONTRIBUTION 35% · THE LARGEST

Plant signals — the one we do not own

Vibration, process history, control-room actions. This is X. Without it there is no prediction, only knowledge.

```
one mid-size unit  x  500 tags  x  1 Hz  x  24 months
                                ~  3 x 10^10 samples

# against which the entire 40-year federal corpus is a rounding error
```

Volume dwarfs everything else combined
By four orders of magnitude. Nothing else is close.

And it is almost entirely useless alone
Unlabelled, and nearly all negative. In 24 months on one unit: perhaps two or three real incidents.

This is why the corpus and the simulator are not optional extras. The customer's own data cannot teach a model what failure looks like — it barely contains any.

SOURCE 4 OF 5 · ESTIMATED CONTRIBUTION 20%

The simulator, and what the evidence actually says

NVIDIA's published sim-to-real results, read honestly rather than enthusiastically:

Synthetic alone is poor. Combined is better than real alone by 0.005. That is a small margin and we should not oversell it. The case for simulation is not that it beats reality.

The case is that it supplies a class of example reality never produces: the near-miss that resolved. Real operations generate very few and never label them. A simulator makes unlimited alarming-looking trajectories that end safely, each with ground truth — the only known way to attack the 100-false-alarms arithmetic at its root.

Caveat recorded: this evidence is robotic manipulation, not industrial degradation. The transfer is assumed, not shown.

AND THE FINDING THAT LOWERS THE BAR

A thousand rough simulations beat one perfect one

Domain randomisation — the most useful result to come out of the Isaac line of work, and the one that decides how much we have to build.

WHAT WE FEARED WE NEEDED

- A validated engineering twin of each asset
- Correct material properties
- Calibrated against real failure data
- Signed off by an engineering authority
- Years of work, per asset class

WHAT THE EVIDENCE SAYS WE NEED

- Thousands of rough simulations
- Deliberately varied and imperfect
- Randomised across plausible ranges
- No single one has to be right
- Weeks, and it reuses the 3D asset model we already hold

Why it works. A model trained on one perfect simulation learns that simulation. A model trained on a thousand imperfect ones cannot rely on any single quirk, so it is forced to learn the thing they have in common — which is the physics. Variation is doing the work that fidelity was supposed to do.

Commercially this is the difference between a research programme and a product. It is also why the asset model matters more than the physics engine.

SOURCE 5 OF 5 · ESTIMATED CONTRIBUTION 20%, RISING

Volume today: zero. That is the argument, not a hole in it.

```
1 technician  x  1 shift        ~   10 guided interactions
200 technicians                 ~  2,000 records / day
                                ~  700,000 / year

# on assets we hold the model for, in a format we designed,
# with the outcome attached to the prediction that preceded it
```

Every other source is static
The corpus will always be 21,386. The Situation Room grows slowly and by hand.

This one compounds
New, labelled, outcome-linked examples every shift, on exactly the assets the customer cares about.

Which is why sequencing matters: Field IQ has to be deployed and useful before the engine is any good. That happens to be the position EON is already in — and no competitor is.

What the list was missing

Eight more. Four of them are free.

Eight inputs that belong in the engine

1

Inspection & integrity data

ILI, UT thickness, hydrotest, CP survey. The highest-resolution degradation measurement that exists — and 63.4% of failures came within 2 years of one

LARGEST OMISSION

2

One-call / excavation tickets

Public, geolocated, timestamped, issued days before the dig. A leading indicator for 17.6% of all incidents

FREE

3

Work orders (SAP PM / Maximo)

The existing, degraded proxy for the intervention variable. Records the job, not the judgement

TIER 1

4

Alarm & event journals

Alarm flooding before an upset is a documented precursor. Compact, structured, rarely sensitive

CHEAPEST WIN

Continued on the next slide. Five sources were on the list; eight more belong.

…and four that cost effectively nothing

5

Published degradation physics

Paris-law crack growth, corrosion rates, bearing L10. Priors that let a model extrapolate from a 30-day window to a 90-day horizon rather than just interpolate

FREE

6

Weather & geohazard feeds

Natural force damage is 6.0% of incidents, and weather also drives corrosion, ground movement and freeze failures

FREE

7

The asset model itself

Not fuel — substrate. It makes simulation possible and lets a warning be shown in space rather than prose. And it is what makes a class of hallucination structurally impossible

HELD

8

Near-miss reporting

Aviation has ASRS — the richest safety dataset in any industry. Industrial operations has no equivalent. Field IQ is an instrumented one

DOES NOT EXIST

Item 5 answers the question raised earlier: the physics is what makes the June reconstruction of a March defect legitimate as an answer key while remaining inadmissible as an input.

The competition

Four layers. Each with a measured ceiling.

THE SHAPE OF THE MARKET

$22bn, and nobody has won it

$22.3 B

asset performance management, 2026

Across all industries

~22%

is oil & gas

About $5 billion, which is the number we have been using internally

28%

held by the top five vendors combined

AspenTech/Emerson, AVEVA, GE Vernova, IBM, Siemens

~7%

held by the leader

After two decades. Nobody is close to winning

A market that large, that fragmented, after twenty years, **is a market where the product has not been found yet.** Two thirds of industrial AI never leaves pilot — and the false-alarm arithmetic on slide 7 is a large part of why.

WHO IS OUT THERE

Four layers, four ceilings

Layer	Who	What they have	The measured ceiling
Condition monitoring	AspenTech/Emerson, GE Vernova, AVEVA, Siemens, Augury, Uptake	Anomaly detection on rotating & pressure gear	Sees the instrumented population only. Corrosion 13%, excavation 0%
Inspection & integrity	Baker Hughes, ROSEN, NDT Global, T.D. Williamson	The best degradation measurement that exists	Reports state, not trajectory. 63.4% failed within 2 years of it
Connected worker	Parsable, Innovapptive, Honeywell Forge, PTC, Librestream	The technician, and 6.4M+ assets	Digitises the form. Captures completion, not judgement
Industrial LLM	SLB Lumi, Baker Hughes Leucipa, Cognite, generic ChatGPT/Claude	The language of the domain	No plant state. Hallucination in safety-critical maintenance is unsolved

They are all good at what they do. The point is not that they are bad — it is that each owns exactly one of the four observations, and the answer needs all four.

THE CLOSEST NEIGHBOUR, AND THE SHARPEST DISTINCTION

A digitised checklist versus a captured judgement

WHAT A CONNECTED-WORKER PLATFORM CAPTURES

“Step 4 marked complete at 09:41 by badge 2213. Photograph attached.”

Real, useful, and sold well. It is a record that the work was done.

WHAT NOBODY CAPTURES

“Looked at the coupling. Thought it sounded wrong. Checked it twice. Decided it holds until the turnaround. Did not write that down — there is no field for it.”

This is the part that predicts.

The distinction matters commercially as well as technically.

Parsable and Innovapptive are not competitors for the prediction engine — they are competitors for the tablet. If the operator standardises on a forms platform, the forms get digitised and the judgement is still lost, because the platform has no model to be curious on its behalf. A checklist cannot ask a follow-up question.

THE HONEST POSITIONING

We are not a better model. We are the join.

Condition monitoring	The signal	Context, procedure, the field, the human
Inspection / integrity	The state	The trajectory, and the date
Connected worker	The technician	A model, and any record of judgement
Industrial LLM	The language	Grounding in a specific plant
EON	**The join — and an instrument already in the technician's hand**	A published performance result

The defensibility is narrow and should be stated narrowly: the missing training data can only be produced by something already standing next to the work, already trusted, already visual. A sensor company has no person. A forms company has no model.

What we need from them

Item by item. And why each one is answerable.

THE ASK, PRECISELY

Four tiers, and only one is a real request

TIER 0

Nothing at all

An equipment class list. Yields generic risk ranking from the public prior. Demonstrable in the first meeting with nothing signed.

TIER 1

Read-only, low sensitivity

Asset register · P&ID or 3D model · historian, one unit, 500 tags, 24 months · work orders · inspection records · alarm journal

TIER 2

Sensitive — and it never leaves

Incident and near-miss reports, operator logs, root-cause files. The model travels to it. Tenant Delta trains inside their boundary and is their property.

TIER 3

Does not exist — we make it together

Field IQ session records. Not really an ask: there is nothing to hand over. It is produced by using the system.

What we do NOT ask for: production volumes, commercial terms, reserves, personnel. If a request cannot be traced to a term in the equation, it is not made.

THE ARGUMENT THAT MAKES TIER 1 ANSWERABLE

We are not asking them to expose anything new

Historian, one unit, 24 months	→	Their PdM vendor already has this exact feed
Inspection records — ILI, UT, thickness	→	Their inspection contractor already holds it
Work-order history	→	Their CMMS vendor already holds it
P&ID or 3D model	→	Already provided to EON, for Field IQ
Asset register	→	Routinely shared with every maintenance supplier
Alarm & event journal	→	Rarely considered sensitive once explained

Every item is already in a supplier's hands. We are asking to be the first party allowed to put the six of them side by side. The novelty is the join, not the disclosure.

Work intelligence

The layer nobody captures. This is the important part.

Why most anomalies “resolve themselves”

In any functioning plant, most anomalies never become incidents. This is usually read as evidence that anomalies are mostly benign. It is not. It is evidence that people are good at their jobs.

```
# what the historian shows

    vibration rises  ->  vibration back to normal  ->  nothing happened

# what actually occurred

    vibration rises  ->  A TECHNICIAN INTERVENED  ->  nothing happened
                         ^^^^^^^^^^^^^^^^^^^^^^
                         invisible in every dataset in the industry
```

Train on that history and the model learns, with perfect internal consistency, that rising vibration resolves on its own. It learns not to alarm.

And it validates beautifully, because in the historical record nothing did happen — precisely because somebody kept intervening.

AND THE CONSEQUENCE

Every PdM system carries a bias toward silence

Trained on a customer's own operating history, and the bias is invisible to standard validation because the validation set contains the same confound as the training set.

This is not a criticism of any particular vendor. It is a structural property of the data everybody in this market is using — and it may be a large part of why the field's real results have persistently disappointed against its published metrics.

To record an intervention you have to be there when it happens.

Not afterwards, when a work order is closed with three words. At the moment, at the asset, while the person is deciding.

Field IQ is there — and not as a data-collection exercise, which nobody would tolerate. It is there because it is guiding the work. The technician is looking at it. It is showing them where to stand and what to check. Capturing what happened next is free.

BE CONCRETE ABOUT IT

What work intelligence actually is

Element	Example	Why it predicts
What was surfaced	Model flagged the outboard bearing at 62%	The prediction, stamped before the outcome — its own report card
What was checked	Tech inspected the coupling and seal, not the bearing	Where a competent human's attention goes is itself a signal
What was found	Coupling guard loose; slight discolouration on the seal	Observations no sensor makes and no form asks for
What was done	Retorqued the guard. Left the seal	The intervention. The missing variable
What was deferred	Seal left — holds until the September turnaround	An explicit human risk judgement, with a horizon attached
Time on task	Nine minutes on a job estimated at three	Duration tracks ambiguity, and ambiguity precedes surprise
The unwritten rule	“Always check that coupling when discharge sounds rough”	Tacit expertise in no SOP anywhere — it retires with the person

The last row is the commercial one. Every operator is losing its most experienced people and has no way to keep what they knew. That mechanism and the engine's fuel are the same substance.

AND THIS IS THE FLYWHEEL

The same instrument warns, and then watches

1

The model warns

Outboard bearing, 62%, on this pump, this shift

→

2

Field IQ shows it

Highlighted in 3D, on the asset, where the tech is standing

→

3

The tech acts

Checks, finds, does, defers — all captured as it happens

→

4

The outcome lands

Right or wrong, on the same device, against the same asset

→

5

The model learns

A graded prediction. Every shift. On assets the customer cares about

↵ *and back to 1, with a better model*

Every other source is consumed by use. This one is created by it. Which is why its contribution rises while everything else stays flat — and why Field IQ has to be deployed and valuable before the engine is any good. That is the position EON already occupies and no competitor does.

READ THIS AS PART OF THE ARGUMENT, NOT A DISCLAIMER

What would show that we are wrong

Risk	Why it is credible	The test
Selection bias in the corpus	Reports exist because assets failed. The cutoff governs values inside a document, not whether it exists	Train a classifier on metadata only. If it separates classes, we are confounded
The contribution % are wrong	They are judgement, not ablation	Ablate each source and measure
Simulation does not transfer	Isaac evidence is robotics, not degradation	Build one asset; evaluate against held-out real failures
Field IQ capture does not happen	Techs may not record under time pressure	Measure capture rate. Under 40% and the flywheel does not spin
The historian is unusable	Tag naming, gaps, resolution, drift	Ask for one unit first. Find out cheaply
The false-alarm problem survives	It has defeated everyone so far	That is what the Shift Test measures. It is unrun

There is no accuracy figure anywhere in this deck. Not withheld — it does not exist. The Shift Test is designed and unrun. Nothing here is evidence the engine works.

IN ONE PARAGRAPH

The stack works. It breaks anyway.

SCADA is running and finds a quarter of incidents. The inspection tool ran, usually within two years. The dig ticket was raised in nearly two thirds of strikes. Every layer fired.

What is missing is not a sensor. It is the join — and one observation no system records: what a competent human perceived, judged and chose to do, at a specific asset, at a specific moment, and why.

Four of our five sources supply labels and priors — necessary, obtainable by anyone, not predictive alone. One supplies the inputs and belongs to the customer, which is why the Tier 1 ask is as small as we could make it. One supplies the intervention record, exists nowhere, and can only be produced by an instrument already in the technician's hand for another reason.

Not that our models are better. We have not measured that and should not claim it.

EON REALITY · PREPARED FOR [CUSTOMER]

Nothing happened.

Which is the whole point, and the hardest thing to sell.

A model that reads your plant and tells your crew what to check — before it matters.

This deck describes a system that is designed and partly built. No performance result exists yet, and every place one would go is left visibly empty.

[DATE] · [PRESENTER]

DRAFT · NOT FOR EXTERNAL USE · NO

WHERE THIS STARTS

Twenty million characters of real failure, already read

Before asking you for anything, here is what has already been pulled apart and counted.

21,386	**real industrial incidents**	Every reported US pipeline incident since 1986 — hazardous liquid, gas transmission, gathering, distribution and LNG.
19,057	**first-hand accounts**	Written by the operator who was there, on a regulatory form, in their own words.
20.7 M	**characters read and counted**	About 83 full-length books. Every one machine-read and cause-coded — not skimmed.
40	**years of record**	1986 to today. Long enough that the reporting rules changed twice underneath it.

Every figure here is reproducible from the same public files, by you, in an afternoon.

AND WHAT WE FOUND IN IT

One column is filled in on every report.

The other does not exist.

WHAT BROKE

21,386 of 21,386

Cause, subcause, equipment, commodity, installation year. Complete, on every single report, for forty years.

COULD IT HAVE BEEN SEEN COMING

0 of 21,386

No regulator anywhere in the world records this. Not PHMSA, not CSB, not HSE, not the EU. The question is never asked.

Everybody records what broke. Nobody records whether anyone could have known. That second number is what this is built to produce.

We are not proposing to do better at something already measured. We are proposing to measure something nobody measures.

A TUESDAY AT ONE OF YOUR UNITS — TODAY

The data was available

06:00	Handover is five minutes and spoken. The oncoming crew gets whatever the night shift happens to remember.
06:15	A work order list. Twelve jobs. Ordered by how old the ticket is.
09:40	Routine check on a pump. Done by the book. Everything within limits. Signed off.
	Vibration has been climbing for nine days. That is on a screen in a different room.
	The lube sample is three weeks overdue. That is in a different system again.
+11 days	Bearing seizure. Unplanned shutdown. And it could have been a person.

"The data was available" is the sentence that appears in almost every investigation report ever written. It is not a criticism of anybody. It is a description of how plants work.

DRAFT · NOT FOR EXTERNAL USE · NO RESULT EXISTS YET

What changes

05:40	Overnight the model reads your plant's own memory — sensors, inspections, work orders, permits, last night's log. 312 assets. NOTHING LEAVES THE SITE.
06:00	The 3D unit is on the wall. Two things are glowing amber. The crew can SEE which pump. No tag number, no drawing.
06:02	Touch it. The bearing housing lifts out and turns. The nine-day trend draws itself. Underneath: likely mechanism, stated confidence, and a real documented case that looked like this.
06:05	ONE thing to do. Take a lube sample — and here is the drain valve, highlighted on the pump, before anyone walks out.
09:40	The sample is cloudy. The photograph attaches to the pump, not to an email.
Saturday	Water confirmed. Bearing changed in planned downtime. No lost production.

Nothing happened. Which is the entire product.

Five things changed

Today	With EON Delta

None of these are about the AI being clever. They are about the right thing being in front of the right person while they can still act on it.

THE SECOND FLAG THAT MORNING

Not everything that goes wrong is a machine

The second amber item was not equipment at all. Nothing was wrong with any metal.

An excavation permit open on Line 22. Work scheduled this morning. Locate verification not recorded for six days.

Across all US pipeline incidents, roughly 39% have nothing wrong with the equipment. Somebody skipped a step.

An anomaly detector will never find this — because there is no anomaly. There is an absence.

Why this matters commercially

Condition monitoring has been sold into plants like yours since the 1990s, and it is good at what it does.

It watches sensors for something unusual.

A permit that was never verified produces no unusual sensor reading. It produces nothing at all — which is exactly why it is missed.

Source for 39%: PHMSA flagged incident files, all systems, 1986 to present. Reproducible from the public download.

THE FAIR QUESTION

You already have monitoring. And you have ChatGPT.

Your condition monitoring

Watches sensors for anomalies on the assets it is connected to. Mature, and good at it.

Cannot see a missing permit verification. Cannot cite a documented case. Cannot tell you what to check — only that something moved.

A general AI model

Read the whole internet, including a great deal about industry.

Never studied the incident record as a discipline. Cannot point at anything in your plant. And you would be sending it your data.

EON Delta

Studied the industrial failure record; reads your plant's own data; answers inside the 3D world your crews already use.

Does not exist as a measured product yet. That is stated plainly on every slide of this deck.

These are not competitors to be replaced. Delta reads what your monitoring already collects — it is a layer above it, not a rip-and-replace.

THE OFFER

We are not asking for your data.

We are offering to leave it where it is.

Picture a locked room inside your building.

1	**Your data stays in the room**	It is never copied, never uploaded, never pooled.
2	**We send in a student**	It reads everything, learns, and stays in the room permanently.
3	**The student is yours**	Delete it whenever you like — and we can prove it is gone, not assert it.
4	**Only the general lesson leaves**	And only after the same lesson has been seen at several unrelated companies, so it points at nobody.

This is not a policy we wrote. It is enforced by the system: training refuses to start in any environment that does not carry your identity.

CONCRETELY

What leaves the room, and what never does

“A pump starved of oil runs hot before it seizes.” ← leaves

“The pump at your Baytown site ran hot last March.” ← never

A dataset is a pipe — whatever goes in comes out.\nA model is a teacher — it carries the lesson without carrying the source.

Why this is the shape that works

You and your competitors cannot pool operating data. That is not caution — in several jurisdictions it is the thing competition law exists to prevent.

Layer	Trained on	Lives	Yours or ours
Your Delta	Only your own material	Inside your environment. Training refuses to run anywhere else.	Yours. Deletable on demand, and the deletion is provable.
Base Delta	Only patterns independently confirmed at several unrelated organisations	Shared	Ours — and licensed back to you free, in perpetuity, including if you leave.

Nothing attributable to you ever reaches the shared layer. A pattern only qualifies once it has been seen at several organisations that have no relationship with each other — at which point it belongs to nobody.

We would expect your counsel to test this properly, and we would rather they did it early than late.

WHAT YOU GET

Four things

1

Your own model

Trained on your plants, your equipment, your history. Runs in your building. Never leaves it.

2

Everyone else's lessons, free

The shared model — everything learned at every other participating operator — licensed back at no cost, in perpetuity, even if you stop participating.

3

A structure counsel can approve

You and your competitors can both take part without either seeing the other's data.

4

It in the plant, not in a report

Delivered inside the 3D world your crews already use. Not another dashboard. Not another inbox.

And every one of these gets better the longer you run it, because the model keeps learning from what it gets right and wrong in your plant.

DRAFT · NOT FOR EXTERNAL USE · NO RESULT EXISTS YET

THE PART MOST VENDORS LEAVE OUT

What we cannot tell you today

There is no accuracy figure in this deck. Not because we are withholding it — because it does not exist.

We can say this today	We cannot say this yet

Every place a number would go in this deck says PENDING. It will keep saying PENDING until a benchmark we published in advance has actually been run — including the parts we lose.

If a vendor shows you a number for this and cannot show you the exam that produced it, ask them the same question you should ask us.

We test AI the way airlines test pilots

Real physics. New situations. No answer sheet to memorise.

Airlines do not certify pilots with a written exam about past crashes. They put them in a simulator and give them a situation they have never seen. Nobody has ever called that unfair.

EON has built simulators of industrial work for twenty-five years. Testing AI inside one is the obvious use of what we already have.

Scored on	Result
Correctly naming what is developing	[PENDING]
Citing a real, checkable comparable case	[PENDING]
Naming a check that would settle it	[PENDING]
Honest confidence — and knowing when to stay quiet	[PENDING]
False alarms on a separate quiet track	[PENDING]

Published openly

The method, the scoring, every result — including the ones we lose.

Your own technical team can run it. That turns a twelve-month evaluation into an afternoon.

A benchmark we always win is marketing, and everyone can smell it.

The exam is being built BEFORE the model is finished — deliberately. Build the model first and you will design an exam it passes without noticing.

WHAT WE WOULD ASK OF YOU

Small, and in this order

1

An hour with two of your engineers

Costs you an hour.

Show them the Tuesday. Ask them where it is wrong. Anyone who has stood in a control room at handover will see something we have not.

2

One unit, read-only

Costs you access, not data.

A single unit's historical data, inside your environment, with nothing leaving. We measure what the model would have flagged against what actually happened.

3

A decision at the end of it

Costs you nothing to walk away.

Either the number is good enough to continue, or it is not. Both are useful, and we will publish either.

We are not asking for a commitment, a budget line, or a data transfer. We are asking for an hour and then a read-only look at one unit.

The bearing gets changed\non a Saturday.

Nobody writes a report. Nobody investigates anything. Nothing happened.

That is the product, and it is the hardest thing in the world to put a number on — which is exactly why we are building the exam before the model, and publishing it.

No performance figure appears anywhere in this deck. Nothing has been trained and nothing has been measured. Every number here is a count from a public file that you can reproduce yourself.

[PRESENTER] · [EMAIL] · [DATE]