The Regional Memory

What becomes possible when Alberta's environmental data goes digital

Mike Duvall
Developer, Phaseview
RemTech 2026

PHASEVIEW

Closing out Alberta's wells

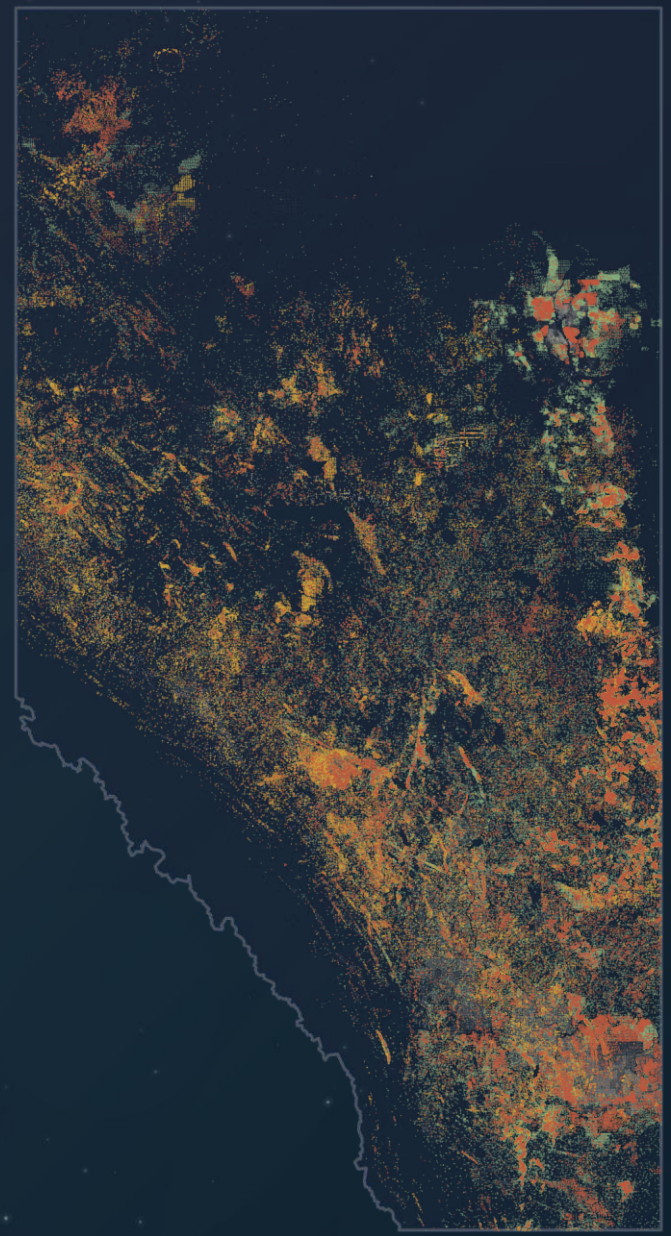
- **484,256** wells licensed in Alberta
- **113,216** reclaimed and certified
- **94,869** abandoned, waiting
- **64,561** suspended

~4,200 a year
reclamation certificates, 2021-25

At today's pace:

22 yrs
to clear abandoned

38 yrs
with suspended



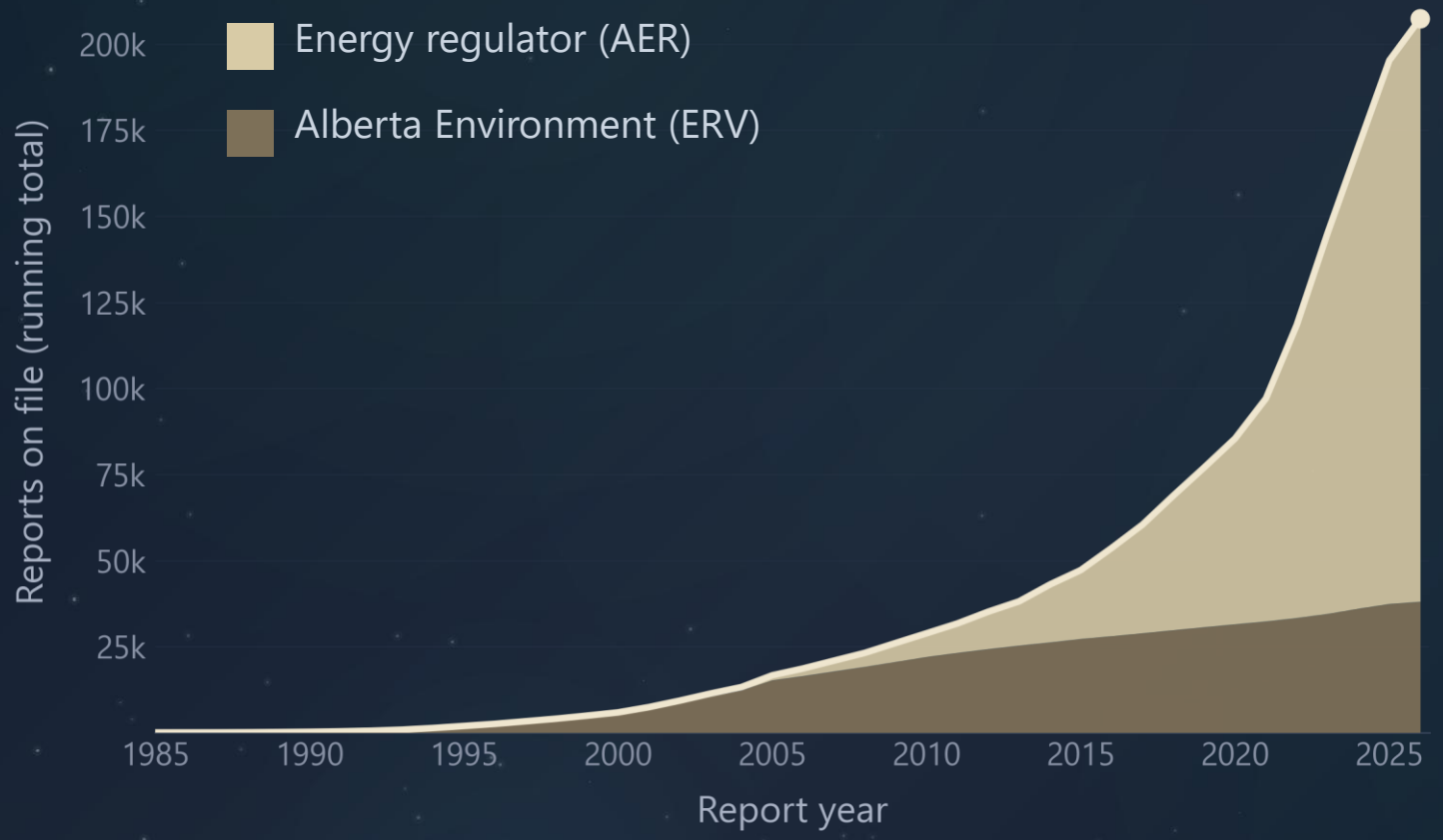
AER ST37 status, 2026-10-03, by licence. A snapshot of today, not a history.

PHASE VIEW

THE SECOND MOUNTAIN

Meanwhile, we built a second mountain

Site assessments, remediation, groundwater, reclamation and site condition records filed with the AER and Alberta Environment (ERV) - cumulative, by report year



~207,000 submissions

~9.4 million pages

Almost all of it filed as PDFs.

All of it hard to interact with.

PHASEVIEW

What does 9.4 million pages look like?

943 m of paper.

About 3 Tunnel Mountains, stacked.



Photo: Harvey Locke, CC BY-SA 3.0, via Wikimedia Commons

PHASE VIEW

What should we be doing with our historical records?

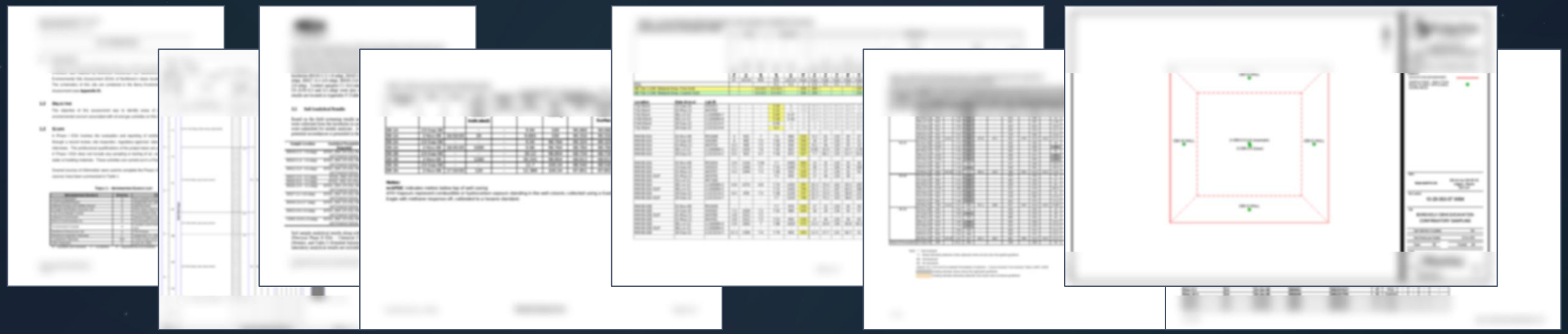
PHASEVIEW

The life of a data point

Field	Lab	Consultant	Report	Regulator
				
PAPER / TABLET	DIGITAL	DIGITAL	PDF	PDF

PHASEVIEW

One wellsite. Twenty-three years. Seven firms.



2002
Phase 1

2005
Phase 2

2007
Phase 2

2009
Groundwater

2014
Groundwater

2019
Groundwater

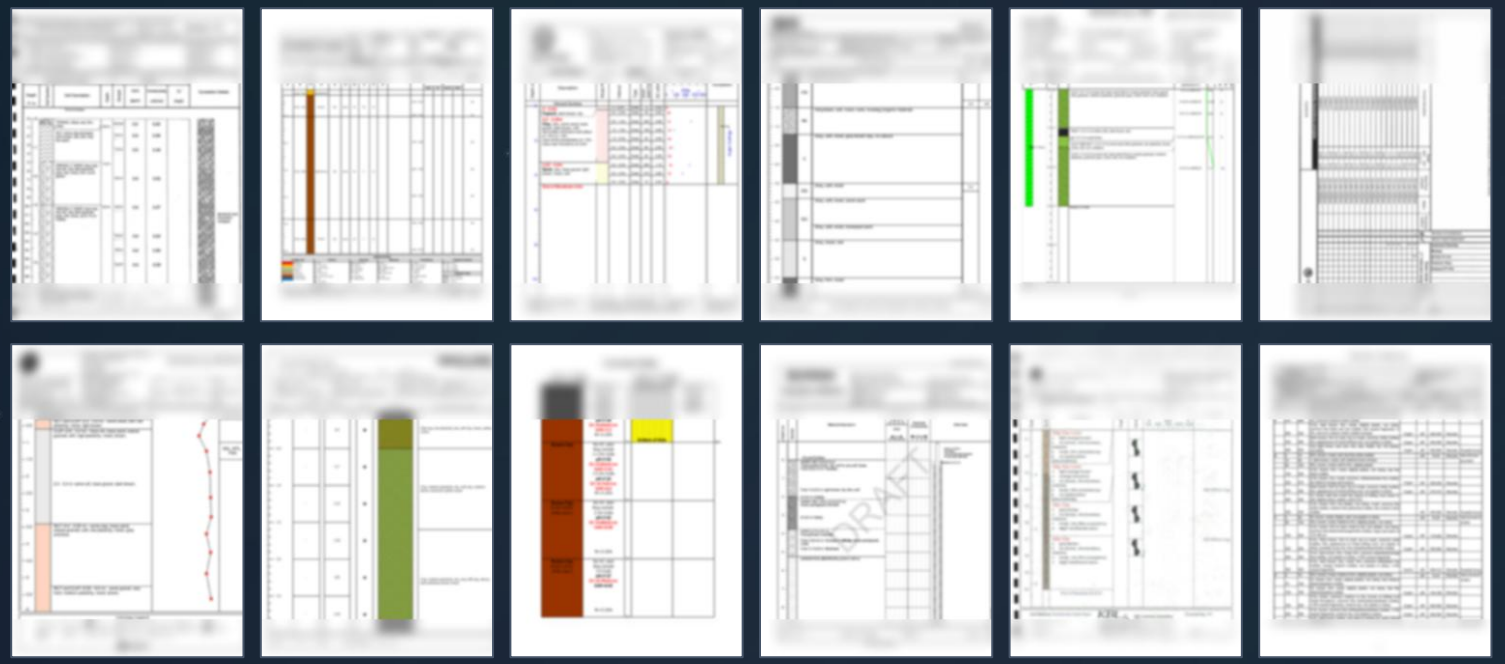
2023
Phase 2

2025
Remediation plan

2026
You

PHASEVIEW

Everyone does it differently



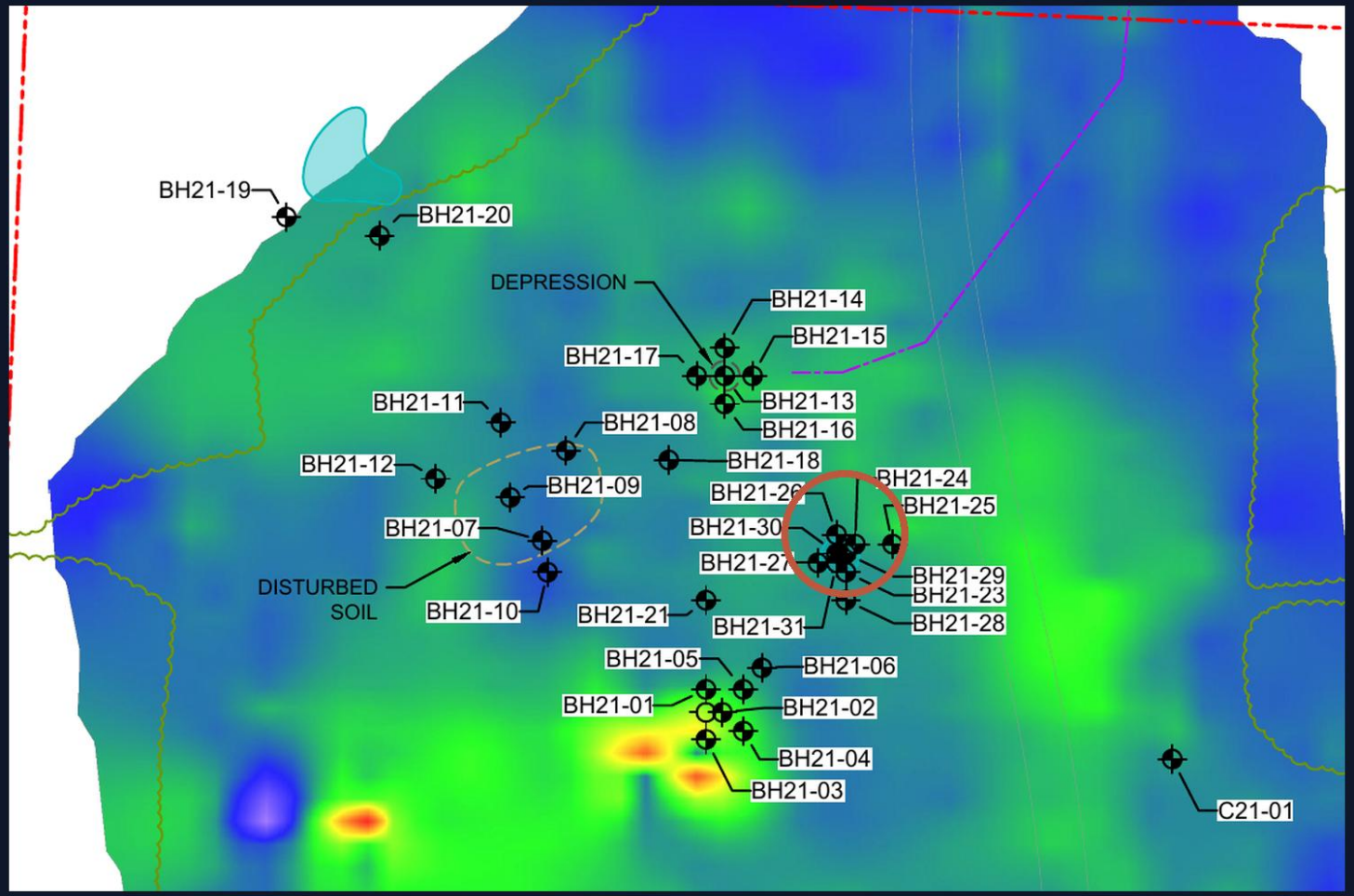
Twelve borehole logs, twelve layouts



Four labs, four certificates

Every difference is time someone spends translating.

What it takes to get one number back



1 Find it

2 Read it

3 Name it

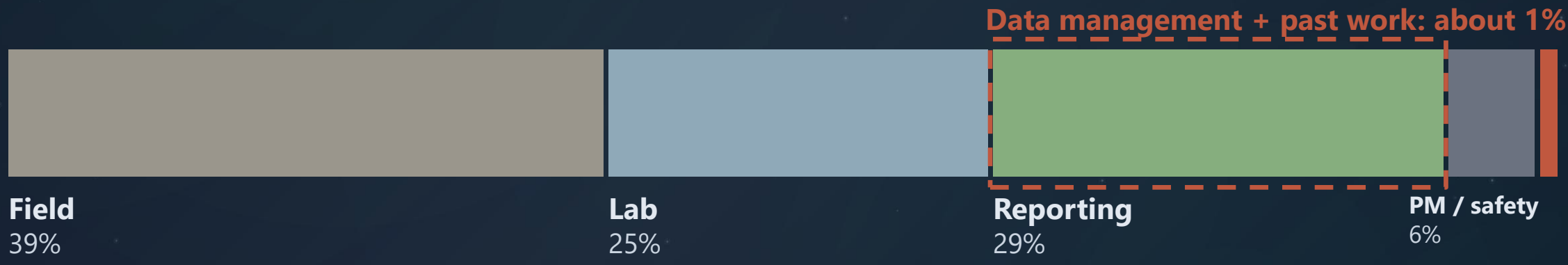
4 Unit it

5 Match it

6 Place it

PHASEVIEW

What does our amnesia cost us?



Re-typing old tables hides in reporting, and nobody prices it.

PHASEVIEW

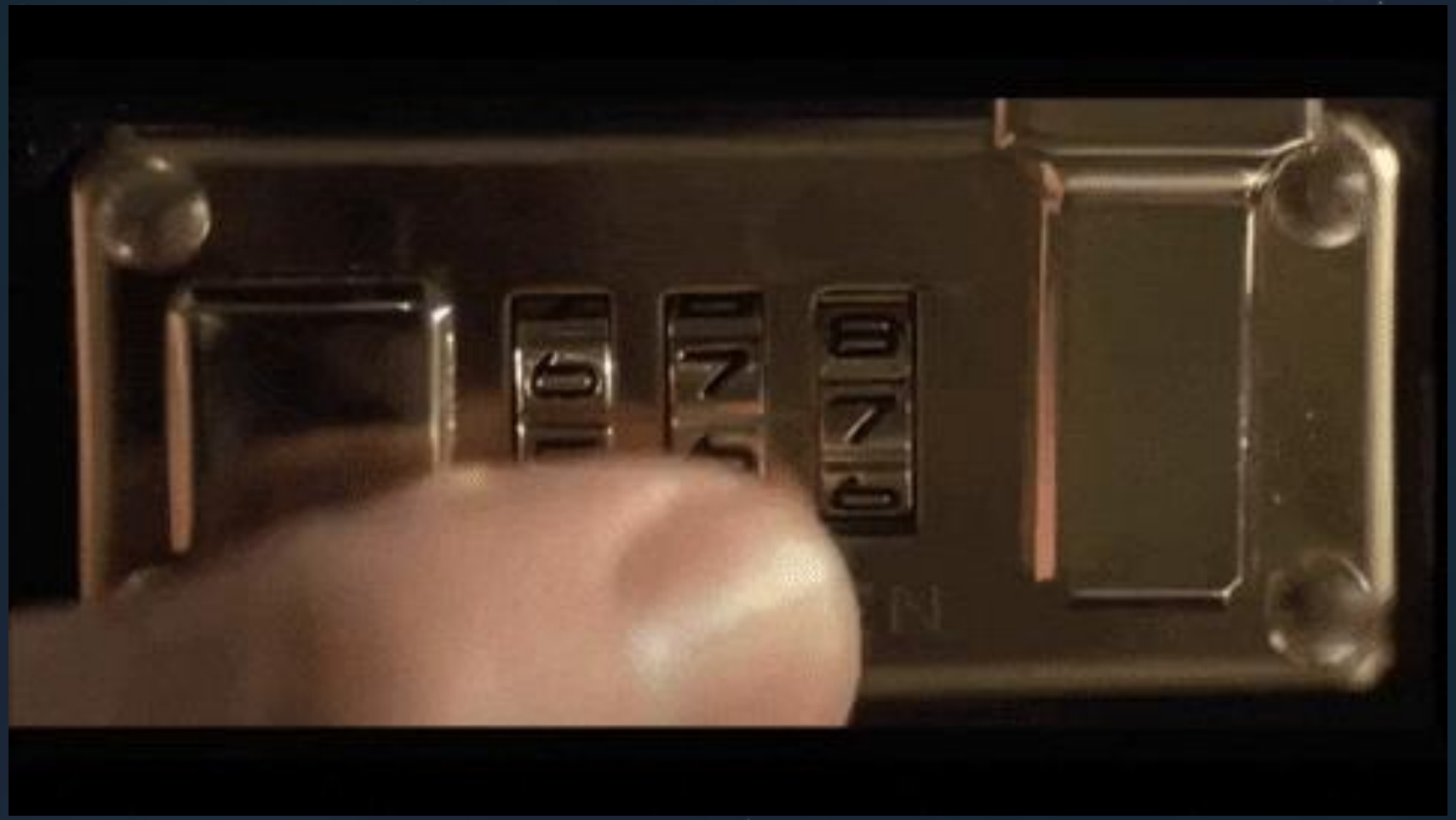
Why are we still making life hard on ourselves?

1. Why do we work digitally, then file PDFs?
2. Why is there no standard?
3. Why do we keep re-typing the same tables?
4. Why is it so hard to see what's been done next door?
5. Why are our tools stuck in the early 2000s?



PHASEVIEW

The Holy Grail: Data Liquidity



Pulp Fiction (1994), via GIPHY

So I tried to build it



every dot is a report

~207,000 submissions

Two regulators

The most relevant reports

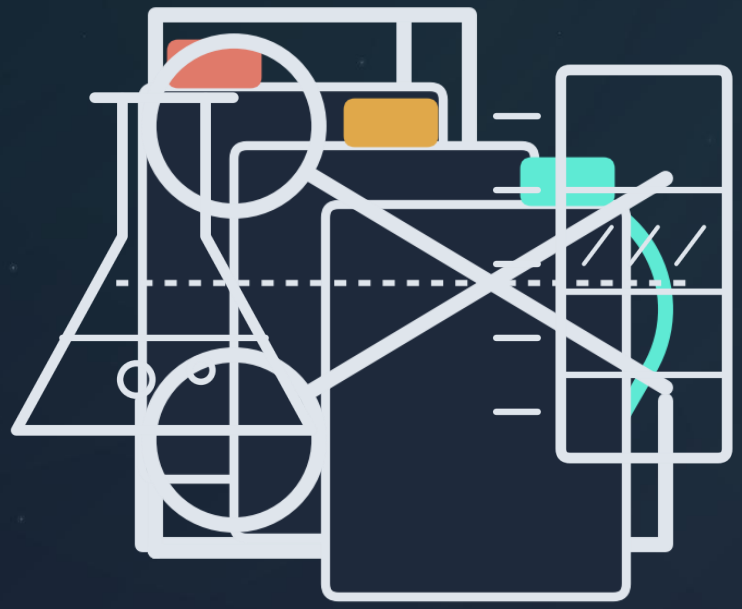
Millions of lab results

Hundreds of thousands of boreholes

And steadily growing...

PHASEVIEW

From public filing to usable data







PHASEVIEW

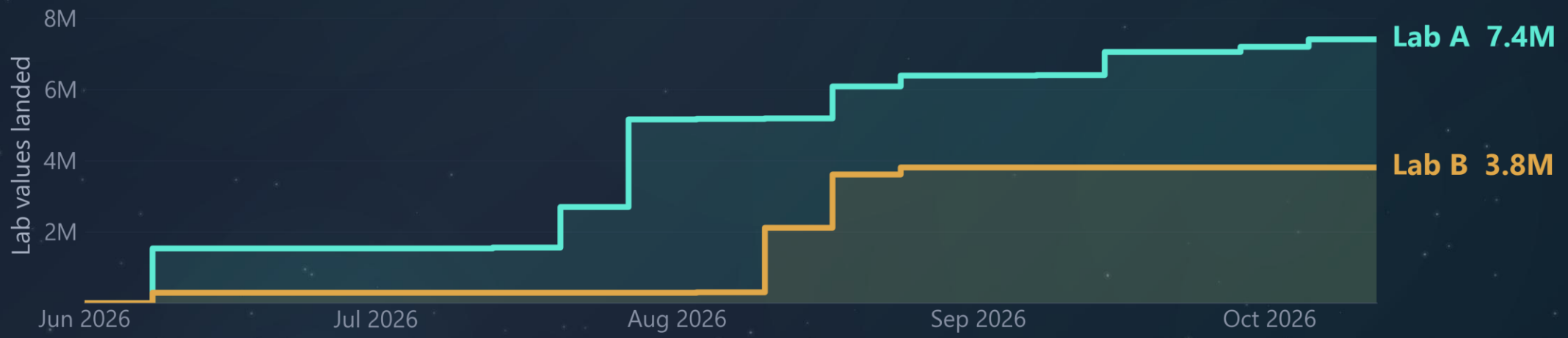
Nothing lands on a guess

READ > CHECK > SPOT CHECK









PHASEVIEW

WHAT THE MEMORY SAYS

What could a regional memory do for us?

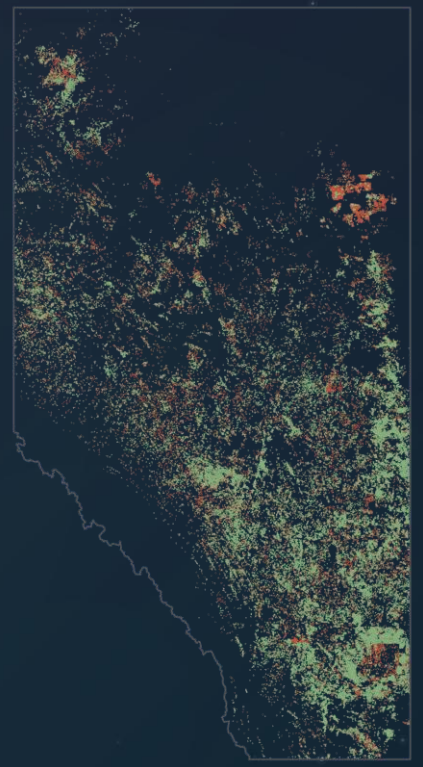
Data liquidity
Every old site in one standard digital format

Accuracy and accountability
No more re-typing; every value traces to its page

A head start on every site
70% of abandoned wells have a report within 1 km

Continuity
A site's history survives a change of firm

The regional picture
Patterns across thousands of sites, not one at a time



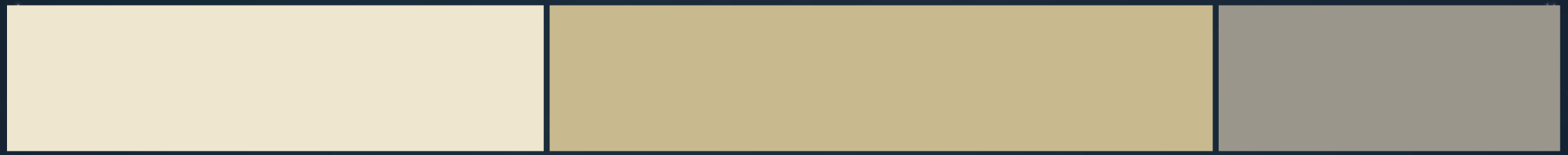
abandoned well, report within 1 km

no report within 1 km

PHASE VIEW

Who writes Alberta's environmental record?

203,112 reports with an identified consultant



Top 10 firms
35% of reports

Next 40
43% of reports

The rest
22% of reports

250+ firms

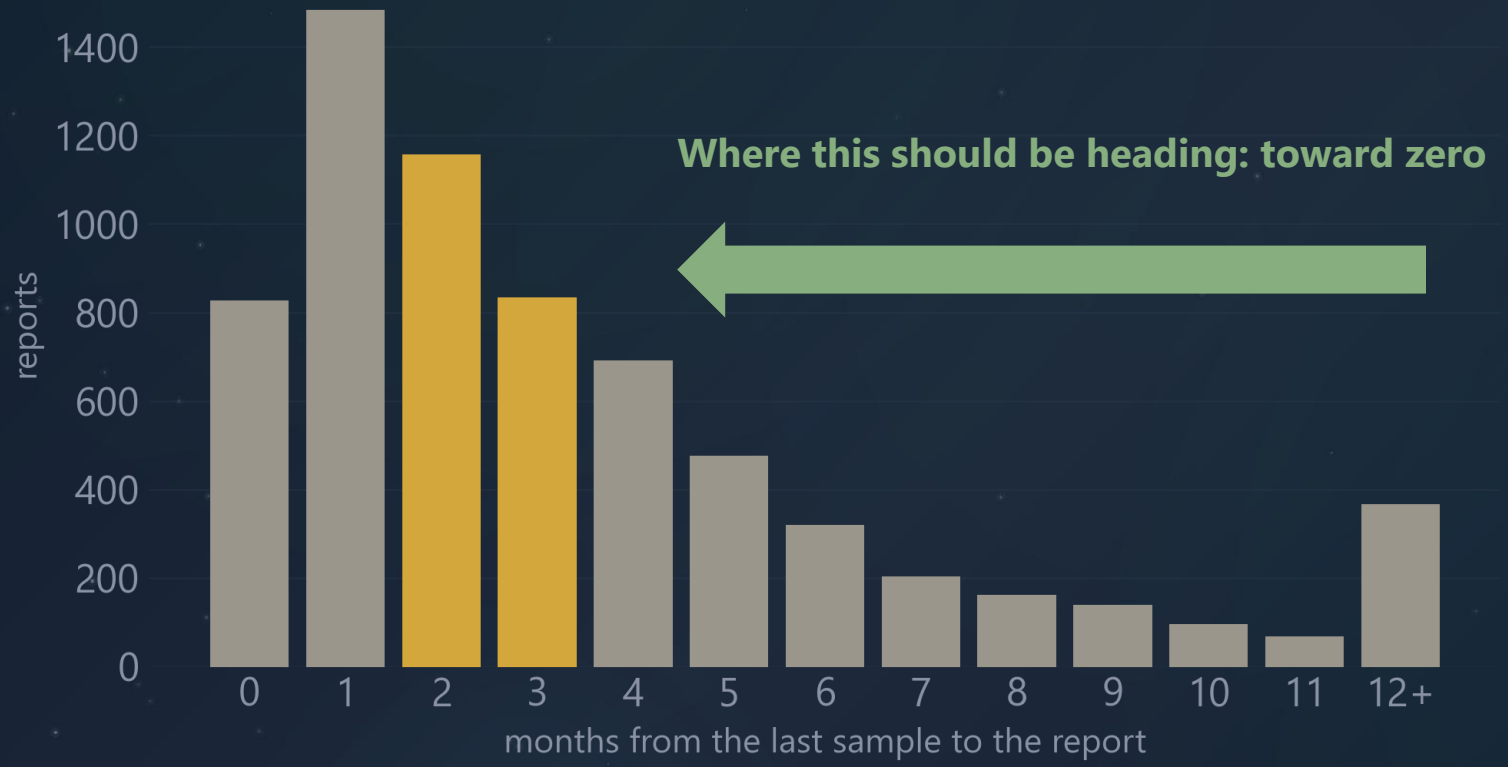
wrote reports in the last three years alone

Alberta's environmental record is written by a diverse group of companies. No single firm can fix this alone.

PHASEVIEW

How long does it take to turn a sample into a report?

6,874 reports: last sample date to report date



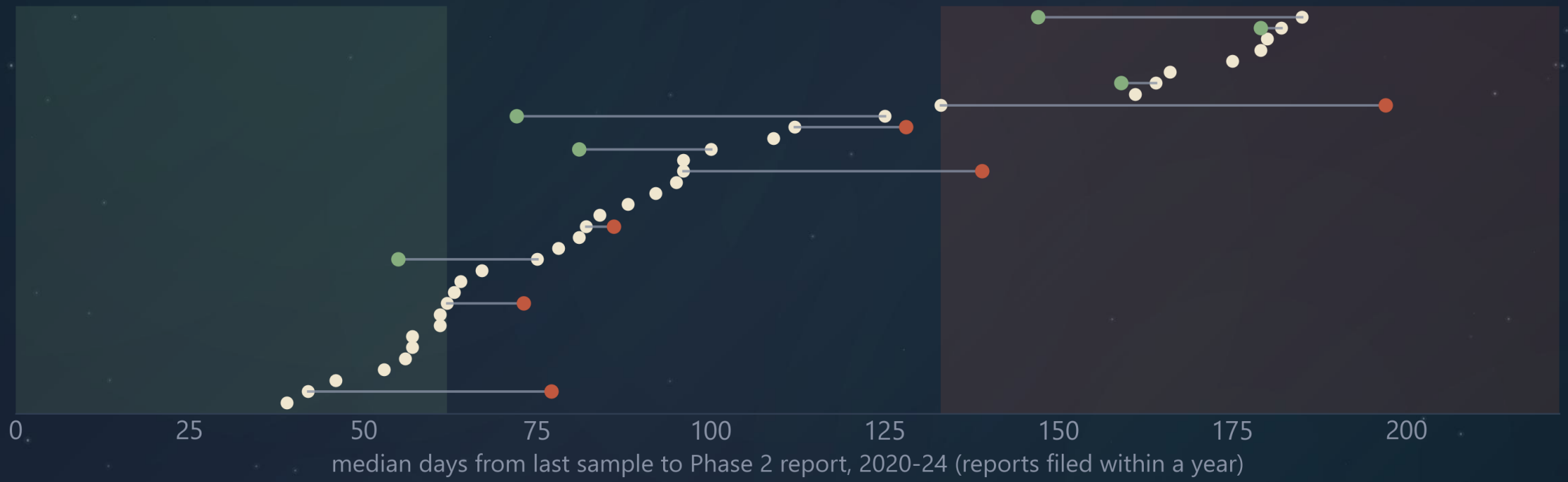
3 to 4 months
typical time from the last sample to the report

1 in 5
takes longer than six months

PHASEVIEW

Is everyone that slow?

36 firms with 10+ Phase 2s in 2020-24 (checked lab data); each dot is that firm's 2020-24 median. Names left off.



Are we getting faster? Not yet.

62 days or less
typical Phase 2 turnaround for the fastest quarter of firms

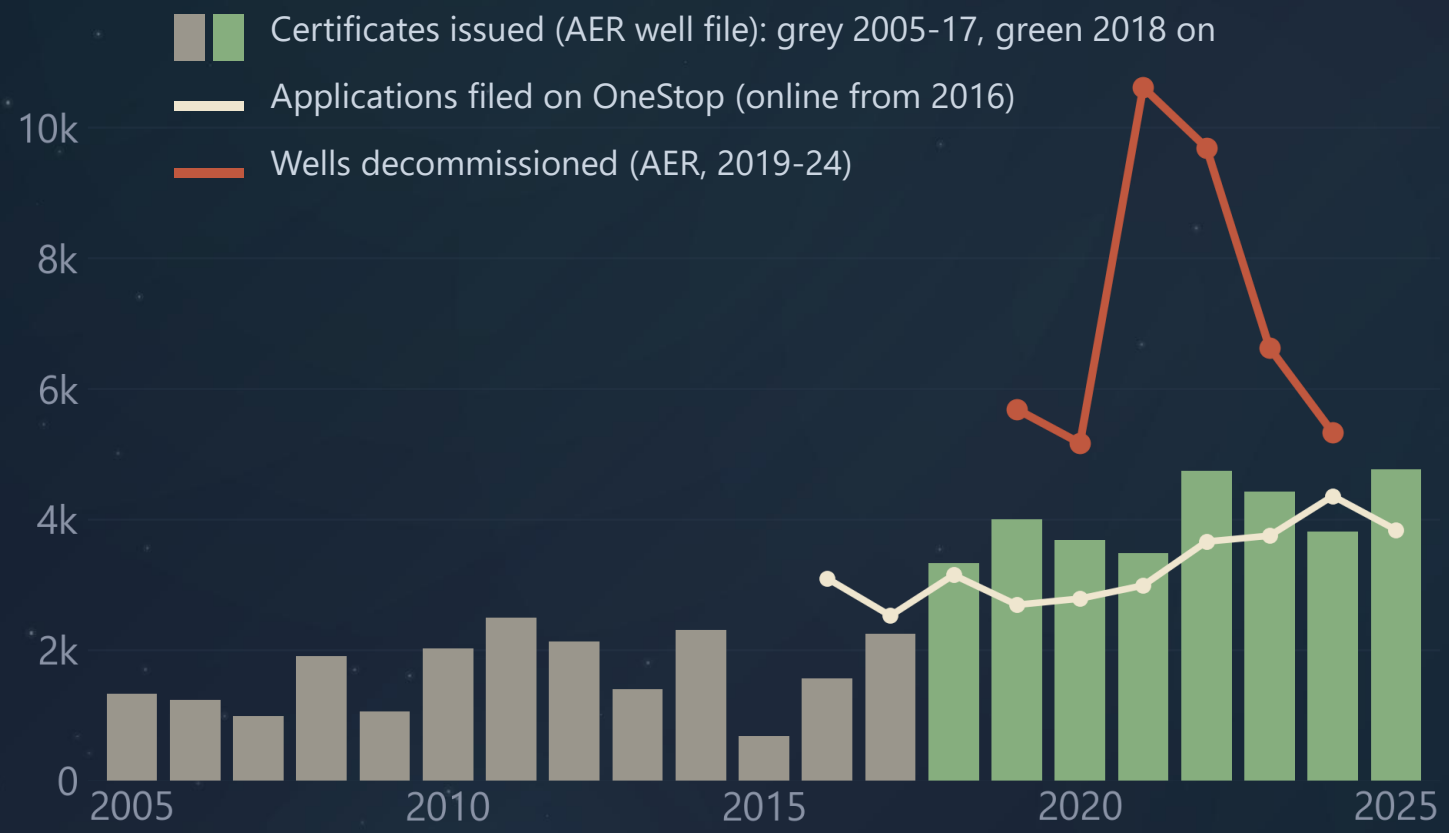
133 days or more
for the slowest quarter

In 2023, 6 of the 12 firms with 5+ Phase 2s were slower than their own 2020-24 median. The typical firm: +4 days.

PHASEVIEW

Are we closing more sites?

Reclamation certificates issued per year, by licence (clerical bulk dates excluded)



Yes. Closures have more than doubled since 2018.

About 1,900 certificates a year from 2010 to 2017; about 4,000 a year since.

OneStop applications (the line) agree.

But we abandon wells faster than we certify them.

About 2 wells decommissioned for every 1 certified, 2019-24.

PHASE VIEW

We've seen what the record says about us. So where does all this friction come from?

Mostly from us. Our hardest data problems are self-inflicted.

The PDF can no longer be the only deliverable.

THE FIX

One thing each of us can do

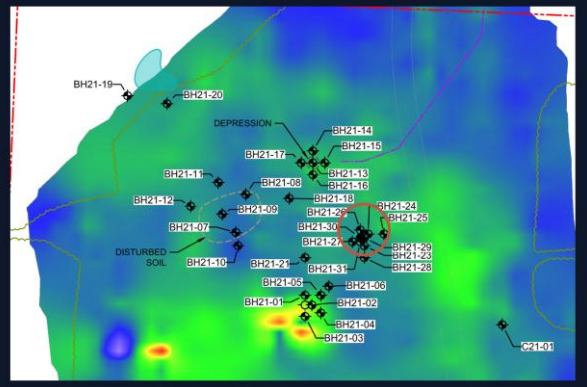
Labs	Back one shared electronic format
Consultants	Put a data deliverable in every scope
Operators	Write data into your contracts, and keep it with the site
Regulators	Accept data files alongside the PDF

PHASEVIEW

Our biggest data sins

1 in 5

reports with borehole logs



Lab certificate

Sample #	35658230713009 @ MW107-11	has a location
BUN817	35658230713001	just a code

Only the consultant's table says what it is

Sample Point	Sample Date	MSI Sample Number	Benzene mg/L
Field Blank	13-Jul-23	35658230713002	<0.0004
Trip Blank	13-Jul-23	35658230713001	<0.0004

No GPS coordinates

About 1 in 5 reports give no GPS coordinates on their borehole logs; many locate the hole only by legal land description. In 2023-2025 reports, about 1 in 12.

Only on a picture

Borehole locations drawn on a site plan image. To use them, someone has to georeference the plan and re-derive every point.

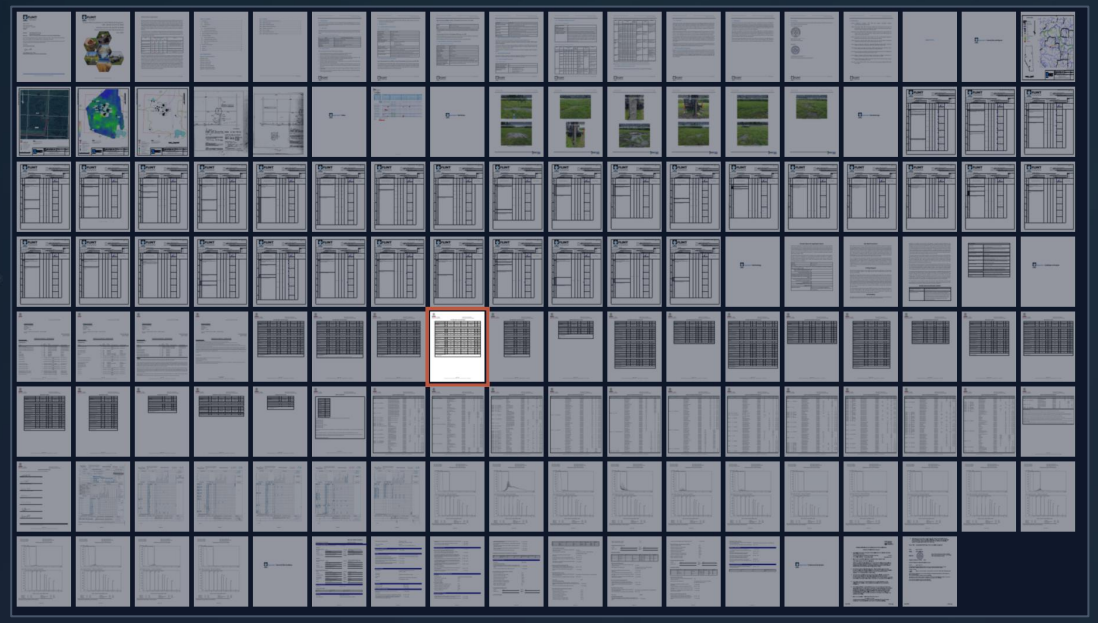
Untraceable sample IDs

On the lab certificate the sample is just a code. Only the consultant's own table says where it came from.

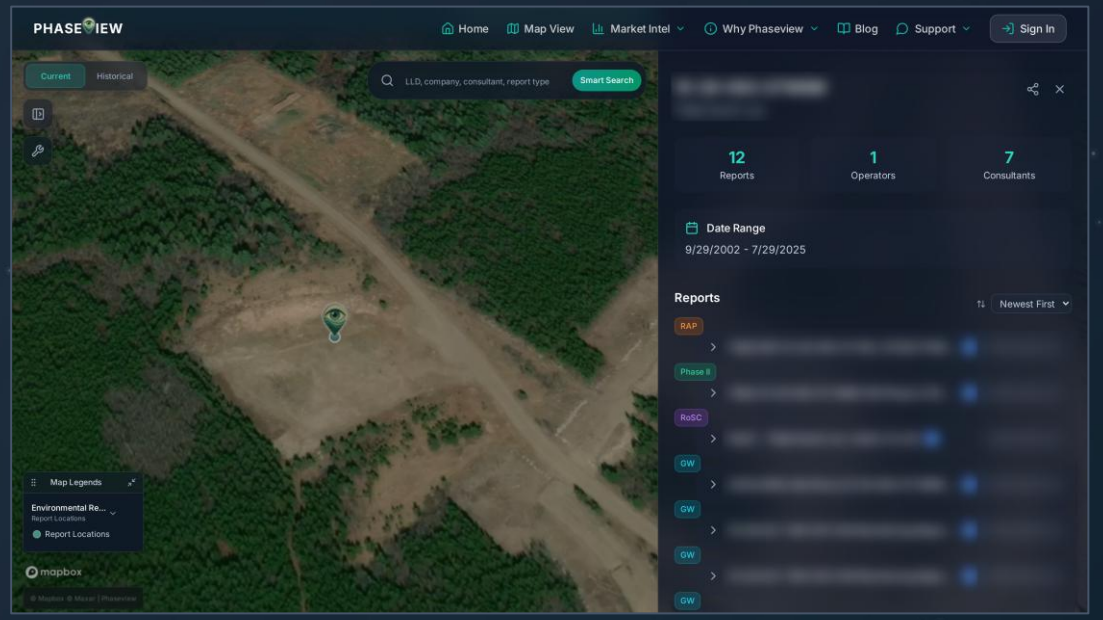
PHASE VIEW

2040. Someone inherits your site.

What do they open?



A box of PDFs



Or the whole history, in one place

We can be better. Phaseview is my humble attempt at that.

phaseview.ca

PHASEVIEW

“If I have seen further, it is by standing on the shoulders of giants.”

Isaac Newton, 1675

What we file today is someone else's starting point.

Thank you

Mike Duvall | Phaseview

phaseview.ca

PHASEVIEW

TWO MOUNTAINS, ONE PROBLEM

Built on old foundations

66%

of 31,992 Phase 2s and later assessments filed in 2021-25 landed where earlier reports were already on file

57%

of Phase 2s alone had earlier work at the same location

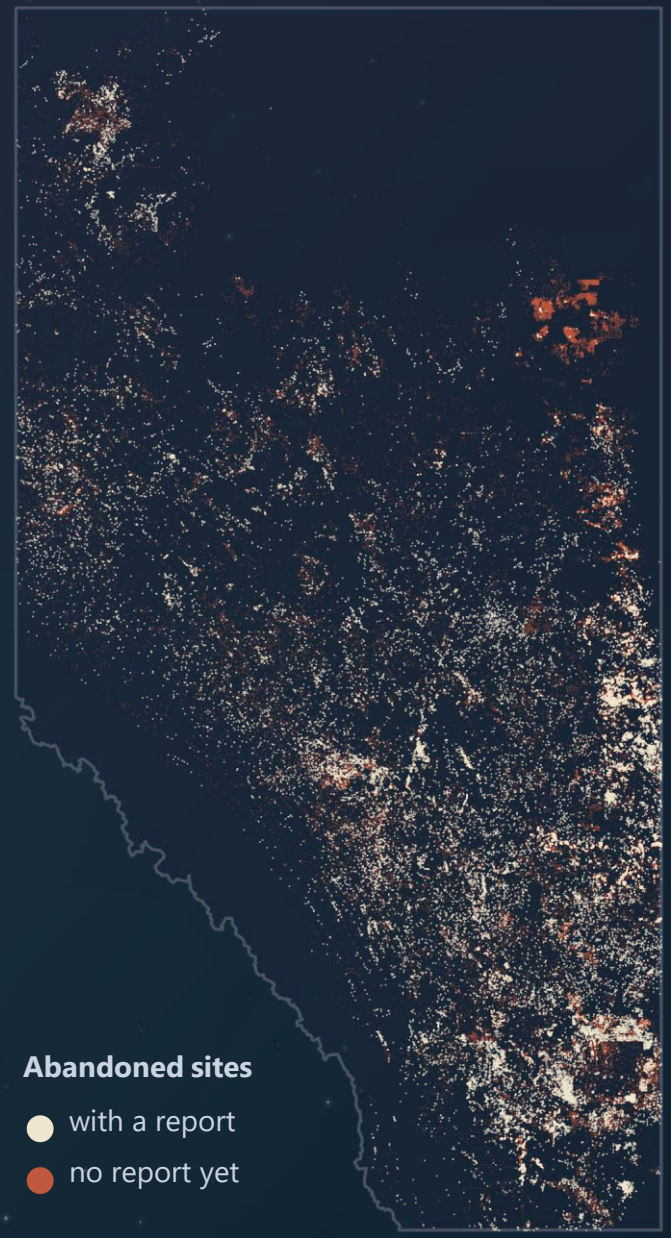
42%

followed work by a different firm, so they inherited someone else's data

Before anyone sets foot on site, part of the answer is already written.

It's just locked in a PDF.

Earlier work = a report at the same legal subdivision, 30+ days before. Phase 1s left out (usually the first look).
Floors: pre-2010 records are thin.



PHASEVIEW

What do we keep finding in the ground?

Share of soil samples analysed, checked lab results: 1.07 million values from 5,581 reports



Salt and heavy hydrocarbons turn up most.

Benzene is detected in about 1 sample in 6.

PHASEVIEW

What do we test for, and where?

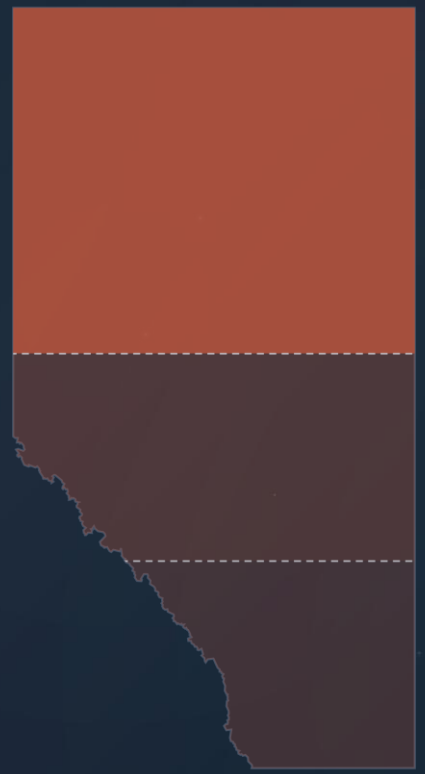
Left: 4,441 reports from 2020-24 with soil chemistry. Right: soil samples by region, split at 52 N and 55 N

88%
of reports test for salinity

87%
of reports test for BTEX

84%
of reports test for F1-F4 hydrocarbons

65%
of reports test for metals



North
F2 detected in 34%, salt above the mark in 48%

Central
F2 detected in 26%, salt above the mark in 32%

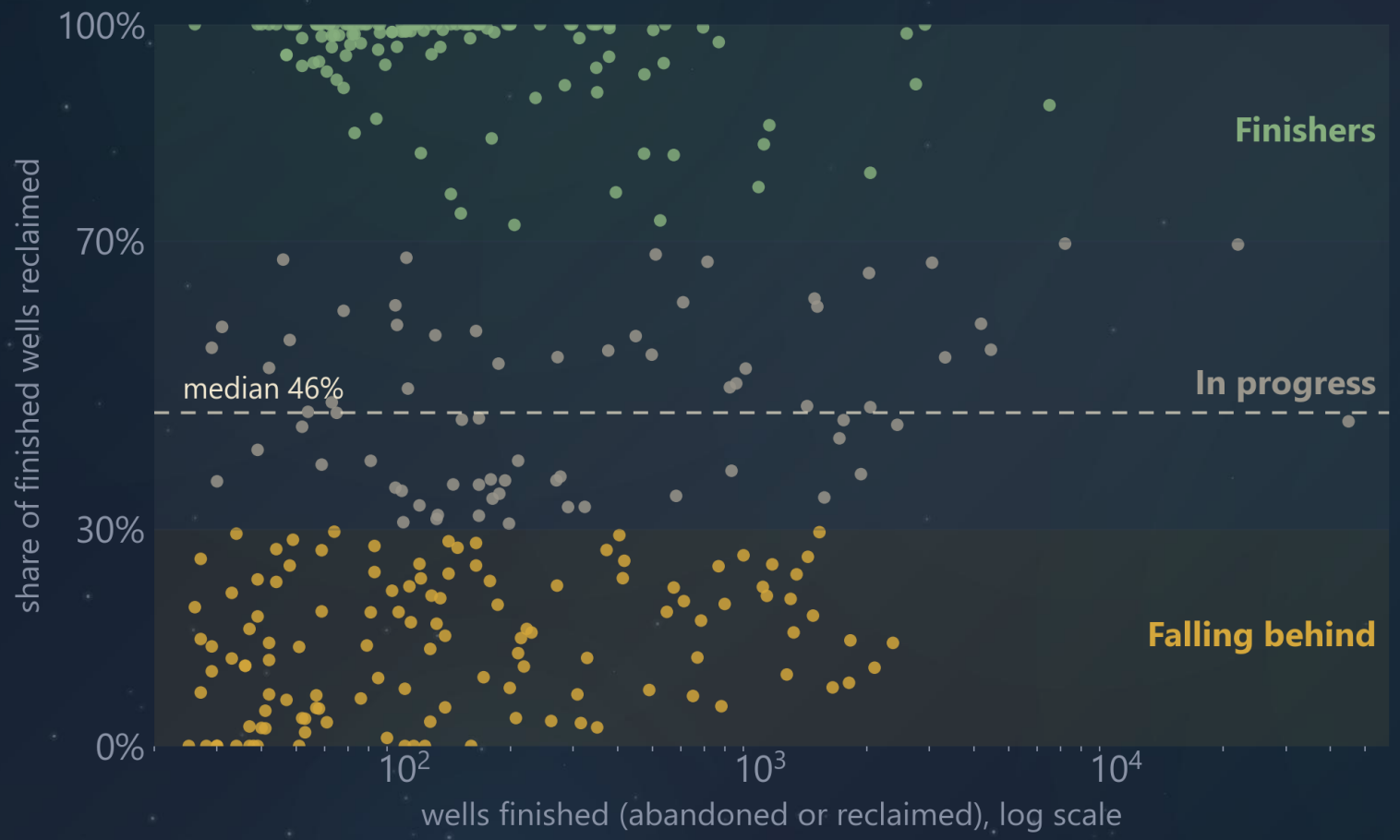
South
F2 detected in 24%, salt above the mark in 52%

We test for nearly everything, nearly everywhere. Hydrocarbons turn up most often in the north.

PHASE VIEW

Who's finishing their wells?

315 operators with 50+ wells: share of finished wells reclaimed vs wells finished, by current licensee



39% falling behind
have reclaimed under 30% of their finished wells

38% finishers
have reclaimed 70% or more

Same rules, same province: as many operators falling behind as finishing.

PHASEVIEW

Backup: likely questions

1. Labs already send electronic data.
2. Can't AI just read the PDFs?
3. Is this data public? Can you use it?
4. How accurate is the extraction?
5. Which format: ESdat, EQuIS?
6. Are you naming firms or operators?
7. Who pays for a standard?
8. Isn't the regulator already doing this?
9. Does this replace consultants?

Labs already send electronic data

Yes. The lab's electronic file reaches the consultant, and it stops there.

The regulator, the operator's next consultant and everyone after them only ever get the PDF.

So the fix is letting that file travel with the report.



The file stops at the consultant.

PHASEVIEW

Can't AI just read the PDFs?

Partly. We use AI, and it helps with reading.

But it isn't accurate enough on its own. We built our hand-labelled test sets with AI's help, and every one needed several rounds of people finding and fixing its mistakes.

If our work has to stand up in court, 'the AI read it' isn't good enough.

Sending the data with the report is cheaper, and it's right the first time.



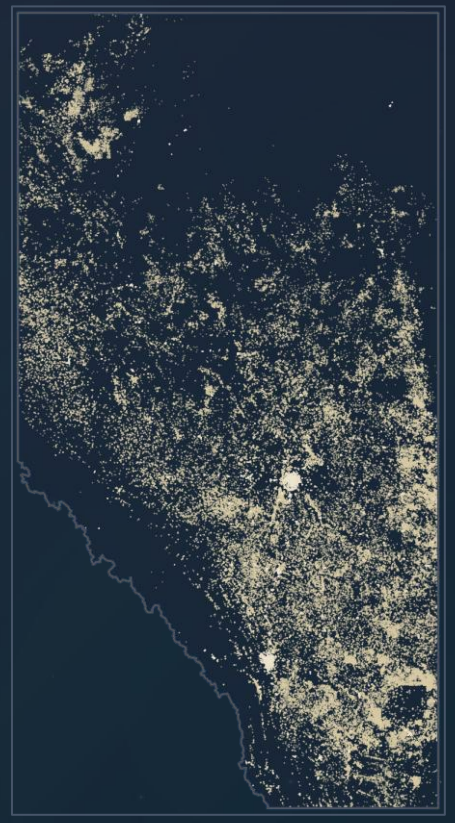
Real benzene units from our data, and what scanned pages turned them into.

Is this data public? Can you use it?

These are public regulatory submissions. Anyone can download them from OneStop and Alberta Environment.

We don't rehost or resell the source PDFs. Everything points back to the regulator's own copy.

Everything on Phaseview is free today.



Every dot is a public filing. Each one links back to the regulator's copy.

PHASEVIEW

How accurate is the extraction?

Two separate steps, built lab by lab: an extractor reads the certificate, then a validator checks it.

The validator re-checks every value against the page: right sample, right parameter, right units and detection limit, and nothing on the page missed.

If any value fails, the whole certificate is held back until it's right. Nothing lands on a guess.

Every value keeps a pointer back to its page and row, so anyone can check it.

100 of 100

randomly drawn values in the database matched their source page: right sample, right parameter, right value. Checked by eye, October 2026.

Lab ID		QX4101	QX4102	QX4103	QX4104		
Sampling Date		2020/09/14	2020/09/14	2020/09/15	2020/09/15		
COC Number		C-000123	C-000123	C-000124	C-000124		
	UNITS	BH-1 @ 1.0-1.5	BH-2 @ 0.5-1.0	BH-3 @ 1.5-2.0	BH-4 @ 0.5-1.0	RDL	QC Batch
Ext. Pet. Hydrocarbon							
F2 (C10-C16 Hydrocarbons)	mg/kg	<10	24	<10	130	10	Q100231
F3 (C16-C34 Hydrocarbons)	mg/kg	<50	210	<50	460	50	Q100231
F4 (C34-C50 Hydrocarbons)	mg/kg	<50	65	<50	150	50	Q100231
Physical Properties							
Moisture	%	12	18	15	21	0.30	Q100198
Volatiles							
Benzene	mg/kg	<0.0050	0.012	<0.0050	0.041	0.0050	Q100244
Toluene	mg/kg	<0.050	<0.050	<0.050	0.090	0.050	Q100244
Ethylbenzene	mg/kg	<0.010	0.015	<0.010	0.022	0.010	Q100244
m & p-Xylene	mg/kg	<0.040	0.051	<0.040	0.110	0.040	Q100244
o-Xylene	mg/kg	<0.020	0.028	<0.020	0.047	0.020	Q100244
F1 (C6-C10)	mg/kg	<10	<10	<10	14	10	Q100244

Illustration, invented values. Every real value keeps a pointer to its page and row.

PHASEVIEW

Which format: ESdat, EQuIS?

Honestly, any open one.

Agreeing on one matters far more than which one we pick.

Start small: one results table, one sample table and one location table, with consistent names and units.

Borehole logs and lab data are the obvious first standards.



F4 alone arrives under 32 spellings today. Any shared format fixes that.

PHASEVIEW

Are you naming firms or operators?

No. Firms and operators only appear as tiers.

We're describing the industry as a whole.

Anyone can look up their own public submissions. We just don't put names on slides.



Top 10: 35% Next 40: 43% The rest: 22%

This is as specific as the talk gets. No names.

Who pays for a standard?

Mostly nobody new. The data already exists in digital form.

The cost is agreeing on the shape of a file and attaching it.

We already pay for the alternative: every inherited site gets re-typed by hand.

We've done this before. In 1989 Calgary's well loggers made a free, open format for well logs (LAS). By 1992 Alberta's regulator was asking for it, and it still requires it today.



About 1% of a typical budget goes to data management and past work.

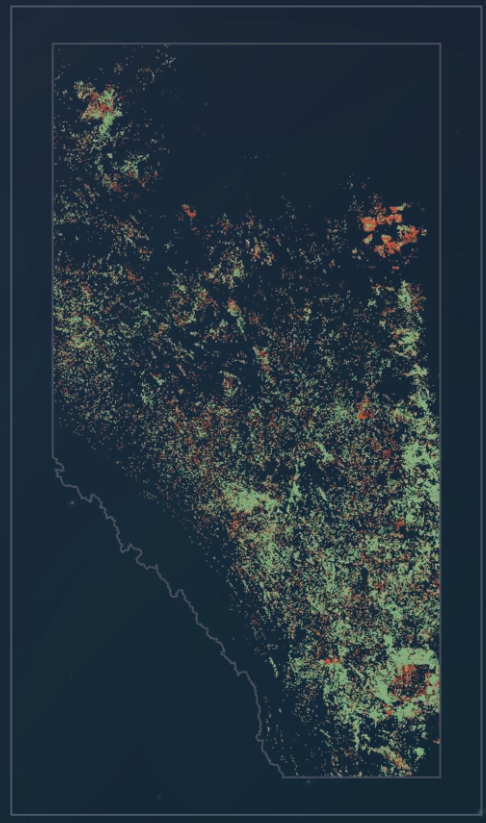
Average of 13 real fee estimates (the budget slide). The re-typing hides in reporting.

Isn't the regulator already doing this?

The regulators keep the reports, and you can find them by location.

What they keep is the PDF. The numbers inside stay locked in it.

They're also the one group that can make a data file stick, by accepting it first and then asking for it.



7 in 10 abandoned wells have a report within 1 km. All of them are filed as PDFs.

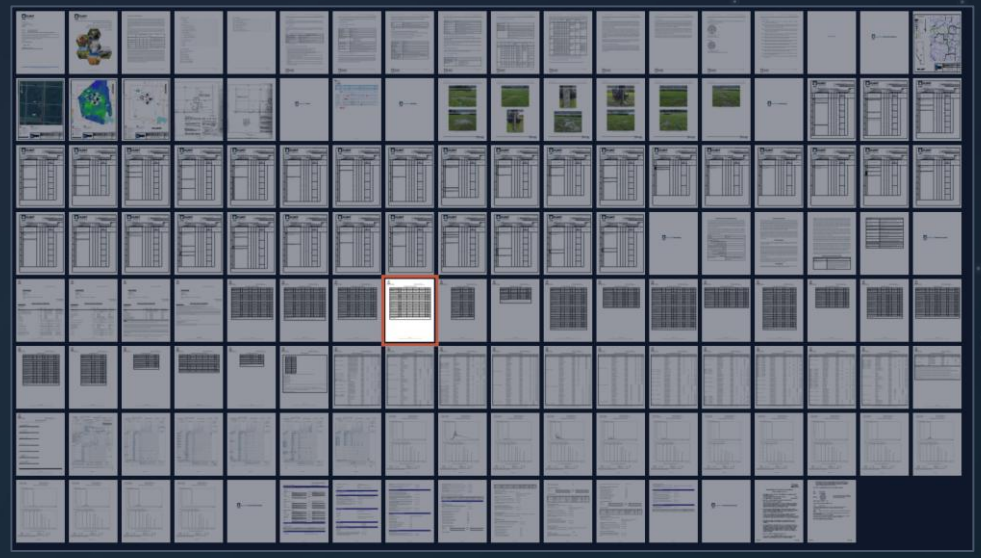
PHASEVIEW

Does this replace consultants?

No. It saves the hours spent re-typing tables. Deciding what the data means is still the consultant's job.

There are more contaminated sites than hands to assess them.

I was the consultant re-typing those tables. I'd have taken the hours back.



Every page of one 142-page Phase 2, to find one table. That's the part we take away.

PHASEVIEW