



ALGOSHRED TECHNOLOGIES



DATA YOU CAN TRUST · CHECKED, LABELLED, OWNED

Know where your numbers came from, and who looks after them.

We help you move your business's numbers into one shared place, through pipes that check them on the way and label each one with where it started. Each important set gets a named owner who decides who gets to see it, so your reports, apps and AI helpers start from data people can trace.

Worth knowing: Checks catch many problems, not all of them. Someone still has to look at what gets set aside.

Move it. Check it. Label it. Own it.

Email contact@algoshred.com with the number your team argues about most, and we'll suggest a first step.

Data engineering & governance, done with your team

Vignesh T.V., CEO & CTO, Algoshred Technologies Pvt. Ltd.

contact@algoshred.com



The problem, and the idea.

Two reports answer the same question with two different numbers.

When a figure looks odd, it takes a long hunt to find where it came from.

Ask who looks after a report and everyone points at someone else.

A customer or the boss spots the mistake before the team does.

People check the main report against their own spreadsheet before they believe it.

Everything went into one big store, and now nobody can find anything in it.

Data engineering and governance means getting your numbers from where they start to the people who use them: moved the same way each time, checked on the way, labelled with where they came from, and looked after by someone named.

The story, decoded. In our story, water is your numbers, the springs are your apps and spreadsheets, the pipes move the numbers into one shared reservoir, the cleaning plant is the checks, Gota's luggage tag is the record of where each number started, the nameplates are named owners, and the private-records tap is who gets to see what. An unnamed bucket is someone nobody has said yes to yet. Tomás is the engineer who looks after the shared pipes; the Mayor stands in for a business leader.



01 Pipes that move your numbers to one place

Numbers from your apps, spreadsheets and partners travel through pipes into one shared reservoir, the same way each time, sorted and labelled. Live where waiting would hurt, once a day where that is fine.

Experts call this: data pipelines, streaming and batch, data platform

Worth knowing “Live” means within seconds or minutes, not at once, and live flows take more effort to run. Use them where waiting would hurt. A big shared store helps only when what goes in is sorted, labelled and owned. The problem is missing labels and owners, not one shared store.

02 A cleaning plant between raw and cleaned

Each delivery is checked before it moves from the raw side to the cleaned side: is it late, is anything missing or doubled, is it the right shape, and does it make sense? What looks wrong is set aside for a named person to look at.

Experts call this: data quality, data observability

Worth knowing Checks catch many problems, not all of them. Someone still has to look at what gets set aside.

03 A tag on each drop

Each set of numbers carries a record of where it came from, each step on the way and which checks it passed, so an odd figure can be followed back to the source that needs fixing instead of argued about.

Experts call this: data lineage, data catalog

Worth knowing Labels show where data came from. They don't make it correct by themselves; that's what the checks, and fixing the source, are for.

04 A name on each set of numbers

Each important set of numbers has an owner from your business, who answers for what it means and decides who can see private records. Engineers look after the pipes themselves.

Experts call this: data ownership, stewardship, data governance, access rules

Worth knowing Rules help when people can follow them easily, so we set them up with the teams who use the data. Helping you see who can open private or personal data is not legal advice, and it does not by itself meet any law's requirements. Your legal advisers have the final say.



Follow one drop: read the tag backwards.

The tag, read backwards

THE STOCK COUNT

- ☐ Cleaned side of the reservoir · caretaker: Tomás
- ☐ Cleaning plant · caretaker: Tomás
- ☐ Raw side of the reservoir · caretaker: Tomás
- ☐ Main pipe · caretaker: Tomás
- ☐ From: Apps spring · owner: the shop manager

THE CUSTOMER LIST

- ☐ Cleaned side of the reservoir · caretaker: Tomás
- ☐ Cleaning plant · caretaker: Tomás
- ☐ Raw side of the reservoir · caretaker: Tomás
- ☐ Hill pipe · caretaker: Tomás
- ☐ From: Old hill tank · owner: nobody yet

Name an owner (here, the Mayor) and the tank can be fixed at its source.

Worth knowing Labels show where data came from. They don't make it correct by themselves; that's what the checks, and fixing the source, are for.

The four checks on the way in

- ☒ **Fresh?** Did it arrive when expected, not days late? **stale**
- ☒ **All there?** Is anything missing, or did something come in twice?
missing bits **double**
- ☒ **Right shape?** Does each value look the way it should, like a date where a date belongs? **odd shape**
- ☒ **Makes sense?** Is it believable, like no stock count below empty?
muddy

All checks on: the glass is clear. Someone still looks at the tray. An illustration, not a measurement. The glass shows clear or cloudy, not a score.

Worth knowing

Checks catch many problems, not all of them. Someone still has to look at what gets set aside.

Who each tap pours for

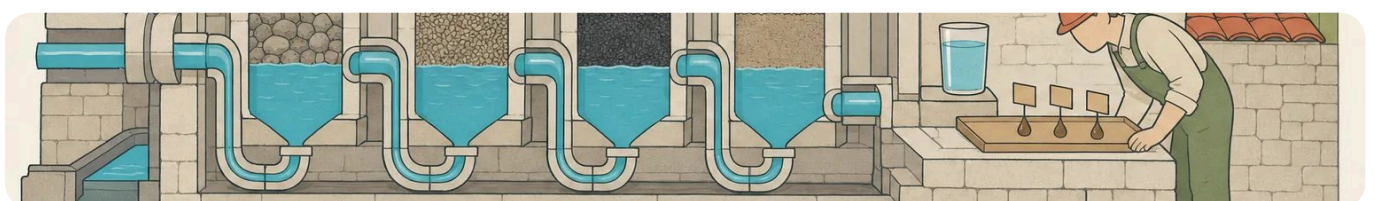
WHO IS ASKING?	BREAD NUMBERS the baker	FIRE-STATION BOARD the fire chief	PRIVATE RECORDS the clinic owner
The baker	● Poured	● Poured	Not on the list. Ask the clinic owner
The fire chief	● Poured	● Poured	Not on the list. Ask the clinic owner
The clinic owner	● Poured	● Poured	● Poured
Someone with an unnamed bucket	Stays dry	Stays dry	Stays dry

Worth knowing

Rules help when people can follow them easily, so we set them up with the teams who use the data.

Worth knowing

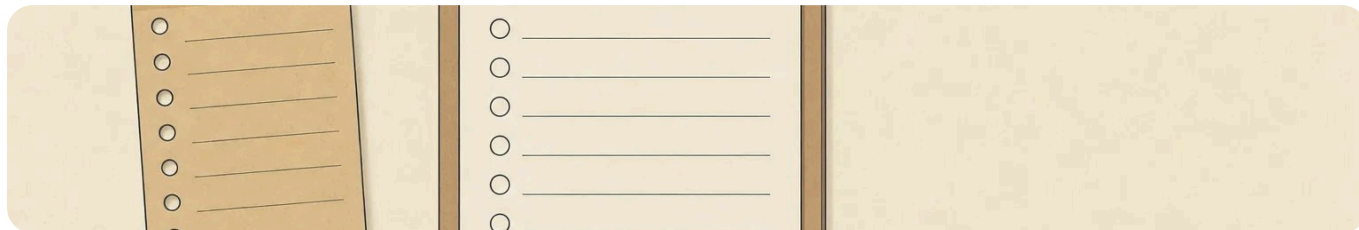
Helping you see who can open private or personal data is not legal advice, and it does not by itself meet any law's requirements. Your legal advisers have the final say.





What we help with.

Start by tracing one number, or bring the pieces together into one checked, labelled, owned supply.



Trace one number back to where it started

We pick one figure that matters to you and follow it back through every spreadsheet, app and step, side by side with the people who use it.

A map of where that number really comes from, where it goes wrong and which source to fix first.

Experts call this: data assessment, lineage discovery

Pipes that run on their own, live or once a day

Automatic routes that bring numbers in from apps, files and partners and deliver them the same way each time, live only where waiting would hurt.

Numbers can arrive without anyone carrying exports by hand.

Experts call this: data pipelines, ETL/ELT (copy, tidy and load, in either order), orchestration (steps run in order), streaming and batch

Tags, a map of your data and agreements between teams

A record of where each set of numbers came from, a findable map of what data you have and who looks after it, and written agreements between the team that sends data and the team that uses it. Written agreements help teams agree before they change things; they don't replace talking to each other.

An odd figure can be followed back, and changes can be agreed before they break a report.

Experts call this: data lineage, data catalog, data contracts

Who can see private records, decided by their owner

Decide who can see which private or personal data, record who opened what, and agree how long it is kept, with laws such as India's Digital Personal Data Protection (DPDP) Act and Europe's General Data Protection Regulation (GDPR) in mind. Helping you see who can open private or personal data is not legal advice, and it does not by itself meet any law's requirements. Your legal advisers have the final say.

Their owner decides, in writing, who may open private records.

Experts call this: data governance, access rules, privacy support

For your technical team: we work with what you already use, for example Databricks, Snowflake, BigQuery, Microsoft Fabric, dbt, Apache Airflow, Apache Kafka, Apache Iceberg and OpenLineage.

One shared reservoir, sorted and labelled

Plan and build one place for your numbers that keeps a raw side and a cleaned side next to each other, moving off old stores step by step. A big shared store helps only when what goes in is sorted, labelled and owned. The problem is missing labels and owners, not one shared store.

Teams can draw from one sorted, labelled place instead of keeping their own copies.

Experts call this: data platform, data warehouse, lakehouse, migration

Checks, and a tray for what looks wrong

Tests on each delivery, before it reaches the cleaned side, for late arrivals, missing or doubled records, odd shapes and values that don't make sense, with what fails set aside for its named owner. Checks catch many problems, not all of them. Someone still has to look at what gets set aside.

Problems can show up at the plant rather than at the tap.

Experts call this: data quality, data observability

A named owner for each set of numbers, and shared meanings

Name an owner and a day-to-day caretaker for each important set of numbers, agree what words like "sale" and "customer" mean, and tidy duplicate records.

Each important question has an agreed meaning and a named person to ask.

Experts call this: data ownership, data stewardship, master data (one record per customer or product)

Getting your numbers ready for AI helpers

Prepare documents and records for AI, tag where they came from, and set things up so AI helpers are meant to read only the material their owners have approved.

AI helpers can start from material its owners have approved. Tagged, checked material gives AI helpers a better start, not right answers. Who may see what is still the owner's call.

Experts call this: AI-ready data, provenance (a record of where it came from)

Capabilities, from tracing one number to pipes your team runs with us.

From following one figure back to its source, to one shared reservoir, to checks and tags on the way in, to named owners and clear rules for private records, and data ready for AI helpers.

Trace and plan

Where your numbers start, and where they go wrong.

- Tracing one important number back to every source
- Talking to the people who make and use the numbers
- A plan tied to one real problem before any rulebook
- A plain map of which sources and pipes to fix first

Check and label it

Checks before the cleaned side, and tags that show where each number started.

- Tests on each delivery: fresh, all there, right shape, makes sense
- Warnings for late, missing or odd data, sent to the named owner
- Where-it-came-from records (lineage) and a findable map of what exists (catalog)
- Written agreements between sending and using teams (data contracts)

Move it

One shared place, fed by pipes that run on their own.

- One shared store with raw and cleaned sides (warehouse, lake or lakehouse), on open file formats where possible, which makes it easier to change tools later
- Automatic pipelines from apps, files and partners
- Live flows where waiting would hurt, daily runs where daily is fine
- Moving off old stores and spreadsheets step by step
- Watching running costs and switching off pipes nobody uses

Own it

A name on each important set of numbers, and clear rules for private records. This is not legal advice; your legal advisers have the final say.

- Named owners and caretakers (stewards)
- Shared meanings for important words, and one record per customer or product
- Who-can-see-what rules, records of who opened what, and how long things are kept (not legal advice)
- Preparing documents and records for AI helpers, with their origin tagged

WHERE IT BECOMES REAL · ILLUSTRATIVE EXAMPLES

Retail and e-commerce

For example, the stock count on the shop floor and the one on the website could come from the same checked pipe, so the two teams compare the same number.

Finance teams

For example, the revenue figure in the board report could carry a tag back to its sources, so an odd number is followed back instead of rebuilt by hand.

Healthcare

For example, records about people could sit behind their own tap, with a named owner deciding who can see them, and a record of who opened them. This is not legal advice; your legal advisers have the final say.

Logistics

For example, delivery updates could flow live where a late answer would hurt, while the monthly summary stays on a simpler daily run.

Manufacturing

For example, machine readings could be checked on the way in, with odd values waiting for the line's owner before they reach the weekly report.

Teams using AI helpers

For example, an assistant that answers questions from company documents could read only tagged, permitted material and show where an answer came from.

Trace first, then build one pipe at a time.



01 TRACE

Follow one number back

We pick one number your team argues about and follow it back through every spreadsheet, app and step, writing down what we find.

02 TOGETHER

Sit with the people who use it

Side by side with the people who make and use the numbers, we agree who owns each important set of numbers, what the important words mean and which source to fix first.

03 ONE PIPE

Build one pipe end to end

We build one pipe from its sources to one report or AI use, with checks, tags, an owner and who-can-see rules in it from the start.

04 NEXT PIPE

Add the next pipe

When the first pipe works for its team, the next one reuses the same pattern, and the pipe map and shared meanings grow with it.

05 RUN IT WITH US

Your team runs it with us

Your owners and engineers run the pipes, look at what the checks set aside and review the rules regularly, with us working beside them. How much your team runs on its own, and how long we stay involved, is agreed with you.

Ways to start

Trace one number. Follow one important figure back to its sources and write up a map of what we find.

Owners workshop. One working session to name owners and caretakers for your most important numbers.

Checks on one pipe. Add checks to one pipe that matters, with what looks wrong set aside for a named person.

What you get

- A map of one number from its sources to the report, with each stop marked
- A short list of sources and pipes to fix first, agreed with you
- One pipe built end to end, with checks, tags and an owner
- A pipe map that says what exists and who looks after it
- Written who-can-see rules for private records, agreed with their owners. This is not legal advice; your legal advisers have the final say.
- Working sessions with the people who will run the pipes



Questions, myths and the next step.

What is data engineering and governance, in plain words?

Getting your numbers from where they start to the people who use them: moved the same way each time, checked on the way, labelled with where they came from, and looked after by a named person who decides who gets to see them.

Do we have to move everything to a new platform?

No. We start with one number that matters and the pipes behind it. A new shared store comes in only where it helps, and moves happen step by step.

Which tools do you work with?

We start from what you already run, for example Databricks, Snowflake, BigQuery, Microsoft Fabric, dbt, Airflow and Kafka. A new tool comes in only where it fills a real gap.

Isn't governance just paperwork that slows people down?

It shouldn't be. We tie the rules to one real problem first, keep them short and build them into the pipes, so the right person's glass gets filled without a chase. Rules help when people can follow them easily, so we set them up with the teams who use the data.

Who looks after the data once it is set up?

People from your business. Each important set of numbers gets a named owner and a caretaker who keeps it running, and your team runs the pipes with us beside them. How much your team runs on its own, and how long we stay involved, is agreed with you.

Does this help with privacy laws like India's Digital Personal Data Protection (DPDP) Act and Europe's General Data Protection Regulation (GDPR)?

It helps you see where personal data goes, who can open it and how long it is kept. Helping you see who can open private or personal data is not legal advice, and it does not by itself meet any law's requirements. Your legal advisers have the final say.

Does everything need to be live?

No. Live flows take more effort to run. We use them where waiting would hurt, such as stock on the shelf or an unusual payment, and keep daily runs where daily is fine. As in the film: bread can wait.

How do you price this?

It depends on scope. We usually start small, by tracing one number, and agree the plan and cost with you before anything bigger.

MYTHS, ANSWERED

"Buy one big data store and it sorts itself out."

A store with no sorting, no labels and no owners turns into a swamp, like the Mayor's giant tank. The shared store is fine; the missing labels and owners are the problem.

"Cleaning it once means we're done."

Water is checked every day, not once. Checks run on each delivery, and someone still looks at what they set aside.

"Data is IT's job."

Building and running the pipes is the engineers' job. What the numbers mean, and who gets to see them, is the business's call, so each important set of numbers has a business owner.

"If it's labelled, it must be right."

A label says where data came from, not that it is right. In the film, Gota's sticker said mountain spring. Her tag said otherwise.

So, would you trust a glass of your numbers?

Email contact@algoshred.com with the number your team argues about most, and we'll suggest a first step.

contact@algoshred.com

WATCH THE FILM

algoshred.com/data-engineering