



THE LAB · ENTRY 04 · PROPOSAL

# Seven things to build, and *one word we have not earned.*

The first thing anybody will ever use is a page that asks what they run, computes what they granted, and hands them the gap. This is its build specification: three tools, two vaults, two documents — and the arithmetic that says a twenty-connector question is not anonymous, which changes what may leave the browser.

---

**Edition** v1 · 12 September 2026

**Live page** [riskmandate.ai/lab-shape-collector.html](https://riskmandate.ai/lab-shape-collector.html)

**Published by** RiskMandate · CC BY 4.0

**This is a dated edition of a page that changes.** The Lab holds our current thinking, and current thinking moves. This PDF does not: it is what we thought on the date stamped below, kept so the reasoning can be followed rather than only its conclusion. The live page may since have been corrected, extended or withdrawn — and if it has, the edition list on it will say so.

## WHAT THIS IS

### The cheapest product we have, and it asks one thing.

Every other route into what an agent can do requires something to be installed. This one asks. Somebody says which assistants they use, on which surfaces, with which connectors switched on; the page works out what that grants; and it hands back the difference between that and what they would have said they intended. No install, no administrator, no procurement.

- **It is stages two to four of the flow next door, collapsed into one screen.** [Lab 02](#) draws twelve stages from a stranger's first question to a delivered policy. Self-select, draft and correct are three of them, and the finding here is that they are not three conversations — they are one page where the picture on the right changes as the answers on the left do, and correcting it is done by changing an answer rather than by writing to us.
- **The output already has a name.** It is *the shape* — the deployment shape, which is the noun the model site already uses for a configuration somebody actually runs. The collector produces a shape record and the grant is derived from it. No new noun.
- **It is also the whole of the smallest paid thing.** Their answers, their measured grant, the delta between them, as a file they keep. That is what the first tier on [the pricing page](#) describes, and until now it has been a questions page waiting for a tool.
- **And it is honest about what it is not.** This finds what people will tell you. That is a different thing from what is on the network, and neither one is complete. That sentence belongs on the product, not in a footnote — see [below](#) for why nobody else says it.

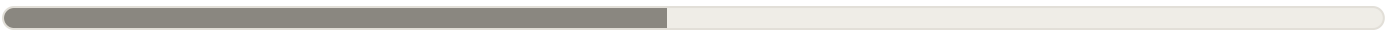
#### The collector asks the shape, computes the grant, and returns the delta.

Everything on this page after that sentence is about one question: what, if anything, is allowed to leave the browser.

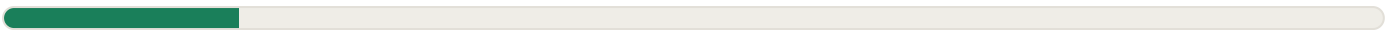
## THE ARITHMETIC

### A twenty-connector question is **not anonymous**.

The instinct is that answers like these are harmless in bulk, so they can simply be collected. That instinct is wrong, and it is wrong measurably rather than arguably — which is the good case, because it means the fix is arithmetic rather than judgement.



2<sup>20</sup> — about a million distinct answers to one question.



4 bands × 2<sup>5</sup> category combinations = 128 answers. A reduction of about 8,000×, and the research loses almost nothing — the interesting finding is *how many* and *of what kind*, not which brand.

**One benchmark makes the first bar concrete.** A browser-uniqueness study of 470,000 samples measured about 18 bits across a whole fingerprint — user agent, plugins, fonts, screen, timezone, the lot — and found 83.6% of visitors instantaneously unique on it. **Our connector question, on its own, carries more entropy than that entire fingerprint.**

| RESPONSES | CELLS THAT COULD HOLD A GROUP OF FIVE | REALISTIC OUTCOME                                                            |
|-----------|---------------------------------------|------------------------------------------------------------------------------|
| 100       | at most 20                            | Effectively every respondent unique                                          |
| 500       | at most 100                           | Effectively every respondent unique                                          |
| 5,000     | at most 1,000                         | Common shapes reach a group of five. <b>Every interesting shape does not</b> |
| 50,000    | at most 10,000                        | The head is safe and the tail still singles out                              |

- **The nominal space is around five trillion cells** once assistants, surfaces, role and company size are multiplied in — roughly 42 bits. Real distributions are skewed, so far fewer cells are ever occupied. The comparison that decides the question is not with the space, it is with the number of respondents, and at any sample we will plausibly reach, the respondents lose.
- **The regulator's test is singling out** — the ability to isolate the records relating to one person — assessed against a motivated intruder assumed reasonably competent with ordinary resources, and it cites groups of five as strong protection. Nothing in the table above reaches that.
- **A second anchor, from outside our field entirely.** Three low-cardinality fields — postcode, gender and date of birth — uniquely identify about 87% of a national population. A configuration is a fingerprint in the literal sense, and fingerprints do not become less identifying because the person typed them in themselves.

- **So the submissions are personal data and must be handled as such.** That is not an argument against building it. It is an argument for building it the way described next, which costs almost nothing and makes the product better.

**Where the banding lands us, honestly.** Applied across every field, the banding cuts the nominal space from about 42 bits to about 20 — a factor of roughly four million, and the single largest reduction available to us. It also leaves a banded submission at *approximately the entropy of that browser fingerprint*. Banding is therefore the necessary first move and not the whole answer: the submission only stops singling people out once [suppression on output](#) and [a written assessment](#) exist alongside it. This paragraph is our own arithmetic, derived from assumptions we have shown rather than measured from data we do not have.

#### THE RESOLUTION

### Compute locally in full. **Submit banded.**

This keeps everything the product needs and gives up nothing the user wants. The full configuration is computed in the browser and never leaves it; what crosses the wire, and only if somebody presses submit, is a banded record.

STAYS ON THE MACHINE

The full shape, and everything derived from it

Every connector by name, every surface, every assistant. The grant is computed from the full shape, the delta derived, the label and the leaflet rendered. The user sees all of it. It is their configuration.

- named** connectors, assistants, surfaces
- computed** grant, delta, barriers
- typed** anything written freely
- kept** in this browser, to be resumed

EXPLICIT  
SUBMIT  
ONLY

MAY CROSS THE WIRE

Bands and categories. Nothing else.

A record with no names in it, no identifier of any kind attached to it, and nothing recorded about the request that carried it.

- band** connector count, 4 values
- categories** mail, files, calendar, code, chat
- coarse** desktop, web, editor, command line
- bands** role (3–4), company size (3)

| FIELD                             | FULL, LOCAL              | SUBMITTED                                                                                             |
|-----------------------------------|--------------------------|-------------------------------------------------------------------------------------------------------|
| connectors                        | The named list of twenty | <b>A count band</b> — none, 1-2, 3-5, 6+ — <b>plus categories</b> : mail, files, calendar, code, chat |
| assistants                        | The named list           | A count band and the categories                                                                       |
| surfaces                          | Each one                 | Coarse — desktop, web, editor, command line                                                           |
| role                              | Free choice              | <b>Three or four bands</b>                                                                            |
| company size                      | Exact                    | <b>Three bands</b>                                                                                    |
| free text                         | Shown back to them       | <b>Never submitted.</b> Free text cannot be anonymised                                                |
| address, user agent, precise time | Not needed               | <b>Never recorded</b> — each one independently defeats the banding                                    |

## **A collector that cannot read back cannot correlate.**

That is the estate's own enforcer test — a control bounds a grant only if it is enforced by something the grant does not include — applied to our own product. A promise not to correlate is bounded only by something the collector does not have, and here that something is read access. It is a structural claim rather than a policy, which is the only kind worth publishing.

- **The write-only lane between two vaults is therefore not a platform demonstration.** It is the correct privacy architecture for this specific job, and saying so turns a capability into an argument. The submitting side holds a write credential and no read key; it cannot join a submission to an earlier one, cannot read across respondents at the point of collection, and cannot reconstruct a session — because it cannot read anything at all.
- **One hard rule follows, and it goes in the build rather than in a policy.** A submission carries *no stable identifier of any kind*. Not a session token, not an install identifier, not a hash of anything, not a salted hash, not a "random" value that persists past the page. The pattern to copy is the install counter that refuses identifiers outright and buckets by week.
- **Suppress on output, and check the subtraction.** Never render a published cell backed by fewer than five submissions, and verify that a suppressed cell cannot be recovered by subtracting the ones around it from a total. The second half is the part that gets forgotten.

### WHAT HAS TO BE BUILT

## **Three tools, two vaults, two documents.**

Nothing below exists. Each card says what the thing is, where it runs, what it holds, what it must never do, and what has to be decided before it can be started. The two documents are on this list because without them the product cannot use the word it wants to use.

## T1 The shape collector

TOOL · BROWSER

The page itself: under fifteen questions, one item per screen on a phone, an honest stated duration, a picture that accumulates on the right as the answers come in on the left, one guess screen before the reveal, and a result the user can keep. Everything it computes, it computes locally.

|            |                                                                                                                                                                                                        |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RUNS       | Entirely in the browser. Static page, no build step, works from a file on disk — the same constraint as every page on this site.                                                                       |
| HOLDS      | In-progress answers in local storage, so a half-finished assessment survives a reload. <b>Strictly necessary</b> — see <a href="#">the two regimes</a> — so no banner.                                 |
| PRODUCES   | A shape record (T2), the grant derived from it, the delta, and a file the user keeps.                                                                                                                  |
| NEVER      | <b>Sends anything on its own. Submission is a separate, explicit, unticked act on its own screen.</b>                                                                                                  |
| BLOCKED ON | Open questions <a href="#">1</a> , <a href="#">2</a> and <a href="#">3</a> — the exact question count, whether connectors are one multi-select or category-then-count, and what the guess screen asks. |

## T2 The shape record schema

DOCUMENT · OPEN DATA

A closed, controlled vocabulary for a *deployed configuration*: which assistant, on which surface, with which connectors granted which scopes — and the mapping from that to the capability primitives it grants. Twice this month the answer to "do we need to build this" has been "it already exists". Not this time.

|            |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WHY        | Every standard in the area describes the <b>artefact</b> and none describes the <b>deployment</b> . The machine-learning component bill of materials, standardised as an international specification in December 2025, describes a model: parameters, task, architecture family, datasets, inputs, outputs, considerations. The other bill-of-materials family has an equivalent profile. <b>Neither says which assistant a person runs, on which surface, with which connectors.</b> That is a genuine gap. |
| REUSES     | The published capability grammar — verb, object, reach — from <a href="#">abp.sgit.ai</a> . The schema's whole job is to map a configuration to a set of those primitives, and nothing more.                                                                                                                                                                                                                                                                                                                 |
| DISCIPLINE | <b>Derive rather than assert.</b> Every row from a connector to the capabilities it grants carries a source, a date, and whether it was measured or derived. That material is the research in <a href="#">Lab 01</a> , and the per-row provenance field it needs is <a href="#">Request 3 on the model site</a> .                                                                                                                                                                                            |
| NEVER      | <b>Gets called an ontology. The value is the closed vocabulary, not the formalism around it.</b>                                                                                                                                                                                                                                                                                                                                                                                                             |
| BLOCKED ON | Open question <a href="#">7</a> — whether this publishes early and invites correction, or stays internal until enough policies exist to know its shape.                                                                                                                                                                                                                                                                                                                                                      |

## V1 The submitting vault

VAULT · WRITE-ONLY

The lane the browser hands a banded record to. It holds a write credential for V2 and no read key for anything, which is what makes the privacy claim structural instead of promissory.

|            |                                                                                                                                                                  |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CAN        | Write one banded record into the collection vault.                                                                                                               |
| CANNOT     | <b>Read. Not its own writes, not V2, not a previous submission. This is enforced by the absence of a read key rather than by code we wrote.</b>                  |
| CARRIES    | The banded record and a week bucket. <b>No stable identifier of any kind.</b>                                                                                    |
| RECORDS    | <b>No address, no user agent, no precise timestamp. Each of those independently defeats the banding, which makes them a build rule rather than a preference.</b> |
| BLOCKED ON | Nothing. This one can be provisioned today, and is the smallest piece of the seven.                                                                              |



## V2 The collection vault

VAULT · APPEND-ONLY

Where banded records accumulate. Read access sits with us and not with the collector, which is the entire point of splitting it from V1.

|            |                                                                                                                                                                                         |
|------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| HOLDS      | Banded records, one per submission, bucketed by week. Nothing that could be joined to a person, a session, or another record.                                                           |
| VERSIONED  | By the vault, so the corpus has a history rather than a current state — the same discipline the rest of the estate claims.                                                              |
| NEVER      | Holds free text. Not in a notes field, not in an "other" box, not in a bug report pasted into it later.                                                                                 |
| BLOCKED ON | Open question 6 — the suppression threshold, which is a property of what we publish from here rather than of the vault itself, but which should be decided before anything lands in it. |

## T3 The aggregator

TOOL · OURS

The thing that turns a collection vault into something publishable, and the only place the suppression rule can actually be enforced. It is listed separately because a rule that lives in prose is not a rule.

|            |                                                                                                                                                                                     |
|------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ENFORCES   | No published cell backed by fewer than five submissions — and a check that a suppressed cell cannot be recovered by subtracting its neighbours from a published total.              |
| PUBLISHES  | Cells, bands, and counts. Where the threshold bites, it says a cell was suppressed rather than showing a zero, because a silent zero is a lie about the data.                       |
| NEVER      | Returns a row. Not to us on a dashboard, not to a customer, not in an export. There is no legitimate use for an individual banded row that a suppressed cell does not serve better. |
| BLOCKED ON | Open question 6, and on V2 holding enough that suppression does not simply hide everything — which it will, early on, and the page should say so rather than wait.                  |

## D1 The motivated-intruder assessment

DOCUMENT · SHORT

A short written assessment naming who the motivated intruder would be, what they would already know, and what the banded corpus would add to it. The obligation is not to reduce the risk to zero; it is to make a reasoned assessment and record it.

|                             |                                                                                                                                                                                                   |
|-----------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| WHY IT IS ON THE BUILD LIST | Because it is a <b>precondition for a word</b> , not a compliance chore. Until it exists, the product may not say <i>anonymous</i> — see <a href="#">below</a> .                                  |
| NAMES                       | For the public mode: somebody who reads the published cells and wants to identify a respondent. For the employer mode: <b>the buyer</b> , which is an uncomfortable sentence and the correct one. |
| LIVES                       | Published, on this site, dated, and revised when the instrument changes. An assessment nobody can read is indistinguishable from one nobody wrote.                                                |
| BLOCKED ON                  | Open question <a href="#">4</a> — who writes it. It is a short document and it gates the language on the product, which makes it the highest ratio of consequence to effort on this page.         |

## T4 The employer instrument

TOOL · SECOND PRODUCT

The same questions, sent by an employer to its own staff, to find out how many agents are in use. It is a real second product and it is *not the same product*, because the law changes when the employer sends the link. Detail in [its own section](#).

|            |                                                                                                                                                                                  |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| DIFFERS BY | Lawful basis, an impact assessment, and the output. Probably not by the instrument — same questions is cheaper, and the consent problem does not care which questions are asked. |
| NEVER      | Returns an individual row to the employer. Only cells above the suppression threshold, only bands — and the page the employee sees says so <b>before</b> they answer.            |
| REQUIRES   | Its own impact assessment before it runs. That is a piece of work, not a paragraph, and it is not D1 with the names changed.                                                     |
| BLOCKED ON | Open question <a href="#">5</a> , and on T1 existing. This is fourth in order and first in revenue, which is a tension worth naming rather than resolving early.                 |

**Order, if only some of it happens.** V1 and V2 are hours and unblock the honest version of everything else. T1 is the product. D1 is a short document that decides what T1 is allowed to say, so it should not trail the build. T3 can wait until there is anything to aggregate, but its rule has to be agreed before T1 ships or the first publication will breach it. T2 is the largest piece and the one with outside value. T4 is the largest revenue and the largest legal surface, and it should not be started while T1 is still moving.

TWO REGIMES

Storing the answers is one thing. Sending them is another.

The analysis splits cleanly, and the design should follow the split rather than cover the whole tool with one consent banner. Three of these four rows need no banner at all. The fourth needs more than a banner.

| OPERATION                   | REGIME                                      | POSITION                                                                                                                                                                                                                                |
|-----------------------------|---------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| storing in-progress answers | Storage rules, in scope                     | <b>Strictly necessary.</b> Assessed from the user's point of view, and somebody who starts a self-assessment plainly wants their answers to persist. <b>No banner</b>                                                                   |
| reading them back           | Storage rules, in scope                     | Same exception, same reasoning                                                                                                                                                                                                          |
| analytics about the tool    | Storage rules, in scope                     | Drop-off point, device class, which question loses people. <b>The statistical-purposes exception fits,</b> with clear information and a simple free means to object — a toggle defaulted on. <b>Browser settings are not sufficient</b> |
| submitting the answers      | Not the storage rules. Data protection only | <b>An explicit, separate, unticked act.</b> The answers are the content the user produced, not statistics about how the service is used                                                                                                 |

**The statistical exception covers *how* a service is used. It does not cover *what* the user told it.**

Those are different things and only the first is exempt. The regulator is explicit that the exception is not a broad one covering all analytics — which means the submit screen is a real screen, with its own unticked control, and not a line in a footer.

**This is research, not legal advice.** The split above is read off the regulator's own published guidance, on a stated date, and the citations are at the bottom of this page. An impact assessment for the employer mode is a piece of work rather than a paragraph, and nothing here substitutes for it.

## COMPLETION

### Do not design it as a game. Design it to finish.

The instinct that the game framing carries over from the teaching work is wrong, and the evidence that applies here is a different literature. This is a data-collection instrument, not a teaching artefact: what matters is completion, and that literature is well developed and mostly contradicts the folklore.

| PROGRESS INDICATOR                    | EFFECT ON DROPPING OUT                        |
|---------------------------------------|-----------------------------------------------|
| constant speed – an honest linear bar | <b>No significant effect</b>                  |
| fast at first, then slowing           | Drop-off odds multiplied by about <b>0.80</b> |
| slow at first, then speeding up       | <b>Drop-off odds up by 56%</b>                |

- **The harmful pattern is exactly what this tool produces by accident.** A short introductory section followed by a long connector-by-connector section makes the bar stall in the middle — which is the slow-then-fast shape, the one that raises drop-off odds by more than half. The meta-analysis behind the table covers nineteen studies and thirty-two experiments and concludes that its findings question the common belief that progress indicators reduce drop-off at all.
- **One question per screen does not reliably buy completion either.** A controlled comparison of a conversational form against an ordinary one found 89.8% against 91.4% completion — no significant difference — and the conversational version took **53% longer**. A 2026 field experiment found it rated more original and more entertaining, harder to navigate, and with no practical data-quality advantage.
- **And the widely quoted commercial figure does not survive inspection.** A conversational-form vendor claims 47.3% completion against a stated industry average of 21.5%, with no methodology, no denominator definition and no independent verification. Its companion claim about rich media is an uncontrolled correlation between finishing a form and how much effort somebody put into building it.

| RANK | LEVER                                                 | EVIDENCE                                                                                                                                                  |
|------|-------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1    | <b>Fewer questions</b>                                | The only lever with both randomised and large observational support, and the observational data shows a cliff above roughly fifteen                       |
| 2    | <b>An honest, short stated duration</b>               | Randomised and replicated. <b>The announcement is itself a treatment</b> , and an indicator only helped when the task was promised short and actually was |
| 3    | <b>Works on a phone, one item at a time, no grids</b> | Strong on quality, good on completion. Most responses will be mobile                                                                                      |
| 4    | A high-value question early                           | One clean experiment, one large observational study. Modest but real                                                                                      |
| 5    | Showing partial results                               | <b>Satisfaction significantly higher, completion roughly unchanged.</b><br>Treat as an untested hypothesis                                                |

- **The accumulating picture is still worth building, for the honest reason.** It makes the thing feel worth finishing and it makes the output legible. It is *not* evidenced as a completion mechanism, so it should be measured rather than assumed — which is the fifth row of the table above, not the first.
- **One game mechanic earns its place, and it is the estate's own.** Before revealing what the selected connectors grant, ask the user to guess the number. Stating a belief before being shown the answer is the mechanic the games site is built on; it costs one screen; and it is what makes the result land rather than scroll past. It is also what turns a form into the experience the game instinct was reaching for, without the 53% time penalty of a conversational interface.

|                                                                           |                           |
|---------------------------------------------------------------------------|---------------------------|
| <b>Which assistants</b> — multi-select, one screen                        | 1                         |
| <b>Which surfaces</b> — coarse, four options                              | 1                         |
| <b>Which connector categories</b> — mail, files, calendar, code, chat     | 1                         |
| <b>How many in each selected category</b> — a band per category           | 2 typical · 5 worst       |
| <b>Role</b> — three or four bands                                         | 1                         |
| <b>Company size</b> — three bands                                         | 1                         |
| <b>The guess</b> — a belief stated before the reveal, not a data question | 1                         |
| <b>Submit</b> — an explicit unticked act, not a question                  | 1                         |
| <b>Screens before the result</b>                                          | 9 typical · 12 worst case |

**That is the budget, and it is the constraint everything else negotiates against.** Under fifteen, an honest stated duration, one item per screen on mobile, a front-loaded progress indicator or none at all, and one guess screen before the reveal. Every feature proposed after this page has to say which screen it is taking, or which one it replaces.

## THE SECOND PRODUCT

### When the employer sends the link, the law changes.

The same instrument, circulated inside a company to find out how many agents are actually in use, is a real product and a different one. Three things move, and the third is the one that decides the design.

- **Consent stops working.** The regulator's guidance on monitoring workers says consent is not usually appropriate in the employment context because of the imbalance of power, and that it must be freely given and withdrawable without detriment. A survey circulated by an employer, whose answers could reveal a policy breach, is close to the worst case for freely given consent. The realistic basis is legitimate interests with a documented assessment.

- **An impact assessment is likely required before it runs.** One must be carried out before processing likely to cause high risk to workers' interests, and a tool that inventories which systems an individual employee uses, on which devices, with which data connectors, is systematic monitoring of workers.
- **And the employer holds the identifying context we do not.** The regulator recognises relative anonymity explicitly: information can be personal data in one organisation's hands and anonymous in the hands of another that lacks the context. The employer has the organisation chart, the device fleet and the sign-on logs.

**"One product person on a desktop assistant with mail, files and code connectors" is anonymous to us and a name to them.**

So the employer mode never returns an individual row. Only cells above the suppression threshold, only bands, and the product says so on the page the employee sees before they answer.

**The privacy constraint and the data-quality constraint are the same constraint.** A version that could return a row is also the version an employee answers carefully rather than honestly. Suppression is not a concession extracted from the product here — it is the only reason the answers would be worth collecting.

## THE MARKET

### **Nobody else asks. That is the opening and the problem.**

A ten-vendor comparison of the discovery market, published 7 September 2026, found that all ten discover through technical telemetry — browser extensions, endpoint agents, network inspection, API integrations, sign-on logs — and none through self-report. One free assessment tool disparages self-report explicitly and recommends usage telemetry instead. Both halves of that matter.

## THE OPENING

### Self-report reaches what telemetry cannot

A personal device, a personal account and a locally run server are all invisible to every method on that list, and all three are where the interesting deployments are. One of the larger vendors implicitly concedes the gap by selling desktop agents to close it.

**no** procurement  
**no** installation  
**no** administrator

○ SAME  
FACT

## THE PROBLEM

### There is no vocabulary of trust for it

No incumbent treats survey data as a legitimate discovery source, so the burden of explaining why it counts falls entirely on us — and it cannot be met by claiming more than the method delivers.

**says** what people will tell you  
**not** what is on the network  
**neither** one is complete

- **The completion standard to beat is a free seven-question organisational assessment** taking about two minutes, with no email gate. That is the bar, and our budget above is nine screens — so the honest stated duration has to be honest.
- **And nothing we found asks an individual which assistants and connectors they personally use and returns them a result.** That is the specific thing being built, and it is the reason this is worth doing at all.

## LANGUAGE

### The honest word, until then, is banded.

One sentence of product copy is doing more damage than any missing feature on this page, and it is the one that says the data is anonymous. It is a claim with a defined test behind it, and we do not currently pass the test.



**The word *anonymous* may not appear in this product until all three of these exist:**

1. **The banding** — no named connector list, no free text, no address, no user agent, no precise time, no stable identifier. *Specified above; not built.*
2. **The suppression** — no published cell below five, and the subtraction check alongside it. *Specified above; not built.*
3. **The assessment** — written down, naming the motivated intruder, published and dated. *Not written.*

- **Until then the product says *banded*,** and explains what that means in a sentence: we ask how many and of what kind, never which brand, and never anything you typed. That is a claim we can currently support, which is the only kind this site publishes.
- **This is the same rule the rest of the site runs on.** No claim about a third party without a source and a date; no score without stated inputs; no capability asserted without saying whether it was measured or derived. A privacy claim is not exempt from the discipline just because it is about us.

#### OPEN QUESTIONS

### Seven things we **cannot settle alone.**

These are real rather than rhetorical, and four of the seven block a card in the build list above. If you have an opinion on any of them, that is more useful to us today than agreement with the rest of the page.

| # | QUESTION                                                                                                                                                                          | BLOCKS  |
|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| 1 | <b>How many questions, exactly?</b> The cliff is around fifteen and the connector question alone could be twenty items if built carelessly                                        | T1      |
| 2 | <b>Is the connector question one multi-select, or a category then a count?</b> The second is the banding made native, and it may lose people who want to see their own tool named | T1      |
| 3 | <b>What does the guess screen ask?</b> One number, or one number per category                                                                                                     | T1      |
| 4 | <b>Who writes the motivated-intruder assessment, and where does it live?</b> It is a short document and it is a precondition for a word                                           | D1      |
| 5 | <b>Does the employer mode need a different instrument, or the same one with a different output?</b> Same instrument is cheaper, and the consent problem does not care             | T4      |
| 6 | <b>What is the suppression threshold?</b> Five is the cited standard, and a small early sample will suppress nearly everything                                                    | T3 · V2 |
| 7 | <b>Does the shape schema become the published vocabulary, or stay internal until the fifth policy?</b> Publishing early invites correction and locks a shape too soon             | T2      |

HONEST TENSIONS

And seven we are choosing to live with.

| TENSION                    | BOTH HALVES ARE TRUE                                                                                                                    |
|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| banding the submission     | It is what makes the data lawful to hold, and it throws away the brand-level detail that would be the most interesting thing to publish |
| computing locally          | It is the strongest privacy position available, and it means we learn nothing at all unless somebody presses submit                     |
| not designing it as a game | The evidence is about completion rather than enjoyment, and the instinct about feel is what will make people start                      |
| the guess screen           | It is the one mechanic with a published argument behind it, and it adds a screen to an instrument whose main lever is fewer screens     |
| employer mode              | It is a real second product, and it carries an impact assessment, a weak lawful basis, and an intruder who is the buyer                 |
| self-report                | It reaches what telemetry cannot, and no incumbent treats it as legitimate                                                              |
| building the schema        | It is a genuine gap, and it is the third schema this estate has taken on in a fortnight                                                 |

**What this page deliberately is not.** It is not a schema — the gap is established and the disciplines are named, but no field is defined. It is not a design — question count, ordering and the accumulating picture are constrained by evidence, and nothing is drawn. It is not legal advice. It is not a privacy claim, for the reason in [the section above](#). And it is not a completion promise: the evidence constrains the design and predicts nothing whatsoever about our numbers.

Written 12 September 2026 from a dev brief of the same date. Sources, all read 12 September 2026 — **completion:** the progress-indicator meta-analysis of 19 studies and 32 experiments, Villar, Callegaro and Yang 2013, *Social Science Computer Review* 31(6); the expectation-matching result, Yan, Conrad, Tourangeau and Couper 2011, *IJPOR* 23(2); the conversational comparison, Kim, Lee and Gweon 2019, CHI; the 2026 field experiment, Cavusoglu Deveci, Fuchs and Metzler, *BMS* 169-170(1), 6 March 2026; the item-by-item finding, Revilla, Toninelli and Ochoa 2015; the personalised-feedback trial, Kühne and Kroh 2018, *SSCR* 36(6); vendor completion claims at [typeform.com](https://typeform.com), treated as marketing with no methodology. **Schemas:** [cyclonedx.org](https://cyclonedx.org), machine-learning component bill of materials v1.7, standardised as an international specification December 2025; the alternative family's profile at [spdx.github.io/spdx-spec](https://spdx.github.io/spdx-spec). Neither describes a deployed configuration. **Identifiability:** the regulator on effective anonymisation, the singling-out test, the motivated-intruder test and groups of five, at [ico.org.uk](https://ico.org.uk); Sweeney on postcode, gender and date of birth; Eckersley 2010 on 470,000 browser samples and 83.6% instantaneous uniqueness. **Storage and submission:** the guidance on storage and access technologies, published 29 April 2026, at [ico.org.uk](https://ico.org.uk), including the strictly-necessary test assessed from the user's point of view, the statistical-purposes conditions, the objection mechanism that may not be a browser setting, and the statement that the exception is not a broad one covering all analytics. **Employment:** the regulator on monitoring workers, October 2023 and last updated 16 June 2026, at [ico.org.uk](https://ico.org.uk). **The market:** a ten-vendor comparison published 7 September 2026; [nudgesecurity.com](https://nudgesecurity.com) for the one transparent price found; the free seven-question assessment at [aona.ai](https://aona.ai), which recommends telemetry over self-report. **Inside the estate:** the capability map and its 23 primitives at [what-can-it-do.games.sgit.ai](https://what-can-it-do.games.sgit.ai); the belief-before-answer mechanic at [games.sgit.ai](https://games.sgit.ai). The entropy arithmetic in [the second section](#) is ours, derived from stated assumptions rather than measured.

---

## ● Lab 04

# Two vaults are hours. One word is a document.

The write-only lane can be provisioned today and it is what makes the privacy claim structural rather than promissory. The assessment is a short piece of writing and it decides what the product is allowed to say. Neither is the hard part, and both are ahead of the hard part.