



Relio

Project memory for the Acme rollout, with a receipt on every claim.

1

2–4

17

<https://relio-ai.eu>

RELEX challenge track "Memory With a Receipt" · AaltoAI Hackathon, 18–20 September 2026

A year in, nobody agrees on what was decided.

The Acme archive: 45 documents, 23 Teams transcripts, 20 email threads, 2 status report threads, March 2024 to July 2026. People have changed jobs. An assistant pointed at the pile produces a clean summary that is wrong in ways nobody catches until a meeting goes badly.

1

Reports a reversed decision as current

Bakery was in the fresh scope in July 2024 and removed at the November 2025 go/no-go. The first answer alone is wrong.

2

Calls a suggestion an agreement

A consultant floating an idea is not the customer agreeing. Repeated back as a commitment, it starts a fight.

3

Repeats a record that was never true

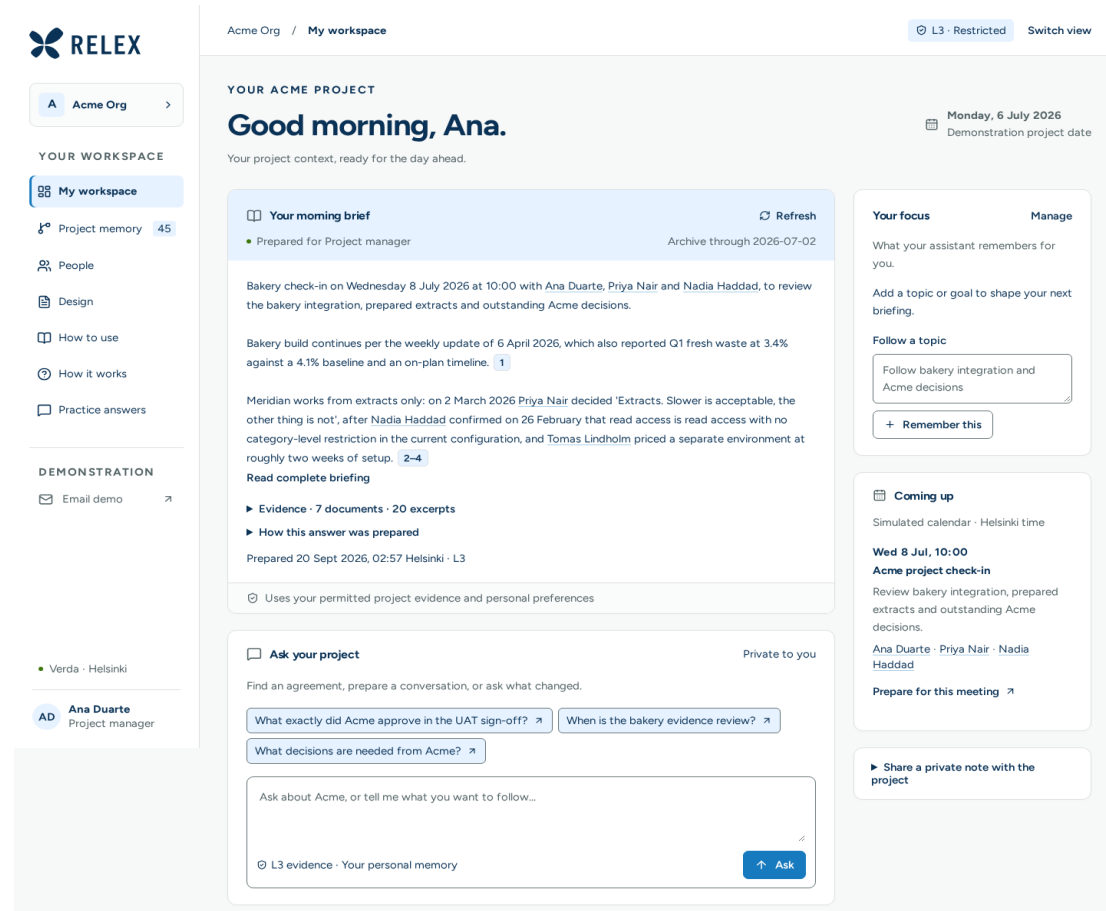
Weekly reports said "nightly extract completed, no errors". The contents had never been checked.

4

Cannot forget anyone

Any named person can ask to be removed, and the removal has to reach the index, the embeddings and every cached summary.

One workspace, shared project memory



The screenshot displays the RELEX workspace interface. On the left, a sidebar shows the RELEX logo, a navigation menu with 'Acme Org', 'My workspace', 'Project memory', 'People', 'Design', 'How to use', 'How it works', and 'Practice answers', and a 'DEMONSTRATION' section with 'Email demo'. The main content area is titled 'YOUR ACME PROJECT' and 'Good morning, Ana.' It includes a 'Your morning brief' section with project context, a 'Your focus' section with a 'Manage' button, a 'Follow a topic' section with a 'Remember this' button, a 'Coming up' section with a 'Simulated calendar', and an 'Ask your project' section with a 'Private to you' button. The interface also shows a date 'Monday, 6 July 2026' and a 'Switch view' button.

Morning brief

Prepared for the employee from permitted evidence, role, saved goals and the calendar. Every sentence has a dated receipt.

Project memory

148 extracted decisions, commitments, questions and corrections, each linked to its source records and shown in a graph.

Email participation

Cc the agent in Outlook and the message joins shared memory. The agent replies only when it has something to correct.

Ask the project

Answers with numbered citations that open the exact message or transcript turn, and the turns around it.

People

Every named person with their role, organisation from email evidence, source connections and an erasure control.

Access levels

L1 customer-shareable, L2 RELEX internal, L3 restricted. Evidence is filtered on the server before the model sees it.

Where each scored dimension lives in the product

Dimension	Weight	What Relio does	Where a judge can check it
Provenance	25%	Each claim cites a record ID: document, author, date, message or turn position. Citations open the exact passage.	Any answer → click a chip → source receipt
Attribution	20%	Claims carry proposed / agreed / corrected / superseded status with who said it and who confirmed.	Project memory → Decisions & observations
Currency	20%	Chronology questions inventory every document, then read original turns. Superseded and never-accurate records are labelled.	Ask "How has shelf-life population changed?"
Deletion	20%	Person erased from runtime source copies, claims, answers, mail and calendar state. In-flight inference cancelled. Receipt with counts.	People → pick anyone → Confirm person erasure
Initiative	15%	Agent in Cc reads new mail, corrects a consequential error unprompted and stays silent on routine mail. Brief refreshes on its own.	Email demo → example 2 → send

Residency: the archive, the application, the database and the model all stay in Helsinki. Details on slide 11.

Every claim points at a record, and the record opens

- 2,974 source records imported into SQLite: 110 email messages, 25 report updates, 2,839 transcript utterances. Each keeps its original ID, document, author, date, position and access level.
- Extracted claims store the record IDs that support them. Saved answers store the exact excerpt text, the source references, the employee, the access level and the workspace version.
- The server binds citation numbers to real records after generation. A reference to a document without a position, or a quote that is not a substring of the record, is rejected and repaired.
- A second model pass checks the draft against the cited passages and the neighbouring turns before the answer is shown.
- Clicking a chip opens the passage with its surrounding conversation. Text and PNG exports keep the same numbering.

Source receipt

Shelf life field mapping - fresh export

email_thread

L1

Author / speaker	Recorded date
Nadia Haddad	2025-10-30T12:02:00
Position	Document
Message 1	emails/01_shelf-life-field-mapping

From: n.haddad@relexsolutions.example
To: ana.duarte@relexsolutions.example
Cc: priya.nair@acme-org.example, tomas.lindholm@relexsolutions.example, lena.fischer@acme-org.example

We proceeded. Nobody objected.

[Nadia Haddad](#)
Solution Consultant [Image removed by sender]
+46 76 552 30 18 | CET
n.haddad@relexsolutions.example
Vasagatan 7, 111 20 Stockholm
[Image removed by sender]
Plan better. Sell more. Waste less
relexsolutions.example

Source receipt for emails/01_shelf-life-field-mapping, message 1

A suggestion is not a commitment

The UAT sign-off thread, January 2025. The question "what did Acme accept?" has one correct scope, and the three messages give it three different flavours.

Ana Duarte, RELEX · 21 Jan 2025

proposed wording

"...completed satisfactorily for the 25 store pilot cohort, covering ambient and dry goods... does not extend to fresh categories... does not constitute acceptance of the overall programme."

1 emails/15_uat-signoff #message-003

Lena Fischer, Acme · 22 Jan 2025

agreed

"Signed and attached. The wording is what I asked for, thank you for being careful with it." The signatory and scope are established; the attachment itself is unavailable.

2 emails/15_uat-signoff #message-002

Kwame Boateng, RELEX · 24 Jan 2025

observation

Two rounding defects fixed and retested, nine of nine scenarios pass. A technical update. It does not widen what was signed.

3 emails/15_uat-signoff #message-001

How: extraction assigns each claim a kind (decision, commitment, question, correction, observation) and a status. Verification rejects a status the cited passage does not support. Answers name the speaker and whether anyone confirmed.

Stale and wrong are different problems

superseded

Decided, then changed

- Shelf-life population: 48% in September 2024 (Kwame, workshop turn 25), still 48% in December 2024, then 61% in December 2025 for the three in-scope categories only. Relio gives the trail and says the denominator changed.
- Bakery: listed in the fresh scope at the July 2024 kickoff, removed by Lena at the 17 November 2025 go/no-go, a separate Meridian workstream from January 2026.
- Marco's 40-store pilot figure: acknowledged wrong on 21 May 2025, the cut to 25 stores had happened in September 2024.

corrected

Wrong when it was written

- "Nightly article extract completed, no errors" appears in the weekly reports of 12 January, 9 February and 9 March 2026. On 16 June Nadia explains the job ran and the file landed, and the contents had never been checked; extract totals were lower than till numbers.
- The 9 February report itself records two DC-2 failures in January. The 3.4% Q1 waste figure was derived from those extract totals and needs review.
- Relio reports the status lines as dated assertions and attaches the June finding, rather than repeating "no errors" at face value.

How: claims are dated and carry a status. Chronology questions build an inventory across every document partition, then answer from the original records and neighbouring turns. Source dates and answer preparation time are shown separately, because a fresh timestamp does not make a fact current.

Erase a person, then keep answering

- An authorised L3 employee picks a person on the People page. Relio resolves the name, known aliases, addresses and identifying details from runtime evidence.
- Removed from the workspace's runtime source copies: authored records, identifying passages, recipient metadata and nearby pronoun references.
- Invalidated: every extracted claim and every generated answer, brief and card in the workspace. Affected mail, calendar and personal state is removed.
- Active inference is cancelled and the erasure epoch increments, so a generation started before the erasure cannot publish.
- Hashed suppression tokens reject reimport after restart. SQLite secure_delete is on and the write-ahead log is checkpointed.
- A receipt records the removal counts and time without repeating the identity.

Remove this person

Erase identifying material from this workspace's source copies and invalidate affected answers, briefs and other derived content.

Review erasure

Remove Ana Duarte from this workspace? This operation persists across restarts. Evidence independently supported by other people can remain.

Confirm person erasure

Worked example: erase Kwame Boateng

His September 2024 assessment (31%, 4,214 articles, 3,105 missing lead times) is gone, so the master-data question can no longer cite it. Priya's 1,300-fix estimate, her 780 of 1,317 progress report, Ana's 48% note and Lena's cut to 25 stores survive with their own dates. Ana, Priya and Lena stay named. Judges can ask about the removed person and about nearby facts afterwards.

This is removal from stored data, not a query-time filter. Boundaries: original challenge fixtures, other workspaces, downloaded copies and provider backups sit outside the mechanism.

Unasked: the agent in Cc speaks up when it matters

[Load an editable example](#) [1. Schedule the bakery review](#) [2. Continue the UAT discussion](#) Demo additions · 6 July 2026 · Helsinki

Mail

ana.duarte@relexsolutions.example Simulated Outlook

Inbox & sent 18

Re: Chain wide rollout - indicative numbers
marco.rossi@relexsolutions.example
Robert, Indicative numbers attached. Summary here so you have it in the mail. Chain wide
2026-06-26 17:15

Acme monthly steering report
marco.rossi@relexsolutions.example
Hi Robert, Lena, Monthly steering report, March 2026. Status: ☐ On Track First full qu
2026-04-10 17:40

Re: Q1 2026 waste reporting - fresh
lena.fischer@acme-org.example
This email originated from outside of RELEX. Be careful of attachments and links from unkn
2026-04-07 13:05

Weekly Acme update
ana.duarte@relexsolutions.example
Hi All, Hereby the weekly update. Update 06-04-2026 (Week 15) ☐ On Track Q1 fresh wast
2026-04-06 17:30


Your conversations become project memory
Select a thread or load an example. Messages addressed to the project agent enter the same memory used by your workspace.
[Try the bakery review example](#)

1 Cc acme-memory@relexsolutions.example

The delivered message joins shared project memory at its audience level. Without the agent in Cc it stays in the mailbox. Quoted history only enters if it is in the message.

2 It decides whether to reply

Ana writes that the January sign-off covered the whole programme including fresh. The agent answers in the same thread with the 25-store, ambient-and-dry scope and Lena's confirmation, from L1 evidence. Measured: 37.8 s.

3 It stays silent on routine mail

A scheduling update for the bakery review produces no reply, a new shared source and a calendar entry for 9 July. Measured: 37.4 s.

4 The brief keeps itself current

New delivered mail and changed personal goals queue a refreshed brief; a scheduler checks daily after 06:00 Helsinki.

Different readers, different correct answers

Level	Who	Sources available
L1	Nadia Haddad, solution consultant · partner firm · customer	Customer-shareable material
L2	Marco Rossi, account executive	L1 plus RELEX internal material
L3	Ana Duarte, project manager · may erase people	L1 and L2 plus restricted material

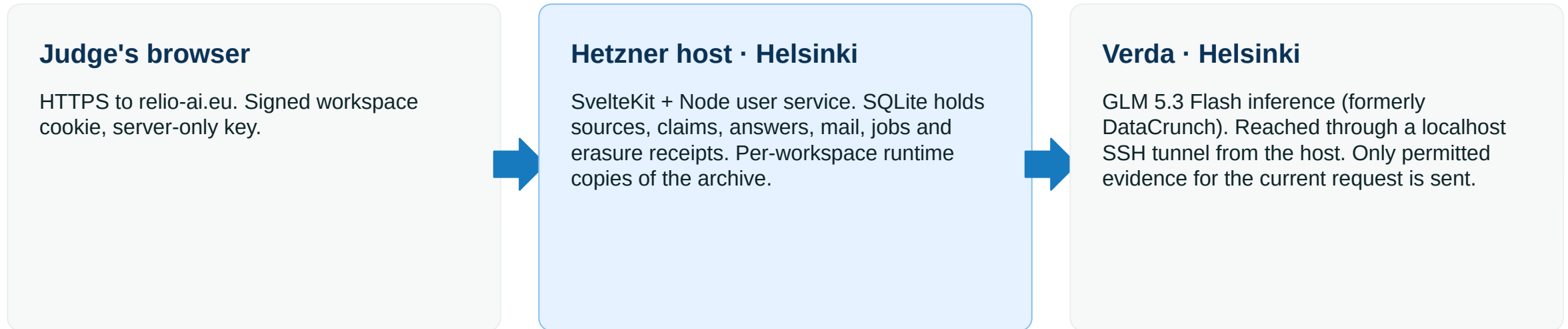
- The server selects evidence before building model input. Search, claims, mailboxes, sources and saved artifacts all apply the selected context.
- Customer evidence cards and externally addressed email are generated fresh from L1. Internal recipients get the least permitted clearance, capped at L2.
- Private memory stays with its owner: goals, preferences and chat history never reach another employee, a customer card or an outgoing email.
- "Switch view" stands in for the identity provider. Historical documents are classified from audiences and internal/restricted metadata; production needs an organisation-owned policy.

A fourth identity belongs to a different project and sees none of the Acme archive.

Share a private note with the project

An employee can publish a private note as an attributed shared source at a chosen level. Preview first; L1 publication is refused when the note holds higher-level evidence.

The archive stays in Helsinki



EU boundary: every component above sits in Helsinki, Finland. No embeddings service, no third-country API.

- Model calls have bounded concurrency and a timeout; jobs persist in SQLite and resume after a restart. Thinking is disabled by default.
- Simulated in this build: sign-in, email transport and calendar. Real in this build: ingestion, access filtering, inference, verification, erasure.
- Outside the application's deletion mechanism: the inference provider's own logging and backup retention. We state it rather than claim otherwise.

Measured limits, kept in the record

The practice questions were run through the live pipeline and reviewed against an independent, source-checked reference. These failures are preserved as findings; a later success does not erase them.

Exhaustive coverage

P1 (September master-data assessment) keeps the exact 30 September counts but still omits the workshop findings and adds unsupported qualifications. 84.9 s, six citations.

Earliest-event claims

P3 (operator ID removal) now attributes the November speaker correctly, misses the September 2024 agreement and wrongly calls November the earliest evidence. 74.8 s, 22 citations.

Follow-up completeness

P8 (uncompleted commitment) selects a real outstanding obligation, the missing location list for Sofia, and omits some later follow-ups.

Identity heuristics

Erasure of aliases, unique roles and nearby pronouns is conservative and can withdraw extra context. Arbitrary coreference needs a human look.

Semantic verification

Source binding is deterministic; interpretation is a model pass. A cited passage can still be misread, so every answer stays inspectable.

Latency

Direct questions take tens of seconds. Partitioned chronology research can take minutes; one P8 run took 327 s under concurrent load.

What has been checked on the deployed build

2,974

source records from all 45 documents, original IDs and positions kept

71

deterministic tests passing: boundaries, erasure, citation binding, stale-job rejection

13

Playwright workflows on the public URL, desktop and mobile, including accessibility

148

prepared memory entries across L1, L2 and L3, covering every document

- Live inference timings recorded: UAT answer 44.7 s, autonomous UAT correction 37.8 s, routine-mail silence with calendar entry 37.4 s, question about a new message 19.3 s, Finnish personal-goal save 9.4 s.
- Erasure verified across a public restart: preference and epoch persist, removed source 404, independently supported neighbour 200, reset opens a separate workspace.
- Independent adversarial review of the whole product, with reproductions and post-fix retests; immutable release directories and an atomic switch came out of it.
- Blind usability review by a reviewer who saw only the RELEX materials and the running application.
- Six persona/access briefs generated through the real pipeline (20 to 62 s each) so a fresh judge session shows cited briefs immediately.

What we will show

0:00

Ana's morning brief

Open a receipt, show the passage, the archive coverage date and the separate preparation time.

0:40

Cc the agent

Send the bakery review scheduling mail. Silence, a new shared source, a calendar entry. Ask what Ana promised to prepare.

2:15

It catches the mistake

Send the "whole programme" sign-off claim. The agent replies in the thread with the 25-store ambient-and-dry scope.

3:30

Boundary and export

Same decision from Nadia's L1 view; customer evidence card with receipts, copied or as PNG.

4:10

Erase Kwame

Confirm erasure, read the receipt, ask the sign-off question again. Ana and Lena's evidence answers it.

Generation runs live. If a step is still processing when the slot ends, we say so; the judge session is separate and can be inspected afterwards.



Memory you can challenge.

<https://relio-ai.eu>

Ask us about

Provenance storage · how a stale fact is caught · what erasure removes and what it does not · where inference runs · what the model still gets wrong

Toward production

Identity provider, real mailbox and calendar integration, organisation-owned access classification, broader alias evaluation, infrastructure retention controls, load testing

In the sidebar

Design (one page), How to use, How it works, Practice answers with exact source links