SmartMemory

WHITEPAPER

Memory for clinical trial operations

How a study team keeps what it learns across monitoring visits, protocol amendments and staff changes, and shows where every answer came from.

THE PROBLEM

Systems of record hold the trial data. The reasons behind it sit in visit reports, memos and people's heads, and they leave when people do.

THE IDEA

A memory layer for the study team. It keeps facts, events, procedures and decisions with their dates, their sources and the alternatives that were rejected.

IN THIS PAPER

How it works, a worked example from a fictional Phase II study, the deployment options, and what it does not do.

1 Knowledge that gets lost between visits

A Phase II study with 120 participants at four sites produces thousands of records. The EDC, the CTMS, the safety database and the trial master file hold them well. What they hold less well is the reasoning that connects them.

Why was screening paused at one site in March? Which deviations were logged late, and what changed afterwards? Why did the team keep the top dose after a serious liver event, and what else did it consider? The facts needed to answer each of these exist. They are spread across a visit report, a deviation log, a decision memo, an IRB letter and a follow-up visit, written by different people months apart.

Context is lost at three predictable moments.

- **Between monitoring visits.** A CRA is on site every six to eight weeks. What the coordinator said about a workaround, or what the CRA meant to check next time, lives in a follow-up letter and one person's head.
- **Across amendments.** Each amendment changes what is true, and not everywhere on the same day. A site with a monthly local IRB can run on an older consent form for weeks. "What applied at that site on that day" needs the history.
- **Through staff turnover.** CRAs, data managers and study leads change during a study. Each handover loses context the next person rebuilds from documents, where it can be rebuilt at all.

ICH E6(R3) asks for oversight that is proportionate to risk, and for essential records that allow the conduct of a trial and the quality of its data to be evaluated. Both depend on reconstructing why things happened, not only what. Today that work is manual, slow and done under time pressure.

2 What memory is

SmartMemory is a memory layer for teams and the AI tools they use. It sits beside the systems of record and holds what connects them: the decision memo and the alternatives turned down, the root cause behind a CAPA, the date a consent form changed at one site.

	SYSTEM OF RECORD (EDC, CTMS, eTMF, SAFETY)	CHAT ASSISTANT	MEMORY LAYER
Holds	Structured data and documents	Text in a conversation	Facts, events, procedures and decisions, with dates and sources
Answers	What is the value of this field?	Whatever sounds plausible	What happened, why, and on what evidence
Time	Current state, plus an audit trail	None	What was true, and what was known, on any past date
Conflicts	Field validation rules	Not detected	New records checked against stored ones, conflicts recorded for review
Sources	The record itself	Often none	Every answer lists the records it used

Not a database: questions are asked in plain language and answered from the relevant records. **Not a chatbot:** every answer is built from stored records and returns them as evidence. If the language model is unavailable, no answer is given, because an unsupported answer looks the same as a grounded one when it is read.

3 How it works in a trial

Kinds of record

The memory holds facts, events, procedures and decisions. Preferences are kept as facts about people.

KIND	EXAMPLES FROM A STUDY
Facts	Protocol versions and their key rules, sites and investigators, consent form version history, data management conventions, how the medical monitor wants hepatic events reported
Events	Monitoring visits, protocol deviations, SAEs, IRB approvals, DMC meetings, re-consents
Procedures	SAE reporting, deviation logging, each with its version and effective date
Decisions	Dose decisions, enrolment caps, site pauses, each with the alternatives rejected and why. Who recommended and who decided are written into the record

As-of questions

Every record carries two times: when it was true in the study, and when it was recorded. That allows two kinds of question: what applied on a given date ("which consent form was in force at site 03 on 3 March?"), and what the team knew on a given date ("what did we know about the liver signal on the day of the DMC meeting?"). When a protocol, a procedure or a decision is replaced, the earlier version is kept and marked as replaced.

CONSENT FORM IN FORCE, ASKED AS OF	ANSWER	WHY
2026-02-25, site 01	ICF 3.0	The central IRB approved it on 2026-02-20
2026-02-25, site 03	ICF 2.0	The site's local IRB had not yet approved 3.0
2026-03-03, site 03	ICF 3.0	Approved locally on 2026-03-02, so a consent signed on 2.0 that day was a deviation

Contradictions surface

When a new fact appears to conflict with stored ones, the conflict is recorded on it, with the records involved, for a person to review. Nothing is silently overwritten. For decisions, supporting and contradicting evidence is recorded against a decision as it arrives, and a decision can be formally replaced or withdrawn while the earlier version stays on record. In the example study, a site's endoscopic subscore that differs from the central read is such a conflict, and the convention that the central read governs is itself a stored fact.

An audit chain behind every answer

Each version of a record is kept with when it was made, and versions are linked by a hash chain, so a later change to that history is detectable. A record that must keep its history is changed by replacing it with a new version, never by editing it in place. On the Enterprise plan, workspace-wide verification against regular checkpoints also reports records removed since the last checkpoint. One call explains any record: where it came from, its versions, what it replaced and what replaced it, and the records and decisions linked to it. An answer to a question lists the records it used, so a reviewer can walk from the answer back to its sources.

Export. An administrator can export a whole workspace as a portable bundle, and any record's change history can be retrieved as structured data.

4 A worked example

Corvane Therapeutics runs Study KVR-417-201, a randomized, double-blind, placebo-controlled, dose-ranging Phase II study of KVR-417 in moderately to severely active ulcerative colitis. 120 participants at four US sites are randomized 1:1:1 to placebo, 75 mg or 200 mg once daily. Meridale Clinical Research is the CRO for monitoring, data management and medical monitoring. All of them are fictional.

The study's memory holds 60 records: the protocol and two amendments, four site initiation visits, three monitoring visit reports, six protocol deviations with their CAPAs, four data queries, two SAEs, six decisions, two procedures, the working preferences of three people, plus reference facts and dated events.

In January a participant at site 02 is admitted to hospital with ALT at 6.2 times the upper limit of normal. The DMC meets, the medical monitor recommends and the study lead decides to keep the top dose with weekly liver tests, and amendment 2 follows. The amendment then trips up the one site with a monthly local IRB: two participants sign consent on the superseded form, and screening there pauses for five weeks. Two deviations are logged late, and an SAE that happens on a Saturday is reported after 49 hours instead of 24. Each problem ends in a CAPA and a written decision. Enrolment falls behind and is recovered by August.

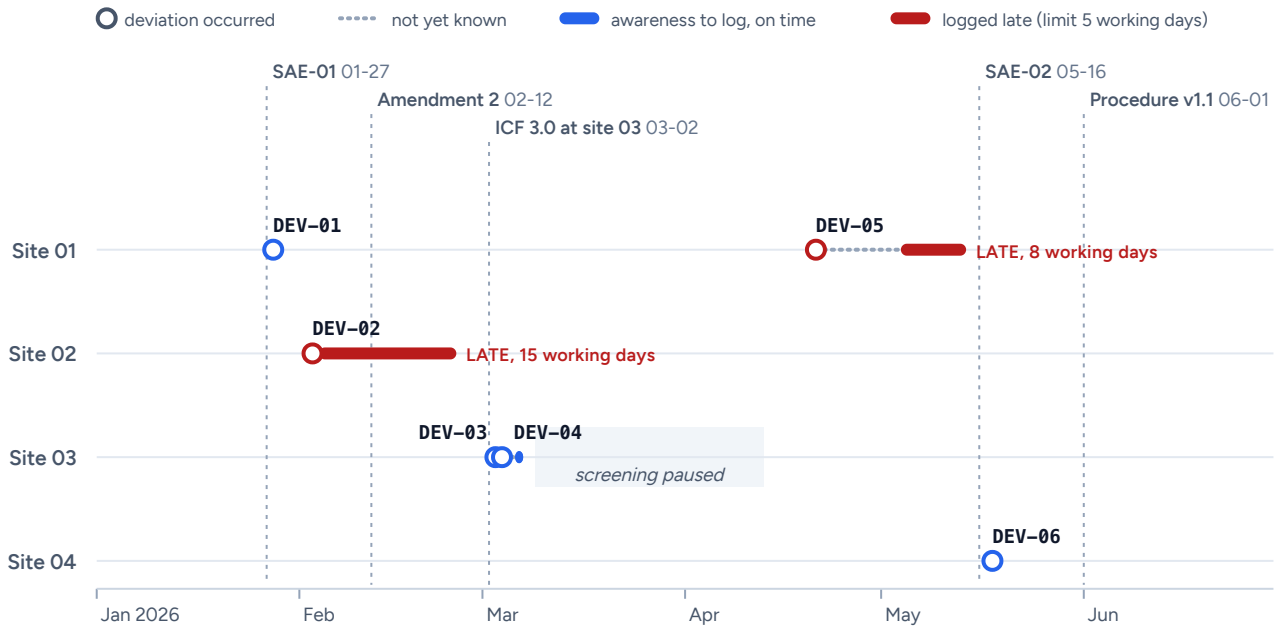


Figure 1. Protocol deviations at the four sites, January to June 2026. A circle marks when a deviation happened, and a bar runs from when the study team became aware of it to when it was logged. Blue is on time, red is late. DEV-02 and DEV-05 missed the 5-working-day limit and led to version 1.1 of the deviation logging procedure, effective 2026-06-01.

A decision, with what was rejected

The medical monitor asks for decision memos that list the alternatives considered. Stored that way, the question "why did we keep the top dose, and what else was on the table?" has a direct answer months later, after the people in the room have moved on.



Figure 2. A decision record: what the team knew, what it chose, who recommended and who decided, the three alternatives it rejected with the reason for each, and what followed. The ids are the example study's record ids.

Four of the twelve test questions

The example comes with twelve test questions, each with an expected answer and the records that ground it. An answer passes only if it states the key facts and cites at least one grounding record.

QUESTION	EXPECTED ANSWER, IN BRIEF	GROUNDED IN
Which deviations were logged late, and what changed afterwards?	DEV-02 (15 working days) and DEV-05 (8), against a limit of 5. Deviation logging procedure version 1.1 from 2026-06-01.	CT-REF-10 CT-DEV-02 CT-DEV-05 CT-DEC-05
Was the 24-hour SAE reporting clock met for both SAEs?	SAE-01 met, in 7 h 35 min. SAE-02 missed, at 49 h 15 min: the on-call sub-investigator had no eSAE account.	CT-SAE-01 CT-SAE-02 CT-DEV-06
What changed in amendment 2, and why?	Weekly liver tests to week 8, hepatic stopping rules, a tighter ALT exclusion and ICF 3.0. Prompted by SAE-01 and the DMC review.	CT-PROT-03 CT-EP-04
How does the medical monitor want hepatic events reported?	Phoned to her within 1 hour, with liver values as multiples of ULN in a table.	CT-PREF-01

An audit question, answered with its sources

The question an auditor or a new study lead asks is rarely about one record. It is about a chain of them. The memory answers from the records it retrieved and returns them with the answer. Each can be opened to see its date, origin and versions.

Q **Why did enrolment at site 03 pause in March, and what had to happen before it resumed?**

A New screening at site 03 was paused on 2026-03-09. Two participants had signed consent on ICF 2.0 on 2026-03-03 and 2026-03-04, after the local IRB approved ICF 3.0 on 2026-03-02. They were re-consented on ICF 3.0 on 2026-03-11 and allowed to finish screening. New screening resumed on 2026-04-13, once the CRA had checked the corrective action on site. Six participants at the site were already randomized.

SOURCES, IN ORDER

DATE	RECORD	WHAT IT SAYS	WHERE IT CAME FROM
2026-02-12	CT-PROT-03	Amendment 2 brings ICF 3.0 with liver risk language	Protocol v3.0, sponsor
2026-03-02	CT-EP-06	Site 03's local IRB approves ICF 3.0 ICF 2.0 superseded for new consents from this date	IRB approval letter to the PI
2026-03-03	CT-DEV-03	Participant 03-007 consents on ICF 2.0 Randomized 03-24 after re-consent. 03-008 consented 03-04 (DEV-04).	Deviation log, important
2026-03-06	CT-MVR-02	Monitoring visit finds DEV-03 and DEV-04 Escalated to the study lead and medical monitor that day	Visit report, Tomas Reyes, CRA
2026-03-09	CT-DEC-04	**Pause new screening at site 03** Rejected: close the site, or retrain and keep screening	Decision, Daniel Okoro, study lead
2026-04-08	CT-EP-08	Follow-up visit verifies the CAPA on site	Visit note, Tomas Reyes, CRA
2026-04-13	CT-EP-09	Pause lifted after 5 weeks	Study lead, Daniel Okoro

Figure 3. Illustration built from the example study's records: the answer, and the seven records behind it in date order with their origin. Any of them can be opened to see its versions, what it replaced and what it was derived from.

5 Deployment options

The same memory runs in three places. Most sponsors start with a cloud pilot on retrospective, coded data and decide on hosting after.

OPTION	WHERE THE DATA LIVES	SUITS
Cloud	SmartMemory's managed service. One workspace per study, access by workspace membership.	Pilots, and teams without their own hosting requirements
Your own AWS account	Being prepared with our first customer installs (Enterprise). The full service runs inside the sponsor's own AWS account. Ask us for current availability.	Sponsors that keep study data inside their own cloud
Lite	One computer. No server and no database to install.	One person trying the approach on non-confidential material

Answering a question uses a language model. On the cloud service you can supply your own model provider keys. Lite can run without any model key, and then returns labelled excerpts from the stored records instead of a composed answer.

6 What it is not

- **Not a clinical decision tool.** It records the decisions people make and the reasons they give. It does not assess causality, expectedness, eligibility or dose. Medical judgement stays with the investigator, the medical monitor and drug safety.
- **Not an EDC, CTMS, eTMF or safety database.** Those remain the systems of record. The memory holds the context around their records and points back to them.
- **Not a home for essential records.** A document that belongs in the trial master file belongs in the trial master file.
- **Not a place for direct identifiers.** Whatever the deployment, put in what trial records already contain: participant numbers, not names. The memory does not need direct identifiers to be useful.
- **Not a reporting system.** It can show that a reporting clock was missed and why. It does not send expedited reports or notify regulators.
- **No regulatory claims.** We make no claim about the regulatory status of SmartMemory. Whether and how to use it in a regulated process is for the sponsor's own quality and computerised system assessment.

7 Next steps

The fastest way to judge a memory layer is to ask it the questions your team already had to answer, on a study you already know.

1. **Pick one completed or late-stage study.** Gather its monitoring visit reports, deviation log, decision memos and IRB correspondence, with coded identifiers only.
2. **Write down ten to fifteen real questions** the team had to answer during that study, with the answer it reached and how long it took.
3. **Load the material into a dedicated workspace** and ask the same questions. Compare correctness, the sources cited and the time taken.
4. **Decide on the evidence.** If the answers hold up, plan a prospective pilot on one ongoing study and choose the deployment option.

We are happy to run the first three steps with your team, and to walk you through this example study live.

SmartMemory sales@smartmemory.ai | smartmemory.ai

Study KVR-417-201, KVR-417, Corvane Therapeutics, Meridale Clinical Research, and the sites, review boards and people in this paper are fictional. Any resemblance to real organisations or people is unintended. Product capabilities are described as of November 2026. This paper makes no regulatory claims.