

What Should Your AI Remember?

Consumer AI Adoption Gap | Eldris Inc | October 7, 2026

A decision guide to memory, personalization, and continued use

© 2026 Eldris Inc. All rights reserved.

Executive brief

AI memory makes a commercial promise: the assistant will become easier to use because the customer will not have to start over. Delivering that promise requires more than preserving information. The assistant must interpret what the information means, recognize when it applies, and respond correctly when the person changes it.

Remembering “I want inexpensive dinners this week” as a permanent preference is different from remembering “Please keep my weekly meal plans within this budget until I change it.” Remembering a guest's preferences as the account holder's preferences creates another problem. Both failures can occur even when the assistant accurately recalls the original words.

Product documentation already distinguishes memory, conversation history, and temporary conversations. Research benchmarks separately examine recall, updates, preference following, and context-sensitive application. Together, these sources establish that memory has several dimensions. They do not establish which memory design consumers will prefer or which will improve commercial retention. [1-6]

The business decision: for a recurring consumer task, choose how preferences become persistent, where they apply, and how people correct or remove them. Compare automatic memory, explicitly approved memory, and memory limited to that task.

The bounded question: in repeated weekly meal-planning sessions, which memory design reduces effort while preserving consumers' understanding and control when preferences change?

The population: US adults who use AI assistants and participate in planning household meals. Household role and prior experience with memory should be recorded. The scope is ordinary preferences, budgets, and preparation constraints, rather than clinical nutrition or medical advice.

Eldris recommends treating memory as a maintained agreement about future use. The product should make important retained preferences inspectable, distinguish durable preferences from temporary exceptions, and test whether corrections actually affect later sessions. That is a design recommendation to evaluate, not a finding that one universal memory setting wins.

What the evidence can establish

This report synthesizes primary product documentation and published research. It contains no original Eldris consumer study, measured retention uplift, or ranking of assistants. The examples and decision frameworks are Eldris analysis. The proposed study has not been conducted.

Evidence	Decision relevance	Evidence limit
Gemini memory documentation [1]	Past-chat use, correction, and deletion are distinct operations.	Documentation does not measure consumer comprehension or successful deletion behavior.
Claude memory documentation [2]	Inspectable entries, editing, pausing, and reset are product choices.	Documented controls do not prove that users discover or understand them.
Claude incognito documentation [3]	Exclusion from history and memory differs from service retention.	A temporary mode does not imply that no information is retained anywhere.
LongMemEval and PrefEval [4, 5]	Recall, updates, abstention, and preference following need separate evaluation.	Benchmark performance is not a measure of consumer acceptance or retention.
BenchPreS [6]	Remembered preferences can be inappropriate in another context.	Its communication tasks do not establish behavior in household meal planning.
Personalization-privacy literature review [7]	Personalization can offer benefits while raising concerns about data use.	Historical digital-marketing research cannot predict current AI memory choices.

Aguirre and colleagues' review describes personalization as capable of improving or weakening consumer engagement, depending on benefits and privacy concerns. It provides background for competing explanations, not a current estimate of consumers' willingness to enable AI memory. [7]

The implication for a product team is to separate technical memory performance from the consumer experience of being remembered. A system can improve recall without improving usefulness. A visible control can exist without enabling successful correction. Neither should be substituted for evidence about repeated use.

Memory is a set of product decisions

For this report, memory means information carried from prior interactions into later personalization. That information might be a structured preference, an inferred summary, or material retrieved from a previous conversation. These mechanisms need not have identical controls.

The product needs to resolve four questions: what to retain, whose information it is, when to use it, and how to revise it. A consumer-facing label such as “personalization” does not answer those questions by itself.

Consider four illustrative statements:

- “I prefer dinners that take less than 30 minutes.” This may be a recurring preference, but the product still needs to establish whether it should persist.
- “This week I have almost no time to cook.” This is a temporary condition with an implied time boundary.
- “My friend dislikes mushrooms.” This is information about someone else, not necessarily a preference of the account holder.
- “I used to avoid spicy food, but I want to try it now.” This updates the preference rather than adding a second equal rule.

A product that stores all four as permanent account-level instructions loses distinctions needed to produce useful future recommendations. Conversely, a product that remembers nothing requires the user to repeat potentially valuable context. The question is which distinctions matter enough to retain, expose, and test.

Memory is a lifecycle

Four decisions shape what the assistant carries into a later interaction.

Capture	What should persist?	Was it stated or inferred?
Scope	Whose preference is it?	Where and when does it apply?
Apply	Is it useful for this task?	Does current context override it?
Revise	Can the user correct it?	Does the next session change?

Useful memory requires the full lifecycle, not recall alone.

Figure 1. Memory is a lifecycle. Eldris conceptual framework. Capture, scope, application, and revision require separate decisions. The figure represents no measured consumer outcomes.

A useful retained preference should identify whose information it is, its source, its task, and its time boundary. Users need a plain-language description of consequential preferences, not a database schema. Test selective explanations against broader disclosure: showing every inference may add effort, while hiding them may prevent correction.

Remembering is not the same as applying correctly

LongMemEval evaluates information extraction, reasoning across sessions, time, knowledge updates, and abstention. PrefEval examines whether models infer, remember, and follow explicit and implicit preferences. These research programs demonstrate that recall and successful personalization warrant more than a single test. [4, 5]

BenchPreS addresses a further distinction: whether a stored preference should be applied in a particular communication context. Its reported results identify over-application as a problem in the tested systems and tasks. This is a March 2026 preprint; its findings are not a benchmark of every deployed assistant or of the meal-planning workflow used here. [6]

Eldris proposes evaluating three outcomes independently. **Recall** asks whether the assistant can retrieve the relevant prior information. **Application** asks whether using it improves the current task. **Revision** asks whether a correction changes later behavior without damaging unrelated preferences.

Illustrative failure	What happened	What to test
Forgotten preference	The plan repeatedly requires equipment the user does not own.	Whether the stated constraint is retained and used when relevant.
Over-applied preference	A temporary low budget limits future plans after the user changes it.	Expiry, update handling, and precedence of current instructions.
Wrong-person preference	A visiting friend's food preference shapes the account holder's later plans.	Person attribution and task scope.
Incorrect inference	A one-time request becomes a presumed enduring preference.	Confirmation, correction discovery, and subsequent behavior.
Failed correction	The interface acknowledges an edit, but later plans use the old preference.	Effective memory state and downstream retrieval after the edit.

These examples are diagnostic cases, not estimates of how often users experience them. A consumer panel can establish responses to deliberately presented errors. A production study is needed to estimate natural incidence in a particular product.

Product teams should avoid optimizing only for how much the assistant remembers. A higher application rate can include more unwanted application. Measure helpful use of a preference and inappropriate use separately, with expected behavior defined for each task.

Make the controls match the promise

Current documentation illustrates why a single “memory off” switch cannot be assumed to settle every data question.

Google explains that a deleted Gemini chat may stop influencing personalization after a delay. Its instructions for information also available in a connected app require both deleting relevant chats and disconnecting that source. This makes the paths through which information can return part of the control design. [1]

Claude's current memory documentation distinguishes pause from reset and allows inspection, editing, and deletion of individual topics. It also states that deleting a conversation does not automatically remove memory entries generated from it. The same page includes legacy behavior, so the current and legacy experiences should not be conflated. [2]

Claude's incognito documentation states that these conversations do not enter chat history or memory and are not used for training, while still being retained for 30 days by default. That documented distinction matters when explaining what “temporary” means. It is not a claim about all assistants' retention policies. [3]

Different controls answer different questions

Describe the actual effect of each action in ordinary language.

Future use	Will this information personalize a later answer?
Stored memory	Has the saved preference been removed or changed?
Chat history	Can the original conversation still be found?
Model training	Can the interaction be used to improve models?
Service retention	What information remains, and for how long?

One switch does not necessarily answer all five questions.

ELDRIS / CONSUMER AI ADOPTION GAP

© Eldris Inc

Figure 2. Five distinct control questions. Eldris framework informed by product documentation [1-3]. Actual features and effects vary by service, account, and version; this is not a vendor feature comparison.

For a consumer product, the practical requirement is to explain the effect of the selected action in ordinary language. “Do not use this in future meal plans” and “remove the saved preference” describe different outcomes. Removing a saved preference may leave an original conversation available to another retrieval path. Excluding a conversation from personalization may still leave it subject to a service retention policy.

Memory use, history visibility, training use, and service retention should have accurately described effects. Avoid making the consumer infer those effects from the word “forget.” Verify behavior across the implemented paths, including retrieval from history, connected sources, and any refreshed summaries.

Compare three designs for one recurring task

The meal-planning workflow gives the product team a concrete decision. A user provides a few ordinary preferences, returns for another plan, changes one preference, and later requests a plan for a different occasion. The assistant must support continuity without silently broadening the original instruction.

Automatic memory

The assistant infers and retains potentially useful information during use. The proposed advantage is lower setup effort. The competing concern is that consumers must identify and repair inferences they never deliberately approved.

Evaluate when automatic capture saves repetition and when it produces unwanted persistence. Keep the relevant controls discoverable at the point where the preference affects a plan. This design is a candidate to test, not an established retention strategy.

Explicitly approved memory

The assistant proposes a specific retained preference and asks whether it should persist. The proposed advantage is clearer agreement about future use. The competing concern is repeated confirmation effort or approval without comprehension.

A useful comparison tests targeted approval of consequential preferences against automatic capture, rather than adding a confirmation to every sentence. Record whether the person understands the proposed scope and time boundary, not just whether they click “save.”

Memory limited to the task

Preferences persist within meal planning, while other tasks remain outside that memory scope. The proposed advantage is a narrower expectation about use. The competing concern is that the person may expect a useful preference to carry into another activity and experience its absence as inconsistency.

Task boundaries must correspond to something users can recognize. A technical distinction between sessions or storage systems is not enough if the experience feels like one continuing assistant.

Compare three memory designs

Competing benefits and burdens to measure in the same recurring task.

Automatic capture	Explicit approval	Limited to a task
Potential benefit: Less setup and repetition.	Potential benefit: Clearer agreement about use.	Potential benefit: A narrower expectation of use.
Question to test: Which inferences should persist?	Question to test: Is confirmation worth the effort?	Question to test: Is the boundary understandable?

Approval and scope can be combined. Define the exact packages being compared.

ELDRIS / CONSUMER AI ADOPTION GAP

© Eldris Inc

Figure 3. Compare convenience, approval, and scope. Eldris design alternatives. These are not exclusive dimensions: a product can combine explicit approval with memory limited to a task. A study must define the actual configurations it compares.

An initial study can compare three complete packages: automatic account-level capture, approved account-level capture, and automatic capture limited to meal planning. Such a comparison estimates the effect of each package. It does not isolate approval from scope. If a sponsor needs that attribution, use a factorial design or a second study varying one dimension at a time.

Measure behavior across repeated use

A credible proprietary extension would follow a returning panel through multiple genuine interactions. This report proposes the design; it does not report results or commit to a sample size without a power analysis and recruitment assessment.

Recruit US adults who help plan household meals and have used an AI assistant. Record their experience with persistent memory, planning responsibility, and relevant technology familiarity. A recruited online panel is not automatically representative of all US consumers; report its coverage and recruitment limits.

Use a functioning study assistant whose memory controls perform the advertised operations. Verify stored state and later behavior mechanically before exposing participants to the conditions. Do not use a nonfunctional control to imply that a memory was edited or deleted.

Randomly assign the initial design packages. Keep the model, task instructions, available meal options, and initial preference opportunities consistent where possible. Memory is expected to change later recommendations, so realized answer quality is an outcome to measure rather than something that can be assumed constant throughout.

The study needs distinct stages:

- **Establish a preference:** observe capture or approval, then test whether the participant understands what will persist.
- **Return for another plan:** measure repetition avoided, task effort, and the appropriateness of the remembered constraint.
- **Change or correct information:** observe discovery and use of controls, then verify the effect in a later interaction.
- **Change the occasion:** ask for a different type of plan and test whether the prior preference is appropriately applied, clarified, or suspended.
- **Choose whether to continue:** provide a real choice about future memory use, with compensation independent of enabling memory.

Scheduled, paid return sessions measure experience under study conditions. They do not establish voluntary product retention. A separate optional return opportunity can provide a behavioral indicator, but its incentives and observation window must be reported. Long-term retention claims require an appropriately long observation period and relevant usage evidence.

To isolate the effect of visibility or control from answer usefulness, a narrower experiment can give participants the same recommendations while varying the explanation and available control. That answers a different question from testing the complete memory-enabled product. State which comparison supports each finding.

Turn the findings into product choices

The study should produce evidence that changes a design choice, not a generic percentage saying consumers “trust AI memory.” Eldris proposes the following decision mapping.

Observed result	Product response to evaluate	Remaining question
Memory saves repetition and is correctly understood.	Retain the useful capture design for that task.	Does the benefit persist outside the study?
Useful personalization creates unwanted durable inferences.	Add selective approval or narrower scope.	Which preferences merit that extra step?
Users find controls but cannot repair later behavior.	Fix the update and retrieval paths before expanding memory.	Does the repair hold across sessions and source refreshes?
Approval adds effort without improving understanding.	Simplify the confirmation or reduce its frequency.	Can a different explanation make approval meaningful?
Task boundaries prevent unwanted use but confuse consumers.	Make the boundary clearer or offer deliberate transfer.	Where do consumers expect continuity?

Track accepted plans, time and effort, preference corrections, unwanted application, successful deletion or suspension, comprehension, and continued memory choices separately. Report counts and denominators for each, including how many participants encountered the relevant event.

Do not interpret unchanged settings as informed acceptance. A person may not have noticed the feature or understood the alternatives. Likewise, disabling memory may be a rational choice about one task rather than rejection of the assistant. Ask short follow-up questions after observing the behavior to distinguish plausible explanations.

Attrition can also change the apparent result. Report participants who did not return and compare available baseline characteristics across groups. Analyze initial assignment as well as actual exposure, rather than reporting only people who completed every stage. Prespecify the main decision outcomes and distinguish subsequent exploratory findings.

A product team can use these results to choose a capture design, identify preferences requiring confirmation, simplify correction, and define the scope of its commercial promise. Commercial uplift remains a separate empirical question; neither memory accuracy nor stated enthusiasm establishes subscription value.

The decision record to require

Before expanding consumer memory, require a concise record of the task, population, selected design, and evidence supporting it. That record should show what the product retains, where it applies, what each control changes, and whether changes propagate to later interactions.

Require separate results for useful personalization and unwanted application. Include the effort required to approve and maintain preferences, successful correction behavior, and the limits of any continued-use measure. State what remains untested, including broader tasks, shared accounts, and longer-term behavior where relevant.

Eldris recommends expanding memory when its demonstrated convenience survives preference changes and consumers can understand and change consequential retained information. If users cannot repair persistent errors, improve the revision process. If the assistant applies information outside its intended context, revisit scope. If memory provides little benefit for a task, continuing without it remains a valid product choice.

The commercial promise should be that the assistant carries forward useful context on terms the consumer can understand and change. Whether a particular design delivers that promise must be established through functioning controls and repeated behavior, not inferred from the existence of a memory feature.

Sources

Sources reviewed October 7, 2026. Product documentation establishes described controls, not independently verified outcomes. Benchmark findings apply to the tested systems, datasets, and conditions. All figures and decision frameworks are Eldris analysis.

1. Google Gemini Apps Help. [Get personalization with memory of your past Gemini chats](#). Past-chat use, correction, deletion delays, and connected-source interactions.
2. Claude Help Center. [Use Claude's chat search and memory to build on previous context](#). Current memory controls, topic editing, and distinction from legacy behavior.
3. Claude Help Center. [Use incognito chats](#). History, memory, training, and retention distinctions.
4. Wu, D., Wang, H., Yu, W., Zhang, Y., Chang, K.-W., and Yu, D. [LongMemEval: Benchmarking Chat Assistants on Long-Term Interactive Memory](#). ICLR 2025; revised March 4, 2025. Evaluation of multiple memory abilities.
5. Zhao, S., Hong, M., Liu, Y., Hazarika, D., and Lin, K. [Do LLMs Recognize Your Preferences? Evaluating Personalized Preference Following in LLMs](#). ICLR 2025; submitted February 13, 2025. PrefEval benchmark for explicit and implicit preference following.
6. Yoon, S., Kim, S., Hong, H., Jeung, W., Kim, Y., Seo, W., Yeen, H., and No, A. [BenchPreS: A Benchmark for Context-Aware Personalized Preference Selectivity of Persistent-Memory LLMs](#). Preprint submitted March 17, 2026. Context-sensitive preference application in communication tasks.
7. Aguirre, E., Roggeveen, A. L., Grewal, D., and Wetzels, M. [The personalization-privacy paradox: implications for new media](#). Journal of Consumer Marketing, 2016. Literature review providing historical context rather than current AI adoption evidence.