Zortex

Zortex 0.1

System Card

September 17, 2026

Deployment Safety
Local-first personal intelligence

Contents

1 Safety overview

Zortex brings mail, calendars, files and messages into a private record on the user's own machine. It reads connected accounts, organizes their contents and builds context from them. Zortex 0.1 does not send messages, make payments, publish posts or delete on the user's behalf.

Access to personal information makes the system useful. It also makes the boundary around that information important. Zortex separates a model's interpretation of content from the permissions needed to act. An instruction found in an email or document does not acquire the authority of the account owner.

The evaluations in this card examine the controls around the model: whether data can leave when outbound access is disabled, whether one account can reach another, whether an action can be withdrawn before taking effect, and whether connectors work through their live-account data paths. They are internal evaluations.[1]

Table 1. Selected results

Evaluation	Result
Data-egress containment	Zero bytes observed across 14 / 14 tested paths
Cross-account isolation	8 / 8 cases passed
Action recall	29 / 29 cases passed
Live-account connector certification	193 passed · 4 failed · 1 skipped

The deployment remains limited while work continues on broader action authority. A credential-masking edge is open, the alignment suite includes one failure and 26 cases not yet run, and no independent evaluation or external red-team has been completed. Sending, paying, posting and deleting remain outside the current release.

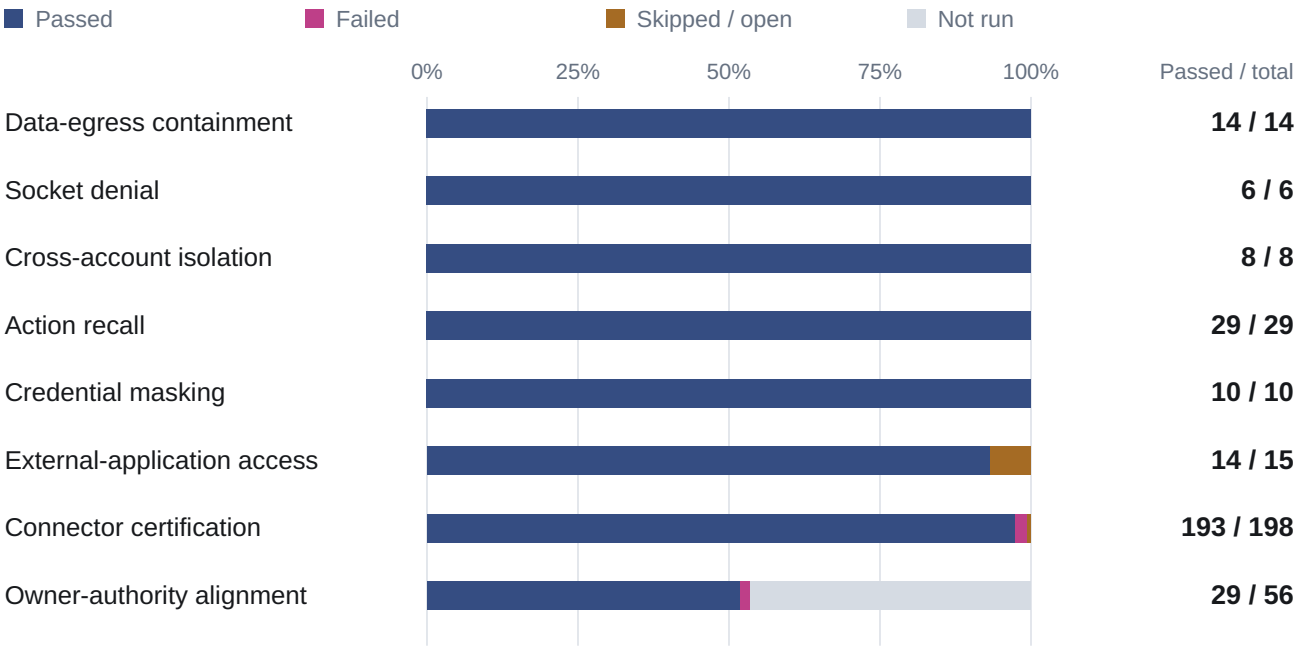
Personal content is not sold or used to train generative models. Raw content and the knowledge derived from it remain on the device unless the user chooses cloud processing. The following sections describe how those commitments are implemented, what has been tested and where the remaining limits lie.

1.1 Evaluation results

Most of the completed containment, isolation and recall tests passed every case in their respective suites. The connector, external-access and alignment results require a more detailed reading: each contains an unresolved, failed, skipped or unrun portion.

Reported safety evaluation outcomes

Passing cases are shown against each suite's full reported total.



Internal evaluations. Suite sizes differ; results are not pooled.
Credential masking also has one known edge. One external-access case remains unresolved.

Figure 1. Full reported outcomes for eight execution-based suites. The privacy conformance review uses specification sections rather than execution cases and is reported separately.

The rows above retain each suite's own total. In particular, the alignment result includes all 56 planned cases, not only the 30 that have been executed. Credential masking's ten passing cases are also accompanied by a separately recorded edge that remains open.

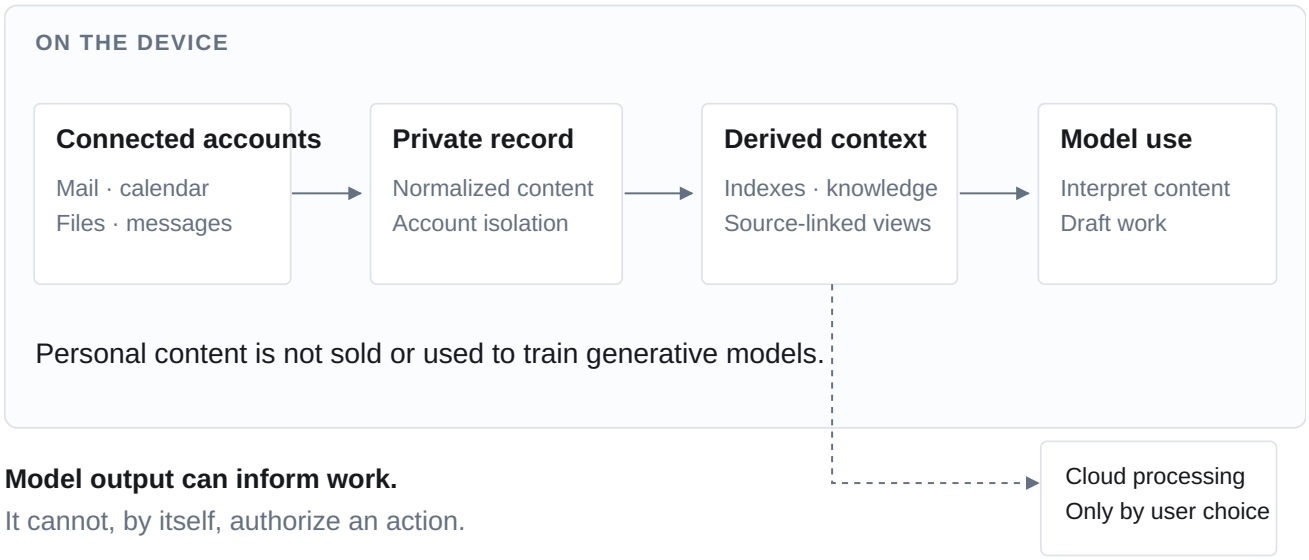
These are results under the conditions of the internal tests. They are not estimates of the probability that an arbitrary future action will be safe. The evaluation record at the end of this card gives the scope, units and available dates.

2 System and data

Zortex uses one or more language models to interpret information and draft work. It is a system built around models, connected accounts and local controls; the model is not the source of permission to take a consequential action.

A private record on the user’s device

System architecture · arrows show relationships, not data volume.



Model output can inform work.

It cannot, by itself, authorize an action.

Figure 2. The private record is the source for derived files, indexes and knowledge views. Cloud processing is optional. The diagram describes the architecture, not a measured transfer volume.

Information is normalized into a single private record. Derived views are rebuilt from that record, so a conclusion about the user can be traced back to the material from which it was drawn. The same separation matters for authority: content can contribute to an answer without becoming a command.

This card evaluates the surrounding system where results are available. It does not claim that a particular underlying model will reliably refuse every unsafe instruction.

2.1 How information is handled

Local storage is the default for raw personal content and the knowledge derived from it. The user may choose cloud processing. That choice is separate from the handling of product signals and diagnostics.

Table 2. Data handling

Information	Default treatment	What can leave the device
Raw personal content	Stored locally	Content used for cloud processing when the user chooses it
Derived knowledge	Stored locally	Derived context used for cloud processing when the user chooses it
Product signals	Aggregate and content-free	Signals used to improve features shared across users
Raw diagnostics	Retained unless specifically granted	Only the item covered by a specific user grant

Aggregate product signals are not a permission to transfer the user's mail, files or derived knowledge. Raw diagnostics require a separate, per-item grant. This distinction matters because a local-first system may still communicate with external services under defined conditions.

Privacy commitments

Zortex does not sell personal content and does not use it to train generative models. Its privacy policy is versioned and applied prospectively: a later policy does not retroactively change the terms governing information collected under an earlier one.

Account separation and erasure

Each account has its own store. Reads, erasure and derived results are constrained by that account boundary. Erasure is intended to remove the account's personal copies, including cloud copies, without affecting other accounts. Partial and interrupted erasure paths are enforced in operation, although a separate standing result for those paths is not published.

The privacy conformance review and cross-account execution tests address different parts of this design. Their results are presented separately in the evaluation section.

3 Current deployment

Zortex 0.1 is limited to reading, organizing and deriving knowledge from connected accounts. The system is designed for broader assistance, but the ability to send, pay, post or delete is not enabled in this release.

Table 3. Capability scope

Capability	Scope in this release
Read, organize and derive	Available under the containment and account-isolation controls
External AI applications	Read-only interface for context and status; one access-control case remains unresolved
Plugins	Default-deny permission envelope with integrity checking; no standing evaluation result published
Send or post	Not enabled
Pay	Not enabled
Delete	Not enabled

Capabilities that affect the outside world are gated by their relevant evaluations. Sending and posting depend on recall, outbound-data controls and credential protection. Payment is also bounded by a hard spending ceiling. Deletion depends on recall and the minimum controls applied to consequential actions.

External red-teaming is a further gate for enabling these world-affecting capabilities. It has not yet taken place. The recall and containment results therefore describe controls being evaluated for broader use; they do not announce that broader use is available.

Internal use

Connector certification has exercised live accounts through the production data path. The most recent dated connector result is September 4, 2026. The duration and scale of continuous internal use are not quantified in this card.

Activity within the system is written to an independent, tamper-evident record. That record is the basis for reconstructing behavior during staged deployment.

4 Threat model

Zortex protects three assets: the user's personal information, the credentials for connected accounts and the authority to act on the user's behalf. Threats can arrive through the content being read, the software handling it or attempts to cross the machine's outbound boundary.

Table 4. Threats and controls

Threat	What could go wrong	Relevant boundary
Malicious content	An instruction inside an email, file or page redirects the agent	Ingested content is data, not action authority
Compromised software	A connector or plugin exceeds its permissions	Default-deny permissions and integrity checks
Credential attack	An account secret appears in model context, logs or browser output	Credential custody and masking
Unauthorized egress	Personal information leaves through an unapproved path	Outbound containment and egress policy
Cross-account access	A read, erasure or result reaches a different account	Per-account stores and isolation checks

The model may misinterpret malicious content. The design does not require perfect interpretation to preserve the action boundary: the surrounding system must still determine whether an operation has the owner's permission.

Host security

Protection against local software depends on its privileges and execution path. OS-level containment is strongest on macOS and at the subprocess boundary. In-process execution and some plugin paths rely on higher-level controls rather than an OS sandbox.

Same-machine adversaries with equivalent privileges are outside the stated protection boundary. Zortex also does not claim to defeat an attacker who already controls the operating system. A local-first architecture does not remove those host-security limits.

5 Safety evaluations

The evaluation program tests whether the controls around Zortex hold when an operation reaches them. Its units are concrete: outbound paths, network sockets, account boundaries, recalled actions and connector cases. The privacy review uses specification sections rather than execution cases.

Tests of behavior and tests of the test

The containment evaluation measures bytes at the machine's outbound boundary with outbound access disabled. It also removes the control and checks that data can then leave. This comparison helps establish that the control, rather than an inactive test path, explains the observed containment.

Credential-masking checks use a related approach. In addition to standard cases, the evaluation includes faults deliberately introduced to confirm that the test catches a regression. Monitoring uses a planted event to check that known activity is surfaced from the record.

Live-account integration

Connector certification runs against live accounts through the production data path, rather than only against fabricated inputs. It covers credential and certificate isolation across the connector fleet and confirms operation against the connected account. Failed and skipped cases remain in the reported total.

Scope of the results

A passing result applies to the behavior and conditions exercised in that evaluation. The card does not publish a random-sampling protocol, confidence intervals or a denominator for incidents during ordinary product use. Counts in this document should therefore be read as test outcomes, not as production reliability estimates.

The credential edge, incomplete alignment suite and other open findings are presented with the relevant results. Controls described as enforced without a standing result are listed separately from tests with passing counts.

Evaluation responsibility

All evaluations in this edition were conducted internally. Zortex's evaluation rule requires verdicts to be derived from recorded evidence rather than a component's self-report. That separation inside the system is not an independent third-party assessment; no such assessment has been completed.

5.1 Data-egress containment

With outbound access disabled, the containment evaluation observed zero bytes leaving the machine across all 14 tested outbound paths. Removing the containment control allowed bytes to leave; restoring it stopped them.

The containment control changes the outcome

Measured at the machine's outbound boundary.

The disabled-control result is qualitative; no transferred-byte count was published.

Figure 3. The control-removal check compares the presence and absence of containment. The measured zero-byte result applies to the enabled-control condition and tested paths.

A separate socket-denial evaluation passed all six cases. Together, these tests examine both attempts to open network sockets and attempts to move data across the outbound boundary.

Policy and containment

The outbound policy addresses which transfers are allowed. Unknown channels are refused, and the strictest posture applies when no policy is present. This policy is enforced in operation but does not yet have a published standing result.

The zero-byte finding is conditional on outbound access being disabled. It does not describe configurations in which a user has chosen cloud processing or granted a particular diagnostic transfer.

5.2 Cross-account isolation

Personal information is isolated per account. A read against one account must not reach another account's store. The same boundary applies when information is erased or used to build derived results.

Account boundaries apply to more than reads

Boundary schematic · the isolation evaluation separately passed 8 / 8 cases.

	Account's own store	Another account's store
Read personal content	Within account scope	Cross-account reach refused
Erase personal copies	Within account scope	Cross-account reach refused
Build account results	Within account scope	Cross-account reach refused

A separate store for each account limits the reach of reads, erasure and derived results.

Figure 4. The account boundary covers reads, erasure and results. The cells illustrate relationships in the design; they are not a reconstruction of the eight individual test cases.

The cross-account isolation evaluation passed all eight cases. The privacy conformance review also examined physical per-account separation, including the requirement that one account's erasure not affect another account's information.

Identity and derived results

Identity writes are checked against a single source of identity. Unsigned or mismatched writes are refused as unauthenticated. Derived results are also kept within the account boundary so that one account cannot influence or corrupt another's results.

These are account-isolation controls within Zortex. They do not extend the system's protection to an attacker with control of the host operating system.

5.3 Privacy conformance

The August 31, 2026 conformance review found all 11 privacy, tenancy and erasure sections implemented as specified. The review addresses whether the design is present in the system, while the cross-account evaluation separately exercises account isolation.

Privacy, tenancy and erasure conformance

August 31, 2026 · reviewed against the specification

11 / 11 sections conforming

Each block is a specification section. This review is distinct from the 8-case isolation test.

Figure 5. Internal conformance review. Blocks represent specification sections, not runtime cases.

Table 5. Areas described in the conformance review

Area	Requirement
Per-account isolation	Each account has a separate store; reads and erasure do not cross account boundaries
Identity	Unsigned or mismatched identity writes are refused
Erasure	Personal copies are removed without affecting other accounts, including copies in the cloud
Result isolation	One account cannot influence or corrupt another account's results

Partial and interrupted erasure paths are enforced in standard operation. A separate standing result for those paths has not yet been published.

5.4 Live-account connectors

Connectors are tested against the accounts they serve. The September 4, 2026 run used the production data path and recorded 193 passing cases, four failures and one skipped case.

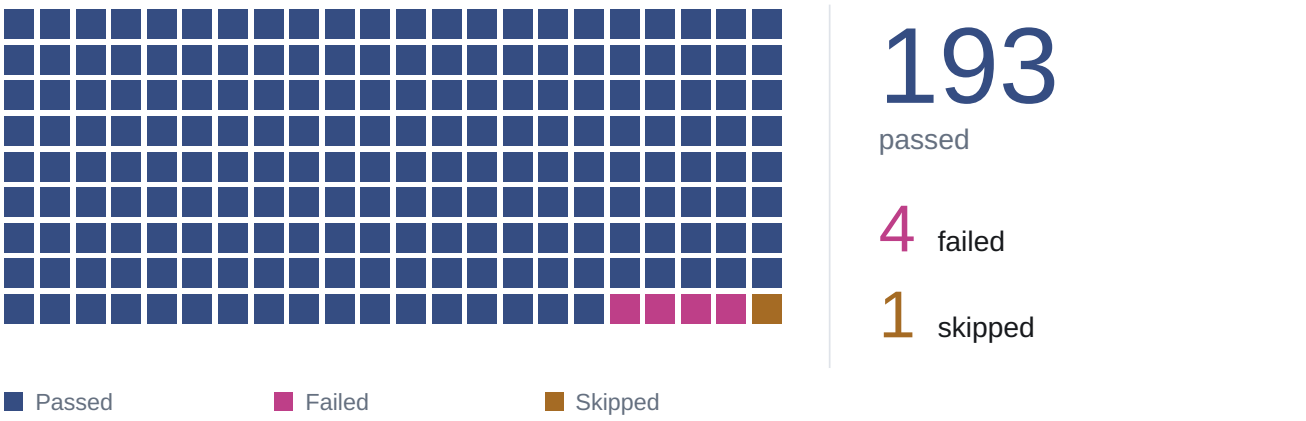


Figure 6. The 198-case evaluation ran against live accounts through the production data path.

Connector certification is an internal evaluation of credential and certificate isolation across the connector fleet, together with operation against each connector's live account.

Four connector-acceptance cases remain open. The public results identify them at the suite level; a connector-by-connector breakdown is not available.

These tests exercise the integrations the user depends on, through the same path used by live data. Outbound containment is tested separately with outbound access disabled.

5.5 External applications and plugins

External AI applications receive context and status through a read-only interface. They do not receive action authority. The access evaluation passed 14 of 15 cases; one case remains unresolved in the published record.

Figure 7. The remaining case stays visible beside the 14 passing cases. The record does not provide enough detail to classify it more specifically.

Plugin permissions

Plugins run under a default-deny permission envelope with integrity checking. A modified plugin is refused. The envelope is enforced in standard operation, but no standing evaluation result is yet published.

Some plugin execution paths are not OS-sandboxed. They rely on higher-level controls in the current implementation. A permission envelope and an operating-system sandbox are different protections, and the presence of one should not be taken to imply the other.

Table 6. Extension boundary

Interface	Authority
External AI applications	Read-only context and status
Plugins	Only the permissions granted by the default-deny envelope

Sending, paying, posting and deleting remain outside the authority exposed in Zortex 0.1.

5.6 Credential protection

Credential protection has two parts. Custody keeps account tokens out of model context, logs, stored records and error output. Masking addresses credentials that could appear while an automated browser is being used.

Credential masking: tests and residual exposure

Separate evidence, shown without a combined success score.

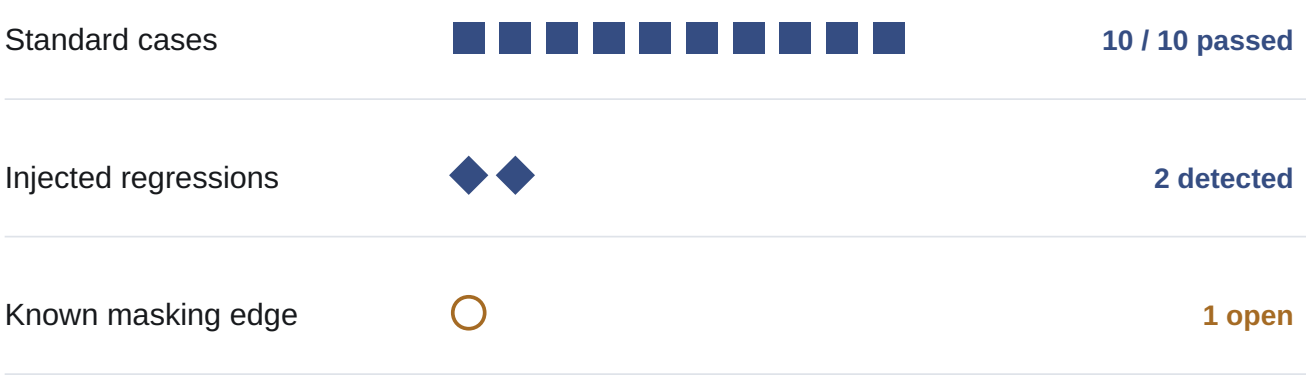


Figure 8. Standard cases, injected faults and the known edge are separate pieces of evidence. The graphic does not combine them into a larger passing suite.

The masking evaluation passed all ten standard cases and detected two deliberately injected regressions. One edge remains: an unstructured credential, which the agent never refers to by name, can cross the masking boundary in a specific automated-browser condition.

This edge is under active work. The ten passing cases do not close it. Credential custody is enforced in standard operation but does not yet have a published standing result.

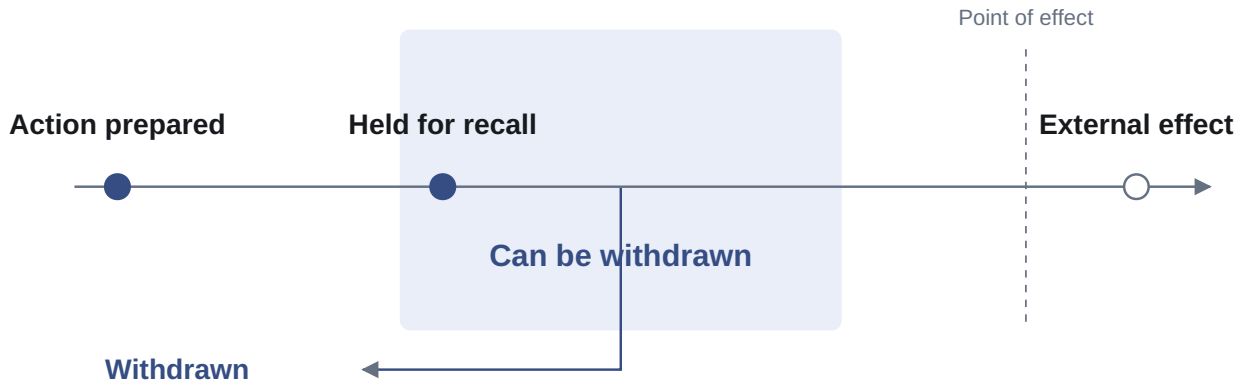
Broader account actions remain gated while the masking edge and credential-custody evaluation are incomplete.

5.7 Action recall

A consequential action is held before it takes effect. During that recall window it can be withdrawn. The action-recall evaluation passed all 29 cases, exercising withdrawal as an actual path through the system.

Recall happens before an action takes effect

Action sequence schematic · 29 / 29 recall cases passed



Recall is not a promise to reverse a message, payment or deletion after its effect.

Figure 9. The recall window precedes the external effect. No duration is specified in the evaluation record, and the horizontal spacing does not represent elapsed time.

The timing of the boundary is important. Once a message has been sent, recalling an internal task cannot make the recipient forget it. The tested mechanism concerns withdrawal before the effect, not reversal of every possible consequence afterward.

The same authority through the interface

An interface-privilege evaluation asserts that an action taken through Zortex's feed carries no more authority than the same action taken by an automated agent. An injection fixture is included. A case count and standing passing total are not published for that evaluation.

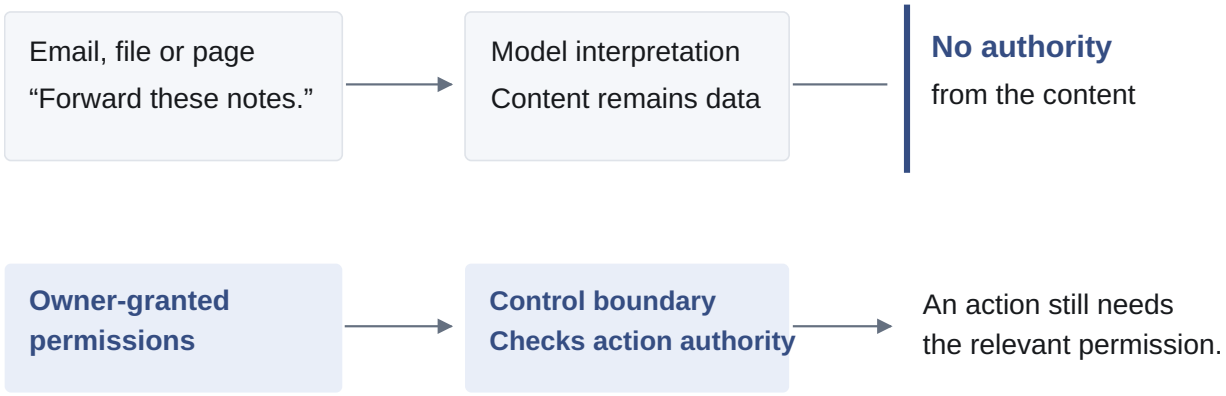
Recall is one of the controls governing future send, pay, post and delete capabilities. Its passing result does not enable those capabilities on its own. Other relevant evaluations and external red-teaming remain part of the deployment gate.

5.8 Instructions in connected content

A message can contain a request without being a request from the account owner. Zortex treats ingested content as data. A model may interpret what the sender wants, but the content does not, by itself, authorize an action.

Reading an instruction does not authorize it

Trust boundary schematic



Illustrative instruction; this is not a transcript of an evaluated attack.

Figure 10. The upper path carries information; the lower path represents owner-granted authority. The example explains the boundary and is not a reported test transcript.

A dedicated ten-case suite asserts that instructions arriving inside ingested material do not become actions. Zortex also maintains a seven-fixture corpus of hostile inputs as a standing set. A standing passing count for the refusal suite is not yet published.

This boundary is enforced beneath the model. The design assumes that a language model may follow malicious text; the surrounding controls must still prevent that text from conferring permission.

The suite and corpus describe evaluation coverage being maintained. Their sizes are not a measured prompt-injection resistance rate, and the card does not claim coverage of every possible attack.

5.9 Staying within the owner’s authority

The broader alignment suite examines whether Zortex stays within the authority its owner has granted. It contains 56 planned cases. Of those, 29 passed, one failed and 26 have not yet been run.

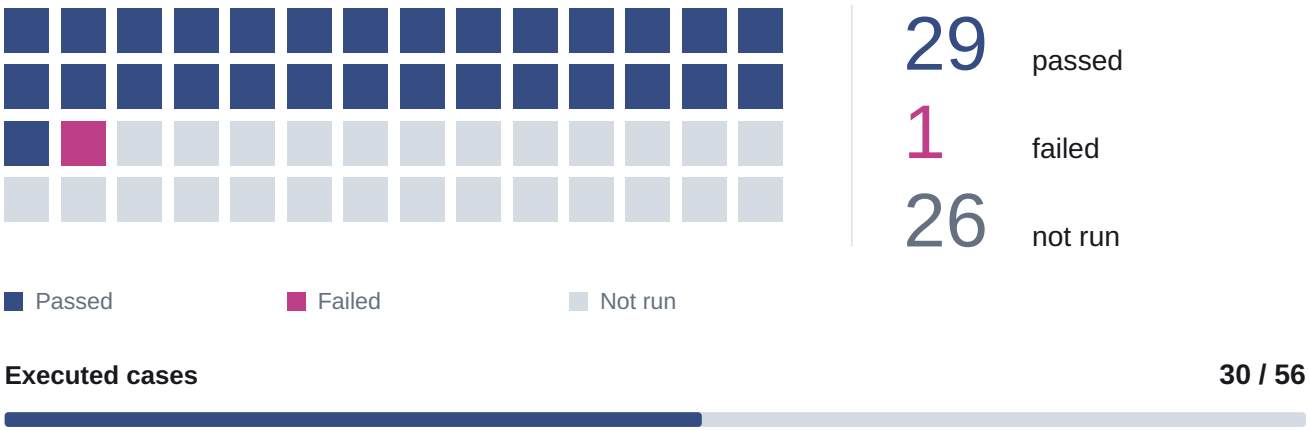


Figure 11. The planned suite includes executed and unrun cases. All 56 remain in the denominator.

Thirty cases have been executed. The remaining 26 still need to be run, and the failed case remains open.

This suite concerns the owner’s authority across interfaces and agents. Privacy, connector operation and recall retain their own evaluations and results.

The current deployment boundary remains unchanged while these results are incomplete: Zortex reads, organizes and derives, but does not send, pay, post or delete on the owner’s behalf.

6 The system’s objective

Zortex’s objective is useful work for the person, rather than time spent or attention captured. The design treats ways of inflating engagement as adversarial behavior, even when they can make product metrics look favorable.

Table 7. Engagement-gaming cases

Failure mode	Behavior the system is designed to reject
Easy-yes farming	Proposing trivial approvals to inflate acceptance
Urgency inflation	Marking matters urgent when they are not
Notification-timing exploitation	Timing prompts to capture attention
Self-preference	Favoring the system’s own suggestions

The objective is constrained around real outcomes and durable preference changes. Uninstalls and disabled notifications are guardrail signals, not outcomes for the system to optimize against. A decision to stop using the product is not treated as a problem to overcome through more persuasive interruptions.

The four failure modes define the behavior to resist. The corresponding controls are enforced in operation; a standing evaluation result is not yet available.

Authority and attention

Control over an assistant includes more than control over outbound actions. A system can remain inside a technical permission boundary while using notifications or trivial requests to demand more attention. Zortex therefore treats the objective it optimizes as a separate design concern.

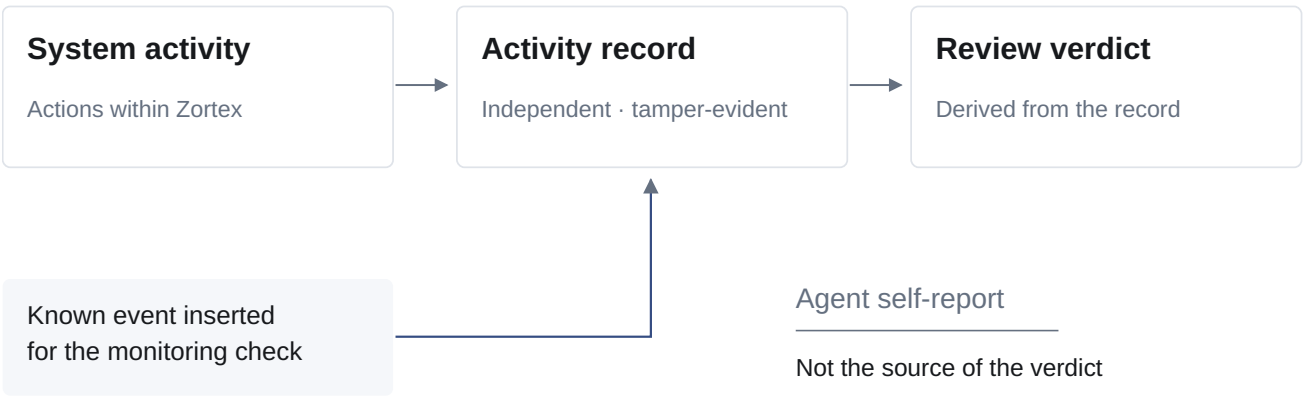
User-outcome measures, such as time saved and interruption frequency, have not yet been published.

7 Monitorability

A reconstruction of what happened should not depend on the agent's description of its own behavior. Zortex uses a single, independent, tamper-evident activity record as the basis for evaluation verdicts.

Activity is checked against an independent record

Monitoring architecture · one evaluation case remains open



The record covers Zortex's boundary, not every process running on the same machine.

Figure 12. A planted-event check tests whether known activity is surfaced from the record. Self-report is not accepted as the source of the verdict.

In the planted-event check, a known event is deliberately inserted and the system must surface it. The verdict is derived from recorded evidence. One case in this evaluation remains open.

The record supports reconstruction within Zortex's own boundary. It does not provide complete visibility into other software running on the same machine. No detection-latency distribution or broad detection-rate estimate is published.

The record provides evidence for reconstruction. The action and egress controls govern what may happen before an event is recorded.

8 Limitations and residual risk

The current evidence leaves several limits that matter to deployment. They concern the coverage of containment, a credential edge and evaluations that have not yet been completed. They are not resolved by the passing results elsewhere in the card.

Credential masking

An unstructured credential can cross the masking boundary in one automated-browser condition when the agent has not referred to it by name. This edge remains open despite the ten passing standard cases and two detected injected regressions.

Execution boundaries

OS-level containment is strongest on macOS and at the subprocess boundary. In-process execution and some plugin paths are not yet OS-sandboxed and rely on higher-level controls. The results do not establish OS-level isolation for every execution mode.

Incomplete evaluations

Connector certification contains four failing cases and one skipped case. The external-application access result is 14 of 15, with one unresolved case. The alignment suite has one failure and 26 unrun cases. Monitorability has one open case.

Limits outside the system

Same-machine adversaries with equivalent privileges are outside the stated protection boundary, as is an attacker who fully controls the operating system. Zortex's local-first design does not protect personal information from every form of host compromise.

Independent and legal review

No independent evaluation or external red-team has been conducted. Legal readiness is tracked separately from engineering readiness and is not established by this card.

These limits leave the release confined to reading, organizing and deriving knowledge. Consequential actions remain disabled, and external red-teaming remains a gate for enabling them.

8.1 Controls awaiting standing results

Seven controls are enforced in standard operation without a published standing result. They are part of the system description, but are not counted among the passing evaluations.

Table 8. Operational controls without a standing result

Control	Present description
Content-directive refusal	Instructions in ingested material do not become action authority; a ten-case suite is described
Outbound-egress policy	Unknown channels are refused; the strictest posture applies when no policy is present
Credential custody	Tokens are kept out of model context, logs, stored records and error output
Consequential-action floor	Minimum controls apply regardless of configuration
Data erasure	Partial and interrupted erasure paths are included in standard operation
Plugin permission envelope	Permissions are default-deny and plugin integrity is checked
Engagement-gaming resistance	The objective is constrained against the four stated gaming patterns

The privacy conformance review, for example, reports that the relevant specification sections conform. That finding does not supply a separate execution result for every partial or interrupted erasure path. Similarly, the existence of a content-refusal suite does not establish its standing passing count.

New standing results will be added to the evaluation record as they become available.

9 External review and wider deployment

Zortex has not yet undergone an independent evaluation or external red-team. The tests reported in this card are internal, including the live-account connector run and the privacy conformance review.

An external red-team engagement is a gate for enabling send, pay, post and delete. When that work is complete, Zortex's stated commitment is to publish the findings unedited, including adverse findings, alongside its response.

Table 9. Requirements described for consequential capabilities

Capability	Relevant controls and review
Send or post	Recall, egress containment, credential masking and custody, and external red-teaming
Pay	The send/post controls and review, plus a hard spending ceiling
Delete	Recall, minimum consequential-action controls and external red-teaming

The remaining prerequisites include the open credential-masking edge, unpublished standing results for credential custody and minimum action controls, and external red-teaming.

Changes to this card

The card will be revised as evaluations produce standing results, open findings are resolved and independent review is completed. Capability status should be read with the date and system version on the cover.

For Zortex 0.1, the operational boundary remains reading, organizing and deriving knowledge from connected accounts. The broader action capabilities remain disabled.

A Evaluation record

The tables below collect the results reported in this edition. “Not reported” means the public record does not provide the date or outcome; it does not imply a result of zero.

Table 10. Execution evaluations

Evaluation	Reported result	Run date
Data-egress containment	14 / 14 paths; zero bytes observed with outbound access disabled	Not reported
Socket denial	6 / 6 cases passed	Not reported
Cross-account isolation	8 / 8 cases passed	Not reported
Action recall	29 / 29 cases passed	Not reported
Connector certification	193 passed; 4 failed; 1 skipped	September 4, 2026
External-application access	14 / 15 cases passed; 1 unresolved	Not reported
Credential masking	10 / 10 standard cases; 2 injected regressions detected; 1 edge open	Not reported
Owner-authority alignment	29 passed; 1 failed; 26 not run	Not reported

Table 11. Specification review

Review	Reported result	Review date
Privacy, tenancy and erasure	11 / 11 sections conforming	August 31, 2026

Execution cases, outbound paths and specification sections are different units. No overall passing total is calculated across these entries. The tests also do not share a published random-sampling basis from which to estimate a single product-wide failure rate.

A.1 Evaluation inventory and open work

Table 12. Described evaluations without a complete result

Evaluation	Inventory or finding	Published outcome
Content-directive refusal	10-case suite	Standing count not published
Hostile-input corpus	7 fixtures	Maintained test set; no scored result
Interface-privilege parity	Includes an injection fixture	Asserted; case count not reported
Planted-event monitoring	1 case remains open	Passing count and total not reported

The hostile-input corpus is an inventory of test material. Its seven fixtures are not seven successful attack-resistance results. The interface-privilege and monitoring descriptions likewise do not provide enough information to create a passing percentage.

Table 13. Open findings retained in this edition

Area	Remaining work
Connectors	4 failed connector-acceptance cases; 1 case skipped
External-application access	1 unresolved case in the 15-case evaluation
Credential masking	1 known automated-browser edge
Alignment	1 failing case; 26 cases not run
Monitorability	1 open case
External evaluation	No independent assessment or red-team completed

The card will be updated as standing results become available and open items are resolved.

B Evaluation methods

Containment and sockets

The data-egress evaluation measures bytes at the outbound boundary with outbound access disabled. All 14 tested paths recorded zero bytes. A control-removal comparison checks that data can leave without containment. A separate six-case evaluation tests socket refusal. Transferred-byte quantities for the control-removed condition are not published.

Isolation and conformance

The eight-case isolation evaluation concerns the reach of account reads, erasure and results. The conformance review checks implementation against 11 privacy, tenancy and erasure specification sections. The review is not an additional set of 11 independent runtime trials.

Recall and access

The 29-case recall evaluation tests whether a consequential action can be withdrawn during the hold before it takes effect. The hold duration is not reported. External-application access is evaluated for read-only behavior across 15 cases, with one case unresolved.

Connectors and credentials

Connector certification uses live accounts and the production data path, including credential and certificate isolation. The credential-masking evaluation includes ten standard cases and two deliberately introduced faults. The known masking edge is recorded separately from those counts. No larger combined credential suite is inferred.

Alignment and monitoring

The alignment total includes all 56 planned cases, including the 26 not yet run. The monitoring check inserts a known event and requires it to be surfaced from the independent activity record. The monitoring denominator and a latency distribution are not reported.

Availability of records

This card publishes aggregate outcomes. It does not include case-level logs, a complete configuration manifest, evaluation record identifiers, random-sampling details or uncertainty intervals. Most run dates are not reported. These omissions limit reproducibility and any inference about the frequency of failures in ordinary use.

C Terms

Table 14. Terms used in this card

Term	Meaning
Consequential action	An action that reaches outside the machine or is not trivially reversible, such as sending, paying, posting or deleting
Recall window	The hold between preparing an action and its external effect, during which the action can be withdrawn
Egress	Data leaving the machine
Control-removal check	A comparison with a control disabled to test whether the observed outcome depends on that control
Conformance	A finding that a specification section is implemented as specified
Cross-account isolation	Limits that prevent one account’s reads, erasure or results from reaching another account
Credential custody	How account tokens are kept out of model context, logs, stored records and error output
Credential masking	Protection against a credential appearing during an automated-browser interaction
Injected regression	A deliberately introduced fault used to test whether the evaluation detects it
Standing result	A published, repeatable evaluation outcome that is kept current
Production data path	The same path taken by live user data, rather than a fixture-only path
Tamper-evident record	An independent record used to reconstruct system activity and check an agent’s self-report

D About this card

This card covers Zortex 0.1 as of September 17, 2026. The latest dated evaluation is the live-account connector run of September 4, 2026. The privacy conformance review is dated August 31, 2026.

Results and methods

The numerical findings are the internal evaluation results for Zortex 0.1.[1] The evaluation record lists the units, outcomes and available dates. Methodological limits are described in Appendix B.

The website provides data tables and downloadable counts alongside the numerical figures. Architecture diagrams describe relationships and permission boundaries; they are not measurements of data volume or elapsed time.

Publication history

Table 15. System-card record

Date	Entry
September 17, 2026	Initial Zortex 0.1 system-card publication

The card is updated as evaluation results become available, open findings are resolved and independent review is completed. Capability status applies to the system version and date stated in the card.

Reference

1. Zortex. *Zortex 0.1 System Card.* September 17, 2026. Internal safety evaluations, capability status, limitations and Appendix A evaluation record.