TECHNICAL BRIEF

Agent Defender

Permission checks before an AI agent acts.

In development · controlled local pilot
Brief revision 1.0 | October 3, 2026

Agent Defender is designed to check an AI agent's proposed action against its permissions and keep a record of the decision. An AI agent is software that uses AI to carry out tasks. The demonstrated pilot covers a small, explicitly connected local workflow.

What it does

A permission check is a checkpoint before an action. It can cover only the requests that actually pass through that checkpoint.

How the intended workflow works

- An agent requests a specific action, such as reading a connected tool's status or creating a branch in a practice project.
- The checkpoint compares the request with the permission granted for that task. An approval means a person has agreed to a particular action; the agent's own text is not approval.
- A connected tool carries out an allowed action. A request can also be refused or sent for human review. The workflow keeps the request, decision and result for later checking.

What the local pilot demonstrates

The recorded pilot supports reading Git status and creating an exactly approved branch in one disposable local project. Git is a tool for tracking changes to files. A branch is a named starting point for work.

A person can prepare a request, inspect its exact details, approve or decline it, and check the saved result without repeating the change. This is a controlled demonstration, not coverage of ordinary computer activity.

Inputs and outputs

Inputs	A proposed action, its connected resource, permissions and any required human approval.
Outputs	An allow, refuse or review decision; a record of the request and result for the connected workflow.
Coverage	Only the connected tools and actions that use the checkpoint. A bypassing action is outside its coverage.

TECHNICAL BRIEF

Agent Defender / Technical brief

In development · controlled local pilot
Brief revision 1.0 | October 3, 2026

Setup and information use

A useful deployment starts by identifying the exact actions to check, the tools involved and who can approve them.

What setup requires

Reviewed environment	Alocal pilot on Windows with a browser interface and a disposable Git project.
Connection work	Each protected tool needs an explicit connection through the checkpoint. There is no universal connection to every agent or application.
Approval rules	Specify allowed actions, resources, limits and who reviews requests. Check refusal and recovery as well as normal use.
Published specifications	Minimum hardware, a list of supported operating systems, response times, capacity and availability of a hosted service are not established for a public release.

What information is involved

The pilot uses action requests, approval information, local permission keys and saved result records. Its existing requests, keys and receipts are retained locally when the protected pilot is stopped or closed.

What a future connected tool reads or sends depends on that tool and its setup. Agree on storage, access, retention and any external transfer before using sensitive information. This brief does not establish encryption at rest or an enterprise retention policy.

Limits to keep in mind

- The pilot does not cover email, banking, GitHub, personal files, other agents or general computer activity.
- Agent Defender does not replace antivirus or fully isolate a computer. Software with other access may act outside the checkpoint.
- This website and PDF describe the product. They do not run permission checks or activate a service.

TECHNICAL BRIEF

Agent Defender / Technical brief

In development · controlled local pilot
Brief revision 1.0 | October 3, 2026

Evidence, limits and next steps

Tested behavior in a small local workflow does not establish effectiveness across other tools, agents or attacks.

What has been checked

Local workflow records describe request preparation, human approval or decline, and checks of saved results. A separate controlled demonstration checked missing approval, changed request content, repeated use of an approval and an unavailable checkpoint.

Focused source checks of permission expiry and compatibility have also been recorded. Those checks do not establish that the updated source is installed in the running pilot or that a complete public release has passed review.

What remains unverified

- Earlier tests with live models did not establish a security advantage and had limits on task completion. They are not evidence of useful everyday protection.
- Broad production readiness, independent external acceptance, security certification, protection accuracy and performance guarantees remain unestablished.
- Connections to customer tools and general interoperability need separate review. No working integration with Swarm Hunter is claimed.

Before evaluating it with your team

- Choose one harmless workflow and list its normal actions, forbidden actions and required approvals.
- Confirm that every covered action passes through the checkpoint. Test refusal, changed requests, repeat attempts and loss of the checkpoint.
- Review the resulting records and recovery process before expanding the scope. Confirm release version, platform support and data handling for your proposed setup.

Review basis: Selected local pilot and engineering records through October 2, 2026. A public product release version has not been established.
Public details: Read the technical details online
Typography updated October 4, 2026. Product evidence and review date unchanged.