MENLO SECURITY

SOLUTION BRIEF

Securing AI Sidebars

Get the Productivity without the Exposure

AI assistants built into leading enterprise browsers, including Microsoft Edge and Google Chrome, have changed how your employees work. But the same capabilities that make them powerful also introduce serious risks, as sensitive enterprise data and regulated personal information are silently streamed to large language models with little visibility or control. This brief explains the threat landscape and shows how Menlo Security closes the gap inside the browser session, where the exposure actually happens.

Your users spend most of their day in the browser. It is where team communication, workflows, SaaS applications, and web resources all live, and it is where most AI requests now originate. So, it was only a matter of time before elements of the browser itself became an integrated AI surface, too.

Chrome and Edge each offer AI-enabled sidebars, referred to as AI assistants, and the two are roughly comparable. For Chrome users, clicking on Ask Gemini opens the sidebar. For Edge users, Copilot opens with the chat icon. Copilot is also embedded across Microsoft 365 applications. In both browsers, the sidebar lets users summarize open tabs and query the underlying LLM without opening a new tab. Both are evolving toward more agentic versions, though what additional capabilities any given user sees today depends on license type, enterprise policy, and location.

The terminology in this instance is important. “Conversational” AI, or prompt-response AI, is the type that most users know. The user goes to the AI tool and poses a question → the browser sends it to the AI provider’s servers → the underlying LLM processes it → a response comes back. It is passive and user-driven. The model knows only what is entered in the user prompt.

AI Sidebars, or AI Assistants, Are Different. They have awareness of what is on the browser screen and in other tabs. In addition to the user’s question, the browser now sends content from the tabs the user has shared with the sidebar to the AI provider’s servers. Moreover, Gemini has access to Gmail/Drive, while Edge has access to content in Office 365 depending on permissions; both may retrieve additional content from these sources even if those tabs are not open in your browser.

Both Google and Microsoft are evolving these sidebars to include agentic capabilities. If conversational AI can be thought of as a question-and-answer model, and AI assistants can be thought of as Q&A in the content of the user's browser tabs, agentic AI is a tool that pursues goals. The user sets an objective, the AI plans the steps to reach it, and then the AI executes those steps. Agentic sidebars are no different.

Examples of things that Gemini auto browse could call upon to complete a task include Gmail for reading emails and drafting responses, and Google Calendar for scheduling. The broader integration set includes Drive, Keep, YouTube, Maps, and Photos, plus a "personal intelligence" layer that draws on Gmail, Search, YouTube, and Photos. So, a task like "find the confirmation email, add the trip to my calendar, and pull up the hotel on Maps" spans three of these APIs. The Browse with Copilot agentic sidebar in Edge can go even further by accessing connected Microsoft 365 connectors, so that agent responses and actions can be grounded in Microsoft 365 data including mail, documents, and calendar. That means that a task like "summarize this vendor's proposal against the contract in my OneDrive and draft a reply in Outlook" reaches Word/OneDrive, the mail store, and the calendar through these connectors.

> There are benefits and risks to any type of AI, and the sidebars are no different.

As agentic capabilities arrive, they layer on top of the previous AI assistant. Auto browse in Gemini, for example, represents agentic capabilities layered on top of Ask Gemini, and Edge works the same way.

Browser	AI Sidebar	Agentic Capabilities
Google Chrome	Ask Gemini sidebar	Auto browse
Microsoft Edge	Copilot in Edge sidebar	Browse with Copilot

Security and IT leaders may be aware of the risk posed by AI generally, but many don't know that AI capabilities are now built directly into the browsers that employees are working with. End users, even those who have been schooled in the use of sanctioned conversational GenAI tools, are almost certainly completely unaware of the risks posed inside their own browsers. And because browser vendors keep changing what these sidebars can do, often shipping innovations without warning, a flexible, forward-looking approach is necessary.

AI Sidebars, Security and Compliance Issues

The sidebars in Chrome and Edge are already taking on agentic behavior, and they change on a timescale measured in weeks, not months. Your security team has to account for what these sidebars can do today and stay current as that changes, which is a tall order. Google and Microsoft are both moving carefully, because the browser has become the way users reach enterprise applications, data, and systems, and a sidebar can inherit the user's permissions to reach all of it. Each vendor takes a slightly different path to the same goal

of enabling AI safely. Whatever the path, every sidebar shares information with the associated LLM, so both today's AI assistants and the emerging agentic ones call for the same caution. This is essential, because while both types of sidebar can offer productivity benefits, both come with risks.

In addition, while you can control browser policy for managed endpoints, BYOD and unmanaged endpoints may use browsers that feature sidebars with different capabilities and may leverage a user's account that is not their corporate account. Once again, this is an argument for a flexible approach that is easy to change as browsers themselves do.

Sidebar Governance Challenges

One of the biggest issues with AI sidebars has followed AI since ChatGPT arrived—data exfiltration. As with most AI use, users are not trying to exfiltrate sensitive data. They are trying to get their work done. What is different here is that users often do not realize they are sharing data at all. Regardless of whether a sidebar is an AI assistant or an agentic one, it is sending data from the user's browser to an LLM, often bypassing visibility and control. A few examples include:

Data Category	Example Exposure Scenario
Protected Health Information (PHI)	An employee opens a patient portal. The sidebar scrapes the record and sends it to the LLM as tab context.
Personally Identifiable Information (PII)	An HR self-service page with employee names, SSNs, and compensation details is shared as a background tab.
Financial and Payment Data (PCI)	A finance team member has a banking or ERP tab shared. Account numbers and balances are transmitted.
Legal and Privileged Communications	An attorney accesses a matter management system. Privileged client communications are sent to the LLM.
Enterprise IP and Trade Secrets	An engineer opens an internal code repository or product roadmap. Source code and strategy are transmitted.
Credentials and Session Tokens	A password manager tab is included in shared context, exposing credentials to the LLM.

The risk grows when considering the tabs themselves. For example, say a user had a tab with the enterprise CRM on it. Then the same user asks who won the World Cup. The browser sidebar may scrape the content of the CRM page to respond, even though it has nothing to do with the question about FIFA. And sidebars are not sending only the content from the foreground tab. For example, say that a user has opened their bank account in a browser tab. The user then opens a new tab to check the status of a project, and asks the sidebar to summarize progress. The content from your bank account tab might be shared along with the tab about the project that you asked to summarize. As AI sidebars continue to emerge and the ramifications of their use become apparent, it is essential to remember that flexible, easy-to-administer policy is essential.

Beyond data exfiltration, LLM-based AI also faces threats from prompt injection. Malicious instructions can get buried in the web content a sidebar uses as context, or in a page an AI browser visits during a task. The AI reads all of it and executes the instructions, including text a human cannot see and commands no legitimate user would issue. Even the model makers acknowledge that prompt injection will never be fully eliminated, which makes browser session visibility more important, not less. These attacks are dangerous because sidebars have access to the user's browsing activity and session authentication. That is why the controls that can stop them have to live at the browser session layer.

Use AI Sidebars Securely, Today, with Menlo

As this brief has shown, the AI sidebars in Chrome and Edge (and those yet to be created) carry capabilities and risks that cannot be controlled from outside the browser session. Securing them takes a solution built to operate inside the session itself. Menlo Security provides a layered control plane that sits between the AI sidebar and the LLM and enforces policy at the point where data would otherwise leave the organization. It addresses four distinct problems that enterprises must ask themselves:

- **Visibility.** What tabs are being shared, and what is being sent to the LLM?
- **Access control.** Which tabs should be allowed to share content at all?
- **Data security.** What sensitive content should be redacted or blocked, even from permitted tabs?
- **Auditability.** How can you get the information that you need to ensure that you have complete visibility, that access control is appropriate, and that data security is ensured?

Visibility

Menlo gives security administrators a single dashboard with real-time and historical visibility into AI sidebar activity across the organization, including:

- **Tab sharing inventory.** Which tabs each user has shared with the sidebar, including the domain and category of each one.
- **Prompt logging.** Full or summarized records of every prompt sent to the LLM, tied to the user's identity and active tab context at the moment it was submitted.
- **File upload tracking.** When a user uploads a file to a sidebar, Menlo captures the file metadata, including name, type, size, and classification, alongside the action the data security policy took.

Menlo correlates these logs with your existing identity and DLP signals, so analysts see the full picture in one place instead of pivoting between tools.

Access Control

With Menlo, administrators set policy for which browser tabs may share content with the AI sidebar. Policy conditions include:

Policy Dimensions	Examples
Domain	Block sharing from *.ehr-vendor.com, payroll.company.internal, or legal-matter-mgmt.com.
URL category	Block all tabs categorized as Healthcare, Finance, Legal, or Credential Management.
User or group	Apply stricter controls to users in regulated business units such as Legal, Finance, and Clinical.
Enterprise IP and Trade Secrets	Ensure that users working with code repositories or product roadmaps cannot share things like source code or strategy.
Credentials and Session Tokens	Block sharing from tools like password managers, which could expose credentials to the LLM.

When a tab is blocked from sharing, the user gets a clear in-browser notification, so the policy holds without leaving the user guessing about what happened. Menlo logs the block event for admin review.

Data Security

Even when a tab is allowed to share, the content scraped from it, and the prompt itself, can carry sensitive data that should not reach the LLM. Menlo's data security layer evaluates and acts on three categories of AI-bound data: user prompts, browser-scraped tab content, and file uploads.

User Prompts

Menlo's inline DLP control engine inspects every prompt before it is transmitted. When a prompt contains sensitive patterns, such as SSNs, credit card numbers, or internal project names, Menlo can take one of three actions:

- Mask. Obfuscate sensitive content with anonymized placeholders before the prompt reaches the LLM. The model still responds in context, but the sensitive values never leave the organization.
- Block. Stop the prompt from being sent and tell the user it contains content that violates policy.
- Log and allow. Let the prompt proceed while creating an audit record for compliance review.

Browser-Scraped Tab Content

When the sidebar pulls page content in as context, Menlo intercepts and inspects it as it is sent to the LLM. The same redact, block, and log actions apply when the sidebar is scraping content as when users are entering prompts. This protects passively scraped data that the user may never have meant to include.

File Uploads

Where a vendor supports file uploads, Menlo inspects the file content before it is transmitted. Administrators can:

- Block uploads to the underlying LLM globally.
- Log file metadata. File name, MIME type, size, and the action taken, whether allowed, redacted, or blocked, are captured in the AI logs.

Logging

Every Menlo action generates a structured log event that can be ingested by your SIEM or data lake. The AI sidebar control plane adds tab sharing context, including the domain, category, and time of any shared tabs.

The Agentic Evolution

The AI landscape is moving faster than most security frameworks can keep pace. What began as a chatbot in a browser panel is turning into a full browser-automation agent. Menlo is building the flexible platform that will let you stay ahead of the curve and safely deliver all of the benefits of AI, regardless of what browser is in use. While this document concerns AI sidebars in browsers, Menlo capabilities extend throughout agent use in general.

Human Versus Agent Identity in Web Traffic

When an AI sidebar agent makes an HTTP request, whether it is navigating to a URL, submitting a form, or uploading a file, it does so through the same browser session as the human user. Without more context, network security tools see a standard browser request and cannot tell it apart from user-initiated activity.

Menlo is extending its web traffic analysis to surface agent identity signals, so logs can carry an explicit indicator of whether each request came from the human user or from an AI agent acting on their behalf. That distinction will let administrators:

- Audit agent actions separately from human actions, so you can see what autonomous workflows are running across the organization.
- Apply differentiated policy. For example, allow a human to upload files to a SaaS platform while blocking an agent from doing the same.
- Detect anomalous agent behavior, such as high-velocity requests, access to unusual domains, or actions outside the agent's declared workflow.

Agent-Specific Policy Controls

As vendor feature sets settle across regions and subscription tiers, Menlo will extend its policy engine to support agent-scoped rules:

- Allow agents to read but not write in specific web applications.
- Restrict agent navigation to an approved list of domains or categories.
- Require user confirmation before an agent submits any form containing sensitive fields.
- Prevent agents from downloading files to local storage or uploading them to external destinations.

Conclusion

AI sidebars are not going away. They deliver real productivity gains, and employees will find ways to use them with or without IT's approval. So the question for your security team is not whether to allow AI sidebars in the browser, it is how to allow them safely.

Menlo Security gives IT and security teams the visibility, access control, and data protection to deploy AI sidebars with confidence, and it does so at the one layer where these risks are actually controllable: the browser session. From seeing what is being shared today, to stopping sensitive data from leaking, to preparing for the autonomous agents arriving tomorrow, Menlo keeps the productivity gain from turning into a compliance nightmare.