TrustModel.ai Exposes Hidden Risks in Top Chrome Extensions and AI Agents—63% Flagged in First Independent Trust Audit

First large scale automated trust assessment finds widespread risk across browser extensions, including AI agents. Only 9 of 108 earn 'Highly Trusted' status

MOUNTAIN VIEW, CA, UNITED STATES, April 16, 2026 /EINPresswire.com/ -- As AI-powered browser extensions rapidly proliferate, a new category of security risk is emerging inside the enterprise browser. Today, TrustModel.ai announced the first large-scale, independent TrustScore analysis of the 100 most-installed Chrome extensions and 10 leading AI browser agents — revealing that the majority introduce meaningful security and data exposure risks.

> Browser extensions are now one of the largest unmanaged attack surfaces in the enterprise. What makes this risk unique is the level of access and the speed of exposure, with frequent extension updates"
>
> *Ketan Nilangekar, Founder and CEO of ThreatWorx*

The analysis comes amid an unprecedented surge in browser extension supply chain attacks. In 2025 alone, over 35 Chrome extensions with a combined 2.6 million users were compromised through phishing attacks targeting extension developers, injecting data-stealing code into trusted extensions. The Cyberhaven breach alone exposed sensitive data from 400,000 users when attackers hijacked the company's Chrome Web Store account through a targeted OAuth phishing campaign.

KEY FINDINGS:

· 43% of the top 100 extensions have access to ALL websites you visit. These extensions can read, modify, and exfiltrate data from every page — including banking, email, and healthcare portals. Users grant this access with a single click during installation.

· 46 extensions monitor keyboard input. Keyboard event listeners were detected in nearly half of all scanned extensions — including extensions that have no functional need to track keystrokes.

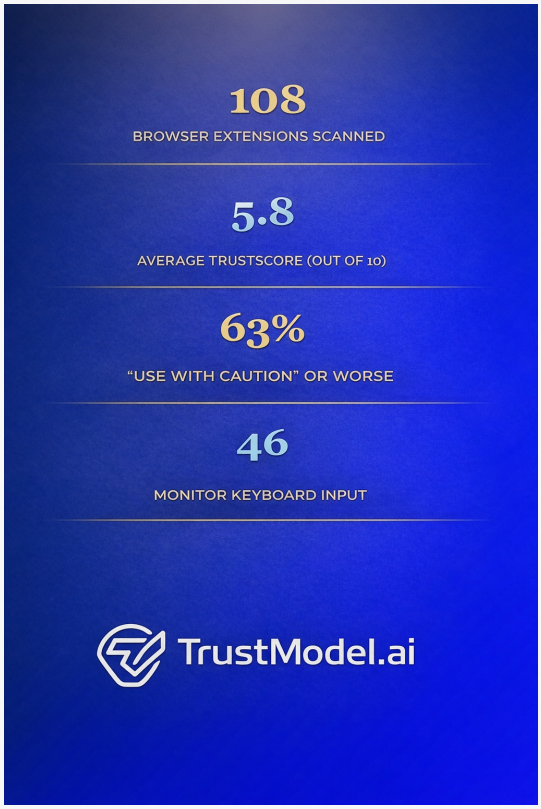
· 27 extensions use eval() — dynamic code execution that can download and run arbitrary code after installation, bypassing Chrome Web Store review. This is the primary vector used in supply chain attacks.

· Only 9 of 108 extensions earned "Highly Trusted" status (TrustScore 8.0+). The majority (68 extensions) fell into the "Use With Caution" tier, meaning they have legitimate functionality but an attack surface that warrants monitoring.

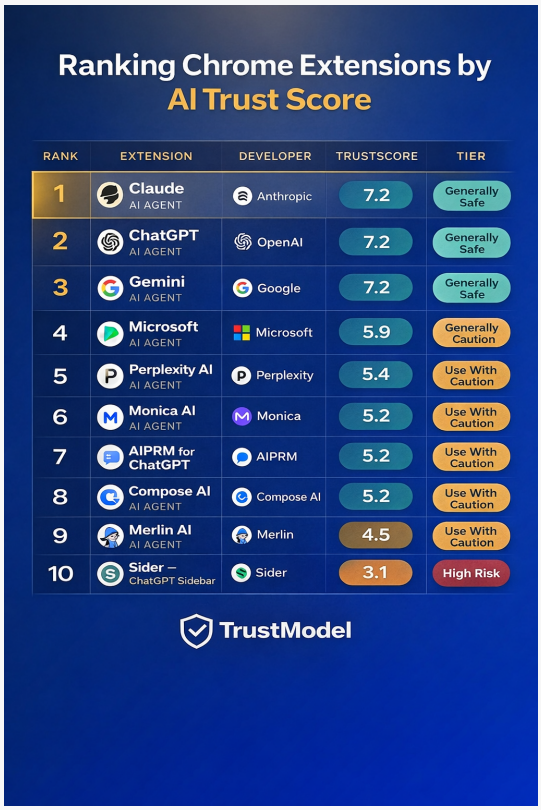
AI Agent Browser Extensions: A New Attack Surface
AI browser agents introduce a new and largely unmonitored attack surface. Unlike traditional extensions, these tools actively process user conversations, documents, and browsing context — dramatically increasing both data exposure and potential impact.

The big three — Claude, ChatGPT, and Gemini — scored at the top of the AI agent category, which reflects the security investment major AI labs are making, according to TrustModel's analysis. However, third-party AI extensions that wrap these models often introduce additional data collection, broader permissions, and less transparent code practices. Sider, the lowest-scoring AI agent at 3.1, requests access to all websites and exhibits code patterns associated with extensive data collection.

The Supply Chain Threat Is Real
Browser extensions represent one of the most under-examined attack surfaces in enterprise security. Unlike mobile apps (reviewed by Apple and Google's app stores with runtime sandboxing), Chrome extensions operate with deep access to the browser and are updated automatically, meaning a trusted extension can become compromised overnight through a supply chain attack, and every user receives the malicious update within hours.
Recent high-profile incidents include:



About the TrustModel Browser Extension Study



Browser Extensions Ranked

· Cyberhaven breach (December 2025): Attackers phished a Cyberhaven developer's Chrome Web Store credentials, pushed a malicious update to 400,000

users that exfiltrated cookies and session tokens for Facebook Ads, ChatGPT, and other platforms.

· 35+ extension campaign (2025): A coordinated attack compromised 35 extensions with 2.6M total users through OAuth phishing emails mimicking Google Chrome Web Store policy violations.

· Honey (PayPal) controversy (2025): Security researchers revealed that the Honey extension (20M+ users) was silently injecting affiliate cookies to claim credit for purchases — a form of revenue hijacking affecting every e-commerce site.

Top-ranked Browser Extensions

“Browser extensions have quietly become one of the largest unmanaged attack surfaces in the enterprise,” said Ketan Nilangekar, Founder and CEO of ThreatWorx. “What makes this risk unique is both the level of access and the speed of exposure — a compromised extension update can propagate to every user within hours, often without detection. As organizations adopt more AI-powered tools inside the browser, the need for continuous visibility and control over what’s running in that environment becomes critical.”

"With AI agents now embedded in the browser, the stakes are even higher,” said Ramesh Chitor, Chief Customer Officer, TrustModel.ai. These extensions don't just see your browsing, they process your conversations, read your documents, and interact with your data in ways that are opaque to the user. Our analysis shows that while the major AI labs and LLMs (e.g., those from Anthropic, OpenAI, Google) are working to build responsibly, the third-party ecosystem wrapping their models introduces significant additional risk. Enterprises need independent trust assessment for every extension in their fleet, which is what TrustModel provides."

Scoring Methodology
TrustModel's Browser Extension TrustScore evaluates each extension across five weighted dimensions using automated static analysis:

· Data Egress Safety: 25% weight (i.e., network communication patterns, external domains, data transmission)
· Permission Scope: 20% weight (i.e., Chrome permissions requested vs. functional need)
· Privacy Alignment: 20% weight (i.e., Privacy policy presence, cookie/keyboard/password monitoring)
· Code Integrity: 20% weight (i.e., eval() usage, obfuscation, encoded payloads, dynamic injection)

· Supply Chain: 15% Dependency vulnerabilities, CDN risk, external code loading

Extensions are classified into five tiers: Highly Trusted (8.0-10.0), Generally Safe (6.0-7.9), Use With Caution (4.0-5.9), High Risk (2.0-3.9), and Critical Risk (0.0-1.9). All extensions are re-scanned weekly as new versions are published.

Full Rankings Available
View the full rankings of all 108 extensions and individual extension reports at: trustmodel.ai/chrome-extensions. Extension developers who believe their score is inaccurate can request a manual review at extensions@trustmodel.ai.

About TrustModel.ai
TrustModel.ai is the AI Assurance platform — the independent trust standard for AI systems. TrustModel evaluates, remediates, and certifies AI across three layers: foundation models, COTS applications (Workday, SAP, Salesforce, and 230+ enterprise systems), and AI agents. The platform provides 10-dimension trust scoring, compliance mapping to EU AI Act, NYC Local Law 144, EEOC, HIPAA, and NIST AI RMF, real-time guardrails, and TrustModel Certified badges. TrustModel was developed by Predixtions, Inc. and backed by StartX (Stanford's startup accelerator). The platform is available on Google Cloud Marketplace and connects to 230+ enterprise systems via one-click OAuth integration. For more information, visit trustmodel.ai

Ramesh Chitor
TrustModel.ai
+1 925-895-6758
press@trustmodel.ai
Visit us on social media:
LinkedIn
