The State of Browser Security

2026

keep aware

内容

前言

在过去几年里，浏览器已成为最重要的企业终端设备。工作不再仅限于企业网络或传统应用程序，而是在标签页、AI副驾驶和生成式AI应用中进行。在SaaS平台，以及在实际上伪装成网页浏览器的桌面应用程序中。我们之前就说过，浏览器现在是工作的操作系统；2025年是这一理念成为现实的年份。

与此同时，安全市场释放了重要信号。我们见证了SASE和EDR领域的重大收购。整个类别都在相互碰撞，试图恢复对工作实际发生方式的可见性和控制权。这些收购不仅仅是整合，更是结构差距的证明。 **今天，浏览器是商业中最不成熟的安保控制点之一。**

既没有网络控制也没有传统的端点代理是为在企业逻辑在浏览器中执行的世界而设计的。它们都没有理解在数据创建、转换和共享点上用户行为的原生上下文。而且它们也没有被构建来管理直接在网络会话中发生的由人工智能驱动的互动。

2025年还见证了AI原生浏览器和AI嵌入式应用的快速崛起。浏览器不再是被动渲染网页的工具：它们是代理人、助手、自动化引擎和数据处理器。它们阅读、撰写、总结、上传、转换、并且……

在机器速度下传输敏感信息，并能自主行动。这种转变从根本上改变了风险模型：安全团队不再仅仅是保护浏览器中的用户，还需要保护浏览器本身可以执行的操作。

过去一年的事件让一点变得明确：浏览器不能仅仅是网络政策的延伸或终端保护的附属品。它需要独立站立，拥有自己的遥测功能、执行模型和数据架构。
浏览器需要自己的数据模型。

安全团队必须能够理解：

- 哪些数据正在被AI工具访问、生成或粘贴？
- 正在使用哪些SaaS应用以及如何使用
- 用户和人工智能代理正在实时采取哪些行动
- 数据如何在标签页、会话和云服务之间流动

没有浏览器原生模型，组织对生产力与风险汇聚的层面对此一无所知。了解浏览器安全状况的起点是了解浏览器本身的状况。今年的《浏览器安全状况报告》探讨了这一转变：AI浏览器如何重塑工作，为什么传统控制机制难以跟上步伐，以及SASE与EDR的融合如何反映出一个行业试图填补浏览器内部的最终差距。

工作的未来在浏览器中运行。

人工智能的未来在浏览器中运行。

安全必须在那里迎接。

Ryan Boerner，首席执行官，
创始人，保持警觉

引言

行业数据和去年安全团队的第一手经验证实了一个根本性的转变：浏览器现在已成为授予访问权限、处理数据和攻击展开的主要环境。

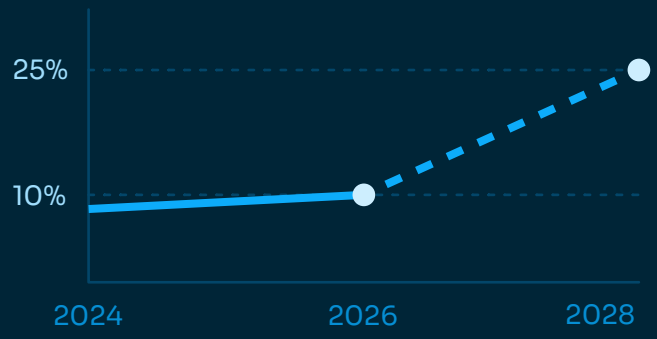
随着网络和电子邮件安全性的成熟，这些威胁转向了“已知安全”的环境：合法的浏览器、受信任的SaaS和已验证的会话。与电子邮件不同，在电子邮件中，组织可以控制流程和检查，而浏览器活动本质上是在数百个传统可见性之外的网页应用和服务中分散的。

这个浏览器盲点暴露了整个业务中的风险区域。凭证盗窃、会话黑客攻击和数据泄露现在可以通过日常工作（文件上传、SaaS登录、嵌入的第三方工具）发生，而不会触发传统的检测。攻击者针对浏览器，因为它是最受信任、面向用户的现代工作界面。

Verizon 2025 年 DBIR 这反映了：60%的安全漏洞与人为因素有关。行业分析师如Gartner同样指出，浏览器是一个关键但受保护不足的控制点。
Gartner预测，到2028年，将有25%的组织部署安全企业浏览器或基于浏览器的安全控制措施。 从今天的不到10%上升。

该报告中的研究验证了IT和安全团队在各个行业长期怀疑的事情： **传统的防御手段已不足以应对在浏览器内运行的劳动力。**

公司部署基于浏览器的安全控制



来源：2026年2月《Gartner企业安全浏览器市场指南》

方法论

本报告基于从生产型企业环境中使用Keep Aware浏览器安全平台收集的匿名遥测数据编制。分析反映了在过去十二个月观察到的真实世界浏览器活动，并通过针对一个月的快照选取的指标来突出近期趋势。结果以汇总和归一化形式呈现，以展示行为模式和比例，而非绝对量。

定义与分类

- **敏感数据** 包括符合企业政策定义的结构化和非结构化内容，如个人信息、健康信息、财务数据、源代码、机密文件和受监管的标识符。

- **企业账户** 是由企业身份提供者联合管理还是由组织验证为所有权的账户域进行管理？

- **个人账户** 身份不是由企业SSO或身份基础设施管理的吗？在许多情况下，这两种方式在同一浏览器会话中同时使用。

- **已确认的钓鱼攻击** 反映通过用户交互观察到的会话期间凭证收集或欺骗性登录流程。

- **扩展风险** 根据权限范围、声誉信号、组织环境、行为指标、源代码分析以及更新特征来决定。

安全挑战

随着工作在浏览器内持续巩固，企业最紧迫的安全挑战也日益凸显。 **生成式人工智能** 工具现已嵌入日常工作中，引入了新的数据暴露风险，这些风险需要实时监控和管理。敏感信息通常会被输入、粘贴和上传到SaaS应用程序中，使得 **浏览器数据泄露防护** 至关重要的在于确保数据移动过程中的可见性。同时，攻击者越来越多地通过浏览器原生技术——如钓鱼、OAuth滥用、恶意扩展和社交工程——进行攻击，这推动了需求……

浏览器检测与响应（BDR） This is translated into Chinese as: 在这之上还有 **广泛使用浏览器插件** 在浏览器内，过度授权或被木马化的插件会创造持续的风险。此外，新的 **AI优先和AI集成浏览器**

通过引入新的执行模型、扩大治理复杂性以及进一步复杂化用户、代理和客户端代码之间行为的归因，放大这些风险。这些趋势共同强化了一个简单的事实：确保现代工作安全需要将浏览器视为一流的安全控制措施。

1. 人工智能在现代社会劳动力监控中的挑战

生成式AI迅速融入日常工作中，浏览器成为主要的交互界面。虽然这些工具带来了生产力的提升，但也引入了新的数据暴露风险，许多组织都在努力理解，更不用说管理了。

我们的2025年数据显示，通用人工智能（GenAI）的应用已经普及，管理不均衡，且经常与敏感信息交织在一起。

广泛的通用人工智能（GenAI）采用在个人和企业账户之间呈现出碎片化态势。

在2025年，**41%的终端用户通过浏览器至少与一个AI网页工具进行了互动。**，带有**平均每位用户1.91个AI工具**这表明，GenAI 的使用不再局限于早期采用者或技术团队——现在它已经成为跨越各个角色和职能的主流活动。

41%

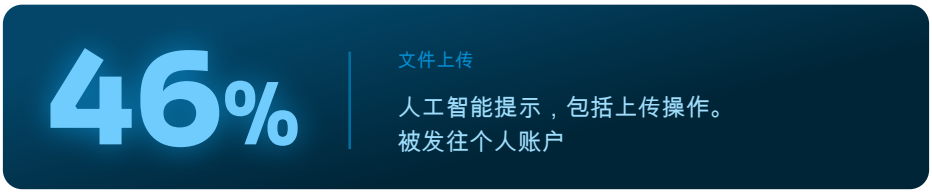
通用人工智能采用

至少有用户通过浏览器与至少一个AI工具互动过

然而，这种采用是分散的。在一月的时间里，**58%的AI提示输入被发送到了个人账户。**，相比于**42% 到企业账户**这种分裂凸显了一个反复出现的挑战：用户经常出于便利、熟悉或无限制访问等理由，默认使用个人AI服务——通常超出组织可见度和政策边界。

文件上传代表了一种特别高风险的交互。虽然只有**15%的人工智能提示包含了上传的内容。**，近一半的人

上传 **46%被发送到个人人工智能账户** ，
剩余（54%）发送到工作账号。



人工智能应用正在推动敏感数据的实质性暴露

风险在审查数据敏感性时更为明显。高达 **12%的AI提示输入涉及敏感信息** 包括个人信息、健康信息、财务信息、企业信息或开发者数据。更为令人担忧的是， **22%的包含文件上传的AI提示中包含敏感数据** ，这表明用户不仅仅是提出一般性问题——而且正在积极与人工智能工具共享内部材料。

在隔离敏感数据泄露时，个人账户使用仍是一个关键问题。 **6% of sensitive AI prompt inputs** 并且
23% 的敏感AI提示上传 经核实，系通过发送 **个人账户** —渠道，在这些渠道中，组织对数据保留、模型训练或下游使用几乎没有或没有控制权。

浏览器内管理人工智能的采用

这些发现凸显了一个核心现实： **人工智能风险并非理论上的** ，而且它是 **不限于授权工具** 管理人工智能的采用需要浏览器级别的可见性，了解正在使用哪些网络工具，用户如何与之互动，以及共享了哪些数据。如果没有在浏览器内实施可执行的政策和持续监控，组织在数据已经脱离其控制之后才会对与人工智能相关的事件做出反应。

2. 浏览器数据丢失防护：浏览器中的敏感数据泄露

随着企业工作持续向SaaS和基于Web的应用程序转移，浏览器已成为敏感数据传输的主要途径。凭证、源代码、客户记录、财务数据和内部文件通常会直接在Web应用程序中输入、粘贴和上传——通常根本不触及受管理的端点或企业网络，这使传统的DLP控制难以可靠地观察到。

我们的数据显示，浏览器中敏感数据泄露并非边缘情况，它是一种日常高量级活动，需要在与用户交互时的可见性。

杰出的敏感数据丢失至个人SaaS账户

在一个月的时间内， **46%的敏感输入到了网络应用的个人账户，而54%输入到了企业账户。** 这种几乎均等的结果凸显了安全团队持续面临的挑战：敏感数据往往通过标准企业身份、政策和监控范围之外的业务渠道流转。

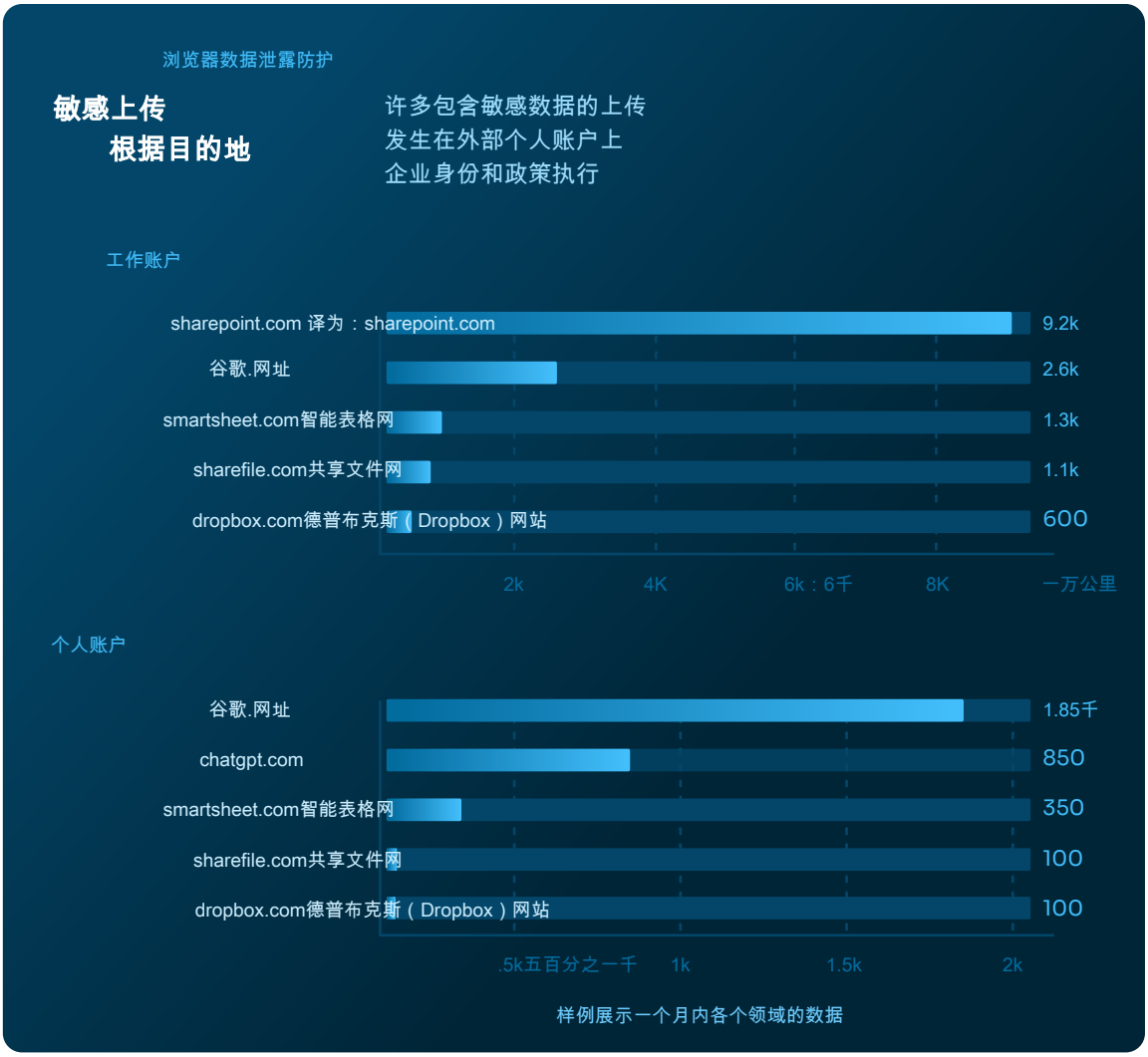
当近一半敏感输入流向个人账户时，仅依赖企业身份和授权应用列表的DLP策略与实际用户行为根本不符。

个人敏感上传与公司应用程序使用融合在一起

敏感上传呈现出更为明显的信号。在相同的一个月窗口期内，包含敏感数据的上传内容被观察到在广泛的网络应用中——涵盖了协作平台、云存储服务、生产力工具以及人工智能驱动的目的地。

敏感的上传，包含敏感内容的上传，在工作账户中在SharePoint、谷歌服务以及其他被授权的协作工具等常见企业平台上高度集中。然而， **敏感的上传到个人账户也显示出类似的模式** 常涉及相同的平台，但访问时脱离企业身份和政策执行。这种重叠使得基于目的地的屏蔽无效，因为…… **风险更多地不是由应用本身决定** 并且通过如何以及 **在哪个账户下使用** .

敏感数据的泄露并非仅限于“未经授权的应用程序”。它经常发生在受信任的平台上，这些平台可能存在身份验证不足或管理不当的情况。



应对现代深度学习平台（DLP）的浏览器原生可见性问题

传统的DLP工具，围绕电子邮件网关、网络检查或端点文件活动设计的，难以检测这些浏览器内的交互。 **输入的文本，复制粘贴的内容** ，并且 **浏览器原生工作流程** 常常完全绕过检查，使组织对敏感数据实际被分享的位置一无所知。

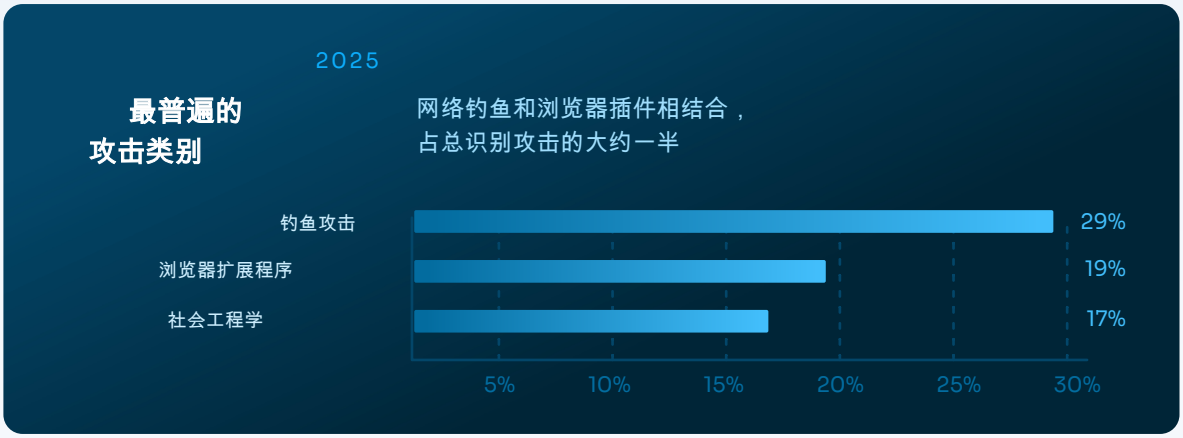
基于浏览器的数据丢失防护（DLP）并非旨在取代现有的控制措施，它更多的是关于…… **缩小可见性差距** 那些传统工具从未被设计来处理。没有对浏览器交互的洞察，组织无法可靠地回答基本问题：

- 敏感数据在哪里被输入或上传？
- 是转入公司账户还是个人账户？
- 哪些网络应用程序是主要的数据出口点？

有效的浏览器数据泄露防护(DLP)需要在浏览器内部直接可见——在浏览器内 **输入，上传** ，并且 **账户上下文** 在实时可见。随着基于SaaS和人工智能驱动的流程继续在现代企业工作中占据主导地位， **数据丢失预防必须存在于数据实际移动的地方：浏览器中。**

3. 基于浏览器的攻击：浏览器作为主要攻击面

网络安全行业持续进化以应对攻击者的创新，但一个关键环境仍然在很大程度上缺乏保护：浏览器。安全网络网关（SWGs）、防火墙和端点检测工具是必不可少的，但它们并非设计用于观察或解释每天在浏览器内部展开的丰富、用户驱动的活动。因此，攻击者越来越多地将他们的重点转向浏览器原生技术，这些技术**绕过外围防御，规避终端检测**，并且**利用受信任用户的互动。**



浏览器检测与响应（BDR）通过在浏览器中直接提供可见性和上下文来填补这一空白——现代钓鱼、社会工程和基于扩展的攻击实际上就在这里发生。我们的2025年数据显示，基于浏览器的攻击者策略主要由以下尝试主导：**凭证盗窃（约41%）和浏览器触发的环境升级**

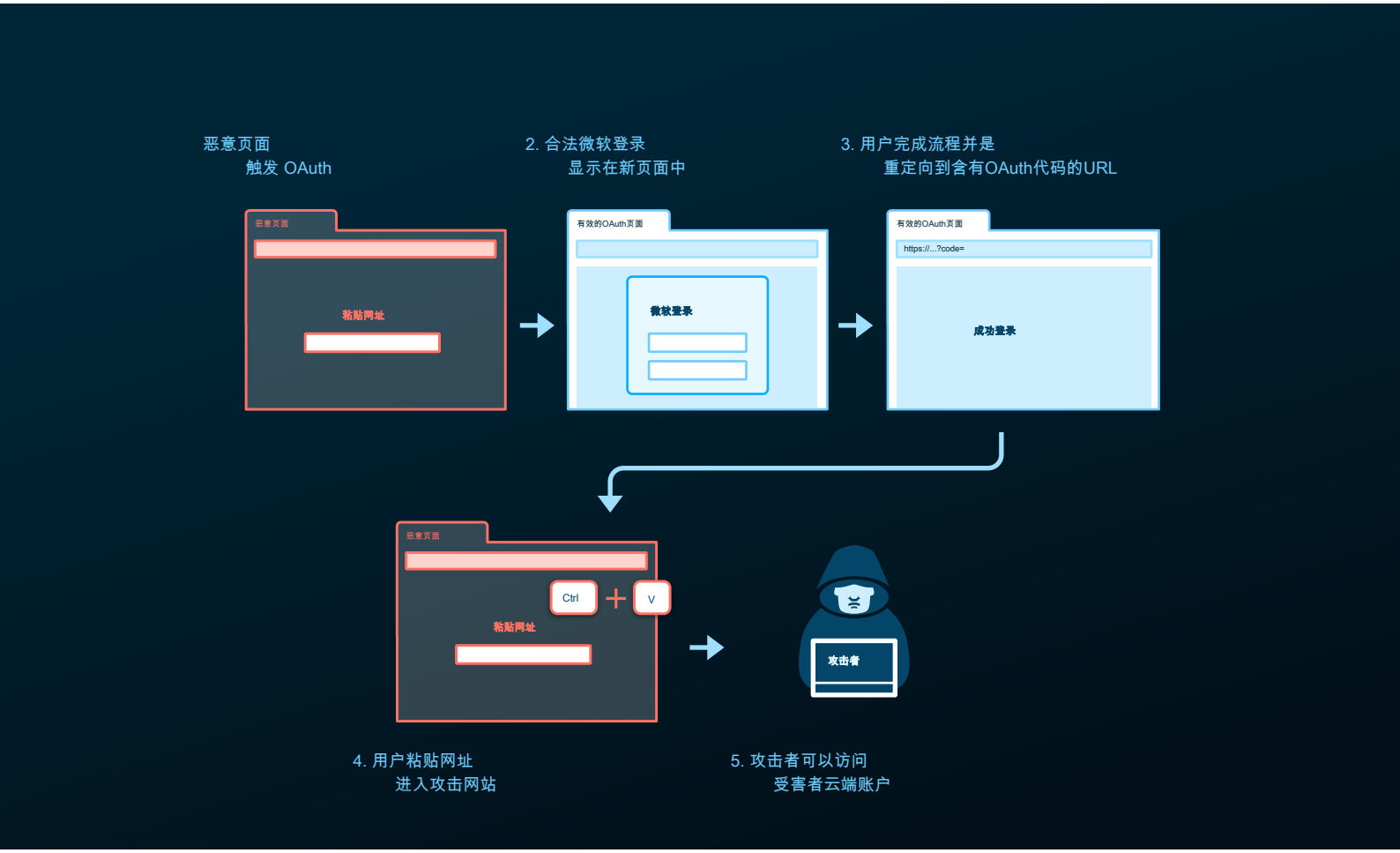
(~31%) 包括将用户从浏览器引导至终端、内部网络、云账户或甚至通过电话支持等非绑定渠道的技术。这强化了浏览器不仅仅是一个接入点，更是实现更广泛妥协的跳板。

新兴基于浏览器的攻击技术

攻击者持续创造和改进利用合法浏览器功能及用户行为的攻击技术。在2025年，三种新的攻击方式尤为突出：

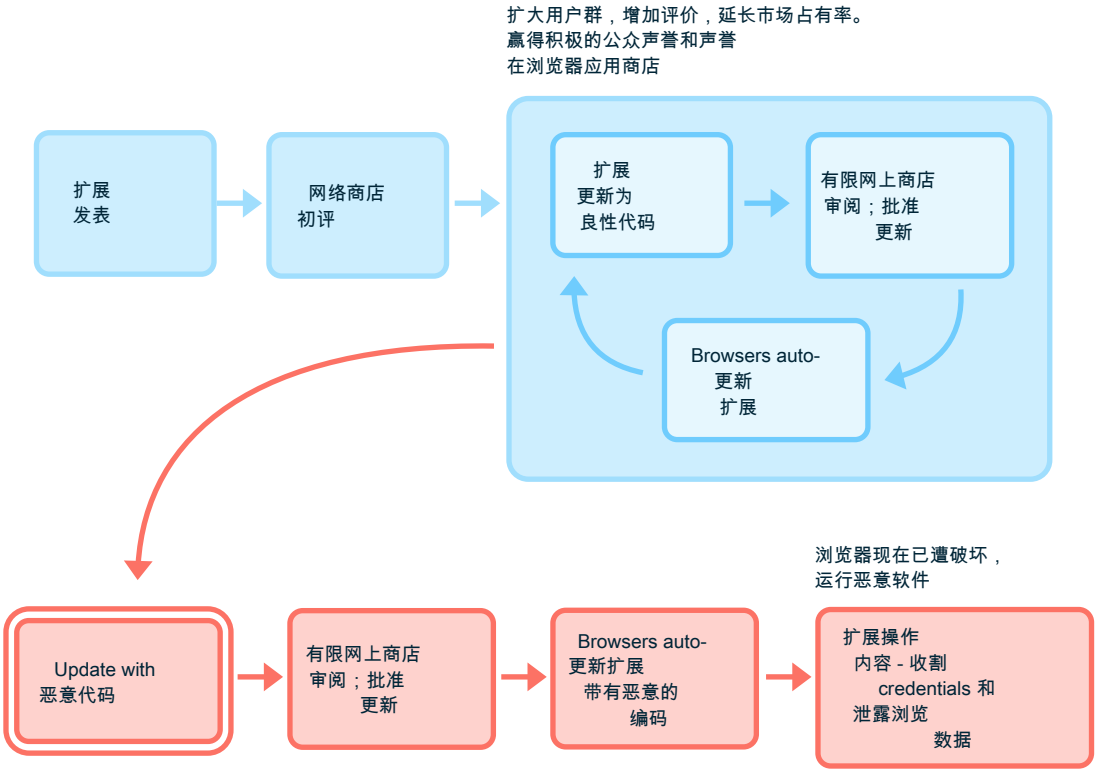
1. ClickFix：剪切板劫持和主机妥协

ClickFix攻击利用恶意网页，在用户不知情的情况下将攻击者控制的命令填充到用户的剪贴板中，然后通过社会工程学手段诱使用户将它们粘贴到终端或命令提示符中。这种技术通常会导致远程访问木马（RAT）或数据窃取器的安装。用户通过SEO中毒的搜索结果、受侵害的网站以及越来越多的途径遭遇ClickFix。 人工智能或大型语言模型生成的回复。



2. ConsentFix：OAuth滥用和云账户接管

同意修复 通过欺骗用户将授权码粘贴到攻击者控制的网站来利用合法的OAuth授权流程。这使攻击者能够访问云账户，而无需凭据或触发MFA提示。因为这个过程使用可信的身份验证机制，所以这些攻击通常会绕过网络、终端和基于身份的安全防御。



3. “隐秘”扩展：具有持久访问权限的木马化插件

睡眠扩展在安装时提供合法的功能，随后通过更新或远程配置激活恶意行为。一旦激活，它们作为持续存在的内部人员在浏览器中操作——篡改内容、收集凭证和窃取浏览数据。其中一些 运动 多年来在受害者环境中保持活跃，悄然规避传统的安全控制。

逃避技术在基于浏览器的攻击中越来越常见。

超越个体技巧，攻击者越来越依赖于
设计规避 构建基于浏览器的攻击结构，以战胜自动化扫描、威胁情报源和带外（OOB）分析。

现代运动通常使用 连锁攻击序列 涉及多个重定向、验证码、条件逻辑或用户输入，才显示恶意内容。这些序列隐藏了真实意图，避免非浏览器工具检测，大大延长了活动的持续时间。

攻击者还严重依赖 **合法或受信任的基础设施**，例如 被入侵的网站 或者攻击者控制的 SharePoint 实例，以绕过电子邮件安全并提高用户信任。 **隐藏**
条件执行 确保只有目标用户看到恶意内容，同时为扫描器和分析师提供无害页面——或者在外部分析期间重定向到像谷歌或亚马逊这样的可信网站。CAPTCHA 关卡进一步阻碍了自动化检测。

这些规避技术专门设计来阻止外部扫描仪、声誉系统、分析师和威胁情报源观察到与实际受害者遭遇的相同恶意内容和行为。结果是显著的检测缺口，这在分析现实世界钓鱼攻击的域和情报覆盖范围时变得显而易见。

域名年龄和威胁信息显示明显的钓鱼防护漏洞

许多组织依赖域龄、黑名单和第三方威胁情报来降低网络风险，但我们的2025客户数据显示，仅依靠这些控制措施是不足够的。

域名涉及已确认的网络钓鱼攻击有 **中位数年龄为6,591天（18岁及以上）**

阻止“年轻域名”并不是在攻击者滥用长期存在且受信任的基础设施时的可靠防御手段。

威胁情报差距进一步加剧了问题。在用户遭遇的以微软为主题的钓鱼网站上：

- **63%未被标记** 由任何VirusTotal供应商
- **77%未被标记** 通过URLScan
- **100% 均可被现有的非浏览器安全工具允许**

传统的工具和饲料一直无法检测出只有在浏览器内部才能明显发现的恶意浏览活动。

如何让安全团队适应：重新思考浏览器中的威胁检测

基于浏览器的威胁是互动的、情境相关的和动态的——这些是传统安全工具从未设计去分析的特质。检测这些攻击需要监控浏览器事件、用户行为、账户环境和攻击展开时的实时行为。

浏览器检测与响应（BDR）使安全团队能够从基于假设的检测转向基于证据的调查 通过直接在浏览器中观察活动，团队可以获取到检测规避攻击所需的可见性和上下文，了解用户影响，并有效地做出响应。

底线： 随着攻击者越来越多地在浏览器内部进行操作，检测和响应也必须转移到那里。BDR已不再是新兴概念——它是保障现代企业环境安全的基础性要求。

4. 扩展管理及安全

浏览器扩展仍然是现代企业浏览器中最缺乏监管和过度信任的组件之一。虽然扩展可以显著提高生产力，但它们也将持续且高度权限的代码引入用户的浏览器，通常缺乏安全团队对可见性或生命周期的监督。我们2025年的数据强调了重新审视扩展管理必要性的需求。

扩展蔓延是常态，非例外

在我们的数据中，扩展使用率始终很高：

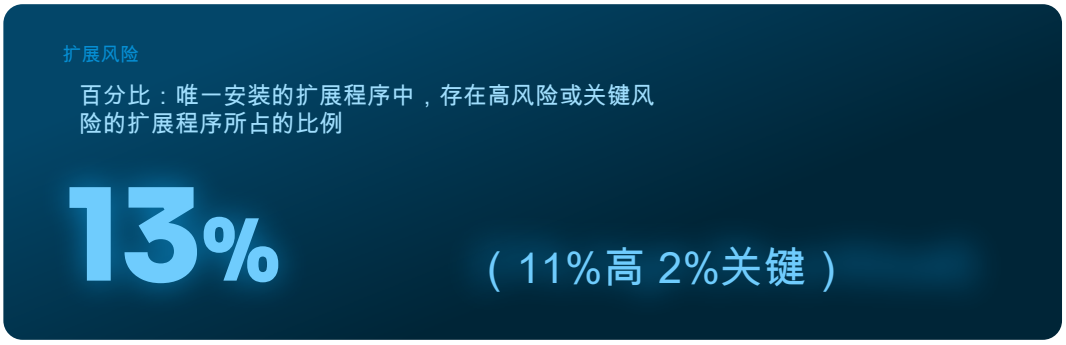
- **平均每位用户的扩展次数：** 4.67
- **中位数：** 4
- **最大观测值：** 40个单一用户的扩展

尽管表面上看起来每个用户分配四个到五个扩展似乎是合理的，每一次额外的扩展都会扩大浏览器攻击面——引入新的权限、新的更新机制以及滥用新机遇。随着时间的推移，无管理扩展蔓延无声侵蚀安全态势，而不会触发传统警报。

高风险扩展已在企业中实施

扩展风险并非理论上的——它已经存在于生产环境中。在我们的客户群中，安装的独特扩展的一个月快照显示：

13% 被归类为高风险或极高风险



尽管一些扩展程序在安装时被阻止，但其他扩展程序在检测时已在企业浏览器中处于活跃状态——通常可以访问敏感数据并拥有持续执行的权限。在许多情况下，风险并非由明显的恶意软件驱动，而是由 **高度授权或低声誉扩展** 那超出传统安全审查流程的部分。

权限是扩展风险的主要驱动因素

权限仍然是评估扩展风险的最强指标。高风险扩展通常会要求广泛访问浏览器功能，例如存储、脚本、标签访问、网络请求、Cookie以及访问所有网站的权限——这些权限能够深入了解用户行为和浏览器会话数据。

值得注意的是，我们的2025年数据显示：

- **30% of rare extensions with high-risk permission sets** 被浏览器市场分类为 **生产工具**

这突显了一个核心挑战： **基于名称、类别或观察到的功能看似无害的扩展仍可能引入重大风险** 根据他们请求的权限。没有权限级别分析，危险扩展程序会无缝融入日常工作中。

市场分类并不能反映实际风险

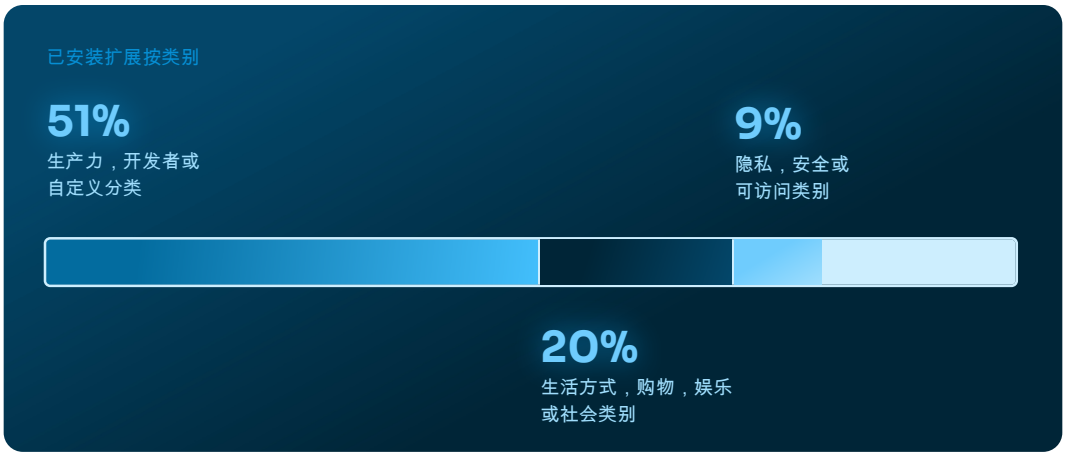
浏览器市场分类提供的安全信号意义不大。我们的数据显示，几乎所有 **声誉不佳的扩展** —那些表现出令人担忧的特征，如过度权限、行为不一致、广告软件活动或隐私问题——被广泛归类为 **生产力** 工具。

在2025年：

- **93% of poor reputation extensions** 被市场标称为生产力扩展程序

查看一个月内的整体安装扩展情况：

- **≥51%** 陷入 **生产力、开发者、定制** 分类
- **≥20%** 进入 **生活方式，购物，娱乐，或社交** 分类
- **≥9%** 进入 **隐私、安全或可访问性** 分类



这种类别与安全态势不匹配导致基于类别的允许列表无效。必须评估扩展的 **他们请求的权限以及他们所做的事情** 不是如何被营销的。

《“隐形”延长服务的真实风险》

“卧底”扩展功能与卧底特工非常相似——在几个月甚至几年内正常运作，然后突然间爆发出惊人的能量。 **在激活恶意行为之前建立信任** 这些扩展程序在安装时通常是良性的，提供合法的功能，并请求看似合理的权限，使它们能够逃避安装时的审查并广泛传播。

在2025年，利用潜伏期延长的活动被报告的频率逐渐增加。一个显著的例子是 Shady熊猫 ，which repeatedly used **被木马化的生产力扩展** 该软件后来通过静默更新或远程配置更改激活了恶意行为。因为扩展会继承先前授予的权限，并且无需用户交互即可更新。 **从合法转变为恶意的行为通常发生在安装之后很久** 当控制最不可能有人在看的时候。

这使 **即时批准和静态允许列表无效** 扩展风险是动态的，防御措施必须考虑安装后的行为变化。

重新思考扩展安全和管理的策略

低信誉、权限过度的扩展可能会 **访问敏感的浏览器数据，并跨会话持续。** 由于它们看起来无害，用户会轻易安装它们，而传统的安全控制很少会标记它们。因此，有意义的扩展风险往往直到造成损害才被发现。

有效的扩展安全需要从静态治理转向持续监管。组织必须评估扩展 **名誉、权限、代码和随时间改变的行为** —不要仅依赖市场分类或安装时的批准。随着浏览器成为主要的工作环境， 扩展管理 必须被视为一项积极的安保纪律，而不仅仅是行政上的事后考虑。

5. 人工智能浏览器与浏览器扩张

十多年来，企业浏览器策略相对简单。在大多数环境中，两种主导浏览器——Chrome和Edge，在我们的客户基础中代表了90%以上——占据了绝大多数的使用份额。治理、兼容性测试、扩展管理以及政策执行都集中在一个狭窄的生态系统内。

这种简单正在消失。 **2025年标志着以人工智能为先导和具备代理功能的浏览器新时代的开始。** ChatGPT Atlas、Dia和Comet作为AI原生浏览器进入市场。同时，Chrome整合了Gemini，Edge推出了Copilot模式，将代理功能直接嵌入主流企业平台。短短一年间，浏览器从以用户为中心的界面转变为代理执行平台。

这次转变实质性地改变了浏览器安全性的本质。

浏览器扩展：更多厂商，更多执行引擎

人工智能浏览器的兴起引入了一种形式的 **现代IT蔓延** 每次新的浏览器平台都会扩大企业的攻击面和管理范围。安全团队现在必须评估：

- 新浏览器供应商和信任模型
- 数据处理规范，适用于新系统和嵌入式人工智能系统
- 具有不同权限范围和治理机制的代理特征

每种浏览器，以及每一项AI添加和集成，都代表着数据处理、自动化和用户交互的独立控制平面。即使一家组织在Chrome或Edge上标准化，AI共飞驾驶员和“自动浏览”功能的添加也是如此。

AI浏览器

人工智能原生浏览器的兴起

在不到18个月内，主要的浏览器供应商已经将AI原生嵌入到浏览体验中扩大浏览器攻击面。

2024

2024年末

\"迪亚\

宣布为Arc的人工智能优先继任者针对2025年初的发布

2025

也许

Chrome + Gemini（Chrome + 双子座）

Gemini在Chrome上的推出；页内助手，摘要，多标签上下文

6月11日

\"迪亚\

仅限邀请的Dia浏览器测试版发布

7月9日

彗星

初版在Windows和macOS上推出，高级用户访问权限有限

7月28日

边缘 + 伙伴

Copilot 模式启动，将 Edge 定位为“AI 浏览器”。

9月17日

Chrome + Gemini（Chrome + 双子座）

更广泛的Gemini在美国Windows/Mac的推广；展示代理功能预览

10月21日

ChatGPT Atlas

公测版Atlas浏览器在macOS上发布

10月22-23日

边缘 + 伙伴

共飞模式被视为直接的AI浏览器竞争对手；“旅程”和更主动的用户体验（UX）被引入。

2025年末

彗星

彗星将免费且在全球范围内提供

11月20日

彗星安卓版，彗星应用程序的安卓版本发布

2026

1月27-28日

Chrome自动浏览

代理“自动浏览”功能在美国桌面版付费Gemini等级中推出

February

边缘 + 伙伴

更具体的代理型Copilot功能陆续推出，多标签页任务自动化

有效地创建新的执行引擎和 **模糊了用户、代理和页面驱动活动之间的责任归属。** 在现有浏览器会话中。

代理浏览改变身份归属和政策执行

传统浏览器安全假设了一个相对简单的模型：人类用户与网页交互。AI集成和代理浏览器通过引入另一个实体到系统中，挑战了这一假设。

现代人工智能集成的浏览器可以自动搜索和浏览网页，执行多步骤工作流程，并与SaaS平台互动——通常无需用户明确点击或同意。

这引入了 **一项新的归因挑战：一个动作是由用户发起、由AI代理执行，还是由客户端页面代码触发的。** 每个演员可能都在相同的认证会话中操作，但具有不同的意图和风险影响。

从政策角度来看，人工智能代理实际上成为了拥有授权的新数字行为者。如果没有对人类、代理和页面驱动活动在会话层面的明确归属，使用传统安全工具在数据泄露防护、威胁检测和扩展治理方面的一致性应用将变得越来越困难。

新型威胁模型：提示注入、过度代理和虚假信息

人工智能在浏览器中的能力扩展也带来了风险，正如文档所记载的那样。 OWASP Top 10 for LLM and GenAI Applications 项目

错误信息 (LLM09:2025) 对操作效率产生影响。自动浏览功能，可检索并执行内容，可能会将不准确的或攻击者控制的内容引入决策及随后采取的行动。正如本例所示： 真实世界案例 由人工智能驱动的响应可以引导用户从看似无害的查询

进入恶意执行链。

提示注入 (LLM01：2025) 在人工智能集成浏览器背景下变得尤为相关。恶意网页或电子邮件可以嵌入隐蔽指令，旨在操纵人工智能代理的行为，使其检索并泄露敏感信息，导航到恶意站点，或执行其他未经用户知晓或明确同意的操作。

过度代理 （LLM06:2025）将这些风险相加。当代理被授予广泛的权限——比如：电子邮件访问、SaaS集成、文件检索能力——注入漏洞可能从误导信息升级到未经授权的操作。

随着AI浏览器的自主性增强，这些风险已经从理论层面和轻微令人担忧的状态转变为实际存在和容易被滥用的状态。

持续差距：人工智能集成时代中的DLP和扩展

引入代理浏览增加了新的浏览器风险，同时加剧了现有的风险。

敏感数据现在可能通过经过验证的浏览器会话内的AI中介处理流程进行流动，这些互动是传统企业DLP控制措施未曾设计用于监视或执行的。

扩展风险依然存在。在AI启用浏览器内部运行的高权限或木马化的扩展可能会获得更广泛的环境和访问权限，从而增加基于扩展的攻击的风险。

浏览器蔓延进一步增加了这些风险必须持续管理和控制的场景数量。

探索人工智能浏览器生态系统

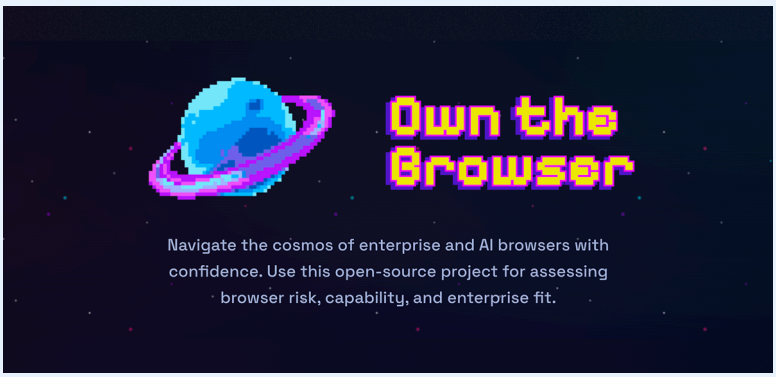
\"键盘与椅子之间存在风险\"这个成语已不再适用，因为 **人工智能代理是有效的新数字员工，在浏览器会话中运行。** AI赋能的浏览器现在必须被视为多用户执行环境，需要对其...

需要区分用户和代理的行为。安全团队还必须评估代理权限、数据保留和培训政策，以及有意义的遥测数据（包括身份归属）的可用性，以实现有效的检测和响应。

人工智能集成的浏览器扩大了现有的浏览器风险。将这些平台视为简单的生产力升级的组织可能会失去对浏览过程中谁或什么在行动的清晰认识，以及有效执行安全措施。

随着浏览器泛滥的趋势持续，以及人工智能的普及速度加快，在不同环境下实现一致的、集中的、浏览器原生的可见性将变得至关重要。没有它， **人工智能驱动的浏览不仅引入了操作复杂性，还带来了新的且难以诊断的安全盲点。**

为了支持结构化评估，Keep Aware 发布了一个开源框架，用于...**评估人工智能和企业浏览器风险、能力及治理因素** **掌控浏览器**。



未来展望：无浏览器可见性下的增加的混乱

浏览器已成为现代企业的执行层。它是输入凭证、授予SaaS访问权限、上传敏感数据、执行AI工作流以及越来越多地发生攻击的地方。到2025年，凭证盗窃和浏览器引发的向其他环境升级的威胁活动占主导地位，同时扩展风险持续存在，个人账户在生成人工智能和Web应用程序中的使用进一步模糊了敏感数据泄露。传统的边界、端点和基于声誉的防御措施并未设计用来解释现在定义企业工作的会话级浏览的动态特性。

AI原生浏览器和代理正将浏览器会话转变为能够代表用户执行多步骤工作流的自动化引擎。随着这一演变的持续，安全团队将面临
日益难以区分由用户发起的操作、客户端脚本行为以及AI执行的任务。 在域级别或在静态用户与应用程序交互基础上构建的日志和策略模型，在无法进行会话级别活动归属和区分人为操作、代理以及页面驱动执行的情况下，将越来越显得不足。

与此同时，边界之间 **个人和企业SaaS使用将继续模糊。** 敏感数据将在企业账户和个人账户之间无限制流动，反映了在SaaS环境中工作的实际操作方式。仅凭网络检查和静态授权应用控制无法解释在浏览器会话内部但超出企业SaaS控制之外发生的输入、剪贴板活动或文件上传。

攻击者正在利用这一被忽视且治理不足的接口。条件执行、隐形、链式攻击序列和基于CAPTCHA的钓鱼流程越来越多地

普通。这些技术确保自动化扫描器、爬虫、频外分析和第三方威胁情报源不会观察到发送给真正受害者的相同内容。随着隐蔽技术的扩展，包括基于AI用户代理的差异化。 **外部声誉系统将继续不足以保护免受新或零日钓鱼攻击。** 对组织持续构成浏览器层威胁。

扩展风险将保持持续。过授权和被特洛伊木马化的插件，包括安装后数月或数年才激活的休眠扩展，表明安装时的审批和静态允许列表是不够的。有效的治理需要持续评估权限、更新和行为。随着更多恶意活动被公开报道， **安全团队将面临越来越大的内部压力，以加强对未得到有效管理的浏览器扩展威胁向量的治理和保护。**

在这个环境中， **浏览器必须被视为执行环境，而不是可信应用。** 检测和响应策略必须结合浏览器级别的可见性、实时会话上下文以及用户、AI代理和客户端代码间的活动明确归因。

浏览器检测与响应（BDR）是浏览器、SaaS依赖和AI驱动型工作团队中坚固的深度防御策略的自然补充。 **提供所需的带内可见性、行为上下文和响应能力** 为了应对数据泄露、凭证盗窃、规避式网络钓鱼、基于扩展的妥协以及日益增长的AI浏览器领域。

攻击类型	定义	示例
凭证盗窃 登录表单 / 钓鱼攻击	社会工程学攻击类型，其中 攻击者创建欺诈性电子邮件， 网站或冒充的消息 可信实体以欺骗用户进入 泄露他们的登录凭证。	攻击者发送一封声称要 来自一家银行的安保团队 链接到伪造的登录页面。当用户 输入他们的用户名和密码，然后 攻击者获取了凭证和 访问真实账户。
ClickFix攻击	社会工程诈骗手段，通过说服…… 用户在本地执行代码，启用 攻击者建立设备持久性 或更广泛的系统访问。	用户访问了一个受损害的网页 指导用户无意中 将恶意代码粘贴到设备中 终端。粘贴后，恶意软件 下载并执行下一阶段 攻击，信息窃贼。
同意修复 攻击	攻击欺骗浏览器用户 将OAuth授权码提供给 攻击者控制的网站，最终 允许攻击者获得访问权限 云账户，未提供 证书或MFA。	一个用户从恶意页面访问 一个谷歌搜索结果。该页面 提示用户提供他们的电子邮件地址 合法登录到微软，并 粘贴生成的OAuth重定向 恶意网站URL。攻击者 然后使用OAuth授权码 在URL中获取访问权限并刷新 代币发放给用户的微软云 账户
横向移动 浏览器发起的 升级至 其他 环境/ 浏览器到X 升级	攻击者使用的提升技术 从浏览器交互到更广泛 访问——无论是用户端点上 在内部网络中，进入外部 账户和系统，或转移到另一个 通讯渠道，例如电话 基于社交工程	在访问一个恶意搜索引擎时 结果，财务术语，用户 遇到ClickFix攻击提示 欺骗用户在他们的设备上运行恶意软件 主机设备。攻击者的影响 开始于浏览器，在网页中， 转向了受害者的机器。
“卧底” 扩展	扩展程序，展示合法的 功能在安装时，然后，之后 数月甚至数年，将激活 通过代码进行恶意行为 更新或远程配置。	攻击者发布了一个良性的“浏览器 \"Chrome浏览器扩展：清洁器\ 储藏室。安装后，用户能够 清理一些大型浏览器缓存 数据。经过四个月的良性 活动，然而，攻击者发布 一项包含恶意代码的更新 监听网络活动并窃取 敏感数据到一个攻击者控制的 领域

关于保持警觉

感谢您抽出时间阅读《2026浏览器安全报告》。我们同时鼓励读者探索 掌控浏览器 这是一个免费的开源项目，它为安全团队提供了一个简单的方法，通过企业浏览器的全面目录来评估浏览器的风险、功能和企业适应性。

Keep Aware的以可观测性为先导的浏览器安全平台提供对每个工作和人工智能浏览器的深度可见性和浏览器检测与响应（BDR），将浏览器转化为一项一流的安全控制工具。

该平台观察真实浏览器行为，以检测现代的浏览器原生攻击。它保护通用人工智能（GenAI）的使用，对浏览器活动实施细粒度控制，并阻止身份威胁、网络钓鱼、社交工程和恶意扩展，所有这些都不会干扰员工的生产力。

如果您团队将浏览器安全视为首要任务，我们邀请您申请演示，了解Keep Aware如何帮助您确保今天的工作方式得到安全保护。

请求演示