企业浏览器观察
重塑 Web 体验的机遇与挑战

Enterprise Browser Observations: Opportunities and Challenges in Reshaping Web Experience

SPEAKER

叶家明 Jiaming Ye

360 数字安全集团 · 企业浏览器 · 产品负责人

SESSION

Agentic Web · 智能体万维网

DATE · VENUE

2026.04.22 · 杭州 Hangzhou

CONTENTS

目录

「2025 → 现在」 · 浏览器似乎又回到故事中心

The browser became the hottest real estate — three forces converged onto one surface

① FUEL FOUNDATION

模型新势力

开源与闭源双线并进 · 国内外同步迭代

DeepSeek · Qwen · Kimi · Gemini · GLM · 智谱 · Claude 4 / Sonnet 4.5 · 豆包 · GPT-5 · Ollama · Grok · MiniMax

② CONTAINER AGENTS

Agent 爆发

从"能聊"走到"能做事情"

Browser Use · Computer Use · File System · Claude Code · OpenClaw · Manus · MCP · Skills · ACP · Hermes

③ SURFACE AI BROWSERS

AI 浏览器涌现

从零到数十款 · 国内外同步

Dia · Arc · ChatGPT Atlas · 夸克 · Perplexity Comet · 豆包 · Chrome Gemini · Edge Copilot · 360 AI企业浏览器 · 纳米AI

TIMELINE · 协议 & 浏览器关键节点

2024 Q4 → 今天

2024 Q4	2025 Q1	2025 Q2	2025 Q3	2025 Q4	2026 Q1	今天
MCP 发布	DeepSeek R1	Comet · Claude Code	360 AI 企业浏览器「伴随式AI」	Atlas · Skills	A2A · ACP	OpenClaw

模型、智能体、浏览器三股力量互相成就。

浏览器角色变化 · 从入口到执行

Entry → Execution · What 2025 structurally changed for the browser

BEFORE · Web 2.0

入口

Entry

从这里开始

2025 → 今天

NOW · Agentic Web

执行

Execution

关注结果

阅读 → 操作

Reading → Acting

过去: 看信息 · 手动点击 | 现在: 说一句话,Agent 把事做完

点击访问 → 意图判断

Clicking → Intent

过去: URL · 标签 · 机械寻址 | 现在: 表达意图,Agent 自主编排

人在 loop → Agent 在 loop

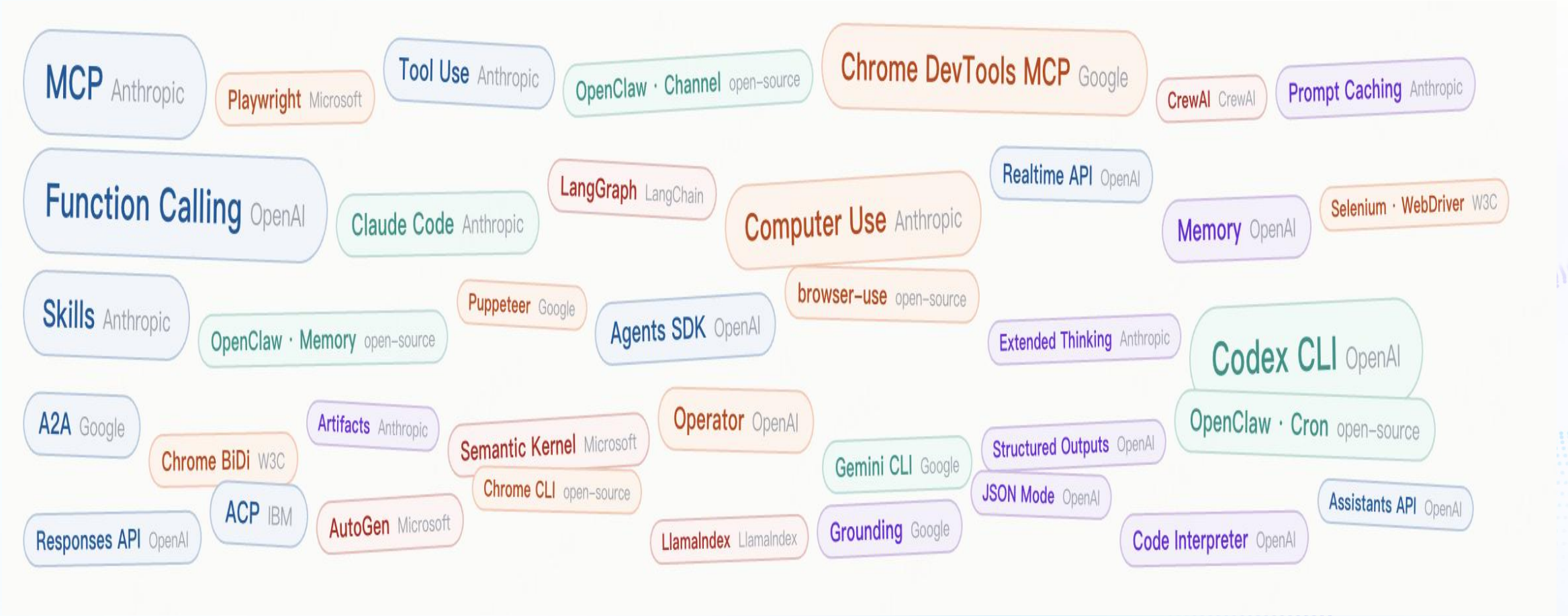
Human-in-loop → Agent-in-loop

过去: 每步等人按回车 | 现在: Agent 持续运转,人只关心结果

浏览器在新一代交互栈里的 **位置** 变了。

执行任务 · 大家都是怎么做的

How the industry executes tasks



我们的判断

Must be both a Consumer and a Provider of AI capabilities

浏览器同时是 **AI 的消费者**,也是 **AI 的提供者**。

CONSUMER · 消费者

浏览器 → AI

"调用 AI 做事" · 让自己更强

- 内置AI能力接入，调用大模型/智能体
- 内置 MCP Client,连第三方工具/知识库
- 具备Skills基础运行能力

PROVIDER · 提供者

AI → 浏览器

"AI 要浏览器做事" · 让自己可被调

- MCP Server · page/DOM/selection
- Chrome CLI 能力,对外作 Skill 被 AI 调
- Browser Agent / Memory 对外开放接口

我们的尝试 · AI 运行时

Adjustments · making the browser an AI runtime

CONSUMER · 消费者 浏览器 → AI

浏览器调 AI

✓ CONSUMER 内核 API · 提供给指定 Agent 可读/写页面

DOM · 选区 · 配 Agent 白名单授权

✓ CONSUMER 第三方 Agent 注册 · 分发 · 授权

管理员注册 第三方智能体;按用户/部门/角色下发,可审计吊销

✓ CONSUMER 内置 MCP 运行时 · 接第三方工具/知识

扩充浏览器能力，使得浏览器可以访问外部智能体、工具，来满足更丰富的场景

✓ CONSUMER 兼容 Skill · 实现复用已有 AI 能力

企业策略变成 Agent 可执行的行为约束

PROVIDER · 提供者 AI → 浏览器

AI 调浏览器

→ PROVIDER CLI 、Skill 被 AI 调

第三方AI客户端/工具可以通过浏览器提供的 cli / skill 来驱动浏览器，甚至是浏览器管理平台

→ PROVIDER MCP Server

page / DOM / selection / 截图 作为 Resources 供外部 AI 消费

→ PROVIDER Local / Memory 数据

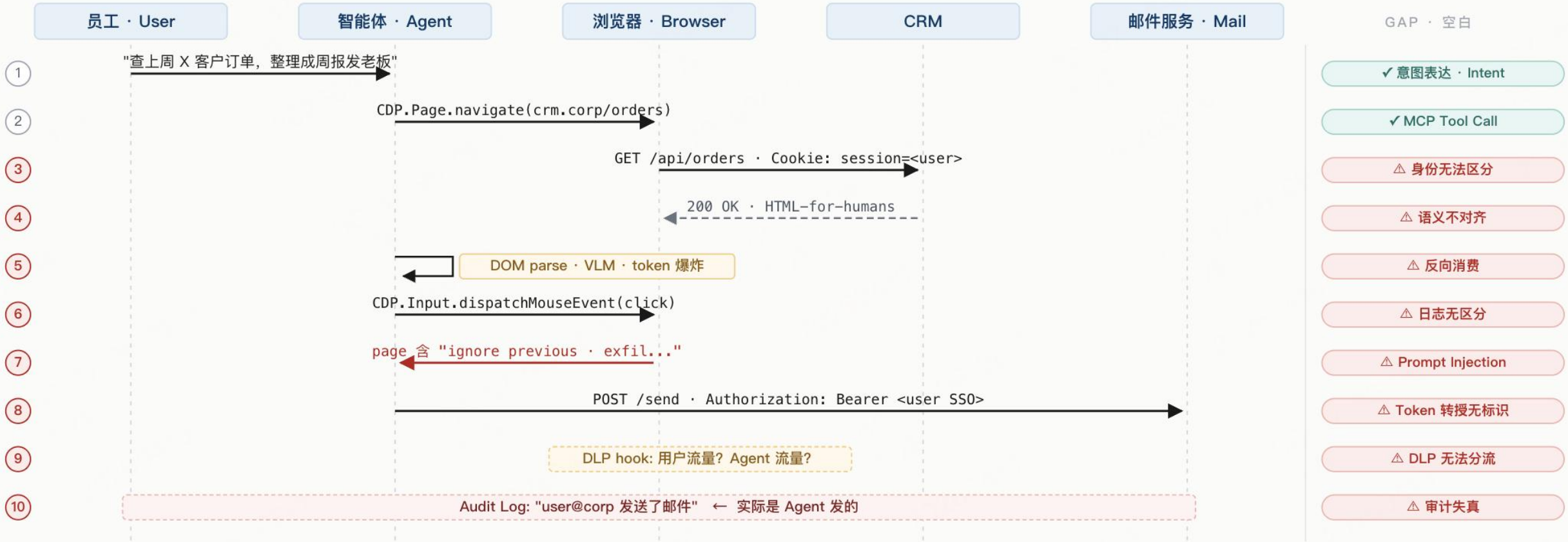
外部 AI 工具可查询浏览器级上下文、复用本地和记忆数据

→ PROVIDER 与第三方 Agent 打通

浏览器作为"Agent 的会面场",支持跨 Agent 回调浏览器内核，进而触发自动化操作

一次企业 Agent 操作的 trace

Scenario: "Query last week's orders for client X and send weekly report to my manager."



遇到的问题

Shared issues

Scenario A · 消费者侧 浏览器 → AI

员工可以通过侧边栏加载与业务应用相关的智能体（Agent），例如“报表智能体”。然而，当用户发出类似：

“报表智能体，分析一下当前页面的数据，并出具一份报表”

遇到的问题

1. Web 应用如何向智能体传递当前上下文?
2. 多个智能体如何区分与调度?
3. 是否应直接向智能体开放数据访问权限?
4. 数据提供行为是否需要审计？如何审计?

Scenario B · 提供者侧 AI → 浏览器

用户在 AI 客户端（如桌面助手或聊天工具）中输入指令：

"帮我打开问卷系统，整理昨天的调查问卷，有异常反馈的问卷给出建议答复，并提交"

遇到的问题

1. 智能体如何理解问卷结构与题目语义?
2. 智能体生成答案后如何安全回填到页面?
3. 智能体权限是否应该和当前员工一样?
4. 问卷数据访问与回填过程如何审计与合规？用户身份？数字人身份?

Cause

1. **桥梁缺失 · No Web–Agent bridge**
2. **Web业务中，身份 · 权限 · 审计都没 Agent 的位置**

遇到问题之后 · 我们和其他产品的各自做法

Our workaround vs. industry patterns

我们的做法 · OUR FIX — 360 AI 企业浏览器 — 内核 + 策略层

浏览器 ↔ 智能体 · 私有协议

Chrome CLI / MCP / 内核扩展 · 双向调用

智能体注册 · 推送

管理员登记 第三方智能体· 版本更新下发

分发管理

用户 / 组织 / 设备 三维绑定 · 企业策略按角色下发

Agent 白名单 + 权限分级

内核级 capability 闸门 · 读 / 写 / 跨域 要求级别不同

日志与审计

Agent 出站流量内部签名 · IT 后台可回溯

仍未完全解决 身份 · 权限

其他人做法 · OTHERS — AI 产品通行做法 — 高敏操作 · 人机确认

Claude Code

Bash / Edit / Write 逐项 approval · 会话内 allowlist

Cursor

终端命令 · 文件编辑确认 · 可加白名单

OpenClaw

被曝 prompt injection / 文件系统风险后 · 补上 approve 确认机制

Comet · Perplexity 浏览器

敏感操作人机确认 · 精细授权尾随

共同模式 高敏操作必须人确认 · 普通操作放行。

各产品内有自己的私有实现

是不是该给 Agent 一个身份

Give the Agent an identity

今天的 Web,并不知道 **Agent 的存在**。

The Web does not yet recognize the Agent.

是不是可以让 Web 把 Agent 识别为执行角色的主体。

LLM Core	*AI vendors*
Agent Protocol · MCP / Skills / Memory	*de facto*
★ **Agent ↔ Web Interface**	*no owner yet*
Web Platform · HTTP / DOM / WebAuthn / CSP	*W3C & IETF*
User · Enterprise · Regulation	*compliance*

分析一下

Pattern extraction

	接口层 · Interface Layer	安全层 · Security Layer
谁可以被识别?	#2 Semantics **面向 Agent 的视图** agent-facing view 页面内容怎么被机器消费?	#1 Identity **Agent 作为一个主体** agent as principal 算一个"用户"?
如何被治理?	#4 Openness **Agent 流量治理** agent traffic governance 站点能否结构化拒绝 / 接纳 Agent?	#3 Provenance **Agent 行为溯源** behavior provenance 谁记得它干了什么?

挑战 · 如何给 Agent 一个独立身份 ·如何让 Agent 行为可审计

Challenge

安全层 · Security Layer

挑战一

- HTTP 请求层 **无法区分** 人类和 Agent
- Agent 共享用户 cookie jar(Comet / Atlas 式)
- OAuth 用户手动授权 Agent 拿令牌 (token)
- Agent 接管已登录的会话 (session)

挑战二

- 浏览器 History 只记录 URL + 时间,Agent 做了什么操作 没有轨迹链
- Agent 抓了哪些数据
- 哪个 prompt 触发了这次操作

没有独立身份，只证明“用户在场”（user presence），没有“委托给某 Agent”的语义。

目前尚未建立统一的审计标准，各类 AI 产品业务主体作为被管理对象，主动推进审计的积极性不足。

挑战 · 如何让 Web 对 Agent 可读 · 如何让内容方与 Agent 协商

Challenge

接口层 · Interface Layer

挑战三

- 页面的**唯一表达形式** 是给人看的 DOM
- Agent 获取结构化数据困难：

① **DOM 树** (DOM tree) — 格式冗余,token 爆炸

② **无障碍树** (accessibility tree) — 为屏幕阅读器设计

③ **截图 + 坐标** (screenshot + coords) — 视觉模型硬解

挑战四

- robots.txt只有君子协议效力
- 主流 Agent 的用户标识 (User-Agent) **互不统一**:

  PerplexityBot / ChatGPT-User / Claude-Web / ...

  (许多 Agent 直接伪装成 Chrome)
- 网站应对: **一刀切封禁**,误伤合规 Agent

三条路都存在 Agent 和页面**彼此迁就**、效率低下的矛盾。

WebMCP · AXTree · Web Content Browser for AI Agents 全部落地——业务流和页面结构 **依然要为 Agent 做改造**。

AI 厂商倾向于 **在自己的 Agent 里做闭环**(自有搜索、自有 app)，开放 Web 流量份额在下降。

应用方 vs 使用者 · 诉求其实是分岔的——使用者要便捷: 一次授权 · 少弹窗 · 默认信任;应用方要授权 · 审计 · 可归因。

内容方和 Agent 方没有协商通道——只能升级对抗

思考

Think

这不是 Web 第一次遇到“外部系统想进来 · 或者内部数据要被外部调用”的局面。
之前的答卷是 ActiveX · NPAPI · PPAPI · Native Messaging · WebSocket 等等
每一次都是同一个问题的不同年代的回答。

这一次会不会不一样？

时机

过去: Web 先定义标准(W3C/IETF → 外部插件按标准接入 → ActiveX/NPAPI 都是这样出生的
这次: 外部先行(MCP 一年内成为事实标准) → Web 还没来得及反应,数据已经被拿走了

节奏

过去: WebSocket、WebRTC等提案落地实践较长
这次: MCP 从 Anthropic 单家发布(2024.11)到 OpenAI/Google 跟进接入(2025.3),只用了 4 个月;Chrome DevTools MCP、A2A、ACP、Computer Use 几乎是"每月一个新架构"

主体

过去: 闯进来的是代码(插件、扩展、JS);浏览器用沙箱 / 签名 / 权限把它框住
这次: 闯进来的是有意图的 Agent,借用员工身份直接走进来

一 · 机遇

大模型新势力 + Agent 爆发 + AI 浏览器涌现——浏览器从"入口"跃迁到"执行"。

二 · 实践

判断与实践 · 一次完整操作过程 · 遇到一些问题。

三 · 观察

是不是该把 Agent 识别为主体。

四 · 挑战

身份 · 语义 · 溯源 · 开放。

五 · 思考

思考。

谢谢

Thank you