

浏览器自动化Skill：Agent Browser究极攻略（含登录态解决方案）

原创 晓哥AI编程 晓哥AI编程 2026年1月30日 10:34

如果你是想 🚧

- 💡 写完代码后，让claudecode、codex，**自己使用浏览器进行UI自动化测试**
- 💡 想做网站 🌐，**却苦于人机验证码阻拦**
- 💡 用 agent-browser，**不知道如何做登录态持久化**，每次都要登录
- 💡 想了解 agent-browser 的落地场景

那么这篇文章一定适合你 🎉

前情提要

agent-browser，是Vercel官方推出的一款 Agent Friendly的浏览器工具。**并以绝对的性能优势，碾压了Playwright MCP 和 Chrome DevTools MCP**



下载&安装 🛡️

```
npm install -g agent-browser
```

```
agent-browser install
```

同时，Vercel官方推出了配套的skills: **Agent Browser Skill**

```
npx skills add vercel-labs/agent-browser
```

有了Agent Browser，就相当于是赋予了命令行，操作浏览器的能力 🤖

以上就是我们基础篇的内容 👇 这个SKILL，自动操作我的浏览器，扒出了我的网购记录 0.0! ?

晓哥也文章里谈到过：**这是迄今为止，晓哥觉得最有潜力的skill**

那么接下来，我们就一起来进一步，挖掘这个skill的能力 🐱

登录态如何持久化?

在上一篇文章中，粉丝提到了登录态持久化问题 👇

广东 1月21日 回复 5
利用截图扫码登录的功能放远程服务器上做每天自动打卡，攒积分，薅羊毛已经搞过好一阵喽。AgentBrowser SKILL 也测试过，不过登录持久化不如Playwright 和好用，很多网站都做不了，每次都要AI截图扫码太麻烦，博主能不能研究一下?
作者赞过

用过Agent Browser的小伙伴一定都会遇到这个痛点：**每次都需要重新登录**

解决方案如下 🧑

第一次先使用 --headed，也就是“有头”模式手动登录下

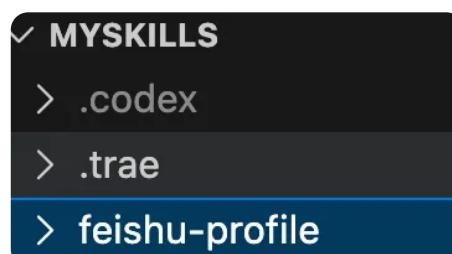
如果是在服务端环境，页可以直让Agent Browser“吐”一下登录二维码，方法见👉 这个SKILL，自动操作我的浏览器，扒出了我的网购记录0.0！？

无论有头，还是没头，都需要先手动登录下

为了进行更加直观的演示，这里晓哥用“有头”模式，并先不急用skill，直接用cli命令进行演示



首先，先随便建一个空文件夹，作为我们持久化登录态信息存储的目录



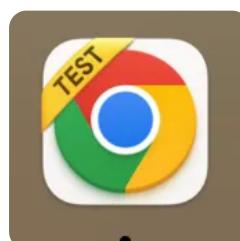
然后最重要的一步 !!

agent-browser --profile 你的目录路径 open 你的链接 --headed

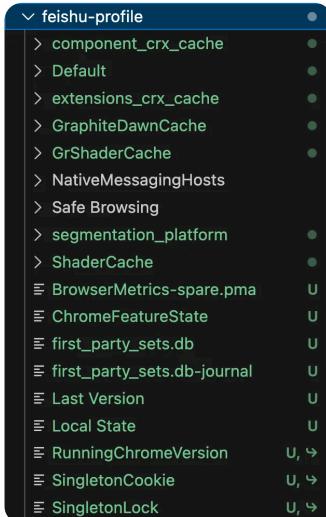
```
GP47R194J9:myskills [redacted] $ agent-browser --profile '/Users/[redacted]/Desktop/myskills/feishu-profile' open https://my.feishu.cn/docx/Vu0udUE[redacted]UGY1ctjnonPe --headed
✓ 飞书 - 登录
https://accounts.feishu.cn/accounts/page/login?app_id=2&no_trap=1&query_scope=all&redirect_uri=https%3A%2F%2Fmy.feishu.cn%2Fdocx%2FVu0udUElKod9LqxUGY1ctjnonPe%3Flogin_redirect_times%3D1
```

👉 可以看到现在打开的页面，由于没有登录态，所以跳到了登录界面

--headed 会呼起浏览器，然后正常手动登录就好 👇



登录成功后，可以看到，原本空空的目录下，多出了很多文件



👉 这就是我们的持久登录态了 😊

下次想要登录，只需要再执行下（这次就不用加 `--headed` 了）

```
agent-browser --profile 你的目录路径 open 你的链接
```

可以看到，我们在关闭了chrome实例的前提下，依旧顺利的登录飞书，拿到了文档中的内容



```
GP47R194J9:myskills [~] $ agent-browser close
✓ Browser closed
GP47R194J9:myskills [~] $ agent-browser --profile '/Users/[REDACTED]/Desktop/myskills/feishu
profile' open https://my.feishu.cn/docx/Vu0udU1ctjnonPe
浏览器自动化 Skill: Agent Browser究极攻略 (含登
态解决方案) - 飞书云文档
https://my.feishu.cn/docx/Vu0udU1ctjnonPe
```

只要登录态不过期，这个profile文件就可以一直用



最佳实践（含人机验证码策略）

相信在做 🐛 的时候，被人机验证拦截，也是一个让人很头疼的点 😠



Agent Browser，可以完美的帮我们解决这个问题 🙌

👉 先用Agent Browser Skill，让Agent熟悉我们目标站点的页面结构

👉 再让Agent Browser Skill，写下 🐛 脚本。注意：骚操作来了 😳 👇

跟Agent说：“如果遇到captcha，则启动 --headed 模式，我来输入验证码”

脚本里就会多出这么一段，大致意思就是：遇到人机验证码，报异常退出，并启动Agent Broswer的“有头”模式

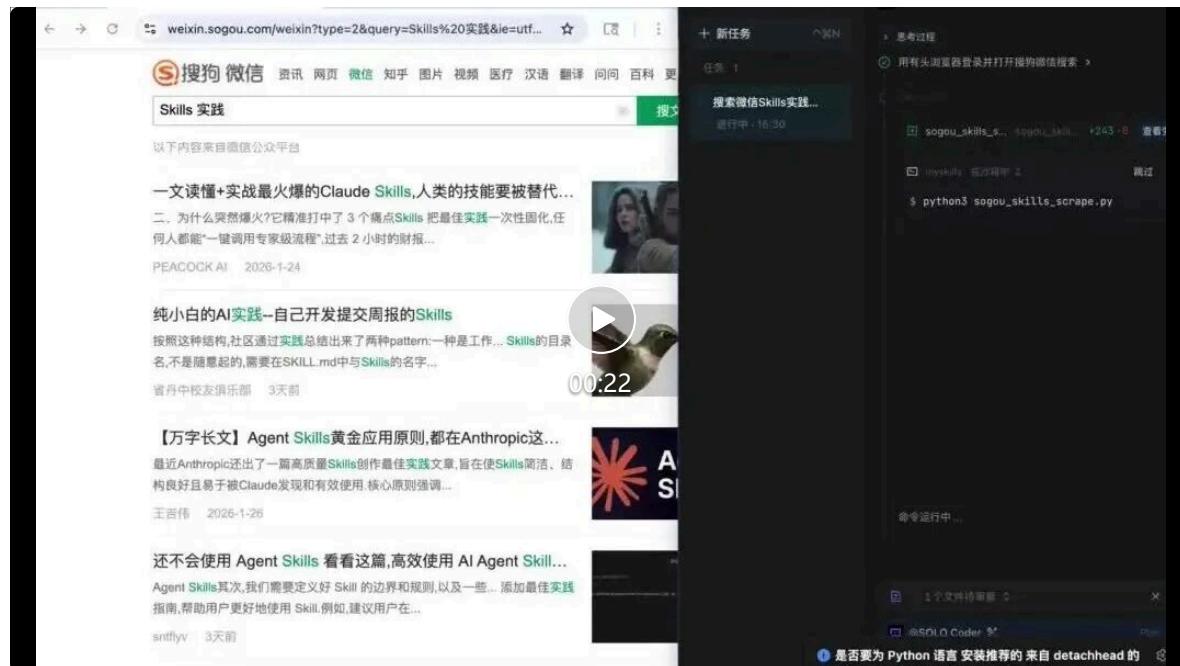
```

try:
    ensure_tab(results_tab)
    # We'll navigate by directly opening each result page (more reliable than clicking pager).
    open_search_page(args.query, args.start_page)
except CaptchaRequired as e:
    print(str(e), file=sys.stderr)
    print(
        "Please solve the captcha in the headed browser window, then rerun the same command."
        "Already collected lines remain in the output file.",
        file=sys.stderr,
    )
raise SystemExit(2)

```

这样我们就可以通过人工的方式，来帮助AI过验证码了 😊

或者一开始就指使启动 --headed 模式也是可以的，这样其实跟我们普通刷网页是一样的，只不过现在是Agent代替我们来看网页，如下 👇



👉 我全程无操作，AI自己在刷网页，非常的鬼畜 🤖

UI自动化验收

ClaudeCode之父Boris，提出的13个让AI编程效率起飞习惯中的，最重要的习惯：给Agent一个能够自我验证的途径👉 让AI编程效率提升2-3倍的秘密，Claude Code创始人的13个习惯



Boris Cherny ✅ @bcherny · Jan 3

···

13/ 最后一个小贴士：要让 Claude Code 取得出色的效果——给 Claude 一个验证其工作的方法。如果 Claude 有这个反馈循环，最终结果的质量将提升 2-3 倍。

Claude 会测试我提交的每一个改动来claude.ai/code 使用 Claude Chrome 扩展。它会打开一个浏览器，测试界面，并不断迭代，直到代码能正常工作且用户体验感觉良好。

不同领域中的验证方式不同。它可能简单到运行一个 bash 命令，或者运行一个测试套件，或者在浏览器或手机模拟器中测试应用。务必投入资源确保这一点非常稳固。

code.claude.com/docs/en/chrome

翻译一下就是：**给AI一个查看代码效果的途径**：如果是服务器代码，提供一个启动服务的方法；如果是网页代码，提供一个查看和交互 UI 的方法等

Agent Browser，绝对是UI自动化验收的不二之选：相当于是直接赋予了claudecode、codex等AI编程工具，感知和操作浏览器的能力 😊

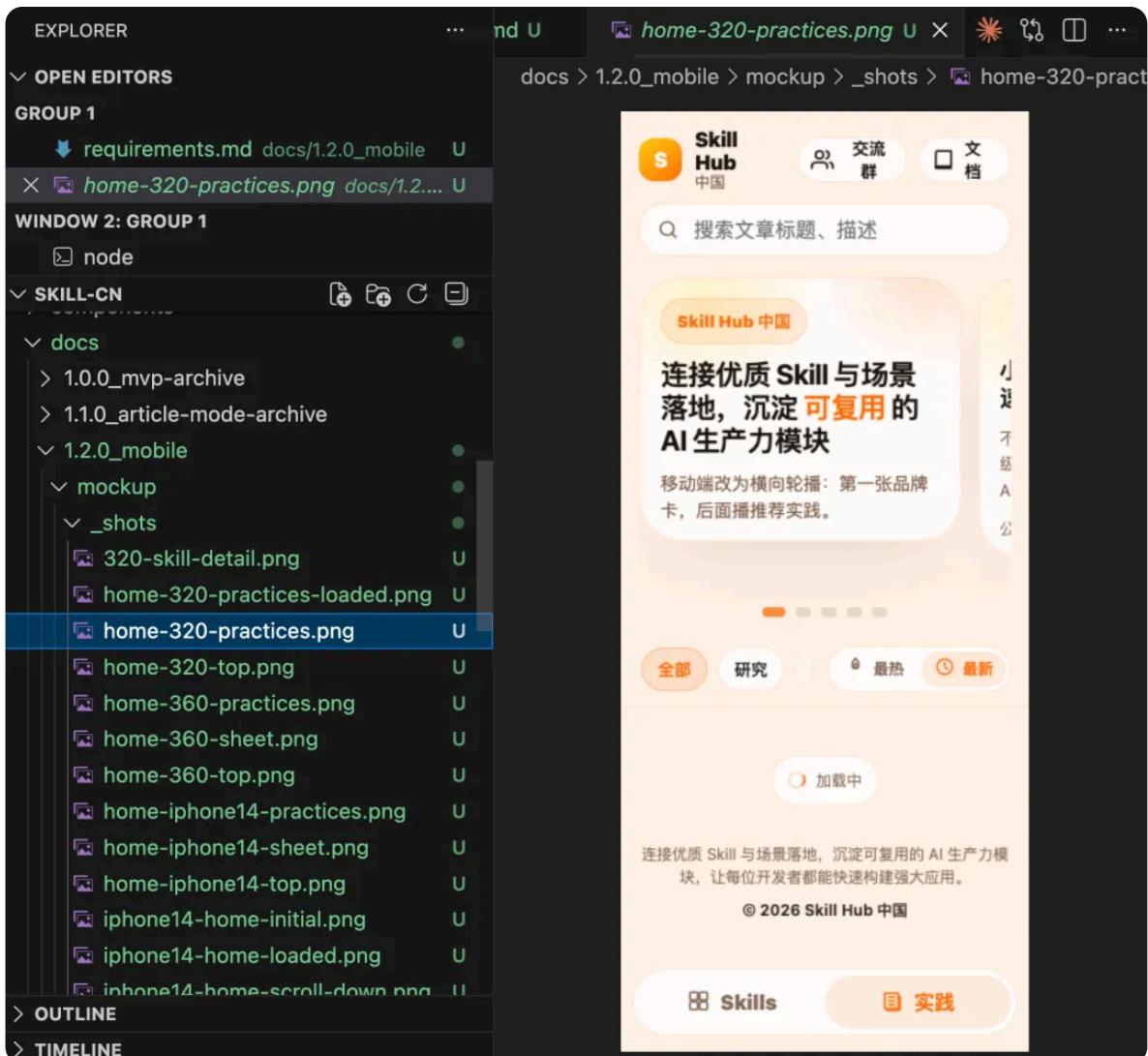
下面这个例子，**就是晓哥让Codex自己进行视觉验收**：这里同时调用了两个skill (ui-ux-pro-max可见👉 回小白如何使用AI编程，快速打造商用级UI)

“站在顶级移动端网站UI/UX设计师的角度，找出并罗列不合理的点，并进行优化”

> \$agent-browser \$ui-ux-pro-max

站在顶级的移动端网站UI/UX设计师的角度，找出所有目前设计不合理的点，罗列并给出优化方案

Agent Browser Skill会对项目界面进行验收和截图 👉



然后给出优化建议

- Worked for 18m 25s
- 下面基于我用 `$agent-browser` 在 iPhone 14 / 360px / 320px 下把首页 (skills/practices)、Bottom Sheet、详情页都了一遍后的结论；按“问题点 → 优化方案”列出。
- P0 (必须改，不然移动端会显得很业余/不顺手)**
 - Header 在 320px 会“挤爆” (Logo/按钮文案换行、按钮变形)
 - 方案：Header 右侧入口改为「纯图标 + 轻提示」或「更多(...)'菜单」；“交流群/文档”不要在按钮内显示长文案（宽度很容易断行）。
 - 方案：Logo 文案在 320 下简化为单行（例如只保留“Skill Hub”，把“中国”隐藏或移到 Hero）。
 - 点击热区偏小 (不适合移动端 44px 触控规范)
 - 现状：分类 chip 34px、排序按钮 28px、header icon-btn 36px（都偏小）。
 - 方案：移动端把这些控件的可点击高度统一提升到 ≥44px（不一定变“更高”，可以用 padding 扩大 hit area）。
 - Bottom Sheet 面板不够“干净” (透底导致底部模式切换条/背景元素穿透)
 - 现状：Sheet panel 用了半透明白，底部 Segmented 会隐约透出来，信息层级变脏。
 - 方案：Sheet panel 改为纯色不透明 (#fff 或 —surface-solid)，遮罩稍微加深一点；打开 Sheet 时强制隐藏/底部模式切换条（避免“两个浮层抢焦点”）。
 - 列表首屏加载体验太“空” (只看到一个“加载中”胶囊 + footer 文案)
 - 方案：首屏加载用 3~5 张骨架卡片占位（与最终卡片高度一致），而不是单个 loader；footer 文案只在真实滚动到底部再出现（或仅在“已到底”时出现）。
 - “实践卡点按弹出 ActionSheet”缺少可发现性提示
 - 现状：卡片看起来像“可点去详情”，但点了是弹操作面板。
 - 方案：在卡片右上角加一个轻量的「.../操作」图标（不改变“点任意区域也能弹”的口径，只是增加提示）；或给卡加更明确的 pressed 状态（背景轻变深/高光一闪）而不仅仅是 translateY(1px)。
- P1 (强烈建议改，能明显提升高级感)**
 - 顶部筛选区 (标签 + 排序) 同一行在小屏会拥挤
 - 方案：保持吸顶不变，但在 360/320 下把排序收敛为一个小的 sort chip（点开小弹层/二选一），或把排序放到第二行，让层级更清爽。
 - 底部模式切换条体积偏大、抢视线，还会遮挡列表信息
 - 方案：缩窄高度/阴影强度；仅激活态显示“图标+文字”，非激活态仅图标；或把它做成更薄的悬浮条（减少占屏）。
 - 回到顶部按钮位置容易压住卡片内容/与底部条打架
 - 方案：与底部模式条做“互斥显示”：模式条出现时，回到顶部按钮上移/隐藏；或把回到顶部按钮放在左侧（避免遮挡阅读量等信息区域）。

并优化

```
562     state.loading = true;
510 -     setFooterState({ type: "loading" });
563 +     // 首屏已用骨架屏表达加载，底部不再重复显示 loader; 后续页才展示
564 +     if (state.page > 1) setFooterState({ type: "loading" });
565     :
572         state.error = true;
573 +         // 避免骨架一直挂着：失败时清掉骨架，让“失败重试”更明确
574 +         if (state.page === 1) clearSkeleton();
575         setFooterState({ type: "error" });
576     :
593
594 +     // 首屏真实数据到达：先清骨架，再渲染真实卡片（避免 DOM 堆叠）
595 +     if (state.page === 1) clearSkeleton();
596 +
597     slice.forEach((item) => {
• Edited docs/1.2.0_mobile/mockup/app.js (+4 -0)
930     icon(`<path d="M4 4h14v16H4z"></path><path d="M18 8h2v12a2 2 0 0 1-2 2H6"></path><path d="M8 8h6"></path><path d="M8 12h6"></path><path d="M8 16h6"></path>`);
931 +     const iconMore = () =>
932 +         // “更多/操作提示”图标：用于移动端实践卡右上角，增加可发现性
933 +         icon(`<circle cx="12" cy="5" r="1.5"></circle><circle cx="12" cy="12" r="1.5"></circle><circle cx="12" cy="19" r="1.5"></circle>`);
934     :
954     iconNews,
955 +     iconMore,
956 });

• Ran bash -lc 'agent-browser open file:///Users/bytedance/Desktop/skill-cn/docs/1.2.0_mobile/mockup/index.html'
└ ✓ Skill Hub 中国 · Mobile Mockup
    file:///Users/bytedance/Desktop/skill-cn/docs/1.2.0_mobile/mockup/index.html

• Working (16m 09s * esc to interrupt)

> Write tests for @filename
74% context left · ? for shortcuts
```

优化后，会再次通过Agent Browser Skill进行验收 🎉

```
- Worked for 23m 33s
-
• 我按 ui-ux-pro-max 的移动端可用性口径，把 mockup 里提到的 P0 问题先落地修了一轮，并用 agent-browser 在 320 / 360 / iPhone14 下回归截图验证了。
    已改动（主要在 mockup 层）：
    - 320px Header 不再“挤爆”：<=360px 自动隐藏“交流群/文档”文案，仅保留图标（aria-label 保留可访问性）
    - 全站关键触控热区对齐 44px：Header 按钮 / 搜索框 / Chip / ids-chip / 底部模式切换 / 排序组整体热区
    - 首屏不再“太空”：第一页加载先出 3 张骨架卡（后续页仍用底部加载中），失败时清骨架并显示“加载失败 + 重试”
    - 实践卡“可操作”提示增强：右上角新增 “::” 提示图标（不增加额外按钮）
    - Bottom Sheet 视觉更干净：遮罩加深、面板改纯白；打开 Sheet 时隐藏底部 mode bar / 回到顶部，避免穿透脏感
    - prefers-reduced-motion 下禁用骨架动画等动效（只保留必要状态变化）

    文件：
    - docs/1.2.0_mobile/mockup/styles.css
    - docs/1.2.0_mobile/mockup/app.js

    新截图（可直接对比旧图）：
    - docs/1.2.0_mobile/mockup/_shots/home-320-top-v2.png
    - docs/1.2.0_mobile/mockup/_shots/home-360-top-v2.png
    - docs/1.2.0_mobile/mockup/_shots/home-iphone14-top-v2.png
    - docs/1.2.0_mobile/mockup/_shots/home-320-skeleton-v2.png
    - docs/1.2.0_mobile/mockup/_shots/home-360-sheet-v2.png

    你先看下这版是否更“高级/顺滑”，如果认可我再继续把剩余 P1（比如底部模式条占位/信息密度进一步收敛策略）也一并改掉。
```

一个小坑

如果大家要起多个agent browser指令，记得使用 `--session` 将不同会话区分开，不然页面会被“顶号” 🚫

（骁哥开始起了两个agent browser，一个爬搜狗，一个验项目UI。就因为没区分 `session`，结果发现最后验的是搜狗的UI。。。 😅）

小结

晓哥觉得agent browser是可以成为playwright完美平替的，只有一个场景：playwright mcp有一个chrome插件，可以直接复用用户浏览器的登录态，这个agent browser暂时还没有比较好的方案

还是那句话，操作浏览器，想象空间很大，大家也可以一起脑爆一下，晓哥这里举几个场景
抛砖引玉

比如可以自动打卡，薅羊毛

晓哥打算把之前所有的公众号文章，都通过agent browser给Agent看，让AI分析晓哥的风格，以后帮晓哥写点文章...

那本期就这样，我是AI编程晓哥，跟上我，一起在AI时代共同进步！

往期Skill实践

AI编程可以直接生成视频了？Remotion Skill 初体验

0基础小白建站：Supabase免费数据库手把手教学

这个SKILL，自动操作我的浏览器，扒出了我的网购记录0.0！？

自媒体取标题太耗时间？我把它做成了一个SKILL

用AI独立开发产品，域名想不出来就用这招！

为什么你用AI做的产品总是不满意？因为跳过了这个环节

小白如何使用AI编程，快速打造商用级UI

如何在Cursor、TRADE中，用Skills 3分钟制作精美PPT

给8岁宝宝的Agent Skill入门教程

骁哥 AI 编程

与我一起，用 AI 编程创造全新可能



技术专家

现任某 T0 大厂资深技术专家

专注领域

AI 编程 / Agent / 副业 / 自媒体



xiaoge_aibreaker

欢迎+V，一起成长～