How Unify Built AI Research Agents at Scale



Subscribe















Unify scaled their AI research agents to handle 36 billion tokens monthly - here's what Connor Heggie (CTO) and Kunal Rai (Engineer) learned along the way. From initial ReAct framework experiments to production deployment across thousands of companies, they share real insights on model performance, cost reduction strategies, and building tools that actually work at scale.

Watch all of our recorded sessions from Interrupt here: https://interrupt.langchain.com/video...

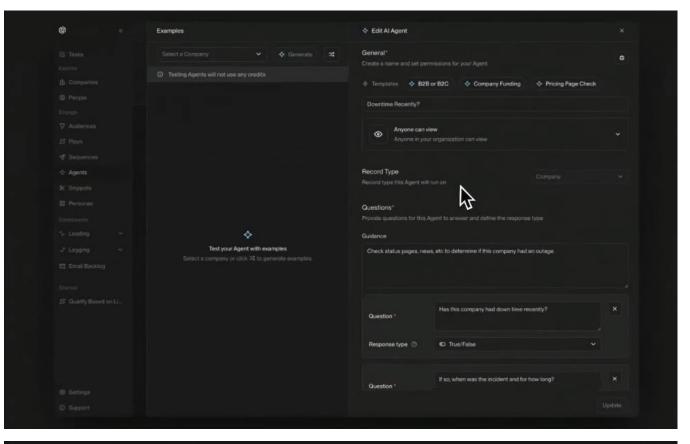
Learn more about LangGraph Platform: https://www.langchain.com/langgraph-p..

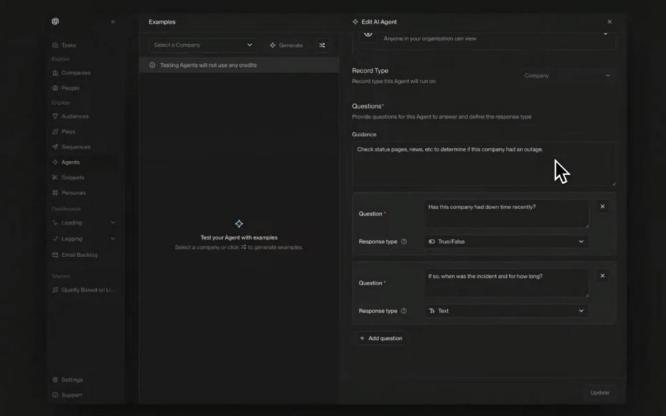
BREAKTHROUGH AGENTS: LEARNINGS FROM BUILDING AI RESEARCH AGENTS **o** unify

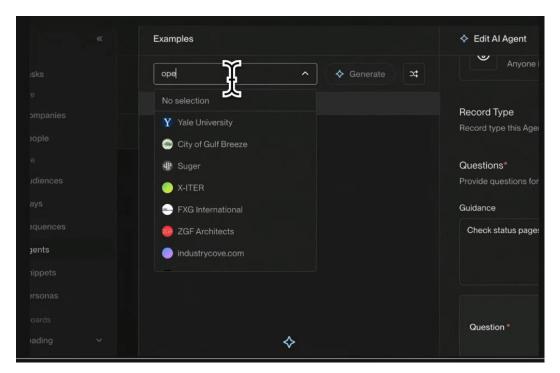
Unify

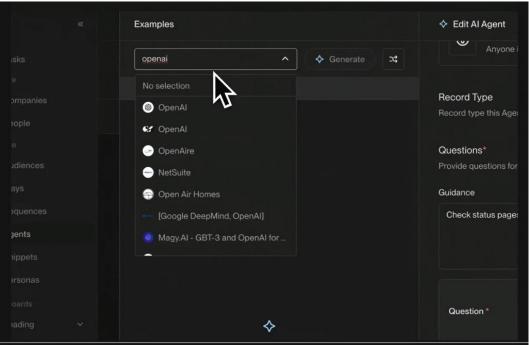
Our mission is to make growth a science and help the best products win.

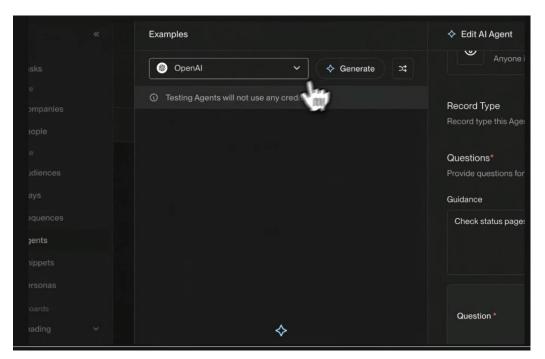
- We treat go-to-market as a search problem:
 - Find people and companies that have a problem your product uniquely solves
- To find those people and companies we run Al research agents

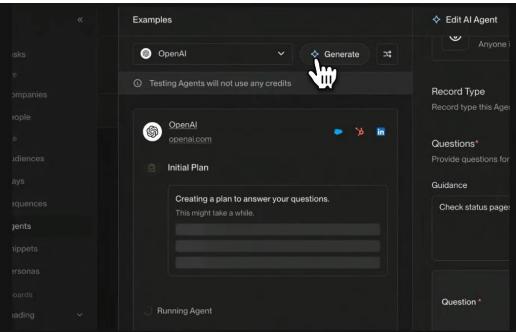


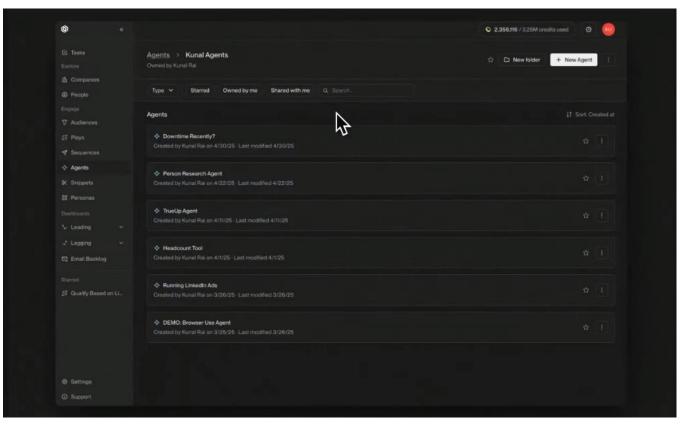


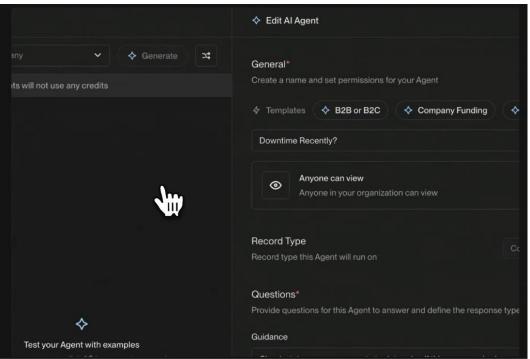


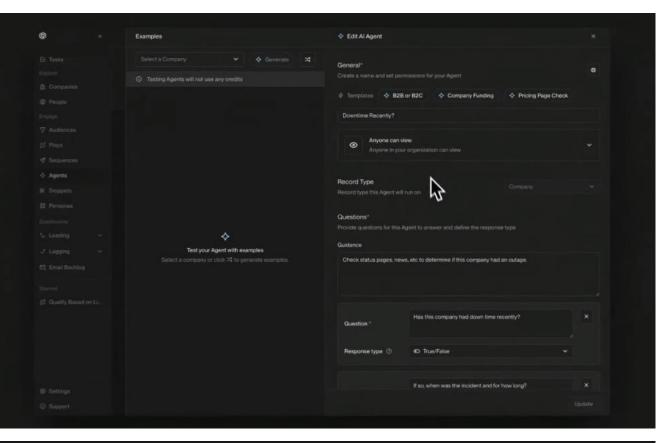


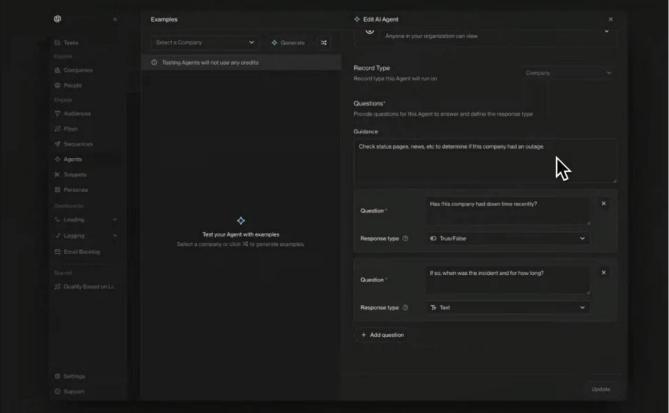


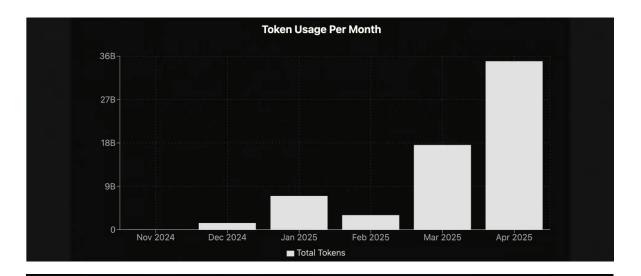








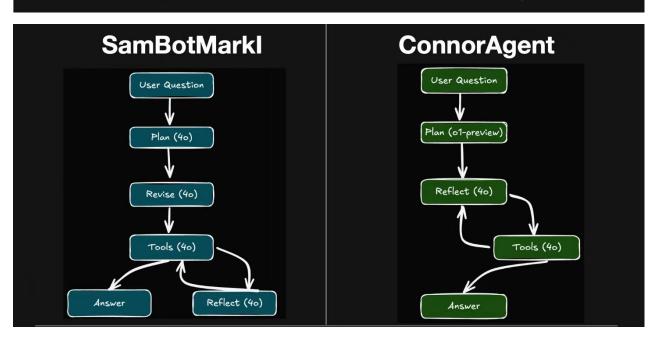




Some learnings from building and running ger ralized research agents at scale

Our V1s: Sam vs Connor

- ReAct = "reason then act" agent framework
 - Allows for corrective and reactive actions over the course of a research path
- Connor and Sam both built versions of this framework
 - Three tools: search the internet, search a website, and scrape a website

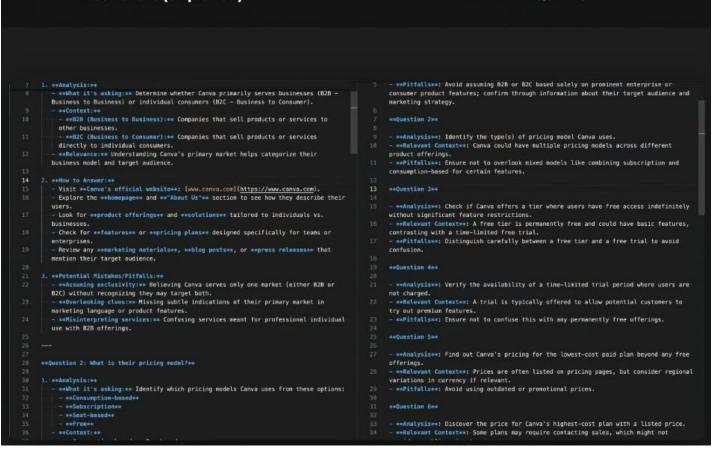


How do we know which is better?

- Before we built any evals or metrics we spent a long time looking through trajectories manually
 - vibe check
- We found that o1 produced much more thorough research plans

1600 Tokens (o1-preview)

600 Tokens (gpt-4o)



600 Tokens (gpt-4o)

```
**Question 2: What is their pricing model?**
                                                                                                                              **Relevant Context**: Prices are often listed on pricing pages, but consider regional
                                                                                                                           variations in currency if relevant.

***Pitfalls**: Avoid using outdated or promotional prices.
1. **Analysis:**
     - **Consumption-based**
- **Subscription**
      - **Seat-based**
                                                                                                                            - **Analysis**: Discover the price for Canva's highest-cost plan with a listed price.
                                                                                                                           - **Relevant Context**: Some plans may require contacting sales, which might not provide a public price to assess.
    - seContext: 4s
       - **Consumption-based:** Pay based on usage.
      - **Subscription:** Pay a recurring fee for ongoing access.
- **Seat-based:** Pay per user or per seat.
- **Free:** Offers services at no cost indefinitely.
                                                                                                                             **Pitfalls**: Do not count plans that require contacting sales unless a specific
2. **How to Answer:**
      Navigate to **Canva's pricing page**: [www.canva.com/pricing](https://www.canva.
                                                                                                                           contact sales for pricing.
    **Relevant Context**: Enterprise plans often come with customized features and
   com/pricing).
      Examine each **plan** available:

- **Free Plan:** Determine if they offer a free tier.
                                                                                                                           annual payments).
                                                                                                                            enterprise pricing.
   - **Enterprise Plan:** Look for **seat-based** pricing details.

- Read the **features and pricing details** to see how users are billed.

- Note any **additional costs** or **consumption-based fees** (e.g., paying for

    **Website Search**: Start by searching Canva's website for terms like "B2B", "B2C",
"target audience", and "pricing". This can quickly yield relevant results for various

   premium images).
   - **Rissing multiple models:** Failing to recognize that Canva may use a combination of pricing models.
                                                                                                                            2. **Internet Search**: Conduct Google searches using queries like "Canva B2B or B2C",
       **Confusing terms:** Misunderstanding the difference between seat-based and
                                                                                                                            "Canva pricing model", "Canva free tier vs free trial", and "Canva enterprise pricing". This will provide supplemental information and sometimes direct access to press
    - ##Overlooking small print: ** Not noticing additional consumption-based charges for
                                                                                                                            releases and interviews where Canva discusses their strategy.
                                                                                                                            3. **Scrape Wabpages**: Scrape Canva's pricing and product/service pages. This will help confirm the different tiers, associated features, and pricing models.
                                                                                                                               ⇒Evaluate Data+*: Analyze the collected information to answer each question
                                                                                                                            accurately:
                                                                                                                                - Determine the primary market focus (828 or 820).
- Identify all pricing models: consumption, subscription, etc.
      **What it's asking:** Confirm if Canva offers a free plan that users can access
    indefinitely without payment.
                                                                                                                                - Confirm the presence of a free tier and/or free trial.
```

1600 Tokens (o1-preview)

600 Tokens (gpt-4o)

	Cuestion 7
2. **How to Answer:**	
- Navigate to **Cenva's pricing page**: [www.canva.com/pricing](https://www.canva	 - **Analysis**: Determine if there is an enterprise plan that requires customers to contact sales for pricing.
- Examine each **plan** available:	- **Relevant Context**: Enterprise plans often come with customized features and
- **Free Plan: ** Determine if they offer a free tier.	pricing.
- **Pro and Team Plans:** Check if these are **subscription-based** (monthly or annual payments).	
- **Enterprise Plan:** Look for **seat-based** pricing details.	
- Read the **features and pricing details** to see how users are billed.	**Overall Plan**
- Note any **additional costs** or **consumption-based fees** (e.g., paying for	
prenium images).	"target audience", and "pricing". This can quickly yield relevant results for variou
3. **Potential Mistakes/Pitfalls:**	questions.
Missing multiple models: ** Failing to recognize that Canva may use a combination	
of pricing models.	2. **Internet Search**: Conduct Google searches using queries like "Canva B2B or B2C
- **Confusing terms:** Misunderstanding the difference between seat-based and	"Canva pricing model", "Canva free tier vs free trial", and "Canva enterprise pricin
subscription models.	This will provide supplemental information and sometimes direct access to press
- **Overlooking small print:** Not noticing additional consumption-based charges for	releases and interviews where Canva discusses their strategy.
certain features.	
	help confirm the different tiers, associated features, and pricing models.
Question 3: Do they have a free tier?	accurately:
1. **Analysis:**	- Determine the primary market focus (B2B or B2C).
- **Mhat it's asking:** Confirm if Canva offers a free plan that users can access	 Identify all pricing models: consumption, subscription, etc.
indefinitely without payment.	- Confirm the presence of a free tier and/or free trial.
Context: A free tier provides ongoing access without a time limit, unlike a free trial.	plans.
	 Verify if any enterprise plans require contacting sales for pricing.
2. **How to Answer:**	
- Check the **pricing page** for a plan labeled ***Free*** or similar Verify that the free plan does not have a **time restriction**.	 **Conclude and Send Response**: Compile the findings and send the final answers to the user, ensuring clarity and completeness of the responses.
- Review the **features** included in the free tier to ensure it's a functional	the user, ensuring coarsty and completeness of the responses.
offering.	By following this systematic approach, the assistant should efficiently gather accur-
diterally.	my following this systematic approach, the assistant should efficiently gather accur- information to answer all questions regarding Canva's business and pricing strategie:
3. **Potential Mistakes/Pitfalls:**	information to answer all questions regarding canva's business and pricing strategie
J. **Potential Mistakes/Fittalls:** - **Confusing with free trial:** Mistaking a time-limited trial for a free tier.	
- **Assuming limitations:** Believing the free tier is too limited without verifying	
its capabilities.	

o1-preview question 4

Question 4

Description: *Do they have a free trial?*

Value Type: BOOLEAN

1. Brief Analysis - A **free trial** is a time-limited period during which users can access services for free. - After the trial, payment is typically required to continue using the service.

2. How to Answer - **Check for Free Trial Offers**: - Look for banners or notices on the website offering a free trial. - **Review Signup Options**: - During the account creation process, see if Canva provides trial credits or time-bound access. - **Read Terms and Conditions**: -Verify the duration and limitations of the free trial.

3. Potential Mistakes or Pitfalls - **Assuming No Trial Without Checking**: Concluding there is no free trial without thoroughly investigating. - **Overlooking Registration Requirements**: Not noting if a credit card is required to start the free trial. - **Confusing Free Tier with Free Trial**: Mistaking ongoing free services for a temporary free trial.

gpt-4o question 4

Question 4

- **Analysis**: Determine if Canva offers a time-limited free access to their services termed as a free trial.
- **Relevant Context**: Free trials are time-bound and a common way to allow new users to test services.
 - **Pitfalls**: Do not confuse free trials with unrestricted free tiers.

Building actual evals

- **Accuracy**: percentage of questions answered correctly by the agent
 - Hand labeled a set of 100 companies for each eval
- We built initial eval sets for expected use cases
 - Business Model (ex. "Is this company B2B or B2C")
 - Firmographics (Find the headquarters, company size, operating locations)
 - Technographics (What technologies does this company use?)

Showdown: ConnorAgent vs SamBotMarkI

	Industry Classification	Firmographics	Account Qualification	Technographics
ConnorAgent	86	48	58	81
SamBotMarkI	75	38	59	72

Learnings from initial evals

- Reasoning models can have an outsized impact on downstream actions for an agent
 - o1-preview was worth the latency and cost to pre-plan
- Accuracy based evaluations were great for establishing a baseline
- No clear insights into how to improve the agents from accuracy based evals alone

Where do we go from here?

- Change the graph of the Agent? No
- Change the models and prompts? Yes
- Add more tools? Yes

Model and Prompt Changes

Model Changes: Performance & Cost

Problem:

o1/o3/o1-pro are \$\$\$ and slow

Solution:

- Try different models and providers to see if we can match performance
- No other models matched or exceeded o1 until gpt4.1

Outcome:

Agent Run with o1: ~\$0.35, Agent Run with gpt4.1: ~\$0.10

Model Changes: Performance & Cost

Model	Industry Classification	Firmographics	Account Qualification	Technographics
о3	0.98	0.40	0.58	0.80
Deepseek R1	0.97	0.34	0.58	0.81
GPT-4.1	0.98	0.40	0.57	0.75
Claude 3.7	0.98	0.32	0.57	0.75
o1	0.98	0.31	0.54	0.78
Gemini-2.5	0.98	0.27	0.57	0.78

Prompt Changes: Date Formatting

Problem:

The agent would incorrectly compare dates unless they were in the exact same format example: 5/14/2025 3:00PM PT vs May 15th 2024

Solution:

Adjust the date format and inject current date into all node prompts

Outcome:

Agents now process temporal data accurately

Prompt Changes: Tool Calling

Problem:

Agents make "throwaway" agent tool calls

Solution:

Change Pydantic models for tools to accept specific configuration settings

Outcome:

Agents select the right tools more frequently and can surface more relevant information with more precise tool inputs

Prompt & Model Change Learnings

- Agent cost from swapping o1 to 4.1 with no quality change
 - Just wait for a new model to make everything better
- There are tons of edge cases (like the date issues) that are difficult to find through metrics
 - Vibe checking is still a top evaluation methodology
- Models tend to spike in different use cases and system wide evaluations don't highlight differences
- Increasing tool variety = more information is pulled into context → context management is critical

What tools to build?

2 criteria to consider for new tools:

- Is there a customer play this could power on day 1 of launching?
- Will this generalize to many customers' future use cases?

We wanted to build the tools that would be step function improvements for everyone using the Unify agent

Tools we built

- Deep Internet Search: upgrade to how we search the internet
- Browser Access: use Chrome to access websites that aren't scrapable
- Search HTML*: ability to scrape and parse raw HTML
- Dataset Access*: provisioning access to internal data feeds

*we don't have slides on these tools, but happy to chat about them after if folks are interested

Deep Internet Search: Internet search is hard

- The internet has all data, but surfacing the right data for entity specific queries is a challenge (i.e. "what is the privacy policy of samllminference.ai")
 - Google deal with SEO
 - OpenAl / Perplexity introduce an uncontrollable variable with Al summarization of search results
- Models were looking at google search previews as a source of truth
- Agents should be able to reproduce human research trajectories
 - the trajectories you used to do before ChatGPT

Deep Internet Search: Upgrading from Google

```
class InternetSearchInput(BaseModel):
      "Input schema for the internet search tool,"""
   -query: str = Field(description="Search terms")
 · · · category: · Optional [ExaCategory] · = · Field(
     · · · default=None.
   ····description="Category to search within",
 · · · livecrawl: · Optional[bool] · = · Field(
  · · · · · default=False
  -----description="If this search requires time-sensitive data or recent data",
 ---includes_phrase: Optional[str] = Field(
  · · · · · · default=None.
 .....description="Return results that contain a specific phrase".
 ... constrain by domain: Optional[str] = Field(
  · · · · · · default=None.
  ·····description="Limit-results-to-a-specific-domain".
  published date: Optional[datetime] = Field(
   ····default=None,
   ····description="Return results published after this date".
```

Deep Internet Search: Upgrading from Google

```
before
."title": . "Canva · Pricing: · The · Complete · Guide · in · 2024 · - · Ecomm · Design",
 "url": "<a href="https://ecomm.design/canva-pricing"">https://ecomm.design/canva-pricing</a>",
"preview": · "Canva·can·be·accessed·for·free, ·or·the·Canva·Pro·plan·will·cost·$14.99
per-month-or-$119-per-year. Canva-for-Teams-has-variable-pricing...
·"title": · "How · Much · is · Canva · Pro: · Canva · Pricing · & · Plans · 2025",
"url": 'https://thimpress.com/Design",
               We'll explore the features you get with Pro, its affordable cost (only
 $13 a · month!), · and · the · 30-day · free · trial · to · see · if · it's · the · right · fit · for...
·"title": ·"How · Much · Does · Canva · Cost · Per · Month · in · 2024?",
·"url": ·"https://hubdigitalcontent.com/tools/how-much-does-canva-cost"
·"preview": "Canva Pro costs $15 a month, If you get an annual subscription, you
save 34% and pay only $120 per year. Canva for Teams costs $10 per user per month.
·"title": · "Canva · Pricing · Guide · 2025: · Free, · Pro, · Teams · & · More",
\cdot "url" : \cdot "\underline{https://www.stockphotosecrets.com/buyers-guide/canva-pricing-guide}",
             ·"Canva·Pro, ·the·paid·option·for·individuals, ·goes·for·between·$10·and
 $15 monthly and is packed with value. Here is what you get with your membership:"
```

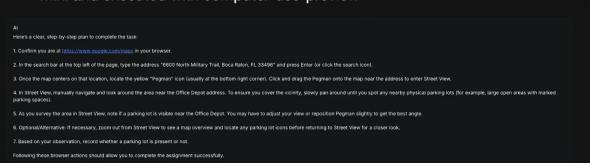
after ·"title": ·"Compare ·Free, ·Pro, ·Teams · and ·Enterprise · plans", "url": "https://www.canva.com/en/pricing/", "text": "Calculate for your team Save from 16% with yearly Design anything and bring your ideas to life. No cost, just creativity. Features you'll love: Easy drag-and-drop editor 2M+ professionally-designed templates 1000+ design types (social-posts, 'docs, 'presentations, 'sheets', 'and 'more) '4.5M+-stock 'photos' and graphics 'AI 'creative partner for 'design, 'image and 'text 'generation' 'Build' $interactive \cdot designs \cdot and \cdot experiences \cdot with \cdot AI^{*+} \cdot Bring \cdot multiple \cdot formats \cdot into \cdot one$ design* Design printing and delivery Canva Pro Unlock premium content, more powerful design tools, and AI features. Everything in Free, plus: Unlimited access to premium templates 140M+ photos, videos, graphics, audio 1000 Brand Kits to manage your brand Quickly resize and translate designs Remove backgrounds in a click Create, resize, translate and remove backgrounds at scale*-Boost-creativity-and-productivity-with-20+-AI-tools^-Plan-and-schedule social-content-Canva-Teams-Transform-teamwork, grow-your-brand, and simplifyworkflows. //year-per-person-Minimum-3-people-(US\$300/year)-Everything-in plus: Scale your brand and centralize assets Ensure brand consistency with approvals Edit, comment, and collaborate in real time Generate on brand text and-visual-assets-with-AI^-Team-reports-and-insights-Import-and-sync-data-from-your-favourite-external-sources--ISO-27001-certification-and-SOC-2-Type-IIcompliance · Canva · Enterprise · Empower · your · organization · with · an · all—in—one workplace solution. Get in touch to learn more Everything in Teams, plus: Centrally manage multiple teams and brands Set individual or group design approvals · Control · access · to · Canva · s · content · and · templates · Indemnification · for AI output* Advanced security and admin controls (including AI permissions) SCIM license provisioning Custom apps and integrations Priority customer support Design support from Enterprise Studio creative team* Dedicated customer success

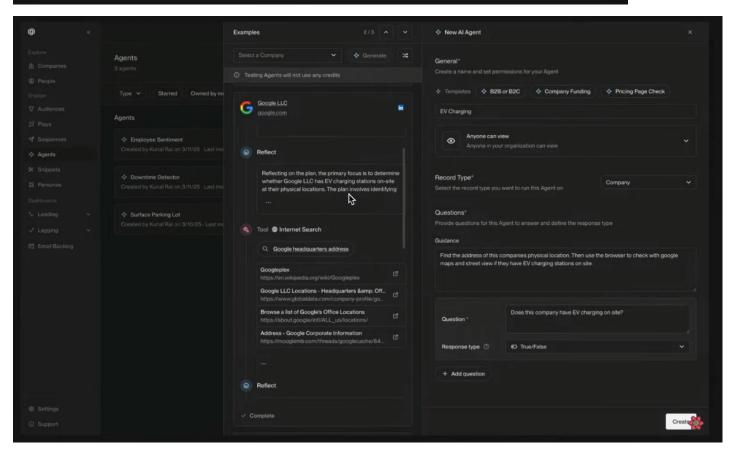
Browser Access: How do humans do it (again)?

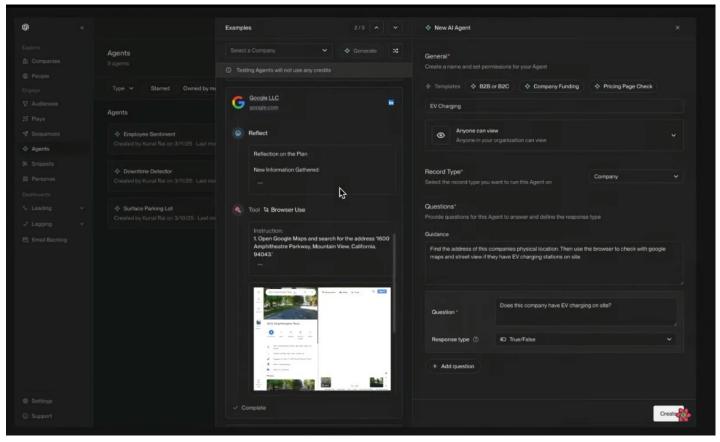
- Problem: There is a lot of rich data online that isn't accessible through basic scraping
 - Examples: Google maps images, interactive search experiences, etc
- Solution: Allow the Unify agent to use the browser the same way we would

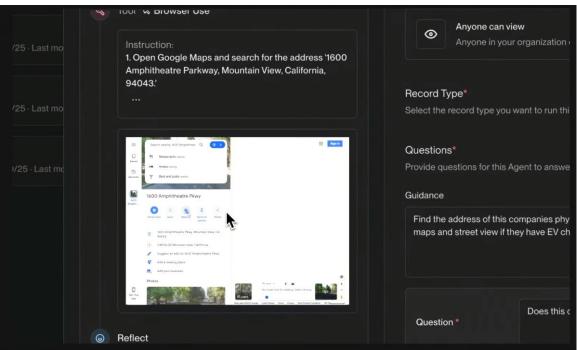
Browser Access: A sub agent

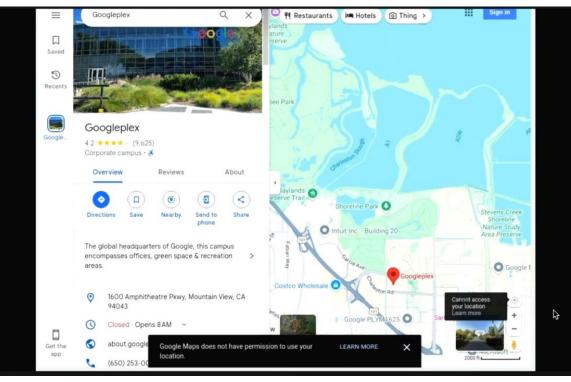
- The browser access tool is implemented as a sub agent
- It is given a task, which it decomposes into a browser trajectory using o4mini and executed with computer-use-preview

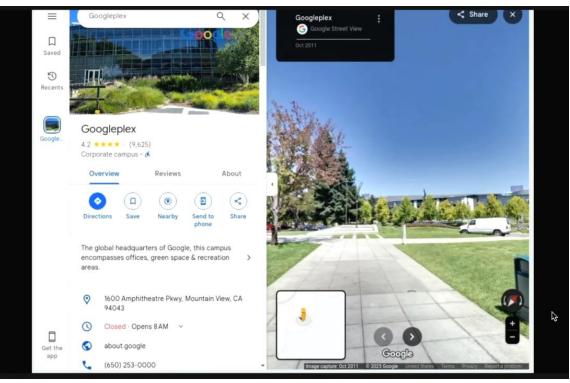


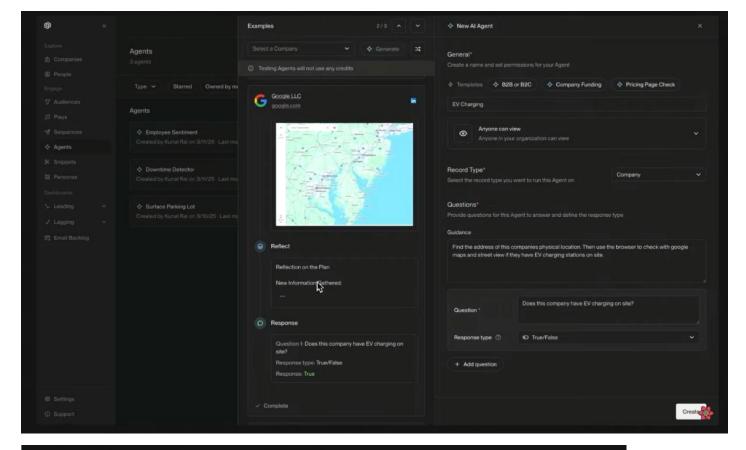












Tool Learnings

- We can't be reliant on naive internet search and scraping
 - SEO, misleading Google previews, and incorrect entity resolution make search very challenging for an agent
 - Deep search massively reduced misinterpretation of internet results
- Some tools unlocked whole new categories of questions the agent could answer
 - The HTML tool massively improved technology detection
- When giving the agent access to more tools, it emergently began cross checking sources with multiple tools

New Champion: KunalBrowserAgent

	Industry Classification	Firmographics	Account Qualification	Technographics
ConnorAgent	86	48	58	81
SamBotMarkI	75	38	59	72
KunalBrowserAgent	98	48	66	86

Next Steps: Improving Evals

- Evaluations should highlight problems and allow for quantitative assessment of gaps
- Accuracy based metrics are helpful but not a good eval for iteration
- Hard to evaluate models at qualitative tasks without using lots of human eval or llm-as-a-judge
- Pulling in more customer use cases as benchmarks for new tools and general agent performance

TL;DR

- Unify is building a system of action to make growth a science and help the best products win
- The more abstract the task, the more important evals are (but vibe checks are still very important)

Questions?

Follow us on Twitter: @unifygtm @heggieconnor @kunal_rye We're Hiring! Join us in building the world's first system of action for growing revenue! unifygtm.com/careers