The Adoption and Usage of AI Agents: Early Evidence from Perplexity*

Jeremy Yang[1,2] Noah Yonack[1]
Kate Zyskowski[1] Denis Yarats[1] Johnny Ho[1] Jerry Ma[1]

[1]Perplexity

[2]Harvard University

December 1, 2025

Abstract

This paper presents the first large-scale field study of the adoption, usage intensity, and use cases of general-purpose AI agents operating in open-world web environments. Our research focuses on Comet, an AI-powered browser developed by Perplexity, and its integrated AI agent, Comet Assistant. Drawing on hundreds of millions of anonymized interactions with Comet and Comet Assistant, we address three fundamental questions: Who is using AI agents? How intensively are they using them? And what are they using them for? Our findings reveal significant variation in adoption and usage across user segments. Earlier adopters, users in countries with higher GDP per capita and educational attainment, and individuals working in digital and knowledge-intensive sectors—such as digital technology, finance, marketing, and academics—are more likely to adopt and actively use the agent. To systematically characterize the substance of agent usage, we introduce a novel hierarchical agentic taxonomy that organizes common use cases across topic, subtopic, and task levels. The two largest topics—*Productivity & Workflow* and *Learning & Research*—together comprise 57% of all agentic queries. The two largest subtopics—*Courses* and *Shopping for Goods*—together represent 22% of queries. The top 10 out of 90 tasks make up 55% of queries. Personal use constitutes more than half of the queries, while professional and educational contexts account for 30% and 16%, respectively. Over time, users shift toward more productivity-oriented use cases. The development and diffusion of increasingly capable AI agents has implications for researchers, businesses, policymakers, and educators, motivating new directions for future research on this emerging class of AI capabilities.

*J.Y. and N.Y. contributed equally. We thank Gustav Lindqvist, Alexis Weill, and many other Perplexity staff for helpful insights, discussions, and technical assistance. All errors are the authors' own.

1 Introduction

2025 is frequently heralded as the year of agentic AI, as the frontier shifts from conversational Large Language Model (LLM) chatbots to action-oriented AI agents.[1] This narrative has emerged as AI agents have progressed from theorized concepts to productized assistants, demonstrating strong potential to transform work and daily life by planning and executing complex tasks in response to high-level human instructions with little supervision [Wooldridge and Jennings, 1995].[2] AI agents could profoundly reshape team, organization, and market structures by increasing productivity and efficiency and lowering transaction costs as autonomous participants in both consumption and production processes [Hadfield and Koh, 2025, Rothschild et al., 2025, Shahidi et al., 2025]. In aggregate, Precedence Research estimates that the global agentic AI market size will grow from $8 billion in 2025 to $199 billion by 2034.[3] PWC forecasts that the economic contribution will reach between $2.6 trillion and $4.4 trillion annually by 2030.[4]

Despite this enthusiasm and far-reaching economic implications, systematic behavioral evidence in the field on how people actually adopt and use AI agents remains limited, often relying on non-representative, firm-level surveys, or focusing on specialized agents such as coding [Sarkar, 2025].[5] Launched in July 2025, Comet by Perplexity is one of the first widely-adopted AI browsers and offers the embedded Comet Assistant as a general-purpose AI agent that performs user-specified tasks across open-world web environments. By studying hundreds of millions of anonymized adoption and usage data points from Comet users through the lens of a novel agentic taxonomy, we address the gap by providing early insights into three fundamental questions: Who is using AI agents? How intensively are they using them? And what are they using them for?

AI Agents

Like Artificial General Intelligence (AGI), agentic AI is another concept that resists a precise definition.[6] For instance, OpenAI defines agentic AI systems as those capable of taking actions that consistently contribute towards achieving goals over extended periods without their behavior being explicitly specified in advance;[7] Anthropic describes agents as systems that dynamically direct their own processes and tool use, maintaining control over how they complete tasks;[8] and Google defines agentic AI as an advanced AI system focused on autonomous decision-making and

[1] https://finance.yahoo.com/news/nvidia-jensen-huang-says-ai-044815659.html
https://x.com/gdb/status/1879327050819104778
https://www.aboutamazon.com/news/company-news/amazon-ceo-andy-jassy-on-generative-ai
https://www.ibm.com/think/insights/ai-agents-2025-expectations-vs-reality

[2] Examples of such agentic AI products or features include Perplexity's Comet browser, OpenAI's ChatGPT Operator, Codex, and Atlas browser, as well as Anthropic's Claude Code and Computer Use.

[3] https://www.precedenceresearch.com/agentic-ai-market

[4] https://www.pwc.com/m1/en/publications/agentic-ai-the-new-frontier-in-genai.html

[5] https://knowledge.wharton.upenn.edu/special-report/2025-ai-adoption-report/

[6] Some researchers also distinguish between AI agents and agentic AI—for example, Sapkota et al. [2025] defines AI agents as single-task, modular entities with tool-based autonomy, while agentic AI refers to distributed, multi-agent systems capable of collaborative, adaptive intelligence across workflows.

[7] https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf

[8] https://www.anthropic.com/engineering/building-effective-agents

action.[9] Despite these variations, the definitions share several common themes: goal orientation, action taking, and autonomy. At its core, an AI agent is guided by a goal, takes a sequence of actions by interacting with its environment to achieve that goal, and directs and redirects its own planning and execution processes. We define agentic AI similarly within the context of Comet Assistant, placing particular emphasis on the agent's capacity not only to exchange information with its environment through its actions but also to directly modify it.

The agentic workflow is typically characterized by three iterative phases—thinking, acting, and observing [Yao et al., 2022].[10] In the thinking phase, the agent interprets the goal from the query[11] and devises a step-by-step plan to achieve it. In the acting phase, the agent executes actions by controlling external tools to interact with its environment. In the observing phase, the agent processes feedback from its environment, returning to the thinking phase to confirm or revise its plan as needed. The agent autonomously moves between the phases to achieve the end-goal.

It is also useful to contrast LLM chatbots and AI agents. Both chatbots and agents build upon LLMs, but agents extend chatbots' capabilities from conversations to autonomous actions. LLMs serve as the "brain" of an agent, functioning as the central reasoning engine that processes information, evaluates options, and makes decisions. Tools are the "hands" that connect the agent's reasoning to the external world, enabling it to act upon its environment. More advanced agent capabilities also include multi-agent orchestration—the ability to interface with and manage the workflow across multiple collaborating agents, and self-evolution—the ability to identify gaps in pre-specified resources and dynamically expand them.[12]

Research Setting: Perplexity and Comet

Perplexity is an AI-powered platform that helps users discover, analyze, and act on information. Instead of requiring users to navigate through pages of results ("blue links"), as traditional search engines do, Perplexity interacts with the web on users' behalf to deliver direct, verifiable, and conversational answers. Each answer includes inline citations and links to original sources, allowing users to verify information and explore topics in more detail.[13]

Comet is a browser from Perplexity that builds an AI assistant directly into the browsing experience, helping users better search, digest, and act on information. Its core feature, Comet Assistant, operates as an autonomous agent that takes actions and completes open-world web-based tasks on behalf of users.[14] To fulfill user requests, Comet Assistant can perform a variety of tasks including scheduling meetings, editing documents, sending emails, booking flights, making purchases, and more.[15]

[9] https://cloud.google.com/discover/what-is-agentic-ai

[10] https://huggingface.co/learn/agents-course/en/unit1/agent-steps-and-structure

[11] We use query and prompt interchangeably.

[12] https://www.kaggle.com/whitepaper-introduction-to-agents

[13] https://www.perplexity.ai/help-center/en/articles/10352155-what-is-perplexity

[14] Besides the Comet Assistant, other agentic features in Comet include a specialized Email Assistant and a Background Assistant that operates in the background without taking control of a browser tab. Our study focuses on the Comet Assistant. https://www.perplexity.ai/hub/blog/comet-is-now-available-to-everyone-worldwide.

[15] We provide some sample agentic queries in Figure 9 and an example of the agent executing a real task in Figure 10 in Appendix B.

Comet was launched on July 9, 2025 on desktops for subscribers of Perplexity's Max tier[16], along with selected users from a pre-launch waitlist.[17] Access expanded to Pro subscribers on August 13, 2025, beginning with users in the United States.[18] On October 2, 2025, Comet became available to all users worldwide.[19] Besides these general cohorts, Comet was opened to all university students globally on September 3, 2025.

Data

Our analysis relies on three samples collected from Comet desktop users between July 9 and October 22, 2025.[20] First, we use anonymized data from the entire population of Comet users and their queries to provide high-level, aggregated statistics on agent adoption and usage intensity; this sample includes millions of users and hundreds of millions of queries. Second, we analyze a random sample of 100,000 Comet users and classify their O*NET occupation clusters and subclusters based on The National Career Clusters Framework to examine variation across occupations.[21] Third, we analyze another random sample of 100,000 agent users and classify all their agentic queries based on a novel hierarchical agentic taxonomy to better understand common use cases at topic, subtopic, and task levels.

Summary of Findings

We report two sets of main results on the adoption and usage patterns of AI agents—the extensive and intensive margins, as well as a comprehensive taxonomy of use cases.

Adoption and usage intensity

Overall, agent adoption and usage intensity demonstrate sustained growth with acceleration following the general availability (GA) of Comet. The post-GA period accounts for 60% of cumulative agent adopters and 50% of agentic queries throughout our sampling period. Earlier Comet adopters (those with pre-GA access) represent a disproportionately large share of agent adopters and query volume relative to their user share—an average user in the first cohort (July 9) is twice as likely to adopt the agent and makes nine times more agentic queries than an average user in the GA cohort (October 2). At the country level, adoption and usage intensity show strong positive correlations with GDP per capita and average years of education. At the occupational level, adopters and queries tend to come more from digital and knowledge-intensive domains.

[16] Perplexity offers three consumer subscription tiers: Free, Pro ($20 per month), and Max ($200 per month).

[17] https://www.perplexity.ai/hub/blog/introducing-comet

[18] https://www.perplexity.ai/hub/blog/the-intelligent-business-introducing-comet-for-enterprise-pro

[19] https://www.perplexity.ai/hub/blog/comet-is-now-available-to-everyone-worldwide

[20] We define Comet users as those who made at least one query on Comet during our study period. We use October 22, 2025 as the cutoff date because a major agent update began rolling out to selected users on October 23, which could affect adoption and usage patterns thereafter. The new agent was launched to all users on November 6. The updated agent performs 23% better than the previous version and has greater capacity for multitasking across multiple tabs. The agent we analyze in our data is focused on operating in a single web environment. For more details, see: https://www.perplexity.ai/hub/blog/the-new-comet-assistant.

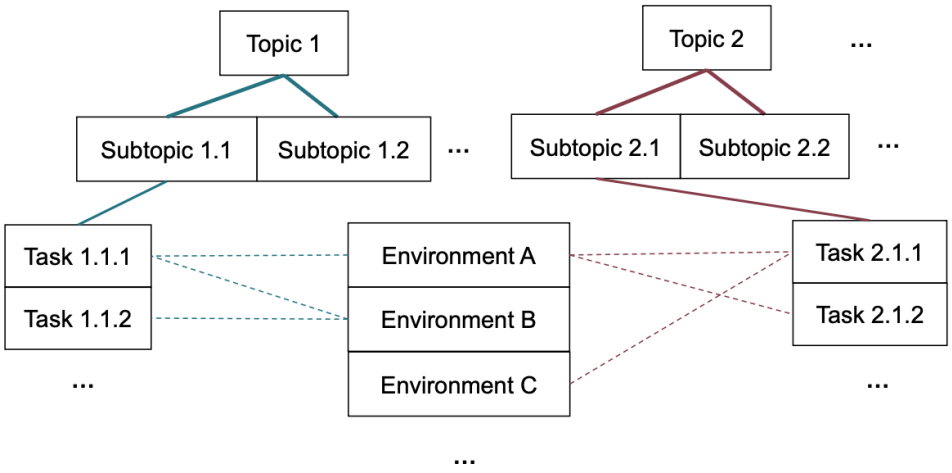
[21] https://www.onetonline.org/find/career?c=0
https://careertech.org/career-clusters/

Digital technology[22] represents the largest occupational cluster, comprising approximately 30% of adopters and query volume, followed by academics (including both the student and education clusters), finance, marketing, and entrepreneurship. These six occupational clusters collectively account for over 70% of total adopters and queries.

Use cases

We illustrate the hierarchical structure of our agentic taxonomy in Figure 1 and report our full taxonomy in Table 1. Productivity is the dominant topic with a 36% share. It is followed by learning (21%), media (16%), and shopping (10%). The most prevalent subtopics with over 5% query share include courses (13%), shopping for goods (9%), research (8%), document editing (8%), account management (7%), and social media (7%). The most frequently observed tasks are course exercise assistance (9%), research information summarization and analysis (7%), document creation and editing (7%), product search and filtering (6%), and research information search and filtering (6%). We also study the use of agents across environments, which are the websites that these tasks are performed on. The concentration of environments varies substantially across subtopics: the top five environments account for 97% of queries in music, 97% in videos, and 96% in professional networking, compared to only 28% in account management, 35% in shopping for services, and 37% in project management. Across all use cases, approximately 55% of agentic queries originate from personal usage context, while 30% stem from professional and 16% from educational settings. Temporally, query shares shift from travel and media topics to productivity, learning, and career topics. Users also show strong within-topic persistence, demonstrating stickiness in use cases; when topic transitions occur, users have higher propensity to migrate toward productivity, learning, or media-related applications.

[22] The Digital Technology Career Cluster focuses on developing digital systems for communication and data storage using critical technologies such as artificial intelligence (AI), data analytics, and cybersecurity. https://careertech.org/career-clusters/digital-technology/.



Note: A topic contains multiple subtopics, a subtopic contains multiple tasks, and a task can be connected to multiple environments and an environment can also be connected to one or multiple tasks within or across different topics and subtopics. The colors indicate different topics, solid lines indicate connections between topic, subtopics, and tasks, and dashed lines indicate which tasks are performed in which environments. Note that task 1.1.1 can be performed in environment A and B, whereas task 1.1.2 can only be performed in environment B. Task 2.1.1 and 2.1.2 indicate similar patterns. Environment A cuts across subtopics 1.1 and 2.1, whereas environment B and C are specific to a subtopic. We substantiate the structure with some examples. The query "unsubscribe me from all promotional emails that I receive more than twice per month" would be labeled as {Topic: Productivity & Workflow, Subtopic: Email Management, Task: Search or filter emails, Delete or unsubscribe emails}. In this case, search or filter emails and delete or unsubscribe emails can both be performed in environments such as Gmail or Outlook. Now imagine another query that gets classified into {Topic: Shopping & Commerce, Subtopic: Shopping for Goods, task: Search discounts, Make product purchase}, both tasks can be performed on Instacart, whereas only search discounts can be performed on SimplyCodes as it only shows discount codes and do not sell products directly. Facebook is one example of cross-topic environments—it could be the environment for Media & Entertainment queries, but also Shopping & Commerce queries when they are about products listed on Facebook Marketplace.

Figure 1: Hierarchical Structure of the Agentic Taxonomy

Topics	Subtopics	Tasks
Productivity & Workflow	Account Management	Register/log in to accounts, Manage settings/profiles, Manage files, Summarize/analyze account information
	Document & Form Editing	Create/edit documents, Search/filter documents, Summarize/analyze documents
	Multimedia Editing	Create/edit multimedia, Search/filter multimedia, Summarize/analyze multimedia
	Email Management	Search/filter emails, Create/edit emails, Send emails, Delete/unsubscribe emails, Summarize/analyze emails
	Spreadsheet & Data Editing	Create/edit spreadsheets, Search/filter spreadsheets, Summarize/analyze spreadsheets
	Computer Programming	Create/edit code, Execute code, Summarize/analyze code
	Investments & Banking	Search/filter stocks, Buy/sell stocks, Summarize/analyze investment information, Summarize/analyze banking information
	Project Management	Create/edit projects, Summarize/analyze project information
	Calendar Management	Create/edit events, Check availability, Search/filter events, Summarize/analyze events
Learning & Research	Courses	Navigate courses, Summarize/analyze course materials, Assist exercises
	Research	Search/filter research information, Summarize/analyze research information
Media & Entertainment	Social Media & Messaging	Search/filter social media posts, Summarize/analyze social media posts, Create social media posts, Engage with social media posts, Send social media/text messages
	Online Games	Search/filter online games, Summarize/analyze online game information, Play online games
	Movies, TV, & Videos	Search/filter videos, Summarize/analyze videos, Play videos, Navigate within videos, Manage playlists
	Music & Podcasts	Search/filter music/podcasts, Summarize/analyze music/podcasts, Play music/podcasts, Manage playlists
	News	Search/filter news, Summarize/analyze news
	Sports	Search/filter match/player information, Summarize/analyze match/player statistic
Shopping & Commerce	Goods	Search/filter products, Search discounts, Summarize/analyze product information, Add products to cart, Make product purchase, Manage orders
	Services	Search/filter products, Search discounts, Summarize/analyze product information, Add products to cart, Make product purchase, Manage orders
Travel & Leisure	Flights & Transportation	Search/filter flights & transportation, Summarize/analyze flights & transportation, Add flights & transportation to cart, Book flights & transportation
	Trip Itineraries	Search/filter destinations, Plan trips, Summarize/analyze trips
	Lodging	Search/filter lodging, Summarize/analyze lodging information, Add lodging to cart, Book lodging
	Restaurants	Search/filter restaurants, Summarize/analyze restaurant information, Book restaurants
Job & Career	Job Search & Application	Search/filter jobs, Summarize/analyze job descriptions, Complete applications
	Professional Networking	Search/filter professional profiles, Summarize/analyze professional profiles, Send professional connection requests/messages, Engage with professional profiles/posts

Note: The table contains all topics, subtopics, and tasks in the agentic taxonomy, except "Other". Topic and subtopics are general goals and tasks are specific tasks the agent is expected to complete in order to achieve that goal. A query is classified into one topic, one subtopic underneath that topic, and one or more tasks underneath that subtopic. Queries that cannot be classified into the taxonomy at a given level are labeled as "Other" at that level and all subsequent levels. For example, query that does not belong to any of the topic would be labeled as "Other" at topic, subtopic, and task levels; query that belongs to productivity but does not belong to any of the subtopics under productivity will be labeled as "Other" at subtopic and task levels; query that belongs to productivity and email management but does not belong to any of the tasks under email management will be labeled as "Other" at the task level.

Table 1: Agentic Taxonomy—Topics, Subtopics, and Tasks

The remainder of this paper is structured as follows. Section 2 reviews related literature and highlights our contributions. Section 3 describes our sampling methodology and data privacy safeguards. Section 4 explains the development of our hierarchical agentic taxonomy. Section 5 presents our main findings regarding the patterns of AI agent adoption, usage intensity, and use cases. Finally, Section 6 discusses the implications of our findings for researchers, businesses, and policymakers, while acknowledging limitations and identifying promising avenues for future research that we aim to pursue. Key figures and tables are included in the main text. Additional figures and tables and other supplementary materials are provided in the Appendices.

2 Related Work

Our paper is directly related to the literature on how people use LLMs and AI agents in real-world settings.[23] Our paper extends recent work on the adoption and usage of LLM chatbots. A few prominent studies in this area are Handa et al. [2025b], which documented how people use Claude, and Zhao et al. [2024] and Chatterji et al. [2025], which examined how people use ChatGPT. Anthropic also extended their analysis to specific user groups such as university students [Handa et al., 2025a], educators [Bent et al., 2025], and geographics and enterprises [Appel et al., 2025]. These papers developed taxonomies to categorize standard Q&A queries. We also develop a taxonomy using internal data from an AI product; however, our key contribution is to move beyond the Q&A queries and instead focus on agentic queries. The main difference is that Q&A queries focus on information exchanges between the user and model in a conversation, whereas agentic queries focus on the agent executing tasks on behalf of the user in an external environment. Evidence on how people use AI agents in the field is more limited, and usually focused on specific use cases such as coding. For example, Anthropic [2025] studied the usage of Claude Code, a coding agent, in software development, and Sarkar [2025] investigated the adoption, usage, and productivity impact of coding agents in Cursor. Our contribution differs in that we analyze a general-purpose agent operating across all common use cases.[24]

3 Data

Sampling

Our analysis leverages three samples collected between July 9 and October 22, 2025—that is, from the launch date to 20 days after general availability. Each sample is tailored to a particular set of research questions.

We define a Comet user as a user who has made at least one query on Comet during the study period. At the user level, we exclude all enterprise users, users under the Perplexity for Government program, users who deleted their accounts during the sampling period, users who

[23] Following our definition of AI agents, we do not discuss papers that do not involve the agent taking actions to manipulate their environments.

[24] It is also worth noting that there is adjacent work in experimental settings that focused on understanding the behaviors of AI agents or human-agent collaboration. For example, Allouah et al. [2025] studied agent decisions when shopping on behalf of users, and Ju and Aral [2025] examined human–agent collaboration in image editing tasks.

opted out of data retention for model training during that period, and logged-out users. At the query level, we define an agentic query as a query that involves the agent taking control of the browser or taking actions on external applications—such as email or calendar clients—through connectors built on the Model Context Protocol (MCP) or API calls.[25] Under this stricter definition, we do not consider all queries with tool use (such as web search or code interpreter) as agentic, since these tools merely exchange information with external environments but do not manipulate them. When users onboard onto Comet, sample agentic queries are shown for demonstration purposes; we remove these queries to focus only on user-initiated ones. In rare cases, a single agentic query might trigger multiple browser-control or MCP and API calls; we exclude such cases to focus on queries that triggered a single call to ensure a clean inference of user intent. Lastly, we exclude queries made in Comet's incognito mode. We describe the three samples we analyze below.

Sample A: The population of Comet users and queries

We use the entire population of millions of users and hundreds of millions of queries on Comet—both agent and non-agent—during the sampling period to understand overall patterns in adoption and usage intensity.

Sample B: A random sample of Comet users and queries

We draw a random sample of 100,000 Comet users and a random sample of their recent queries—both agent and non-agent—during the study period to infer their O*NET occupation clusters and subclusters, enabling us to examine variation in adoption and usage intensity across occupations.[26] The sampling includes two stages. First, a random set of users are selected, then for each user, a random set of queries from recent dates are selected. The sampled queries are then concatenated into a single string and labeled with a classifer against the O*NET occupation taxonomy. We include students as a separate cluster; the student status is verified through a third-party vendor.

Sample C: A random sample of Comet agent users and all their agentic queries

We draw another random sample of 100,000 agent users and classify all their agentic queries using a novel agentic taxonomy to identify common use cases. The sampling is performed only at the user level: once a user is selected, all their agentic queries are included in the analysis. This procedure allows us to track within-user agent usage trajectories and uncover longitudinal patterns. In large user samples, the queries are also representative of query-level estimands such as common use cases. For the same sample of users, we further infer their O*NET occupation clusters and subclusters.

[25] https://www.anthropic.com/news/model-context-protocol

[26] https://www.onetonline.org/find/career?c=0

Data Privacy

We follow industry standards and adopt multiple safeguards to ensure that no personally identifiable information (PII) is used by any human in any part of the analysis.

First, as noted above, enterprise users, users under the Perplexity for Government program, users who deleted their accounts during the sampling period, users who opted out of data retention for model training during that period, logged-out users, and queries made in incognito mode are excluded from the analysis. Second, our analysis does not use any demographic information, names, email addresses, or other real-world identifiers; all user-level matching is performed through internal numerical user IDs. Third, we employ automated classifiers to label occupations and use cases. The classifier input is not the raw query text but rather a reformulated description of the underlying intent, enriched with context such as prior queries in the same conversation and the website on which the query was made. Lastly, all results reported in the paper are presented only in highly aggregated form.

4 Agentic Taxonomy

We develop a hierarchical agent use case taxonomy guided by two principles. First, it should comprehensively capture common agentic intents so that it can generalize to other agentic products beyond Comet. Second, it should have a hierarchical structure that reveals higher-level goals while distinguishing specific lower-level tasks and actions.

To achieve these goals, we adopt a bottom-up approach consisting of three phases: exploration, refinement, and classification. In the exploration phase, we draw a random sample of agentic queries, extract their embeddings, and apply K-means clustering to group them based on semantic similarity. Queries are then sampled from each cluster and concatenated into a single string representing that cluster. We then summarize each concatenated string to interpret the meaning of each cluster. In the refinement phase, we manually examine the cluster labels identified in the previous step to merge, split, trim, or expand them, following the guiding principles. When a significant share of queries is labeled as "Other", suggesting that they are not well captured by the provided taxonomy, we re-classify the queries in that cluster using the bottom-up approach in the first step to identify clusters missing from the taxonomy and update it. Finally, we classify agentic queries within the finalized taxonomy using a query classification model.

Our final taxonomy consists of three hierarchical levels—*topic*, *subtopic*, and *task*—as illustrated in a stylized diagram in Figure 1, along with their connections to the environments the tasks are performed in.[27] The full taxonomy is summarized in Table 1.

Topic and subtopic are top and mid-level use cases of the agent, indicating the overall goal, and tasks are the specific tasks the agent is expected to complete in order to achieve that goal. Each query is classified into one topic, one subtopic, and one or more tasks. For instance, the

[27] Queries that cannot be classified into the taxonomy at a given level are labeled as "Other" at that level and all subsequent levels. For example, a query that does not belong to any of the topics would be labeled as "Other" at the topic, subtopic, and task levels; a query that belongs to productivity but does not belong to any of the subtopics under productivity will be labeled as "Other" at the subtopic and task levels; a query that belongs to productivity and email management but does not belong to any of the tasks under email management will be labeled as "Other" at the task level.

query "unsubscribe me from all promotional emails that I receive more than twice per month" would be labeled as {Topic: Productivity & Workflow, Subtopic: Email Management, Task: Search or filter emails, Delete or unsubscribe emails}.

The environments the agent operates in are observed in the data and can be connected to our taxonomy. Tasks in a particular subtopic are performed in a particular set of environments, and each environment might involve one or more of these tasks. For instance, under the subtopic *Email Management*, tasks such as *Search or filter emails* and *Delete or unsubscribe emails* can both be performed in environments such as Gmail or Outlook; under *Shopping for Goods*, *Search discounts* and *Make product purchase* can both be performed on Instacart, whereas only *Search discounts* can be performed on SimplyCodes as it only shows discount codes and does not sell products directly. Furthermore, an environment might cut across multiple topics and subtopics. For instance, Facebook could be the environment for *Media & Entertainment* queries, but also *Shopping & Commerce* queries when they are about products listed on Facebook Marketplace. We further categorize the usage context into *personal*, *professional*, and *educational* domains.

We validate the classification accuracy against a golden dataset of 1,000 anonymized and desensitized queries. The classifier-assigned labels agree with the topics, subtopic, tasks, and usage context in the golden dataset 92%, 89%, 80%, and 86% of the time, respectively. More details about the validation are provided in Appendix D.

5 Main Results

We first discuss the results on the adoption (extensive margin) and usage intensity (intensive margin), and then the use cases (agentic taxonomy).

5.1 Adoption and Usage Intensity

We define agent adopters as users who had at least one agentic query in the sampling period. The results below are all based on Sample A, except for occupation, which is based on Sample B. Figure 6 shows that agent adoption and overall usage as measured by agentic query volumes are growing steadily over the period studied, with an increased pace after Comet became generally available. About 60% of agent users are acquired and 50% agentic queries occurred post-GA. The query volumes grow at a slightly higher rate than adopters.

We analyze the adoption and usage patterns of user segments defined by cohort, country, and occupation. To capture the magnitude of adoption and usage of a user segment relative to their user share, we define Perplexity Agent Adoption Ratio (AAR) and Agent Usage Ratio (AUR) as the percentage of agent adopters and agentic queries of a segment relative to their overall user shares.[28] An AAR or AUR greater than one indicates that a segment is over-represented in the adopters or queries relative to their population base, and vice versa.

[28] These ratio-based metrics are often used to quantify the relative degrees of adoption and usage (e.g., Appel et al. [2025])

By cohort

Table 2 shows that among the three cohorts by access dates, earlier adopters (those with access before GA) account for about 30% of total users, but about 50% of agent adopters and 70% of query volumes. The disparity is more pronounced in usage intensity than in adoption—an average user in the first cohort (July 9) is twice as likely to adopt the agent but makes nine times more agentic queries than an average user in the GA cohort (October 2). AAR and AUR both decrease in the order of access cohorts.

In general, the composition of early adopters is endogenous to the rollout plan; in our case, it is based on the subscription tiers. Nonetheless, these results are consistent with general patterns in the adoption and usage of new technologies: early adopters disproportionately drive initial usage, and further diffusion may require additional educational efforts. With improved agent-to-agent collaboration capabilities and standardized protocols, one might expect stronger network effects that could accelerate adoption and usage in the future.[29]

Cohort	User Share (%)	Agent Adopter Share (%)	Agentic Query Share (%)	AAR	AUR
July 9	4.3	7.7	18.9	1.79	4.40
August 13	28.3	38.1	48.5	1.35	1.71
October 2	67.4	54.2	32.7	0.80	0.49

Note: The table shows the agent adoption and query volume by cohort. User share is the number of users in each cohort divided by total users. Agent adopter and query share are the number of adopters and agentic queries in each cohort divided by the total adopters and agentic queries. AAR (Agent Adoption Ratio) is the ratio between agent adopter share and user share. AUR (Agent Usage Ratio) is the ratio between agentic query share and user share. AAR and AUR greater (less) than 1 indicate that a cohort is over-represented (under-represented) in agent adopters and queries relative to their user base.

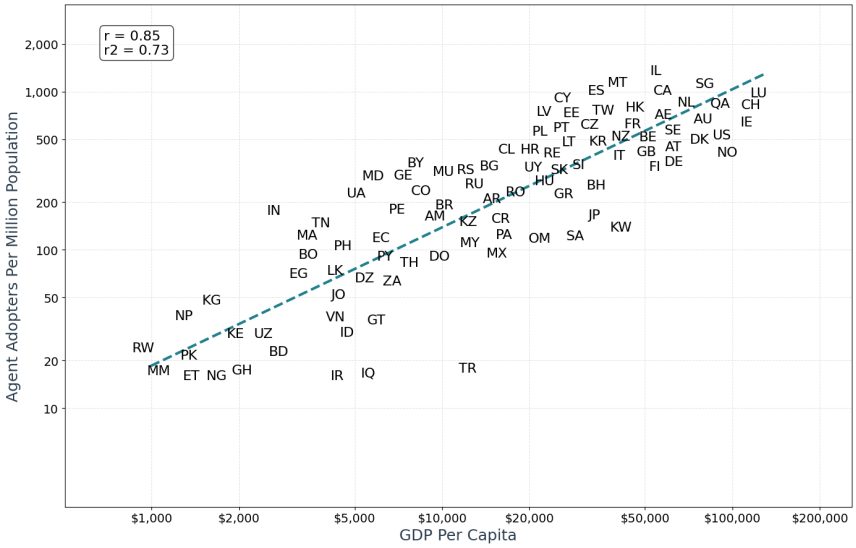
Table 2: Agent Adoption and Agentic Query by Cohort

By country

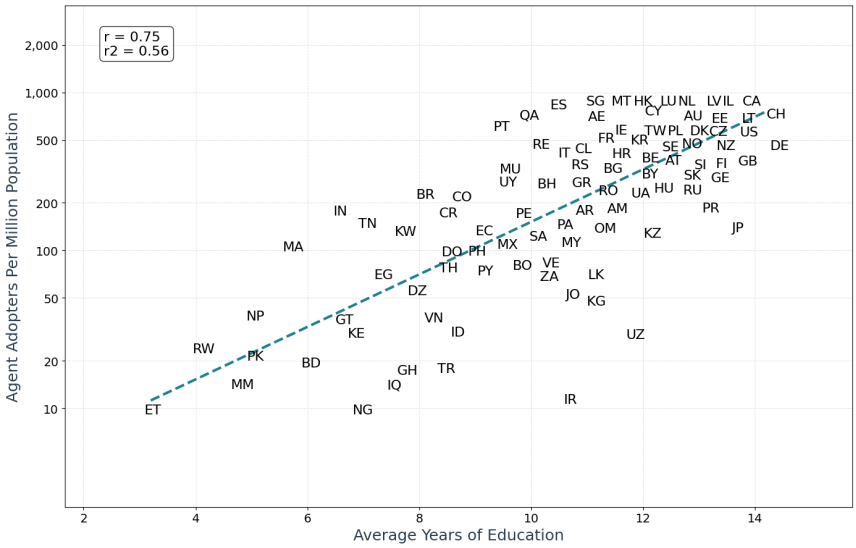
Figure 2 and 3 show that there is strong positive correlations between log agent adopter count per million population and log GDP per capita ($r = 0.85, p < 0.001, R^2 = 0.73$) and years of education ($r = 0.75, p < 0.001, R^2 = 0.56$), where r is the correlation coefficient, p is the p-value of the correlation coefficient, and R^2 is the R-squared of the regression lines. The correlations between log agentic query count per million population and log GDP per capita ($r = 0.86, p < 0.001, R^2 = 0.74$) and years of education ($r = 0.75, p < 0.001, R^2 = 0.57$) follow a similar pattern. Together, they suggest that relatively more economically developed and educated countries tend to adopt and use the agent more.[30]

[29] For instance, Agent2Agent Protocol (A2A). https://a2a-protocol.org/latest/.

[30] Results remain consistent when population is replaced with working population.



(a) Log GDP Per Capita vs. Log Agent Adopter Per Million Population



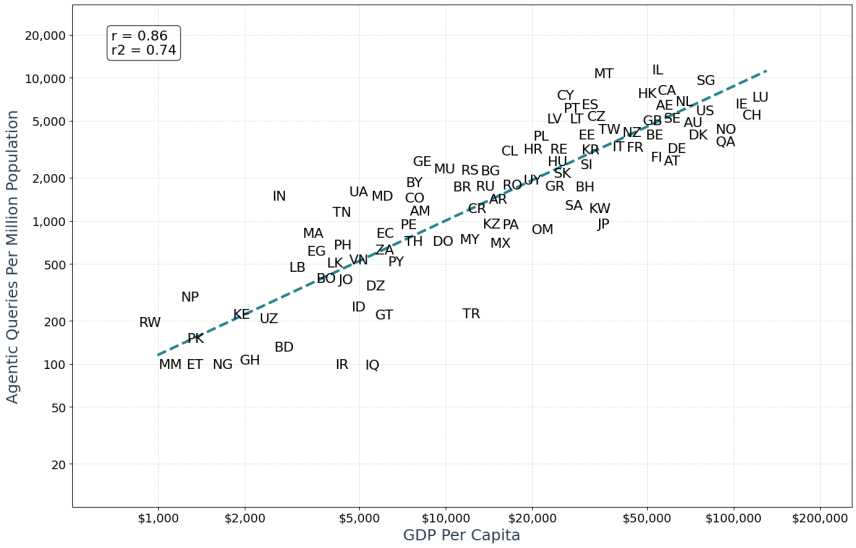
(b) Average Years of Education vs.
Log Agent Adopter Per Million Population

Note: The plots show the scatterplots and best-fitting lines of log GDP per capita and average years of education vs. the log of agent adopter per million population for the top 100 countries by agent adopter count. The plots are on log scale but labels are in absolute values for better readability. Jitter is applied to the country labels to provide better visual separation. r is the correlation coefficient, p is the p-value of the regression coefficient, and R^2 is the R-squared of the regression lines. The GDP and population data is from World Bank World Development Indicators (2024)[a] and the average years of education data is from UNDP Human Development Report (2024)[b].

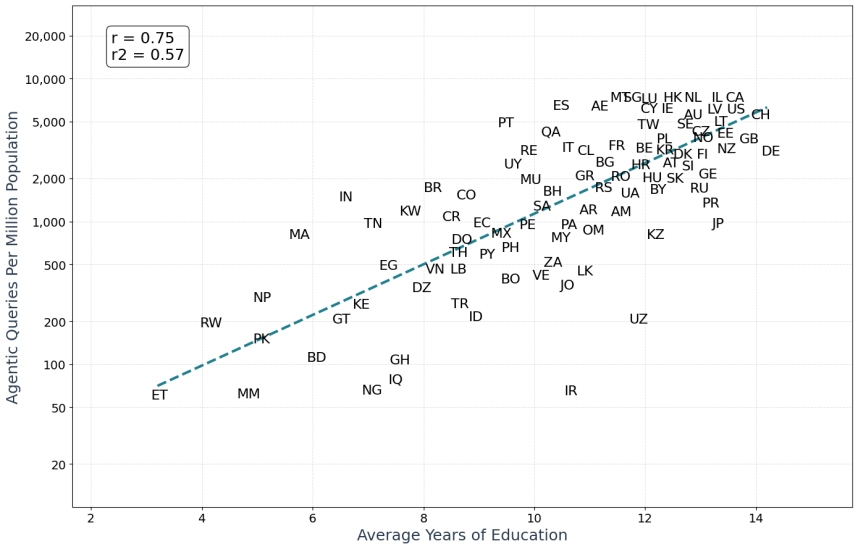
[a] https://data.worldbank.org/

[b] https://hdr.undp.org/content/human-development-report-2023-24

Figure 2: Log GDP Per Capita and Average Years of Education vs.
Log Agent Adopter Per Million Population by Country



(a) Log GDP Per Capita vs. Log Agentic Query Per Million Population



(b) Average Years of Education vs.
Log Agentic Query Per Million Population

Note: The plots show the scatterplots and best-fitting lines of log GDP per capita and average years of education vs. the log of agentic query per million population for the top 100 countries by agentic query count. The plots are on log scale but labels are in absolute values for better readability. Jitter is applied to the country labels to provide better visual separation. r is the correlation coefficient, p is the p-value of the regression coefficient, and R^2 is the R-squared of the regression lines. The GDP and population data is from World Bank World Development Indicators (2024)[a] and the average years of education data is from UNDP Human Development Report (2024)[b]

[a] https://data.worldbank.org/

[b] https://hdr.undp.org/content/human-development-report-2023-24

Figure 3: Log GDP Per Capita and Average Years of Education vs.
Log Agentic Query Per Million Population by Country

By occupation

Table 3 ranks occupation clusters in O*NET plus the student cluster by agent adopter shares and AAR.[31] Digital technology is by far the largest cluster and accounts for 28% of adopters, slightly higher than its user share. Academics (including the student and education clusters) and financial workers have more than 10% adopter share. Workers in marketing, design, and entrepreneurship have more than 5% adopter share. Clusters with lower adopter shares are typically those require interacting with the physical environment. The hospitality cluster has the highest AAR at 1.36, although the sample size is significantly smaller; it is followed by marketing and entrepreneurship at 1.24 and 1.17, respectively.

Table 4 ranks occupation clusters by agentic query share and AUR. The patterns track adoption closely—top clusters remain the same with slight changes in the rank. Comparing the AUR and AAR for the same cluster reveals patterns in the degree of usage intensity relative to the degree of adoption: students and workers in entrepreneurship, marketing, and digital technology all have AUR / AAR greater than one, suggesting that their tendency to use the agent conditional on adoption is even stronger than their tendency to adopt.

Tables 8 and 9 show the top ten occupation subclusters by agent adopter shares and AAR, and agentic query share and AUR, respectively. Patterns at the subcluster level are mostly in line with patterns at the cluster level, with software engineers being the largest subcluster that accounts for 14% of adopters and 15% of queries, and their AAR and AUR are around 1.1 and 1.2. Other subclusters are all below 6% in both adopter and query shares. Subclusters in marketing—such as business development and sales, digital marketing, and market research, and subclusters in entrepreneurship—such as information management, operation, and strategy, tend to have the highest AAR and AUR.

These results could reflect the specific task composition in each occupation and the degree to which it aligns with common agent use cases, which is the focus of the next section.

[31] O*NET clusters are defined for workers so they do not include students.

Agent Adoption by Occupation Cluster: By AAR					
Cluster	**User Share (%)**	**User Share Rank**	**Agent Adopter Share (%)**	**Agent Adopter Share Rank**	**AAR**
Hospitality, Events, & Tourism	2.5	11	3.4	9	1.36
Marketing & Sales	7.2	6	8.9	4	1.24
Management & Entrepreneurship	6.5	7	7.6	6	1.17
Digital Technology	26.4	1	27.7	1	1.05
Supply Chain & Transportation	2.2	13	2.3	12	1.05
Financial Services	10.0	3	10.1	3	1.01
Student	12.4	2	12.4	2	1.00
Construction	2.3	12	2.2	13	0.96
Energy & Natural Resources	0.9	14	0.8	14	0.89
Arts, Entertainment, & Design	9.1	4	8.0	5	0.88
Education	7.7	5	6.4	7	0.83
Healthcare & Human Services	4.9	8	4.0	8	0.82
Advanced Manufacturing	3.5	9	2.8	10	0.80
Public Service & Safety	3.4	10	2.7	11	0.79
Agriculture	0.9	15	0.7	15	0.78

Note: The table shows agent adoption by O*NET occupation cluster. "Other" category is removed. We put students in a separate cluster and educators in the Education cluster. User share is the number of users in each cluster divided by total users. Agent adopter share is the number of adopters in each cluster divided by the total adopters. AAR (Agent Adoption Ratio) is the ratio between agent adopter share and user share. AAR greater (less) than 1 indicates that a cluster is over-represented (under-represented) in agent adoption relative to their user base.

Table 3: Agent Adoption by Occupation Cluster

Agentic Query by Occupation Cluster: By AUR					
Cluster	**User Share (%)**	**User Share Rank**	**Agentic Query Share (%)**	**Agentic Query Share Rank**	**AUR**
Marketing & Sales	7.2	6	10.5	3	1.46
Management & Entrepreneurship	6.5	7	9.0	4	1.38
Student	12.4	2	15.6	2	1.26
Digital Technology	26.4	1	29.6	1	1.12
Hospitality, Events, & Tourism	2.5	11	2.6	9	1.04
Supply Chain & Transportation	2.2	13	2.0	10	0.91
Financial Services	10.0	3	8.6	5	0.86
Arts, Entertainment, & Design	9.1	4	6.9	6	0.76
Education	7.7	5	5.6	7	0.73
Construction	2.3	12	1.6	13	0.70
Healthcare & Human Services	4.9	8	3.3	8	0.67
Agriculture	0.9	15	0.5	15	0.56
Energy & Natural Resources	0.9	14	0.5	14	0.56
Advanced Manufacturing	3.5	9	1.9	11	0.54
Public Service & Safety	3.4	10	1.8	12	0.53

Note: The table shows usage intensity by O*NET occupation cluster. "Other" category is removed. We put students in a separate cluster and educators in the Education cluster. User share is the number of users in each cluster divided by total users. Agent query share is the number of agentic queries in each cluster divided by the total agentic queries. AUR (Agent Usage Ratio) is the ratio between agentic query share and user share. AUR greater (less) than 1 indicates that a cluster is over-represented (under-represented) in agent usage relative to their user base.

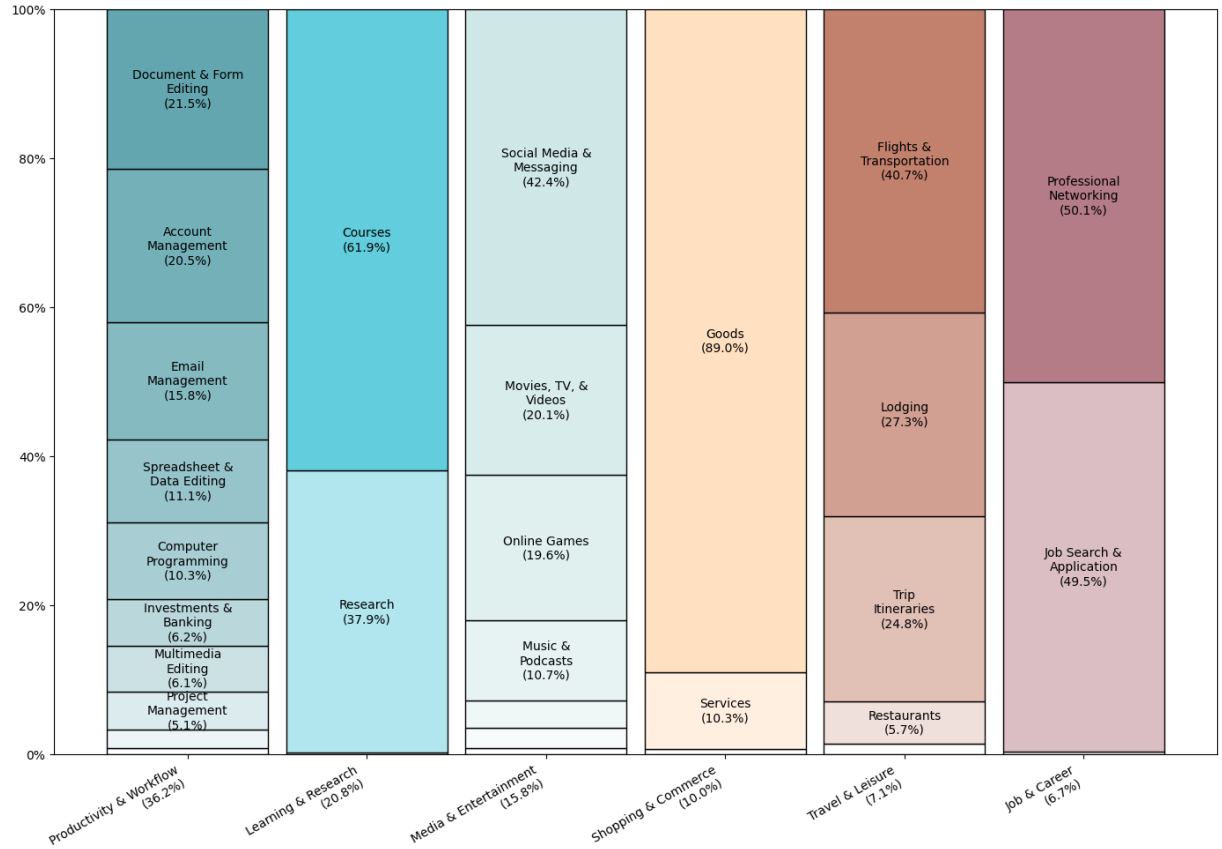
Table 4: Agentic Query by Occupation Cluster

5.2 Use Cases

We document the use cases by topic, subtopic, task, environment (the websites on which the tasks are performed), and usage context. All results are based on Sample C.

Topics and subtopics

Topics and subtopics capture the high-level goals of an agentic query. Figure 4 presents the topic share and subtopic share breakdown by topic. Productivity is the largest category that accounts for 36% of all agentic queries. Learning, media, and shopping are the other topics with over 10% query share. The two biggest topics—productivity and learning—together account for 57% of all queries. Some topics, such as learning, shopping, and career, have a dominant sutopic that accounts for more than half of the queries in that topic. Table 10 also shows the overall query share of each subtopic. Courses account for 13% of all queries, followed by shopping for goods, document editing, account management, social media, and email.



Note: The plot shows the percentage shares of subtopics within each topic. "Other" category is removed. The labels for subtopics that account for less than 5% within a topic are suppressed. Topic shares are shown in the labels on the x axis. The subtopics within each topic are sorted by percentage in decreasing order from top to bottom. The darker shades within a topic represent subtopics with higher percentages.

Figure 4: Topic Breakdown by Subtopic Percentage

Table 5 shows the topic distribution by occupation cluster. Topic-wise, productivity remains the largest topic for most occupations, whereas learning and reseach is the largest for students and educators, and travel is the largest for the hopitality cluster. Occupation-wise, workers in finance have the largest query share in productivity, students have the largest share in learning, workers in design have the largest share in media, workers in advanced manufacturing have the largest share in shopping, workers in entrepreneurship have the largest share in career, and workers in hospitality have the largest share in travel.

Cluster / Topic	Productivity & Workflow (%)	Learning & Research (%)	Media & Entertainment (%)	Shopping & Commerce (%)	Job & Career (%)	Travel & Leisure (%)
Digital Technology	41.0	19.9	14.7	9.3	9.1	6.0
Student	29.1	43.3	10.6	5.3	8.1	3.7
Management & Entrepreneurship	45.8	13.7	12.3	9.8	12.2	6.2
Marketing & Sales	37.5	12.1	23.8	14.0	8.2	4.5
Financial Services	46.7	15.2	14.0	9.8	5.7	8.6
Education	34.1	37.0	13.8	6.6	2.6	5.8
Arts, Entertainment, & Design	39.4	12.1	25.0	11.3	6.0	6.2
Healthcare & Human Services	38.6	23.3	14.3	10.2	5.9	7.7
Advanced Manufacturing	30.4	19.8	11.8	25.4	5.4	7.1
Public Service & Safety	39.7	26.2	17.7	7.2	3.2	6.0
Hospitality, Events, & Tourism	29.9	6.6	13.4	12.2	2.7	35.2
Supply Chain & Transportation	40.3	13.4	12.5	18.6	5.3	10.0
Construction	39.1	14.9	14.1	16.3	7.7	7.9
Energy & Natural Resources	42.9	18.8	13.1	10.4	4.9	9.8
Agriculture	41.5	20.5	13.9	12.4	5.4	6.3

Note: The table shows the distribution of topics by occupation cluster. Topic percentage ($P(\text{Topic} \mid \text{Occupation})$) is the topic share among all agentic queries from a given occupation cluster. Percentages may not sum to 100% due to rounding.

Table 5: Topic Distribution by Occupation Cluster

Figure 7 contrasts the topic distribution among the first agentic query for each user with the overall distribution. Over time, the query shares of productivity, learning, and career topics are increasing, suggesting a shift toward more productivity-oriented use cases. We also examine the topic transition patterns from query to query. Figure 8 shows the transition probability matrix from the previous query to the next query among all query pairs aggregated from the user level. The results show that in most cases topics transition into themselves, suggesting the stickiness of agent use cases. Productivity, learning, and career topics are the most sticky, whereas travel is the least sticky; media & and shopping are in between. When cross-topic transitions occur, they most likely transition into productivity, learning, or media-related topics.

Tasks

Tasks under topics and subtopics capture the low-level tasks the agent is expected to complete in order to achieve the end-goals. We show the top ten tasks in Table 6. Half of the top ten is in learning, including various learning and research assistance. The other five are split across productivity (edit documents and manage account settings), shopping (search or summarize product information), and media (search social media).

Topic	Subtopic	Task	Overall (%)
Learning & Research	Courses	Assist exercises	9.41
Learning & Research	Research	Summarize/analyze research information	6.71
Productivity & Workflow	Document & Form Editing	Create/edit documents/forms	6.58
Shopping & Commerce	Goods	Search/filter products	6.43
Learning & Research	Research	Search/filter research information	5.95
Shopping & Commerce	Goods	Summarize/analyze product information	5.18
Productivity & Workflow	Account Management	Manage settings/profiles	4.33
Learning & Research	Courses	Summarize/analyze course materials	3.69
Learning & Research	Courses	Navigate courses	3.31
Media & Entertainment	Social Media & Messaging	Search/filter social media posts/messages	3.29

Note: The table shows the top 10 tasks among all agentic queries. $P(\text{Task Overall}) = P(\text{Topic, Subtopic, Task}) = P(\text{Topic}) \times P(\text{Subtopic} \mid \text{Topic}) \times P(\text{Task} \mid \text{Topic, Subtopic})$.

Table 6: The Top 10 Tasks

Tables 11, 12, 13, 14, 15, and 16 show the main tasks under each topic and subtopic with over 5% query shares within a subtopic. Note that because query share measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic can add up to over 100. A few subtopics contain a dominant task that appears in over 80% of all queries in that subtopic. For instance, search flight and search lodging both appear in 93% of queries in the flight and lodging subtopics, respectively; edit document and summarize research information both appear in 85% of queries in document editing and research subtopics, respectively. Some subtopics show more dispersed task distributions: for instance, search videos—the top task in the video subtopic—appears in 48% of all queries in that category; search email—the top task in the email subtopic—appears in 49% of all queries in that category.

Table 17 shows the top five tasks in each occupation cluster. In genenal, research, document editing, and shopping-related tasks appear consistently across clusters. Some occupation clusters feature a prominent task. For instance, search products appears in 21% of queries from the advanced manufacturing cluster. Other clusters, in contrast, have a more diffused task composition. For instance, summarize research information and search products—the top tasks in the entrepreneurship and design clusters—only appear in less than 8% of their queries. The top tasks by occupation also shed light on why certain occupations tend to adopt and use the agent more—for instance, knowledge-intensive sectors such as digital technology, finance, entrepreneurship, and academics tend to use the agent for research and learning-related tasks, whereas highly digitized sectors such as marketing and design tend to use the agent for media-related tasks.

Environments

Environment refers to the external world with which the agent interacts while performing a task in order to achieve its goals. In our context, the environment is the specific website on which the agent operates for a given query. We show the top environments by overall query shares in Table 7. They are typically the top websites in their corresponding domains.

We break down the top five environments under each subtopic and their query shares in Tables 18, 19, 20, 21, 22, and 23. Some subtopics are dominated by a single environment. For instance, linkedin.com alone accounts for 93% of queries in professional networking, and the query shares

of youtube.com and docs.google.com[32]—the top environments in video and spreadsheet editing subtopics—are more than twenty times larger than the share of the second environments in those subtopics. On the other hand, there is only a 2% difference between the shares of coursera.org and netacad.com under courses, and a 3% difference between instagram.com and x.com under social media. Table 24 shows the combined shares of the top five environments in each subtopic. A higher share indicates the agent usage is more concentrated in a few environments. The level of concentration varies significantly: the top five environments account for 97% of queries in music, 97% in videos, and 96% in professional networking, compared to only 28% in account management, 35% in shopping for services, and 37% in project management. Lastly, Table 25 shows the top 5 environments for each occupation cluster and they are closely connected to the main use cases for each occupation.

Environment	**Overall (%)**
docs.google.com	11.97
email services combined	11.23
linkedin.com	9.42
youtube.com	7.03
amazon.com	3.46
instagram.com	2.56
messenger services combined	2.47
maps.google.com	2.20
coursera.org	2.04
x.com	2.00
github.com	1.85
facebook.com	1.77
netacad.com	1.75
canva.com	1.49
canvas.com	1.44
notion.so	1.13

Note: The table shows all environments with a query share above 1% among all agentic queries. docs.google.com includes Google Docs, Sheets, Slides, and Forms. All email domains are grouped into "email services combined" and all online messengers are grouped into "messenger services combined".

Table 7: Top Agent Environments

Usage context

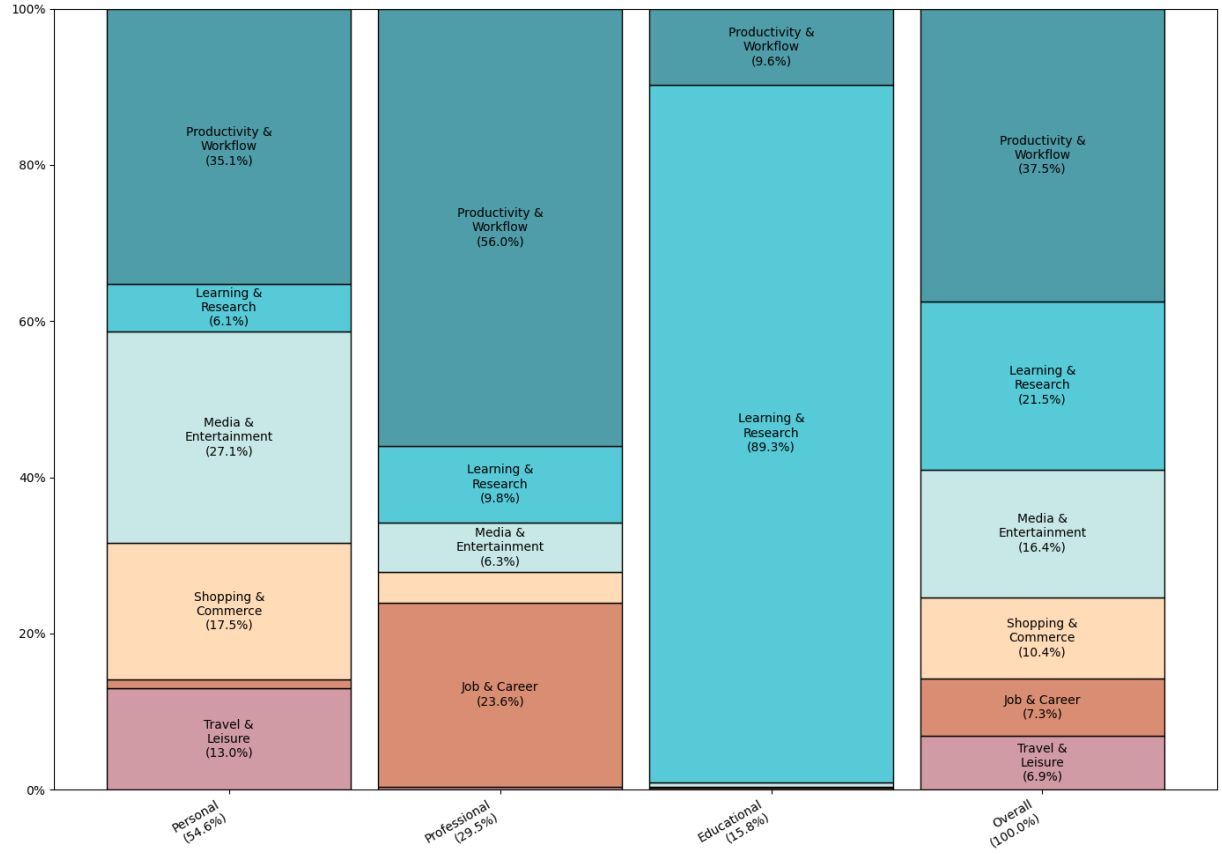
Lastly, we investigate agent usage across personal, professional, and educational contexts.[33] Personal use comprises about 55% of total agentic queries, with professional and educational

[32] docs.google.com includes Google Docs, Sheets, Slides, and Forms.

[33] Note that our use case taxonomy is largely orthogonal to the usage context. For instance, a user may ask the agent to reply to an email from a friend (personal), a colleague (professional), or a professor (educational). Similarly, users may ask the agent to shop for personal items, workplace equipments, or school supplies.

contexts representing 30% and 16%, respectively. There is a slight increase in the share of educational use and a slight decrease in personal use over time, while the share of professional use remains stable. This could be driven by the launch time of Comet overlapping with the beginning of the fall semester and promotional efforts on campus such as early access for university students, rather than a systematic shift in the underlying composition of user groups and use cases.

We show the distribution of topics by usage context in Figure 5. For personal usage, productivity and media together account for 62% of all agentic queries. For professional usage, 80% of agentic queries come from productivity and career-related queries. Educational usage is dominated by learning, comprising 89% of agentic queries. We also show the distribution of subtopic, task, and environment by usage context in Tables 26, 27, and 28. The top subtopics for personal, professional, and educational usage are shopping for goods, document editing, and courses, respectively. The top environments for personal, professional, and educational usage are emails, linkedin.com, and docs.google.com, respectively.



Note: The plot shows the percentage shares of topics within each usage context. "Other" category is removed. The labels for topics that account for less than 5% within a context are suppressed. Context shares are shown in the labels on the x axis. The same topic is shown in the same color across context. The topic percentage overall is shown as a baseline for comparison. The topics within each context are sorted by the same order as in overall for easier comparison across context. The topics in overall are sorted by topic percentage in decreasing order from top to bottom.

Figure 5: Topic Distribution by Usage Context

6 Discussion

Our paper provides the first systematic evidence on the adoption, usage intensity, and use cases of general-purpose AI agents using large-scale behavioral data from Comet by Perplexity. Our findings reveal substantial differences in the propensity to adopt and use the agent across user segments. Earlier adopters, users in countries with higher GDP per capita and higher average years of education, and individuals working in more digital and knowledge-intensive fields—such as digital technology, finance, marketing, and academics (including both students and educators)—tend to adopt and use the agent more actively. Agent use cases span a broad range of categories. The two largest topics—productivity and learning—together comprise 57% of all agentic queries. The two largest subtopics—courses and shopping for goods—together account for 22% of all agentic queries. The top 10 out of 90 tasks represent 55% of all agentic queries. We also document heterogeneity in use cases across occupation clusters that reflect the degree to which they align with each occupation's task composition. Topics such as productivity, learning, and career exhibit higher stickiness, as users are more likely to make consecutive queries within these categories. Over time, users also shift toward more productivity-oriented tasks. In addition, the environments in which the agentic queries are made show significant variation in their levels of concentration across topics and subtopics.

Although our paper is descriptive in nature and does not make normative claims, the methods and findings hold significant implications for researchers, businesses, policymakers, and educators. For researchers, we contribute to a nascent but rapidly expanding literature on the adoption and usage of LLMs and AI agents; and our agentic taxonomy provides a structure that future analysis can build on and extend. For firms developing AI agents, our results offer guidance on target user segments and high-frequency use cases. For businesses providing the environments in which agents operate, our findings suggest opportunities to streamline interfaces to better serve users interacting through AI agents. For both policymakers and educators, a central concern is that unequal adoption and usage of AI agents could exacerbate existing productivity and learning disparities. Consequently, equipping citizens and students with the skills to effectively leverage AI agents, and preparing them for a near future in which such agents are embedded in work and everyday life, will become increasingly important.

We note a few important caveats of our dataset. First, because Comet is a new product, our sample primarily reflects early adopters, who may skew toward more tech-savvy users. We characterize these early adopters using an internal survey in Appendix C. Relatedly, given the short time span of our data, we do not systematically investigate changes in usage patterns over time, and any longitudinal results should be interpreted within this context. Second, the classification of an agentic query depends on internal query understanding modules that trigger the agent based on predicted query intent. These intent predictors show high prediction accuracy in internal validation studies; nonetheless, the data may include both false positives (when a non-agentic intent triggers the agent) and false negatives (when an agentic intent does not trigger the agent).[34]

[34]Two sets of classifiers determine agent activation: one for browser control and one for specific apps. The browser control classifier is a supervised model trained on user queries labeled as either showing or not showing agentic intent. Five features are used for prediction: the current query, the currently viewed page, the user's previous queries in the same conversation, enabled connectors, and the number of attachments. The classifier achieves a ROC-AUC of 0.95 and both precision and recall of 0.90 at the optimal threshold. Each connector additionally has its own

Similarly, the classification into our agentic taxonomy, usage context, and occupation clusters also contain noise. Third, although AI agents' autonomy and task horizons continue to expand, our results should not be interpreted as suggesting any particular balance between automation and augmentation in use cases. For instance, an agentic session may appear to be automation, but it is possible that users are breaking a task into smaller pieces and only delegating some subtasks to the agent, which is closer to an augmentation case. A comprehensive treatment on such a topic would require having a full picture of how users manage their workflows outside of Comet.

There are several natural extensions to this study that we aim to pursue. First, with the expansion of Comet to mobile devices and other environments, it will be valuable to document cross-platform differences in how users interact with AI agents.[35] In particular, whereas agentic queries on desktops are predominantly text-based, the voice mode on mobile may offer a more natural interface. Second, while our sample does not capture enterprise users directly, the substantial share of professionally-oriented agentic queries suggests the need for complementary research on related topics in organizational settings. Third, the adoption and usage of AI agents is closely tied to their performance across tasks; we plan to investigate agent evaluation, common failure modes, and strategies for improvement. Fourth, which types of tasks are best suited for delegation to the agent and how to design optimal human-agent collaborative workflows is another important question. For example, tasks that users can easily complete manually may not warrant delegation, while high-stakes or irreversible tasks might require exceptionally reliable agent performance, a high degree of user trust, and an increased level of human supervision. Fifth, besides awareness and performance, another key barrier to adoption and usage is measuring and substantiating value and impact; we seek to quantify how much economic value users derive from agent usage, an important dimension of the downstream impact of AI agents.

General-purpose AI agents represent one of the most consequential technological advancements of our time. Understanding its real-world adoption and usage with large-scale behavioral data has become both urgent and essential for informing its development and deployment. We hope this work catalyzes further investigation in this rapidly evolving domain.

classifier, following a similar process.

[35] The Android version of Comet was launched to everyone worldwide on November 20, 2025, with the iOS version scheduled to be released in December 2025. https://www.perplexity.ai/hub/blog/comet-for-android-is-here

References

A. Allouah, O. Besbes, J. D. Figueroa, Y. Kanoria, and A. Kumar. What is your ai agent buying? evaluation, implications and emerging questions for agentic e-commerce. *arXiv preprint arXiv:2508.02630*, 2025.

Anthropic. Anthropic economic index: Ai's impact on software development, 2025. URL https://www.anthropic.com/research/impact-software-development.

R. Appel, P. McCrory, A. Tamkin, M. Stern, M. McCain, and T. Neylon. Anthropic economic index report: Uneven geographic and enterprise ai adoption, 2025. URL www.anthropic.com/research/anthropic-economic-index-september-2025-report.

D. Bent, K. Handa, E. Durmus, A. Tamkin, M. McCain, S. Ritchie, R. Donegan, J. Martinez, and J. Jones. Anthropic education report: How educators use claude, 2025. URL https://www.anthropic.com/news/anthropic-education-report-how-educators-use-claude.

A. Chatterji, T. Cunningham, D. J. Deming, Z. Hitzig, C. Ong, C. Y. Shan, and K. Wadman. How people use chatgpt. Technical report, National Bureau of Economic Research, 2025.

G. K. Hadfield and A. Koh. An economy of ai agents. *arXiv preprint arXiv:2509.01063*, 2025.

K. Handa, D. Bent, A. Tamkin, M. McCain, E. Durmus, M. Stern, M. Schiraldi, S. Huang, S. Ritchie, S. Syverud, K. Jagadish, M. Vo, M. Bell, and D. Ganguli. Anthropic education report: How university students use claude, 2025a. URL https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude.

K. Handa, A. Tamkin, M. McCain, S. Huang, E. Durmus, S. Heck, J. Mueller, J. Hong, S. Ritchie, T. Belonax, et al. Which economic tasks are performed with ai? evidence from millions of claude conversations. *arXiv preprint arXiv:2503.04761*, 2025b.

H. Ju and S. Aral. Collaborating with ai agents: Field experiments on teamwork, productivity, and performance. *arXiv preprint arXiv:2503.18238*, 2025.

D. M. Rothschild, M. Mobius, J. M. Hofman, E. W. Dillon, D. G. Goldstein, N. Immorlica, S. Jaffe, B. Lucier, A. Slivkins, and M. Vogel. The agentic economy. *arXiv preprint arXiv:2505.15799*, 2025.

R. Sapkota, K. I. Roumeliotis, and M. Karkee. Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges. *arXiv preprint arXiv:2505.10468*, 2025.

S. K. Sarkar. Ai agents, productivity, and higher-order thinking: Early evidence from software development. *Available at SSRN 5713646*, 2025.

P. Shahidi, G. Rusak, B. S. Manning, A. Fradkin, and J. J. Horton. The coasean singularity? demand, supply, and market design with ai agents. Technical report, National Bureau of Economic Research, 2025.

M. Wooldridge and N. R. Jennings. Intelligent agents: Theory and practice. *The knowledge engineering review*, 10(2):115–152, 1995.

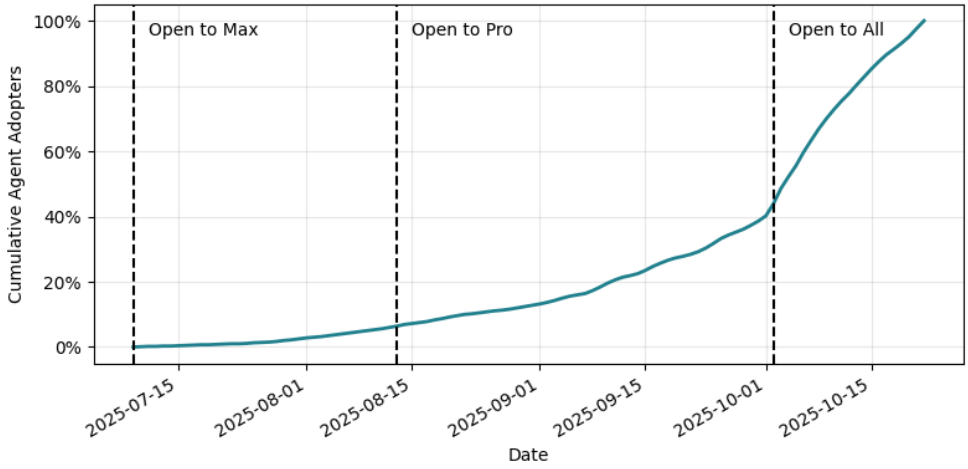
S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. R. Narasimhan, and Y. Cao. React: Synergizing reasoning and acting in language models. In *The eleventh international conference on learning representations*, 2022.

W. Zhao, X. Ren, J. Hessel, C. Cardie, Y. Choi, and Y. Deng. Wildchat: 1m chatgpt interaction logs in the wild. *arXiv preprint arXiv:2405.01470*, 2024.

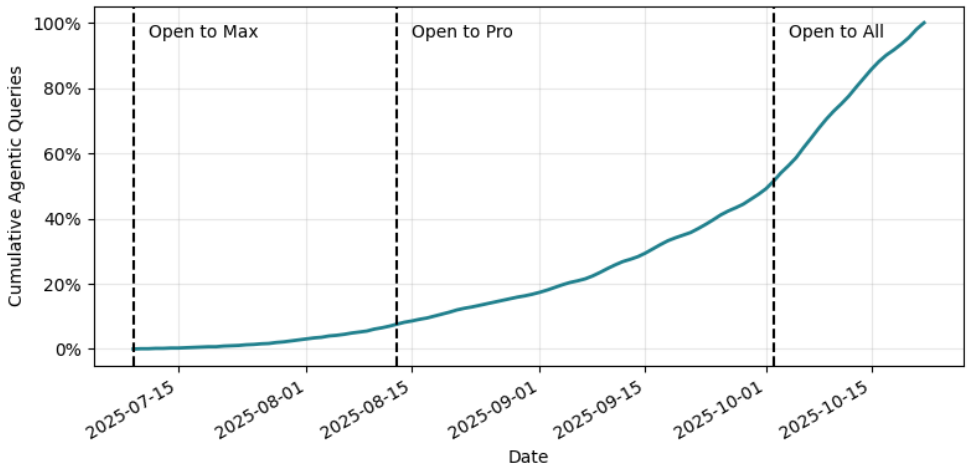
Appendices

A Figures and Tables

A.1 Adoption and Usage Intensity



(a) Cumulative Agent Adopter



(b) Cumulative Agentic Query

Note: The plots show the cumulative number of agent adopters and agentic queries. The exact numbers on the y-axis are masked and the percentages show the relative magnitude of adopters and queries relative to the end date as the baseline. Adopter and query numbers grow steadly over time with a noticable jump when Comet became generally available. The three dashed vertical lines mark the dates for change in access. July 9: launch date and opent to Max subscribers and selected users on a waitlist. August 13: extended access to Pro subscribers. October 2: extended access to everyone worldwide.

Figure 6: Cumulative Agent Adopter and Query

Table 8: Agent Adoption by Occupation Subcluster

Agent Adoption by Occupation Subcluster: The Top 10 Subclusters by Agent Adoption Share				
Cluster	**Subcluster**	**User Share (%)**	**Agent Adopter Share (%)**	**AAR**
Digital Technology	Software Development & Engineering	13.0	13.9	1.07
Marketing & Sales	Digital Marketing & Social Media	3.9	4.7	1.21
Digital Technology	It Support & Infrastructure	4.0	4.3	1.08
Arts, Entertainment, & Design	Design & Digital Arts	5.0	4.3	0.86
Financial Services	Financial Planning & Analysis	4.1	4.3	1.05
Digital Technology	Data Science & AI	3.9	4.0	1.03
Education	Teaching & Instruction	4.9	4.0	0.82
Financial Services	Financial Strategy & Investments	3.8	3.9	1.03
Management & Entrepreneurship	Business Information Management	3.2	3.8	1.19
Management & Entrepreneurship	Leadership & Operations	2.6	3.1	1.19
Agent Adoption by Occupation Subcluster: The Top 10 Subclusters by AAR				
Cluster	**Subcluster**	**User Share (%)**	**Agent Adopter Share (%)**	**AAR**
Marketing & Sales	Business Development & Sales	1.2	1.6	1.33
Marketing & Sales	Digital Marketing & Social Media	3.9	4.7	1.21
Management & Entrepreneurship	Business Information Management	3.2	3.8	1.19
Management & Entrepreneurship	Leadership & Operations	2.6	3.1	1.19
Marketing & Sales	Market Research, Analytics, & Ethics	1.1	1.3	1.18
Management & Entrepreneurship	Strategy & Consulting	2.2	2.5	1.14
Supply Chain & Transportation	Planning & Logistics	1.1	1.2	1.09
Digital Technology	It Support & Infrastructure	4.0	4.3	1.08
Digital Technology	Software Development & Engineering	13.0	13.9	1.07
Marketing & Sales	Brand Management & Strategy	1.4	1.5	1.07

Note: The tables show the top 10O*NET occupation subclusters by adoption. "Other" category is removed. When a user appears in multiple subclusters, their data is used in all relevant subclusters. User share is the number of users in each subcluster divided by total users. Agent adopter share is the number of adopters in each subcluster divided by the total adopters. AAR (Agent Adoption Ratio) is the ratio between agent adopter share and user share. AAR greater (less) than 1 indicates that a subcluster is over-represented (under-represented) in agent adoption relative to their user base. AAR rank is among subclusters with over 1% of user share.

Table 9: Agentic Query by Occupation Subcluster

Agentic Query by Occupation Subcluster: The Top 10 Subclusters by Agentic Query Share				
Cluster	**Subcluster**	**User Share (%)**	**Agentic Query Share (%)**	**AUR**
Digital Technology	Software Development & Engineering	13.0	15.4	1.18
Marketing & Sales	Digital Marketing & Social Media	3.9	5.9	1.51
Digital Technology	It Support & Infrastructure	4.0	5.3	1.32
Management & Entrepreneurship	Business Information Management	3.2	4.7	1.47
Financial Services	Financial Planning & Analysis	4.1	3.9	1.05
Digital Technology	Data Science & AI	3.9	3.9	1.00
Management & Entrepreneurship	Leadership & Operations	2.6	3.9	1.50
Arts, Entertainment, & Design	Design & Digital Arts	5.0	3.8	0.76
Education	Teaching & Instruction	4.9	3.6	0.73
Financial Services	Financial Strategy & Investments	3.8	3.9	1.03
Agentic Query by Occupation Subcluster: The Top 10 Subclusters by AUR				
Cluster	**Subcluster**	**User Share (%)**	**Agentic Query Share (%)**	**AUR**
Marketing & Sales	Business Development & Sales	1.2	2.1	1.75
Marketing & Sales	Digital Marketing & Social Media	3.9	5.9	1.51
Management & Entrepreneurship	Leadership & Operations	2.6	3.9	1.50
Management & Entrepreneurship	Business Information Management	3.2	4.7	1.47
Marketing & Sales	Brand Management & Strategy	1.4	1.9	1.36
Management & Entrepreneurship	Strategy & Consulting	2.2	3.0	1.36
Digital Technology	It Support & Infrastructure	4.0	5.3	1.32
Marketing & Sales	Market Research, Analytics, & Ethics	1.1	1.4	1.27
Digital Technology	Software Development & Engineering	13.0	15.4	1.18
Supply Chain & Transportation	Planning & Logistics	1.1	1.1	1.00

Note: The tables show the top 10O*NET occupation subclusters by usage intensity. “Other” category is removed. When a user appears in multiple subclusters, their data is used in all relevant subclusters. User share is the number of users in each subcluster divided by total users. Agent query share is the number of agentic query in each subcluster divided by the total agentic query. AUR (Agent Usage Ratio) is the ratio between agentic query share and user share. AUR greater (less) than 1 indicates that a subcluster is over-represented (under-represented) in agent usage relative to their user base. AUR rank is among subclusters with over 1% of user share.

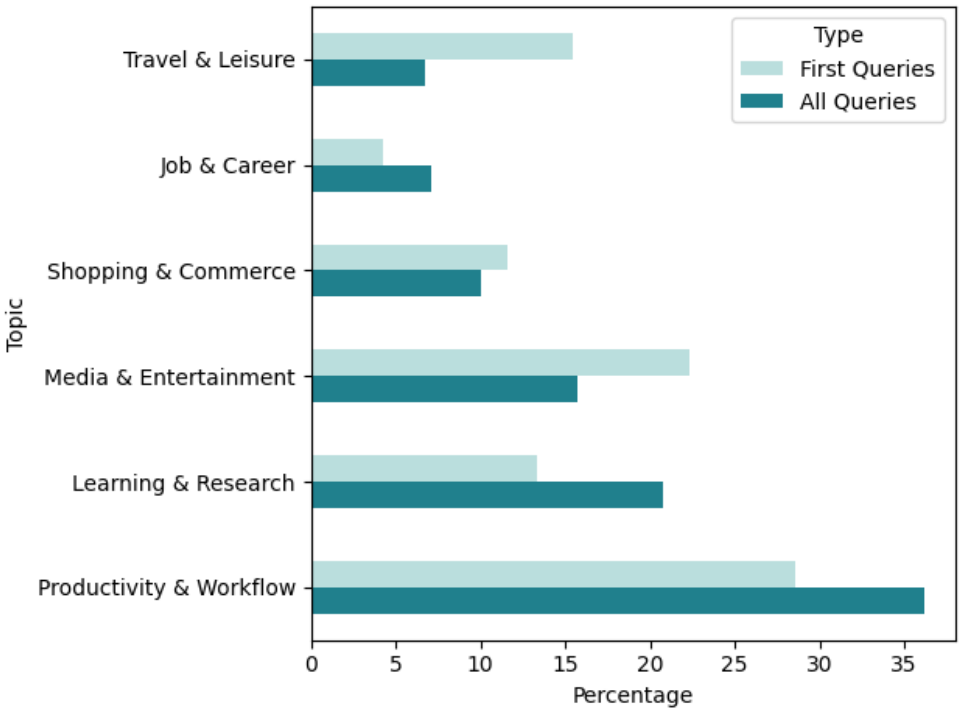
A.2 Use Cases

A.2.1 Topics and Subtopics

Topic	Topic (%)	Subtopic	Subtopic (%)	Overall (%)
Productivity & Workflow	36.2	Document & Form Editing	21.5	7.78
		Account Management	20.5	7.43
		Email Management	15.8	5.73
		Spreadsheet & Data Editing	11.1	4.01
		Computer Programming	10.3	3.73
		Investments & Banking	6.2	2.25
		Multimedia Editing	6.1	2.22
		Project Management	5.1	1.85
		Calendar Management	2.5	0.91
		Other	0.8	0.30
Learning & Research	20.8	Courses	61.9	12.86
		Research	37.9	7.88
		Other	0.2	0.04
Media & Entertainment	15.8	Social Media & Messaging	42.4	6.69
		Movies, TV, & Videos	20.1	3.17
		Online Games	19.6	3.08
		Music & Podcasts	10.7	1.68
		News	3.8	0.59
		Sports	2.7	0.42
		Other	0.8	0.13
Shopping & Commerce	10.0	Goods	89.0	8.94
		Services	10.3	1.03
		Other	0.7	0.07
Job & Career	7.1	Professional Networking	50.1	3.56
		Job Search & Application	49.5	3.52
		Other	0.4	0.03
Travel & Leisure	6.7	Flights & Transportation	40.7	2.73
		Lodging	27.3	1.83
		Trip Itineraries	24.8	1.66
		Restaurants	5.7	0.38
		Other	1.5	0.10
Other	3.4	Other	100.0	3.42

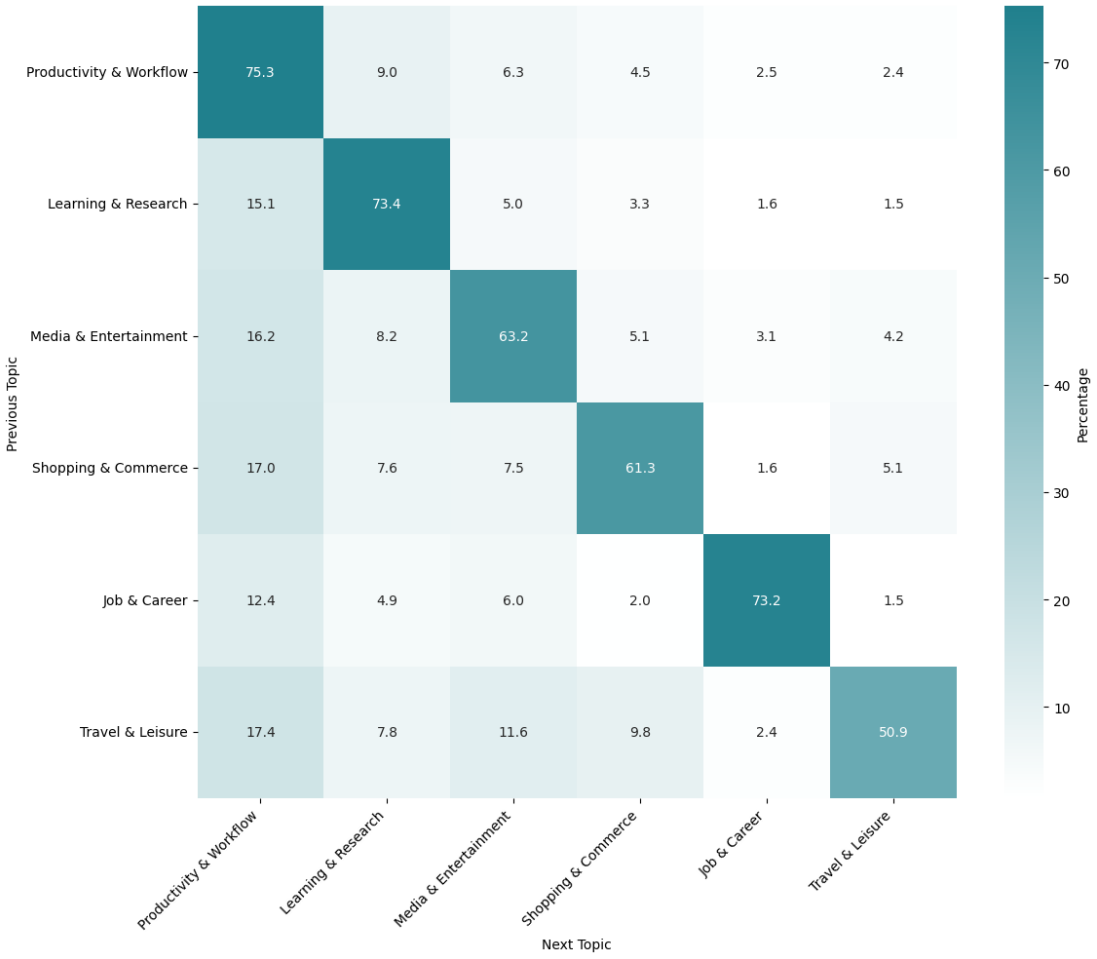
Note: The table shows the distribution of topics and subtopics. Topic percentage ($P(\text{Topic})$) is the topic share among all agentic queries. Subtopic percentage ($P(\text{Subtopic} \mid \text{Topic})$) is the percentage of a subtopic within a topic. Overall percentage ($P(\text{Subtopic}) = P(\text{Topic}, \text{Subtopic}) = P(\text{Topic}) \times P(\text{Subtopic} \mid \text{Topic})$) is the subtopic share among all agentic queries. Percentages may not sum to 100% due to rounding.

Table 10: Topic and Subtopic Distribution



Note: The plot shows the topic distribuiton among users' first queries versus all queries. Over time, the query shares are shifting from Travel & Leisure, and Media & Entertainment to Productivity & Workflow, Learning & Research, and Job & Career. Shopping & Commerce share stays relatively stable.

Figure 7: Topic Distribution by First vs. All Queries



Note: The plot shows the transition matrix from the previous query to the next query, aggregated from the user level. Most query topics transition into themselves (the off diagonal). Other than themselves, other topics are most likely to transition into Productivity & Workflow (the first column). Productivity & Workflow, Learning & Research, and Job & Carrer are the most sticky with the highest self-transition probabilities. Whereas Travel & Leisure is the least sticky, and Media & Entertainment and Shopping & Commerce are in between. The steady state probability distribution based on this transition matrix and the observed share are closely matched—39% vs 37% respectively for Productivity & Workflow, 24% vs 22% for Learning & Research, 16% vs. 16% for Media & Entertainment, 10% vs. 10% for Shopping & Commerce, 7% vs 7% for Job & Career, and 7% vs 7% for Travel & Leisure. Note that to be aligned with the transition matrix, the observed shares used here do not include the "Other" category in Table 10, the shares used here are based on the topic shares in that table divided by 0.96.

Figure 8: Topic Transition Matrix: Previous vs. Next Agentic Query

A.2.2 Tasks

Topic	Subtopic	Task	Task (%)	Overall (%)
Productivity & Workflow (36.2)	Document & Form Editing (21.5)	Create/edit documents/forms	84.6	6.58
		Summarize/analyze documents/forms	24.1	1.87
		Search/filter documents/forms	11.4	0.89
	Account Management (20.5)	Manage settings/profiles	58.3	4.33
		Register/log in to accounts	31.2	2.32
		Summarize/analyze account information	24.0	1.79
		Manage files	15.8	1.18
	Email Management (15.8)	Search/filter emails	49.1	2.81
		Create/edit emails	32.8	1.88
		Delete/unsubscribe emails	30.6	1.75
		Summarize/analyze emails	22.8	1.31
		Send emails	9.6	0.55
	Spreadsheet & Data Editing (11.1)	Create/edit spreadsheets/data	72.5	2.91
		Summarize/analyze spreadsheets/data	38.7	1.55
		Search/filter spreadsheets/data	27.5	1.10
	Computer Programming (10.3)	Create/edit code	63.8	2.38
		Summarize/analyze code	48.5	1.81
		Execute code	20.3	0.76
	Investments & Banking (6.2)	Summarize/analyze investment information	75.0	1.69
		Search/filter stocks	28.8	0.65
		Summarize/analyze banking information	11.7	0.26
		Buy/sell stocks	7.8	0.18
	Multimedia Editing (6.1)	Create/edit multimedia	81.0	1.80
		Summarize/analyze multimedia	23.6	0.52
		Search/filter multimedia	15.3	0.34
	Project Management (5.1)	Create/edit projects	64.6	1.19
		Summarize/analyze project information	48.0	0.89
	Calendar Management (2.5)	Create/edit events	71.0	0.64
		Search/filter events	24.6	0.22
		Summarize/analyze events	22.0	0.20

Note: The table shows all tasks under Productivity & Workflow with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage (P(Task | Topic, Subtopic)) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. P(Task Overall) $=$ P(Topic, Subtopic, Task) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Task | Topic, Subtopic)). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 11: Task Distribution for Productivity & Workflow

Topic	Subtopic	Task	Task (%)	Overall (%)
Learning & Research (20.8)	Courses (61.9)	Assist exercises	73.2	9.41
		Summarize/analyze course materials	28.7	3.69
		Navigate courses	25.6	3.29
	Research (37.9)	Summarize/analyze research information	85.2	6.71
		Search/filter research information	75.6	5.95

Note: The table shows all tasks under Learning & Research with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage (P(Task | Topic, Subtopic)) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. P(Task Overall) $=$ P(Topic, Subtopic, Task) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Task | Topic, Subtopic)). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 12: Task Distribution for Learning & Research

Topic	Subtopic	Task	Task (%)	Overall (%)
Media & Entertainment (15.8)	Social Media & Messaging (42.4)	Search/filter social media posts/messages	49.5	3.31
		Summarize/analyze social media posts/messages	35.3	2.36
		Create social media posts/messages	34.0	2.28
		Engage with social media posts/messages	29.3	1.96
		Send social media/text messages	20.7	1.39
	Movies, TV, & Videos (20.1)	Search/filter videos	48.4	1.53
		Summarize/analyze videos	43.8	1.39
		Play videos	27.1	0.86
		Navigate within videos	18.4	0.58
		Manage playlists	6.2	0.20
	Online Games (19.6)	Play online games	76.8	2.37
		Summarize/analyze online game information	30.7	0.95
		Search/filter online games	14.5	0.45
	Music & Podcasts (10.7)	Search/filter music/podcasts	75.4	1.27
		Play music/podcasts	61.2	1.03
		Manage playlists	25.5	0.43
		Summarize/analyze music/podcasts	9.7	0.16
	News (3.8)	Search/filter news	70.7	0.42
		Summarize/analyze news	56.6	0.34
	Sports (2.7)	Summarize/analyze match/player information	77.0	0.32
		Search/filter match/player information	67.7	0.28

Note: The table shows all tasks under Media & Entertainment with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage (P(Task | Topic, Subtopic)) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. P(Task Overall) $=$ P(Topic, Subtopic, Task) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Task | Topic, Subtopic)). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 13: Task Distribution for Media & Entertainment

Topic	Subtopic	Task	Task (%)	Overall (%)
Shopping & Commerce (10.0)	Goods (89.0)	Search/filter products	71.9	6.43
		Summarize/analyze product information	57.9	5.18
		Add products to cart	19.8	1.77
		Search discounts	10.2	0.92
	Services (10.3)	Search/filter products	54.5	0.56
		Summarize/analyze product information	45.1	0.46
		Make product purchase	20.2	0.21
		Search discounts	12.5	0.13
		Add products to cart	7.6	0.08
		Manage orders	7.0	0.07

Note: The table shows all tasks under Shopping & Commerce with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage (P(Task | Topic, Subtopic)) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. P(Task Overall) $=$ P(Topic, Subtopic, Task) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Task | Topic, Subtopic)). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 14: Task Distribution for Shopping & Commerce

Topic	Subtopic	Task	Task (%)	Overall (%)
Job & Career (7.1)	Job Search & Application (50.1)	Complete applications	65.7	2.31
		Search/filter jobs	57.1	2.01
		Summarize/analyze job descriptions	26.3	0.93
	Professional Networking (49.5)	Search/filter professional profiles	53.8	1.92
		Summarize/analyze professional profiles	39.9	1.42
		Send professional connection requests/messages	31.4	1.12
		Engage with professional profiles/posts	24.0	0.85

Note: The table shows all tasks under Job & Career with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage (P(Task | Topic, Subtopic)) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. P(Task Overall) $=$ P(Topic, Subtopic, Task) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Task | Topic, Subtopic)). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 15: Task Distribution for Job & Career

Topic	Subtopic	Task	Task (%)	Overall (%)
Travel & Leisure (6.7)	Flights & Transportation (40.7)	Search/filter flights & transportation	93.4	2.55
		Summarize/analyze flights & transportation	63.1	1.72
		Book flights & transportation	9.5	0.26
	Lodging (27.3)	Search/filter lodging	92.9	1.70
		Summarize/analyze lodging information	67.5	1.23
		Book lodging	5.5	0.10
	Trip Itineraries (24.8)	Plan trips	87.4	1.45
		Summarize/analyze trips	48.9	0.81
		Search/filter destinations	48.8	0.81
	Restaurants (5.7)	Search/filter restaurants	73.8	0.28
		Summarize/analyze restaurant information	54.2	0.21
		Book restaurants	25.0	0.09
		Manage bookings	5.7	0.02

Note: The table shows all tasks under Travel & Leisure with more than 5% share within the subtopic. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Task percentage ($P(\text{Task} \mid \text{Topic, Subtopic})$) is the task share within the subtopic. Overall percentage is the task share among all agentic queries. $P(\text{Task Overall}) = P(\text{Topic, Subtopic, Task}) = P(\text{Topic}) \times P(\text{Subtopic} \mid \text{Topic}) \times P(\text{Task} \mid \text{Topic, Subtopic})$). Note that because task percentage measures the fraction of queries in which a task is present and a query might contain multiple tasks, the task percentages under each subtopic do not sum to 100.

Table 16: Task Distribution for Travel & Leisure

Cluster	Task	Task (%)
Digital Technology	Assist exercises	9.1
	Search/filter products	6.4
	Create/edit documents/forms	6.2
	Summarize/analyze research information	6.2
	Manage settings/profiles	5.7
Student	Assist exercises	26.5
	Summarize/analyze course materials	10.7
	Navigate courses	7.8
	Create/edit documents/forms	7.1
	Summarize/analyze research information	5.4
Management & Entrepreneurship	Summarize/analyze research information	7.7
	Create/edit documents/forms	7.5
	Search/filter research information	6.8
	Search/filter products	6.3
	Create/edit spreadsheets/data	5.2
Marketing & Sales	Create social media posts/messages	8.1
	Search/filter social media posts/messages	8.1
	Summarize/analyze product information	8.1
	Summarize/analyze research information	8.0
	Search/filter products	7.6
Financial Services	Summarize/analyze investment information	11.9
	Summarize/analyze research information	7.0
	Search/filter products	6.7
	Search/filter research information	6.1
	Summarize/analyze product information	5.2
Education	Assist exercises	19.6
	Create/edit documents/forms	9.8
	Summarize/analyze course materials	8.6
	Summarize/analyze research information	7.0
	Search/filter research information	6.9
Arts, Entertainment, & Design	Search/filter products	7.8
	Create/edit documents/forms	7.4
	Create/edit multimedia	7.3
	Summarize/analyze product information	6.1
	Summarize/analyze research information	5.8
Healthcare & Human Services	Summarize/analyze research information	9.5
	Search/filter research information	9.4
	Create/edit documents/forms	7.9
	Search/filter products	6.6
	Assist exercises	5.5
Advanced Manufacturing	Search/filter products	20.9
	Summarize/analyze product information	18.1
	Summarize/analyze research information	7.8
	Search/filter research information	7.5
	Assist exercises	6.6
Public Service & Safety	Search/filter research information	15.7
	Summarize/analyze research information	15.1
	Create/edit documents/forms	10.2
	Search/filter products	5.1
	Assist exercises	4.3
Hospitality, Events, & Tourism	Search/filter flights & transportation	11.9
	Search/filter lodging	9.6
	Search/filter products	8.3
	Summarize/analyze flights & transportation information	8.1
	Summarize/analyze lodging information	7.6
Supply Chain & Transportation	Search/filter products	12.7
	Summarize/analyze product information	10.6
	Create/edit documents/forms	7.4
	Summarize/analyze research information	7.2
	Search/filter research information	6.4
Construction	Search/filter products	11.0
	Summarize/analyze product information	9.1
	Summarize/analyze research information	8.8
	Search/filter research information	8.3
	Create/edit documents/forms	5.6
Energy & Natural Resources	Create/edit documents/forms	10.5
	Summarize/analyze research information	10.4
	Search/filter research information	10.4
	Search/filter products	7.7
	Summarize/analyze product information	6.0
Agriculture	Create/edit documents/forms	10.3
	Search/filter products	8.5
	Summarize/analyze research information	8.4
	Search/filter research information	8.4
	Summarize/analyze product information	7.2

Note: The table shows the top 5 tasks under each occupation cluster. Task percentage ($P(\text{Task} \mid \text{Cluster})$) is the task share among all agentic queries within a cluster.

Table 17: The Top 5 Tasks by Occupation Cluster

A.2.3 Environments

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Productivity & Workflow (36.2)	Document & Form Editing (21.5)	docs.google.com	66.6	5.18
		notion.so	6.4	0.50
		canva.com	2.5	0.19
		overleaf.com	2.0	0.16
		perplexity.ai	1.3	0.10
	Account Management (20.5)	perplexity.ai	10.4	0.77
		docs.google.com	7.6	0.56
		settings	4.0	0.30
		github.com	3.2	0.24
		linkedin.com	3.1	0.23
	Email Management (15.8)	mail.google.com	69.9	4.00
		outlook.office.com	10.8	0.62
		outlook.live.com	2.9	0.17
		mail.yahoo.com	1.5	0.09
		mail.yandex.ru	0.5	0.03
	Spreadsheet & Data Editing (11.1)	docs.google.com	78.9	3.17
		notion.so	4.4	0.18
		airtable.com	2.3	0.09
		excel.cloud.microsoft	1.4	0.06
		app.powerbi.com	1.1	0.04
	Computer Programming (10.3)	github.com	30.7	1.14
		colab.research.google.com	5.3	0.20
		leetcode.com	4.9	0.18
		aistudio.google.com	4.2	0.16
		script.google.com	3.1	0.12
	Investments & Banking (6.2)	tradingview.com	47.3	1.06
		binance.com	5.7	0.13
		kite.zerodha.com	4.9	0.11
		groww.in	4.6	0.10
		perplexity.ai	4.4	0.10
	Multimedia Editing (6.1)	canva.com	42.9	0.95
		figma.com	8.6	0.19
		docs.google.com	5.3	0.12
		youtube.com	5.3	0.12
		aistudio.google.com	3.5	0.08
	Project Management (5.1)	app.clickup.com	9.6	0.18
		trello.com	8.1	0.15
		notion.so	7.3	0.13
		linear.app	6.5	0.12
		adsmanager.facebook.com	5.4	0.10
	Calendar Management (2.5)	calendar.google.com	50.3	0.45
		outlook.office.com	7.7	0.07
		meet.google.com	3.7	0.03
		mail.google.com	3.5	0.03
		teams.microsoft.com	1.4	0.01

Note: The table shows the top 5 environments under Productivity & Workflow. The share of topic among all agentic queries and the share of subtopic within a topic are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) = P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 18: The Top 5 Environments Distribution for Productivity & Workflow

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Learning & Research (20.8)	Courses (61.9)	coursera.org	18.0	2.32
		netacad.com	15.6	2.01
		canvas.com	12.6	1.62
		learning.mheducation.com	8.3	1.07
		docs.google.com	6.4	0.82
	Research (37.9)	youtube.com	17.8	1.40
		perplexity.ai	6.1	0.48
		github.com	5.8	0.46
		maps.google.com	5.4	0.43
		docs.google.com	3.9	0.31

Note: The table shows the top 5 environments under Learning & Research. The share of topic and subtopics among all agentic queries are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) = P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 19: The Top 5 Environments Distribution for Learning & Research

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Media & Entertainment (15.8)	Social Media & Messaging (42.4)	instagram.com	21.3	1.43
		x.com	18.0	1.21
		whatsapp.com	13.6	0.91
		facebook.com	10.1	0.68
		linkedin	6.1	0.41
	Movies, TV, & Videos (20.1)	youtube.com	89.9	2.85
		netflix.com	4.1	0.13
		in.bookmyshow.com	1.3	0.04
		twitch.tv	0.7	0.02
		tiktok.com	0.6	0.02
	Online Games (19.6)	chess.com	32.5	1.01
		store.steampowered.com	15.0	0.46
		nytimes.com	14.2	0.44
		roblox.com	6.8	0.21
		humanbenchmark.com	5.4	0.17
	Music & Podcasts (10.7)	open.spotify.com	46.1	0.78
		youtube.com	39.9	0.67
		suno.com	6.2	0.10
		soundcloud.com	2.7	0.05
		music.apple.com	2.4	0.04
	News (3.8)	youtube.com	21.0	0.13
		trends.google.com	13.6	0.08
		nytimes.com	8.7	0.05
		perplexity.ai	7.9	0.05
		x.com	5.8	0.03
	Sports (2.7)	youtube.com	28.3	0.12
		fantasy.espn.com	20.1	0.09
		sleeper.com	7.8	0.03
		sofascore.com	3.3	0.01
		livescore.in	1.8	0.01

Note: The table shows the top 5 environments under Media & Entertainment. The share of topic and subtopics among all agentic queries are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) = P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 20: The Top 5 Environments Distribution for Media & Entertainment

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Shopping & Commerce (10.0)	Goods (89.0)	amazon.com	43.2	3.84
		flipkart.com	6.2	0.55
		admin.shopify.com	5.3	0.47
		alibaba.com	3.7	0.33
		ozon.ru	3.4	0.30
	Services (10.3)	perplexity.ai	12.2	0.13
		maps.google.com	8.7	0.09
		amazon.com	5.1	0.05
		fiverr.com	5.1	0.05
		avito.ru	4.1	0.04

Note: The table shows the top 5 environments under Shopping & Commerce. The share of topic and subtopics among all agentic queries are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) $=$ P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 21: The Top 5 Environments Distribution for Shopping & Commerce

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Job & Career (7.1)	Professional Networking (50.1)	linkedin.com	92.5	3.29
		upwork.com	1.3	0.05
		app.apollo.io	0.8	0.03
		naukri.com	0.6	0.02
		instagram.com	0.5	0.02
	Job Search & Application (49.5)	linkedin.com	60.2	2.12
		naukri.com	6.3	0.22
		ziprecruiter.com	2.7	0.09
		indeed.com	2.5	0.09
		dice.com	2.4	0.08

Note: The table shows the top 5 environments under Job & Career. The share of topic and subtopics among all agentic queries are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) = P(Topic) $\times$ P(Subtopic | Topic) $\times$ P(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 22: The Top 5 Environments Distribution for Job & Career

Topic	Subtopic	Environment	Environment (%)	Overall (%)
Travel & Leisure (6.7)	Flights & Transportation (40.7)	skyscanner.com	35.6	0.97
		maps.google.com	18.4	0.50
		makemytrip.com	7.2	0.20
		irctc.co.in	6.1	0.17
		expedia.com	4.0	0.11
	Lodging (27.3)	booking.com	54.9	1.00
		airbnb.com	19.5	0.36
		expedia.com	4.6	0.08
		maps.google.com	4.0	0.07
		agoda.com	3.5	0.06
	Trip Itineraries (24.8)	maps.google.com	85.2	1.42
		docs.google.com	4.2	0.07
		yandex.ru	1.3	0.02
		skyscanner.com	1.2	0.02
		booking.com	0.8	0.01
	Restaurants (5.7)	maps.google.com	54.7	0.21
		opentable.com	6.5	0.02
		map.naver.com	5.2	0.02
		swiggy.com	4.6	0.02
		ubereats.com	3.9	0.01

Note: The table shows the top 5 environments under Travel & Leisure. The share of topic and subtopics among all agentic queries are shown in the parentheses. Environment percentage (P(Environment | Topic, Subtopic)) is the environment share within the subtopic. Overall percentage is the environment share among all agentic queries. P(Topic, Subtopic, Environment) $= P$(Topic) $\times P$(Subtopic | Topic) $\times P$(Environment | Topic, Subtopic). Note that unlike tasks, an environment is not unique to a subtopic, so P(Topic, Subtopic, Environment) is the share of an environment when it is used under that subtopic and does not equal to P(Environment), which is the share under all subtopics.

Table 23: The Top 5 Environments Distribution for Travel & Leisure

Topic	Subtopic	Sum of the Top 5 Environments (%)
Productivity & Workflow	Spreadsheet & Data Editing	88.1
	Email Management	85.1
	Document & Form Editing	78.8
	Investments & Banking	66.9
	Calendar Management	66.6
	Multimedia Editing	65.6
	Computer Programming	48.2
	Project Management	36.9
	Account Management	28.3
Learning & Research	Courses	60.9
	Research	39.0
Media & Entertainment	Music & Podcasts	97.3
	Movies, TV, & Videos	96.6
	Online Games	73.9
	Social Media & Messaging	69.1
	Sports	61.3
	News	57.0
Shopping & Commerce	Goods	61.8
	Services	35.2
Job & Career	Professional Networking	95.7
	Job Search & Application	74.1
Travel & Leisure	Trip Itineraries	92.7
	Lodging	86.5
	Restaurants	74.9
	Flights & Transportation	71.3

Note: The table shows the sum of the top 5 environments shares by topic and subtopic. A higher (lower) share indicates agentic queries are more (less) concentrated in a small number of environments. It can be interpreted as the agent use market share of environments among agent adopters on Comet.

Table 24: Sum of the Top 5 Environment Shares by Topic and Subtopic

Cluster	Environment	Environment (%)
Digital Technology	linkedin.com	6.5
	email services combined	5.4
	docs.google.com	4.0
	youtube.com	3.3
	amazon.com	2.4
Student	docs.google.com	7.8
	linkedin.com	6.1
	email services combined	4.3
	canvas.com	3.4
	youtube.com	3.0
Management & Entrepreneurship	linkedin.com	10.3
	email services combined	8.8
	docs.google.com	7.7
	youtube.com	2.0
	amazon.com	1.8
Marketing & Sales	linkedin.com	7.8
	docs.google.com	6.8
	instagram.com	6.5
	x.com	5.2
	email services combined	4.5
Financial Services	email services combined	6.6
	docs.google.com	5.0
	linkedin.com	4.3
	youtube.com	3.6
	tradingview.com	2.6
Education	docs.google.com	9.9
	email services combined	6.5
	youtube.com	4.4
	canvas.com	2.4
	amazon.com	2.1
Arts, Entertainment, & Design	youtube.com	7.0
	email services combined	5.1
	docs.google.com	4.9
	linkedin.com	3.6
	instagram.com	3.2
Healthcare & Human Services	email services combined	7.3
	docs.google.com	5.8
	linkedin.com	4.7
	youtube.com	3.0
	amazon.com	2.5
Advanced Manufacturing	email services combined	5.5
	linkedin.com	4.6
	docs.google.com	4.3
	youtube.com	3.7
	amazon.com	3.4
Public Service & Safety	email services combined	6.2
	docs.google.com	4.9
	youtube.com	4.1
	trends.google.com	3.9
	linkedin.com	2.8
Hospitality, Events, & Tourism	email services combined	6.4
	maps.google.com	5.6
	docs.google.com	5.4
	booking.com	3.6
	skyscanner.com	2.9
Supply Chain & Transportation	email services combined	8.0
	docs.google.com	4.9
	linkedin.com	3.7
	amazon.com	3.0
	maps.google.com	2.5
Construction	email services combined	8.2
	linkedin.com	6.3
	docs.google.com	4.3
	youtube.com	3.1
	amazon.com	2.8
Energy & Natural Resources	email services combined	7.4
	docs.google.com	6.6
	linkedin.com	3.8
	amazon.com	3.4
	youtube.com	2.9
Agriculture	docs.google.com	7.4
	email services combined	6.9
	youtube.com	3.6
	linkedin.com	3.6
	amazon.com	2.4

Note: The table shows the top 5 environments by occupation cluster. Environment percentage ($P(\text{Environment} \mid \text{Cluster})$) is the environment share among all agentic queries in that cluster. docs.google.com includes Google Docs, Sheets, Slides, and Forms. All email accounts are grouped into "email services combined".

Table 25: The Top 5 Environments by Occupation Cluster

A.2.4 Usage Context

Context	Topic	Subtopic	Subtopic (%)
Personal	Shopping & Commerce	Goods	15.6
	Media & Entertainment	Social Media & Messaging	9.9
	Productivity & Workflow	Account Management	8.0
	Productivity & Workflow	Email Management	7.6
	Media & Entertainment	Online Games	6.0
Professional	Productivity & Workflow	Document & Form Editing	13.3
	Job & Career	Professional Networking	12.5
	Job & Career	Job Search & Application	11.0
	Productivity & Workflow	Account Management	10.2
	Learning & Research	Research	8.9
Educational	Learning & Research	Courses	83.9
	Learning & Research	Research	5.3
	Productivity & Workflow	Document & Form Editing	5.0
	Productivity & Workflow	Account Management	1.1
	Productivity & Workflow	Computer Programming	0.9

Note: The table shows the distribution of the top 5 subtopics by usage context. Subtopic percentage ($P(\text{Subtopic} \mid \text{Context})$) is the subtopic share among all agentic queries in a given usage context.

Table 26: The Top 5 Subtopic Distribution by Usage Context

Context	Topic	Subtopic	Task	Task (%)
Personal	Shopping & Commerce	Goods	Search/filter products	8.6
	Shopping & Commerce	Goods	Summarize/analyze product information	6.5
	Travel & Leisure	Flights & Transportation	Search/filter flights & transportation	3.4
	Media & Entertainment	Social Media & Messaging	Search/filter social media posts/messages	3.4
	Productivity & Workflow	Document & Form Editing	Create/edit documents/forms	3.2
Professional	Productivity & Workflow	Document & Form Editing	Create/edit documents/forms	8.1
	Learning & Research	Research	Summarize/analyze research information	5.9
	Job & Career	Job Search & Application	Complete applications	5.4
	Productivity & Workflow	Account Management	Manage settings/profiles	5.1
	Job & Career	Professional Networking	Search/filter professional profiles	4.8
Educational	Learning & Research	Courses	Assist exercises	48.1
	Learning & Research	Courses	Summarize/analyze course materials	18.8
	Learning & Research	Courses	Navigate courses	16.0
	Productivity & Workflow	Document & Form Editing	Create/edit documents/forms	3.6
	Learning & Research	Research	Summarize/analyze research information	3.4

Note: The table shows the distribution of the top 5 tasks by usage context. Task percentage ($P(\text{Task} \mid \text{Context})$) is the task share among all agentic queries in a given usage context.

Table 27: The Top 5 Task Distribution by Usage Context

Context	Environment	Environment (%)
Personal	email services combined	14.5
	youtube.com	10.8
	docs.google.com	10.7
	amazon.com	6.3
	maps.google.com	3.8
Professional	linkedin.com	29.6
	docs.google.com	11.4
	email services combined	9.6
	github.com	3.8
	admin.shopify.com	2.8
Educational	docs.google.com	14.8
	coursera.org	14.6
	netacad.com	12.7
	canvas.com	10.2
	learning.mheducation.com	6.8

Note: The table shows the distribution of the top 5 environments by usage context. Environment percentage ($P(\text{Environment} \mid \text{Context})$) is the environment share among all agentic queries in a given usage context. docs.google.com includes Google Docs, Sheets, Slides, and Forms. All email domains are grouped into "email services combined".

Table 28: The Top 5 Environment Distribution by Usage Context

B Agent Demo

B.1 Sample Agentic Queries

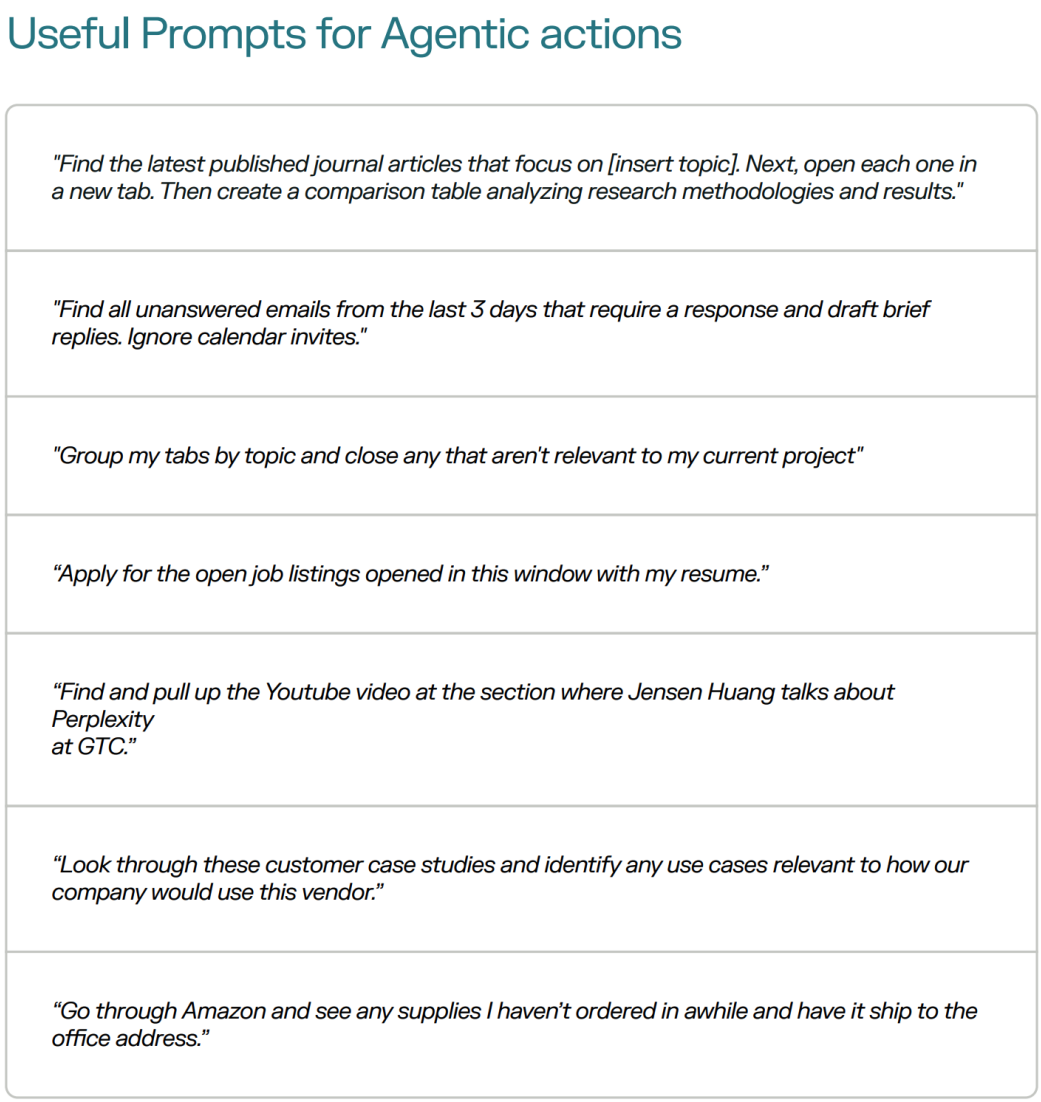


Figure 9: Sample Agentic Queries from Perplexity at Work

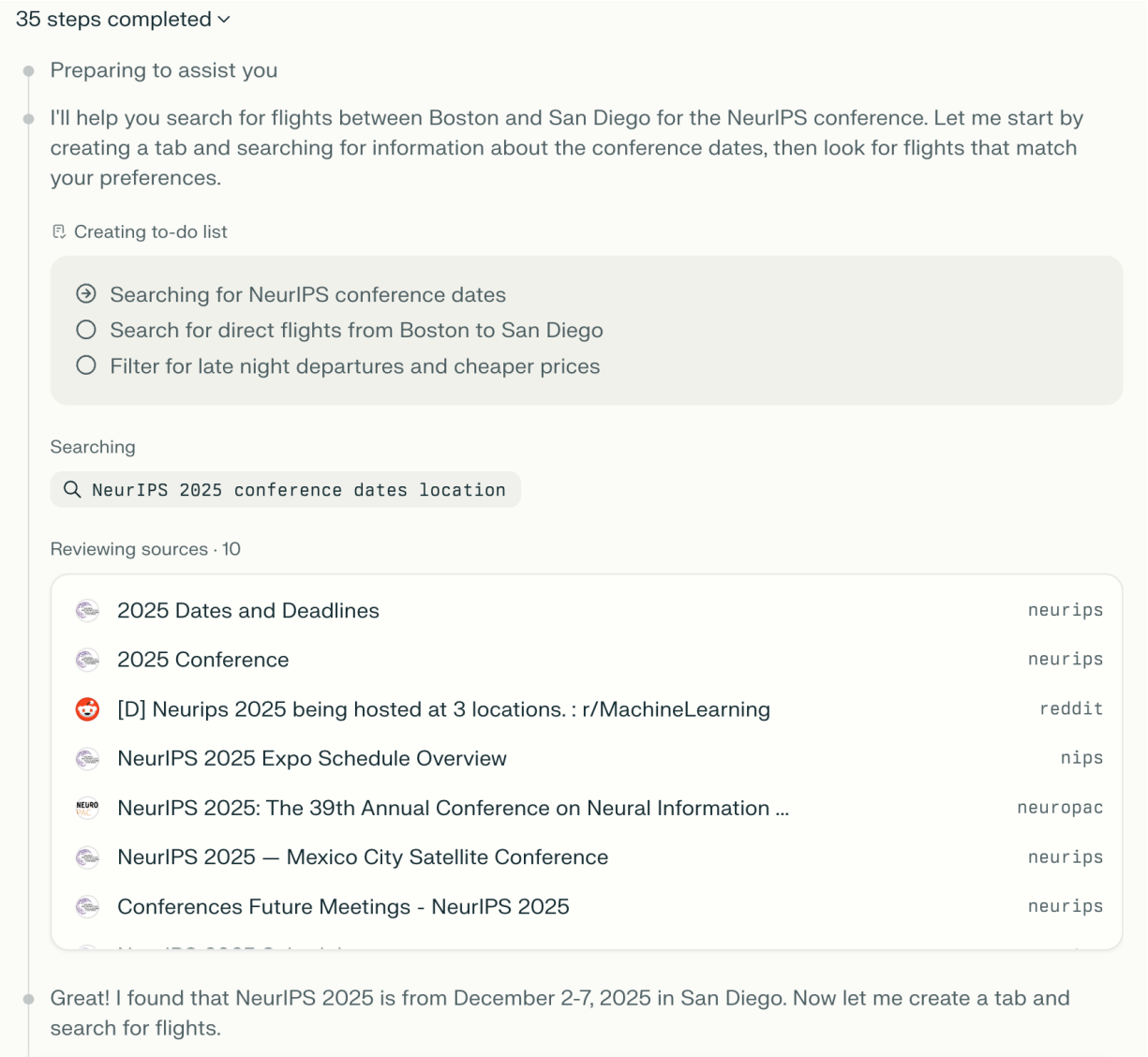
Sample Query	Topic	Subtopic	Task
Find the latest published journal articles...	Learning & Research	Research	Search/filter research information, Summarize/analyze research information
Find all unanswered emails...	Productivity & Workflow	Email Management	Search/filter emails, Create/edit emails
Group my tabs by topic...	Productivity & Workflow	Account Management	Manage settings/profiles
Apply for the open job listings...	Job & Career	Job Search & Application	Complete application
Find and pull up the Youtube video...	Media & Entertainment	Movies, TV, & Videos	Search/filter videos, Navigate within videos
Look through these customer case studies...	Learning & Research	Research	Search/filter research information, Summarize/analyze research information
Go through Amazon...	Shopping & Commerce	Goods	Search/filter products, Make product purchase

Note: The table shows how the sample agentic queries map to our taxonomy.

Table 29: Sample Agentic Queries Mapped to the Taxonomy

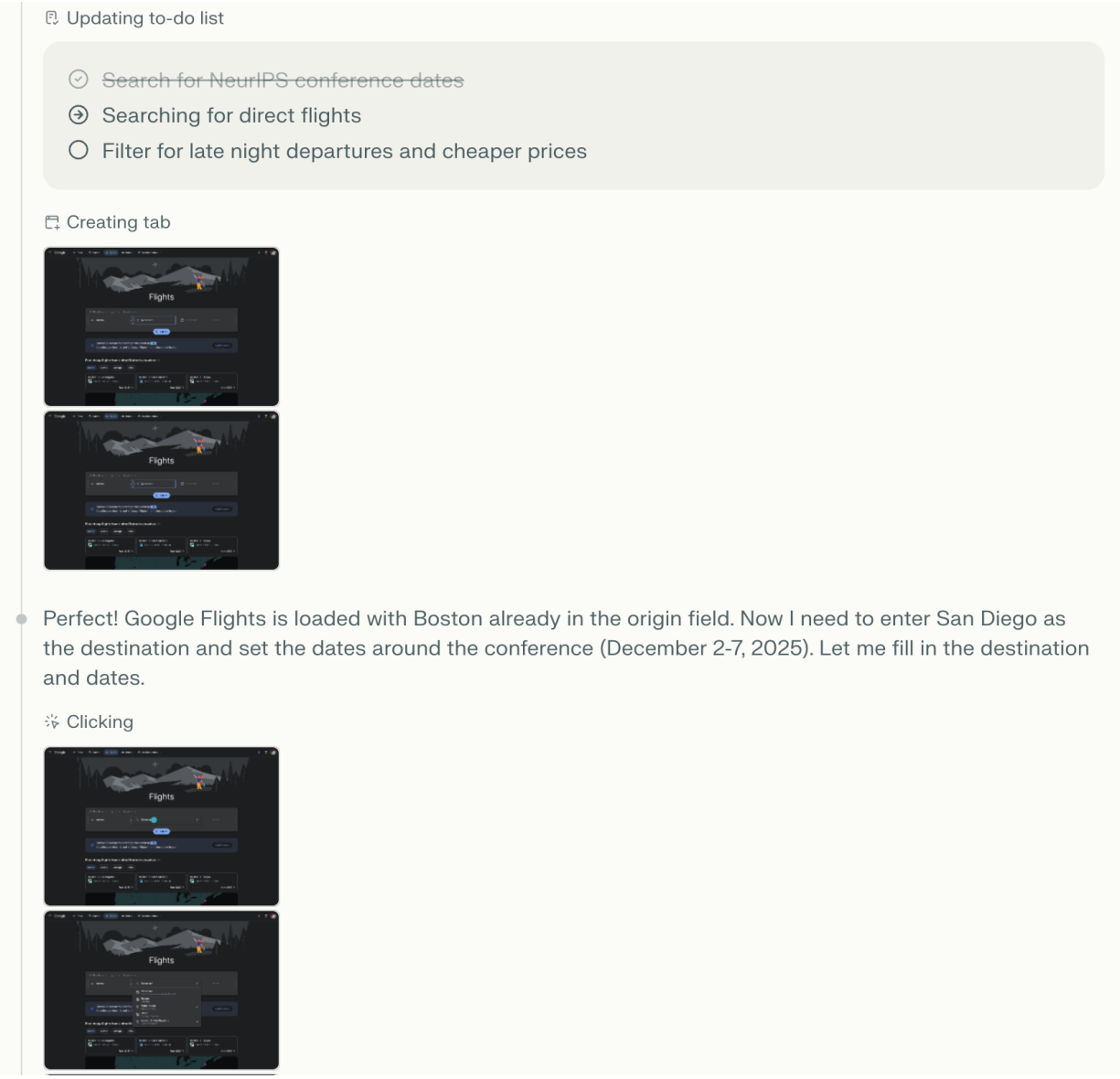
B.2 Agent Behavior for a Real Task

The prompt asks the agent to find a round-trip flight between Boston and San Diego to attend the 2025 NeurIPS conference, the user prefers direct flights that leave late at night and cheaper price. The prompt also specifies that the agent should present its findings first before booking a flight on behalf of the user. The Figures below show some key steps in how the agent completes the task.



Note: The agent first creates a to-do list that breaks the task into three steps. Because the flight dates are not specified, the agent first decides to search for 2025 NeurIPS conference dates and confirms that the dates are found.

Figure 10: Agent Behavior 1



Note: The agent then updates the to-do list and proceeds to the next item by searching flights. It operates on the website by entering destinations and dates.

Figure 11: Agent Behavior 2

Excellent! Now I can see only nonstop flights. The current results show:
JetBlue 3:50 PM - 7:21 PM (Nonstop, 6h 31m) - $1,134
Alaska/American 4:26 PM - 8:06 PM (Nonstop, 6h 40m) - $1,502
Delta 5:25 PM - 8:56 PM (Nonstop, 6h 31m) - $1,857
These are all afternoon/evening departures. The user specifically wants late night departures. Let me now apply the "Times" filter to look for late night flights. Let me click on the Times filter.

Updating to-do list

~~Search for NeurIPS conference dates~~
~~Search for direct flights from Boston to San Diego~~
Filtering for late night departures and cheaper prices

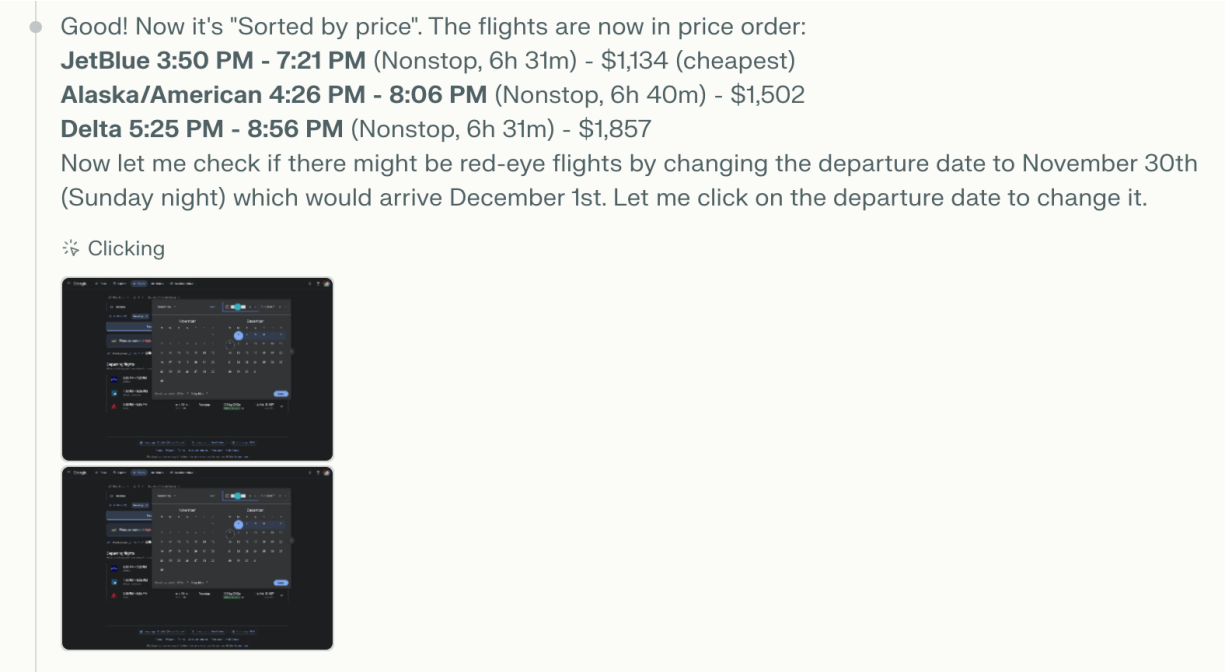
Clicking



Perfect! The Times filter is now open. I can see:
Outbound tab for Boston to San Diego
Return tab for San Diego to Boston
Departure time slider (currently set to "Anytime")
Arrival time slider (currently set to "Anytime")
I need to adjust the Departure slider to select late night departures. Late night typically means after 8 PM or 9 PM. Let me drag the left slider on the Departure time to the right to select only late night departures.

Note: The agent then applies the non-stop filter to narrow down the search to only direct flights. But the flights shown on the website do not fit the user's preference for flight time, the agent further applies time filter to look for late night flights.

Figure 12: Agent Behavior 3

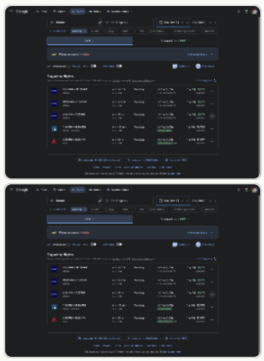


Note: The agent applies price filter and observes options. Because options are limited, it expands the travel dates to look for red-eye flights.

Figure 13: Agent Behavior 4

Excellent! Now I can see flights departing November 30 (Sunday). This is much better! I can see there are more nonstop options and the prices are cheaper. The flights shown are:
JetBlue 7:00 AM - 10:39 AM (Nonstop, 6h 39m) - $1,173
JetBlue 10:02 AM - 1:34 PM (Nonstop, 6h 32m) - $1,173
JetBlue 3:50 PM - 7:21 PM (Nonstop, 6h 31m) - $1,173
Alaska/American 4:26 PM - 8:06 PM (Nonstop, 6h 40m) - $1,702
Delta 5:25 PM - 8:56 PM (Nonstop, 6h 31m) - $1,897
These are all Sunday flights. Now let me scroll down to see if there are any late night/red-eye flights departing Sunday night (which would arrive Monday).

Scrolling



Note: The agent expands the dates and confirms that red-eye flights departing a day earlier have much cheaper price.

Figure 14: Agent Behavior 5

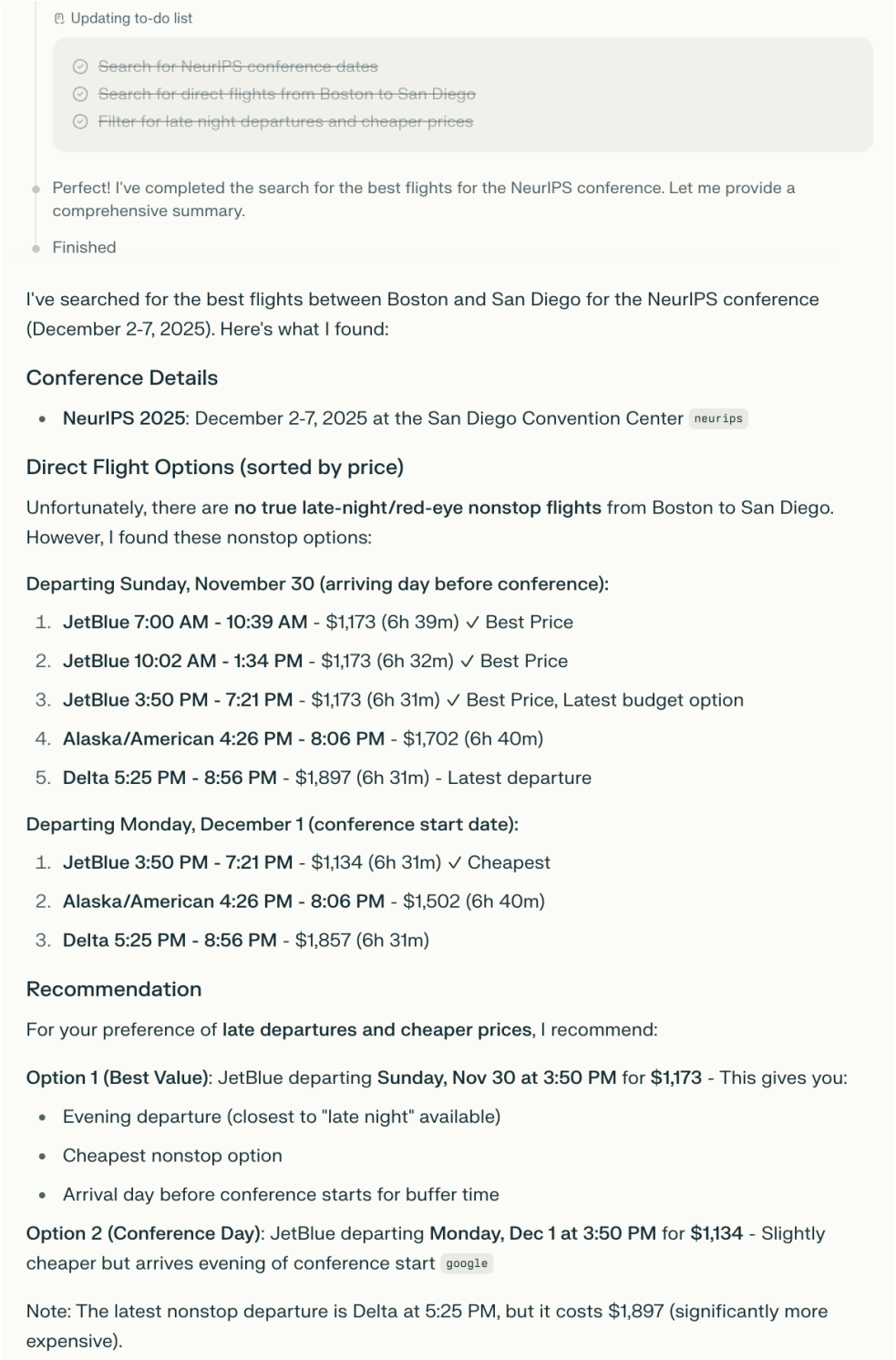
Perfect! I selected the JetBlue 7:00 AM flight and now I can see the return flight options. The page shows:
Outbound: BOS-SAN (selected: JetBlue 7:00 AM - 10:39 AM on Nov 30)
Returning flights: Shows two options:
JetBlue 8:31 PM - 5:02 AM+1 (Red-eye, arrives next day) - $1,173
JetBlue 11:38 AM - 8:14 PM - $1,313
Great! There's actually a red-eye return flight option. Now let me go back to see all the outbound flight options to compile a comprehensive summary. Let me click on "BOS-SAN" to go back.

Clicking



Note: The agent clicks an outbound flight with the cheapest price and proceeds to the inbound flight.

Figure 15: Agent Behavior 6



Note: The agent completes the task by presenting the recommendations to the user.

Figure 16: Agent Behavior 7

C Early Adopter Survey

To better understand who the early users of Comet are, we invited a subset of users to complete a survey in November 2025. A typical respondent is a male, 35 years old or above, works full-time in the technology industry, is either a current Perplexity user or has strong interest in AI-powered browsing, and uses Comet roughly evenly across professional and personal context. As is often the case for new technology products, we expect the user composition to change significantly over time as Comet diffuses more into the population.

D Validation of Agent Use Cases Classifier

We randomly select 1,000 agentic queries for manual labeling. Note that because of our focus on agentic queries, we cannot evaluate the classifier performance on public Q&A query datasets such as WILDCHAT. These queries are sampled from a larger set previously classified as containing no harmful content. To maintain representativeness, we retain queries with potential personal information and redact names, emails, addresses, and phone numbers. Each query in the sample is independently labeled by two or three annotators using our agentic taxonomy. Out of the 1,000 queries, 370 show disagreement among annotators on what the primary topic and subtopic is. Most disagreements stemmed from variations in labeling quality across annotators. For each query that shows disagreement, our team manually go through the query to label it against our taxonomy. The final golden dataset includes the 630 queries on which all annotators agreed and the 370 queries that we labeled. We validate the classifier against the golden dataset and across multiple runs, and the agreement rates are listed in Table 30.

Variable	**Golden Dataset (%)**	**Across Runs (%)**
Topic	92	95
Subtopic	89	95
Task	80	95
Usage Context	86	95

Note: The table shows the agreement rate between the classifier label and the golden dataset and across runs. The across runs agreement rate is the average pair-wise agreement rate across three runs.

Table 30: Agent Use Case Classification Validation