0. Start Here

ChatGPT Product Recommendations
Ian D'Silva | September 19, 2025

Context

This document is intended to chart the path I believe ChatGPT should take, providing food for thought and insight into how I think and work. I'd love to hear your thoughts at iandsilva96@gmail.com.

Contents

This Google Doc contains two primary documents:

1. **Product Strategy Doc:** A document that details a product strategy for ChatGPT, including:
    a. Where To Play Key user problems, why now, ChatGPT's right to win
    b. How to Win A high-level, two-part flywheel
    c. Initial Focus Areas: Two product categories to prioritize development focus towards
    d. How We Know We're Winning: A north star metric to guide focus and measure success
    e. Where to Start: The first product to build
2. **Morning Brief Product Brief:** A Product Brief for the first product recommended to build from the strategy doc, including:
    a. Problem Alignment: User pain point, key personas
    b. Solution Alignment: High-level approach, goals, non-goals
    c. Execution Alignment: Key features, future features, key flows, mockups, risks
    d. Launch Plan: Eligible users, ramp plan
    e. Measurement & Program Evaluation: Success metrics & KPIs, experimentation plan

Notes

- This is the first tab in this document. To view the other tabs, navigate at the left (web) or bottom (mobile).

1. Product Strategy Doc

ChatGPT Product Strategy
Ian D'Silva | September 19, 2025

Summary

- Where to Play: Navigating the fragmented digital world is taxing, but ChatGPT can be a simplifying super-assistant given progression in model intelligence.
- How to Win: We win through a two-part flywheel: (i) becoming the control center for users' lives by turning intents into outcomes and (ii) leveraging the high-intent users to create a developer ecosystem, culminating in ChatGPT as the digital world's orchestration layer.
- Initial Focus Areas: We recommend prioritizing a Personal Assistant Hub that helps manage users' days to solidify our role as a central, top-level orchestrator and a Developer Canvas, a widget intelligently surfaced in ChatGPT, to lay the foundation for turning intent into outcomes.
- Where To Start: Kickstarting our flyloop starts with Mission Control, so our initial product should be an MVP for the Personal Assistant Hub; we recommend building Morning Brief, a bespoke chat session surfaced every morning to help manage today's tasks, to build a daily habit of using ChatGPT as a personal assistant to reduce cognitive load.
- How We Know We're Winning: We will know if we are winning if using ChatGPT as a top-level digital orchestrator is a daily habit for users, measured by our platform's Daily Active Users.

1. Where to Play: Navigating the fragmented digital world is taxing, but ChatGPT can be a simplifying super-assistant given progression in model intelligence.

OpenAI's mission is to ensure that AGI benefits all of humanity. The internet is a messy, fragmented place and navigating it is taxing for users. Leveraging OpenAI's foundational technology, ChatGPT can simplify our interface with the internet by being a personal super-assistant that seamlessly turns intents into completed outcomes.

The Pain Point: *Completing a task (e.g., planning a trip) requires traversing a fragmented digital world, building cognitive toll.*

Accomplishing everyday tasks is burdensome:

- **The digital world is fragmented.** To get restaurant delivery you go to Doordash or Uber Eats. To book flights you go to Kayak or Delta. To learn about a topic, you go to Wikipedia, Reddit, Google, Substack, etc.
- **Navigating this is taxing.** To navigate the digital world, we pay a *cognitive toll* of friction-filled steps (tedious page clicking, switching back and forth from apps and pages, filling forms, authenticating, etc.).
- **A typical task requires significant navigating**: A single goal like "I need to plan a weekend trip to see my parents" requires a dozen discrete digital tasks (creating an itinerary, booking flights, scheduling transportation, last minute purchases, etc.) for users to complete.

The burden has fallen on users to connect the fragmented digital world. This friction drains our energy and gets in the way of getting stuff done.

See Appendix 1 - User Research for further insights.

Why Now? *Improvement in model intelligence can now be harnessed to navigate the digital world's complexity.*

Until now, the only solution to this was expensive human capital: a personal assistant who could understand intent and manage complexity on your behalf. This was a privilege reserved for the 0.01%.

Today's AI models possess the sophisticated reasoning and planning capabilities required to manage multi-step, complex tasks that were previously impossible for traditional software to handle. Model reliability can be an issue, but good agents can mitigate this for certain use cases while this improves.

ChatGPT can leverage these models to build the intelligent, connective layer that has been missing from the internet. Acting as a personal super-assistant that understands your goals and orchestrates digital tools to achieve them, ChatGPT can help users accomplish tasks without the cognitive toll.

See Appendix 2 - Model Capability Analysis for more background.

Why Us? *We have earned trust with users, have world-class product polish, and scale to attract a developer ecosystem; but, competitors can quickly close this gap if we don't move with urgency.*

We are very well positioned to solve this problem for users.

- First, we are currently in the lead with users. With 700m WAUs / 10x the web traffic of all other AI assistant providers *combined* and industry-leading retention, we have earned trust and permission to embed ourselves in the daily lives of users.
- Second, we have world class researchers and application teams who have proven ability to create industry leading AI experiences by deeply connecting model and product capability.
- Third, our distribution potential will attract developers, further compounding our product advantage.

But, this race has just started and competitors will bring compelling offerings to market. Google's deep vertical integration across the value chain (compute, data, models, applications) make it a formidable challenger. Perplexity's early product focus on this (Perplexity Assistant) can gain traction if uncontested. Apple, if / when they choose to participate, can provide a compelling offering. Further still, Claude, Meta AI, and Grok bring unique approaches to the consumer AI chatbot space.

See Appendix 3 - Market Positioning for more.

2. How to Win (The Master Plan): We win through a two-part flywheel: (i) becoming the control center for users' lives by turning intents into outcomes and (ii) leveraging the high-intent users to create a developer ecosystem, culminating in ChatGPT as the digital world's orchestration layer.

Our strategy is to be the intelligent orchestration layer for the digital world, providing a simple interface for users to achieve their goals. This requires native integration with existing digital services to maximize reach, impact, and reliability.

Our goal is a two-part flywheel: deeply embed with users by becoming their digital orchestrator and leverage that role to convert user intent seamlessly into outcomes with developer integration, further solidifying our central role as a digital orchestrator:

1. **Host "Mission Control" For Completing Tasks:** Position ChatGPT as the user's central console in the digital world — one place where intents turn into outcomes. Users state intent and ChatGPT creates a plan, orchestrates the right tools, remembers context and progress, and

explains/confirms actions. By hosting “Mission Control”, ChatGPT reduces friction, concentrates intent, and reliably solves users’ goals.

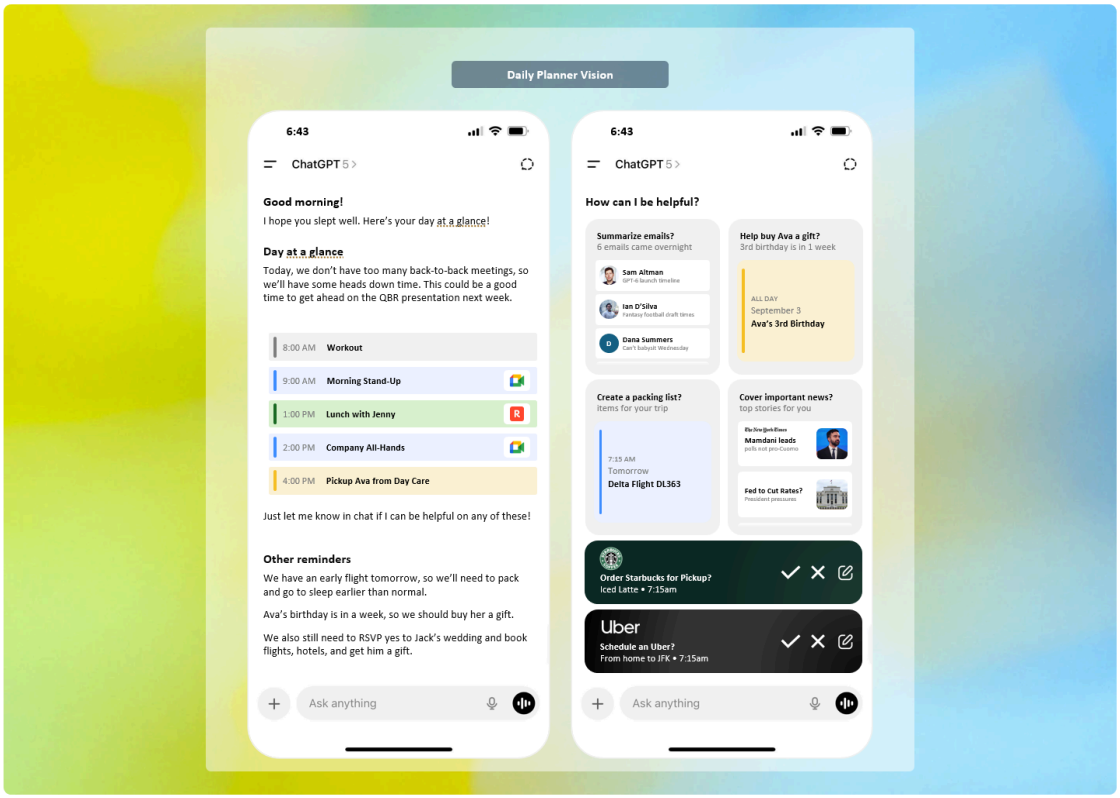
2. **Leverage Peak Intent to Build an Ecosystem:** By hosting “Mission Control”, we can surface developer integrations for full outcome resolution without leaving the app. This would give developers a premier distribution channel to high-intent users – integrations appear exactly when a user is trying to complete the task they solve, driving seamless conversion and incentivizing them to build.

As more developers plug in, users see broader coverage and better outcomes, reinforcing ChatGPT’s role as the orchestration layer and spinning a two-sided flywheel of user intent and developer growth.

3. Initial Focus Areas: We recommend prioritizing a Personal Assistant Hub that helps manage users’ days to solidify our role as a central, top-level orchestrator and a Developer Canvas, a widget intelligently surfaced in ChatGPT, to lay the foundation for turning intent into outcomes.

For the “Mission Control” focus area, we recommend building a **Personal Assistant Hub**, a persistent, top-level orchestrator that makes “Mission Control” tangible by scheduling, prioritizing, and executing tasks from one place. This establishes ChatGPT as the user's trusted, top-level orchestrator but defers complete end-to-end tasks until reliability is sufficient. Furthermore, it transforms ChatGPT from a transient tool to a persistent platform that users can offload cognitive burden to.

For the Developer Ecosystem focus area, we recommend building a **Developer Canvas**, an in-ChatGPT widget that is intelligently surfaced to assist with intent-completion. This product offers users a direct interface to the external digital world to complete tasks and reduce cognitive load, is scalable, and balances developer integration with quality control as we learn the right balance.



Together, these will position ChatGPT as a foundational platform for turning intent to outcomes by helping users navigate the digital world from one, integrated and simple interface.

See Appendix 4 - Top-level Ideas for more.

4. Where to Start: Kickstarting our flyloop starts with Mission Control, so our initial product should be an MVP for the Personal Assistant Hub; we recommend building Morning Brief, a bespoke chat session surfaced every morning to help manage today's tasks, to build a daily habit of using ChatGPT as a personal assistant to reduce cognitive load.

Kickstarting our flyloop starts with Mission Control, so our initial product should be an MVP for the Personal Assistant Hub.

For our initial product, we must build a lightweight MVP that establishes ChatGPT as a personal assistant, serving as the foundation on which to become a top-level digital orchestrator. To do this, we believe this is best accomplished with a product that offers proactive assistance and has a time-based (e.g., start of day) recurring trigger to drive habit-formation.

Accordingly, we recommend building a **Morning Brief**, a dedicated, auto-generated, and bespoke chat session surfaced every morning to centralize users' daily obligations and provide contextual chat suggestions to help with executing tasks, as the initial product of Personal Assistant Hub.

By launching this lightweight MVP, we can gather critical feedback to guide the development of the more robust Personal Assistant Hub.

PRD for Morning Brief available here.

See Appendix 5 - Personal Assistant Hub Ideas for more context.

5. How We Know We're Winning: We will know if we are winning if using ChatGPT as a top-level digital orchestrator is a daily habit for users, measured by our platform's Daily Active Users.

With an established user base of 700 million Weekly Active Users (WAUs), ChatGPT's primary focus should be on cultivating the highest form of habitual engagement and retention: **Daily Active Users (DAUs)**. Since there is still a significant opportunity to expand reach and engagement, monetization can be prioritized later, once a deeply engaged user base is solidified.

To prevent low-quality growth, a company-wide contribution profitability guardrail should be implemented. Team-specific metrics should be paired with their own quality guardrails to guide their local optimizations.

See Appendix 6 - North Star Metrics for more.

A1: User Research

ChatGPT Product Strategy POV: UXR
Ian D'Silva | September 19, 2025

Summary

Ian's user research shows that:

- **Background:** Ian conducted UXR across the following personas: a busy professional, a parent / household COO, a product manager, and a student, which can be used inform product development.
- **Pain Points:** Across the personas, the cognitive toll of daily life creates a burden and barrier from performing and being their best selves.
- **Opportunities:** This creates a significant opportunity to build a proactive, intelligent assistant that centralizes information, automates tasks, and understands context to offer a more holistic solution.

Users Research Takeaways

The Core Problem

Our users' core need is to automate the "work about the work" — the cognitive and logistical toil of daily life — so they can reclaim their time and energy for what truly matters.

- In our **personal and social lives**, it creates decision fatigue and stress around planning, discovery, and connection (e.g., having to constantly remember things, finding the right gift, organizing a group trip).
- In our **family lives**, it becomes a relentless, 24/7 mental load that directly erodes personal well-being and leaves no space for self-care.
- In our **professional lives**, it is a barrier to high-impact work, consuming time in coordination and manual synthesis that should be spent on strategic and creative thinking.

This creates an opportunity to build an intelligent assistant that moves beyond simple task management to offer a holistic solution.

- **Centralize & Synthesize:** Creating a single hub that captures and organizes the user's fractured information — tasks, notes, plans, and context.
- **Automate & Generate:** Automating routine tasks like scheduling and reminders but also more complex needs like discovery and recommendations (e.g., gifts, travel, activities).
- **Be Proactive & Context-Aware:** The ultimate goal is an assistant that anticipates needs. By understanding the user's context, it can proactively offer help, streamline coordination, and ultimately help unlock their potential.

See next page for further analysis.

User Research

Source Document: *JTBDs and Pain Points*

The Takeaways

THE BUSY PROFESSIONAL

Ian

"My job is really busy right now. I can't find time to run errands."

JOB TO BE DONE

For needs to seamlessly manage the logistics and mental load of his personal life and social obligations, so he can be fully present for important events without feeling overwhelmed by the planning and coordination required.

KEY PAIN POINTS

- Time-consuming research for travel, gifts, etc.
- Forgetting key dates, tasks, and conflicts.
- Decision fatigue when making plans or purchases.
- Inefficient group coordination.
- Lack of proactive reminders and context-aware assistance.

KEY OPPORTUNITIES

- Automate scheduling, research, and purchases.
- Provide personalized recommendations for gifts, travel, and activities.
- Offer proactive reminders and communication aids.
- Simplify group planning with streamlined coordination tools.

PARENT & HOUSEHOLD COO

Sharon

"Taking care of my baby, working a full-time job, and keeping the house together is exhausting."

JOB TO BE DONE

For Sharon needs to offload the mental and logistical burden of managing her household and family, so she can feel more organized, reduce decision fatigue, and reclaim time and energy for herself.

KEY PAIN POINTS

- Overwhelming mental load causes exhaustion and forgetfulness.
- Scattered information with no central place for managing tasks.
- Decision fatigue from countless daily choices.
- No time or mental space for personal needs or self-care.

KEY OPPORTUNITIES

- Centralize & consolidate all lists/notes/tasks.
- Proactively assist (e.g., suggestions, reminders, answering questions).
- Optimizing schedules based on constraints.
- Recommend personalized self-care activities.
- Automate routine tasks like shopping restocks and family updates.

PRODUCT MANAGER

Melanie

"I spend more time coordinating and stitching context together than shaping the product."

JOB TO BE DONE

For Melanie needs to translate fragmented ideas, data, and conversations into a cohesive product strategy, by automating the manual work of information synthesis and coordination so she can focus on high-impact, creative thinking.

KEY PAIN POINTS

- Scattered information and lost context across multiple tools.
- Excessive time spent on the tedious tasks.
- Difficulty connecting disparate insights into a cohesive strategy.
- Inefficient coordination & admin. overhead.
- Lacking a clear, objective summary of weekly progress and decisions.

KEY OPPORTUNITIES

- Centralize all notes, decisions, and tasks in a single place.
- Synthesize information automatically.
- Automate the first draft of deliverables.
- Streamline coordination by auto-logging decisions and action items.
- Provide automated daily briefings and weekly progress summaries.

STUDENT

Mark

"Between classes, labs, and social events, it's easy to fall behind."

JOB TO BE DONE

For Mark, the job is to integrate the fragmented demands of his academic and social life, so he can manage his time effectively and stay ahead of his work without feeling constantly overwhelmed.

KEY PAIN POINTS

- Difficulty prioritizing a complex academic and social schedule, leading to procrastination.
- Information overload from scattered course materials, notes, and group chats.
- Constant context-switching between different subjects erodes focus and efficiency.
- Pervasive anxiety and overwhelm from the feeling of constantly being behind.

KEY OPPORTUNITIES

- Proactively build an intelligent, adaptive schedule, integrating deadlines and social plans.
- Synthesize course materials into actionable study aids like summaries and quizzes.
- Streamline group collaboration by summarizing conversations and identifying action items.
- Provide smart coaching and automation to reduce procrastination and build better habits.

Ian – The Busy Professional

Ian

THE BUSY PROFESSIONAL

"My job is really busy right now. I can't find time to run errands."

JOURNEY SCENARIO

Ian is invited to a friend's wedding in SF and needs to plan his attendance.

LOCATION
New York & SF

AGE
29 Years Old

CHATGPT USAGE
Daily

PHASE	Receiving Invitation	Booking Flights	Booking Lodging	Buying a gift and outfit	Making Plans	Packing for the Trip	Returning Home
STORY	Ian checks his mail and opens a wedding invitation from his friend. Excited, Ian checks his calendar. He often forgets this step. Oh no! It's his niece's birthday that weekend. He needs to let his sister know he can't make it and figure out an alternative way to celebrate. He also needs to RSVP and save the date.	The reception is Saturday with welcome drinks on Friday night. Ian needs to find a flight – he can't miss the big leadership meeting Friday at 2pm, so he needs to fly after that and he doesn't want to get back home too late on Sunday. He spends 30 minutes looking for flights trying to balance price, time, and loyalty perks.	He looks at the wedding website and finds the wedding hotel block. He remembers many of his b-school friends are going and a group airbnb would be fun! He pitches the idea – everyone agrees, but no one seems to want to spend the time planning and coordinating. He ends up doing it.	Ian's bringing a +1 this time, which he hasn't done before. Do you give double the gift then? **He asks ChatGPT.** He picks a gift and drafts a note, asking **ChatGPT to refine it in his voice.** He edits it and sends the gift. He sees there's an Indian reception. He needs a kurta. **He asks ChatGPT for advice** and looks on Amazon for an appropriate item.	Ian checks which of his friends, cousins, and former colleagues are still in SF to try to make plans. He reaches out to see who is around. **He uses ChatGPT to figure out plans and where to go.** He decides on coffee with an ex-coworker, drinks with a friend, and a hike with his cousins.	Ian uses **ChatGPT to help come up with a packing list.** He reminds it that he needs to have stuff for the wedding and his other plans. He missed that it's going to be raining this weekend, so he forgot an umbrella for the outdoor ceremony.	Ian had a blast at the wedding and is returning home. Uh oh, Ian almost forgot to get his niece a gift! **He looks up great baby gifts on ChatGPT to buy on Amazon.** He looks it up and it's out of stock. He goes with option #2 and sends it to his sister's house.
PAIN POINTS	• *Forgets existing calendar conflicts (e.g. niece's birthday)* • *No smart detection of schedule clashes*	• *Manual research to balance time, price, perks* • *No help synthesizing constraints*	• *Group coordination is clunky and time-consuming* • *No shared planning or delegation tools*	• *Uncertain what to buy, lacking personal guidance* • *Writing gift note feels like extra work*	• *Forgets to contact key people* • *Needs help narrowing options for limited time*	• *No context-aware reminders on the wedding* • *Had to remind ChatGPT of full itinerary*	• *No proactive reminders on niece's gift* • *Recommendation was out of stock*
JOBS TO BE DONE	Remember & Remind Prioritize & Plan	Answer & Advise Prioritize & Plan Execute & Automate	Coordinate & Mediate Prioritize & Plan Execute & Automate	Create & Synthesize Answer & Advise Execute & Automate	Remember & Remind Remember & Remind Coordinate & Mediate	Remember & Remind Execute & Automate	Remember & Remind Execute & Automate
OPPORTUNITIES	• **Smart detection of calendar conflicts (e.g. birthdays, work events)** • **One-click RSVP and calendar integration** • **Help with schedule coordination**	• **Summarized flight options and recommendations based on constraints (price, loyalty, timing)** • **One-click purchase automation**	• **Summarized Airbnb options and recommendations based on group constraints** • **One-click purchase automation**	• **Personalized gift suggestions based on recipient and event** • **Auto-drafted gift messages with tone refinement** • **Outfit planning suggestions based on event type**	• **Context-aware suggestions on who to connect with, what to do, and where to go.**	• **Context-aware (e.g., weather aware, itinerary understood) recommendations**	• **Proactive reminders for missed tasks (e.g. baby gift)** • **Native integration with retailers for better suggestions**

Sharon – The Parent and Household COO

Sharon

PARENT & HOUSEHOLD COO

"Taking care of my baby, working a full-time job, and keeping the house together is exhausting."

JOURNEY SCENARIO

It's Saturday – Sharon is balancing caring for her 1-year-old, running errands, and carving out time for herself.

LOCATION
Long Island, NY

AGE
33 Years Old

CHATGPT USAGE
Weekly to Monthly

PHASE	Waking Up	Morning Errands	Pediatrician	Mid-Day (Nap Time)	Afternoon Personal Time	Dinner & Evening Routine	Wind Down (Late Evening)
STORY	Sharon wakes early to feed and change her baby. She logs the sleep and feed details while rushing to grab a coffee. The baby didn't sleep well and isn't too interested in breakfast today. She looks at her calendar and to-do list. It's long. The day feels hectic before it's even begun.	Since it's Saturday, it's time to run errands. There was not time during the week with work. Her partner watches the baby. She heads to the grocery store and pharmacy, picking up food and medication. She orders the remaining items she needs on Amazon. The list? In her head. There was not time to write it down. Shoot, she forgot eggs.	Next, she's off to the pediatrician – thank goodness her friend recommended a doctor with weekend appointments. She books the next appointment and adds it to her calendar. The visits ends and Sharon heads home, trying to remember everything the doctor recommends while keeping her baby content on the go.	They get home and it's nap time. During the baby's nap, Sharon squeezes in laundry and tidying before diving into daycare research. **She tries ChatGPT** at her brother's insistence. It's quite good! But, it doesn't have all the context on requirements: close to home / work, allergy friendly, etc. She needs to prompt that.	The baby wakes up from her nap and her partner is coming home from running his share of errands. Sharon passes baby duties to him and goes for a run. Teamwork makes the dream work! She opens up the Peloton app for a guided run. Overwhelmed by all the choices she spends 15 minutes looking at them, just to pick a random one.	It's now time for dinner and the nightly decision: cook or takeout. They're feeling uninspired so they **look up restaurants and recipes on Google**. They input the ingredients they have and ChatGPT recommends an awesome stir fry. They feed and bathe the baby, make their dinner.	With the baby down, it's time to tidy up and prep for tomorrow: dishes, cleaning, and bottle prep. She facetimes the grandparents, racking her brain for updates and the doctor's advice and her thoughts on daycares. They finish catching up and she goes to sleep. She's exhausted, but survived another day. And she'll do it all again tomorrow!
PAIN POINTS	• *Starting the day already tired from the burdens of the previous day/week* • *Immediately juggling feeding and logging.*	• *Carries a mental load of remembering lists, which is tiring and can lead to mistakes* • *Each place is its own process. There's no one central place to manage everything.*	• *Pediatrician advice is another mental note to keep*	• *Very limited time window – she needs to carefully pick what she starts* • *Research feels incomplete without all context, so she needs to spend another day confirming*	• *The mental load of deciding and remembering things is catching up.* • *She's frustrated the app doesn't just know which one she should do*	• *Constant decision fatigue ("cook or takeout?") – can't someone help her decide?*	• *Still carrying cognitive load (grandparent updates, scheduling, doctor's notes).* • *Exhaustion leaves little space for self-care.*
JOBS TO BE DONE	Remember & Remind Prioritize & Plan	Prioritize & Plan Remember & Remind Execute & Automate	Remember & Remind	Prioritize & Plan Answer & Advise Coordinate & Mediate	Coordinate & Mediate Answer & Advise	Remember & Remind Execute & Automate	Remember & Remind Execute & Automate Coach
OPPORTUNITIES	• **Auto-log baby's feeding/sleep data from smart devices.** • **Morning digest: surface today's priorities**	• **List consolidation (combine grocery, pharmacy, Amazon into one checklist).** • **Auto-detect forgotten items** • **Route optimization across errands with real-time updates.**	• **Real-time transcription + summary of doctor's advice.** • **Pediatric Q&A companion and personalizer**	• **Chore planner: suggest which tasks fit the nap window.** • **Personalized research assistant** • **Auto-schedule daycare tours that fit her availability.**	• **Recommend the best workout based on context (time, energy, baby's mood).** • **Sync schedules and surface next best step.**	• **Smart meal suggestions based on what's in the fridge + dietary needs.** • **One-click grocery restock for missing ingredients.**	• **Summarize the day's key takeaways (doctor, daycare, errands).** • **Auto-generate family update texts.** • **Proactively create tom. priority list so she can unplug.**

Melanie – The Product Manager

Melanie

PRODUCT MANAGER

"I spend more time coordinating and stitching context together than shaping the product."

JOURNEY SCENARIO

It's Monday. Melanie must craft a big product idea and present it to her team on Friday.

LOCATION
Sydney, Australia

AGE
38 Years Old

CHATGPT USAGE
Daily to Weekly

PHASE	Commute & Daily Plan	Morning Stand Up	Data Deep Dive	Market Research	Prototyping	PRD & Deck Assembly	Reflecting on the Week
STORY	On her drive, Melanie mentally plans the week. Her to-dos and obligations live across apps. She can't use her phone or ChatGPT while driving, so ideas and to-dos go uncaptured until she's at her desk.	In standup, people discuss priorities and blockers. Melanie works with them to schedule reviews and workshops. They go back and forth on times. She types notes while talking. Trying to lead the meeting & taking notes is hard! If only ChatGPT integrated. She tries to remember her priorities from the car ride. She almost forgets to mention a key priority!	To size the opportunity, she thinks through what matters. There are so many different cuts of data that would be interesting to view. She uses **ChatGPT Enterprise** to help write queries but she still needs to manually load them into the dashboard. She wishes her team could get HC for a product analytics associate to help.	Melanie reviews customer interviews, competitor offerings, and reviews pricing notes. **ChatGPT Enterprise** helps her research and synthesize the information. But she still needs to do some work to put everything into a cohesive narrative.	The opportunity is large and customers want this! There's a real opportunity here. She wants to create a small prototype to illustrate the vision based on all her quantitative and qualitative work. She uses Cursor to generate an illustrative prototype. She iteratively prompts it until it expresses her vision. It actually takes quite a while to get it just right.	It's now time to put it all together. With **ChatGPT Enterprise**, she drafts faster. But, she still has to copy/paste Figma links, charts, and research. And, she still needs to fully synthesize and create a narrative for the disparate parts.	She delivers the presentation and it goes well! Execs love the idea. Melanie is growth-oriented, so she celebrates the win but also reflects on what she could've done better. The week was such a blur, though, so she can only reflect so much.
PAIN POINTS	• *To-dos are across calendars, workflow apps, and her brain.* • *It's hard to use ChatGPT in the car.* • *Wrangling everything means, she can't proactively plan the ideal week.*	• *Remembering key points can be difficult.* • *Decisions / owners / dates get lost.* • *Tedious work to create prep points and recaps.*	• *Hand-writing SQL is cumbersome and repetitive* • *She needs to spend time thinking through what matters – a thought partner would be super helpful.*	• *Insights are not mapped to data analysis and strategic frameworks.*	• *She is having a trouble getting cursor to build the front-end design. She has to keep tweaking the prompt until it's right. She wishes someone else could do this on her behalf.*	• *Insights still live in silos and she has to think deeply to connect the dots*	• *She can't remember everything* • *She's tired after the long week* • *She doesn't have an objective view of everything that happened*
JOBS TO BE DONE	Remember & Remind Prioritize & Plan	Coordinate & Mediate Execute & Automate Remember & Remind	Answer & Advise Execute & Automate	Create & Synthesize Answer & Advise Execute & Automate	Create & Synthesize Execute & Automate	Create & Synthesize Execute & Automate	Remember & Remind Coach
OPPORTUNITIES	• **A morning brief and agenda setting: top 3 priorities, obligations, risks, time blocks, etc.** • **CarPlay voice daily planner with offline reliability.**	• **Google Meet integration with project memory; auto decisions / owners / due dates posted to Slack / email.** • **Searchable decision log with rationale.**	• **Reasoning assistant to help scope and design how to tackle the problem.** • **An integrated generator that understands the database to generate queries.**	• **Auto-transcribe customer calls.** • **Improved synthesis capability.**	• **Agentic iterator until vision is achieved.**	• **Improved synthesis capability.**	• **Progress summaries from the week.** • **Moments of feedback from meetings and documents.** • **Priority growth areas for next week.**

Mark – The Student

Mark

STUDENT

"Between classes, labs, and social events, it's easy to fall behind.

JOURNEY SCENARIO

It's midterms week. Mark must prep for an Econ exam, finish a CS lab, and make a group presentation by Friday, while juggling his social network.

LOCATION
Cambridge, MA

AGE
20 Years Old

CHATGPT USAGE
Weekly User

PHASE	Morning Download	Late Morning Dread	Econ Prep	CS Lab	Group Presentation Text	Dinner	Back at the Dorm
STORY	Mark wakes up to calendar alerts. He opens ChatGPT and prompts: "Summarize my week from my connected calendar." It lists the Econ exam, CS lab, and group presentation, all due Friday. Seeing the organized list just confirms how much work he has.	He tries to make a plan. "Create a study schedule for me," he types. The AI gives a generic template. He tries to refine the prompt but isn't sure how, gets frustrated by the vague output, and starts scrolling TikTok instead.	He snaps a picture of his Econ syllabus. "Explain the key concepts for this exam." The AI gives a solid list, but he's still not absorbing it. He asks it to "explain 'elasticity' like I'm five," which helps. One concept down, many more to go. Why didn't he review the topics each week like he was supposed to…	Ditching Econ, he moves to the CS lab. He asks ChatGPT to generate boilerplate code, then spends the next hour asking it to help him customize his code editor theme—a perfect form of productive procrastination.	The group chat for the presentation is a mess of memes and vague plans. Mark screenshots it and asks ChatGPT to "find the action items and suggest meeting times." He pastes the clean summary back to the group.	A friend texts "Dinner?" Annoyed with his lack of progress, Mark, looking for a distraction says yes. The break is a relief but completely kills his momentum for the night. They spend 30 minutes just deciding where they should go, wasting time Mark can't really afford to lose – oh well.	Back in his dorm, Mark has an idea. He gives ChatGPT his syllabus, the project specs, and his calendar constraints. "Act as an expert academic coach and create a detailed, hour-by-hour plan for Tues-Fri." The actionable, specific output finally makes him feel in control.
PAIN POINTS	• *Seeing all tasks listed at once increases anxiety.* • *Still feels passive, like a recipient of bad news.*	• *Unsure how to write effective prompts to get desired results.* • *Generic AI output is not actionable.*	• *AI can explain concepts, but can't force retention.* • *Information overload, even when summarized.*	• *Easy to use AI for low-impact, adjacent tasks.* • *Avoiding the core, difficult part of the work.*	• *Digital tools don't solve human communication problems.*	• *Decision paralysis / slow decision making for simple decisions*	• *N/A*
JOBS TO BE DONE	Remember & Remind Create & Synthesize	Prioritize & Plan Create & Synthesize Coach	Answer & Advice Create & Synthesize	Create & Synthesize Execute & Automate	Coordinate & Mediate Create & Synthesize Prioritize & Plan	Answer & Advice Prioritize & Plan	Coach Prioritize & Plan Create & Synthesize
OPPORTUNITIES	• **Proactive weekly summary with suggested starting points.** • **Frame tasks with motivational language.**	• **Prompt suggestions and templates for goals like "build a study plan."** • **Interactive planning that asks clarifying questions.**	• **Auto-generate flashcards or quizzes from uploaded materials.** • **Link concepts to real-world, engaging examples.**	• **AI-powered "first step" suggestions for complex logic.** • **Detect and flag "procrastination" patterns in usage.**	• **Group chat integration to auto-suggest meeting polls.** • **Draft messages to help focus the group on decisions.**	• **Recommendations with context-awareness**	• **Save as template" feature for highly effective prompts.** • **Turn a successful plan into a repeatable routine.**

Core Jobs to Be Done

1. **Remember & Remind** — externalize memory and surface it at the right time.
    a. Examples: “Remember my KTN,” “Remind me on the 1st business day each month.”
2. **Answer & Advise** — provide clarity, comparisons, and rationale to speed decisions.
    a. Examples: “Which headset is best for noisy offices?” “Explain QLED vs. OLED for daylight use.”
3. **Create & Synthesize** — ideate, draft, edit, summarize, translate, and reformat artifacts.
    a. Examples: “Draft the vendor email,” “Turn these notes into a 1-pager,” “Brainstorm 10 taglines.”
4. **Prioritize & Plan** — choose what to do and sequence it against goals and constraints.
    a. Examples: “Plan my week,” “Build a 3-day Tokyo itinerary,” “Create an interview prep plan.”
5. **Coordinate & Mediate** — reach alignment across people, orgs, or systems.
    a. Examples: “Find a time for Alice/Bob/me,” “Propose an agenda both teams can accept.”
6. **Execute & Automate** — perform actions in external systems so work gets done.
    a. Examples: “Book JFK→CDG under $400,” “File the expense,” “Send calendar invites.”
7. **Coach** — help the human build skills, habits, and outcomes.
    a. Examples: “Hold me to a daily writing habit,” “Quiz me on SQL,” “Give feedback on my answer.”

A2: Model Capability

ChatGPT Product Strategy POV: Model Capability Analysis

Ian D'Silva | September 19, 2025

Summary

- **Background:** Product use case development is extremely sensitive to model capability. This document helps gain a high-level understanding of model general-purpose capability to inform where to focus.
- **Models Are On Precipice of Solving Many Real-World Problems, AI Agents Can Close The Gap:** AI capabilities are advancing exponentially, reliability remains the primary bottleneck for complex tasks. Research suggests AI agents can meaningfully improve performance by deconstructing broad problems into granular, well-defined steps and improving context.

Insights

- **Models are improving rapidly, but reliability is the key bottleneck:** AI capabilities are advancing exponentially. GPT-5 can solve complex tasks humans take ~2 hours on, with 50% success. For an 80% success rate, the task duration drops to ~23 minutes, highlighting reliability as the main hurdle.
- **Performance is highest on granular tasks, and providing context is crucial.** Models perform better on constrained and granular tasks (e.g., interpreting results) but struggle with broad challenges (e.g., creating or reproducing a full paper). Providing context and background boosts performance, indicating domain knowledge is crucial.
- **The agent layer is a critical factor in performance.** Agent design significantly impacts performance, as demonstrated by one team improving GAIA benchmark completion from 50% to 83% by iterating on agent design, not the underlying GPT-4.1 model.
- **Web navigation remains a major challenge, highlighting the need for better tool integration.** Models struggle with web tasks like flight booking (40-65% success) due to incorrect tool use, domain misunderstanding, navigation errors, and incomplete steps. Computer-empowered agents like Operator improve performance but not enough, indicating developer-led tool integration may be necessary to achieve real-world reliability demands.

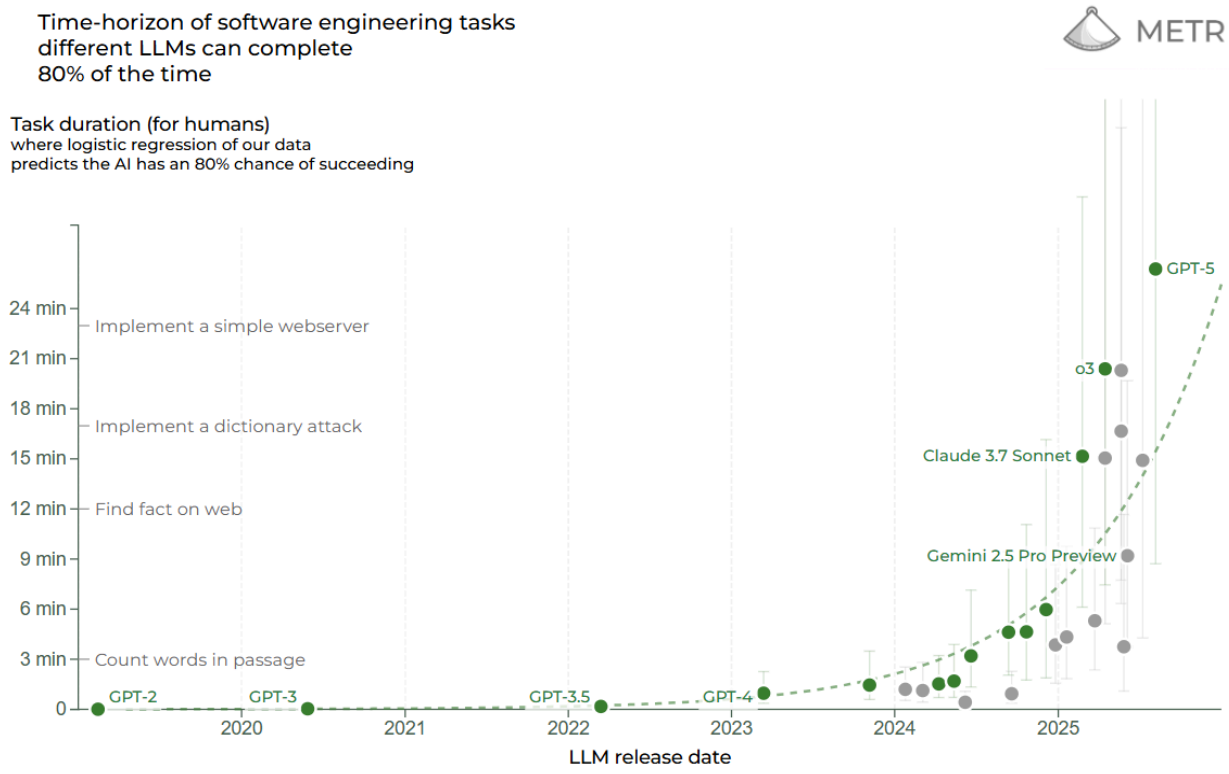
See next page for further analysis.

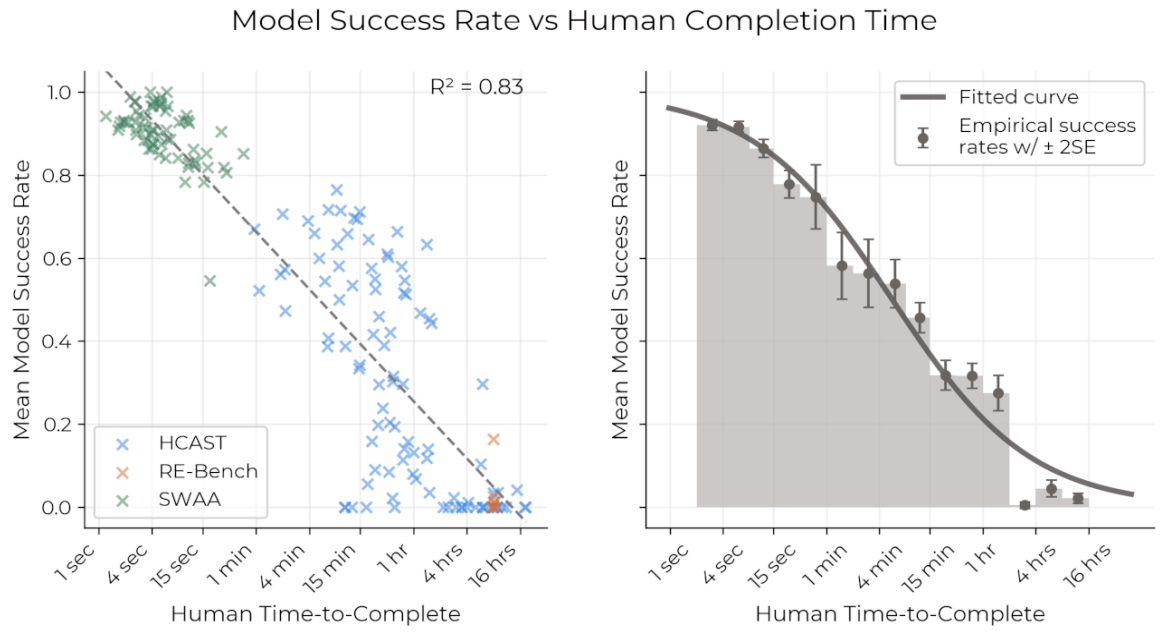
Eval Review

Measuring AI Ability to Complete Long Tasks (Link)

>> Models are increasingly solving longer tasks, but reliability still is the key bottleneck

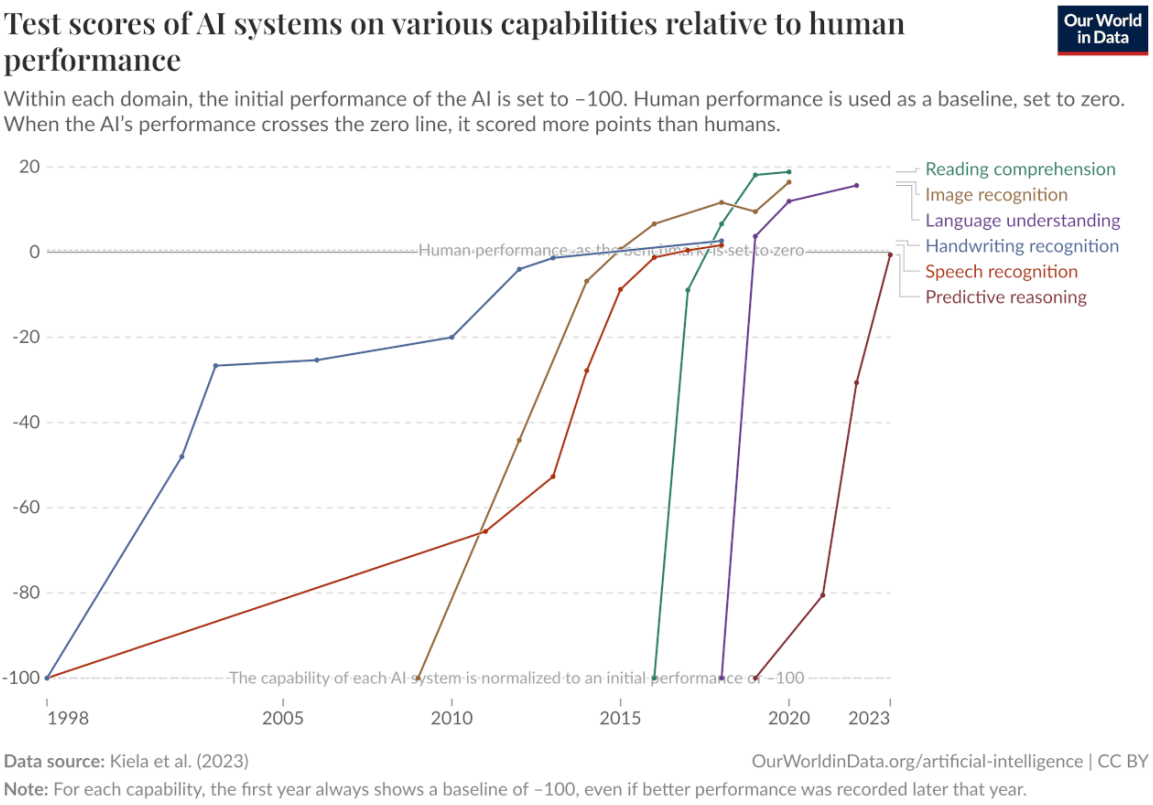
- **Model Reliability Bottleneck:** Models are increasingly able to solve tasks, but reliability remains the largest gap.
    - GPT-5 can perform a ~2 hour task autonomously with a 50% pass rate but ~23 minutes with an 80% pass rate.
    - Models can solve coding tasks that would take a human ~5-10 seconds with ~90%+ reliability
- **Progress is Exploding:** Progress has been improving at an exponential rate with massive jumps in 2025 over 2024.





>> Models are increasingly solving longer tasks, but reliability still is the key bottleneck

- **Parity with Humans:** Models can now perform at the same level as humans in many key knowledge tasks.



CORE Bench – Reproducing Scientific Papers (Link)

Eval Background: CORE-Bench evaluates the ability of agents to computationally reproduce the results of published scientific papers.

>> Models are not able to reproduce full scientific papers, but have more success on more granular tasks like interpreting results

- **Models Can Only Solve Granular Tasks:** The top models are only able to solve 40-50% of tasks. Tasks range from interpreting results (Easy) to full paper reproduction (Hard), suggesting more granular tasks like interpreting results are only solvable by models.
- **Reliability Is Still an Issue:** GPT-4o could accomplish 60% of easy tasks, 60% of medium tasks and 20% of hard tasks (fig. 6 of paper). While the easy tasks should be more performant, there was no improvement in the solve rate.
    - This may partially due to cost constraints, as pass rate went up for 4o (but not 4o-mini) with a higher cost cap (fig. 7 of paper).

SciCode – Generating Code for Scientific Papers (Link)

Eval Background: SciCode evaluates AI agents' ability to generate code for realistic scientific research tasks. It is made up of 65 main problems decomposed into 338 subproblems across 16 subfields in six natural science domains (Mathematics, Physics, Chemistry, Biology, Material Science, and Computational Mechanics).

- Level 1 questions generally require no tools, or at most one tool but no more than 5 steps.
- Level 2 question generally involve more steps, roughly between 5 and 10 and combining different tools is needed.
- Level 3 are questions for a near perfect general assistant, requiring to take arbitrarily long sequences of actions, use any number of tools, and access to the world in general.

>> Models cannot solve highly complex coding tasks for research problems, though hand-feeding important context improves performance

- **Models Fall Short on Sub- and Main Problems:** For the easier subproblem-level evaluation, the state-of-the-art models we test solve 14-26% of the subproblems. However, all models perform much worse on the more realistic and challenging main problem evaluation with the best model (Claude) solving 4.6% of total main problems
- **Context Improves Performance:** All models substantially improve performance (c. +10pp uplift) for both subproblem and main problem evaluations when given background like text authored by scientists and generated solutions to previous subproblems
    - However, mid-20-percent performance is well below satisfactory performance

GAIA – General AI Assistants Benchmark (Link)

Eval Background: GAIA is a benchmark for General AI Assistants that requires a set of fundamental abilities such as reasoning, multi-modality handling, web browsing, and tool-use proficiency. It contains 450 questions with unambiguous answers, requiring different levels of tooling and autonomy to solve and has three levels.

>> Model performance is improving rapidly

- **Models Show Reliability:** The best agents are now able to reliably solve Level 1 problems with 93%+ accuracy and Level 2 with 80%+ accuracy, demonstrating meaningful real world usability. Level 3 problems are at 60%+ pass scores, suggesting reliably solving could be near.
- **Agent Layer Matters:** Agents using the same model can generate outperformance and agents using multiple models may perform better.
    - Leveraging the same model, GPT-4.1, one research group's agent improved from 50% completion (Agent v0.0.2) to 83% completion (Agent v0.1.4)
    - Multi-model agents seem to perform well but don't necessarily outperform single-model agents (primarily GPT-4.1)
-

TauBench – Real World Use Cases with Tool-Agent-User Interaction (HAL Link, Paper, Retail)

Eval Background: TAU-bench is a benchmark for Tool-Agent-User Interaction in Real-World Domains. TAU-bench Airline and Retail evaluates AI agents on tasks in the respective domains.

>> Models still cannot solve a (seemingly basic) task like booking a flight reliability but are closer to making simpler purchases like retail

- **Passable Retail Performance:** Top models can make retail purchases with just ~82% accuracy.
- **Sub-par Airline Performance:** Top models can book flights with just ~60-65% accuracy.
- **Retail Outperforms:** Removing the domain policy from the prompt hurts performance much more in the τ-airline domain than in τ-retail, indicating that τ-retail tasks rely more on commonsense and less on complex rules
- **Variety of Failure Modes:** Models messed up the tasks by (i) passing an argument incorrectly to a tool call, (ii) failure to understand the domainspecific knowledge or rules and making the wrong type of tool call, (iii) failure to comply to all user requests.

Mind2Web – Generalist Agents Navigating the Web (HAL Link)

Eval Background: Mind2Web is a dataset for developing and evaluating generalist agents for the web that can follow language instructions to complete complex tasks on any website. Online Mind2Web is the live, online version of Mind2Web. It does not rely on cached pages and allows for real-time testing against dynamic, evolving web interfaces.

>> Models still cannot solve a (seemingly basic) task like booking a flight reliability

- **Sub-par Performance:** Top models can book flights with just ~40% accuracy. This improves with Claude Computer Use 3.7 and OpenAI Operator with 56% and 61% success rates, respectively.

- **Variety of Failure Modes:** Models failed for the following reasons, including filter and sorting errors, incomplete steps (e.g., not submitting a form or opening detailed pages), navigation errors, and misunderstanding the task's main goal.
- **Operator Excels:** Operator outperforms other agents by using filters and structured search more effectively than broad keyword queries. It has a flexible action space with various tools, including "Ctrl+F" for web navigation. A key feature is its self-verification and self-correction, allowing it to recheck task requirements and independently rectify missing or incorrect filters.
- **Limitations of Operator:** Operator has two main limitations: it often fails to meet numerical and temporal constraints and despite its exploratory nature, it occasionally misses niche website features needed for tasks.

Mind2Web – Generalist Agents Navigating the Web (HAL Link)

Eval Background: Mind2Web is a dataset for developing and evaluating generalist agents for the web that can follow language instructions to complete complex tasks on any website. Online Mind2Web is the live, online version of Mind2Web. It does not rely on cached pages and allows for real-time testing against dynamic, evolving web interfaces.

>> Models don't current have great results interfacing with the web.

- **Sub-par Performance:** Top models can book flights with just ~40% accuracy. This improves with Claude Computer Use 3.7 and OpenAI Operator with 56% and 61% success rates, respectively.
- **Variety of Failure Modes:** Models failed for the following reasons, including filter and sorting errors, incomplete steps (e.g., not submitting a form or opening detailed pages), navigation errors, and misunderstanding the task's main goal.

>> Computer equipped agents are interesting, providing better performance, but still have limitations.

- **Operator Excels:** Operator outperforms other agents by using filters and structured search more effectively than broad keyword queries. It has a flexible action space with various tools, including "Ctrl+F" for web navigation. A key feature is its self-verification and self-correction, allowing it to recheck task requirements and independently rectify missing or incorrect filters.
- **Limitations of Operator:** Operator has two main limitations: it often fails to meet numerical and temporal constraints and despite its exploratory nature, it occasionally misses niche website features needed for tasks.

Assistant Bench – Generalist Agents Navigating the Web (HAL Link)

Eval Background: AssistantBench evaluates AI agents on realistic, time-consuming, and automatically verifiable tasks. It consists of 214 tasks that are based on real human needs and require several minutes of human browsing.

>> Models still cannot solve a (seemingly basic) task like booking a flight reliability

- **Sub-par Performance:** Top models can complete tasks with just a ~40% passing rate.
- **Variety of Failure Modes:**
    - **Task difficulty & length:** Expert-authored tasks are tougher for closed-book models, while web agents struggle on very short or very long trajectories, performing best at ~10 steps.

- **Failure modes by system:** Web agents fail from navigation/grounding errors or giving no output; closed-book models mostly hallucinate or give outdated info; retrieval models break down when retrieval is irrelevant or incomplete.
- **Overall capability:** Even ChatGPT with search + code tools often over-relies on snippets, hallucinates, or abstains—highlighting that no current approach solves AssistantBench reliably

A3: Marketing Positioning

ChatGPT Product Strategy POV: Market Positioning
Ian D'Silva | September 19, 2025

Summary:

- **Background:** Market positioning can help us understand relative strengths to press on and weaknesses to mitigate.
- **Focus Areas:** ChatGPT's biggest strength is its user scale / affinity and product polish, but its lack of a proprietary application ecosystem is an area to mitigate.
- **Places to Watch:** We should keep an eye out for:
    - Google given their vertical integration (compute, model, application) capabilities
    - Perplexity given their digital orchestrator focus
    - Meta / Social given their inability to develop a dominant social product yet.

Market Positioning Key Takeaways

- **ChatGPT's Strength is User Affinity & Scale:** ChatGPT's biggest advantage is its scale, with nearly 10x the combined web site visits of Gemini, Claude, Meta AI, Grok, and Perplexity.
- **ChatGPT's Weakness is No Native Ecosystem:** Unlike key competitors like Gemini and Meta, ChatGPT doesn't have a native application base that is already integrated into user's lives. It has a partnership with MSFT, but that isn't a truly native ecosystem. ChatGPT will need to figure out how to create value despite that by creating its own native applications, porting into competitors, or solving user problems in a different way for a truly proprietary and synergistic ecosystem.
- **Google is the Scariest Threat:** Google's deep vertical integration across the value chain (compute, models, applications) make it a formidable challenger. ChatGPT will need to win by leveraging its scale to demand API access to key applications and build its own proprietary ecosystem.
- **Perplexity is the Nearest Threat:** Perplexity's focus on Perplexity Assistant gives it a headstart on central orchestration tasks, but its lack of scale makes it difficult to challenge ChatGPT in the long-run.
- **Meta's Stumble Provides Temporary Opportunity in Social:** With Meta AI not having figured out a dominant social AI product yet, there is a small window to re-think social with superior AI capabilities.

See next page for further analysis.

Market Positioning

Source Document: *JTBDs and Pain Points*

Product Strengths / Weaknesses

Name	Strengths	Weaknesses	Takeaways for ChatGPT / OpenAI
ChatGPT	• **User Affinity:** Best-in-class brand awareness / strength, user base, retention • **Best Product UX**: Industry leading product design and polish	• **Native Ecosystem:** No proprietary ecosystem; reliant on the Microsoft partnership for ecosystem	➢ **Identify ways to leverage lead in user affinity** ➢ **Identify ways to mitigate lack of proprietary ecosystem**
Gemini	• **Native Ecosystem:** Consumer and Enterprise use cases with Android, Chrome, Search, Work Suite, Maps	• **Core Product:** Gemini does not have much product edge. • **Ecosystem Usage:** Though changing, Google has not capitalized on its ecosystem to win user love.	➢ **Leverage scale to demand API access to key ecosystem applications (e.g., mail, calendar)** ➢ **Build proprietary ecosystem applications**
Claude	• **Coding Specialization:** Has taken a focus on coding use cases. • **Brand Positioning:** Has garnered industry perception as the “ethical” and “safest” AI provider	• **Native Ecosystem:** Does not have a proprietary ecosystem. • **Distribution:** Does not have the same reach as competitors.	➢ **Improve coding-focused capabilities to not lose the use case** ➢ **OR Forego blinded specialization in coding in favor more general-purpose capability and intelligence**
Meta AI	• **User Reach:** Massive global reach with its social ecosystem • **Open-Source:** Natural developer incentives with open-source model.	• **Social Ecosystem:** The ecosystem is limited to social contexts; while important, does not cover many critical general use cases. • **Model Gap:** Currently has a gap in model capability with the slip up of Llama 4.	➢ **Unique potential opportunity to explore social AI with category leader currently hamstrung**
Grok	• **Twitter Distribution:** X / Twitter provides a large reach and unique social data stream • **Distinctive Personality:** Cultivate an unfiltered personality, appealing to a sub-segment of the market.	• **Volatile & Unfiltered:** Its strengths also poses risks; it’s personality and X data stream can make it susceptible to misinformation, toxicity, bias, etc.	➢ **Custom personalities or modes could be beneficial, as shown by Grok’s distinctive approach** ➢ **ChatGPT could plug in to more real-time data streams for recency relevancy**
perplexity	• **Answer Engines:** Has carved out a strong identity as a reliable, cited "answer engine“. • **Perplexity Assistant:** Has a headstart in developer integration and agentic use cases with Perplexity Assistant	• **Search Positioning:** Primarily focused on search use cases, though changing • **No Native Ecosystem:** Does not have a proprietary ecosystem, including foundation models, relying on fine-tuning external models	➢ **Leverage scale and product development chops to own the assistant use case**

Product Comparisons

Name	Total MAUs	Annual Website Visits	Model	Core Product Retention	Proprietary Ecosystem	Developer AI Ecosystem
ChatGPT	~1B MAUs / 700M WAUs	46.6B	Frontier	70%	Microsoft Partnership	Nascent
Gemini	400M MAUs	2.7B	Frontier	60%	Search, Work Suite, Maps, Android	Nascent
Claude	19M MAUs	1.2B	Frontier	>60%	Amazon Partnership	Nascent
Meta AI	1B MAUs	130M	Near-Frontier	N/A	Facebook, Instagram, Whatsapp, Messenger	Nascent
Grok	35M MAUs	687M	Frontier	N/A	Twitter / X	Nascent
perplexity	22M MAUs	1.5B	N/A	<50%	N/A	Nascent

Market Leader Contender

A4: Top-Level Ideas

ChatGPT Product Strategy POV: Top-Level Ideas
Ian D'Silva | September 19, 2025

Summary:

- We recommend building a **Personal Assistant Hub**, a persistent, top-level orchestrator that makes "Mission Control" tangible scheduling, prioritizing, and executing tasks from one place. This establishes ChatGPT as the user's trusted, top-level orchestrator and transforms ChatGPT from a transient tool to a persistent platform that users can offload cognitive burden to.
- We recommend building a **Developer Canvas**, an in-ChatGPT widget that is intelligently surfaced to assist with intent-completion. This product offers users a direct interface to the external digital world to complete tasks and reduce cognitive load, is scalable, and balances developer integration with quality control as we learn the right balance.

1. Host "Mission Control" For Completing Tasks:

We recommend building a **Personal Assistant Hub**, a persistent, top-level orchestrator that makes "Mission Control" tangible scheduling, prioritizing, and executing tasks from one place.

The first priority is creating a persistent solution over a transient one. Transient solutions, which ChatGPT largely is currently, requires users to bear the burden of context handling, which is against our goal of reducing user cognitive toll.

The second priority is to establish ChatGPT as the user's trusted, top-level orchestrator. This means that we will focus on establishing general planning capabilities, deferring end-to-end task completion. We believe this is the right approach since users will likely only have one trusted top-level orchestrator and end-to-end task completion is still not quite reliable enough for mass adoption. Not winning the top-level role implies getting layered, reducing the scope of value we can create for users. With that role established, we can add end-to-end task completion as reliability improves.

A Personal Assistant Hub met this criteria and would deliver tangible and meaningful value for users at scale, resulting in our prioritization.

2. Convert User Intent into Outcomes via a Developer Ecosystem:

We recommend building a **Developer Canvas**, an in-ChatGPT widget that is intelligently surfaced to assist with intent-completion, since it keeps ChatGPT in a central role, can help solve tasks end-to-end, and is a scalable solution.

The first priority is developing a product that helps users interface with the external digital world to complete tasks via developer integrations. This is a key lever to reducing cognitive load for users.

The second priority is creating a solution that can do this in a scalable way. There are more developers than we can possibly natively integrate with in bespoke ways.

The final priority is to come up with a solution that provides developers significant distribution to high-intent users, but does not cede too much real estate and control so as to potentially undermine the user experience. As we better learn how to create developer-oriented solutions, we can determine how to properly manage the tradeoff between developer customization and integration and quality control.

A Developer Canvas achieves all 3 of these criteria, resulting in its prioritization.

Appendix

Ideas & Prioritization Document: Idea Prioritization

Mission Control

ID	Idea	Description	25% Strategic Fit	20% Reach	25% User Impact	10% Effort	20% Confidence	100% Score	Rank	Comments
1	**Personal Assistant Hub**	Make Mission Control tangible by scheduling, prioritizing, and executing tasks from one place like a Personal Assistant would.	5.0	3.5	5.0	3.0	4.0	4.3	1	• **Strategic Fit:** Strong strategic fit with placing ChatGPT as central role • **Reach:** 70% of adults rely on a digital calendar to manage their lives (source) • **User Impact:** Strong ability to reduce cognitive load with reminders and task completion. • **Effort:** Moderate effort to establish new UX / window, but modularity deloads effort. • **Confidence:** Risk lies in user adoption and UX execution, derisked by bite-sized execution steps.
2	**Project Dashboard**	One place to see all active plans, blockers, and next actions; pause/resume across devices.	5.0	2.5	4.0	3.0	2.0	3.5	3	• **Strategic Fit:** Strong strategic fit with placing ChatGPT as central role • **Reach:** Many users may not be accustomed to the project medium and may not think of their lives in such a manner. • **User Impact:** For those using projects, there is strong ability to reduce cognitive load. • **Effort:** Moderate effort to establish new UX / window, but modularity deloads effort. • **Confidence:** Risk lies in user adoption and UX execution, derisked by bite-sized execution steps.
3	**Plan Composer Tool**	Turn any request into a living plan: steps, dependencies, cost/time tradeoffs, and a "Run" button with checkpoints	2.0	4.0	2.0	5.0	2.0	2.7	4	• **Strategic Fit:** Helps ChatGPT play a central planning role in user's lives but doesn't set up task completion. • **Reach:** Applicable to all ChatGPT users, making it easy for them to leverage. • **User Impact:** Is limited in seamlessly scoping beyond the use case to broader task completion. • **Effort:** Simple to develop; functionality already exists, just need to formally package it. • **Confidence:** Risk lies in consumers not finding it to be valuable above and beyond normal chat
4	**Playbooks Tools / Library**	Reusable, editable templates for common jobs (trip planning, moving, returns, event planning), surfaced at the right moment.	3.5	3.0	4.0	4.0	3.5	3.6	2	• **Strategic Fit:** Strong strategic fit with placing ChatGPT as central role in end-to-end task completion but isn't a scaled, robust solution. • **Reach:** While applicable to many users, users still need to discover the playbooks to use them. • **User Impact:** Will help users solve certain tasks in a more seamless way than navigating through ChatGPT and developers can port into the playbook; users still need to find the right playbook. • **Effort:** Scaling playbook by playbook lowers initial effort but requires significant effort to scale; successfully implement end-to-end is difficult • **Confidence:** Risk lies in effectively scaling and user discovery.
5	**Workflow Recorder**	"Record" a multi-step web flow once; MC replays it safely with variables (name, date, budget).	2.0	1.5	3.0	2.5	2.0	2.2	5	• **Strategic Fit:** Puts ChatGPT in a central role, but adds minimal value add from ChatGPT (relative to other offerings) • **Reach:** Only relevant to a select set of digital power users who would be willing to record workflows. • **User Impact:** Highly impactful for power users who want to automate a specific workflow, but solutions exist • **Effort:** Moderate effort to be able to control computer use. • **Confidence:** Low confidence in user adoption.

Developer Ecosystem

ID	Idea	Description	20% Strategic Fit	20% Reach	20% User Impact	20% Effort	20% Confidence	100% Score	Rank	Comments
1	**Developer Canvas**	An in-ChatGPT widget that is intelligently surfaced to assist with intent-completion.	5.0	5.0	5.0	2.0	3.0	4.3	1	• **Strategic Fit:** Aligns with vision of putting ChatGPT at the center of task completion. • **Reach:** Relevant to all users at scale. • **User Impact:** Integrating developers can help close the loops on solving tasks, reducing cognitive load. • **Effort:** Requires creation of a developer kit to integrate, approval / moderation process, and management to ensure reliability. • **Confidence:** Moderate to High execution risk given difficulty in making agentic tool use reliable.
2	**Individual Integrations**	Creating custom integrations in ChatGPT that can be surfaced depending on the task.	4.0	4.0	5.0	1.0	4.0	4.0	2	• **Strategic Fit:** Aligns with vision of putting ChatGPT at the center of task completion, but as scalable and open • **Reach:** Relevant to all users at scale, but deployment rate will be slower limiting use early on. • **User Impact:** Integrating developers can help close the loops on solving tasks, reducing cognitive load. • **Effort:** Requires bespoke integrations alleviating individual edge cases, but requires many custom integrations to scale. • **Confidence:** Less execution risk with native integrations.
3	**Deep Links**	Typed, signed links that open a specific plan or step in a developer's app with prefilled inputs/constraints and a callback to deliver the final receipt.	3.0	4.0	2.0	4.5	4.0	3.3	3	• **Strategic Fit:** Aligns with vision of solving tasks end to end, but core problem of reducing cognitive load is only partially solved. • **Reach:** Relevant to all users at scale, but deployment rate will be slower limiting use early on. • **User Impact:** Helps users navigate solving tasks, but doesn't complete it for them. • **Effort:** Low effort as developers can create a simple framework for ChatGPT to transfer users through with relevant context. • **Confidence:** High confidence in ability to execute; risk in getting developers on board despite low effort from meager value
4	**App Store**	Creating an app store that let's developers run full applications within the ChatGPT app environment.	2.0	4.0	4.0	3.0	3.0	3.2	4	• **Strategic Fit:** Gets developers involved but doesn't put ChatGPT at the center of solving problems. • **Reach:** Relevant to all users at scale, but adoption rate might be a limiting factor. • **User Impact:** Ability to solve problems end to end with the ChatGPT engine, unclear why it needs to be in ChatGPT app. • **Effort:** Moderate effort to create a platform for developers to build on. • **Confidence:** Risk of UX feeling clunky and unintentional.
5	**ChatGPT Embedding Kit**	An SDK that helps integrate ChatGPT's capabilities into your app to solve tasks.	2.0	2.0	4.0	4.0	2.0	2.7	5	• **Strategic Fit:** Doesn't place ChatGPT at the center of problem solving. • **Reach:** Limited reach since access point starts at developer properties. • **User Impact:** Ability to deliver high impact with ChatGPT engine. • **Effort:** High difficutly determining how to properly provide capabilities (e.g., context) above API capability • **Confidence:** High execution risk in establishing core value proposition

A5: Personal Assistant Hub Ideas

ChatGPT Product Strategy POV: Personal Assistant Hub Ideas
Ian D'Silva | September 19, 2025

Summary:

- The goal of this product is to build the habit of using ChatGPT as a personal assistant. This is best accomplished with a product that offers proactive assistance, has time-based (e.g., start of day) triggers, is habit-driving, and is lightweight for an MVP.
- Accordingly, we recommend building a **Morning Brief**, which generates a morning brief from a user's integrated calendar and email to centralize their daily obligations and provide contextual chat suggestions for executing tasks, as the initial product of Personal Assistant Hub.

Personal Assistant Hub Initial Product

Kickstarting our flyloop starts with Mission Control, so initial focus is on the Personal Assistant Hub.

The goal of this initial product is to associate ChatGPT as a true personal assistant above and beyond a Chatbot text box, serving as the foundation on which to build a top-level digital orchestrator.

To accomplish this, we must focus on four priorities for this initial product:

- **Proactive Assistance:** The product should focus on proactive value. ChatGPT already has a strong foundation for reactive assistance and personal assistants' highest value comes from taking proactive action.
- **Time-Based Assistance:** A proactive assistant can be triggered by events (e.g., receiving a text message) or by time (e.g., the start of the day). An event-based approach, while powerful, requires deep access to user data — a high-friction step for a new capability. A time-based approach, however, provides a more natural and predictable entry point to proactively assist users without significant initial permissions.
- **Habit-Driving:** Given its proactive and assistive nature, the product should be designed to create a daily habit. By providing predictable value at a recurring, time-based interval, we can create a reason for users to engage daily. This transforms ChatGPT from a transient tool into a persistent platform that users instinctively rely on to manage their day and reduce their cognitive toll.
- **Lightweight:** Lastly, as with all MVPs, it needs to provide customer delight in a lightweight form factor to serve as a foundation for iteration and expansion.

For these reasons, the first feature we will launch is the **Morning Brief**. The Morning Brief is an initial, lightweight form of the Personal Assistant Hub that focuses on managing and executing today's tasks. This is the ideal starting point:

- **It builds a persistent, recurrent, time-based daily habit:** It creates a recurrent, predictable reason for users to engage with ChatGPT at the start of their day, ingraining the behavior of using our platform as a central orchestrator.
- **It builds trust through value:** It delivers immediate utility by helping users plan their day, proving our value and earning the trust required to graduate to more context-aware, event-based features in the future.

By launching this lightweight MVP, we can gather critical feedback to guide the development of the more robust Personal Assistant Hub.

Appendix

Ideas & Prioritization Document: *Idea Prioritization*

Personal Assistant Hub									
			25%	20%	25%	10%	20%	100%		
ID	Idea	Description	Strategic Fit	Reach	User Impact	Effort	Confidence	Score	Rank	Comments
1	**Daily Planner**	Generates a morning brief from a user's integrated calendar and email to centralize their daily schedule, reminders, and contextual chat suggestions for prioritizing and executing tasks.	5.0	3.5	5.0	3.0	4.0	4.3	1	• **Strategic Fit:** Strong fit with proactiveness, time-based recurrence, and habit formation • **Reach:** Applicable to users with calendar / email integration • **User Impact:** High user impact for target users • **Effort:** Fairly lightweight; integrations adds complexity • **Confidence:** Moderate confidence in providing user value
2	**Nightly Debrief**	Engages the user at the end of the day to help them unload cognitive burden, reflect on accomplishments, and prepare for the next day.	4.0	3.5	3.5	3.0	3.5	3.6	2	• **Strategic Fit:** Strong fit with proactiveness, time-based recurrence, but weaker habit formation • **Reach:** Applicable to users with calendar / email integration • **User Impact:** End of day is less impactful than start of day • **Effort:** Fairly lightweight; integrations adds complexity • **Confidence:** Moderate confidence in providing user value
3	**Weekly Planner**	Proactively engage the user at the beginning of each week to help them define their top-level priorities.	3.0	3.5	3.5	3.0	3.0	3.2	4	• **Strategic Fit:** Proactive, time-based recurrence, but weak habit formation with weekly cadence • **Reach:** Applicable to users with calendar / email integration • **User Impact:** Moderate impact to plan week ahead • **Effort:** Fairly lightweight; integrations adds complexity • **Confidence:** Medium confidence in providing user value
4	**Reminders**	Chat-based and push notification based reminders based on context	3.0	4.5	2.5	4.5	4.0	3.5	3	• **Strategic Fit:** Is only partially proactive, doesn't drive recurrent habit • **Reach:** Applicable to many people • **User Impact:** Non-game-changing impact • **Effort:** Fairly lightweight • **Confidence:** Moderate confidence in providing user value

A6: North Star Metrics

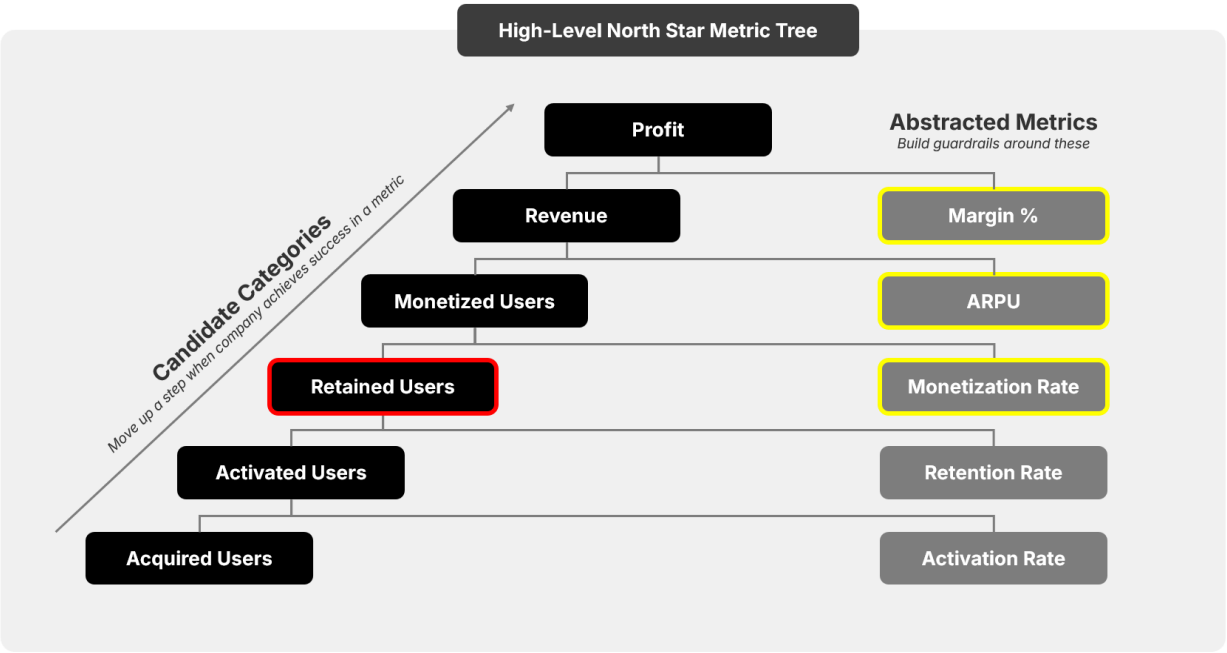
ChatGPT Product Strategy POV: North Star Metric
Ian D'Silva | September 19, 2025

Summary

- **Proposed NSM:** We recommend DAUs as the NSM, prioritizing building deep user engagement and expanding reach now, deferring an aggressive focus on monetization to the future.
- **Implementation & Guardrails:** To prevent unprofitable growth, a company-wide contribution profitability guardrail should be installed. Individual teams are encouraged to use specific engagement NSMs and quality guardrails to guide their local optimizations in a way that aligns with the top-level DAU goal.

North Star Metric

This North Star Metric (NSM) tree outlines a hierarchy of potential focus areas, starting from user growth at the bottom and culminating in business value at the top. Each level represents a candidate NSM, and a company's choice depends on its maturity and strategic goals. As a company achieves success in one metric, it can "move up a step" to the next. The progression typically moves from building a user base (Acquired and Activated Users) , to fostering engagement (Retained Users) , and finally to capturing value (Monetized Users, Revenue, and Profit).



With an established base of 700 million Weekly Active Users (WAUs), ChatGPT's primary focus should be on cultivating the highest form of retention: **Daily Active Users (DAUs)**. Since there is still a significant opportunity to expand reach and engagement, monetization should not be the main priority at this time. That can be prioritized later, once a deeply engaged user base is solidified.

Guardrails

By focusing on Retained Users as the North Star Metric, we are deliberately deferring a primary focus on optimizing for ARPU and Margin %. To prevent the risk of low-quality growth that this might entail, it is

essential to install a contribution profitability guardrail to monitor unit economics. Because the underlying economics may improve over time, this guardrail should be calculated assuming a steady-state variable cost per query.

Team Level Guardrails

While **DAUs** is the headline company NSM used for alignment and reporting, individual teams and products may benefit from more specific metrics. All team efforts should ladder up to the top-level goal, but product- or team-level NSMs can help steer surfaces more effectively.

Similarly team-specific metrics should be paired with their own quality guardrails to guide their local optimizations.

A7: FAQs

ChatGPT Product Strategy POV: FAQs

Ian D'Silva | September 19, 2025

Q1 **Why does building a top-level orchestrator make more sense than building agentic primitives (e.g., flight agent, shopping agent, etc.)?**

A We believe starting with a top-level orchestrator makes the most sense for a few reasons:

1. Starting Early on Top-Level Orchestration Matters

We believe that early advantages will compound. By working with users across all their top-level tasks, we'll develop a personalized knowledge-base that isn't reproducible, earning user loyalty. Users are likely to only have one top-level orchestrator and once a platform has met a critical personal advantage, switching costs can become insurmountable.

2. Building end-to-end use cases is really hard and requires bespoke development

Building an agent that can perform end-to-end task completion is really hard. They are unreliable, so they require a lot of tuning and monitoring to ensure they are acting as expected. It's impractical for ChatGPT to individually build each of infinite possible workflows until model capability improves to not require meticulous oversight across a broad set of use cases.

Furthermore, there still needs to be a mechanism to intelligently surface end-to-end workflows, otherwise users are still left to summon them, keeping ChatGPT in a reactive role that still places a burden on the user. The top-level orchestrator serves this role.

3. We'll be better at building end-to-end use cases

By taking a top-down approach, we'll end up building better end-to-end use cases when it comes time for it.

1) We'll be better because we have context. Assuming agents can work without reliability issues, their true power comes from incorporating context. There's no point if an agent can autonomously book me a hotel if it's completely on the wrong side of town. The top-level orchestrator's top-down approach brings the most context to solving use cases. Interpreting personal intent and creating an appropriate plan is the first step of end-to-end task completion after all, and that requires significant context.

2) We'll be better because we know which use cases matter. By building a top-level orchestrator, we'll gain real user feedback on which use cases are most sought after and valuable to users. Without it, users will be performing tasks outside of ChatGPT's purview, limiting our view into where we can add the most value.

2. Morning Brief Product Brief

ChatGPT Product Strategy: Morning Brief (Personal Assistant Hub MVP) PRD
Ian D'Silva | September 19, 2025

Context

See the Product Strategy Doc, which recommends the development of a personal assistant hub, for more context. This document is the PRD for the MVP of the Personal Assistant Hub, a Morning Brief.

Problem Alignment

Navigating the digital world is taxing. Completing a task (e.g., planning a trip) requires traversing a fragmented digital world, which incurs a cognitive toll to navigate it.

The Pain Point

Accomplishing everyday tasks is burdensome:

- **The digital world is fragmented.** To get restaurant delivery you go to Doordash or Uber Eats. To book flights you go to Kayak or Delta. To learn about a topic, you go to Wikipedia, Reddit, Google, Substack, etc.
- **Navigating this is taxing.** To navigate the digital world, we pay a *cognitive toll* of friction-filled steps (tedious page clicking, switching back and forth from apps and pages, filling forms, authenticating, etc.).
- **A typical task requires significant navigating**: A single goal like "I need to plan a weekend trip to see my parents" requires a dozen discrete digital tasks (creating an itinerary, booking flights, scheduling transportation, last minute purchases, etc.) for users to complete.

The burden has fallen on users to connect the fragmented digital world. This friction drains our energy and gets in the way of getting stuff done.

Personas

This spans multiple types of personas:

- **Ian, the Busy Professional:** I am Ian, a busy professional and I need help planning travel for my friend's wedding. The hustle and bustle of everyday life makes it difficult to properly prepare for the wedding (e.g., saving the date, booking my flight and hotel, buying a gift, figuring out plans, etc.), making me forgetful and tired.
- **Sharon, the Parent and Household COO:** I am Sharon, a parent and household COO and I need help surviving the day. From running errands, taking care of the baby, ordering groceries, making dinner, and finding time for myself, I never have enough time in the day, making me always feel overwhelmed and stressed.
- **Melanie, the Product Manager:** I am Melanie, a product manager and I need help turning big product ideas into cohesive outputs. My week is filled with scattered to-dos, meetings, research, and synthesis, leaving me juggling information across tools and struggling to pull it all together into a clear vision, making me feel drained and less effective.
- **Mark, the Student:** I am Mark, a college student and I need help staying on top of midterms. Between Econ exams, CS labs, group projects, and social commitments, my tasks are scattered across apps and I lose focus easily, leaving me feeling overwhelmed and stretched too thin.

See Product Strategy Appendix 1 - User Research for further insights.

Solution Alignment

High Level Approach

ChatGPT can simplify our interface with the internet by being a personal super-assistant that helps simplify turning intents into completed outcomes.

We can help solve this by building a **Personal Assistant Hub** in ChatGPT that helps with scheduling, prioritizing, and executing tasks from one place. The Personal Assistant Hub establishes ChatGPT as the user's trusted, top-level orchestrator but defers complete end-to-end tasks until reliability is sufficient.

The **Morning Brief** is an initial form of the Personal Assistant Hub, which focuses on managing and executing today's tasks. By launching a lightweight MVP, we can gather feedback on how to progress development towards a more robust Personal Assistant Hub.

Goals

The goal is to assist users with planning, prioritizing, and executing tasks for the day in one place, much like a personal assistant would. This entails:

- Reminding users of their obligations so they do not forget
    - e.g., "Today you have your big product review and, after work, a dinner date with your partner at Indienne."
    - e.g., "Your niece's birthday is next week. We need to get her a gift."
- Providing recommendations, advice, and assistance to complete tasks
    - e.g., "You need to book a flight for the wedding. This option on Friday at 6pm seems like a good option after your last meeting at 3pm"
    - E.g., "You had quite a few emails come in last night. Would you like me to summarize them for you?"

Non-goals

1. **Become a to-do list:** While we are surfacing things the user should be acting on, the goal is not to become a comprehensive to-do list manager for users. We are instead there to help take away burdens from the user, like a personal assistant would.
2. **Become a calendar tool:** While calendars will be helpful tools and contexts, the goal is not to simply be calendar manager. The goals are to determine actions the user needs to be prioritizing and assisting them with execution.

Execution Alignment

Key Features

We will create Morning Brief, which is a dedicated, auto-generated (each morning), and bespoke chat session that is served as a suggested prompt and includes:

- **Greeting**: A cheerful greeting to start your day on a positive note.
- **Day at a Glance:** Pulls in your context (calendar) to outline obligations today

- **Other Reminders:** Pulls in your context (calendar, mail, memory) to remind you about obligations in the future we may want to be thinking about today (e.g., getting ready for a flight tomorrow).
- **Smart Chat Suggestions:** Recommends chats to start based on your context to assist with accomplishing tasks that provide a prompt to start a new chat session. For example:
    - Calendar based recommendations
        - Prepare for meetings, buy a gift, etc.
    - Email based recommendations
        - Summarize overnight emails, prepare for a meeting, etc.
    - Memory based recommendations
        - Reminders
    - Other non-contextual recommendations
        - Weather outlook
        - News
- **Start a Chat:** Start a chat session with the plan as the chat context.

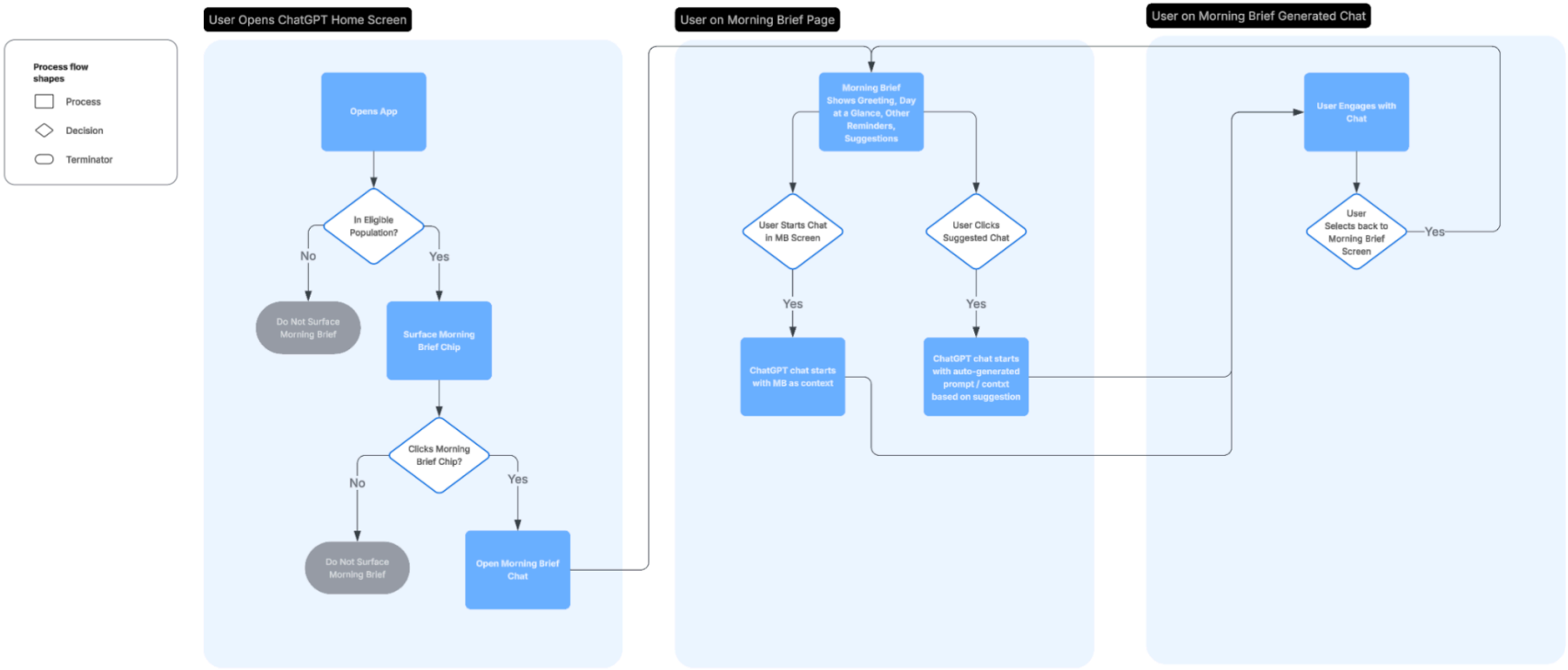
Currently, we will not build (but will consider for the future):

1. **Deprioritized Functionality**
    a. Extending to more general planning generation capability (e.g., weekly planner, etc.)
    b. Complementing “read” functionality with “write” functionality to Mail, Calendar integrations
    c. Adding in more context-providing integrations
    d. Developer integrations that help accomplish common tasks externally like making purchases (e.g., booking a hotel / flight, scheduling an uber, placing an e-commerce order for items or groceries), making a reservation, etc.
    e. More customized recommendation screens beyond providing a prompt for a chat session
2. **Deprioritized Surfacing**
    a. Making Morning Brief (part of) a default screen on main chat
    b. Suggest non-eligible users add mail / calendar integrations

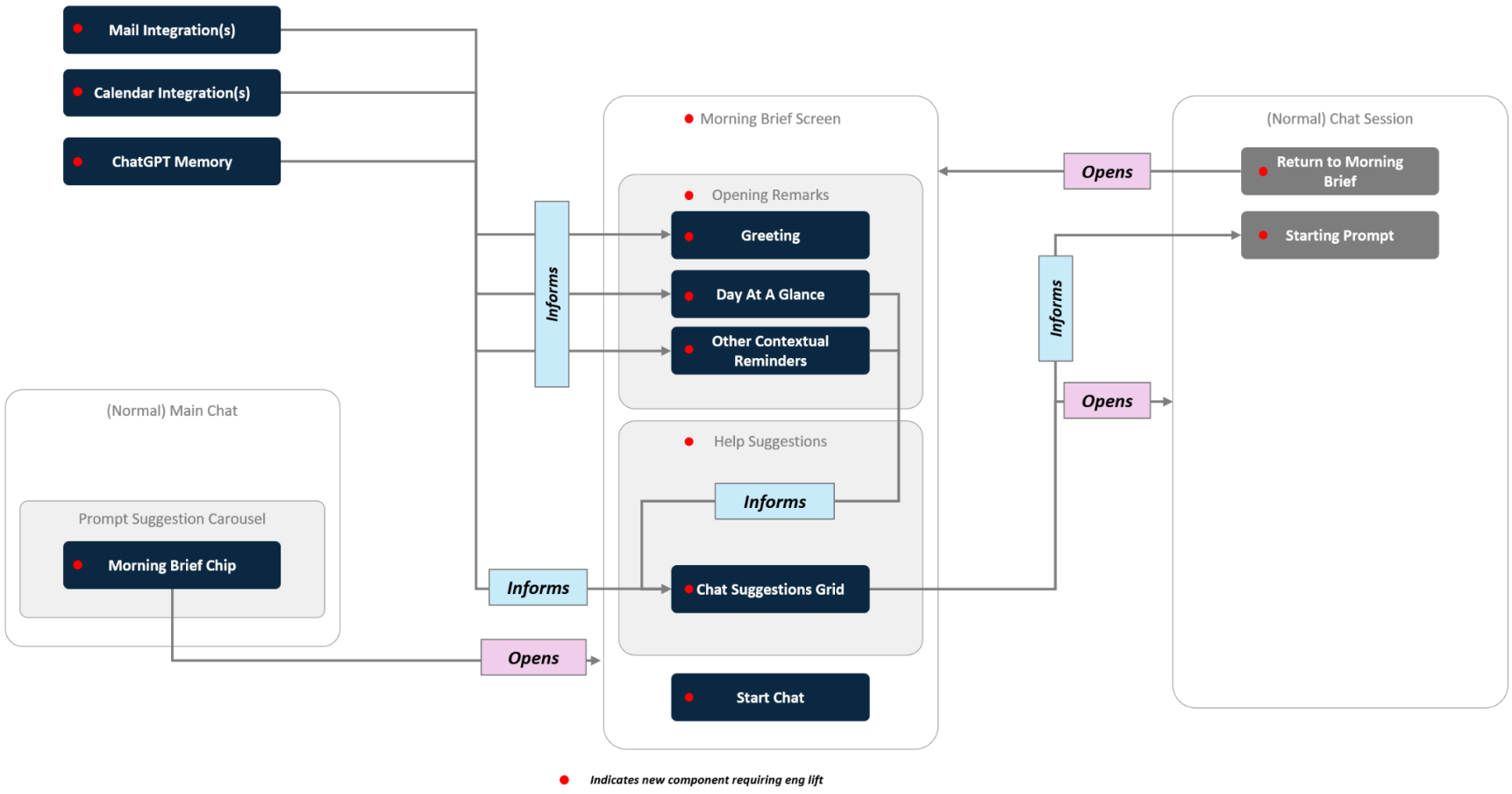
Key Flows

>> User Flows

Lucidchart document



>> High-Level Components and Architecture



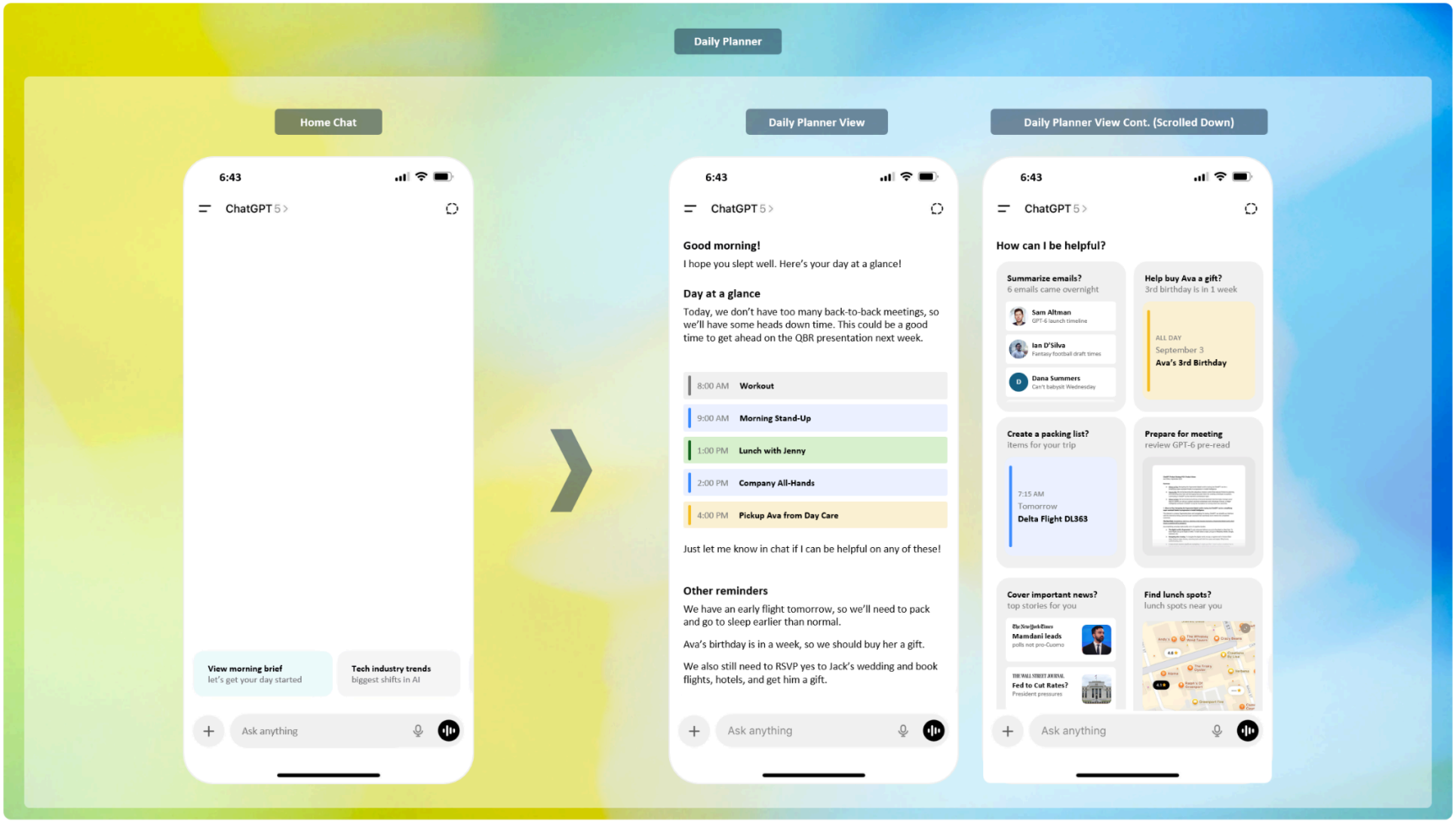
- **Platforms:** All ChatGPT primary surfaces (web, mweb, iOS, MacOS, Android, Windows)
- **Performance:** Morning brief is batch auto-generated early each morning (e.g., 6am)
- **API Integrations (via ChatGPT connectors):** Google Mail, Google Calendar, Google Drive, Outlook Mail, Outlook Calendar, Sharepoint

High-Level Requirements - Full product requirements to be developed after working with eng / design on feasibility and UX details.

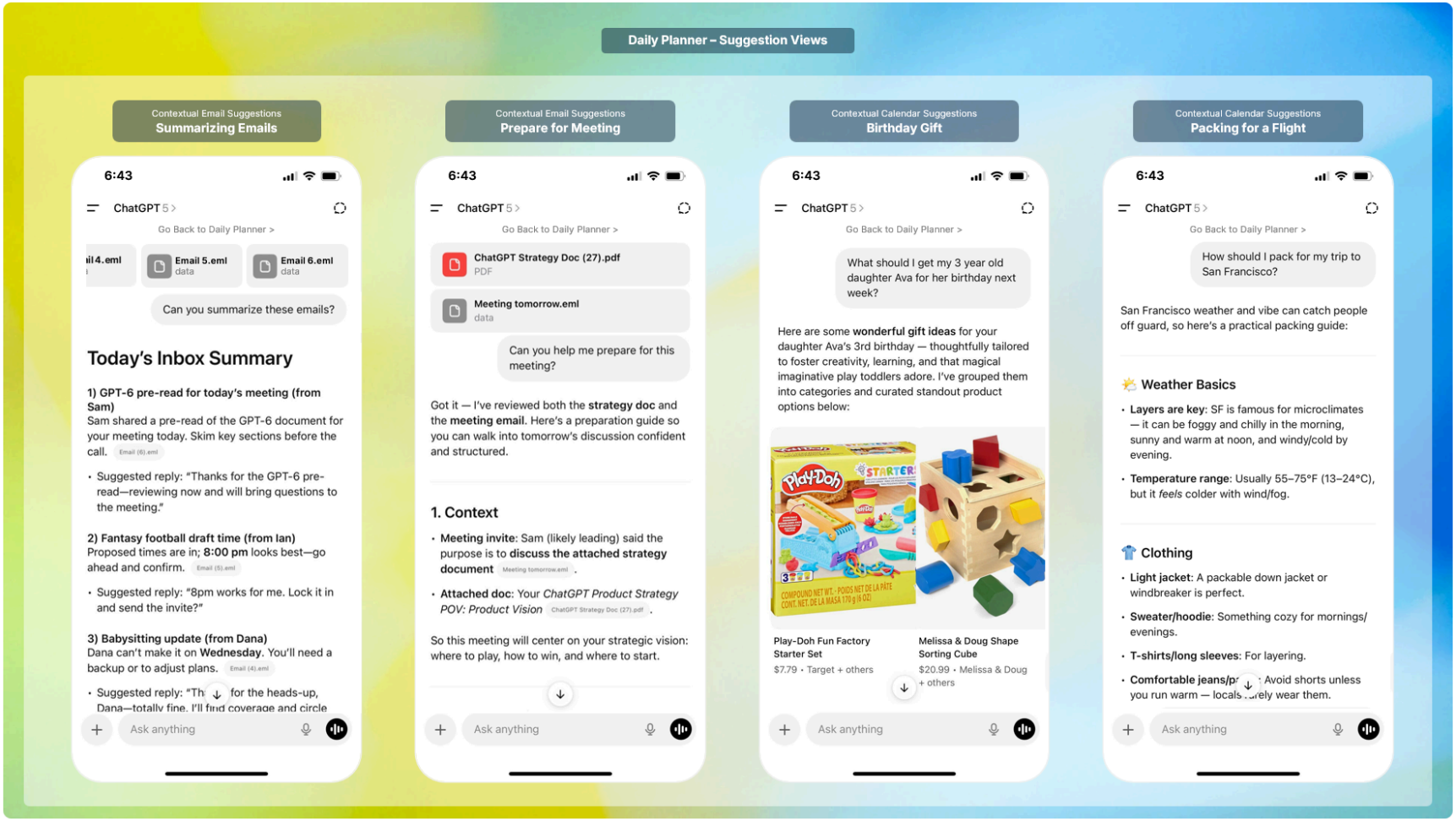
Priority	Requirement	Description
	Morning Brief Chip	**Build a chip that serves as the entry point to the Morning Brief screen**
P0	Morning Brief Prompt Suggestion Entry Point	Opens the Morning Brief screen. Leverage prompt suggestion chip infra with different destinations.
	Integrations	**Integrations that serve as context for creating the Morning Brief**
P0	GMail integration	Build an integration that pulls in a user's Gmail emails. Leverage ChatGPT's "Connector" infra.
P0	GCal integration	Build an integration that pulls in a user's GCal calendar events. Leverage ChatGPT's "Connector" infra.
P1	Outlook Mail integration	Build an integration that pulls in a user's Outlook emails. Leverage ChatGPT's "Connector" infra.
P1	Outlook Calendar integration	Build an integration that pulls in a user's Outlook calendar events. Leverage ChatGPT's "Connector" infra.
	Morning Brief Screen	**Start a custom chat session that provides an overview of a user's tasks, events, and other reminders to help them manage their day.**
P0	Greeting	Use model API to generate a simple, warm greeting to the user.
P0	Day at a glance – text	Leverage mail, calendar, and memory as context to AI generate a text summary of the user's day ahead with important notes and suggestions.
P1	Day at a glance – cal	Leverage mail, calendar, and memory as context to generate a calendar visual of the user's day ahead.
P1	Other Contextual Reminders	Leverage mail, calendar, and memory as context to AI generate a text summary of other reminders the user should be aware of (beyond today).
P0	Suggested Chats	Leverage mail, calendar, and memory as context to AI generated suggested chats. Leverage prompt suggestion chip infrastructure. Clicking stars a normal chat session with the morning brief as context
P1	Start Morning Brief Chat	Start a Chat with Morning Brief as context
	Suggested Chat Sessions	**Custom chat sessions based on Morning Brief chat suggestions, leveraging normal chat session infrastructure.**
P1	Back to Morning Brief Affordance	Text / button that allows users to go back to the Morning Brief screen.

>> Illustrative Mockups

From Chat to Morning Brief View



From Morning Brief View to Suggested Chats



Key Risks

- Users may not have integrate calendar and mail, limiting value
    - Mitigant: Do not surface to users without integration
    - Mitigant: Encourage users to integrate mail and calendar
- Users may have sparse mail or calendar based recommendations.
    - Mitigant: If a user has limited calendar or email data for the day, the planner will automatically prioritize non-contextual but still helpful "other recommendations".
    - Mitigant: Monitor users with integration but limited context for experience
- Reliability and quality of suggestions may vary or be harmful
    - Mitigant: Install recommendation guardrails and observability dashboards once live.

Launch Plan

Eligible Users

This will be eligible to all users with a mail and calendar integration on ChatGPT.

Ramp Plan

TARGET DATE	MILESTONE	DESCRIPTION	EXIT CRITERIA
Weeks 1+2	**Alpha**	**Internally test to bug hunt and gather feedback**	**Squash all P0/P1 bugs**
Week 1	Alpha (Working Team)	Internal testing with working team only	All P0/P1 bugs are resolved. Core user flow is functional without crashes.
Week 2	Alpha 2 (All Employees)	Internal testing with all company employees	No new P0/P1 bugs identified. No negative impact on app performance. Positive qualitative feedback from employees.
Week 3-6	**Canadian Beta**	**Test Product Market Fit**	**Achieve target adoption rate, success metrics, and no violation of guardrails.**
Week 3	Exposure Ramp (1%, 10%, 25%, 50%, 75%, 100%)	Public release ramp	System stability maintained at each stage. No spikes in latency, API errors, or crash rates
Week 6	Experiment Readout	Analyze experiment results, summarize findings, and present a formal recommendation on the feature's future (e.g., graduate, iterate, or deprecate).	A final ship/no-ship decision is made and communicated to all stakeholders.
Week 7-8	**(Optional) Rework**	**Use feedback to make p0 critical updates that threaten viability of the**	**p0 updates made**

	Development Period	product	
Week 9-12	**Public Launch**	**Evaluate Product Market Fit**	**Achieve target adoption rate, success metrics, and no violation of guardrails.**
Week 9	Exposure Ramp (1%, 10%, 25%, 50%, 75%, 100%)	Public release ramp	System stability maintained at each stage. No spikes in latency, API errors, or crash rates
Week 12	Experiment Readout	Analyze experiment results, summarize findings, and present a formal recommendation on the feature's future (e.g., graduate, iterate, or deprecate).	A final ship/no-ship decision is made and communicated to all stakeholders.

Measurement & Program Evaluation

Success Metrics and KPIs

Target Metric	Why We Care	Target
Success Metrics		
iDAUs	DAUs is our north star metric and we want to operate in service of that metric	Positive incrementality
Morning Brief DAUs	If the feature is valuable it will be adopted recurringly	Feature adoption in line with or stronger than other prompt suggestion adoption.
Guardrails and Quality Metrics		
Cost per User	We want to make sure unit economics are intact	In line with guardrail
Platform retention metrics	We want to make sure we aren't harming the rest of the platform	Not negatively incremental
Platform engagement metrics	We want to make sure we aren't harming the rest of the platform	Not negatively incremental
KPIs		
Morning Brief % Activated Users	Feature adoption is necessary to driving behavior change	N/A
Suggested Chat % Activated Users	Feature adoption is necessary to driving behavior change	N/A

Morning Brief Sessions Started	Feature engagement is indicative of user delight	N/A
Suggested Chat Sessions Started	Feature engagement is indicative of user delight	N/A
Suggested Chat Messages Sent	Feature engagement is indicative of user delight	N/A
Morning Brief DAU to MAU	Indicates user stickiness and power user ratio	N/A
Qualitative Customer Feedback	Fills in the narrative behind the metrics	N/A

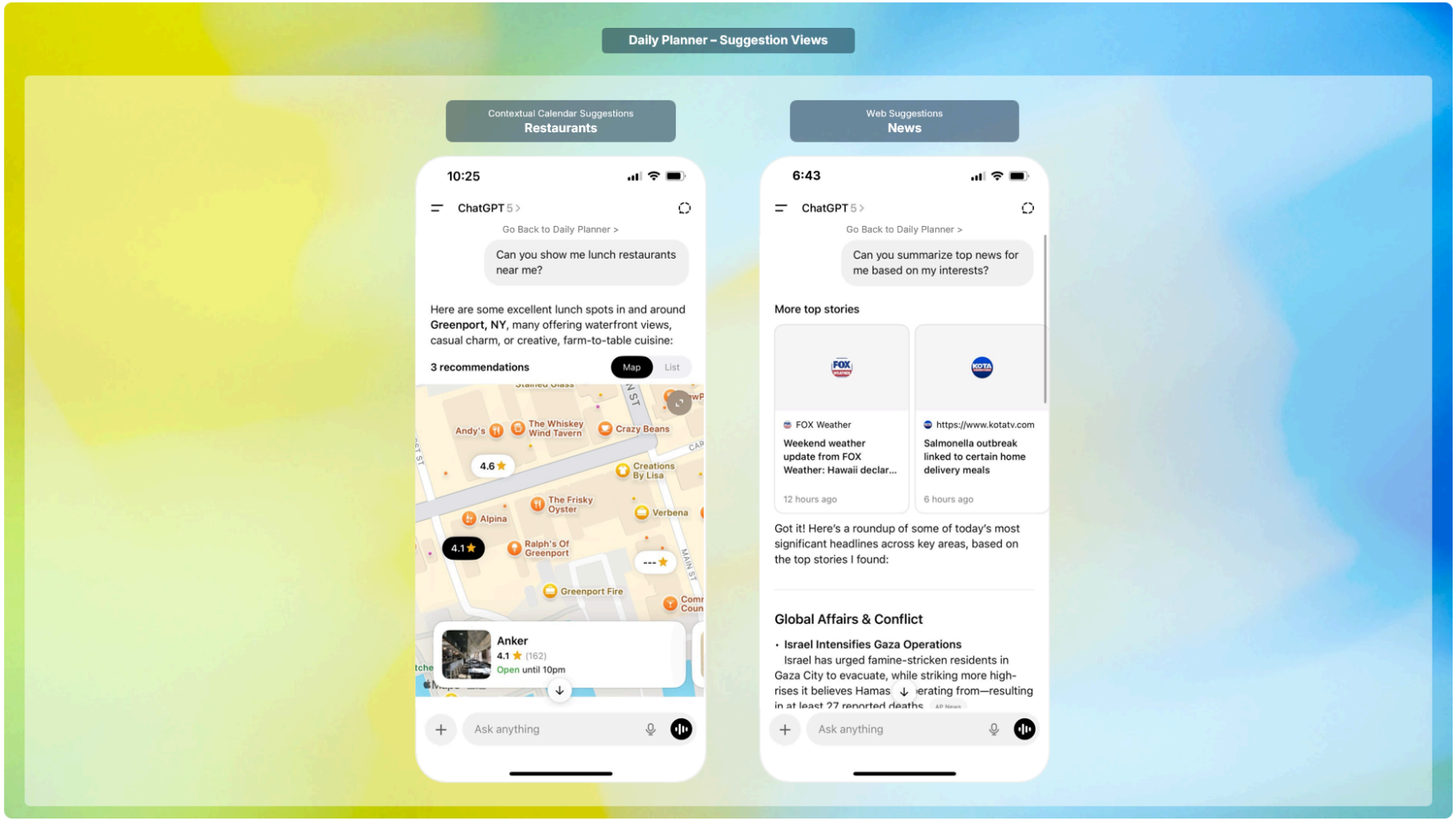
Experimentation Plan

- Experiment type:
    - A/B split test
- Assignment:
    - Eligible population is assigned at product launch by user_id (i.e., user_ids adding mail / calendar integration after launch are not part of eligible population)
- Initial Experimentation
    - Holdout group: 20%
    - Treatment Duration: 3 weeks
    - Attribution Window: 3 weeks
- Long term experimentation
    - Holdout group: 5%

Appendix 1

>> Suggested Chats Continued

From Morning Brief View to Suggested Chats



>> Broader Vision (Illustrative)

Morning Brief with Developer Integrations