

T H E C O M P L E T E

# CHATGPT

## BEGINNER'S GUIDE

---

*From First Login to Power User — Everything You Need to Know*

Custom GPTs • Deep Research • Canvas • Agent Mode  
Advanced Voice • Codex • Sora • Projects • Tasks

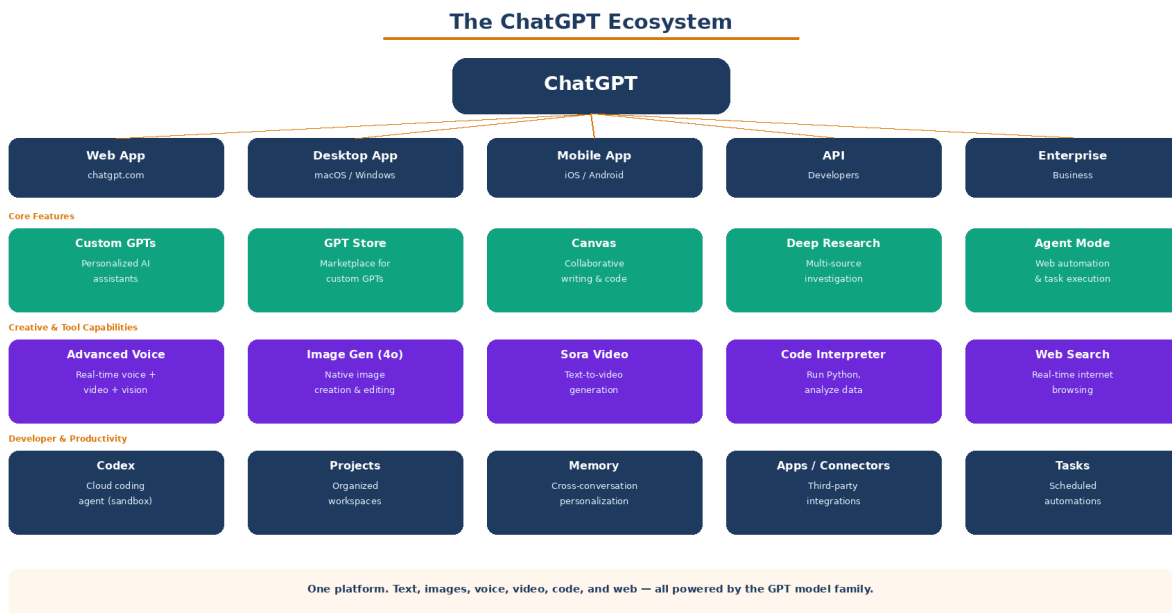
2026 Edition • Updated March 2026

# Chapter 1: What Is ChatGPT?

ChatGPT is an AI assistant built by OpenAI, a San Francisco-based AI research company. Since its launch in November 2022, it has become the fastest-growing consumer application in history, surpassing 300 million weekly active users by early 2025. ChatGPT can write, analyze data, generate images, create videos, browse the web, execute code, hold voice conversations, and even take actions on websites on your behalf.

What sets ChatGPT apart from a simple chatbot is the breadth of its capabilities. It's not one tool but an entire platform. You can have a natural voice conversation while it sees through your camera, ask it to research a topic by reading 50+ web sources, generate professional images, write and execute Python code on your data, or create a custom AI assistant tailored to your exact workflow. All of this happens inside a single interface.

## The ChatGPT Ecosystem



The diagram above shows just how much ChatGPT has grown. It spans five access points (web, desktop, mobile, API, and enterprise), a layer of core features like Custom GPTs and Canvas, creative capabilities like image and video generation, and a developer/productivity layer with Codex, Projects, and scheduled Tasks. Let's briefly introduce each access point.

## How to Access ChatGPT

**ChatGPT Web App (chatgpt.com)** — The browser-based interface. Works on any device with internet access. Create a free account with your Google, Microsoft, or Apple ID and start chatting immediately. This is the most feature-complete version and where new features typically appear first.

**Desktop App (macOS & Windows)** — A native app that mirrors the web experience with added conveniences: a system-wide keyboard shortcut (Option+Space on Mac, Alt+Space on Windows) lets you summon ChatGPT instantly from any application. Supports Advanced Voice Mode and screen sharing for voice conversations.

**Mobile App (iOS & Android)** — The mobile app includes everything the web version offers, plus Advanced Voice Mode with live camera and screen sharing. On Android, ChatGPT can replace Google Assistant as your default AI assistant. Conversations sync seamlessly across all your devices.

**API Platform** — For developers who want to build ChatGPT's capabilities into their own applications. Pay-as-you-go pricing with a generous free tier. Access all models including GPT-4o, o3, and specialized models like GPT-4.1.

**Enterprise & Team Plans** — Business-grade versions with SOC 2 compliance, SSO/SCIM, admin controls, shared workspaces, and a contractual guarantee that your data is never used for model training. Enterprise adds encryption key management and custom data retention policies.

## Creating Your Account

1

### Go to chatgpt.com

Open your browser and navigate to chatgpt.com. Click "Sign up" in the top right corner.

2

### Choose your sign-up method

Sign up with Google (fastest), Microsoft, Apple, or email. Google and Apple sign-ups skip email verification entirely.

3

### Verify if needed

If you signed up with email, check your inbox for a verification link. Click it to activate your account.

4

### Land in the chat interface

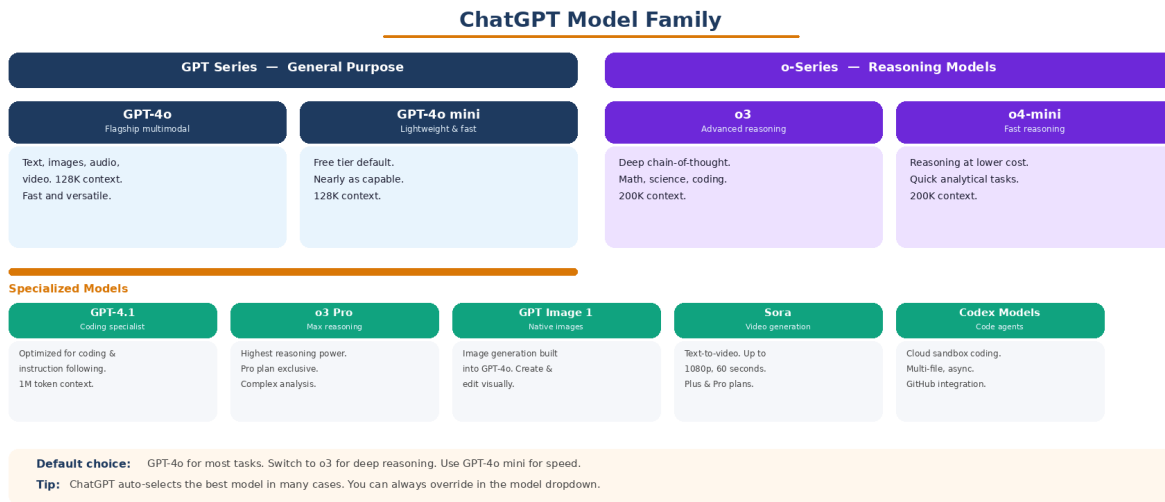
You're dropped into the main chat. On the free plan, you're using GPT-4o mini by default with limited GPT-4o access. Type anything to begin.

**TIP: Your Very First Prompt**

Skip “Hello” and go straight to something useful. Try: “I just got a new job as a marketing manager at a fintech startup. Help me create a 30-60-90 day plan with specific milestones for each phase.” The more context and specificity you provide, the more useful ChatGPT's response will be.

## Chapter 2: Understanding the Models

ChatGPT is powered by a family of AI models from OpenAI. These models fall into two distinct families: the GPT series (general-purpose, fast, multimodal) and the o-series (reasoning models that think step-by-step before answering). Understanding which model to use and when is one of the most impactful skills you can develop.



### The GPT Series — General Purpose

#### GPT-4o (“Omni”) — The Flagship

GPT-4o is OpenAI’s flagship model and the default for Plus and Pro users. The “o” stands for “omni” because it natively processes text, images, audio, and video in a single model. It has a 128K token context window (roughly 300 pages of text) and excels at virtually everything: writing, analysis, coding, creative work, and multimodal tasks. GPT-4o also powers ChatGPT’s native image generation, meaning you can ask it to create and edit images directly in conversation. This is the model you should use for roughly 80% of your work.

#### GPT-4o mini — Fast and Free

GPT-4o mini is the model that powers ChatGPT’s free tier. It’s a smaller, faster version of GPT-4o that delivers roughly 90% of its quality at a fraction of the cost. It shares the same 128K context window and handles everyday tasks like email drafting, Q&A, summarization, and light coding with ease. For most casual users, GPT-4o mini is genuinely sufficient.

#### GPT-4.1 — The Coding Specialist

Released as a developer-focused model, GPT-4.1 is specifically optimized for coding, instruction following, and long-context understanding. It outperforms GPT-4o on programming benchmarks and can maintain strong performance with context lengths up to 1 million tokens. If

you're a developer working with large codebases, GPT-4.1 is worth knowing about, though it's primarily available through the API and Codex.

## The o-Series — Reasoning Models

### o3 — Deep Reasoning

o3 is OpenAI's most advanced reasoning model. Unlike GPT-4o, which responds quickly, o3 spends extra time "thinking" through problems using chain-of-thought reasoning before giving you an answer. This makes it dramatically better at complex math, formal logic, scientific analysis, multi-step coding problems, and tasks that require careful sequential reasoning. It has a 200K token context window. The tradeoff is speed: o3 is slower than GPT-4o because it's literally thinking more.

### o4-mini — Fast Reasoning

o4-mini brings reasoning capabilities to a faster, more cost-effective model. It's ideal when you need analytical depth beyond GPT-4o but don't need o3's full reasoning power. Think of it as the middle ground: more thoughtful than GPT-4o, faster than o3.

### o3 Pro — Maximum Reasoning (Pro Plan Only)

o3 Pro is the absolute ceiling of reasoning capability, available exclusively on the \$200/month Pro plan. It uses the most compute per query of any ChatGPT model and is designed for the hardest problems: advanced mathematics, complex scientific reasoning, and intricate multi-step analysis where getting the right answer matters more than getting a fast one.

#### WHICH MODEL SHOULD YOU USE?

GPT-4o for 80% of tasks (writing, analysis, images, general work). o3 or o4-mini when you need step-by-step reasoning (math, logic, complex coding). o3 Pro only for the absolute hardest problems. GPT-4o mini when you want maximum speed and the task is simple.

Model	Type	Best For	Context	Plans
GPT-4o	General	Writing, analysis, images, multimodal	128K	Free (limited), Plus, Pro, Team
GPT-4o mini	General	Everyday tasks, speed priority	128K	Free (default), all plans
GPT-4.1	Coding	Programming, long-context code	1M	API, Codex
o3	Reasoning	Complex math, logic, science	200K	Plus, Pro
o4-mini	Reasoning	Moderate reasoning, faster	200K	Plus, Pro, Team
o3 Pro	Reasoning	Hardest problems, max accuracy	200K	Pro only

## Chapter 3: Pricing & Plans

ChatGPT's pricing has five tiers. The free tier is more capable than most people realize, and each paid tier unlocks progressively more powerful models, higher usage limits, and exclusive features. Here's what you get at each level.

**ChatGPT Plans & Pricing**

Free \$0/mo	Plus \$20/mo	Pro \$200/mo	Team \$25/user/mo	Enterprise Custom
<ul style="list-style-type: none"> <li>✓ GPT-4o mini (primary)</li> <li>✓ Limited GPT-4o access</li> <li>✓ Basic image generation</li> <li>✓ Web browsing</li> <li>✓ File uploads</li> <li>✓ GPT Store access</li> <li>✓ Limited voice mode</li> </ul>	<ul style="list-style-type: none"> <li>✓ Full GPT-4o access</li> <li>✓ o3-mini, o4-mini reasoning</li> <li>✓ DALL-E 3 + GPT-4o images</li> <li>✓ Advanced Voice + video</li> <li>✓ Deep Research (limited)</li> <li>✓ Canvas (full access)</li> <li>✓ Custom GPT creation</li> <li>✓ Sora video (50/mo)</li> <li>✓ Codex coding agent</li> <li>✓ Tasks (40 scheduled/mo)</li> </ul>	<ul style="list-style-type: none"> <li>✓ Everything in Plus</li> <li>✓ o3, o3 Pro reasoning</li> <li>✓ Unlimited Deep Research</li> <li>✓ Unlimited image gen</li> <li>✓ Sora video (500/mo HD)</li> <li>✓ Highest usage limits</li> <li>✓ Priority access to all</li> <li>✓ Pro-exclusive features</li> </ul>	<ul style="list-style-type: none"> <li>✓ Everything in Plus</li> <li>✓ Shared GPT workspace</li> <li>✓ Admin console</li> <li>✓ Higher message limits</li> <li>✓ Data NOT used for training</li> <li>✓ SOC 2 compliant</li> </ul>	<ul style="list-style-type: none"> <li>✓ Everything in Team</li> <li>✓ SSO / SCIM / RBAC</li> <li>✓ Custom data retention</li> <li>✓ Encryption key mgmt</li> <li>✓ Priority support</li> <li>✓ Dedicated account team</li> </ul>

Start Free → Upgrade to Plus when you want full GPT-4o + voice → Pro for unlimited reasoning

### Free — \$0/month

The free plan gives you GPT-4o mini as your primary model with limited access to GPT-4o. You get basic web browsing, limited image generation, file uploads, access to the GPT Store (use custom GPTs that others have built), and limited voice mode. There's no time limit on the free plan, but you'll hit message caps during peak usage. For exploration and casual use, it's a strong starting point.

### ChatGPT Plus — \$20/month

Plus is the sweet spot for most serious users. You get full GPT-4o access with significantly higher message limits, o3-mini and o4-mini reasoning models, full DALL-E 3 and GPT-4o native image generation, Advanced Voice Mode with video and screen sharing, Deep Research (limited monthly queries), Canvas for collaborative writing and coding, the ability to create and publish Custom GPTs, Sora video generation (50 videos/month at 480p), Codex coding agent access, and up to 40 scheduled Tasks per month. Plus subscribers also get priority access to new features.

### ChatGPT Pro — \$200/month

Pro is for power users who push ChatGPT to its limits daily. Everything in Plus, plus: o3 and o3 Pro (the most powerful reasoning models), unlimited Deep Research, unlimited image generation, Sora video at 1080p with 10x the limits, the highest message caps across all

models, and exclusive Pro-only features as they launch. If you rely on ChatGPT professionally for hours every day, Pro removes all friction.

### ChatGPT Team — \$25–30/user/month

Team is built for workplaces. Everything in Plus, plus: a shared workspace where team members can collaborate on Custom GPTs, an admin console for user management, higher message limits than individual Plus, and a contractual guarantee that your data is never used for model training. SOC 2 compliant. Billed at \$25/user/month annually or \$30 monthly.

### ChatGPT Enterprise — Custom Pricing

Enterprise adds the security and compliance features large organizations need: SSO via SAML, SCIM for identity provisioning, role-based access controls (RBAC), encryption key management (EKM), custom data retention policies, priority support, and a dedicated account team. Pricing is custom (typically estimated around \$60/user/month with annual contracts and minimum seat requirements). Contact OpenAI's sales team for a quote.

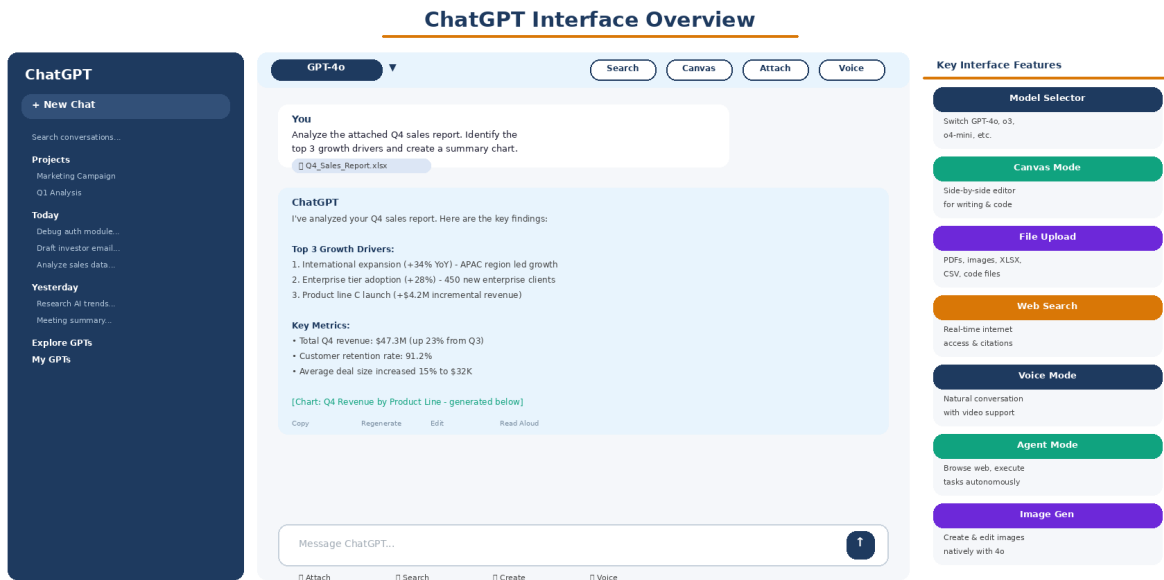
Feature	Free	Plus (\$20)	Pro (\$200)	Team (\$25–30)
Primary model	GPT-4o mini	GPT-4o	GPT-4o + o3 Pro	GPT-4o
Reasoning models	Limited	o3-mini, o4-mini	o3, o3 Pro, o4-mini	o3-mini, o4-mini
Image generation	Limited	DALL-E 3 + 4o native	Unlimited	DALL-E 3 + 4o native
Advanced Voice	Limited daily	Full access	Priority access	Full access
Deep Research	No	Limited/month	Unlimited	Limited/month
Canvas	Limited	Full	Full	Full
Custom GPTs	Use only	Create + publish	Create + publish	Create + share team
Sora video	No	50/mo (480p)	500/mo (1080p)	50/mo (480p)
Codex agent	No	Yes	Yes (priority)	Yes
Scheduled Tasks	No	40/month	Unlimited	40/month
Data training	Opt-out available	Opt-out available	Opt-out available	Never used

#### RECOMMENDATION

Start with Free to explore. Upgrade to Plus the moment you want full GPT-4o, voice mode, or Deep Research. Pro is only for heavy daily users who need unlimited everything. Team is the right call when 3+ people at your company are using ChatGPT.

## Chapter 4: Navigating the Interface

ChatGPT's interface looks simple on the surface, but it's packed with functionality that's not always obvious to new users. This chapter walks through every important element so you can work efficiently from day one.



### The Sidebar

The left sidebar is your conversation hub. Every chat you've ever had is saved here, organized by date (Today, Yesterday, Previous 7 Days, etc.). You can search conversations by keyword, pin important ones to the top, rename them for easy reference, or archive/delete old ones. The sidebar also gives you access to your Projects, the GPT Store ("Explore GPTs"), and your Custom GPTs ("My GPTs").

### The Model Selector

At the top of each new conversation, you'll see a model dropdown. Click it to switch between available models. The models you see depend on your plan. Free users see GPT-4o mini and limited GPT-4o. Plus users see GPT-4o, o3-mini, o4-mini, and more. Models with a "reasoning" badge (like o3) will think step-by-step before responding. In many cases, ChatGPT now auto-selects the best model for your query, but you can always override this.

### The Input Bar

The input bar at the bottom is where you type your messages, but it also houses several powerful tools:

**Attach files** — Click the paperclip/+ icon or drag-and-drop files into the chat. ChatGPT handles images (PNG, JPG, WebP, GIF), documents (PDF, DOCX, PPTX), spreadsheets (XLSX, CSV), code files, and more. You can upload multiple files at once.

**Search the web** — Click the globe icon to enable real-time web search. ChatGPT will browse the internet, read sources, and cite them in its response. This is on by default for queries about current events.

**Create images** — Click the image icon or simply describe what you want in your message. ChatGPT generates images natively using GPT-4o's built-in image generation.

**Voice mode** — Click the microphone/waveform icon to start an Advanced Voice conversation (Plus/Pro/Team plans).

**Canvas** — Click the Canvas icon to open the side-by-side editor for writing or coding tasks.

## Working with Files

File handling is one of ChatGPT's most underused capabilities. When you upload a file, ChatGPT can read, analyze, and transform it. Upload a spreadsheet and ask for a pivot table analysis. Upload a PDF contract and ask it to identify key terms and risks. Upload a photo and ask it to extract text, describe what it sees, or edit the image. Upload code and ask for a review, bug fixes, or documentation.

1

### Upload your files

Click the attachment icon or drag files directly into the chat window. You can upload multiple files at once across different formats.

2

### Be specific about what you need

Don't just say "look at this." Instead: "Analyze this sales spreadsheet. Identify the top 3 products by revenue growth rate and create a bar chart comparing them."

3

### Use Code Interpreter for data

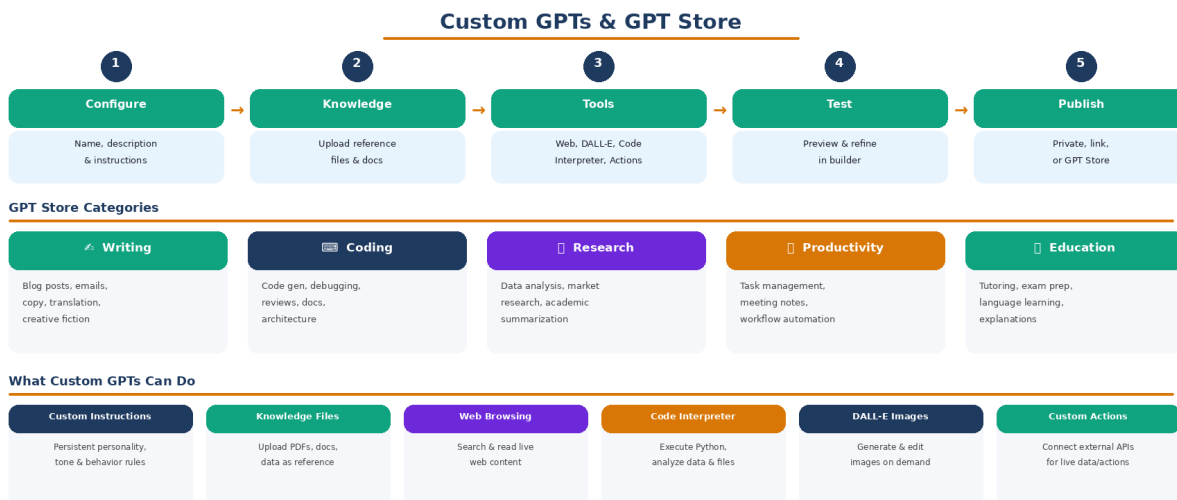
When you upload data files (CSV, XLSX), ChatGPT can use Code Interpreter to run Python on your data. It can create charts, run statistical tests, clean messy data, and generate new files you can download.

### TIP: What Code Interpreter Can Do

Code Interpreter is essentially a Python sandbox running inside ChatGPT. It can: analyze data with pandas, create visualizations with matplotlib, process images with PIL, parse PDFs, convert file formats, run statistical analysis, train simple ML models, and generate downloadable output files. Think of it as having a data scientist available on demand.

## Chapter 5: Custom GPTs & the GPT Store

Custom GPTs are personalized versions of ChatGPT that you configure with specific instructions, knowledge, and tools. They're one of the platform's most powerful features because they let you encode your expertise, preferences, and workflows into a reusable AI assistant that works exactly how you want, every time.



### Why Custom GPTs Matter

Without a Custom GPT, you're starting every conversation from scratch. You have to re-explain your role, your preferences, your context, and your desired output format each time. A Custom GPT remembers all of this permanently. It's the difference between hiring a new assistant every day vs. having one who already knows your business.

### Anatomy of a Custom GPT

**Instructions** — The core of any GPT. This is a text prompt that tells the GPT who it is, how it should behave, what tone to use, what to focus on, and what to avoid. The more detailed your instructions, the better the GPT performs. Think of it as a comprehensive job description.

**Knowledge Files** — Upload up to 20 files (PDFs, documents, spreadsheets, code) that the GPT uses as reference material. Upload your company's style guide, product documentation, training materials, or any other reference the GPT should consult when answering questions.

**Capabilities** — Toggle on or off: Web Browsing (search the internet), DALL-E Image Generation (create images), and Code Interpreter (run Python, analyze data). Enable only what your GPT needs.

**Actions** — Connect your GPT to external APIs. This is the advanced feature: you can make your GPT fetch live data from your CRM, post to your CMS, query your database, or interact with any REST API. Actions transform a GPT from an advisor into an agent that takes action.

## Building Your First Custom GPT

1

### Open the GPT Builder

Click your profile in the sidebar, then “My GPTs” → “Create a GPT.” Or go directly to [chatgpt.com/create](https://chatgpt.com/create).

2

### Use the Create tab (easy mode)

Describe what you want in plain English: “I want a GPT that reviews my Python code for bugs, follows PEP 8, and suggests performance improvements.” The builder will generate the configuration for you.

3

### Fine-tune in the Configure tab

Switch to Configure for precise control. Edit the name, description, instructions, conversation starters, knowledge files, and capabilities manually. This is where you make the GPT truly yours.

4

### Upload knowledge files

Add any reference documents your GPT should know about. A style guide, a product spec, a FAQ document, code examples, or training data. The GPT will reference these when answering.

5

### Test and iterate

Use the Preview panel on the right to test your GPT in real time. Ask it questions, check that its tone and behavior match your expectations, and refine the instructions until it's dialed in.

6

### Publish

Choose your audience: “Only me” (private), “Anyone with a link” (shareable), or “Everyone” (published to the GPT Store). You can change this anytime.

## The GPT Store

The GPT Store is a marketplace with thousands of Custom GPTs built by individuals, companies, and OpenAI partners. Browse by category (Writing, Coding, Research, Education, Productivity, Lifestyle) or search for specific use cases. Popular GPTs include Consensus (searches 200M+ academic papers), Canva (design directly in chat), Zapier (automation), and

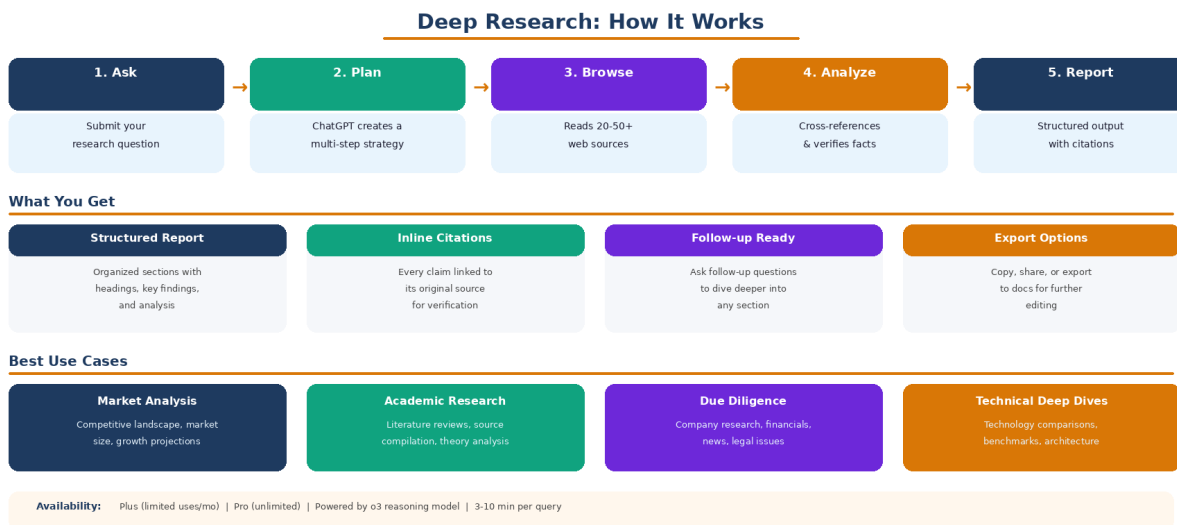
Wolfram (computational knowledge). OpenAI offers revenue sharing to builders whose GPTs get high usage.

### **EXAMPLE GPT INSTRUCTIONS**

Here's a template for a high-quality Custom GPT: "You are [role]. Your job is to [primary task]. When the user provides [input type], you should [specific action]. Always [behavioral rule]. Never [restriction]. Format your output as [format]. If the user's request is ambiguous, ask one clarifying question before proceeding. Use a [tone] tone."

## Chapter 6: Deep Research

Deep Research is ChatGPT's autonomous research agent. When activated, it creates a multi-step research plan, browses dozens of web sources, cross-references information across them, and compiles a comprehensive report with inline citations. It replaces hours of manual searching with 3–10 minutes of automated investigation.



### How It Works Under the Hood

Deep Research is powered by the o3 reasoning model, which gives it the ability to plan, evaluate, and synthesize information at a level far beyond standard web search. Here's the actual process:

- 1. Planning** — ChatGPT analyzes your question and creates a structured research plan: which aspects to investigate, what types of sources to look for, and how to break the question into sub-questions.
- 2. Browsing** — It opens a browser and systematically visits 20–50+ web pages. It reads articles, reports, documentation, and other sources relevant to each sub-question.
- 3. Analysis** — It cross-references facts across multiple sources, identifies consensus and disagreements, evaluates source credibility, and filters out unreliable information.
- 4. Synthesis** — It compiles everything into a structured report with clear sections, key findings, and inline citations that link back to the original sources so you can verify any claim.

### Using Deep Research

1

Write your research question

Be specific. Instead of “tell me about EVs,” try: “What is the current state of the solid-state battery market? Cover key companies (QuantumScape, Toyota, Samsung SDI), timelines to commercialization, performance benchmarks vs. lithium-ion, and major technical challenges remaining.”

2

### Activate Deep Research

Select Deep Research from the model/mode dropdown at the top of the chat, or click the Deep Research icon if available.

3

### Review the research plan

ChatGPT shows you its planned approach. You can modify it (add focus areas, remove sections, change scope) before it begins.

4

### Wait for results

The research typically takes 3–10 minutes depending on complexity. You'll see a progress indicator showing which sources ChatGPT is currently reading.

5

### Explore the report

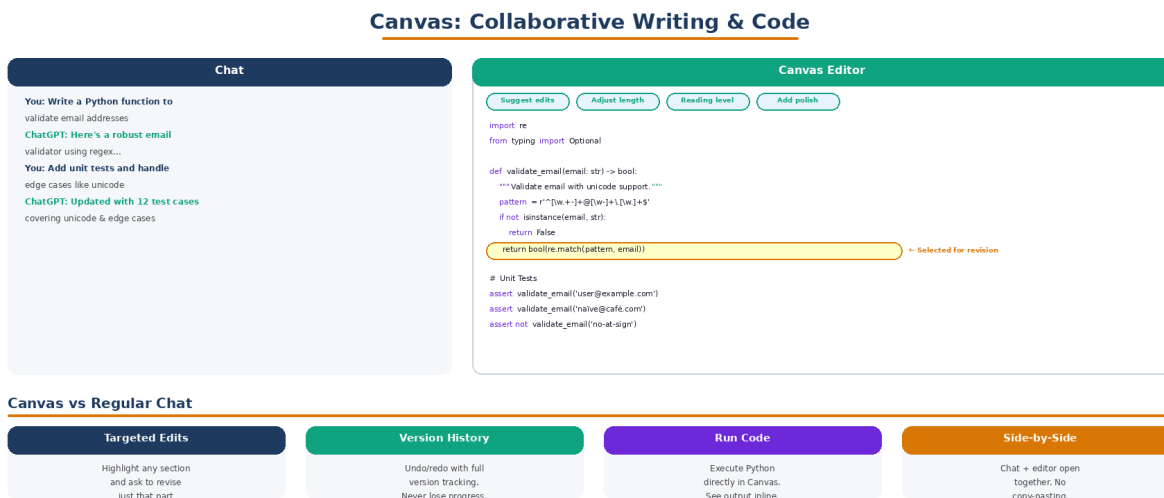
Read the structured report with clickable citations. Ask follow-up questions to drill deeper into any section. Export to Google Docs or copy for further use.

## USAGE LIMITS

Plus subscribers get a limited number of Deep Research queries per month. Pro subscribers get unlimited Deep Research. Each query uses significantly more compute than a regular message because it involves extended browsing and reasoning.

## Chapter 7: Canvas — Collaborative Writing & Coding

Canvas is ChatGPT's side-by-side editing environment. Instead of reading a response in the chat and copy-pasting it elsewhere, Canvas gives you a dedicated workspace where you and ChatGPT collaborate on the same document or code file in real time. It's a fundamentally different way to work with AI.



### Why Canvas Changes Everything

In regular chat, if ChatGPT writes 500 words for you and you want to change one paragraph, you have to describe the change and hope ChatGPT understands which paragraph you mean. In Canvas, you simply highlight that paragraph, type “make this more concise,” and ChatGPT updates just that section while leaving everything else untouched. Meanwhile, you can directly edit other parts by typing, just like a Google Doc.

### Canvas for Writing

When working on written content in Canvas, you have access to a set of built-in shortcuts:

**Suggest edits** — ChatGPT reviews your text and proposes improvements with track-changes-style suggestions you can accept or reject individually.

**Adjust length** — Make text shorter or longer with a single click. Useful when you're over a word count or need to expand a thin section.

**Reading level** — Adjust the complexity from elementary to graduate level. Instantly rewrite your content for a different audience.

**Add polish** — Improve grammar, flow, and clarity across the entire document without changing the substance or meaning.

**Translate** — Translate the entire document into another language while maintaining formatting and nuance.

## Canvas for Code

Canvas is equally powerful for coding. You get a syntax-highlighted editor where you can:

**Fix bugs** — ChatGPT identifies and fixes bugs in your code while explaining what was wrong.

**Add comments** — Auto-generate inline comments and docstrings for your code.

**Convert language** — Convert code from one programming language to another (Python to JavaScript, etc.).

**Add logging** — Insert logging statements throughout your code for debugging.

**Run code** — Execute Python directly in Canvas and see the output inline. No need to switch to a terminal.

**Highlight & revise** — Select a specific function and ask ChatGPT to “debug this,” “optimize this,” or “convert to async.” It changes only what you selected.

### WHEN TO USE CANVAS

Use Canvas whenever you're working on something longer than a paragraph that you'll iterate on: blog posts, reports, emails, code, scripts, documentation. For quick Q&A or short answers, regular chat is faster.

## Chapter 8: Advanced Voice Mode

Advanced Voice Mode transforms ChatGPT from a text-based tool into a real-time conversational partner. It's not like talking to Siri or Alexa where you give one command and get one response. It's a genuine, flowing conversation where you can interrupt, think out loud, change topics, and even share what you're seeing through your camera.

**Advanced Voice Mode + Vision**

Real-Time Voice	Emotional Tone	Live Video	Screen Share
Near-zero latency. Natural back-and-forth. Interrupt anytime.	Expressive voices. Adjusts mood & energy. 9+ voice options.	Share camera feed. ChatGPT sees & responds in real time.	Show your screen. Get help with apps, docs, or code.

**What You Can Do**

Interview Prep	Language Learning	Brainstorming	Visual Help	Hands-Free
Practice with AI interviewer. Real-time feedback on answers.	Conversational practice. Pronunciation help. Any language.	Think out loud. Get instant pushback and new ideas.	Point camera at anything for instant identification/help.	Cook, drive, exercise. Full AI assistance without typing.

**Background Mode**

Voice conversations continue even when your phone is locked or you're using another app. Perfect for hands-free workflows: get directions while driving, follow recipes while cooking, or take notes during walks. ChatGPT stays active in the background and responds naturally.

**Available on:** Plus, Pro, Team, Enterprise | Mobile (iOS & Android) + Desktop | Free users get limited daily access

### What Makes It Different

**Near-zero latency** — ChatGPT responds almost instantly, making conversations feel natural rather than turn-based. There's no awkward pause while it "processes."

**Emotional range** — The voice model can express excitement, concern, humor, and warmth. It adjusts its tone and energy based on the conversation. It can whisper, speak with enthusiasm, or maintain a calm, soothing tone.

**Interruptible** — Jump in at any point. You don't have to wait for ChatGPT to finish speaking. It handles interruptions gracefully and adjusts its response on the fly.

**9+ voice options** — Choose from multiple voice personalities in Settings, each with a distinct character and tone. OpenAI occasionally adds seasonal voices (like Santa during the holidays).

**Background mode** — Voice conversations continue even when your phone is locked or you're using another app. Perfect for hands-free help while cooking, walking, driving, or exercising.

### Vision + Voice: ChatGPT Can See

The most powerful aspect of Advanced Voice Mode is vision integration. You can share your phone's camera or your screen during a voice conversation:

**Camera sharing** — Point your camera at anything and ChatGPT sees it in real time. Ask “What plant is this?” while pointing at a flower, or “Is this rash something I should worry about?” while showing your arm. It can read handwritten notes, identify objects, explain diagrams, and describe scenes.

**Screen sharing** — Share your device’s screen so ChatGPT can see what you’re looking at. It can walk you through settings menus, help debug code on your screen, explain UI elements, give feedback on designs, or help you navigate unfamiliar software.

## Best Use Cases

**Interview preparation** — Have ChatGPT role-play as an interviewer for any position. It asks realistic questions, listens to your answers, and gives constructive feedback on content, delivery, and confidence.

**Language learning** — Practice conversation in any language. ChatGPT adjusts difficulty to your level, corrects your grammar in context (not mid-sentence), and explains cultural nuances.

**Brainstorming** — Think out loud. ChatGPT riffs with you, pushes back on weak ideas, builds on strong ones, and helps you organize scattered thoughts into structured plans.

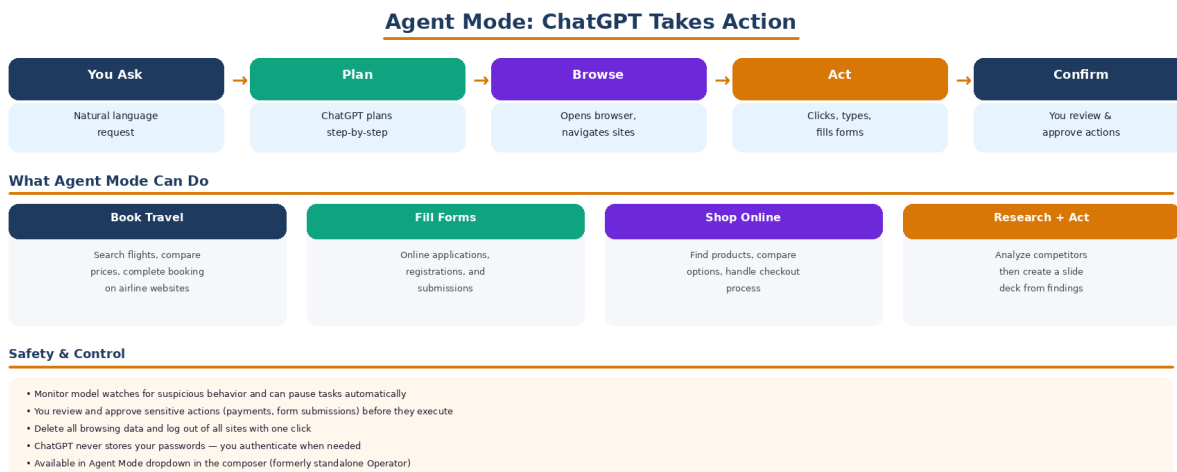
**Real-world visual help** — Get help identifying ingredients, reading labels in foreign languages, diagnosing car dashboard lights, understanding electrical panels, or navigating unfamiliar transit systems.

### AVAILABILITY

Advanced Voice Mode is available on Plus, Pro, Team, and Enterprise plans on mobile (iOS/Android) and desktop apps. Free users get limited daily access. Vision (camera + screen sharing) requires a paid plan.

## Chapter 9: Agent Mode — ChatGPT Takes Action

Agent Mode is one of ChatGPT's newest and most significant capabilities. It evolved from the standalone “Operator” product and is now integrated directly into ChatGPT. In Agent Mode, ChatGPT can open a web browser, navigate to websites, click buttons, fill out forms, and complete multi-step tasks on the internet on your behalf.



### How It Works

Agent Mode is powered by the Computer-Using Agent (CUA) model, which combines GPT-4o's vision capabilities with reinforcement learning to interact with graphical user interfaces. It literally sees web pages as screenshots and uses mouse clicks and keyboard input to interact with them, just like a human would. This means it works on any website without needing custom API integrations.

**1**

#### Select Agent Mode

From the model dropdown in the composer, select “Agent Mode.” Enter your request in natural language.

**2**

#### ChatGPT plans the task

It breaks your request into steps and begins executing. You can see what it's doing in real time as it navigates web pages.

**3**

#### It browses and acts

ChatGPT opens websites, clicks links, fills forms, compares options, and processes information autonomously.

4

### You review and approve

For sensitive actions (payments, form submissions, account changes), ChatGPT pauses and asks for your confirmation before proceeding.

## What Agent Mode Can Do

**Book travel** — “Find and book the cheapest direct flight from NYC to London in June, departing on a Friday.” ChatGPT searches airline sites, compares prices, and walks you through booking.

**Shop online** — “Find a 4K monitor under \$400 with USB-C and good reviews for Mac.” It searches retailers, reads reviews, compares specs, and can complete checkout.

**Fill out forms** — “Fill out this job application using my resume.” It navigates the application, fills fields, and lets you review before submission.

**Multi-step research + action** — “Analyze 3 competitors in the project management space, then create a comparison slide deck.” It researches, synthesizes, and produces deliverables.

## Safety and Privacy

OpenAI built several safety layers into Agent Mode. A dedicated “monitor model” watches for suspicious behavior and can pause tasks automatically. You review and approve any sensitive actions before they execute. ChatGPT never stores your passwords. You can delete all browsing data and log out of all sites with one click in Privacy settings.

## Chapter 10: Image Generation & Sora Video

ChatGPT can both create and edit images natively, and with Sora integration, it can now generate videos too. These aren't separate tools you switch to; they're built directly into the chat experience.

### GPT-4o Native Image Generation

In March 2025, OpenAI launched native image generation within GPT-4o itself. This was a major upgrade over the previous DALL-E integration because GPT-4o understands your conversation context. If you've been discussing your brand's visual identity for 10 messages, the image it generates will reflect that entire conversation, not just the last prompt.

#### Key Capabilities

**Text rendering** — GPT-4o can accurately render text within images, a historically difficult problem. It can create posters, social media graphics, memes, and diagrams with readable, correctly spelled text.

**Image editing** — Upload any image and ask for specific edits. "Change the sky to sunset," "Remove the person on the left," or "Change my shirt to blue." GPT-4o changes only what you ask while preserving lighting, composition, and other elements.

**Inpainting (Select Tool)** — A 2025 addition that lets you select a specific region of an image to regenerate, rather than editing the whole image. Fix a hand, change a background element, or adjust one detail.

**Style transfer** — Upload a reference image and ask ChatGPT to generate new images in that style. Great for brand consistency.

**Full commercial rights** — You own complete commercial rights to every image ChatGPT generates for you. Use them in products, marketing, websites, or anywhere else.

### Sora Video Generation

Sora is OpenAI's video generation model, now integrated directly into ChatGPT for Plus and Pro users. You describe a scene in text, and Sora generates a video clip.

**Resolution** — Up to 1080p for Pro users, 480p for Plus users.

**Duration** — Up to 20 seconds per clip (longer for Pro), extendable up to 60 seconds with Sora 2.

**Aspect ratios** — Widescreen (16:9), vertical (9:16), and square (1:1).

**Monthly limits** — Plus: 50 videos/month at 480p. Pro: 500 videos/month at 1080p with longer durations.

To generate a video, simply describe the scene you want: “A golden retriever running through autumn leaves in a park, cinematic slow motion, warm afternoon light.” Sora handles the rest.

**TIP: Combining Image + Video**

Use GPT-4o image generation to create a still frame you love, then ask Sora to animate it into a video. This gives you much more control over the final result than text-to-video alone.

## Chapter 11: Projects, Memory & Tasks

ChatGPT has evolved beyond one-off conversations. With Projects, Memory, and Tasks, it can now maintain organized workspaces, remember your preferences over time, and automate recurring workflows. These features transform ChatGPT from a reactive tool into a proactive assistant.

### Projects — Organized Workspaces

Projects are smart workspaces that group related conversations, files, and custom instructions in one place. Instead of having 50 scattered conversations about the same initiative, you create a Project that keeps everything organized and contextual.

**What you get** — Each Project has its own custom instructions (applied automatically to every conversation in the project), its own uploaded files (accessible across all conversations), and a dedicated space in your sidebar for easy access.

**Custom instructions per project** — This is the key differentiator. You can set project-level instructions like “All responses should use our company’s brand voice,” “Always reference the product spec I uploaded,” or “Format code in TypeScript with our team’s conventions.” These apply on top of your global custom instructions.

**File context** — Upload files to a Project and they become available to every conversation in that project. Upload a design system doc, and every new chat in that project can reference it without re-uploading.

**Sharing** — Projects can be shared with teammates (on Team and Enterprise plans). When shared, project-only memory activates automatically to maintain privacy between collaborators.

### Memory — Cross-Conversation Personalization

Memory allows ChatGPT to remember information about you across separate conversations. It works in two ways:

**Saved memories** — You explicitly tell ChatGPT to remember something: “Remember that I prefer concise bullet-point summaries” or “Remember that my team uses Next.js with TypeScript.” These are stored and applied to all future conversations.

**Chat history insights** — ChatGPT automatically surfaces relevant context from your past conversations. If you mentioned your job title three weeks ago and now ask for a cover letter, it already knows your background.

You have full control over Memory. Go to Settings > Personalization > Memory > Manage to view, search, and delete any memory. For conversations you don’t want stored, use Temporary Chat mode (toggle at the top of a new conversation).

## Tasks — Scheduled Automations

Tasks let you schedule ChatGPT to do things automatically on a recurring basis. This turns ChatGPT from something you use reactively into something that works proactively in the background.

**What you can schedule** — Any prompt that ChatGPT can execute: “Every Monday at 9am, pull the latest metrics from my dashboard and create a weekly summary.” “Every Friday, generate a motivational quote image for our team Slack.”

**Connected apps** — Tasks can connect to external services like Gmail, Slack, Notion, GitHub, Linear, and 20+ other apps. This means ChatGPT can pull data from these services as part of a scheduled task.

**Limits** — Plus users get 40 scheduled tasks/month. Pro users get unlimited. Tasks combine web browsing, data analysis, and agent capabilities.

### EXAMPLE TASKS

Every Monday 8am: Summarize unread emails from my team and draft replies. Every day at 6pm: Check my calendar for tomorrow and brief me on who I'm meeting. Every Friday: Generate a weekly project status report from our GitHub activity.

## Chapter 12: Codex — Cloud Coding Agent

Codex is OpenAI's cloud-based coding agent, available to ChatGPT Plus and Pro subscribers. Unlike coding assistance in the regular chat (which helps you write code interactively), Codex works asynchronously in the background. You give it a task, and it works on it in a sandboxed cloud environment while you do other things.

### How Codex Works

**Sandboxed environment** — Each task runs in its own isolated cloud sandbox preloaded with your repository. Codex can read files, write code, run commands, execute tests, and lint its own work. It cannot affect your production environment.

**Multi-threaded** — You can assign multiple tasks simultaneously. Kick off a bug fix, a refactor, and a set of unit tests at the same time. Codex works on all of them in parallel.

**GitHub integration** — Codex can create branches and open pull requests directly. Review the code it wrote in a standard PR workflow, just as you would with a human teammate.

**Self-validation** — Codex runs linters and tests to check its own work before presenting results. It doesn't just write code; it verifies that the code works.

### What You Can Ask Codex to Do

1

#### Write features

"Add a password reset flow to the auth module. Include email verification, rate limiting, and tests."

2

#### Fix bugs

"The checkout page throws a 500 error when the cart is empty. Find and fix the bug."

3

#### Refactor

"Refactor the data access layer to use the repository pattern. Maintain backward compatibility."

4

#### Answer codebase questions

"How does the authentication flow work in this project? Trace it from login to session creation."

## Codex App

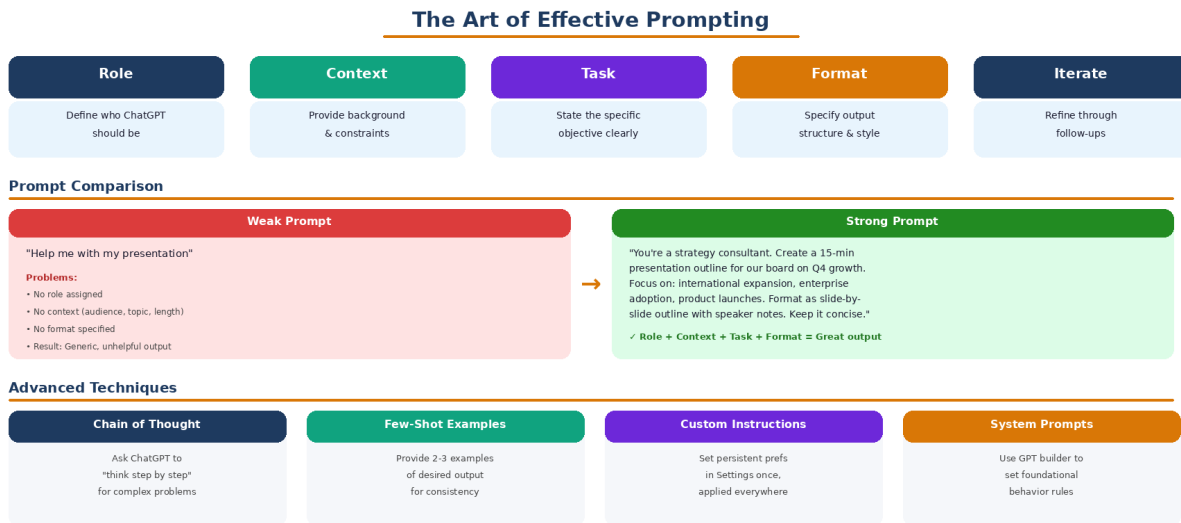
OpenAI released the Codex App for macOS (and Windows as of March 2026), a dedicated desktop interface for managing Codex tasks. It lets you run multiple agents simultaneously, manage parallel workstreams, and collaborate on long-running tasks. The latest model, GPT-5.1-Codex-Max, is specifically optimized for agentic coding with improved speed, intelligence, and token efficiency.

### **TIP: Internet Access**

Since June 2025, Codex can access the internet during task execution. This means it can install packages, check documentation, and pull dependencies as part of building your code.

## Chapter 13: Prompting Like a Power User

The quality of ChatGPT's output is directly proportional to the quality of your input. A well-structured prompt can transform a mediocre response into a brilliant one. This isn't about memorizing tricks; it's about communicating clearly with an AI the same way you'd brief a smart human colleague.



### The Five Pillars of Effective Prompting

#### 1. Assign a Role

Tell ChatGPT who to be. Roles calibrate vocabulary, assumptions, depth, and communication style. "You are a senior tax attorney" produces fundamentally different output from "You are a personal finance blogger." Roles work because they activate specific knowledge patterns within the model.

#### 2. Provide Rich Context

Context is the single most impactful element of a prompt. Tell ChatGPT: Who is the audience? What's the situation? What have you already tried? What constraints exist? "I'm a Series A startup founder preparing a pitch for healthcare-focused VCs. Our product is an AI diagnostic tool for rural clinics." This context shapes every word of the response.

#### 3. Be Specific About the Task

Vague prompts get vague answers. Compare: "help me with my resume" vs. "Rewrite the experience section of my resume for a Senior Engineering Manager role at a FAANG company. Emphasize people leadership, system design decisions, and quantified impact. Use strong action verbs. Keep each bullet under two lines." The second prompt gives ChatGPT everything it needs.

## 4. Specify Output Format

Tell ChatGPT exactly how you want the output structured. “Give me a comparison table with these columns: Feature, Tool A, Tool B, Winner.” “Write this as a numbered list ordered by priority.” “Format as a professional email with subject line.” “Keep it under 200 words.” ChatGPT follows formatting instructions remarkably well.

## 5. Iterate, Don't Start Over

Treat every conversation as a collaboration. Don't accept the first response if it's not perfect. Say “make the tone more casual,” “add quantitative evidence,” “cut the introduction in half,” or “rewrite section 3 to focus on cost savings.” Each iteration builds on your shared context. The conversation usually gets better with each turn.

## Advanced Techniques

### CHAIN OF THOUGHT

For complex reasoning tasks, add “Think step by step” to your prompt. This forces the model to work through the problem systematically rather than jumping to a conclusion. Or use a reasoning model (o3, o4-mini) which does this automatically.

### FEW-SHOT PROMPTING

Provide 2–3 examples of the exact output format you want. “Here's an example of a good product description: [example]. Now write one for my product.” This is one of the most reliable techniques for consistent, high-quality output.

### CUSTOM INSTRUCTIONS

Go to Settings > Personalization > Custom Instructions. Set your preferences once and they're applied to every conversation: your role, preferred format, tone, constraints, and any context ChatGPT should always know about you. This eliminates the need to repeat your context in every new chat.

## Chapter 14: Apps & Connectors

In December 2025, OpenAI renamed “Connectors” to “Apps” to unify third-party integrations under a single concept. Apps let ChatGPT connect to external services and data sources so it can pull in context, take actions, and work with your existing tools directly from the chat.

### How Apps Work

When you enable an app, ChatGPT gains the ability to authenticate with that service and access your data within it. For example, connecting Gmail lets ChatGPT read your emails, draft replies, and search your inbox. Connecting Slack lets it read messages, summarize channels, and send messages. All of this happens through natural language in the chat.

### Available Apps

App	What It Does	Example Use
Gmail	Read, search, draft, and send emails	“Summarize unread emails from this week”
Slack	Read channels, send messages, search history	“What did the engineering team discuss yesterday?”
Google Drive	Access, search, and summarize documents	“Find the Q4 budget spreadsheet and summarize it”
Notion	Read pages, search workspace, access databases	“What’s the status of the product roadmap?”
GitHub	View repos, issues, PRs, code	“Show me open PRs that need review”
Linear	Access issues, projects, cycles	“What’s my team’s sprint progress?”
Zapier	Connect to 5000+ apps via workflows	“When I say ‘log lead,’ add to my CRM”

OpenAI is opening submissions for new app integrations throughout 2026 and plans to introduce monetization options for developers, including the Agentic Commerce Protocol for in-chat purchases.

#### **TIP: Apps + Tasks = Automation**

Combine Apps with scheduled Tasks for powerful automation. “Every Monday morning, check my Gmail for invoices, extract the amounts, and update my expense tracking spreadsheet in Google Drive.” This is where ChatGPT starts functioning as a true personal assistant.

## Chapter 15: Troubleshooting Common Issues

Even experienced ChatGPT users hit friction. Here are the most common problems and their solutions.

Problem	Cause	Solution
Generic, unhelpful responses	Prompt is too vague	Add role + context + format + constraints to your prompt
ChatGPT refuses a request	Content policy triggered	Rephrase with context about your legitimate use case
Hit a message limit	Rate limit reached	Wait for reset or upgrade plan. Use GPT-4o mini for bulk work.
Response cuts off mid-sentence	Output length limit hit	Say “continue” or “finish your previous response”
Outdated information	Training data cutoff	Enable web search. Ask: “Search the web for current info on...”
Code doesn't work	Missing context	Share the full error + all relevant files, not just a snippet
Image has wrong text	Complex text rendering	Simplify the text. Break into shorter phrases. Try again.
Canvas won't open	Browser compatibility	Use Chrome or Edge. Clear cache. Disable conflicting extensions.
Voice mode disconnects	Network instability	Switch to Wi-Fi. Close bandwidth-heavy apps.
Custom GPT performs poorly	Instructions too vague	Rewrite instructions with specific rules, examples, and constraints
Agent Mode stuck on a page	Complex site / CAPTCHA	Guide it manually or try a different approach
Memory wrong/outdated	Stale saved memory	Go to Settings > Memory > Manage and delete incorrect items

### Getting Help

**OpenAI Help Center** — [help.openai.com](https://help.openai.com) for official docs, billing, account management, and troubleshooting.

**Community Forum** — [community.openai.com](https://community.openai.com) for user discussions, tips, use cases, and peer support.

**Status Page** — [status.openai.com](https://status.openai.com) to check for outages or degraded performance.

**OpenAI Blog** — [openai.com/blog](https://openai.com/blog) for feature announcements, model launches, and research papers.

**@OpenAI on X** — Follow for real-time updates on new features, model releases, and event announcements.

## Chapter 16: What's Coming Next

OpenAI ships updates at a pace that's hard to keep up with. Here's what's on the horizon as of March 2026 and what trends to watch.

### Confirmed and In-Progress

**Sora in ChatGPT** — Full Sora video generation integration into the chat interface is rolling out, allowing users to generate HD video clips directly in conversation alongside text and images.

**Expanded Apps ecosystem** — OpenAI is opening third-party app submissions in 2026, dramatically expanding the number of services ChatGPT can connect to. The Agentic Commerce Protocol will enable in-chat purchases.

**GPT-5 and beyond** — The next generation of foundation models promising significant jumps in reasoning, reliability, multimodal understanding, and agentic capability. GPT-5 variants are already available in some contexts.

### Trends to Watch

**Everything becomes agentic** — The line between “chat” and “action” is blurring. Agent Mode, Tasks, and Codex all represent a shift from ChatGPT as an advisor to ChatGPT as a doer. Expect this to accelerate.

**Multimodal by default** — Text, images, voice, video, code, and web actions are converging into a single, fluid interface. You'll increasingly talk to ChatGPT while showing it things and having it take action simultaneously.

**Personalization deepens** — Memory, Projects, and Custom Instructions are the first steps. Future updates will make ChatGPT increasingly aware of your work patterns, preferences, and history, becoming more useful the longer you use it.

**Enterprise expansion** — Expect deeper integrations with business tools, more granular admin controls, compliance certifications, and industry-specific solutions.

### How to Stay Current

1

#### Follow the OpenAI Blog

[openai.com/blog](https://openai.com/blog) for official announcements. This is always the most authoritative source.

2

#### Check the ChatGPT release notes

OpenAI publishes detailed release notes for every update at [help.openai.com](https://help.openai.com).

**3**

**Join communities**

r/ChatGPT on Reddit, the OpenAI Community Forum, and AI-focused YouTube channels are excellent for tips, tutorials, and early looks at new features.

**FINAL THOUGHT**

ChatGPT in 2026 is not the ChatGPT of 2023. It's not a chatbot anymore. It's a platform that can see, hear, speak, browse the web, write and run code, take actions on your behalf, generate images and video, remember your preferences, and automate your recurring workflows. The users who get the most value from it are the ones who integrate it into their daily work: start with one workflow, build Custom GPTs for your recurring tasks, explore Agent Mode, and keep experimenting. The more you use it, the more indispensable it becomes.