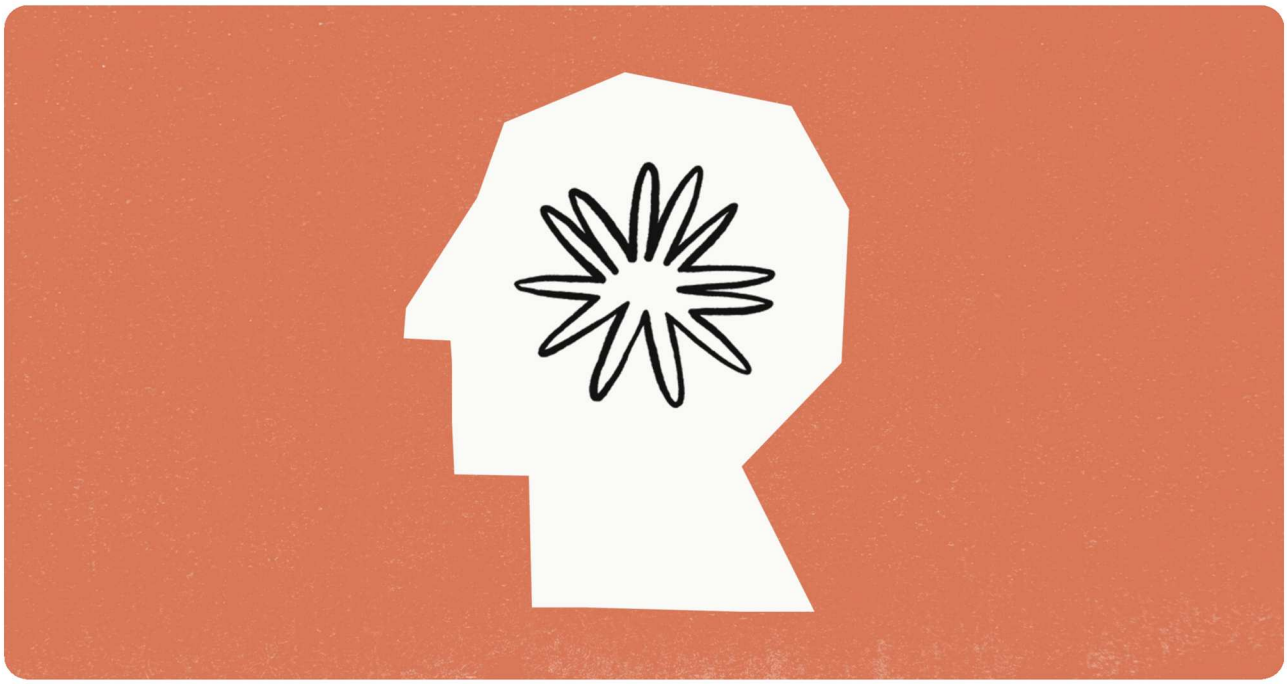
EXHIBIT 174

Announcements

Claude 3.5 Sonnet

Jun 20, 2024

Try on Claude.ai

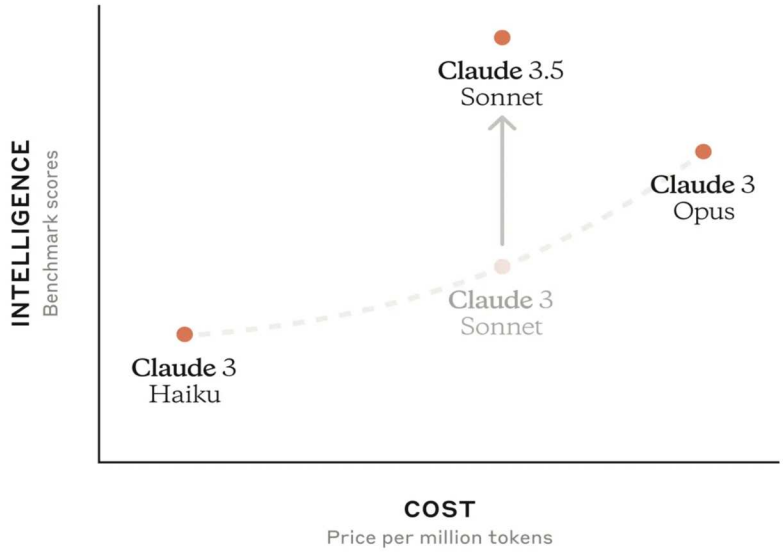


UPDATE Consumer Terms and Privacy Policy Aug 28, 2025

Today, we're launching Claude 3.5 Sonnet—our first release in the forthcoming Claude 3.5 model family. Claude 3.5 Sonnet raises the industry bar for intelligence, outperforming competitor models and Claude 3 Opus on a wide range of evaluations, with the speed and cost of our mid-tier model, Claude 3 Sonnet.

Claude 3.5 Sonnet is now available for free on Claude.ai and the Claude iOS app,

while Claude Pro and Team plan subscribers can access it with significantly higher rate limits. It is also available via the Anthropic API, Amazon Bedrock, and Google Cloud's Vertex AI. The model costs $3 per million input tokens and $15 per million output tokens, with a 200K token context window.



Frontier intelligence at 2x the speed

Claude 3.5 Sonnet sets new industry benchmarks for graduate-level reasoning (GPQA), undergraduate-level knowledge (MMLU), and coding proficiency (HumanEval). It shows marked improvement in grasping nuance, humor, and complex instructions, and is exceptional at writing high-quality content with a natural, relatable tone.

Claude 3.5 Sonnet operates at twice the speed of Claude 3 Opus. This performance boost, combined with cost-effective pricing, makes Claude 3.5 Sonnet ideal for complex tasks such as context-sensitive customer support and orchestrating multi-step workflows.

In an internal agentic coding evaluation, Claude 3.5 Sonnet solved 64% of problems, outperforming Claude 3 Opus which solved 38%. Our evaluation tests the model's ability to fix a bug or add functionality to an open source codebase, given a natural language description of the desired improvement. When instructed

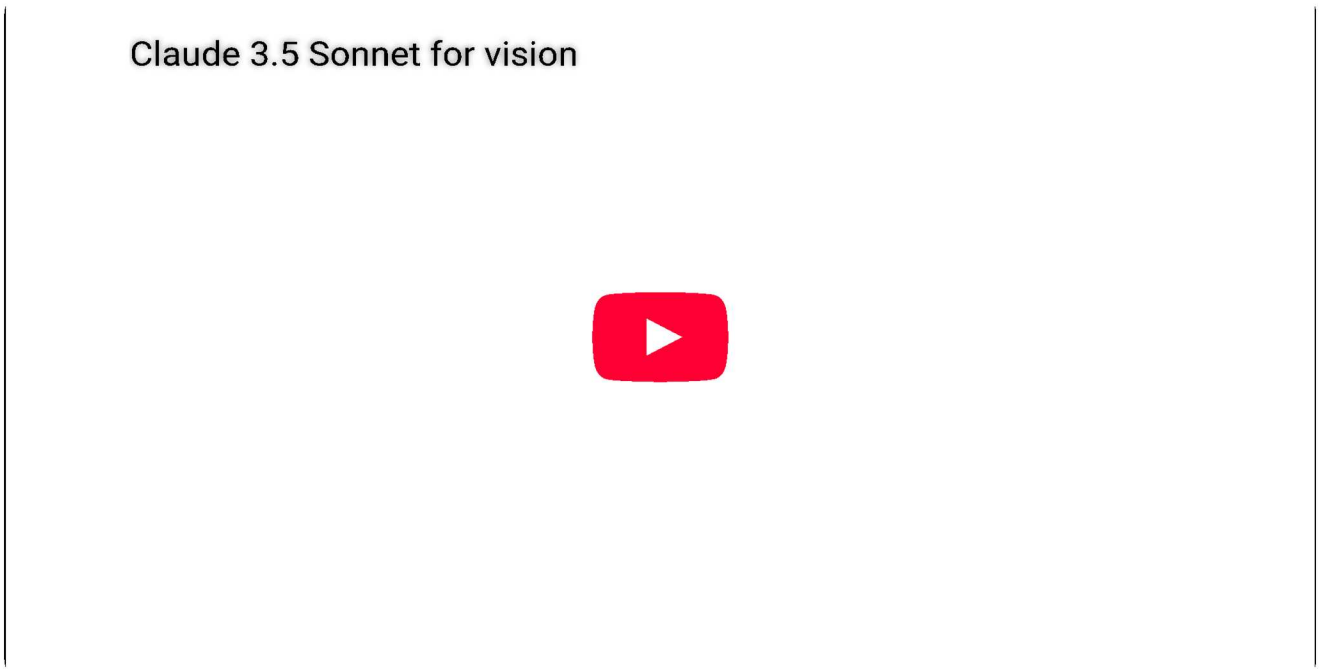
and provided with the relevant tools, Claude 3.5 Sonnet can independently write, edit, and execute code with sophisticated reasoning and troubleshooting capabilities. It handles code translations with ease, making it particularly effective for updating legacy applications and migrating codebases.

	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro	Llama-400b (early snapshot)
Graduate level reasoning *GPQA, Diamond*	59.4%* 0-shot CoT	50.4% 0-shot CoT	53.6% 0-shot CoT	—	—
Undergraduate level knowledge *MMLU*	88.7%** 5-shot 88.3% 0-shot CoT	86.8% 5-shot 85.7% 0-shot CoT	— 88.7% 0-shot CoT	85.9% 5-shot —	86.1% 5-shot —
Code *HumanEval*	92.0% 0-shot	84.9% 0-shot	90.2% 0-shot	84.1% 0-shot	84.1% 0-shot
Multilingual math *MGSM*	91.6% 0-shot CoT	90.7% 0-shot CoT	90.5% 0-shot CoT	87.5% 8-shot	—
Reasoning over text *DROP, F1 score*	87.1 3-shot	83.1 3-shot	83.4 3-shot	74.9 Variable shots	83.5 3-shot Pre-trained model
Mixed evaluations *BIG-Bench-Hard*	93.1% 3-shot CoT	86.8% 3-shot CoT	—	89.2% 3-shot CoT	85.3% 3-shot CoT Pre-trained model
Math problem-solving *MATH*	71.1% 0-shot CoT	60.1% 0-shot CoT	76.6% 0-shot CoT	67.7% 4-shot	57.8% 4-shot CoT
Grade school math *GSM8K*	96.4% 0-shot CoT	95.0% 0-shot CoT	—	90.8% 11-shot	94.1% 8-shot CoT

* Claude 3.5 Sonnet scores 67.2% on 5-shot CoT GPQA with maj@32
** Claude 3.5 Sonnet scores 90.4% on MMLU with 5-shot CoT prompting

State-of-the-art vision

Claude 3.5 Sonnet is our strongest vision model yet, surpassing Claude 3 Opus on standard vision benchmarks. These step-change improvements are most noticeable for tasks that require visual reasoning, like interpreting charts and graphs. Claude 3.5 Sonnet can also accurately transcribe text from imperfect images—a core capability for retail, logistics, and financial services, where AI may glean more insights from an image, graphic or illustration than from text alone.



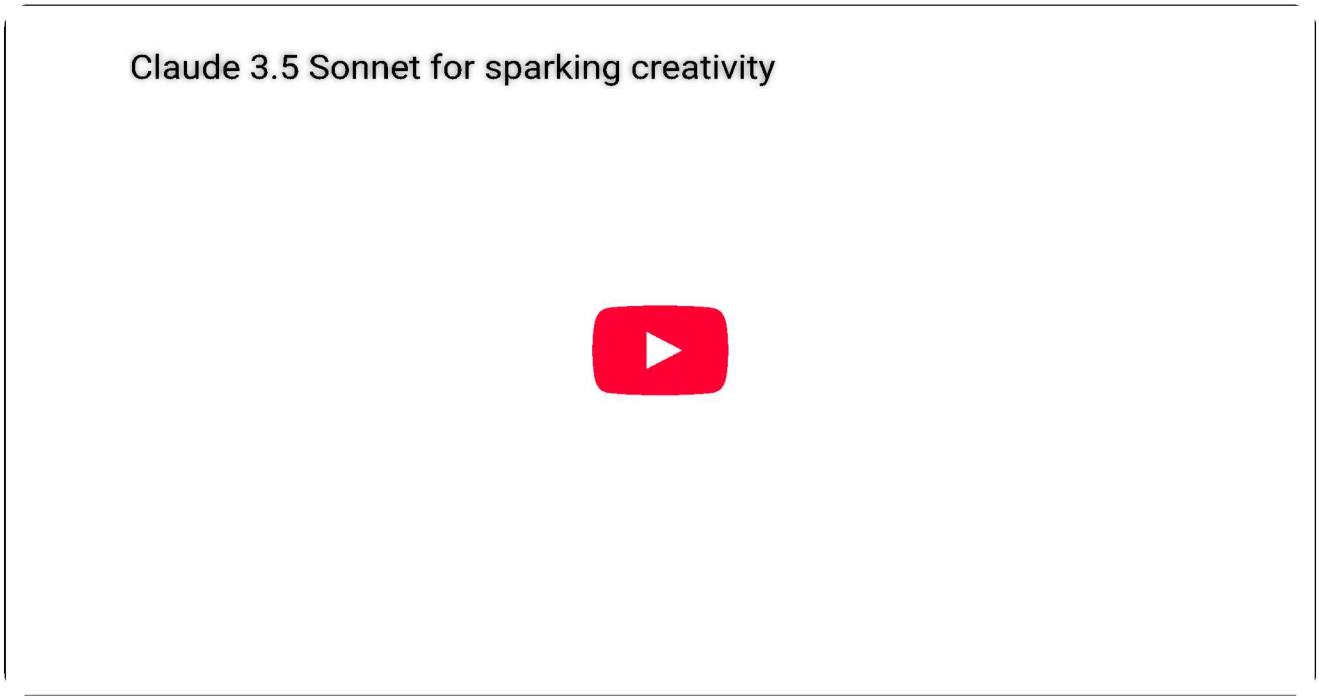
	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro
Visual math reasoning *MathVista (testmini)*	**67.7%** 0-shot CoT	**50.5%** 0-shot CoT	**63.8%** 0-shot CoT	**63.9%** 0-shot CoT
Science diagrams *AI2D, test*	**94.7%** 0-shot	**88.1%** 0-shot	**94.2%** 0-shot	**94.4%** 0-shot
Visual question answering *MMMU (val)*	**68.3%** 0-shot CoT	**59.4%** 0-shot CoT	**69.1%** 0-shot CoT	**62.2%** 0-shot CoT
Chart Q&A *Relaxed accuracy (test)*	**90.8%** 0-shot CoT	**80.8%** 0-shot CoT	**85.7%** 0-shot CoT	**87.2%** 0-shot CoT
Document visual Q&A *ANLS score, test*	**95.2%** 0-shot	**89.3%** 0-shot	**92.8%** 0-shot	**93.1%** 0-shot

Artifacts—a new way to use Claude

Today, we're also introducing Artifacts on Claude.ai, a new feature that expands how users can interact with Claude. When a user asks Claude to generate content like code snippets, text documents, or website designs, these Artifacts appear in a dedicated window alongside their conversation. This creates a dynamic workspace where they can see, edit, and build upon Claude's creations in real-time, seamlessly

integrating AI-generated content into their projects and workflows.

This preview feature marks Claude's evolution from a conversational AI to a collaborative work environment. It's just the beginning of a broader vision for Claude.ai, which will soon expand to support team collaboration. In the near future, teams—and eventually entire organizations—will be able to securely centralize their knowledge, documents, and ongoing work in one shared space, with Claude serving as an on-demand teammate.



Commitment to safety and privacy

Our models are subjected to rigorous testing and have been trained to reduce misuse. Despite Claude 3.5 Sonnet's leap in intelligence, our red teaming assessments have concluded that Claude 3.5 Sonnet remains at ASL-2. More details can be found in the model card addendum.

As part of our commitment to safety and transparency, we've engaged with external experts to test and refine the safety mechanisms within this latest model. We recently provided Claude 3.5 Sonnet to the UK's Artificial Intelligence Safety Institute (UK AISI) for pre-deployment safety evaluation. The UK AISI completed tests of 3.5 Sonnet and shared their results with the US AI Safety Institute (US

AISI) as part of a Memorandum of Understanding, made possible by the partnership between the US and UK AISIs announced earlier this year.

We have integrated policy feedback from outside subject matter experts to ensure that our evaluations are robust and take into account new trends in abuse. This engagement has helped our teams scale up our ability to evaluate 3.5 Sonnet against various types of misuse. For example, we used feedback from child safety experts at Thorn to update our classifiers and fine-tune our models.

One of the core constitutional principles that guides our AI model development is privacy. We do not train our generative models on user-submitted data unless a user gives us explicit permission to do so.

Coming soon

Our aim is to substantially improve the tradeoff curve between intelligence, speed, and cost every few months. To complete the Claude 3.5 model family, we'll be releasing Claude 3.5 Haiku and Claude 3.5 Opus later this year.

In addition to working on our next-generation model family, we are developing new modalities and features to support more use cases for businesses, including integrations with enterprise applications. Our team is also exploring features like Memory, which will enable Claude to remember a user's preferences and interaction history as specified, making their experience even more personalized and efficient.

We're constantly working to improve Claude and love hearing from our users. You can submit feedback on Claude 3.5 Sonnet directly in-product to inform our development roadmap and help our teams to improve your experience. As always, we look forward to seeing what you build, create, and discover with Claude.

Related content

Donating the Model Context Protocol and establishing the Agentic AI Foundation

Read more →

Accenture and Anthropic launch multi-year partnership to move enterprises from AI pilots to production

Read more →

Snowflake and Anthropic announce $200 million partnership to bring agentic AI to global enterprises

Read more →

Products

Claude

Claude Code

Claude and Slack

Claude in Excel

Skills

Max plan

Team plan

Enterprise plan

Download app

Pricing

Log in to Claude

Models

Opus

Sonnet

Haiku

Solutions

AI agents

Code modernization

Coding

Customer support

Education

Financial services

Government

Life sciences

Nonprofits

Claude Developer Platform

Overview

Developer docs

Pricing

Amazon Bedrock

Google Cloud's Vertex AI

Console login

Learn

Blog

Courses

Use cases

Connectors

Customer stories

Engineering at Anthropic

Events

Powered by Claude

Service partners

Startups program

Company

Anthropic

Careers

Economic Futures

Research

News

Responsible Scaling Policy

Security and compliance

Transparency

Help and security

Availability

Status

Support center

Terms and policies

Privacy choices

Privacy policy

Responsible disclosure policy

Terms of service: Commercial

Terms of service: Consumer

Usage policy

© 2025 Anthropic PBC

