

— 2026-06-11 最新整理

Claude Fable 5 新手完全攻略

从模型介绍、模型对比，到上手教程、提示词技巧和避坑建议，一份给小白看的简洁演示文稿。

1M

默认上下文窗口

128k

单次最大输出

\$10/\$50

每百万输入/输出
token



Claude Fable 5

官方发布视觉素材，来源：Anthropic。

— 01 / 模型定位

它不是“更会聊天”，
而是更会“独立干大活”。

Fable 5 面向复杂推理、长时间 Agent 工作、代码工程、文档分析、视觉理解和企业办公任务。

把它想成：一个能先看现场、再拆任务、再执行、最后自查的高级 AI 助手。

长任务

复杂代码

看图分析

办公文档

Fable 5 是 Mythos 级能力的公开版本。

04

Mythos Preview

Anthropic 先把更强能力放在 Project Glasswing，主要给少数可信用户试用。

06

Fable 5 发布

2026-06-09，Fable 5 面向公众发布，API 和主要云平台可用。



加上安全限制

遇到高风险网络安全、生物等请求，可能拒绝或转到 Opus 4.8。



透明化争议

发布后社区关注“被降级/转接是否透明”，Anthropic 表示会改进提示和处理方式。

先记住这 6 个数字和规则。

API ID

claude-fable-5

开发者调用模型时使用这个 ID。

上下文

1M tokens

能读很长的材料，适合大文档、大项目。

最大输出

128k tokens

一次能输出很长结果，但要控制成本。

价格

\$10 / \$50

每百万输入 / 输出 token，比 Opus 4.8 更贵。

思考方式

Adaptive thinking

自适应思考常开，不能像旧模型一样关闭。

数据保留

30 天

官方文档说明：不支持 Zero Data Retention。

不是所有任务都要用最贵的模型。

模型	适合做什么	优点	注意点	新手建议
Fable 5	复杂代码、长文档、长期 Agent、困难分析	最强公开 Claude，长任务更稳	贵，慢，部分领域会触发安全限制	只给难题用
Mythos 5	可信机构的高敏感专业任务	同等能力，限制更少	非公开自助模型	普通人不用考虑
Opus 4.8	高质量推理、代码、专业任务	能力强，成本低于 Fable 5	长任务能力弱于 Fable 5	多数高级任务够用
Sonnet 4.6	日常写作、普通代码、客服、总结	速度和价格更平衡	超复杂任务不如 Fable	默认首选

官方 benchmark 显示： 强项集中在“长任务”和 “工程类任务”。

重点看 SWE-Bench Pro、FrontierCode Diamond、Terminal-Bench、工具使用和知识工作。它不是每个场景都便宜，但在难题上更可能一次做对。

注：benchmark 只能说明测试表现，真实效果还要用自己的任务验证。

	Claude Mythos 5 / Fable 5	Claude Mythos Preview	Claude Opus 4.8	GPT 5.5	Gemini 3.1 Pro
Agentic coding SWE-Bench Pro	80.3%	77.8%	69.2%	58.6%	54.2%
Agentic coding FrontierCode (Diamond)	29.3% xhigh	—	13.4% xhigh	5.7% xhigh	—
Knowledge work GDPval-AA	1932	—	1890	1769	1314
Knowledge work vision GDPpdf	29.8% no tools	—	22.5% no tools	24.9% no tools	16.7% no tools
Spatial reasoning Blueprint-Bench 2	38.6%	—	14.5%	36.2%	26.5%
Tool use AutomationBench	17.4%	—	15.5%	12.9%	9.6%
Computer use OSWorld-Vorthead	85.0%	85.4%	83.4%	78.7%	76.2%
Legal Legal Agent Benchmark	13.3%	—	10.4%	2.1%	0.0%
Multidisciplinary reasoning Humanity's Last Exam	59.0%* no tools	56.8% no tools	49.8% no tools	41.4% no tools	44.4% no tools
	64.5%* with tools	64.7% with tools	57.9% with tools	52.2% with tools	51.4% with tools
Biology BioMysteryBench	46.1%* hard	29.6% hard	40.0% hard	—	—
	83.9%* human solved	82.6% human solved	80.4% human solved	—	—
Agentic coding Terminal-Bench 2.1	88.0%*	—	82.7%	83.4% Codex CLI	70.7% Gemini CLI
Cybersecurity ExploitBench (Cap%)	78.0%*	69.0%	40.0%	34.0%	—
Health HealthBench Professional	66.0%*	64.7%	56.9%	51.8%	—

Methodology: Reported scores are within a 1-3 percentage point difference for Claude Mythos 5 and Claude Fable 5. This table shows the higher score of the two. Starred (*) benchmarks show a larger difference due to our blocking safeguards for cybersecurity and biology-related questions. For these benchmarks, Claude Fable 5 performs closer to Claude Opus 4.8 due to fallbacks. See the system card for details.

官方发布页 benchmark 图。本演示稿保留原图，用于学习和对比。

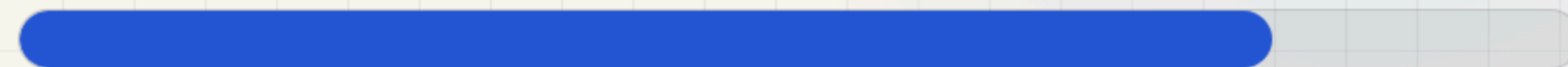
— 06 / 新手看跑分

不用看懂全部榜单，先看这 4 项。

它们对应小白最容易理解的能力：修代码、做复杂项目、处理知识工作、操作电脑环境。

Agentic Coding: SWE-Bench Pro

80.3%



FrontierCode Diamond

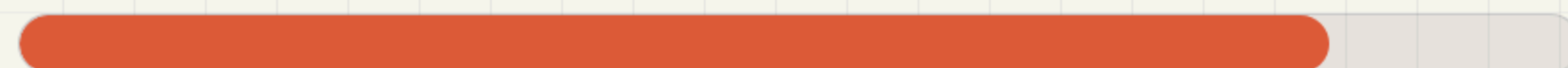
29.3%



这个测试更难，所以百分比低，但相对 Opus 4.8 提升明显。

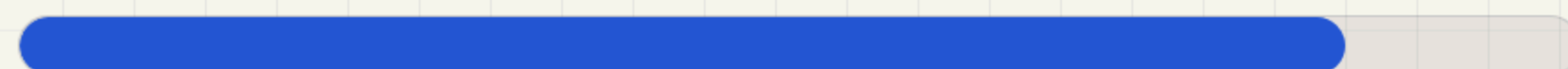
GDPval-AA 知识工作

1932



OSWorld 电脑操作

85.0%



Fable 5 最适合“便宜模型容易反复改”的任务。

值得上 Fable 5

大型代码库改造、疑难 bug、复杂需求实现
几百页资料的研究、报告、合同或财务分析
需要看截图、网页、表格并给出判断
要它持续推进、拆分任务、自我检查

不建议上 Fable 5

改一句文案、写普通邮件、简单摘要
批量低价值内容生成，成本容易失控
需要 Zero Data Retention 的敏感业务
生物、网络安全等容易触发安全限制的请求

小白可以按 4 个入口理解。

1

Claude.ai

像聊天一样使用。适合写作、阅读、方案、截图分析。能选模型时选择 Fable 5。

2

Claude Code

给开发项目用。适合修 bug、改功能、跑测试、整理代码。

3

Claude API

给软件接入。开发者用 model ID 调用，处理拒绝、超时、成本。

4

云平台

AWS Bedrock、Vertex AI、Microsoft Foundry 等企业入口，适合公司管理。

— 09 / API 教程

开发者最小调用示例。

如果你不写代码，只需要记住：API 是“把 Fable 5 接进自己的软件”的方式。

真实项目要补上错误处理、超时、费用记录和拒绝处理。

```
import Anthropic from "@anthropic-ai/sdk";

const client = new Anthropic({
  apiKey: process.env.ANTHROPIC_API_KEY
});

const msg = await client.messages.create({
  model: "claude-fable-5",
  max_tokens: 1200,
  effort: "high",
  messages: [{
    role: "user",
    content: "帮我检查这份产品方案的风险，并给出修改建议。"
  }]
});

console.log(msg.content);
```

— 10 / 小白提示词

Fable 5 更吃“背景 + 目标 + 验收标准”。

不要只说“帮我做一下”。告诉它为什么做、给谁看、怎么算完成，它会更像人在做项目。

可直接套用

Prompt

我正在做【更大的目标】，对象是【谁】。
他们需要【最终结果有什么用】。

请你完成：【具体任务】。

要求：

1. 先判断你还缺什么信息。
2. 能直接行动时就行动，不要空泛分析。
3. 每完成一个阶段，自己检查一遍。
4. 最后用小白能看懂的话总结结果。
5. 不要添加和任务无关的功能或内容。

effort 越高，越聪明，但越慢也可能越贵。

low

快一点

适合普通改写、简单分类、低风险小任务。

medium

日常高级任务

适合一般代码、文档分析、方案梳理。

high

推荐默认

官方建议多数任务从 high 开始，再根据速度和成本调整。

xhigh

最难任务

用于复杂工程、关键分析、需要极高可靠性的工作。

提醒

要设边界

明确“不做额外重构、不加新功能、不发散”，防止它做太多。

成本

先小样本试

先让它做一小段，满意后再扩大到完整任务。

Fable 5 可能会“拒绝”或“转给 Opus 4.8”。

这不是报错。官方文档说，安全分类器拦截时，API 可能返回 `stop_reason: refusal`，并可设置 `fallback`。

会触发的方向

高风险内容

如攻击性网络安全、生物实验细节、模型思考提取等。正常请求也可能误触发。

怎么处理

准备备用模型

在 API 里处理 `refusal`，把可回答的任务转给 Opus 4.8 或其它 Claude 模型。

费用

早期拒绝不收费

官方说明：输出前被拒绝的请求不计费；中途拒绝会计已生成部分。

小白理解

像安全闸门

不是模型不会，而是公开版被加了更谨慎的安全门槛。

别把 Fable 5 当默认模型，把它当“专家会诊”。

先筛选

简单任务先用 Sonnet

摘要、改写、普通问答、简单代码，先用更便宜模型。

再升级

卡住再上 Fable

当便宜模型反复失败、结果需要大量人工修，才切到 Fable 5。

控范围

限制输入和输出

只给必要资料，要求先给计划，确认后再生成完整结果。

能力强，不代表可以无脑信。

坑 1

太贵

长文档和长输出会快速消耗 token。先小样本试，再扩大。

坑 2

太慢

高 effort 可能跑几分钟。产品里要做进度提示和超时处理。

坑 3

会误触发拒绝

生物、网络安全、化学等边界更敏感，业务流程要有备用方案。

坑 4

会做太多

长任务里要写清楚“不要额外重构、不要扩展范围”。

坑 5

仍要验收

代码要跑测试；数据要查来源；法律医疗财务不能当最终建议。

坑 6

隐私要求

官方说明 Fable 5 有 30 天数据保留，不适合所有敏感数据场景。

5 天把 Fable 5 用明白。

Day 1 问答与总结

让它总结 3 篇资料，并要求“用小白能懂的话解释”。

Day 2 对比与决策

让它比较 3 个方案，输出“推荐、理由、风险”。

Day 3 看图与表格

上传截图、数据表，让它找问题并给改进建议。

Day 4 代码任务

给一个真实 bug，让它先定位，再修改，再跑测试。

Day 5 长任务 Agent

给一个完整目标，让它拆计划、执行、检查、交付总结。

Fable 5 的正确用法：把难题交给它， 把边界写清楚，把结果认真验收。

它最像“高级项目助手”，不是免费万能工具。用在高价值、复杂、长周期任务上，才最容易值回成本。

难题才用

明确目标

控制成本

保留验收

资料来源与核对说明

Anthropic 发布公告

Claude Fable 5 与 Mythos 5 的官方发布信息、benchmark 图和安全说明。

Claude API Docs

模型 ID、价格、1M 上下文、128k 输出、拒绝和 fallback。

Prompting Fable 5

官方提示词、effort、长任务和沟通风格建议。

Simon Willison 体验记录

独立开发者对速度、成本、能力和限制的早期观察。

The Verge 报道

社区关注的生物相关请求拒绝问题。

Anthropic Fable 页面

模型定位、客户反馈、官方视觉素材。

模型总览

Fable 5、Mythos 5、Opus、Sonnet、Haiku 的官方对比。

AWS Bedrock Model Card

采样参数、知识截止、拒绝处理、区域与模型 ID。

Anthropic Webinar

如何把 Fable 5 用在自主 Agent 和 Claude Code workflows 中。

Business Insider 报道

安全策略透明度争议和后续调整。