AI 能干什么、不能干什么

资料编号：L1-1 ｜ 免费资料包 · 认知地基 ｜ 来来自媒体工作室出品

适用：刚接触 AI 的人，先把边界搞清楚再用

先讲个真事

我做过一个 AI 客服样板，测试时问它"你们能代开发票吗"，它一本正经答了一长串，里面 **"发票可以代开"这句话是我知识库里根本没写过的**——它自己编的。

这不是它坏，是它的工作原理决定的：AI 不是查资料，是**猜下一个词最像什么**。没学过的东西，它不会说"我不知道"，而是给你编一个"最像答案"的东西。行话叫**幻觉**。

所以用 AI 第一条铁律：**它说的是"像不像"，不是"对不对"。** 搞清它擅长什么、不擅长什么，你才不会踩坑。

AI 擅长的 5 件事（放心用）

事	大白话举例	靠谱程度
写初稿	一篇文章、一段文案、一条朋友圈，让它先起个头	★★★★★ 拿它当草稿机，改着用
总结归纳	长文章、长聊天记录，让它提炼要点	★★★★★ 它最拿手
改写润色	大白话改正式、口语改书面、语气改客气	★★★★☆ 比我快，比我工整
出点子	想标题、想选题、想卖点，一次给 10 个你挑	★★★★☆ 数量管够，质量你筛
做初稿素材	出图、出视频脚本、写代码雏形	★★★★☆ 初稿够用，精修靠自己

共同点：AI 是"第一稿"，不是"最终稿"。 凡是"先有个东西出来，我再改"的活，都放心交给它——这是它最省你时间的地方。

AI 干不了的 4 件事（别指望它）

事	为什么不行	怎么办
替你做决定	它不知道你的处境、你的钱、你输不输得起	它给选项和利弊，**拍板永远是你**
保证信息真实	它会一本正经编（幻觉），尤其数字、书名、法条、人名	关键信息**自己核一遍**再信

事	为什么不行	怎么办
碰钱和隐私的事	它不懂你的风险，出了事它不担责	密码、验证码、付款，**永远自己来**
当真人陪伴	它再贴心也是程序，不承担感情责任	重要的人和事，找真人

一句话记住：**AI 是你的实习生——干活麻利、偶尔闯祸、不能背锅。** 你是那个审核把关的老板。

幻觉是什么、为什么会瞎编

- - **什么是幻觉**：AI 说出一段听着很真、其实是编的内容。你问它一本书的作者，它可能编一个"听起来很像"的作者名。
- - **为什么会有**：它的训练方式是学"文字接龙"，目标是"接得顺"，不是"说得对"。没有的知识，它不会留白，会填一个最顺的。
- - **哪些领域最爱编**：具体数字、冷门书名/法条/人名、没公开的小众信息。越冷门，越爱编。
- - **多严重**：研究里大模型在有把握的问题上也可能错 10% 上下；冷门问题错得更多。**所以别拿它当百科全书的唯一来源。**

防幻觉三招（我的保命招）

第一招：提示词里写清楚"不知道就说不知道"

> 例：……不知道或不确定的，直接说"这个我不确定"，不要编。

就这一句，能把很多瞎编挡回去。

第二招：要它干活时，喂真资料

把你自己整理的真东西（文档、笔记）放进它的知识库，告诉它"只准根据我给的资料回答"。资料库分类清楚点，别一锅烩——我吃过亏：一股脑全塞进去，它答非所问。

第三招：关键信息自己核

涉及数字、价格、法规、别人名字的，花 30 秒自己查一下。**它说的每一句，都默认"待核实"。**

> 这套叫"防幻觉双保险"，我后来做客服、知识库类的 AI 应用时全给装上了，好用。详见教程 04 的实战文档。

一条底线

别把密码交给 AI。 不管它多像人、多贴心，任何要你账号密码、验证码、付款确认的环节，都自己来。这不是信不信任的问题，是规矩。

记住三句话

1. AI 擅长"先有个东西"，不擅长"保证这是对的"
2. 拍板、核真、花钱，这三件事永远自己来
3. 把"不知道就说不知道"写进你的每次提问里

> 边界搞清了，下一步学怎么把它问明白——详见资料包另一份《怎么跟 AI 说话：提示词四要素》。
>
> 完整版资料 + 每周新课 + 一群同路人，都在共学群里。
>
> 进群方式：从来来任意一篇内容找到入口，或直接私信问。欢迎转发；请勿商用。

完整版 + 每周更新，进共学群领取

加「AI 小白·进阶学习交流群」→ 群公告下载全套资料

欢迎转发给需要的朋友，请勿商用
