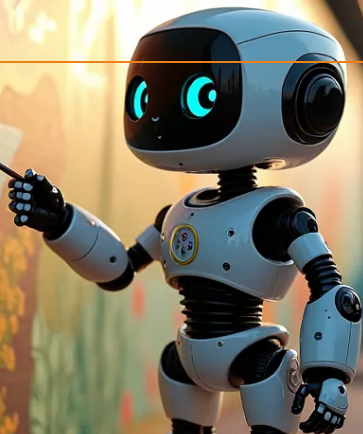


# AI 可以为我们做什么？

---

暂坐读行 & Xing Shi Cai

2024-11-02



😨 AI 会威胁人类吗?

🤖 什么是 AI?

🤔 该怎样看 AI?

🤖 AI 可以做什么?

医疗健康

✎ 写作助手

🔍 信息获取

😓 学习知识

💻 编程助手

📈 数据分析

🖼️ 图像识别

🎨 艺术创作

🗣️ 语音处理

💡 AI 学习指南

 AI 会威胁人类吗?

---

 AI 会威胁人类吗?

---

 什么是 AI?

# 什么是 AI?

AI 是人工智能(Artificial Intelligence)的缩写。广义来说, AI 包含了许多不同的科技 —

-  机器学习 (Machine Learning): **Netflix** 电影推荐系统。
-  自然语言处理 (NLP): **Siri** 和 **小爱同学** 这样的语音助手。
-  计算机视觉 (Computer Vision): AI 帮助相机识别面部、物体和场景。
-  自动驾驶 (Autonomous Driving): **百度** 的自动驾驶汽车。
-  机器人技术 (Robotics): **Tesla** 工厂流水线上的机器人。
-  医疗影像分析 (Medical Imaging): AI 用于识别 X 光片中的疾病迹象。
-  金融预测 (Financial Forecasting): **银行** 用 AI 预测市场趋势。

今天我们讨论的 AI 只限于最近两年快速发展的 两种 Generative Model (生成模型).

今天我们讨论的 AI 只限于最近两年快速发展的 两种 **Generative Model (生成模型)**.

第一种是 **LLM (大语言模型)** 是一种能够理解并生成类似人类的自然语言文本的 AI 系统。

给出一个提示或部分文本，LLM 可以根据上下文生成符合语境且连贯的内容。

今天我们讨论的 AI 只限于最近两年快速发展的 两种 **Generative Model (生成模型)**.

第一种是 **LLM (大语言模型)** 是一种能够理解并生成类似人类的自然语言文本的 AI 系统。

给出一个提示或部分文本，LLM 可以根据上下文生成符合语境且连贯的内容。

## 😊 让 LLM 完成一个笑话

输入 — “为什么数学书总是很难过？”

LLM 输出 — \_\_\_\_\_



今天我们讨论的 AI 只限于最近两年快速发展的 两种 **Generative Model (生成模型)**.

第一种是 **LLM (大语言模型)** 是一种能够理解并生成类似人类的自然语言文本的 AI 系统。

给出一个提示或部分文本，LLM 可以根据上下文生成符合语境且连贯的内容。

## 😊 让 LLM 完成一个笑话

输入 — “为什么数学书总是很难过？”

LLM 输出 — “因为它的问题太多了！”

今天我们讨论的 AI 只限于最近两年快速发展的 两种 Generative Model (生成模型).

第二种是 Diffusion Model (扩散模型) 是一种能够根据提示来生成图像的 AI 系统。

“扩散” 就是慢慢加上噪声 然后再把噪声去掉, 这样一步步地生成图像。



图 1: 生成一张热狗的照片— DALL-E 3

今天我们讨论的 AI 只限于最近两年快速发展的 两种 Generative Model (生成模型).

第二种是 Diffusion Model (扩散模型) 是一种能够根据提示来生成图像的 AI 系统。

“扩散” 就是慢慢加上噪声 然后再把噪声去掉, 这样一步步地生成图像。



图 1: 再热一些— DALL-E 3

今天我们讨论的 AI 只限于最近两年快速发展的 两种 Generative Model (生成模型).

第二种是 Diffusion Model (扩散模型) 是一种能够根据提示来生成图像的 AI 系统。

“扩散” 就是慢慢加上噪声 然后再把噪声去掉, 这样一步步地生成图像。



图 1: 再热一些— DALL-E 3

今天我们讨论的 AI 只限于最近两年快速发展的 两种 Generative Model (生成模型).

第二种是 Diffusion Model (扩散模型) 是一种能够根据提示来生成图像的 AI 系统。

“扩散” 就是慢慢加上噪声 然后再把噪声去掉, 这样一步步地生成图像。



图 1: 再热一些— DALL-E 3

目前在世界范围内，LLM 四巨头是如下几个公司 —

- OpenAI — ChatGPT
- Anthropic — Claude
- Google DeepMind — Gemini
- Meta (Facebook) — LLaMA, 开源为主导

目前在世界范围内，LLM 四巨头是如下几个公司 —

- OpenAI — ChatGPT
- Anthropic — Claude
- Google DeepMind — Gemini
- Meta (Facebook) — LLaMA, 开源为主导

还有一些其他的重要参与者 —

- Grok — xAI
- Mistral AI — Mistral
- 百度 (Baidu) — Ernie
- 阿里巴巴 (Alibaba) — Tongyi Qianwen
- 腾讯 (Tencent) — Hunyuan

 AI 会威胁人类吗?

---

 该怎样看 AI ?



## 🤔 对 AI 的几种不同态度

对 AI 的看法,大致以下几种 —

☹️ 和我有什么关系?

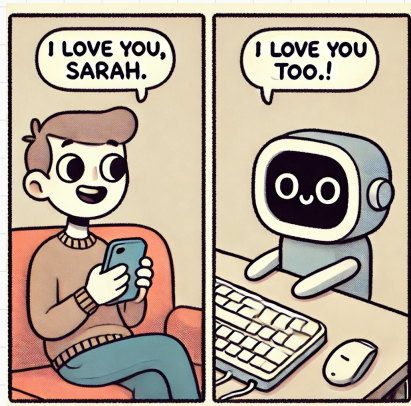


图 2: 你也许没有注意到,但 AI 已经融入你的日常生活。

## 🤖 对 AI 的几种不同态度

对 AI 的看法,大致以下几种 —

🤔 和我有什么关系?

😄 AI 真好!

- AI 能大幅提升生产力
- AI 解放人类劳动
- AI 定制化的艺术创作
- AI 推动科学进步
- AI 解决环境问题

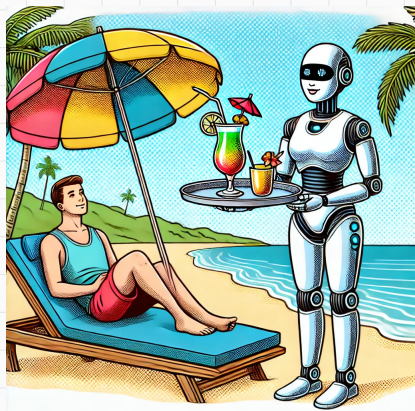


图 2: 我以为 AI 会让我做一点更有趣的工作

## 🤖 对 AI 的几种不同态度

对 AI 的看法,大致以下几种 —

😐 和我有什么关系?

😍 AI 真好!

😡 AI 很讨厌!

- AI 很笨,完全没有实际用途
- AI 只是科技公司的炒作手段
- AI 侵犯版权问题
- AI 可能导致虚假信息泛滥
- AI 可能引发大规模失业
- AI 急剧增加自然资源消耗



图 2: AI 如何帮助我们实现未来的美好愿景?

## 🤖 对 AI 的几种不同态度

对 AI 的看法,大致以下几种 —

🤔 和我有什么关系?

😍 AI 真好!

😡 AI 很讨厌!

😱 AI 好可怕!

- AI 侵蚀数据隐私
- AI 有失控的风险
- AI 威胁人类存在



图 2: AI 会威胁到人类的未来吗?

## 🤖 对 AI 的几种不同态度

对 AI 的看法,大致以下几种 —

🤔 和我有什么关系?

😍 AI 真好!

😡 AI 很讨厌!

😱 AI 好可怕!

😄 其实没有人知道未来!



图 2: 大家其实都在乱猜!

# AI 可以做什么？

---

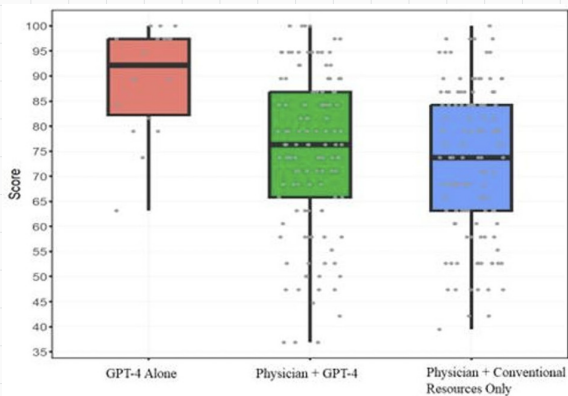
# AI 可以做什么？

---

医疗健康

在最近的一项研究发现 LLM 在诊断疾病时,准确度远超人类医生。

😲 其表现甚至超过了人类医生和 AI 合作。





 AI 可以做什么？

---

 写作助手

接下来我们会做几个小游戏。  
您可以扫描二维码参与



也可以直接访问

[wooclap.com](https://wooclap.com)

然后输入代码

**DISHZR**



下面两首诗,一首是 AI 写的,一首是唐寅写的。

🗨️ 您能猜到哪一首是 AI 写的吗?

### 《秋夜》

深庭明月照华亭,  
几片黄花落远汀。  
凭阑遥望千山静,  
清笛一声度翠屏。

### 《菊花》

故园三径吐幽丛,  
一夜玄霜坠碧空。  
多少天涯未归客,  
尽借篱落看秋风。



下面两个文本,一个是 AI 写的,另外一个小说《我的鸵鸟先生》的开头。您能否猜出哪一段是 AI 写的呢?

### 文本 1

记忆里也有这样一个雪天,那是 1995 年的冬天,雪下得特别特别大。电视新闻里说 E 市碰到了三十年来最大的一场雪,提醒居民们要注意安全,小心出行。


### 文本 2

那是 2010 年开春,我刚从南方调到 H 市工作。三月的北方还带着寒意,早上上班要穿厚外套,中午太阳出来又觉得多余。记得那天新闻说最近天气多变,提醒大家适时增减衣物,小心着凉。

## LLM 的其他写作功能

 您对 LLM 的哪一项文字功能有兴趣, 我们可以试试看!

 翻译 — 实时多语言文本翻译。

 校对与润色 — 检查拼写、语法, 提供改进建议。

 文案撰写 — 生成吸引人的营销文案和产品描述。

 信息提取 — 从文本中提取信息, 回答问题。

 专业化邮件 — 使文字更为正式。

 咨询服务 — 提供职业、学习、心理健康等建议。





## AI 可以做什么？

---

 信息获取

## 孰真孰假？

最近在互联网上看到一些消息,不知真伪—

-  在日本和韩国,西瓜属于奢侈品,一般人吃不起西瓜。
-  鱼类不再被认为是正式的生物分类。
-  12 星座起源于中国,最早可以追溯到辽朝。
-  两名女性可以使用她们的卵子生育一个分享她们双方一半 DNA 的后代。

LLM 能帮助我们吗？

## 🤔 孰真孰假？

最近在互联网上看到一些消息,不知真伪—

- 🍉 在日本和韩国,西瓜属于奢侈品,一般人吃不起西瓜。
- 🐟 鱼类不再被认为是正式的生物分类。
- ⭐ 12 星座起源于中国,最早可以追溯到辽朝。
- 🧬 两名女性可以使用她们的卵子生育一个分享她们双方一半 DNA 的后代。

LLM 能帮助我们吗？

关于 🐟 我们可以试试问 LLM 下面的问题—

- 你是一个世界顶级的博物学家。我**听书**鱼类不再被认为是合理的生物学分类。这是真的吗？





## 把我当成小孩子

当我们在遇到不明的名词时,我们可以要求 LLM 以适当的方式来帮助我们理解。

- 输入 — “请把我当成 5 岁小孩,解释为什么鱼类作为一个生物学分类不存在?”



## 把我当成小孩子

当我们在遇到不明的名词时,我们可以要求 LLM 以适当的方式来帮助我们理解。

- 输入 — “请把我当成 5 岁小孩,解释为什么鱼类作为一个生物学分类不存在?”
- 输入 — “请把我当成青少年,解释为什么鱼类作为一个生物学分类不存在?”



## 把我当成小孩子

当我们在遇到不明的名词时,我们可以要求 LLM 以适当的方式来帮助我们理解。

- 输入 — “请把我当成 5 岁小孩,解释为什么鱼类作为一个生物学分类不存在?”
- 输入 — “请把我当成青少年,解释为什么鱼类作为一个生物学分类不存在?”
- 输入 — “请把我当成大学学生,解释为什么鱼类作为一个生物学分类不存在?”

 AI 可以做什么？

---

 学习知识



## LLM 可以当“举例小助手”

你可以把 LLM 当成一个“举例小助手”，比如 —

- 🐟 “请给我五个我们认为是鱼，但其实和陆地生物关系更近的例子？”
- 🍏 “请给五个日常生活中用到概率的例子？”
- ⚠️ “请提供五个常见的逻辑错误例子，比如生活中常听到的错话？”
- 📈 “请举五个用数据来误导人的例子？”
- 🧠 “请给我五个容易记住事情的小窍门？”



## LLM 可以当“耐心的辅导老师”






你可以把 LLM 当成一个“耐心的辅导老师”，比如一

- 🤔 “理性是不是就是自私？请解释一下。”
- ⚖️ “为什么科学实验要有对照组？”
- 🐱 “薛定谔的猫是什么？请用简单的方式说明。”
- 📖 “请一步一步教我怎么解一个简单的方程？”
- 🔗 “什么是因果关系？它和相关性有什么不同？”



## LLM 可以当“出题老师”

你可以把 LLM 当成一个“出题老师”，比如 —


-  “请出 15 道简单的加减乘除题，数字不要超过 100。”
-  “请给我五道简单的逻辑推理题。”
-  “请出 10 道概率的选择题，适合初学者。”
-  “请给五道高中物理的基础题，比如简单的力学问题。”
-  “请出五个简单的阅读理解练习。”



## LLM 可以“解答难题”

我们可以给 LLM 复杂的数学题 —

在满足以下条件的情况下,教授 Y 有多少种方法可以将 6 本不同的

 分配给 4 位不同的学生:

- 最贵的  必须分给学生 X。
- 每位学生至少要分到一本 ?



## 🤔 LLM 可以“解答难题”

我们可以给 LLM 复杂的数学题 —

在满足以下条件的情况下,教授 Y 有多少种方法可以将 6 本不同的  
📖 分配给 4 位不同的学生:

- 最贵的 📖 必须分给学生 X。
- 每位学生至少要分到一本 📖?

但我们需要选择适合的模型 —

- ✗ ChatGPT 4-turbo — 540
- ✓ ChatGPT o1-preview — 390
- ✗ Claude Sonnet 3.5 — 600
- ✗ Google Gemini 1.5 Pro — 792

## 🤖 LLM 可以“解答难题”

### 各个模型的长处 —

- OpenAI
  - ChatGPT 4o — 综合能力最佳,适用于一般性和深度问题。
  - ChatGPT o1-preview — 擅长处理复杂逻辑和数学推理问题。
- Anthropic
  - Claude Sonnet 3.5 — 编程能力较强,解释代码清晰。
  - Claude Haiku 3.5 — 速度快,能力强。
- Google
  - Gemini 1.5 Pro — 价格实惠,适合对预算敏感的任务。
  - Gemini 1.5 Flash — 面向快速响应的轻量模型,速度优先。
- Meta LLaMA 3.2 — 可以本地部署,适合隐私需求高的场景。

## 长文总结

很多文章(比如新闻)篇幅很长,但我们只需要其中的重要部分。

或者我们读完一本长书,我们希望有一个提纲帮助记忆。

LLM 可以帮助我们做到这一点。

即使是很长的一章,它也可以一次性处理。

很多文章(比如新闻)篇幅很长,但我们只需要其中的重要部分。

或者我们读完一本长书,我们希望有一个提纲帮助记忆。

LLM 可以帮助我们做到这一点。

即使是很长的一章,它也可以一次性处理。

### How Not To Die Chapter 1 – 5200 单词

CHAPTER 1 How Not to Die from Heart Disease

Imagine if terrorists created a bioagent that spread mercilessly, claiming the lives of nearly four hundred thousand Americans every year. That is the equivalent of one person every eighty-three seconds, every hour, around the clock, year after year. The pandemic would be front-page news all day, every day. ...

很多文章(比如新闻)篇幅很长,但我们只需要其中的重要部分。

或者我们读完一本长书,我们希望有一个提纲帮助记忆。

LLM 可以帮助我们做到这一点。

即使是很长的一章,它也可以一次性处理。

### LLM 的总结

心脏病是美国头号杀手,每年致死 40 万人,但其实可以通过改变饮食来预防和逆转。研究发现,以植物性食物为主的饮食方式可以有效降低胆固醇,预防和治疗心脏病。在一些传统饮食以植物为主的地区,心脏病极为罕见。关键是要将 LDL 胆固醇控制在低于 70mg/dL 的水平。这可以通过服用他汀类药物实现,但采用植物性饮食同样有效,且没有副作用。

# AI 可以做什么？

---

 编程助手

在编程论坛网站 Stack Overflow 的 2024 年度调查中,

- 61.8%的受访者正在使用 AI
- 13.7%的受访者计划 AI
- 24.4%的受访者没有计划使用 AI

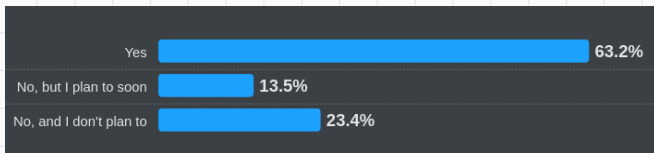


图 3: 你目前在编程中使用 AI 吗?

在编程论坛网站 Stack Overflow 的 2024 年度调查中,

- 61.8%的受访者正在使用 AI
- 13.7%的受访者计划 AI
- 24.4%的受访者没有计划使用 AI

🙄 这即说明了 LLM 的强大能力,也反映了 LLM 经常出错的缺点。

💣 在使用 LLM 的过程中,必须意识到它们经常会出错!

😏 应该把 AI 当成经常犯错的实习生而不是万能的 🤖!



🤖 但有了 AI 的帮助,人人可以写一些简单的程序。

🕒 原本不值得花费时间写的程序,现在可以使用 LLM 生成。

😊 我们可以很多定制化的程序。

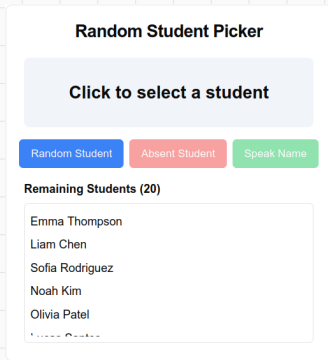


图 3: 花了 5 分钟由 LLM 生成的随机学生选择器

## AI 可以做什么？

---

 数据分析



LLM 使数据分析变得很简单。

比如,在餐厅吃饭的时候,我们可以用  把收据拍下来。

Yunnan Yao Dai tribe stir fried beef with herb	丽江风味爆炒猪	1	份	88
Liliang flavour stir fried pork	云酱烤鲈鱼	1	份	158
Grilled bass with herbs & spices	米饭	3	份	30
普娜	[酒水饮料]			689
Acqua Panna (750ml)	普娜	1	瓶	70
藏式艾尔	藏式艾尔	1	瓶	65
Tibetan pale Ale	干姜水	1	瓶	30
干姜水	干姜水	1	瓶	50
Ginger Ale	青岛	1	杯	75
青岛	天堂梵趣	1	瓶	50
Tsing Tao	LOST HEAVEN PUNCH	1	瓶	70
普娜	普娜	1	瓶	70
Acqua Panna (750ml)	干姜水	1	瓶	30
干姜水	干姜水	1	瓶	70
Ginger Ale	榛子黑巧克力蛋	2	份	84
糕	糕			
椰奶草莓冰激凌	椰奶草莓冰激凌	3	份	165
合计:	合计:			1754

品项	数量	单位	小计
[大厅菜品]			1065
蔬菜春卷	1	份	58
Spring rolls			
云南野菜饼	1	份	58
Yunnan wild vegetable cakes			
云南茄子拌豆腐	1	份	68
Yunnan eggplant and tofu salad			
景颇鬼鸡	1	份	88
Jin bo-style ghost chicken salad			
傣族风味番茄拌	1	份	68
茄子			
Eggplant with tomato chilli sauce			
缅甸虾酱炒四季	1	份	68
豆			
Burmese shrimp paste with String beans			
炒花椰菜	1	份	58
Stir-fried broccoli with tomatoes			
小炒黄瓜花	1	份	68
Stir fried cucumber flower			
云南白蓼炒蛋	1	份	79
Scrambled eggs with yunnan white mushrooms			
傣味鸡肉柠檬	1	份	88
撒			
Dai tribe style chicken lemon sauc			
花腰傣风味炒牛	1	份	88
肉			

图 4: 账单数据分析



然后我们可以要求 AI 把这些收据转换成表格。

品项	数量	单位	价格	小计
蔬菜春卷	1	份	58	58
云南野菜饼	1	份	58	58
云南茄子拌豆腐	1	份	68	68
景颇鬼鸡	1	份	88	88
傣族风味番茄拌茄子	1	份	68	68
缅式虾酱炒四季豆	1	份	68	68
炒花椰菜	1	份	58	58
小炒黄瓜花	1	份	68	68
云南白蔘炒蛋	1	份	79	79
傣味鸡肉柠檬"撒"	1	份	88	88
花腰傣风味炒牛肉	1	份	88	88
丽江风味爆炒猪肉	1	份	88	88
云南烤鲈鱼	1	份	158	158
米饭	3	份	10	30
普娜 (750ml)	2	瓶	70	140
藏式艾尔	1	瓶	65	65
干姜水	2	瓶	30	60
青岛	2	瓶	50	100

图 4: 账单数据分析



接下来我们可以要求 AI 把这些收据分析一下。

## 账单分类统计

素菜	¥525 (29.9%)
荤菜	¥510 (29.1%)
主食	¥30 (1.7%)
酒水	¥689 (39.3%)
甜点	¥249 (14.2%)

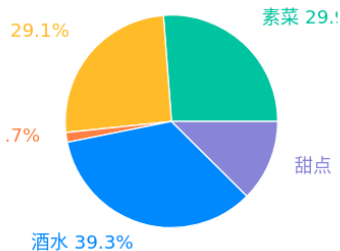


图 4: 账单数据分析

# AI 可以做什么？

---

 图像识别

在讨论 AI 的时候,右边这幅发表于 2014 年的漫画时常会被引用。

如果您不熟悉计算机科学,可能会有点困惑。

LLM 能够帮助我们理解它吗?



图 5: 漫画来源 — XKCD

在讨论 AI 的时候,右边这幅发表于 2014 年的漫画时常会被引用。

如果您不熟悉计算机科学,可能会有点困惑。

LLM 能够帮助我们理解它吗?

😊 其实不止是静止的图像, LLM 还可以识别视频。



图 5: 漫画来源 — XKCD



## AI 可以做什么？

---

 艺术创作

## 🎨 一分钟成为“摄影师”

💬 你能猜出哪一张“照片”是 AI 生成的吗？



图 6: 图片 1 & 图片 2

## 🎨 一分钟成为“摄影师”

💬 你能猜出哪一张“照片”是 AI 生成的吗？



图 6: 图片 1 (By rdne on pexels) & 图片 2 (By Flux)

## 🎨 一分钟成为“画家”

💬 你能猜出哪一张照片是 AI 生成的吗？



图 7: 图片 1 & 图片 2

## 🎨 一分钟成为“画家”

💬 你能猜出哪一张照片是 AI 生成的吗？



图 7: 图片 1 (By Flux) & 图片 2 (By Rembrandt van Rijn)

## 目前主要的 Diffusion Model —

- Stable Diffusion — 开源免费,可操控性最佳
- DALL-E 3 — 和 ChatGPT 集成,使用方便; 微软 Copilot 直接提供访问
- Midjourney — 高质量艺术风格生成效果,适合创意需求
- Flux — 实验性前沿技术,专注生成速度和资源效率
- Imagen (Google) — 注重图像细节与质量,适用于高保真需求
- Firefly (Adobe) — 集成于 Creative Cloud,直接支持创作者 workflow

在使用 Diffusion Model 时,有很多艺术风格可以选择。



图 8: Realistic Photo (照片) — Flux

在使用 Diffusion Model 时,有很多艺术风格可以选择。



图 8: Traditional Chinese Painting (水墨画) — DALL-E 3



在使用 Diffusion Model 时,有很多艺术风格可以选择。



图 8: Hyperrealistic Comic (漫画) — DALL-E 3

在使用 Diffusion Model 时,有很多艺术风格可以选择。



图 8: Cyberpunk (赛博朋克) — DALL-E 3

很多 Diffusion Model 都支持添加文字。



图 9: 宇宙中最热的热狗

很多 Diffusion Model 都支持添加文字。



图 9: 可惜都不支持中文

😞 在 Google 搜索 Baby Peacock (孔雀宝宝), 会得到什么结果?

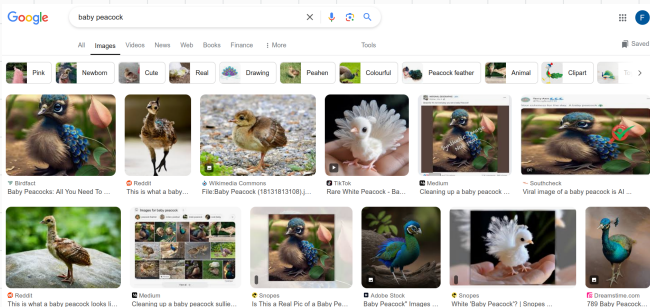


图 10: Google 搜索结果

# 👁️ 眼见为实？

😞 在 Google 搜索 Baby Peacock (孔雀宝宝), 会得到什么结果？

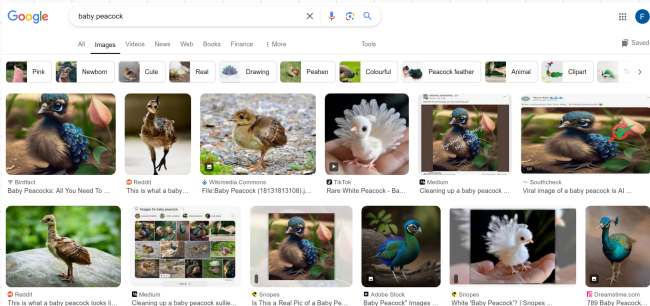


图 10: Google 搜索结果

😱 几乎全部结果都是 AI 生成的图像！

🕒 眼见为实已经是过去式。寻找可信的信息来源最重要。

## ⚠️ 小心骗子

今年5月,跨国公司 Arup 的一位财务人员参加了一个视频会议。

会议中骗子利用深伪技术伪装成公司 CFO 和多名公司员工。

骗子要求该财务人员转账 2.5 亿美元。

他相信了会议中的人是真实的,并同意转账。

🚚 全球诈骗案件增多,大家必须保持警惕!



图 11: 深伪图像示例 📷

# 🎵 一分钟成为“音乐家”

Diffusion Model 也可以生成 🎵。

目前比较流行的音乐生成服务有 Suno 和 Udio。

歌词和歌曲都可以生成。

🎧 我们来听一首 AI 歌曲吧！

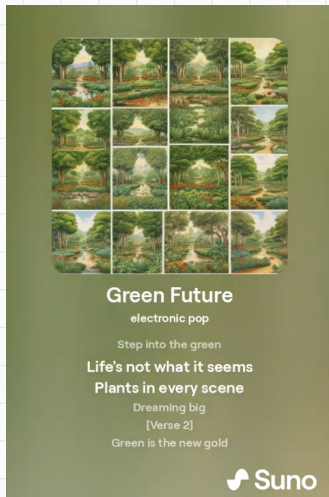


图 12: Green Future (绿色未来) — By Suno





## 一分钟成为“导演”

Diffusion Model 也可以用来视频,例如 — [Sora](#) (OpenAI), [Veo](#) (Google), [Movie Gen](#) (Meta)。



图 13: Sora — 一位时尚女性走在东京的街道上 ...

有兴趣可以尝试 [RunwayML](#)。

## AI 可以做什么？

---

 语音处理

ChatGPT 集成了一个很好的语音识别系统 Whisper。

在 ChatGPT 使用语音输入非常的方便。

ChatGPT 也可以将它的输出语音化。

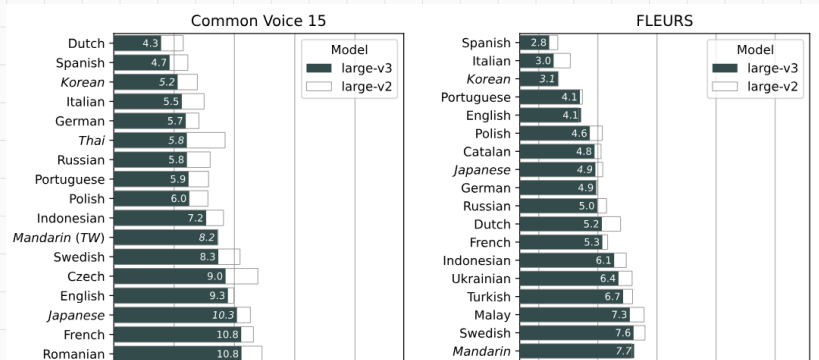


图 14: Whisper 语音识别的错误率

与机器进行像人类般自然的交谈曾是科幻小说中的情景。

但随着 ChatGPT 语音模式和 Gemini Live 的推出,这一愿景几乎成为了现实。

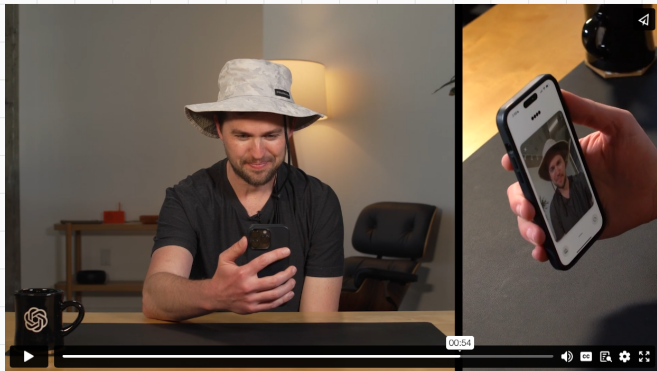


图 15: 和 GPT-4o 的视频聊天演示视频

与机器进行像人类般自然的交谈曾是科幻小说中的情景。

但随着 ChatGPT 语音模式和 Gemini Live 的推出,这一愿景几乎成为了现实。

这为众多应用场景打开了大门 —

- 实时外语教学与翻译
- 虚拟客户服务与技术支持
- 辅助学习与个性化教育
- 智能个人助理
- 心理健康支持与陪伴

 不过实时通话还是不太稳定,延时较多,也经常中断。

Google NotebookLM 允许用户上传文件,然后询问一个机器人关于这些文件的问题。

但它最有趣的一个功能是可以生成一个播客一样的语音总结。

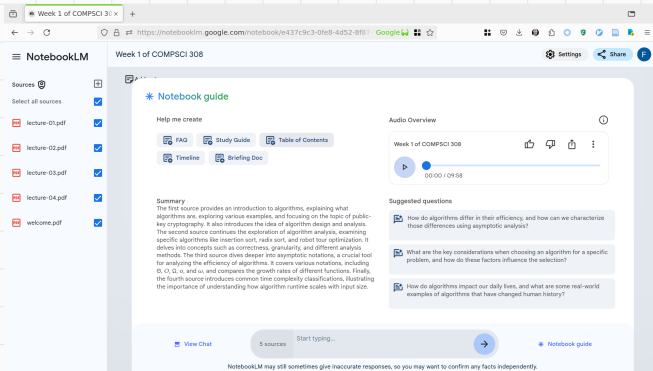


图 15: Google NotebookLM

- 👁️ AI 技术的进步也许可以为视觉障碍的人提供帮助
  - 图像到语音描述: 帮助视力障碍用户获取视觉内容
  - 手语翻译: 实时翻译手语为文字或语音
  - 协助阅读: 通过语音朗读书籍或网页内容
  - 有声书制作: 通过语音生成有声书

### AI 技术的进步也许可以为视觉障碍的人提供帮助 —

- 图像到语音描述: 帮助视力障碍用户获取视觉内容
- 手语翻译: 实时翻译手语为文字或语音
- 协助阅读: 通过语音朗读书籍或网页内容
- 有声书制作: 通过语音生成有声书

### AI 还可以做更多语音相关的事情 —

- 在文字转语音中添加情绪
- 语音克隆
- 语音切换
- 情绪识别和分析
- 实时翻译字幕
- 会议秘书



# AI 学习指南

---



## LLM 是如何写作的？

想象一下 LLM 像一个玩文字接龙的超级高手 —

- 你写下：“今天天气真...”
- LLM 会猜测：“好”“糟糕”“晴朗”等可能的词

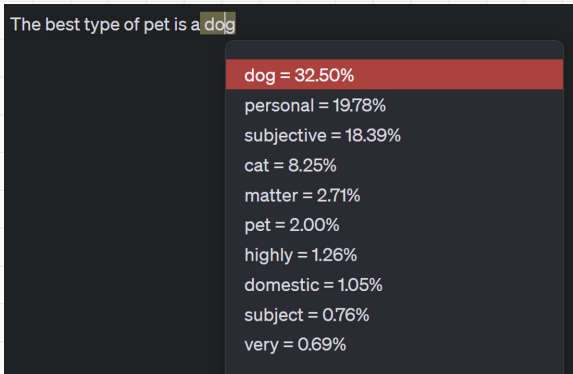


图 16: LLM 预测下一个词的例子



## LLM 如何写出一段话？

想象 LLM 像在写故事 — 一次一个词，

- 每个词都取决于之前的内容
- 内容变化会导致不同的故事走向
- 这解释了同一个问题可能有不同答案

the best type of pet is a **dog** because they are loyal, loving, and make great companions. Dogs are known as "man's best friend" for a reason. They are always happy to see you, and their unconditional love and affection can greatly improve your mood and overall well-being. Dogs are also great for providing companionship and can help reduce feelings of loneliness. They are also highly trainable and can be taught to do a variety of tasks, making them useful in many different ways. Additionally, dogs come in all shapes and sizes, so there is a breed that can fit into any lifestyle. Overall, dogs make wonderful pets and can bring so much joy and




the best type of pet is a **subjective** question and depends on individual preferences and lifestyle. Some people may prefer dogs for their loyalty and companionship, while others may prefer cats for their independence and low maintenance. Other popular pets include birds, fish, rabbits, and reptiles. Ultimately, the best type of pet is one that fits well with your lifestyle and brings you joy and companionship.

图 17: LLM 写作的过程



# LLM 从哪里学习知识？

想象一个博览群书的“学霸” — LLM 从以下来源获取知识：

-  网络文章
-  各类书籍和杂志
-  科学研究和论文

It was the White Rabbit, trotting slowly back again, and looking anxiously about as it went, as if it had lost something; and she heard it muttering to itself "The Duchess! ~~The Duchess on my dear paws!~~ On my fur and whiskers! She'll get me executed, as sure as ferrets are ferrets! Where can I have dropped them, I wonder?" Alice guessed in a moment that it was looking for the fan and the pair of white kid gloves, and she very good-naturedly began hunting about for them, but they were nowhere to be seen — everything seemed to have changed since her swim in the pool, and the great hall, with the glass table and the little door, had vanished completely.

Duchess = 98.55%

Duch = 1.28%

Du = 0.04%

Dutch = 0.03%

Duc = 0.02%

D = 0.01%

= 0.01%

Dou = 0.01%

Duke = 0.01%




\n\n = 0.01%

Total: -0.01 logprob on 1 tokens  
(99.98% probability covered in top 10 logits)

图 18: LLM 展现阅读理解

## 🧠 LLM 的记忆有多长？

LLM 的记忆类似于短期记忆：

-  每次对话重启，LLM 相当于“重新开始”
-  只能记住当前对话内容
-  超过记忆限制的内容会遗忘

LLM 模型	Context Window (记忆长度)	相当于
GPT-4	12.8 万 tokens	 约 300 页书
Claude 3.5	20 万 tokens	 约 500 页书
Gemini 1.5	200 万 tokens	 约 5000 页书

表 1: 不同 LLM 的“记忆力”对比

## 💡 如何让 LLM 更好帮助我们？

使用小技巧和未来前瞻：

- 🔄 LLM 卡住时,重启对话
- 🎭 独特提问方式激发创意回答
- ⚖️ 牢记 LLM 有局限,不要期望无所不知
- 🙌 多试多用,熟能生巧

展望未来：

- 📈 LLM 的 Context Window (记忆长度) 会逐步扩大
- 🧠 未来更能理解和记住用户偏好
- 🔗 更强的数学和编程能力

📖 推荐阅读：[Thinking Like an AI by Ethan Mollick](#)

## 用 10 小时掌握 AI 的核心

The most effective way to understand AI is through hands-on experience. For about 10 hours, immerse yourself in experimenting with AI—try out tasks you typically do for work or fun, explore its quirks, and ask it unexpected questions. This practical exposure will teach you far more than reading articles alone. Through this, you'll gain an in-depth understanding of AI's strengths and limitations, potentially uncovering surprising insights along the way.

— Ethan Mollick, *Thinking Like an AI*

😊 如果你不喜欢读英文，LLM 可以帮你翻译成中文。—

真正理解 AI 的方法是多多使用它。用大约 10 个小时来完成你的工作或娱乐时的各种操作，尝试用 AI 完成这些任务。探索它，触碰它，向它提出奇怪的问题。观察它的亮点和不足。亲身体验会教会你比任何文章都多的知识（即使是这篇长文）。你会发现许多关于如何有效使用 AI 的方法，甚至可能会被自己的发现所惊讶。

— 艾森·莫里克，《像 AI 一样思考》



 如果这段话太长，LLM 可以帮你总结。—

要真正理解 AI，就需要亲身实践。通过大约 10 个小时的实际使用，探索它的优缺点，你会学到比任何文章更多的知识，甚至会对自己的发现感到惊讶。

 如果还是太长，LLM 可以帮你压缩成一句话。—

用 10 小时亲身体验 AI, 比读任何文章更有效。

# 🗨️ 你希望 AI 还可以做什么？



安多与机器人 — 电视剧《星球大战：安多》(2022)