# **从简单问答到复杂决策：大模型工具重新定义问题解答V1**

**作者**：游清宁

**校对**：左文杰

**关键词：**AI工具、GPT、达克效应、乔哈里沟通视窗、5W1H、头脑风暴、角色扮演、文生图、文生视频

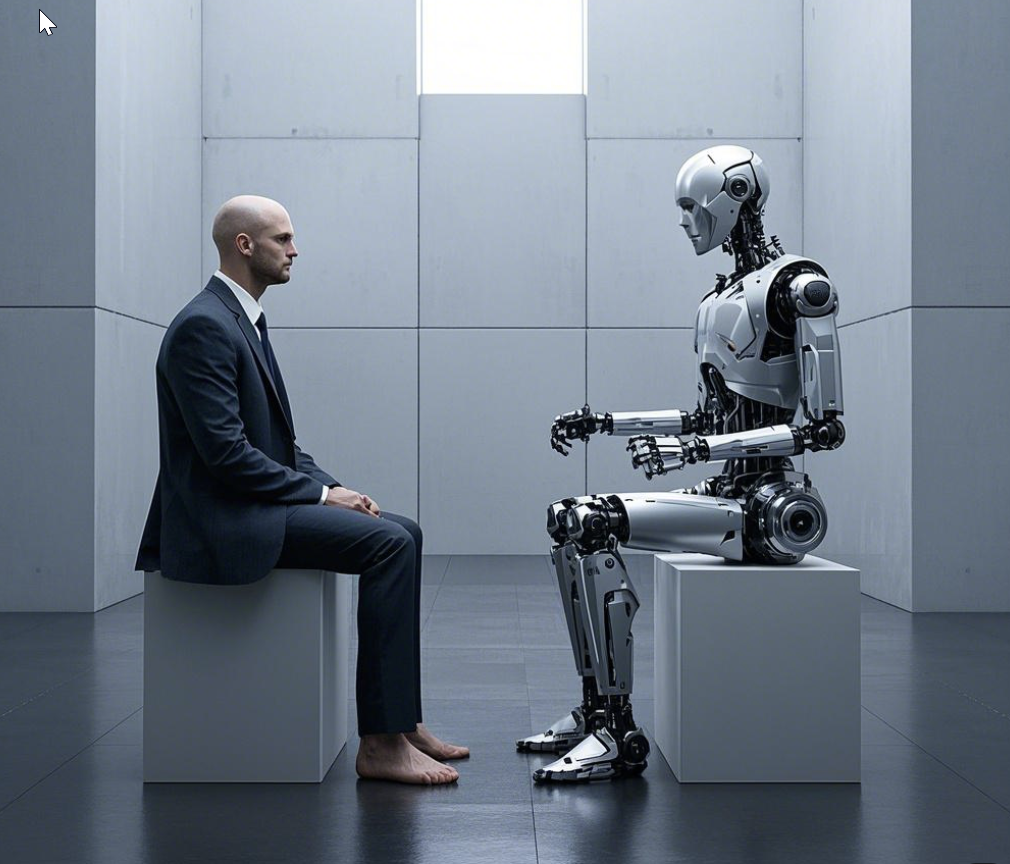
# **1 引言**

近两年，AI 世界仿佛一夜之间变得五彩斑斓。KIMI、豆包、智谱、天工、混元、夸克、千问、文心、盘古、讯飞、Gemini这些名字就像雨后春笋般冒了出来，每一个新工具的发布都伴随着“颠覆”、“吊打”、“改写”、“重新定义”等震撼人心的词汇，标题党们更是不遗余力地制造焦虑感，仿佛不跟上这波潮流就会被时代抛弃。但话说回来，这到底是AI的泡沫，还是我们还没摸到AI的法门？在有限时间内局限性地综合了网络上的各种信息和自己的亲身体验后，提炼了一些关于如何在AI时代精进自我，灵活运用这些大模型工具解决复杂问题的小心得，希望能给大家的AI之旅提供点有意思的参考。

# **2 简介及概念**

## **2.1 未来科技幻想：人机共处**

先来看一个有趣的示例，让大模型工具根据描述画一幅画。这里使用的是大模型工具“豆包”，仅通过输入提示词，我们得到了如下图片，画面细节处理得当，人类西装纹理褶皱自然，机器人复杂的机械部件因光线的巧妙折射，呈现出一种机械美学。周围光滑反光的墙壁与地板，进一步强化了空间的冷峻感。这不仅仅是一幅画，更是对未来人机共处场景的生动畅想，让我们直观地感受到科技发展之下，人类与AI智能之间可能会碰撞出的奇妙火花。（篇幅原因本文不对文生图作过多说明，有兴趣的读者可自行了解）



提示词（Prompt）：画一幅画，在一个极简主义的未来场景中，两个实体相对而坐。视角是从自身向前的视角，坐在自己一侧的是具有生物特征的人类，穿着西装，姿态随意但深思熟虑，赤脚放在冰冷的金属地板上。视角的前方是高度精密的机器人，由金属和复杂的机械部件构成，关节和线路在光线的反射下闪闪发光。它们坐在同样的白色立方体上，中间是一个空旷的空间，周围的墙壁和地板光滑且反光，营造出一种冷峻、理性和充满科技感的氛围。光线从上方照射下来，形成清晰的阴影，凸显了两者的对比。

## **2.2 基本概念**

描述中提到的一些名词，如，提示词、大模型工具、AI等，下面仅对几个名词作简要介绍。

1. **AI（Artificial Intelligence，人工智能）**

AI是一门研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的新兴技术科学。目标是让机器能够模仿人的行为，能够看、听、说、想、做。其涵盖了多个领域，包括机器学习、深度学习、自然语言处理、计算机视觉等。

1. **大模型工具**

大模型工具指的是那些基于大规模数据集训练而成的人工智能系统，能够处理复杂的任务并展现出类似人类的认知能力。这些模型通过深度学习算法从大量的文本、图像、音频等数据中学习模式，并能够根据新的输入做出预测或生成新的内容。大模型工具广泛应用于自然语言处理（NLP）、计算机视觉、语音识别等多个领域，如通义、豆包、文心一言、GPT系列等都是知名的大模型实例。

1. **提示词（Prompt）**

提示词是指用户输入给AI模型的文本或指令，用以指导AI生成特定类型的输出。它可以是一句话、一个问题、一段描述或任何其他形式的文本内容，目的是让AI根据提供的信息来创作故事、诗歌、图像、音乐等，或是回答问题、提供解释、建议解决方案等。对于不同的应用场景，提示词的设计非常重要，它不仅决定了AI输出的内容方向，也影响着输出的质量和相关性。

# **3 基于5W1H基本问题解答**

在了解了大模型工具的基本概念之后，我们不禁好奇，这些工具在面对日常生活或工作中的基本问题场景时，表现又会如何。接下来，我们主要以ChatGPT为例进行测验，其作为当下备受瞩目的大模型代表之一，其强大的语言处理能力使其能够应对各式各样的问题，特别说明的是，作为主要以英文体系为知识沉淀的工具，问答最好都基于英语。它能支持包含搜索、编程、写作、翻译、推理及插件体系的功能，号称具有深入剖析及多维解答的能力，能给我们的办公及生活带来各种超乎想象的作用，我们将从运用 5W1H 开始，测试它所能解答的问题范畴并评估其反馈内容。基于网络原因，下述相关内容都可在如”豆包“之类的国内应用中得到类似的反馈和体现。

## **3.1 简单的问题（**聚焦于 what、who、when、where**）**

对于4W（what、who、when、where）问题，主要解决的是事物的基本定义、主体信息、时间节点以及地理位置等方面的疑问，此时，相关工具可作为百科全书，轻松满足对基础信息的需求。

这里先尝试了几个问题，例如， “唐朝是哪一年建立的，国祚延续了多少年”，“什么是 PDCA”，“XX 大学的地址在哪里，有几个校区”，它都能清晰准确地给结果，并整理出相关的知识内容，满足我们的要求。

## **3.2 复杂的问题（**聚焦于 why、how**）**

面对“why”和“how”这两类复杂的问题，关键在于基于事实进行深度思考，挖掘现象背后的根本原因，进而依据这些本质原因找到解决问题的办法。

### **3.2.1 why的问题**

回答why的问题是基于事实为基准：

* 为什么动物都有寿命限制；
* 为什么战国七雄中，最终是秦国统一天下，而不是其他六国；
* 请用逻辑推导，为什么日不落帝国的人，自从形成了喝茶的习惯之后，人口就大量增加了；
* 为何中医无法像西医一样，在全世界得到推广；
* 为什么像集成芯片这样的技术先出现在美国而不是中国，如果以后的新技术要首先出现在中国，中国需要做哪些改进？

类似问题的回答可以推动我们对一些已有事实的思考。

### **3.2.2 how的问题**

回答how的问题，这类问题主要希望由目标回答出路径和工具/方法：

* 已经20多岁，如何快速提升自己的审美；
* 在答案过剩的时代，如何提出有价值的好问题，如何锻炼自己提问的能力；
* 如何在3天内，快速了解一个行业；
* 我是做电梯出口的，刚从大学毕业加入公司，很多行业相关的概念都不懂，请帮我列出20个必须掌握的行业相关的概念；
* 或者，再复杂一些的问题：（比如目标不是很明确的类型）我作为一名电子电路设计的新员工，如果我想在2年内掌握足够的电子电路知识，从而有更好的职业前景，你能在哪些方面给我提供帮助
  + --> 可进一步追问（通过前面的回答，我们可以继续追问）：要通过什么样的方法和路径，帮我提升电子电路相关能力
  + -->（还可继续）可以提供给我相关的网址及相关课程或是工具路径。

上述这些问题，在大模型工具出现之前我们也可以通过传统的搜索引擎去获取，但大模型工具的出现展现出诸多独特优势：

1. 高效信息获取。能够直接理解用户的问题，并给出精炼、准确的答案。

2. 双向交互。不仅能够回答用户的问题，还能根据上下文进行反向提问，进一步引导对话，挖掘更深层次的信息。

3. 办公提效。无论是搜索、编程、写作、翻译还是推理，甚至是excel公式及复杂操作，都能提供有力的支持。它们能够理解复杂的指令，协助用户完成一系列实际的工作任务。此外，部分如 GPT 这类先进模型还构建有插件体系，比如报表工具、联网、特定接口等，如同为其赋予 “超能力”，进一步拓展了功能边界，解锁更多应用场景，全方位满足用户多样化需求。

# **4 如何提问才能得到高质量的答案**

当前，尤其是大模型工具的应用中，提问工程已成为一项至关重要的技能，它不仅需要对提问模型要有深刻的理解，还需要创造力和战略思维的结合。下面我们将从多种提问体系入手，借用作者上网查得和测试的各种实例，帮助引导大模型工具提供准确、相关且有用的回答。

## **4.1 乔哈里沟通视窗**

首先，为了避免大语言模型出现“垃圾输入——垃圾输出”现象，我们可借助乔哈里（Johari）沟通视窗这一工具。通过了解视窗中的四个象限，我们能明确自己与GPT之间的认知差异，提升提问质量。



### **4.1.1 盲区的提问推荐**

盲区提问的核心在于通过提问帮助发现自己不知道的知识点，或者掌握一些领域中的基础概念。有几种推荐的提问方式：

* **问法1（元问题提问法）**：我想快速了解XXX行业，我应该问你哪些问题？
* **问法2**：请列出XX领域/行业相关的最常用50个概念，并简单解释。若有英文缩写，请给出完整英文解释。

参考例：

* 请详细介绍川普的生平事迹，精确到年月，并给出相关的影响。
* 请详细介绍Tesla的发展历程。
* 最近一年，美国对进口新能源车有什么新的政策，对中国的新能源车出口有何影响？中国应该如何应对？  
  → 进一步追问：汽车品牌出海已有许多成功先例，比如欧美、日韩等国家。能帮我找几份相关的汽车品牌出海发展历程的研究报告吗？
* 关于孩子的心理健康教育，我已经了解了一些概念，比如‘习得性无助、路径依赖、延迟满足、认知失调、自我效能感、家庭动力学’。除此之外，家长还应该掌握哪些其他重要概念？

### **4.1.2 公开区的提问推荐**

公开区的提问，主要用来检验认知和扩充新的知识。

首先，检验自己对某个主题的理解程度和深度：

* **问法1**：对于XX主题/技能（比如掌握GPT的提问技巧），哪些是我必须理解和掌握的核心要点？”  
  → 进一步追问：“你提到了XXX，针对这个，你认为我还需要掌握哪些关键内容？
* **问法3**：我对XX有一些想法（成为数据分析师），你能帮我批判性地分析这些想法的优缺点吗？
* **问法4**：我正在考虑XXX的决定（自媒体创业），你能帮我分析一下可能的结果和影响吗？  
  例如：我正考虑辞去固定的工作，成为自媒体博主，做一个数字游民。你能帮我分析一下可能的结果和影响吗？

其次，扩展认知，寻求更深入、更广泛的信息：

* **问法1**：我正在考虑学习XXX，你能给我一些关于这个领域未来发展的观点吗？
* **问法2**：我知道XXX（AI）的概念，能否进一步了解更多关于XXX的信息？
* **问法3**：我在XX问题上遇到了困难，你能提供一些可能的解决方案或建议吗？
* **问法4**：我想深入了解学习XXX，能推荐一些进阶的学习资源或学习路径吗？
* **问法5**：我想在XX领域有所创新（开发新的APP），你能提供一些启发或思路吗？
* **问法6**：（背景信息）我要做关于XXX的研究，当前已了解了XXX的内容，是否还有其他可能的原因？能给出一些假设吗？”
* **问法7**：我是一个XX行业新手（例如人工智能、创业新手），马上要采访某个行业的资深大佬，我应该向他请教哪些有价值的问题？”（这个是比较常用的句式）  
  → 进一步：我们能再聚焦一下，围绕这个主题还能提出哪些有价值的问题？

### **4.1.3 隐私区的提问推荐**

通常是在自己有明确观点或专业知识的情况下，请求更多的洞见与反馈。

* **问法1**：这会对XXX产生什么样的影响？
* **问法2**：可能的原因有哪些？
* **问法3**：你怎么看待这种现象？
* **问法4**：“你觉得XXX应该怎么做？”

参考例：

* openAI发布新产品，你认为，ChatGPT这类产品的出现，可能对谷歌、百度之类的传统搜索产生什么样的影响？而作为传统的中型制造业公司，我们应该采取哪些必要措施，去迎接这轮技术革新？(多问题同时提问)

### **4.1.4 未知区的提问推荐**

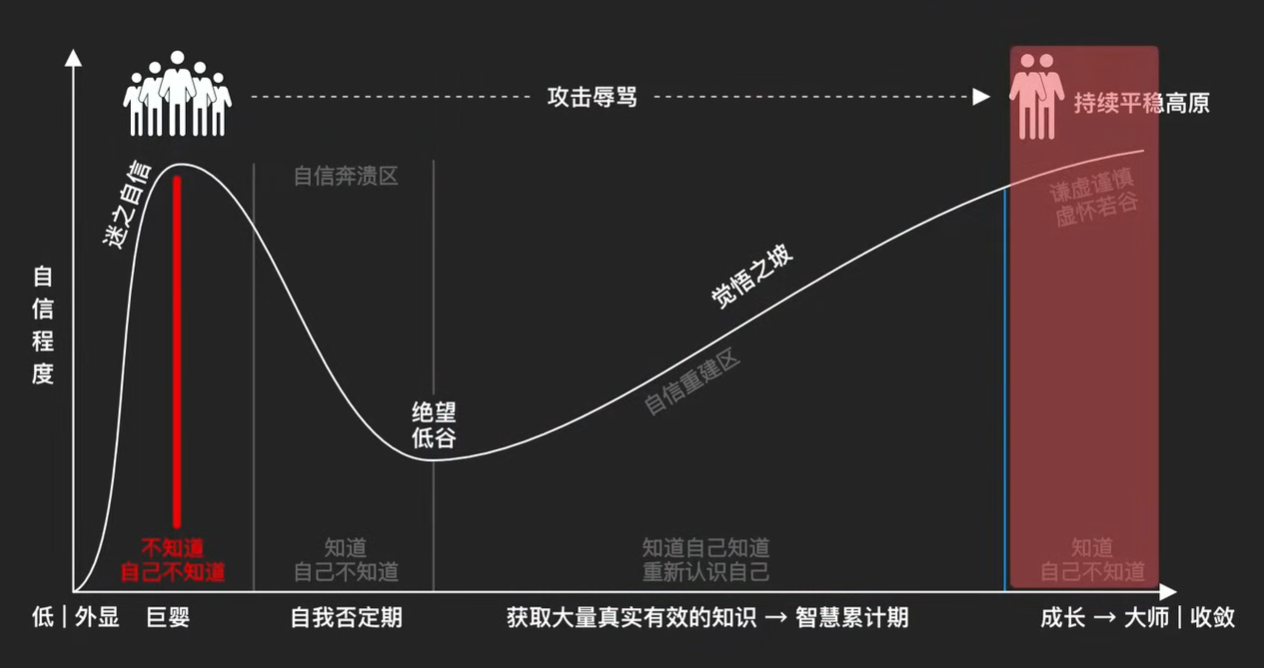
通常涉及到假设性的问题，探索自己尚未接触过的领域和未来可能发生的事件。

* **问法1**：如果XXX（假设发生了某事），这对社会会有什么影响？

例如：如果GPT在整个社会大范围使用，那会对英语辅导、少儿编程、写作辅助、文员办公、财务记账等行业产生什么样的影响？具体影响的点是什么？对于个人而言，可能需要做哪些改变和应对？

## **4.2 达克效应**

在前面的基础上，我们又要引入一个问题：你怎么知道自己真的理解某个领域？在这个领域的认知水平是否客观准确？为此，找到了一个原先成熟的模型——达克效应（Dunning-Kruger Effect）。这一效应揭示了人们在认知过程中可能遇到的盲点和误区。因为其中的第二和第三阶段比较容易理解，我们将重点讨论第1阶段和第4阶段，并通过合适的提问技巧，规避认知陷阱。



### **4.2.1 针对第一个阶段**

在这个阶段，可通过反向提问来检测客观的专业水平：

* **问法1**：为了测试我对xxx（明朝历史）的了解程度，你会问我什么问题来检验我的水平，最少10个
* **问法2**：我是xx领域（动画片领域）的专家，你会问我哪些问题来检验我的专业水平?最少10个
* 我是一个养狗领域的专家，为了测试我的水平，你会问我什么问题来检验我的专业水平，然后可进一步追问-->这些我都知道，换一波更有深度、更有细节的。

还可以玩一些更高阶的游戏，比如你问我答的游戏模式。

* **问法3**：现在我们玩个“你问我答”的游戏，目的是测试我在XXX的专家（机器狗方面），你负责提问，我负责回答，你要根据我的答案进行反馈、评价、补充。如果我说不知道，你就直接输出正确答案，然后你继续提问，直到我提交不想玩为止。

### **4.2.2 针对第四个阶段**

如果你已达到融会贯通的水平，还可以这么玩：

* **问法1**：我已经很精通xxx了,我想知道我是否还有需要学习的地方?
  + -->进一步追问，这些我都懂，还有呢，还有呢。

通过这样的对话，可以进一步扩充你的认知范围，探索你觉得可以探索的地方。

## **4.3 知道和做到的推进**

此外，我们还可以再进一步，从知道向实现级别的做到推进，这里可以有3类提问方式可供参考：

* **第1类**： 知道自己要干什么，但是不知道其能帮你做什么。
  + 提问方式推荐：我想XXX（提升英语水平），你能给我提供什么帮助?
* **第2类**： 知道需要帮你做什么，但是你不知道给它输入什么具体的信息。（这里回到了之前的“元问题”）
  + 提问方式推荐：我想要你帮我做xxx，我应该给你输入什么信息?
  + 我想要你帮我写一篇精彩的演讲词 / 颁奖词，我需要给你输入什么信息。
  + -->根据反馈，请把内容进行结构化，再次提问，然后会得到想要基本答案，再在这个基础上进行相关改进
  + -->请用第三人称改写颁奖词，改成莫言的作品风格。
* **第3类：** 任务目标、背景信息、输出要求都非常清晰时，可直接给指令：
  + 请推荐（书名、电视剧、电影、网站、博主、专家、学习资源）等等。请翻译... 请总结... 请润色...，甚至可直接导入图片，然后提问图片中的文字内容。

也可进一步推进内容，例如下

* 请推荐5部儿童题材的高分电影，按照评分从高到低排列，并列出主演、评分、上映年份，以表格形式输出。
* 请推荐5个专注时尚穿搭领域的B站博主，按照粉丝量从高到低排序，并给出他们的B站频道链接。
* 请将以下内容翻译成英文，要求用高级的、优雅的词汇。（可用一些比较级的语言）
* 也可翻译成其他格式，比如：请将以下内容翻译成中国的文言文。

也可以把一些复杂的面试题给它做，然后顺着思路一起拓展**：**

* 有8个台球，其中一个比其他的7个都要重一些。如果仅仅是使用天平而不称出具体的重量请问最少几次能找出那个最重的台球?

再进一步，

* 有10个台球，其中一个比其他的9个都要重一些。如果仅仅是使用天平而不称出具体的重量请问最少几次能找出那个最重的台球?
* 如果你看到时钟上面的时间是1:05，那么其时针和分针之间的角度是多少?
* 3个5，中间可以加各种数学符号，包含根号平方阶乘等，最后得到25。

或是代码教学：

* 我想要编写一个chrome浏览器插件程序。让我在使用英文网页时，能通过bing翻译或是百度翻译，把页面翻译成中文，并把展示内容分割成一段英文加一段中文的方式。请教我整个步骤。（注：涉及实际的代码实现可能需要有一定的基础，但通过它可以大体搭建整个插件的通用框架及相关环境，也能产生不少的实际的生产力。）

## **4.4** COSTAR提问框架

在这里还没完，为了进一步优化提问，确保人工智能能够准确理解我们的需求，我们又参见了CO-STAR提问框架（背景信息C、目标O、风格S、语气T、受众A和回答R）。这是由新加坡政府科技局数据科学与AI团队创立的提示构建工具，它考虑了大语言响应模型响应效果和相关性的关键因素，能帮助我们获得更优的反馈：

1. 上下文（Context）：背景信息，即上下文。
2. 目标（Objective）：具体任务/目标。
3. 风格（Style）：回答时采用的特定写作方式。可模仿名人的风格，或符合特定行业标准的风格。
4. 语气（Tone）：回答内容的情感色彩，可以是正式的、幽默的、鼓舞人心的等。
5. 受众（Audience）：框定目标群体/受众。（比如本文的受众可框定为受过良好高等教育且有长文阅读能力的人群）
6. 回复（Response）：返回的格式或结构。可以是文本、列表、图表或其他任何形式。

参考例：

* 假设你是一名电梯公司的媒体运营，需要为公司的一款新型智能电梯，写一篇公众号文章作为宣传。

（如果你对电梯产品不了解，那我们又可以从上面介绍的“元问题”出发）

* 我要写一份公众号文章，用于介绍一个电梯公司的新型号电梯，我可以基于什么思维模型去撰写，给我推荐几个。

得到相关信息后，我们进一步套用COSTAR模型的提问：

* + 我想为我公司的新产品做广告。我的公司名叫上海三菱电梯，产品名叫LEHY-PLUS，是一种新型智能电梯。特点是支持屏幕定制，一屏千面；支持语音和触摸按钮；智能儿童关怀、防暴告警。超薄超小双曳引机驱动、直线电机门机体系、智能LED照明。# 上下文 #
  + 是一篇公众号文章，目的是让人了解新智能电梯的卓越性能和高可定制化和高性价比，产生购买意愿。# 目标 #
  + 效仿Telsa汽车的宣传文稿的写作风格，带有科技感。# 风格 #
  + 专业又富有感染力。# 语气 #
  + 有希望了解电梯产品的客户和潜在最终用户，以及潜在的家用别墅梯用户。# 受众 #

这里只是简要的进行提问框架套用，实际情况时，可以把相关格式进行进一步的扩充和精细化。相信通过一系列的提问训练和练习，我们会得到相对满意的回答。

# **5 更复杂的提问场景**

在前面的探讨后，作者觉得还不太够，为此，进一步提升了提问情景的复杂度，通过角色扮演和头脑风暴的方式，让大模型工具试着处理更丰富的场景，提供更深入的思考，展示更多元化的互动。帮助获取更具创意和实际应用价值的答案。

## **5.1 角色扮演**

角色扮演是通过模拟特定人物或群体的思维方式，来拓展思维的边界。通过扮演历史人物、知名人物、虚拟角色等，我们可以从多角度进行问题分析，并获得更有趣、丰富的反馈。以下是一些常见的角色扮演提问示例：

* **问法1**：假如你是《灌篮高手》里的樱木，你后来成功当选了日本队的主力 ，新的联赛即将召开，你会对新的球员说什么？之后还要进一步追问，例如：演讲稿里没有必胜信念的传承，没有同伴的激励，没有讲到晴子，请重新回答。
* **问法2**：如果樱木在高三终于获得了全国冠军，在夕阳下的小酒馆，遇到了三井，三井会说什么，樱木的第一反应是怎么样？-->可进一步追问，没讲到安西教练，请重新回答。
* **问法3**：如果你是巴菲特，咱们一起讨论世界的发展，未来10年，你最看好未来哪几个行业，现在我有1个小目标可以用来投资，你觉得我应该如何做好规划。
* **问法4**：如果你是莫言，请基于你的风格，写一着短诗，歌颂XXX（邓工）的伟大功绩。
* **问法5**：模拟一个心理学专业的教授，大白，是好朋友，我俩交流工作。大白，最近真是烦呀，工作怎么做都做不好。。。以此类推。

除了扮演具体的人，具体的性格，还可扮演一类群体，或是扮演几类人：

* **问法6**：如果你是一位年长的智者，经历过生活的跌宕起伏，希望跟小朋友谈谈一生中最后悔的几件事，你觉得会是什么?

## **5.2 头脑风暴**

在单角色扮演的基础上，试着加入多人场景，赋予其更丰富的应对场景，比如把包含虚拟的、活着的、已故的或是多个角色纳入同主题讨论的框架：

* **用例1：**先推荐5名世界级投资大师，最好包含中国人，比如雷军，沈南鹏，段永平等，并进一步提问他们是否有自己的代表著作
  + -->进一步：请详细介绍《彼得·林奇的成功投资》的核心内容和结论，并结合他的公开演讲、访谈，总结一下他的投资理念和思想。
  + -->再进一步，推荐5位著名学者，要求他们对中国近40年的经济发展有深刻的认知和洞察。尽量推荐了解中国大陆市场经济的学者。
  + -->最后，把自己的问题提出来：讨论主题是：在2025年-2030年之间，小李有1个小目标可投资金，住在中国上海，有什么好的投资行业、投资渠道、以及投资标的。
  + 要求：私董会讨论组成员，是前面提到过的:巴菲特、雷军、马斯克、段永平、马化腾、彼得·林奇。一共6个人，轮流进行讨论发言，先进行5轮讨论。要求发言内容代表发言人的投资理念和认知。

这种情况下，大概率会得出结论：听君一席话，那还是一席话。那怎么进行进一步的深化讨论，并能对实际场景作出有意义的见解呢，可以通过模拟专业类别的人群进行，按多个角色来进行头脑风暴式的提问，会得到比较精确的反馈：

* **用例2：**现在我们模拟一个问答场景：小李是毕业了10年的公司职员，当前有一次与公司管理者面对面访谈的机会。访谈对象包括他的直属部门经理、业务总监、公司领导。现在这3个人是小李的智囊团，小李可以轮流向他们提问。目的是如何帮助小李如何在当前的岗位上成长，进一步提升自己能力和认知，获得更好的职业发展。（实际背景可自行扩充）。请问，小李应该轮流向这3个人问哪些有价值的问题。
  + -->继续，还有吗
  + -->基于以上各领导的所有回答，现在请分别基于这3个人的不同视角，给小李制定一份为期1年的IDP，IDP分4个阶段，每个阶段3个月。要求明确列出，IDP每个阶段的目标、具体的任务要求、以及可量化的评价指标。
* **用例3：**我现想设计一个帮助家人快速整理家中物品清单的App，主要功能是实现家里物品的出入库及物品年份清单，包含且不限于电器类的周期性运维，物品品牌等详细信息，定期食品出清等，请你分别模拟各个角色，互联网产品经理，研发工程师、项目经理，提供有创新价值的设计思路。逐个发言，进行10轮讨论。
  + -->再改进下，把家里群体分成知识层次较高的人城市人群，接受到基本教育的城市人群，以及农村受教育人群，农村基本没受过教育的人群，
  + -->还有，物品可包含私人物品和可公开的物品，互相认识的人之间会产生交集，希望能产生新的社交和互动，产生新的商业模式，请上面的5拨人，继续围绕app的创新设计思路讨论。
  + -->现在，我们暂停App的功能设计讨论，转而讨论App的运营问题。必须考虑到App可行的商业模式，如何收费，如何获利。请站在不同的角度继续讨论可能的方案，也包括他们各自接受的付费方案。请继续讨论。
  + -->好，基于以上的功能和商业模式讨论，我们不可能在App一上线的时候就实现所有功能。请结合前面的讨论，给出合理的开发节奏。分为3个大的版本去选代，每个版本周期为1年，为期3年，请给出详细的开发计划，并估算实现相应功能需要的研发工程师工作量。
  + -->如果第三阶段实现后，有3亿用户。结合预估的功能和用户沉淀的数据，请评估下每年的服务器费用，包含但不限于域名、安全、CDN成本和存储成本等。
  + -->结合前面的讨论，现在市面上有类似的产品或者服务吗?
  + -->接下来我要做一个全面的市场调查和竞品分析，请给我一个完整的调研分析大纲。

通过上述复杂场景的模拟和角色扮演，我们不仅可以从多个角度深入分析问题，还能预测不同角色的反应，从而获得更为多元化的解决方案。这种方法能够有效拓展思维，帮助用户从全新的视角审视问题，进而获得更加精准且具备实际操作性的建议。

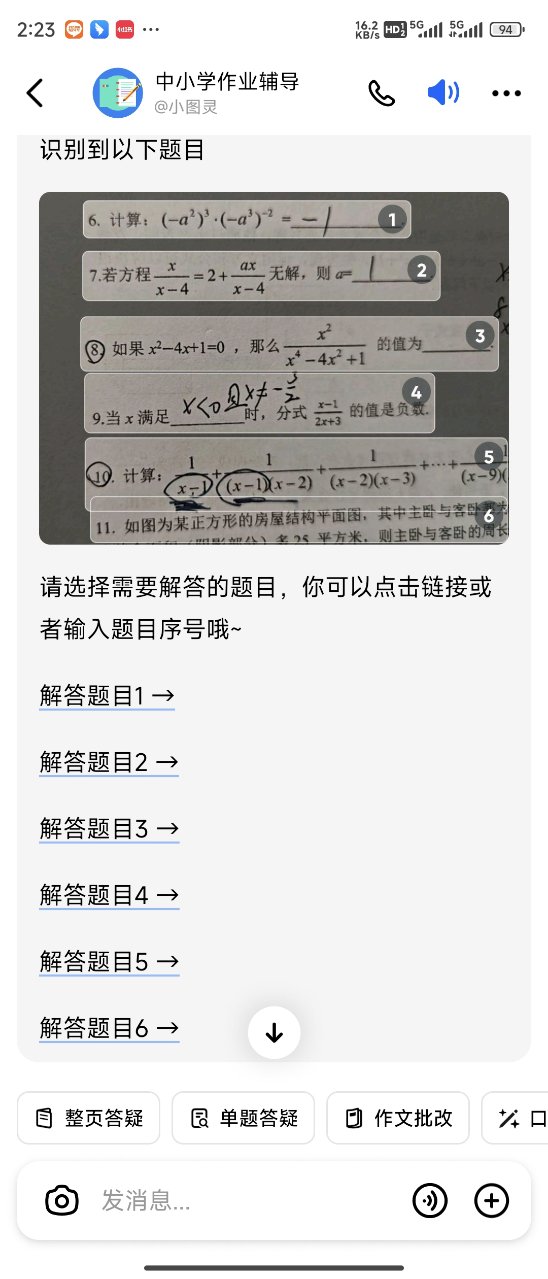
# **6 其他日常AI场景**

在日常生活中，AI技术的应用已经渗透到了各个领域，帮助我们高效处理各种繁琐任务。以下是一些典型的应用场景，包括微信翻译、图片文字识别、语音识别以及作业辅助工具等，这里给出几个常用的推荐工具及使用场景。

* **微信识别**
  + 在国外旅行过程中，你可以用微信直接扫描餐厅菜单、街头广告或路标上的外语文本，快速获得翻译结果，解决俚语类复杂文字的识别。

* + 在办公中，拍摄文档或会议记录后，微信可自动提取文字，便于快速记录或后续处理。
* **百度云盘的语音识别**
  + 会员账号下的百度云盘语音识别，有强大的识别能力，并能进行分类汇总，相比其他的语音识别工具，是让作者惊艳的产品，特此作为给读者的推荐。
* **豆包拍照解题**
  + 能快速解决日常生活中的学生作业辅导，详细的过程和基本准确的回答，能大大提升效率，改善家庭环境，提升家庭幸福。

* **蜜蜂试卷**
  + 一键去手写字，错题成卷，也是老父亲和老母亲的心头好。

等等这些，我们看到了AI技术在生活领域的各种应用。随着相关技术的发展，未来将会有更多的创新场景涌现，使我们的日常生活更加便捷和高效。

# **7** 跨模态扩展——**全领域的应用**

篇幅问题，介绍大体也到此，但大语言模型的应用不仅能够提供文字层面的反馈，还在多个领域和形式中为我们提供各种支持，涵盖的一系列主要场景包含如下：

* 文本text to text
* 代码text to code/code to text
* 音频text to speech/speech to text
* 图片text to image/image to text
* 视频text to video
* 办公text to PPT/APP

以下是两上图生文和文生视频的参考例：



* 提示词：这是张什么风格的图片，里面主要描绘了什么场景



* 提示词：请识别出图片中出现的所有物品清单

[请至钉钉文档查看附件《科幻广角镜头拉镜头高质量电影风格展现地球在宇宙中的孤独旅程，缓缓穿越星空，周围是漂浮的陨石和太空碎片。画面以地球为中心，地球上又多个推进器推进着地球在宇宙中航行，周围环境安静而神秘，太空的寒冷与孤独感透过屏幕传递。影像采用低饱和度色彩，强调宇宙的广袤与宁静。.mp4》](https://alidocs.dingtalk.com/i/nodes/AR4GpnMqJzMLAONPHoDjblnaVKe0xjE3?cid=57217525816&corpId=dingb41cd8fb919e5a6b&doc_type=wiki_doc&iframeQuery=anchorId%3DX02m565240y93s7h67zplj&utm_medium=im_card&utm_scene=person_space&utm_source=im)

* 提示词：科幻广角镜头拉镜头高质量电影风格展现地球在宇宙中的孤独旅程，缓缓穿越星空，周围是漂浮的陨石和太空碎片。画面以地球为中心，地球上有多个推进器推进着地球在宇宙中航行，周围环境安静而神秘，太空的寒冷与孤独感透过屏幕传递。影像采用低饱和度色彩，强调宇宙的广袤与宁静。

（注：需要通过主体&动作、镜头&动作、场景&美感进行镜头描述）

[请至钉钉文档查看附件《在健身房里，一个中国长发美女，穿着明黄色的健身服在跑步机上跑步，45度前脸视角，中景镜头，慢动作镜头，白天室外的光自然照进来，写实专业的光线.mp4》](https://alidocs.dingtalk.com/i/nodes/AR4GpnMqJzMLAONPHoDjblnaVKe0xjE3?cid=57217525816&corpId=dingb41cd8fb919e5a6b&doc_type=wiki_doc&iframeQuery=anchorId%3DX02m568ecmpziyeks2jcmq&utm_medium=im_card&utm_scene=person_space&utm_source=im)

* 指令：在健身房里，一个中国长发美女，穿着明黄色的健身服在跑步机上跑步，45度前脸视角，中景镜头，慢动作镜头，白天室外的光自然照进来，写实专业的光线。

# **8 结论**

未来已来，大模型工具已经渗透到各行各业，在潜移默化中已能助力我们剖析并解答日益复杂的问题，提供多元化的解决方案。从如何提问到掌握不同情境下的提问技巧，大模型工具的应用让我们能更加高效地利用其智能力量，回答我们关心或是认知外的内容，为我们提供更精准、全面的建议，帮助我们理性评估和思考。此外，它还在文本、代码、音视频等多种领域不断拓展认知和创意的边界。

然而，使用过程中也不可忽视一些潜在的挑战与风险。尽管提供了前所未有的便利，但可能受限于数据偏差或模型训练的局限性，其输出的答案并非总是完全可靠。同时，过深的依赖也会影响独立思考和创新的深度。更重要的是，随着人工智能技术的普及，如何确保其公正性、透明度和道德性，仍是一个亟待解决的社会问题。

总的来说，享用技术进步带来的便利，并保持警惕，了解其潜在的风险，才能最大限度地发挥其正向作用，成为我们高效解决问题、推动创新和探索未知的可靠伙伴。

# **9 推荐的开源参考项目及自搭建案例（已亲测）**

1. Ollama：开源的LLM服务工具，开源、免费、本地部署、支持windows系统，结合docker的WebUI即可使用，<https://github.com/ollama/ollama>
2. GLM-4-XB：国货之光，可适配Ollama，<https://github.com/THUDM>，linux环境服务器部署供参考
   1. [《01-私人部署开源GLM-4 Demo全过程》](https://alidocs.dingtalk.com/i/nodes/R1zknDm0WR3eEzv0SajZM1a9VBQEx5rG?utm_scene=person_space)
   2. [《02-私人部署开源LangChain全过程》](https://alidocs.dingtalk.com/i/nodes/93NwLYZXWyglArjmIoKjgOmdJkyEqBQm?utm_scene=person_space)
3. QwenLM：阿里千问2.5，可适配Ollama，<https://github.com/QwenLM/Qwen2.5>
4. SD：stableDiffusion：文生图经典体系，linux环境的部署过程供参考：
   1. sd部署：[《03-私人部署开源stable diffusion全过程》](https://alidocs.dingtalk.com/i/nodes/14lgGw3P8vvlAwDYHoxDw27o85daZ90D?utm_scene=person_space)
   2. sdWebUI部署：[《04-私人部署开源stable diffusion webui全过程》](https://alidocs.dingtalk.com/i/nodes/QG53mjyd80Rj9gvmCE0pooz6V6zbX04v?utm_scene=person_space)
5. SoVITS：声音克隆工具，TTS应用。<https://github.com/RVC-Boss/GPT-SoVITS>
6. awesome-chatgpt-prompts：专用的gpt提示词示例集合，115k的stars足以证明其被认可的程度。