【知识福利社】你的私人书院，微信：312005650 全员热门课程众筹共享！

14 实战：如何拥有你的AI助理？

万维钢 AI前沿
2023年4月15日

进入课程 >

时长：12:05分 大小：8.29MB

00:00 / 12:05

转述：怀沙

你好，这里是《万维钢·AI前沿》课。这一讲咱们来聊聊，如何拥有你的AI助理。

我们课程前面说过GPT有个痛点，就是它不能精确处理比较繁杂的数学计算。其实它还有个众所周知的缺点是作为一个语言模型，它的训练语料是截止日期的。比如GPT-4的语料截止到2021年9月，这就使得它没有在此之后的新知识；而且它的语料早竟是有限的，它不具备所有的知识，有时候爱胡编乱造。

这些问题都可以通过调取外部信息的方式解决——OpenAI正是这么做的。3月23日，OpenAI突然宣布ChatGPT的几项重大更新：

* 一个是可以上网的GPT，这就解决了实时获取最新知识和信息的问题。

* 一个是可以直接调练编程的GPT。ChatGPT给提供了一个虚拟机，相当于是一个安全环境，GPT可以在里面编译和运行程序，这就让编程更方便了。

* 一个是推出了安装第三方插件的功能。

目前这些都处于测试状态，只对少量用户开放。我一看到新闻就提出申请，目前只得到了插件功能。

ChatGPT的插件商店已经有11家应用——

包括订飞机和酒店的Expedia、几个购物和订餐的应用、特别是有一个能调用其他网站应用的应用叫Zapier。

我最感兴趣的是它提供了Wolfram——也就是上一讲说的沃尔夫勒姆的公司的应用插件，因为这就解决了GPT做数学题的问题。这也是沃尔夫勒姆在他书中的设想，把ChatGPT和Wolfram语言结合起来用。不过沃尔夫勒姆本来想的是在他自家网站做一个ChatGPT的插件，结果是他的东西变成了ChatGPT上的插件……

这个插件搞法超出了很多人的预料。我本来也以为用AI就是各家公司调用OpenAI的API，把AI引入到自家产品中，就像微软的Bing：你还是去Bing网站搜索，只不过你可以在Bing网站上用GPT。现在看来，那只能算是「应用解决方案」——直接在ChatGPT上装插件，才是「系统解决方案」。

这个意义在于，ChatGPT成了一个统一的入口，通往各种应用。有人打比方说，ChatGPT就相当于是苹果手机，插件就相当于苹果应用商店里的应用——这么说不太准确，你打开手机还得再打开应用才能工作，而ChatGPT则让你只跟它说话就行，你可以忘记插件。

这次更新让GPT增加了三方面的能力。

*

第一，AI有了调用力。

我在精英日课专栏中提过，Meta推出了一个算法叫Toolformer，允许语言模型调用外部工具。OpenAI似乎并没有直接用那个方法，它的办法是把你的各种需求翻译成外部应用能听懂的语言，从而调动那些应用帮你办事。

以前你想办什么事，得考虑去哪个网站、用什么工具，还得知道具体用法。现在你不用管那些了。你有事儿直接告诉GPT，它看看适合什么工具就调用什么工具。你自己不需要记东，你可能都不需要知道它用了什么工具……AI是你的助理。

而且你只需要这一个助理。

你要订飞机票，把时间地点告诉ChatGPT，它会去查询Expedia。你要订餐，它会先调用OpenTable帮你选定餐馆和菜单，再调用Wolfram给你计算这一餐的卡路里。你要给准发个什么邮件，它会先帮你把邮件写好，你过目后说可以，它就调用Zapier，Zapier会自动访问你的邮件账号，帮你把邮件发出。

你要做的只是提需求和确认准许，一切技术性操作都有AI领着一帮插件帮你完成。

*

第二，GPT能实时浏览互联网了。我没拿到这个功能的内测，但我看了Twitter上好几个人的演示，觉得体验比Bing Chat要好。你可能需要重新思考「上网」这个行为。

有人问GPT本届奥斯卡的最佳男演员、最佳音效、最佳影片和最佳男配角都是谁，并且要求它把这四者用一首诗串联在一起。

你可以看到GPT整个行动过程。它先去搜索了相关信息，像人一样有选择地点击了几个网页，迅速读完后，根据搜集到的信息写了一首诗。

你想想这意味着什么。我们所谓上网搜索，其实浏览网页不是目的——我们的目的是用网页中得到的信息做事情！现在GPT提供了一站式服务，先替你浏览两帮你做事。

还有个例子，你看看GPT的综合调研能力。有人问它，特斯拉在过去四个季度中的收入都是多少？

GPT先搜索了“特斯拉过去四个季度的收入”，可能没有找到全面的信息；它就文专门搜索2022年第二季度和第三季度的收入……这样几次搜索之后，它调读了相关网页内容，最后给你提供了简洁明了的四个数据——

复杂的浏览，一次找不到就再找，然后生成直指要点的信息，这是最干净利索的调研。

还有个人是这么干的，他让GPT去浏览OpenAI的网站以及博客，看看今天发布了什么重要新闻。GPT就去了，点击了几个网页，阅读之后，把新闻总结成了三段话返回。

这个插件搞法超出了很多人的预料。我本来也以为用AI就是各家公司调用OpenAI的API，把AI引入到自家产品中，就像微软的Bing：你还是去Bing网站搜索，只不过你可以在Bing网站上用GPT。现在看来，那只能算是「应用解决方案」——直接在ChatGPT上装插件，才是「系统解决方案」。

这个意义在于，ChatGPT成了一个统一的入口，通往各种应用。有人打比方说，ChatGPT就相当于是苹果手机，插件就相当于苹果应用商店里的应用——这么说不太准确，你打开手机还得再打开应用才能工作，而ChatGPT则让你只跟它说话就行，你可以忘记插件。

这次更新让GPT增加了三方面的能力。

*

第一，AI有了调用力。

我在精英日课专栏中提过，Meta推出了一个算法叫Toolformer，允许语言模型调用外部工具。OpenAI似乎并没有直接用那个方法，它的办法是把你的各种需求翻译成外部应用能听懂的语言，从而调动那些应用帮你办事。

以前你想办什么事，得考虑去哪个网站、用什么工具，还得知道具体用法。现在你不用管那些了。你有事儿直接告诉GPT，它看看适合什么工具就调用什么工具。你自己不需要记东，你可能都不需要知道它用了什么工具……AI是你的助理。

而且你只需要这一个助理。

你要订飞机票，把时间地点告诉ChatGPT，它会去查询Expedia。你要订餐，它会先调用OpenTable帮你选定餐馆和菜单，再调用Wolfram给你计算这一餐的卡路里。你要给准发个什么邮件，它会先帮你把邮件写好，你过目后说可以，它就调用Zapier，Zapier会自动访问你的邮件账号，帮你把邮件发出。

你要做的只是提需求和确认准许，一切技术性操作都有AI领着一帮插件帮你完成。

*

第二，GPT能实时浏览互联网了。我没拿到这个功能的内测，但我看了Twitter上好几个人的演示，觉得体验比Bing Chat要好。你可能需要重新思考「上网」这个行为。

有人问GPT本届奥斯卡的最佳男演员、最佳音效、最佳影片和最佳男配角都是谁，并且要求它把这四者用一首诗串联在一起。

你可以看到GPT整个行动过程。它先去搜索了相关信息，像人一样有选择地点击了几个网页，迅速读完后，根据搜集到的信息写了一首诗。

你想想这意味着什么。我们所谓上网搜索，其实浏览网页不是目的——我们的目的是用网页中得到的信息做事情！现在GPT提供了一站式服务，先替你浏览两帮你做事。

还有个例子，你看看GPT的综合调研能力。有人问它，特斯拉在过去四个季度中的收入都是多少？

GPT先搜索了“特斯拉过去四个季度的收入”，可能没有找到全面的信息；它就文专门搜索2022年第二季度和第三季度的收入……这样几次搜索之后，它调读了相关网页内容，最后给你提供了简洁明了的四个数据——

复杂的浏览，一次找不到就再找，然后生成直指要点的信息，这是最干净利索的调研。

还有个人是这么干的，他让GPT去浏览OpenAI的网站以及博客，看看今天发布了什么重要新闻。GPT就去了，点击了几个网页，阅读之后，把新闻总结成了三段话返回。

你看这几个例子，还是你不用动手。你不用自己去阅读那些网页，你甚至都不用搜索和点击——GPT觉得需要搜索就搜索，它替你完成所有调研工作，最后返回给你的是最终结果。

搜索引擎只是望远镜，GPT则是供察兵。

*

第三个能力升级是GPT终于有“计算器”了……其实我是想通过GPT对Wolfram插件的调用再给你演示一下什么叫「自然语言编程」。

我把底牌与命门这一讲中的数学题再交给GPT，它二话不说直接调用Wolfram插件，这回的答案是绝对正确的。

1231*434523+323*34636

Used Wolfram

计算结果为：

$$1231 \times 434523 + 323 \times 34636 = 546,085,241$$

它还可以解方程、画函数曲线图、算积分，包括各种更复杂的数学，那些不必细说。

Wolfram网站有一项被沃尔夫勒姆多次鼓吹的功能是它有关于当今世界的各种精确的实时数据，ChatGPT也非常理解这一点。

我先向不带插件的GPT-4提问：从美国波特兰到中国贵阳的距离是多少？它回答11300公里。

你看我既没有自己去搜索数据，也没有做任何编程，我只是说话，而且还是中文。试想如果GPT彻底解决了跟Wolfram以及各种公共数据库的对接问题，你想看什么数据想怎么画图都是一句话的事儿，你还可以随意把各种数据放在一张图上作对比，你的思考效率将会提到多少？

还有个挺有意思的经历。我先让GPT画三个圆，代表三个集合，大圆中有两个小圆，两个小圆有一部分重叠。它画出来了——

我立即要求修改，去除图形的坐标，给三个圆加上标题，结果它画出来是这样的——

标题被放在了圆的中间，没法读。我又要求它把标题放在各个圆的上方，它也做到了。

然后我接着改，又提出如下要求——

* 给三个圆分别填充蓝色、红色和黄色

* 圆的边界线都是黑色，上述三种颜色用于圆的内部

* 颜色做透明处理

* 标题字体稍微大一点

它全做到了，生成了最终用的图——

这个过程的有意思之处在于一开始我并没想好图应该是什么样的。我是在互动中一边看结果一边有新想法一边修改设计，而GPT完全理解。它每一次都把我的要求完整翻译成Wolfram能理解的数学语言，交给插件，返回结果——

你可以看一眼Wolfram实际说的指令——

{

"input": "Graphics[{EdgeForm[Black], Opacity[0.5], Yellow, Disk[{0, 0}, 2], Red, Disk[{-0.5, 0}, 1], Blue, Disk[{0.5, 0}, 1], Text[Style[\" 一 切 计 算 \", FontSize -> 20, Opacity[1]], {0, 2.2}], Text[Style[\" 神 经 计 算 \", FontSize -> 16, Opacity[1]], {-0.5, 1.2}], Text[Style[\" 形 式 逻 辑 \", FontSize -> 16, Opacity[1]], {0.5, 1.2}]}, Frame -> False, Axes -> False}"

}

以前要画这样的图，我就得自己写这些指令才行。我要是不熟悉这种语言就得查手册或者上网搜索，可是如果那么费力你说我还能有多大心思画这张图？

*

ChatGPT正在演变成你的AI助理，它对我们的解放是双重的。

首先它去除了做事的障碍。其实很多人说「不会做」什么事，并不是真不会，只是有个界面障碍。

比如如果让你去瑞典生活，你可能说我不知道怎么在瑞典生活，因为我完全不懂瑞典语，甚至不懂英语——但作为一个成年人，你本来可以在任何一个地方生活乃至工作，你需要的只是打破语言这个界面。

一位老奶奶想给孙子买个礼物，那东西只在网上有卖，她说我不会买——她明明懂得什么是好东西，懂得货比三家，怎么能说不会买东西呢？她只是不会上网、不懂电子支付而已。

世间有多少所谓的不会，其实是界面问题？

AI助理把所有界面回归到了自然语言。你只要懂得事情的本质，然后会说话就行。它负责把你的意思翻译成各个应用所能理解的语言，你无需看见那些应用，你甚至无需知道那些应用都是什么……AI助理把我们从界面背后解放了出来。

而对平时没有界面问题的人来说，更大的一个解放是效率。很多事你知道怎么做，但是你一想到它做起来很麻烦就不想做了。实东西得去超市，画个图还要编程，写文章需要做调研……阻碍你的不是大思想而是这些技术细节。现在你可以把细节都交给AI助理。

最近我家的打印机有点不好使，想买个新的。因为之前国了几个墨盒，不想浪费，所以我想买一个能用同款墨盒的打印机，我还希望它的型号比较新，并且有AirPrint的功能。我把这些需求直接告诉Bing Chat，它去搜索了一番，帮我找到了三款打印机。我又问价格分别是多少，它又列出了相应的价格区间。

总共不到一分钟，我就选好了。而不久之后，ChatGPT会用插件把这样的事做得更漂亮。

AI允许我们做看不见的计算，或者说是无需看见的计算。

好，以上就是这一讲的全部内容。从这一讲的演练中可以看出，你只需要提供意图就行；而你的意图表述得越清楚，它完成得就越好。

下一讲，咱们来说说怎么用好ChatGPT，有哪些咒语心法。

本课取材自：
万维钢·精英日课5
帮你和全球精英大脑同步
万维钢 · 科学家
“相当于我们的AI助理”
划重点
添加到笔记
GPT升级了三方面的能力。
第一，AI有了调用力。
第二，GPT能实时浏览互联网了。
第三，GPT终于有“计算器”了。
首次发布：2023年4月15日
上一节 13 实战：怎样用ChatGPT编程？
下一节 15 实战：ChatGPT有哪些咒语心法？
可免费赠送给20朋友
微信
朋友圈
用户留言
最新
最热
只看作者回复
关注
老毕 昨天
插件机制像是GPT手和脚，正在快速延伸能力的GPT正在重构互联网的入口逻辑。我们只接触用户界面，界面后的黑箱里满是重新组织的、高度智能的、不断丰富的功能模块，多数人不再需要知道那些功能如何被做到，甚至不再知道那些功能的存在。
这样推演下去，应该被强化的能力是思考并得出“我要做什么”——这和现行的主流教育相悖——主流教育的多数优胜者更擅长“告诉我要做什么，我努力做好”。
教育可能会分化，GPT 完全可以动态评估每一个人类个体的教育水平、知识结构、GPT素养（与 GPT 协作的能力等级），如此一来，通过中短期冲刺的方式在传统知识性选拔中（比如高考）胜出然后努力进入爵位式分配制度的获胜手段，就不再奏效了，因为我们不需要再通过这种形式的选拔来甄别人才，它太低效了。基于GPT已经展现的内容生产能力和即将展现的协作连接能力，用不了多久，就会出现效率更高的甄别方式，更棒的是，它可以把人类个体的差异填平。其实，一直以来，我们需要的都不是“这个也行那个也行”的全才，我们更需要那些“这个不行那个很行”的专才，以前做不好甚至做不到，甄别的成本太高了，我觉得以后是真行了，当个性化选拔能够便宜到和整体性选拔没有差别的时候，高考这种人才甄别动作就可以每分钟都搞上几万次，因为你和GPT的互动，本身就是在不停地学习和检验，现学现考，随学随考，考完立马分配任务，任务完成立马结账，稍作休息，就可以再次主动选择或者被动分配下一项生存任务。
这些可能性正在发生，想想都让人觉得开心。
收起
转发 评论 赞 分享
进化笛 昨天
关注
期待 AI 助理的到来！
回过头来想，如果人人都有一个 AI 助理，想要做什么都问一下，然后自己来决策，这样到底会带来什么呢？我们或许可以做更多事情了，但所有人都能做，我们会做些什么才显得自己的独特呢？
GPT学习小组
转发 评论 2 分享
遥望十米开外 昨天
关注
一个通用的界面，可以返回我们所需的“所有”信息。这背后的GPT引擎，帮助我们处理的所有。这是趋势，也是必然，且已经来临。
转发 评论 1 分享
Fiona(我的窒息是SJ) 昨天
关注
很受启发。我感觉这段时间会源源不断出现更多插件甚至整合得更好更好更好的工具出现，进一步解放生产力🙏期待万老师继续为我们解锁🙏
转发 评论 赞 分享
封印的瘦女 昨天
关注
感觉自己将会拥有一个万能助理，开心
它轻松搞定我现在做的工作，泪丧
我改怎么做才能不被它替代？忧虑
赶紧学习，不能落后！振奋
转发 评论 赞 分享
得到用户 昨天
关注
感觉一场思维革命已经到来，跟不上形势就会被淘汰。
转发 评论 6 分享
以上留言由作者筛选显示

【知识福利社】你的私人书院，微信：312005650 全员热门课程众筹共享！