



懂图片的 GPT-4 来了，国内 AI 企业如何接招？



作者：盘和林 浙江大学国际联合商学院数字经济与金融创新研究中心
联席主任、研究员，工信部信息通信经济专家委员会委员

3月15日，OpenAI 发布了 GPT-4，并迅速成为热点，引发大量关注。GPT-4 核心提升可总结为两点：一点是基本智能水平提升。举例说，GPT3.0 在全球多项考试中排名在倒数 10%，而如今 GPT-4 在这些考试中能够战胜 90% 的人类，这种进步源自于数据训练和参数的积累。另一点是多模态模型，ChatGPT 过去只能文本交流，现在用户可以发张图片给 ChatGPT，而 ChatGPT 将以文字回应。另外，GPT-4 的参数量是 GPT3.0 的 10 倍，参数达到百万亿级别，这为 GPT-4 的用户体验优化提供了保障。

之前上市的 ChatGPT 实质上等同于 GPT3.5，未来，GPT-4.0 的功能将以 plus 订阅的方式开展。新版本激发了用户热情，付款订阅渠道一度被挤爆。但实际上 OpenAI 在技术仍有保留，外界不经感慨：OpenAI 到底领先对手几个身位？

之所以说 OpenAI 保守，主要是因为两个原因：

其一是 ChatGPT 在去年年底推出的时候本身是旧版本，很多人发现 ChatGPT 无法回到 2021 年以后的问题，原因是 ChatGPT 2021 年就发布了 GPT3.0，而 ChatGPT 只是在 GPT3.0 的基础上略做优化的 GPT3.5 版本，而 2021 年以后新数据都去训练 GPT-4，当然就无法覆盖这部分新知识。而 GPT-4.0 其实在 2022 年年中就基本完成，组建 ChatGPT 的发布是一个保守做法。

其二是 GPT-4 也是有保留的。原因是 OpenAI 有另一个热点 AI 产品，也就是 DELL-E2。DELL-E2 的功能是你输入一段文字，他生成一张图片。而 GPT-4 的功能是，你输入一张图片，他生成一段文字答案。如果两项 AI 功能结合，GPT 将同时获得图片理解能力和图片表达能力。可以说 NLP 自然语言处理只是 OpenAI 在人工智能领域的一块业务，未来整合旗下的多个 AI，功能内容还将进一步丰富。

当前 GPT-4 的识图能力应用前景也非常广泛。比方说医学影像资料本身是图片，未来普通人可以通过 ChatGPT 来实现医学影像自检，自己发现疾病，当然，这只是其中一项可能性，在很多行业领域，看懂图片是需要专业能力的，而 GPT-4 降低了这种识图门槛，比如医学影像、设计图纸、数学模型等等，让普通人能够通过 AI 辅助介入到专业领域当中，所以 GPT-4 未来很可能成长为生产力工具。

中国大模型生成式 AI 企业怎么办？

首先，国内生成式 AI 企业要走积累的路子。生成式 AI 需要持续大规模算力投入，微软为 GPT 提供超过 3 万张 A100 芯片的算力支持，一张 A100 价格要 20 多万人民币，而算力训练每次都要消耗大量能源。GPT-4 百万亿的参数也需要投入大量人力财力来完成积累。而在数据方面投入的价值也很高，GPT 从一开始就重视“小数据”，也就是高质量数据，而高质量数据是 GPT 领先于其他生成式 AI 的核心竞争力，而与之相对应的，数据质量提升是需要钱的。综上可以看出，算力、数据、算法调参这三个

方面，OpenAI 前期烧钱的规模是海量的，而 GPT-4 也是一步步技术积累的最终结果，所谓台上一分钟，台下十年功，一点也不夸张。

其次，国内生成式 AI 企业要在基础科学上做积累。大模型算法的道理并不复杂，复杂的是算法细节，自从微软收购 OpenAI 后，OpenAI 的算法不可能开源，所以“等、靠、要”的做法已经行不通，国内 AI 企业要在算法技术上投入研发，构建中国 AI 企业自身的算法能力，然而算法的精益求精也要依靠基础科学的进步，比如基础数学研究。所以要积极在学术上加强投入，而不应该仅仅聚焦在应用科学层面。

再者，国内生成式 AI 企业不要盲目的推出对标产品。前车之鉴就是谷歌，在 OpenAI 推出产品后，谷歌匆忙应战，结果发布会的若干个小瑕疵导致谷歌股价暴跌。国内生成式 AI 要清楚的认识到的，这一次 ChatGPT 的成功和过去其他概念有所不同，这次 ChatGPT 的成功基础在于用户的认可，用户是会拿着 ChatGPT 和其他竞品做比较，这就让后来者实际上更难做，要求也更高。如果生成式 AI 方面没有达到 ChatGPT 的效果，则往往会暴露自己能力上的短板。而反过来说，国内生成式 AI 企业也可以另辟蹊径，

预览已结束，完整报告链接和二维码如下：

https://www.yunbaogao.cn/report/index/report?reportId=1_53373

