

外热内冷

2023世界人工智能大会的开幕式上，华为轮值董事长胡厚崑说，通用人工智能正带领我们走向下一个黄金十年。

几乎是同一天，网络分析公司Similarweb发布的数据显示，6月ChatGPT全球访问量出现了推出以来的首次环比负增长，降幅达到9.7%。

尽管访问量可能并不具有实际意义，但却几乎瞬时掀起了人们对于ChatGPT流量见顶以及大模型降温的探讨。于资本层面，马千里称，一级市场讨论火热，但并没有特别多的机构真正参与了这个领域的投资，“实际情况远不如当初团购、共享单车的时候”。

AIGC的故事“编不下去”了吗？今年6月，昆仑万维实控人前妻李琼减持的消息，一度带崩整个AI板块。同样的情况还有三六零集团创始人周鸿祎的“90亿”分手费。

余欢称，在最近的这些大模型相关宣传中，有一些比较重要的概念被混淆或者弱化了。如果单单只是以参数达到百亿级别来衡量的话，那么GPT2到GPT3之间的2020前后就已经是大模型了。但这一波“生成式AI大模型浪潮”的重点在于智能涌现能力、多轮对话以及更强的生成能力。这些才是催生新一波生成式AI浪潮或者说LLM大语言模型浪潮的核心要素。

自今年3月，在全球科技大厂中首

发大模型产品文心一言以来，百度对于底层的大模型做了数次的迭代和升级。今年6月，文心大模型从3.0版本升级到3.5版本。百度当下的重点任务，也放在了持续优化生成效果方面，例如优化多轮对话的记忆能力等。

对于如火如荼的“百模大战”，余欢断言，在通用大模型领域，至少往前看两年以上，如果一家公司没有相关领域布局的信息，而是在当下突然宣布自己有大模型产品了，多半都是“套壳”或者接入了其他人的模型，通用大模型产品是很难一蹴而就的。绝大部分的公司和团队都是在“投机”，毕竟资本市场还是认概念、认故事的，而当下的A股市场又恰恰缺乏好标的和新概念。

商业及战略管理专家、她力量创始人&CEO卢步云也对北京商报记者分析称，AI技术在实际应用中的推广和落地，并不是单纯数据整合和效果呈现的事情，而是需要充分考虑业务需求、数据质量、模型定制等与商业模式相结合。除了模型要喂养学习外，现在AI技术人才也相对稀缺，企业同样需要投入大量资源培养。

“其次，AI技术的应用仍处于初期摸索，并未达到成熟阶段，当中存在很多的不稳定性和误差，需要不断优化和改进。另一个关键点在于AI技术的应用和数据处理，涉及隐私和安全风险，需要谨慎对待，并待有关的政策和方案出台。”卢步云称。

时间、周期与成本

大浪淘沙、优胜劣汰是市场经济的规律，对大模型来说仍是如此。

马千里分析称，当企业开始谈场景、谈应用的时候，反映出来的其实就是希望尽快将其商业化，因为大模型的开发成本太高了。人们普遍都已经意识到，宛如当年的“百团大战”，真正能够活下来的大模型可能只有个位数。

在他看来，当下大模型竞争早已超过了技术的范畴，更多是一种生态层面的竞争，具体表现在大模型基础上有多少应用、有多少插件、有多少开发者以及有多少用户。“如果没有用户使用，模型就没有数据进行提升，进而形成一种恶性循环。”

大模型不再唯技术论。余欢也提到了类似的观点，技术本身是门槛，是产业落地的充分不必要条件。大模型的to B服务需要一套非常成熟的运作体系，企业服务领域的生意三分拼的是技术，七分拼的是综合实力，比如长期服务和运维保障能力。

更重要的是，当产业落地真正实现规模化后，大模型头部效应将会越来越明显。“客户的选择通常会聚焦在几个主流模型之内，并着重参考同行友商的选择。PoC阶段并不会把所有模型都试个遍，费钱费力，且不现实。这意味着大模型创业企业的机会可能被进一步压缩。”余欢称。

从投资人和厂商的角度，华院计算数智人商业化副总裁林莱尼对北京商报记者分析称，过去半年的大模型就像是Gartner曲线一样，经历了从惊喜、盲从到趋于理性的过程。这种理性表现在商业变现的时间、周期、成本上。

“不管是投资人还是厂商，我们共同关注的是什么时候能够变现，变现的规模营收是多少。具体的应用场景在哪里，谁会为此付费。而商业应用方面，我们更加关注的是容错率，因为有些行业错误的回答会引发相当高昂的赔付成本。”林莱尼称。

百度集团科技与社会研究中心主任

余欢

大模型产业化 不在训练在推理

Q: 您曾提到，当下的大模型正处于第一波海选结束、正赛开始的阶段，那您认为正赛期间大模型又将经历一种怎样的竞争？

A: 正赛依然是个不断淘汰的过程，而且这个完整的正赛可能要持续5-10年以上。未来半年到一年，很有可能看到当初一窝蜂闯进来的投机团队被甩开，经历破产清算或者被收购。

近期圈内一种比较热门的观点是，选大模型要选大厂而不是创业公司，这其实反映出个很现实的问题：在技术层面，创业公司的大模型即便在某个方面或者某个时点达到了大厂水平，但在未来的竞争中还是难以胜出。

压力一方面来自于资金。大模型是一个需要持续烧芯片的重投入行业，对于创业公司而言，如果没有做好长期、每年数十亿甚至上百亿的投入准备时，最好不要轻易涉足。除通用数据外，很多专有数据集或者商业数据集，需要耗费大量资金购买。在技术不相上下的前提下，如果数据获取不占优势，也很难追上大厂的“进化”速度。

另外一方面是人密度。大厂有着成建制的算法开发和工程交付人员。此外还有大量的运维保障人员。而创业公司，受限于自身资源和条件，团队规模通常不大，往往要一人身兼多职，一旦客户逐渐增多，考验的就是企业的项目管理、人员交付、运维保证等综合能力。

Q: 您认为中国会不会跑出自己的“OpenAI”？

A: 未来的通用大模型应该还是几家厂商之间的游戏，出现真正意义上的中国版OpenAI的机会已经不复存在了。

OpenAI能够“杀”出来，一是因为GPT的路线选择。当时所有人都不知道GPT有智能涌现的功能，将其看作一种纯烧钱的投入，OpenAI一根筋地投入更像是一种乌托邦式的做法。而中国的融资环境、企业考核机制等现实，都决定了企业不可能以这种方式“孤注一掷”地坚持在这件事情上，大厂也是一样。

OpenAI的出现类似于一种田忌赛马机制，当大厂在犹豫和做对比的时候，投入的可能只有8-10个人，这时候“小厂”押注三五十人，可以实现局部领先。但当下，所有企业都已经明白，这件事甚至不是三五十个人就能完成的，因此不计成本地进行投入，在这种情况下就不太可能出现此前的差异化打法，大家都是明牌的情况下，就很难再跑出黑马了。

Q: 您认为大模型商业化的关键在哪？

A: 在商业化之前，人们通常更关注大模型的训练成本，但却忽略了真正的商业化和规模化之后，最大的成本来自于推理。最后的商业化一定是效果不断提升、成本不断下降的平衡点。一方面，模型的“聪明程度”要不断地满足日益提升的需求，解决行业的实际问题。另一方面，还要充分考虑客户的实际付费能力。长期来看，大模型是基础设施，商业的生态繁荣主要还是要看应用和插件的井喷。只有大家都基于大模型做开发，把大模型的应用都用起来，才是真正的商业化落地。