

上海近 40 度的高温，并没有阻止人们参会的热情——相反，7 月 4 日于上海举办的 2024 年世界人工智能大会暨人工智能全球治理高级别会议，比去年更为火爆了。

「就是一个火爆。」一位现场工作人员说。据了解，仅今天一天，此次大会的曝光量就有超过去年总和的趋势。

大会第一天，19 场论坛、数百位 AI 圈最瞩目的学术界、产业界代表展开了密集的讨论。可无论是图灵奖得主、科技公司一把手、还是展台前被 AI 裹挟的普通人，又都在说同一件事——AI 应用如何落地。

这里没有「大模型颠覆世界」的夸夸其谈，也没有「AI 应用落地不及预期」的断言，聊的都是：当下、今天，如何摘取大模型的果实。就着大模型「木桶效应」的最短板，怎么做出今天能用得上、用得好的 AI 应用。Transformer 和 Scaling Law 未必通向 AGI，但一个 LUI（自然语言交互）的新世界已经在日常生活的方方面面徐徐展开。

大会的议程设置往往反映了行业的普遍趋势。除了大模型之外，具身智能、机器人、芯片等领域也延续了去年大会的热度。论坛上，多位 AI 行业大咖都表达了自己的思考，从中可以一窥行业目前的大趋势。Enjoy。

周伯文-

上海人工智能实验室主任、

衔远科技创始人

对 AI 安全的投入远落后于 AI 性能的投入

目前，从算法研究、人才密集度、商业驱动力、甚至包括算力的投入等方面来看，我们对 AI 安全的投入远远落后于对 AI 性能的投入。现在，世界上只有 1% 的（资源）投入在对齐或者安全考量上。

姚期智-

图灵奖得主、中国科学院院士、清华大学交叉信息研究院院长

控制好 AI，但又不能破坏它

AI 风险来自于三个方面：一是网络风险延伸和扩大。现在，我们觉得管理数据安全已经是很困难的，出现了 AI 会困难 100 倍。二是没有意识到的社会风险，比如说 AI 非常强大，而且是可以有很多方式去使用，所以颠覆现在社会结构在短时间内发生的可能性，这是存在的。比如说有人提到，AI 可能带来大规模未来的失业。三是最有意思的层面，生存或者存

在的风险。以前也面临过，当火车或者蒸汽机发明的时候，就有人有这样的担忧。

作为计算机科学家看到了最有深度的问题，一方面我们把 AI 控制好，毕竟这是我们设计出来的；另外一方面，也不希望它被我们给破坏了，这样权衡是非常困难的。正如图灵所说，这是无法预测的，预测不了机器有了足够算力之后会做什么。

李彦宏-

百度创始人、董事长兼首席执行官

我们要避免掉入「超级应用陷阱」

一定要出现一个 DAU10 亿的 APP 才叫成功，这是移动时代的思维逻辑。其实不一定，AI 时代，「超级能干」的应用比只看 DAU 的「超级应用」恐怕要更重要，只要对产业、应用场景能产生大的增益，整体价值就比移动互联网要大多了。

智能体是最看好的 AI 应用方向，搜索是智能体分发的最大入口

热衷于「用 AI 写一个作文能得多少分」，其实这个使用价值是不大的，人家不会让你带一个大模型去参加高考。但是真正的需求是，大量的考生在考完之后要报志愿、要选择学校、选择专业，他们对一所大学、一个专业，会有各种各样的问题，而每一个考生的情况又是不一样的，这时候就需要有一个智能体来回答每一个考生专有的问题。

制作一个好的智能体通常并不需要编码，只要用人话把智能体的工作流说清楚，再配上专有的知识库，一般就是一个很有价值的智能体，这比互联网时代制作一个网页还要简单。

未来，将会形成庞大的智能体生态，而搜索是智能体分发的最大入口。

开源模型在特定场景有价值，但并不适用于大多数应用场景

开源大模型在学术研究、教学领域等特定场景下有存在的价值，但并不适用于大多数应用场景。在激烈竞争的环境中，需要让业务效率比同行更高、成本比同行更低时，商业化的闭源模型是最能打的。

（业界可能会）混淆模型开源和代码开源这两个概念。模型开源，你拿到的是一大堆参数，还是要去做 SFT，还是要去做安全对齐。你不知道这些参数是怎么来的，是无法做到众人拾柴火焰高的，即使拿到对应的源代码，也不知道他用了多少数据、用了什么比例的数据去训练这些参数。所以拿到这些东西，并不能够让你站在巨人的肩膀上去迭代和开发。

傅盛-

猎豹移动董事长兼 CEO，
猎户星空董事长

智能涌现，是一个灰盒状态

今天我们讲的生成式人工智能，是不是人工智能核心的道路？按照这条道路走，是不是能够创造通用人工智能？还是说也会有别的路线？

今天业界、学术界有不同的看法，认为仿人类的智能、仿人类的框架才能真正发现智慧的过程。因为今天的生成式人工智能，我们叫智能涌现，其实对中间的原理并不是特别清楚，是一个灰盒状态。

智能的涌现，可能是多条路线的。对于我们创业者来说，今天看到了这一波人工智能带来的巨大的生产力变革，就应该是坚定地投入。在人工智能技术构建上，去完成一系列产业变革。

安德雷斯·韦思岸

(Andreas Weigend) -

社交数据实验室创始人、亚马逊前首席科学家

人和 AI 不一样在于：好奇心和审美

到底推动人们发展的是什么？我们和 AI 不一样的地方在哪里呢？我有一个想法，那就是好奇心。

我还想再加一句，上周世界上最好的交响乐团——爱乐乐团来到了上海进行表演，他们演奏的音乐会特别棒。这个时候已经不是好奇心让人类疗愈人类了，而是欣赏、审美、热爱。

王坚-

中国工程院院士、之江实验室主任、阿里云创始人

GPT 的潜力还没有被完整地探索

80 年代初，发明了人工智能这个词的十个人之一——赫本山姆 (Herbert A. Simon) 跟我们讲，人工智能是未来。你想一个大学生被讲了以后，可以想象是多么热血。但等了几十年，(人工智能) 也没有来。所以很多问题的探讨，困难要超出我们想象。

但这一次可以说，人工智能是有别于过去的人工智能。不好说它是不是通用人工智能，但是它一定跟过去五六十年是不一样的。

General Intelligence 被翻译成通用人工智能，我个人觉得有一点点混淆，翻译成普通人工智能会更加确切，它是一个最最基本的东西，而不是从通用不通用的这个角度来讲。

人工智能的机会属于谁？

(考虑到大模型所需要的算力、算法、数据门槛，) 人工智能对大公司会友好一点，但是友好并不意味着宽容，一定会有新的大公司出现，也一定会有大公司烈火重生。大企业(可能) 会觉得 AI 是工具的革命，小企业一定会觉得这是革命的工具。我想大企业也要意识到这是革命的工具，那这个变化就来了。

张平安-

华为常务董事、华为云 CEO

AI 创新，离不开「在云端释放算力需求」的创新

中国的 AI 发展，离不开算力基础设施创新。这条 AI 创新道路，包括把端侧硬件 AI 算力的需求，释放到云端。就像华为手机的拍照功能已经可以把图片呈现得逼真、有丰富的细节，但如果把图片处理放在云端计算，图片会更立体清晰，像单反一样，（像所展示的这张图片里）蚂蚁绒毛清晰可见。

在云端，通过云网端芯架构上的协同创新，来构建可持续发展的 AI 算力基础，包括芯端算力上云、面向 AI 的网络架构升级、云基础设施系统架构创新三个方面。

井贤栋-

蚂蚁集团董事长兼 CEO

大模型落地三大难

业界普遍认为，通用大模型落地严谨产业，面临着三个「能力短板」：领域知识相对缺乏、复杂决策难以胜任，以及对话交互不等于有效协同。

通过专业智能体的深度连接，AI 会像互联网一样带来服务的代际升级

为了破解这些难题，专业智能体是通用大模型落地严谨产业的有效路径。通过专业智能体的深度连接，AI 会像互联网一样，带来服务的代际升级。（在移动互联网时代 APP 小程序是服务载体，未来是要往专业智能体方向推进。）

未来智能化的用户体验，一定不是只靠一个大模型，而是需要全行业深度协作，需要很多的专业智能体共同参与、各司其职。蚂蚁坚持走开放道路，和行业共建专业智能体生态。

徐立-

商汤科技董事长兼 CEO

应用是 AI「超级时刻」的关键

我一直的观点是其实虽然我们的行业非常的热，包括像 GPT 带来的聊天式的应用，Sora 带来的视频应用，但它还没有到「超级时刻」，是因为它没有真正走进到一个行业垂直应用中、引起广泛的变化。

可是这两天，我忽然感觉有点变化的想法。因为我的中学的退休的老师不停的在群里面问我，怎么样用人工智能去写文案、生成祝福的图片，发到他的退休群里等等。

我突然想，其实超级时刻和应用是互相成就的。只有超级时刻带来的认知的变化，最后才能推动应用。倒推回来，如果有应用作支撑，那么我们现在这个时刻就是「超级时刻」。所以，应用是「超级时刻」的关键。

走向广泛应用的三大突破点：高质量数据、流畅的交互、可控性

如果要推动人工智能超级时刻的到来，需要大模型可以展现出卓越的深度思考的能力。那么合成的人工数据，特别是高级思维的数据非常重要。所以越是有应用的场景，才能形成更好的、高质量的数据。

第二，自然的没有延迟的交互。端侧其实也是一个非常重要的突破点，推动模型的端侧优化，实时的交互变化会更加流畅。用好端和云两部分的计算机资源，才能够成为一种全自然的交互模式。

第三，所有的生成都要可控，你不需要做得很好，但你需要知道你哪里做得不好，并且说做哪里进行一些修改，有了这样的一个边界，那么实际上才能做到真正的可控技术以及可持续

的技术发展。

孟樸-

高通公司中国区董事长

终端与云端的紧密结合，将成为推动生成式 AI 规模化扩展、加速数字化转型的关键。虽然当前生成式 AI 的研发和应用主要集中在云端，并且云计算仍将发挥重要作用，但如果将 20% 的生成式 AI 工作负载转移到终端侧，预计到 2028 年将节省 160 亿美元的计算资源成本。

这种终端与云端的紧密结合，将成为推动生成式 AI 规模化扩展、加速数字化转型的关键所在。同时，为了推动生成式 AI 的广泛应用，我们也需要将其能力延伸到日常使用的智能设备上，如智能手机、移动 PC 和智能网联汽车等。

实现生成式 AI 在终端上的落地，需要在终端侧打造高性能 AI 处理器；还需要对生成式 AI 模型进行训练优化，使其体量越来越小，效率越来越高。

IDC 预测，预计 2027 年中国新一代 AI 手机出货量将达到 1.5 亿台，市场份额将超过 50%。在 PC 方面，咨询公司预计 AI PC 的渗透率将从 2024 年的 2% 上升到 2028 年的 65%。

闫俊杰-

MiniMax 创始人兼 CEO

应用的关键是，要降低模型错误率

我觉得目前大模型最核心的问题是，错误率还是比较高的。比如 GPT4 在很多测试指标上，正确率只有 60%、70%，也就是有 30%、40% 的错误率。国内的模型整体有 60% 到 70% 的错误率。

为什么大模型的产品都是一个对话的形式？因为对话的容错率比较高。为什么它不能成为一个独立的 Agent？是因为如果是个 Agent，需要进行多步、错误率就更高，就没有办法用。所以我觉得最核心的问题是，怎么把大模型的错误率从 30%、40%，降到 3%、4%、或者 2%。AI 从一个辅助人类的工具到能独立完成工作的最核心标志，就是错误率整体的降低。这对于更大的社会价值是至关重要的。

提升模型正确率的关键路径

比如为什么我们要做合成数据？比如为什么我们在不停提高训练的效率？比如我们为什么要研究新型的、比 Transformer 更好的网络结构？比如说我们为什么要研究各种各样的算法？比如说为什么要做更好的对齐……所有这些事都是围绕着怎么把这些技术加在一起、或者乘在一起，能够让我们半年或一年之后，有一个错误率个位数的模型。

大模型降价是一件好事

我觉得大模型的价格持续走低整体来说，是一个非常正向的事。因为它本来就应该降低。同时它降低的同时，效果就应该更好。

但是对于大模型企业来说，价格降低的好处是你可以有更多的用户、有更多在线的用户使用时长、有更多的流量，产生更大的价值，更多是基于这种流量价值来找到一种好的商业模式。

张鹏-

智谱 AI 首席执行官

大模型的核心突破是多模态

我们现在有一个很重要的点是去突破大模型的多模态。为什么要多模态？是因为真正的人在现实世界中解决问题的时候，他需要的、输入的信息本身就是多模态的。

除了自然语言以外，还有视觉、听觉、触觉，还有常识，所有这些需要综合起来，才能解决现实世界中很多常见的问题。

大模型降价不能长久

价格的下降是因为技术驱动，因为本身技术越来越好，成本越来越低、价格持续走低。但是这个事情过度就不好，真正的价值应该是逐级去做的。

我们给大家提供更优质的服务，大家能够用这个服务创造更大的价值，然后我们创造这一部分价值应该反向再传递回来，大家各自得到自己价值的部分，这是一个正常合理的市场的价值链。从这个角度来说，降价这种事是不长久的。

应用落地，顺着大模型最本质的变化——认知能力和泛化能力

过去一些 AI 的方法，比如说人脸识别，它已经可以在指标级上超过人类的水平了，为什么大家觉得这不是我们 AI 的终极答案呢？

比如说感知时代的 AI，能够产生实际的效能，但是它是受限的，泛用性不够、成本太高、需要垂直化去做很多事情、ROI 算不平。

而大模型带来了全新的能力，我们称它为「类人」的认知能力，让机器能够像人一样去思考。并不是让机器成为一个机器、一个工具，而是让机器像人一样去思考，思考这个能力所带来的效能提升是更重要的。

今天我們希望通过「让机器思考」演进的方向去赋能实体经济，大模型能够提供泛用化的能力，解决一系列的场景和应用需求，从而来解决成本和收益平衡的问题，这是它本质的特点。所以一定要从这个方向去思考，我们怎么去用新一代生成式 AI 和大模型的技术赋能实体经济。要构建更通用、更基础的一个能力，去解决多项的问题。这个过程中，模型本身的能力水平是最核心的。