

AI Insider Monthly#006

云计算的成本与方向、Deepfake、两万亿市值的微软与「后 CPU 时代」的英特尔 | AI Insider Monthly#006

Editor's Note

「AI Insider Monthly」以月度总结的形式，展现过去四周全球 AI、云、自动驾驶等领域的产业变化，通过分析行业事件与梳理技术趋势，勾勒出人类迈向数字化的若干方向与线索。

「AI Insider Monthly」来自于付费邮件通讯产品「[AI Insider](#)」，付费会员每周都会收到一封专属会员邮件，内容包括 AI 领域最新产业变革与前沿技术的深度解析，欢迎试读、订阅。

2021 年第 6 期「AI Insider Monthly」，将关注以下议题：

- 焦点：如何认识云计算成本与方向；
- 未来驾驶：软银调整机器人业务、融资烧钱游戏继续、百度与滴滴自动驾驶新进展；
- 行业：Deepfake 改变了什么、企业市场的语音、低代码；
- 巨头：「后 CPU 时代」的英特尔、Google 重组机器学习部门、微软市值突破两万亿美元、阿里云与腾讯云的海外扩张等；
- 趋势：当 AI 系统获得诺奖、中国超大规模语言模型；
- 报告：AI 伦理、机器学习信任链条；

请查收 2021 年 6 月份的 AI 与云计算产业发展报告。

云计算·成本·方向

作为长期跟踪云计算发展的观察者，过去几年我会深刻感受到越来越多的行业开始接受云计算的技术与商业模式，而在另一个维度上，不管是美国的 AWS、Azure 还是国内的阿里云、腾讯云、华为云，云服务的营收都在持续上涨，也成为支撑各自公司营收以及股价（或者未来增长）的重要部门。

特别是疫情期间，全球主要国家的数字化转型速度加快，大量工作流迁移到线上，企业开始重度使用云服务，整个世界也被迁移到了「云上」。

可以毫不夸张地说，2021 年，已经没有人质疑云计算带给企业的价值。

但没有质疑，或许是最应该警惕的时刻，正是在这样的背景下，投资公司 a16z 此前发表的一篇质疑云计算的[分析文章](#)才有了讨论价值，这篇文章像极了皇帝穿着「新衣」上街，只有一个「小孩儿」提出问题的情形，它试图指出一种可能：公司，特别是创业公司，不要被云计算的低成本入门产品所迷惑，稍有不慎，云计算成本就会吞噬公司的利润。

接下来结合我对行业的理解，提供几个思考角度。

其一，**云计算并不代表低价**。云计算的本质是计算资源的解绑与重新分配，云计算公司将传统意义上只存在于某个服务器上的计算、存储、网络能力单独「抽离」出来，随着抽离的资源越多，形成了一个类似「资源池」的虚拟存在，然后通过软件的调度，根据不同用户的采购需求，实现计算资源的再分配。

这个过程里，所谓的云计算低价，只是展示了最低限度的计算资源费用，一如某些购物网站会将零配件与主要产品的价格混合在一起，并通过低价吸引客户。在云的模式里，低价所代表的是几乎最低限度的 CPU、最小容量的硬盘以及俗称「小水管」的网络传输能力，（比如 1M 带宽），而对企业来说，这些低价的产品几乎就是一堆玩具，根本不能支撑起其基本的业务，因此会花更多钱去采购更贵的云产品。

由于企业的需求不同，目前没有权威的统计显示企业在传统 IT 采购与云计算采购中的费用对比，不久前，亚马逊创始人贝佐斯在写给股东的[年度信件](#)提出一个数字，企业采用 AWS 云服务获得成本节约为 190 亿美元，但这个数字过于笼统且具有极强的 PR 属性，所以也并不可信。

其二，**云计算带给客户的价值是灵活**。Instagram 的早期成长离不开使用 AWS 提供的弹性扩展能力，很长一段时间里，阿里云都支撑了微博因为某个明星出轨/婚变而导致的流量暴涨，这些国内外的案例也被云计算公司不断加工、包装，形成了云服务如此灵活的宣传话术。

与此同时，云计算的灵活还表现在企业能够根据自己的需求不断调整使用云服务的方式，如果你或你所在的企业使用过 AWS 的服务，一定会对 AWS 计费方式之复杂而感到震惊，在「简单云服务」与「复杂计费模式」之间有着巨大的理解和认知鸿沟，由此也形成了一个新行业：帮助企业客户优化 AWS 或其他云服务账单。

通过分析 AWS 以及其他云服务的账单，企业能够相对清晰地了解到云服务的真正成本，这对于很多 SaaS 类公司来说至关重要，行业里一个非常知名的案例，也是 a16z 这篇文章所提及的案例，就是网络存储公司 Dropbox。

和很多 2010 年代的互联网创业公司一样，Dropbox 一开始就使用 AWS，简单来说，这个

时期的 Dropbox，就是一个帮助用户将数据更好/更快存储到 AWS 的产品，而随着 Dropbox 用户数量的不断增长，数据存储、管理的成本直线上升，2016 年，Dropbox 公开表示将搭建自己的数据中心，并逐步将此前存储在 AWS 的数据迁移到自建的数据中心。

此举引发外界广泛关注，虽然我们可以判断出这种迁移一定是基于成本的考虑，但并不清楚具体数字，直到 2017 年 Dropbox 上市，在其提交的 S-1 文件里，Dropbox 展示放弃 AWS 带来的巨大优势，两年累计的成本节约高达 7500 万美元。

Dropbox Infrastructure Optimization Initiative Impact

Dropbox Historical Financials

	2015	2016	2017
Revenue	\$604	\$845	\$1,107
<i>Annual Growth Rate</i>		40%	31%
Infrastructure Optimization Cumulative Net Savings	N/A	40	75
Cost of Revenue	407	391	369
Gross Profit	\$196	\$454	\$738
<i>Gross Margin</i>	33%	54%	67%
Free Cash Flow	(\$64)	\$137	\$305
Incremental Margin vs. 2015 (% Pt)		+21%	+34%

Source: Dropbox S-1, a16z analysis

正如 a16z 两位分析师所言，Dropbox 的案例既具有代表性，又不具备通用性，不同行业、不同规模的企业，其是否使用云计算与成本节约之间并没有直接的因果关系，不能一概而论，事实上，Dropbox 现在也还在使用 AWS 的一些服务，但这个案例展示了一种可能性：企业需要考虑云计算的真正成本，时刻需要关注自己的利润是否被这些云计算公司吃掉了。

其三，对于云计算公司来说，新的挑战也带来新的机遇。很长一段时间里，AWS、阿里云都是公有云的忠实拥趸，像极了各自母公司的电商主营业务，公有云平台为企业提供了系列云计算的产品选择，企业客户只需动一动鼠标，就能快速使用这些云服务。

但在最近三年里，企业对于云计算的需求变得更多元化，公有云的标准化、一站式选择已经无法覆盖企业的多元化需求，再加上复杂的成本计算方式、对云服务商锁定的担忧以及数据隐私的考量，企业在私有云、混合云、多云等产品方面的需求变得越来越多。

由此我们也看到了市场的一些新变化，一方面，AWS、阿里云等公司加速在非公有云产品上的开发力度，从 AWS 的 EKS Anywhere 到阿里云不久前推出的飞天本地 Region，这些产品无不在进一步丰富企业的选择，帮助企业构建更灵活的工作流。另一方面，国内主要云计算公司高举云原生的大旗，也是从技术和产品层面进一步提升云服务的使用效率，进

而降低云计算的成本，这是未来一年都值得关注的趋势。

与此同时，结合一份 PwC 刚刚发布的[调查报告](#)可以看到，超半数的受访者（比例高达 53%）表示并没有从云服务中获取价值。

这也为从「上云」到「云上」提供了新的机会，各大云计算公司押注的无服务器概念与产品，连同虚虚实实的「云原生」，成为云计算公司提供给客户的新方案，其目的就是要让云上的客户感受到价值。

传统意义的云计算（IaaS、PaaS、SaaS）的分类已经变得不再重要，越来越多的创新出现在接近企业客户的产品侧，**Lambda** 会成为 **AWS** 未来最大的机会，**钉钉**也是**阿里云**持续增长的基础，**Office** 会与 **Azure** 共同构建其微软的增长飞轮，这些产品会重现定义云的产品分类与云的价值。

最后还需要再强调一下：上云依然是企业数字化转型的关键路径，通过上云完成 IT 架构、业务流程与成本模式的优化与升级，而在上云的过程之中，必须时刻关注云计算成本与利润的关系，苹果可以一个季度在 AWS 花费 **3000 万美元**，但苹果一个季度的净利润高达 **200+ 亿美元**，你的企业呢？



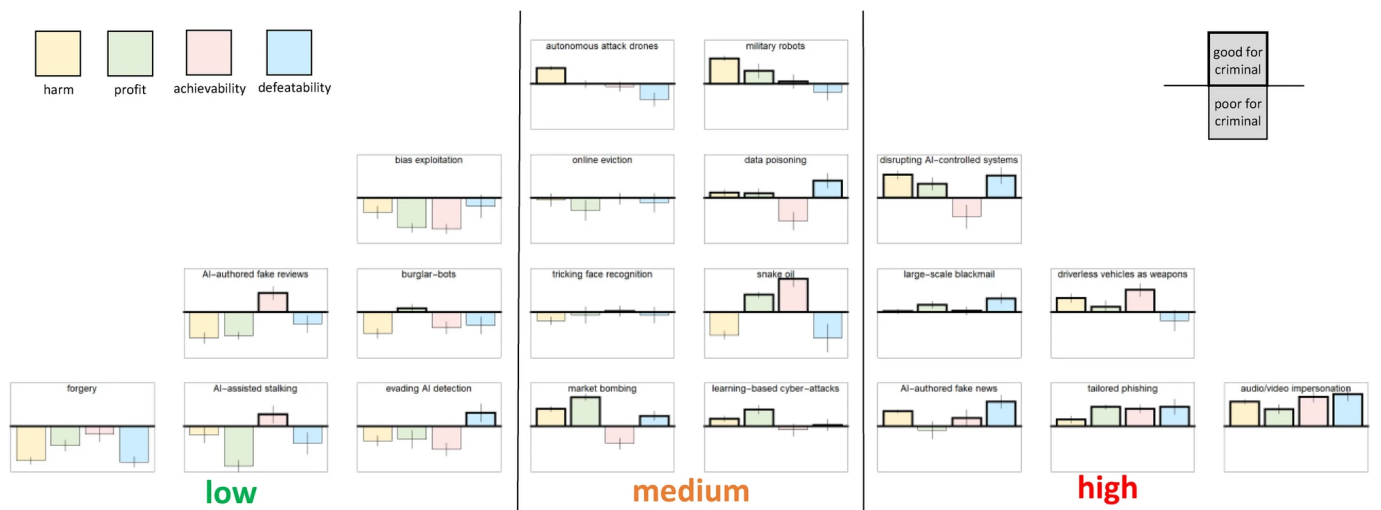
Deepfake·语音·低代码

由 Deepfake 衍生的争议、职业以及新应用场景

作为当下最具争议的技术之一，**Deepfake** 深刻改变着人们对于人物影像真实性的认知，今年 3 月份的时候，一位日本 Twitter 男用户利用 Deepfake 技术生成女性图片，很快积累大量粉丝，你可以在[这里](#)查看这位不存在的美女用户，而该账号的运营者是一位 50 岁的日本大叔（下图的对比），他在一档电视节目里[解释](#)了一切。



更进一步，基于 Deepfakes（不限于图像，包括音频视频）的犯罪已经成为 AI 犯罪的主要类型，加州大学洛杉矶分校去年发布了一份[报告](#)，按照危害等级、资金损失等四个维度做了区分，如下图所示：



目前来看，基于音视频的 Deepfake 正在成为主要威胁。这也是相对开放/开源的 AI 研究成果与越来越便宜的计算能力相结合的「产物」。而通过自然语言相关的犯罪行为还很少，但我们也不能过分乐观，毕竟 GPT-3 已经展示了足够的潜力。

与此同时，在另一个阵营里，围绕 Deepfake 技术及其应用，还有几个值得关注的方向。

其一，用技术对抗 Deepfake。比如 Google 开源了 3000 个 AI 生成的[视频数据库](#)，用于对抗 Deepfake。

社交媒体是 Deepfake 的泛滥地，过去几年 Facebook 不遗余力地对抗 Deepfake 技术，从举办挑战赛到推出检测 Deepfake 造假的 [Deepfake-detecting software](#)，FB 希望在技术与品牌层面成为抵制 Deepfake 的「模范公司」。

但正如斯坦福大学一个检测 Deepfake 项目研究者所言，随着 Deepfake 技术的持续提升，

检测技术将长期处于追赶的阶段，我们应该通过「非技术的方式」，比如培养一定的媒体素养来识别这些虚假视频或图像。

斯坦福大学的这个机器学习模型，通过识别说话人的嘴形与所说的声音进行匹配检测，从而找到细微的不同，可以实现对 Deepfake 视频 80% 的检测成功率，你可以在[这里](#)查看这项研究。

其二，Deepfake 衍生出的新职业。虽然包括微软、FB 等公司都在大力研发检测 deepfake 的技术，但当下仅仅依靠技术并不能完全应对不断升级的 Deepfake 影像/音频。

因此，从新闻机构到大公司，Deepfake 分析师的需求越发明显，这类工作所需要的技能，既包括对 GAN（生成对抗网络）的技术认知，同时还需要新闻记者的事实核实技能，从而为其他部门或客户提供相应的技术支持。

某种意义上，这也构成了围绕 AI 取代人类工作的一个观察维度——当机器「消灭」一部分职业的时候，还在创造另一些新职业。

其三，Deepfake 之于电影制作及更大层面的应用。[Slate](#) 本月初报道了电影制片人 Garrett Gilchrist 利用 Deepfake 技术修复上世纪电影的心得体会，这篇长长的访谈并没有谈论太多技术，更多还是从电影制作的角度展开叙述。

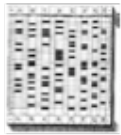
与此同时，[Fortune](#) 此前报道过一家名为 [Metaphysic](#) 的 Deepfake 技术供应商，这家公司通过类似 PaaS 的形式将 Deepfake 技术卖给不同行业，其客户包括广告服务商、电影工作室等，而在其营销的过程中，Metaphysic 高管反复强调该公司对应用领域的「控制」，包括禁止成人内容以及政治宣传等。

最后，回到 Deepfake 的关键技术 GANs，关于 GANs 是什么的解释里，我一直非常喜欢微软亚洲研究院的这篇[文章](#)，感兴趣的读者可以先读一下。而在 GANs 的应用中，除了「换脸」、「脱衣」这类噱头应用外，还有大量颇具潜力的应用场景，比如[这里](#)有 18 个典型应用，绝大多数停留在图像领域，但即便如此，也可以看到 GANs 对于整个 AI 领域的颠覆意义。

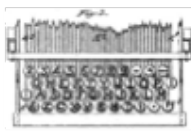
为什么语音的方向是在企业市场

从 2016 年到 2018 年，我会在不同的场景下看到下图或类似下图的变形：

Human-Computer Interaction (1830s – 2015)



Punch Cards for
Informatics
1832



QWERTY
Keyboard
1872



Electromechanical
Computer (Z3)
1941



Electronic Computer
(ENIAC)
1943



Paper Tape Reader
(Harvard Mark I)
1944



Mainframe Computers
(IBM SSEC)
1948



Trackball
1952



Joystick
1967



Microcomputers
(IBM Mark-8)
1974



Portable Computer
(IBM 5100)
1975



Commercial Use of
Window-Based GUI
(Xerox Star)
1981



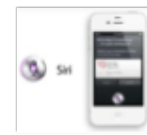
Commercial Use
of Mouse
(Apple Lisa)
1983



Commercial Use
of Mobile
Computing
(PalmPilot)
1996



Touch + Camera -
based Mobile
Computing
(iPhone 2G)
2007



Voice on Mobile
(Siri)
2011



Voice on Connected /
Ambient Devices
(Amazon Echo)
2014

Source: University of Calgary, "History of Computer Interfaces" (Saul Greenberg), & IOPCB - Internet Trends 2015

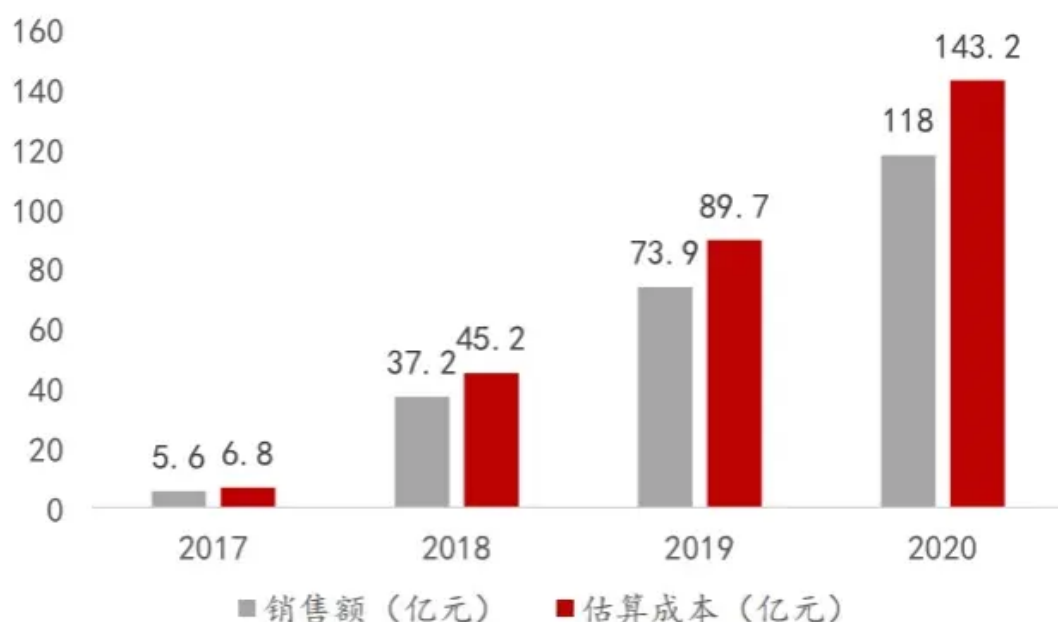
这是有着「互联网女皇」之称的 Marry Meeker 在其互联网报告里一张图，就像这幅图的标题所言，在人类人机交互的时间线上，基于语音的交互正在成为未来。

2016 年开始，中国智能音箱市场正式开启，无论是先入场的阿里、百度还是后续的小米、华为，上面这幅图所阐释的理论、类亚马逊 Echo 的产品设计、低廉的价格以及对万物互联的各种承诺，都让语音与智能音箱牢牢绑定在一起，迅速成长为一个颇为热闹的市场。

但热闹的市场背后，消费者与厂商们并没有获得「应有的价值」，一方面，智能音箱天然是一种固定设备，这极大限制了其应用场景；另一方面，语音技术依然很傻，用户对智能音箱的期待与技术不合格之间的差距，使得用户很难建立起使用语音交互的习惯。

而对于厂商来说，赔钱赚吆喝的行业规则根本无法持续。

2017-2020 年中国智能音箱整机销售额及成本估算



Meeker 指出了技术发展的未来趋势，但现状是，从苹果、亚马逊到阿里巴巴，这些当下中美最强大的技术公司，即便可以做出一台台所谓的智能音箱，但它们更像是语音数据收集工具，还需要更强大的机器或者廉价的人类劳动力，去解决「听见」、「听懂」的问题。

语音技术本质上不仅是要解决机器获取人声，更是要「识别」与「理解」，就目前来看，机器的语音识别准确率已经超过了人类，这既是智能音箱厂商前几年反复强调的技术突破，也为接下来语音在其他场景的应用提供了基础。

从宏观层面上分类，语音技术在消费市场和企业市场的应用场景与需求截然不同。

首先，在消费市场，从智能音箱到手机语音助手，这些语音交互场景都是开放式的，换句话说，机器所需要处理的是一个完全没有规则的交互逻辑，由此也给语音语义理解带来巨大困难。

比如第一代 Siri 出现的时候，人们迫不及待地想知道这个被乔布斯「钦定」的应用到底有多大威力，「调戏 Siri」一度成为全球性的话题，但很快，Siri 就被「玩坏」了，越来越多的用户发现，Siri 实在是太「傻」了，不仅功能单一，而且还无法真正理解人类复杂的语言体系。

What can I help you with?

“ Siri I'm bleeding really bad can you call me an ambulance ”

From now on, I'll call you 'An Ambulance'. OK?

Cancel

Yes

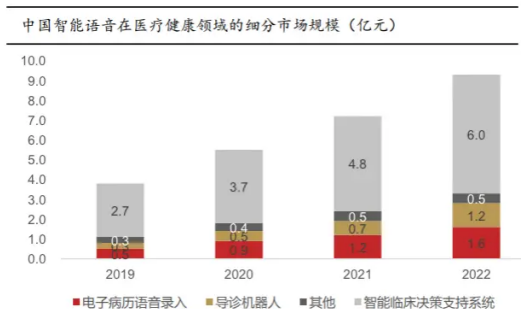
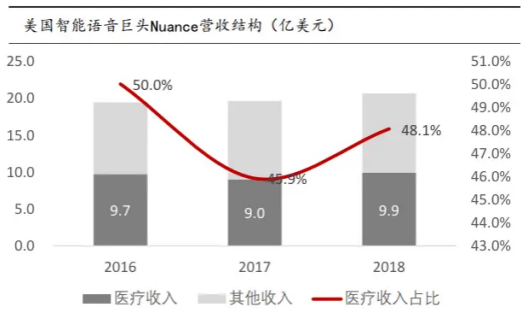


其次，企业市场的语音应用场景更垂直化，以在线教育为例，利用语音合成和 VR 技术，可

以构建起虚拟名师的形象，这个过程里的语音技术，其处理的需求都是集中在在线授课场景中，相对简单。

另一个值得关注的场景是医疗领域的语音电子病历，这里涉及到语音输入、识别等过程，最终形成结构化的病例数据，电子病历之外，包括导诊机器人在内的众多医院设施都具有语音交互的潜力。

这个市场有多大呢？下面是两张对比图，今年被微软 197 亿美元收购的语音技术公司 Nuance，其 2018 年医疗业务营收达到 9.9 亿美元，而 2019 年中国语音市场中的医疗健康细分领域规模只有 4 亿元人民币，预计 2022 年才能达到 10 亿元。

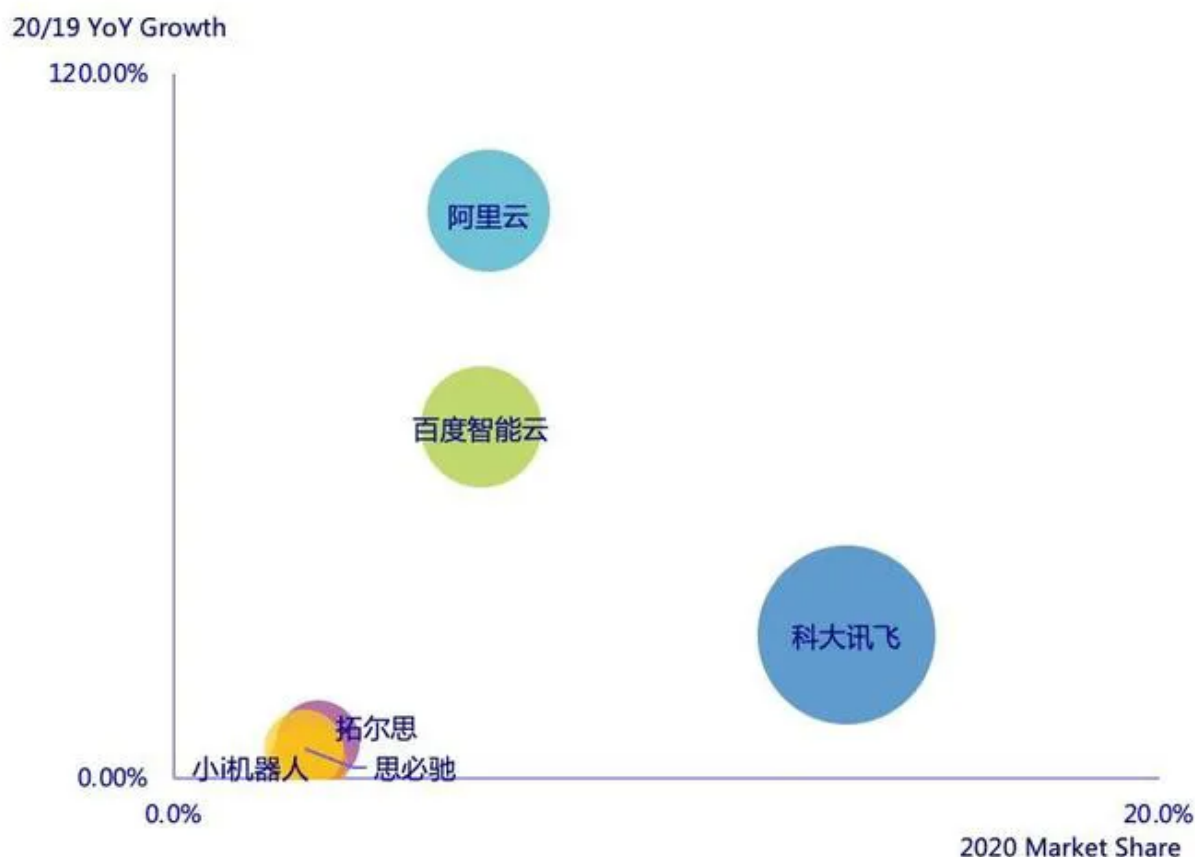


资料来源：iResearch、华西证券研究所

呼叫中心与客服也是一个典型场景，我曾在微软收购 Nuance 之后推断微软会将呼叫中心作为语音技术的发力点；在中国，大量的语音客服出现在电商、电信的业务推广甚至售后服务之中。

与其他企业市场类似，企业市场的语音玩家们也高度集中，上月，IDC 的一份报告也透露，2020 年疫情冲击下，语音语义软件的需求进一步向头部公司集中，其中阿里在该领域增长态势最为迅猛，年增速高达 96.6%，超过百度 59.8% 和科大讯飞 24.3% 的增幅。

2019-2020年中国语音语义市场变化



综上所述，语音技术还在不断发展之中，并且更多的应用创新出现在企业市场，通过「语音+行业知识」的结合，在降低成本的同时，提升客户体验，才是一项 AI 技术落地的关键，至于智能音箱或语音助手，距离成为你我心目中的萨曼莎（电影《Her》里的语音机器人）还有很长一段路。

低代码的偶然与必然

低代码正在成为一个热点。6 月初，国内老牌企业软件公司用友计划以 1.51 亿元人民币全资收购低代码创业公司 APICloud，股权转让完成后，用友拟以 7900 万对 APICloud 进行增资。

公开资料显示，APICloud 成立于 2014 年，2019 年发布低代码开发平台 Plus Mode，收购完成后，用友会将该平台与自己的低代码平台 YonBIP 相融合，进一步提升低代码产品的开发能力。

而在 5 月底的阿里云北京峰会上，阿里云总裁张建锋也透露了钉钉低代码平台的成绩单：在今年 1 月上线低代码开发平台后，钉钉平台应用总数超百万，3 个月增长了近一倍，其中低代码应用 3 个月时间增长了近 38 万个。

海外市场，AWS 去年发布了 [Honeycode](#)、Google 整合收购的 Appsheet 产品并推出[商业应用平台](#)，其目的都是要让客户尽可能减少对于代码的关注，专注快速构建应用。

低代码的流行有其必然性和偶然性。

首先，低代码也是软件开发与使用的必然。自上世纪 50 年代以来的计算机产业革命，持续且深刻地影响着人类社会的变革，在这个过程中，摩尔定律定义了硬件的发展轨迹，而「安迪·比尔定律」又在另一个轨道上加速了软件的突破，一代代软件「吃掉」过剩的硬件资源，推动硬件发展进入下一个循环，新的硬件反过来又支撑起更复杂、更强大的软件。

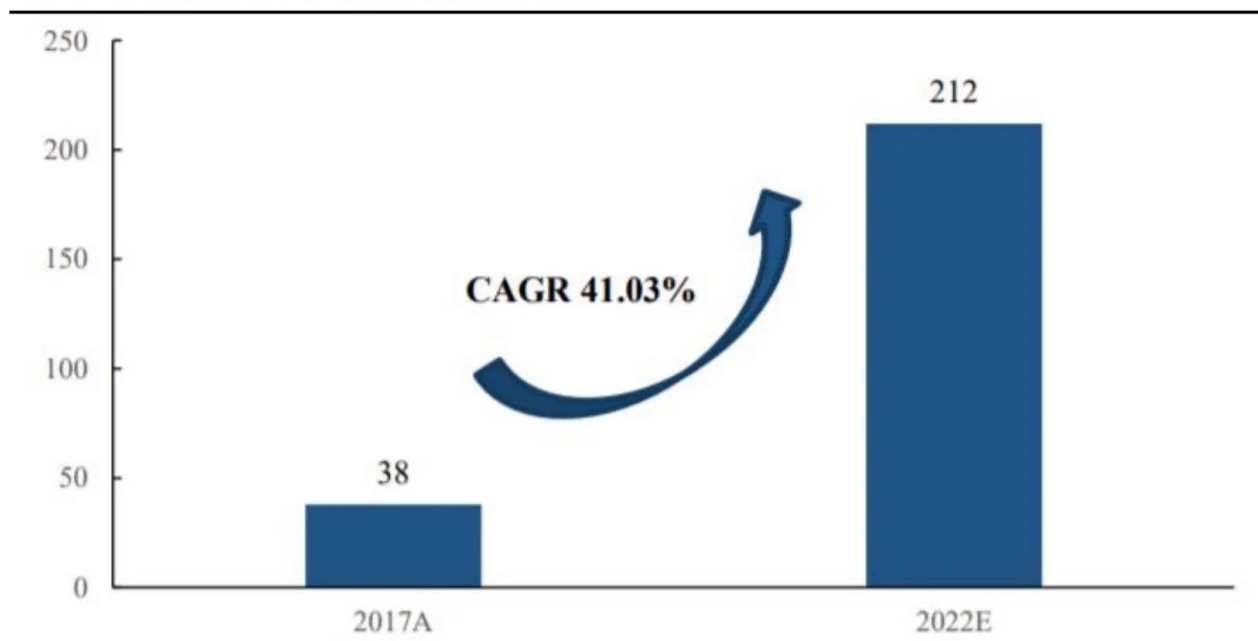
一代代软件的迭代过程，本身就是不断降低软件使用、开发门槛的过程，软件工程的本质是「封装」，由此形成的，便是一个不断变简单的开发、使用平台，换句话说，低代码开发的需求与供应一直存在，比如 Excel 长期以来就是一个低代码的数据库产品。

其次，当下低代码的流行有其偶然性，特别是在疫情的冲击下，数字化成为全球各地企业的必选题，由此也产生了巨大的数字化转型需求，如何利用数字化工具快速降低成本进而创造效益，俨然成为全球各地企业老板的共同需求。

利用低代码工具或平台，企业业务人员可以快速开发、部署业务应用，既减少了与开发人员的沟通成本，又能利用这些应用快速推进业务，实现「降本增效」，市场的需求，极大刺激了低代码平台的快速发展，巨头、创业公司以及资本也在快速进场。

根据 Forrester 的一组数字，全球低代码平台的市场规模将在两年后达到 212 亿美元。

图 5：全球低代码平台市场规模（亿美元）



数据来源：Forrester，东吴证券研究所

艾瑞咨询也对中国市场做了一番预测，如下图所示，年复合增长率达到 50% 以上，2025 年的市场规模达到 131 亿人民币。

图 6: 2017-2025 年中国低代码行业市场规模及增速



数据来源: 艾瑞咨询, 东吴证券研究所

上述这些预测数字更多还是给低代码开发平台或公司所关注的, 对于这个市场的客户而言, 「低代码」三个字背后, 还有诸多需要思考的方向。

其一, 低代码是一个非常笼统的概念, 企业没有必要搞清楚这个概念背后的具体含义, 而应该结合自己的业务需求, 寻找合适的工具和平台。

其二, 低代码不是灵丹妙药, 它不过是企业解决业务需求的一种方式, 当企业业务越发复杂, 不可能指望某个低代码工具或平台就能解决这些问题, 而且更重要的是, 由于很多低代码平台的开发逻辑被完全封装, 某种意义上就是一类软件黑盒, 企业客户往往无法真正了解其实现的原理, 当出现问题的时候很可能无法快速锁定问题。

其三, 低代码不仅会影响企业的应用开发与部署, 也会改变个体对于软件的认知, 当下出现了一大批同时面向企业和消费者的低代码平台, 比如 Airtable、Coda、Zapier 等, 普通用户也可以将这些工具应用到自己的工作和生活之中, 利用拖拽等可视化的方式构建专属于自己的「应用」。



机器人·未来驾驶

路透社援引消息人士的话称, 软银将调整其机器人业务, 并不再生产 Pepper 人型机器人, 该机器人此前由富士康在中国代工生产, 预计生产了 27000 台, 去年已经停产。

与此同时, 软银还将在全球范围内裁减机器人相关的工作职位, 除了已经在英国、美国开始的裁员计划外, 预计还会在法国裁掉当地 300 人团队的一半员工, 另外日本的机器人业务也会进行重组。

值得一提的是, 软银此前还把收购的波士顿机器人公司大部分股权出售给韩国现代集团, 上月, 现代集团也宣布将获得波士顿动力 80% 的股权。

自动驾驶的烧钱游戏还在持续, 先看融资。通用旗下的自动驾驶公司 Cruise 上月获得来自

母公司的 50 亿美元贷款，这笔贷款将帮助 Cruise 生产自动驾驶电动车 Origin，Origin 不会设计方向盘与刹车踏板，这也是 Cruise [第一辆](#)完全自动驾驶的汽车，通用预计将在今年小规模生产 100 辆 Origin，大规模生产将在 2023 年初开始。前不久，Cruise 已经宣布将于 2023 年在迪拜测试 Origin 自动驾驶车辆。

Waymo 也在上月完成一轮 25 亿美元融资，此前市场的传言已经成真，[投资方](#)来自 Alphabet（Waymo 的母公司）、银狐资本、a16z 等。

外界将 Waymo 称之为「Waymo Bank」，该公司已经成为自动驾驶市场融资最多的公司之一。产品层面，Waymo 目前提供 Waymo One，这是一个类似于滴滴或 Uber 的出行服务，所不同的是，Waymo One 车辆是自动驾驶车辆，该服务现在仅限于亚利桑那州的凤凰城。

再看巨头们新投资。多家媒体[披露](#)，**亚马逊**计划买下自动驾驶卡车创业公司 Plus 20% 的股份。Plus 公司瞄准的是 L4 级别的自动驾驶运输业务，该公司的核心产品是软件系统 PlusDrive，此前，亚马逊刚刚向这家公司采购了 1000 辆配置 PlusDrive 系统的自动驾驶系统卡车，中国顺丰也是其客户之一。

亚马逊去年已经全资收购了自动驾驶公司 Zoox，Zoox 也在日前披露了该公司一系列保障驾驶安全的措施，虽然是一个市场营销层面的展示，但可以作为了解自动驾驶安全的一个切入口，感兴趣的朋友可以在[这里](#)详细查看。

福特汽车收购了一家软件 AI 公司 Electriphi。这家公司[利用](#) AI 和机器学习技术为电动车公司提供电力能源软件开发和运营，其软件能够大幅优化电池使用效率，并通过路线规划、天气和司机匹配等方式，进一步提升车队的运营效率。

特斯拉披露其在超级计算机领域的技术投资成果，在上月召开的计算机视觉学术会议 CVPR 上，特斯拉 AI 高级总监 Andrej Karpathy 分享了该计算机的部分细节，整个计算集群拥有 720 个节点，每个节点采用 8 块英伟达 A100 80GB 的 GPU，共计 5760 块 GPU，理论的算力突破 1.8EFLOPS，这个数字在全球超级计算机排行榜上位列第五。

值得一提的是，这个演讲进一步强调了特斯拉在自动驾驶方面的路线：继续坚持基于神经网络的软件路线，而非以激光雷达为代表的硬件路线。你可以在[这里](#)观看 CVPR 会议期间的自动驾驶部分演讲，全长 8 个半小时。

再来看看几家中国自动驾驶创业公司的新进展。

滴滴 IPO 招股说明书披露自动驾驶业务规模。根据滴滴向 SEC 提交的[招股说明书](#)，自动驾驶是其四大核心战略板块，其他三个分别是共享出行平台、车服网络、电动车。

其二，关于滴滴自动驾驶业务的估值，2020 年 5 月，分拆后的滴滴自动驾驶业务完成 A 轮 5 亿美元的[融资](#)，投资方包括软银愿景基金和滴滴，估值达到 34 亿美元，而如果结合最近的一轮融资来看，目前滴滴的自动驾驶业务估值至少在 50 亿美元。

其三，截至到 2021 年一季度，滴滴自动驾驶团队人数超 500 人，自动驾驶汽车车队规模为 100+ 辆车，其主要测试和试运营重心在上海嘉定。

其四，滴滴在上海拥有 530 公里的道路可以进行测试，并将在北京、加州进行道路测试。

在 Uber、Lyft 先后出售自动驾驶业务之后，滴滴也成为行业唯一一家还在推动自动驾驶的按需出行公司，不过需要注意的是，Uber、Lyft 现在已经是上市公司，而滴滴此番正式提交 IPO 并释放如此多的自动驾驶利好信号，比如招股说明书强调将会把募集的 30% 资金用于自动驾驶相关业务研发等，这些安排多少也有助推其市值的意味。

从另一个维度去看，在中国当下的语境里，科技公司不能也不敢宣传所谓「取代人类工人（司机）」，但自动驾驶与按需出行的结合，其本质就是降低人类司机的成本，这也会成为未来滴滴自动驾驶发展的众多不确定性之一。

百度上月发布了第五代 Robotaxi 产品「Apollo Moon」，这款车是与北汽旗下的极狐合作，官方的新闻稿写道：

依托百度 Apollo 全球领先的自动驾驶技术积累，Apollo Moon 采用“ANP-Robotaxi”架构，不仅让共享无人车套件轻量化，还可与智能驾驶汽车数据共生共享，打造超强数据闭环。Apollo Moon 具备全传感器及计算单元冗余，完善的失效检测及降级处理策略，支持 5G 云代驾、V2X 等功能。整体相较上一代车型能力有 10 倍提升，复杂城市道路送达成功率高达 99.99%，在领航辅助驾驶 ANP 独立闭环的基础上，Apollo Moon 还搭载了来自禾赛的定制激光雷达和相应无人驾驶冗余，拥有完全无人驾驶能力。

简单来说就是两点：其一面向按需出行场景的数据、体验优化；其二，车辆成本，百度给出的数字是 48 万元（车辆+后期加装所有成本），这个价格大概是目前 L4 自动驾驶车辆成本的三分之一。

这对市场以及消费者来说都是一个重要的信号，短期来看，无论是价格还是体验，L3 以及更高阶段的自动驾驶车辆不会成为消费者购车的主要选择，基于按需出行的 Robotaxi 将继续承载自动驾驶企业的核心业务并定义技术发展方向。

值得注意的是，百度董事长李彦宏借助「留言」的方式表示：Apollo Moon 共享车「必须要比打车便宜」。

巨头·资本

「后 CPU 时代」的英特尔

英特尔上月发布了一个名为「IPU」（infrastructure processing unit）的新产品，正如官方新闻稿对该产品的定位：「.....a programmable networking device designed to enable

cloud and communication service providers to reduce overhead and free up performance for central processing units (CPUs)」，其作用也是减少某些特殊场景——比如机器学习——对 CPU 的依赖。

某种意义上是说，这也是英特尔承认 CPU 越来越无法满足多样化计算需求的一项举措，过往通过晶体管数量从而提升计算能力的摩尔定律早已失效，过去 5 年的时间里，英伟达已经向行业展示了基于 GPU 带来的巨大计算红利，特别是在面向机器学习的场景中，GPU 的效率惊人，英伟达也在打造「[黄仁勋定律](#)」。

与此同时，包括 AWS、Google、阿里云等，也在将自研的处理器部署到数据中心或云服务上，英特尔 CPU 不再是唯一选择。

数据中心业务营收放缓、制造工艺停步不前、缺乏进取精神，成为英特尔这家公司留给市场的印象，2021 年，该公司高层更迭，前 VmWare CEO Pat Gelsinger 正式成为英特尔 CEO，Pat Gelsinger 上任后不久就出人意料地[宣布](#)加码芯片代工，我曾在第 86 期「AI Insider」里指出英特尔芯片代工的两大困境：

- 内部文化困境：如何构建「客户为中心」代工文化？
- 客户困境：谁会成为英特尔的客户？中国公司会选择英特尔吗？

与此同时，[Bloomberg](#) 报道称，英特尔或以 20 亿美元收购 RISC-V 创业公司 SiFive，后者是一家位于加州的创业公司，目前正在和其他潜在买家洽谈收购，两家公司都拒绝评论该传闻。公开资料显示，SK 海力士、高通以及英特尔都参与了 SiFive 去年的融资，该轮融资使得 SiFive 估值达到 5 亿美元。

RISC-V 是一种开源的芯片架构指令集，自 2010 年出现以来，已经成为 X86、ARM 之外的新选择，特别是在英伟达收购 ARM 与中美持续技术对抗的背景下，RISC-V 架构更引发关注。

产品与收购之外，英特尔还在进行架构调整，该公司[宣布](#)计划成立两个新业务部门：

- 计算加速与图形计算部门
- 软件和高新技术部门

前者聚焦高性能计算和图形技术发展，由英特尔副总裁 Raja Koduri（曾在 AMD 和苹果工作过）负责；后者则是通过软件的方式「驱动英特尔的愿景目标」，负责人是前 VMware CTO Greg Lavender。

另外，负责英特尔网络业务的资深高管 Sandra Rivera，也将成为英特尔数据中心业务的负责人，这个业务包括 Xeon 芯片、FPGA 等面向数据中心的芯片产品。

从产品到架构，我们也可以相对清晰地看到，短期内的英特尔，其战略会紧紧跟着市场需求，IPU 的产品定位与英伟达此前发布的 GPU 有异曲同工之处，在资本市场与竞争对手的巨大压力面前，承认 CPU 的不足、不再过多谈及摩尔定律、并面向市场推出更多针对性产

品，英特尔已然进入到「后 CPU 时代」。

阿里云与腾讯云的海外扩张

上月两家公司相继公布在海外市场的一系列进展。

腾讯宣布，泰国曼谷、德国法兰克福、日本东京以及中国香港的四个数据中心开服，不同数据中心里还有多个可用区域选择，根据腾讯官方提供的数据，目前腾讯云拥有 66 个全球可用区，覆盖 27 个地理区域。

阿里云也宣布了其海外市场新突破，一方面，阿里云将在印尼和菲律宾新增两个数据中心，其中印尼的数据中心已经在 5 月底投入使用，而菲律宾的首座数据中心将在今年年底开通。另一方面，阿里云还将在当地投入 60 亿元人民币用于基础设施建设以及人才培养。

腾讯云与阿里云竞争的另一个重要赛道就是数据库。

上月，蚂蚁集团的分布式数据库 OceanBase 更新至 3.0，根据官方新闻稿，新版本已同时具备事务处理和数据分析能力，并已升级为一款支持 HTAP 混合负载的企业级分布式数据库，和之前的版本相比，事务处理性能提升 50%，数据分析性能提升 10 倍。

OceanBase 的目标是成为中国金融机构数字化转型的首选，随着核心软件国产化进程加快，数据库领域的竞争也会日益激烈，根据蚂蚁集团透露的信息，目前 OceanBase 已经在包括中国工商银行、山东移动、福建移动、数字江西、中国石化、中华财险、人保健康、浙商证券等在内的多家金融、电信机构落地。

此前，OceanBase 在数据分析型基准测试（TPC-H）中以 1526 万 QphH 的性能总分排名 30000GB 第一，针对该测试，OceanBase 团队成员分享了自己的心得以及对数据库未来的看法，值得一读。

	Cloud OceanBase		TPC-H Rev. 3.0.0 TPC Pricing Rev. 2.7.0	
			Report Date: 20-May-21	
Total System Cost	Composite Query per Hour Metric		Price/Performance	
69,336,912 CNY	15,265,305.7 QphH@30000GB		4,542.13 CNY Price/kQphH@30000GB	
Database Size	Database Manager	Operating System	Other Software	Availability Date
30,000 GB	OceanBase V3.2	Alibaba Cloud Linux Server 2.19	OceanBase File System V1.0	07/31/2021
7.3 - Power Test Geometric Mean of Query & Refresh Times				

OceanBase 的核心代码已经开源，你可以在 Github 上查看相关代码。

而在数据库领域，腾讯云自研的数据库产品 TDSQL 在富途、安心财险、和泰人寿等金融机构中应用，未来也会成为 OceanBase 的重要竞争对手。

Google 机器学习部门调整

商业媒体 [BI](#) 从一份 Google 内部的备忘录获悉，Google 正在将内部 AI 研究部门的员工整合到一个新部门，新部门的主要工作是研究如何将机器学习应用到 Google 的产品之中。而根据该备忘录上的 OKR，Google 计划未来两到三年在 AI 领域取得实质性进展，以改善用户与知识的互动方式。

这一方面凸显出机器学习之于 Google 的巨大价值，但另一个方面，我看到了 Google 对 AI 落地的焦虑，在经过了这么多年炒作之后，无论是机器学习还是语音、计算机视觉等，这些技术在实验室环境和真实环境中的成绩还存在巨大差距，Google 现在所做的，更像是面对终端产品——包括硬件、软件、服务——倒逼机器学习的研究与应用，以此形成机器学习落地实践。

机器学习也在改变着 Google 内部对于芯片设计的认知，上月，Google 的一篇论文展示如何利用强化学习设计芯片，这是 Google 一直在探索的领域，即通过算法优化硬件特别是芯片的设计，此前，我曾在「AI Insider」里介绍过 Google AI 负责人 Jeff Dean 的一篇非常通俗的论文，聚焦在深度学习如何影响计算架构与芯片设计，Jeff Dean 提供了一个认识深度学习、计算架构与芯片的认知框架，不需要太多技术背景，便能理解其发展方向，你可以在[这里](#)查看该论文。

回到 Google 的这个新论文，研究者们瞄准的是芯片设计图制作过程中的复杂问题，引入强化学习的思路，结果显示，可以在六个小时内完成芯片设计的关键指标设计，这是一个巨大的突破，并且该论文还表示，将在接下来的 Google 自研 TPU 的设计过程中采用该方法。

这篇论文已经发表在 Nature 网站，感兴趣的朋友可以通过 Nature 的这个[网址](#)免费查看。

Alphabet 旗下的 **DeepMind** 上月与非营利药物公司 DNDi 建立[合作](#)，这项合作的关键是，双方将把 DeepMind 去年震惊世界的 AI 模型 AlphaFold 应用于对众多疑难疾病的研究之中。

2020 年 11 月，DeepMind 宣布，AlphaFold 成功破解了生物学 50 年来的巨大难题「蛋白质折叠」。

蛋白质是人类理解生命组成的关键要素，目前已知的众多疾病，比如各种癌症都与蛋白质相关，但是蛋白质 3D 结构由于复杂且庞大的折叠方式，如何预测这些折叠方式成为长期困扰生物学的难题，AlphaFold 在蛋白质结构预测评估机构（Critical Assessment of protein Structure Prediction, CASP）组织的测试中获得 94 GDT，这意味着该系统具备了预测蛋白质结构的能力。

正如 DNDi 一位负责人所言：AI 可以改变（药物发现）的游戏规则，AlphaFold 打开了研究的新视野。而在去年，《[自然](#)》杂志也这样评论 AlphaFold 的意义：「It will change

everything」。FT 援引诺奖得主、英国皇家学会会长 Venkatraman Ramakrishnan 的话称：「它将从根本上改变生物学的多种方式」。

我们也期待 AlphaFold 尽快有新的消息。

AWS 收购与合作

AWS 上月[收购](#)加密通讯服务公司 Wickr，后者的客户包括政府、军事机构以及企业，这被认为是 AWS 正式进入加密信息服务的关键步骤。

根据 AWS 透露的信息，未来将会把 Wickr 的服务提供给 AWS 上的客户。不过交易金额并未公布。

AWS 上月与 Salesforce 达成重要[合作](#)，合作的要点是通过低代码甚至无代码的方式，打通 AWS 与 Salesforce 之间的应用开发流程，比如直接在 Salesforce 平台使用 AWS S3 或其他数据存储产品，或者使用 AWS 开发应用程序的开发者也可以访问 Salesforce 的数据，实现数据的统一。

需要提醒一点，Salesforce 还和微软、Google 建立了类似的合作关系。

法拉利汽车公司成为 AWS [新客户](#)，AWS 将为这家汽车公司提供存储、机器学习、数据分析等服务。

整个合作还包括车辆设计、测试、车主会员运营以及法拉利 F1 车队的数字化运维等多个方面，从合作的范围来看是一个非常大的合同，不过双方没有透露具体的合同金额。

与硅谷科技公司的公开政治游说类似，这几年越来越多的科技公司也在招募前政府官员，关注白宫动向的媒体 Politico 本月初做了一篇[调查报道](#)，焦点是 AWS 如何利用前政府官员的关系快速「抢占」美国联邦政府以及其他关键部门的云计算需求。

AWS 母公司亚马逊上月[发布](#)四款用于仓库的机器人，其中 Ernie 可以帮助工人节省从机器人货架上搬运物品的时间；Bert 是一个可搬运物品的自主移动机器人；而 Scooter 和 Kermit 则是运送车辆的机器人。

关于这些机器人的工作方式，可以参见亚马逊[官方博客](#)的介绍。

值得一提的是，亚马逊特别强调引入机器人的目的是让人类员工做更多有技能挑战的工作，并且该公司此前还公布了一组数字：自 2012 年亚马逊开始布局仓库机器人以来，该公司已经增加了 100 万个工作岗位。

微软市值突破两万亿

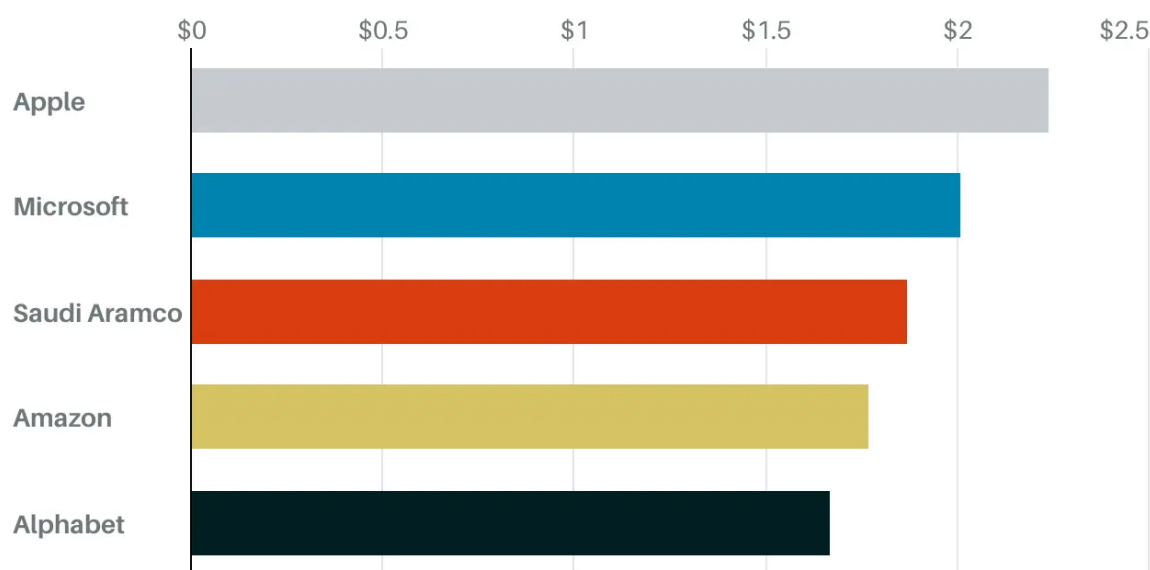
随着微软股价上涨到每股 265.79 美元，该公司市值也正式突破 2 万亿美元大关，成为全球

第二个 2 万亿+ 市值的科技公司，另一家公司为苹果。

The Golden Five

Tech dominates the top-five market caps, with Saudi Arabia's state-owned oil company the outlier.

Top 5 Market Capitalizations (tril)



Source: MarketWatch

而在 Wedbush 分析师 Daniel Ives 看来，得力于 Azure 的增长潜力，微软未来的市值还会继续增长，《巴伦周刊》援引他的话指出，企业数字化转型的需求还在加速，将会成为一个一万亿美元的巨大市场机会，而微软在帮助其客户转向 Azure 方面的工作仅仅完成了 35%，这也意味着，Azure 未来还有非常强劲的增长。

微软也在加大对华投资，[Bloomberg](#) 披露，微软将在 2022 年初增加四座中国数据中心，3 月份的时候，微软已经表示将扩大位于北京的数据中心规模。

目前微软的核心业务，包括 Azure、Office 365、Dynamics 365 等，都是由世纪互联运营，这是海外云计算公司在华运营的基本模式。

微软此前斥资 197 亿美元收购 AI 语音公司 Nuance 已经[获得](#)美国反垄断监管部门的许可，这笔收购也是微软自 2016 年 260 亿美元收购 LinkedIn 之后的最大一笔收购，此前我也在会员通讯里提及，利用这笔收购，微软一方面将继续优化企业市场（如 Teams）的语音交互功能，另一方面，呼叫中心的巨大市场诱惑，或许会成为微软下一个攻克的重要市场。

华为云调整或已结束

华为云接连不断的调整或许将落下帷幕，[多家媒体](#)称，余承东已不再担任华为云 CEO，华为云业务目前的两位负责人是徐直军和张平安，前者任董事长，后者出任 CEO。

过去半年时间，华为云业务从组织架构到人事都进行了多次调整，其中的关键人物就是余承东，作为华为内部炙手可热的高管，余承东过去几年带来华为消费者业务取得诸多令人

瞩目的成绩，而在美国多项技术封锁之下，他在今年先后成为 Cloud&AI BG 总裁、华为云 CEO，并且还以华为云 CEO 的身份上台发布了多款产品。

我在今年 4 月的一封邮件通讯里谈到，余承东出任 Cloud&AI BG 总裁的目的是解决原来 Cloud&AI BG 的利益博弈，最终实现「友好分家」，并推动张平安实现华为内部消费云与企业云的协同。

如今再去看，华为云业务组织架构已经调整完毕，云业务与传统 IT 业务分离，张平安将成为华为云业务发展的重要推手，现在的一个问题是，华为如何实现消费者业务的云与企业业务的云协同？

本质上说，消费云与企业云是两个完全不同的业务，行业里几乎没有哪家公司会推进两种云业务的协同，即便是同时拥有两种云业务的微软，面向消费者的云业务放在了「生产力与业务流程」部门，而企业云业务归于「智能云服务」部门，华为如何整合、协同两种不同云业务之间的关系，值得长期关注。

苹果强化机器学习

上月的 2021 WWDC 上，苹果展示了其各个平台的最新操作系统，目前这些系统还处于测试阶段，并将于今年秋天陆续发布正式版。

机器学习依然是此次大会的重要方面，iOS/iPadOS 15 的一系列新特性都离不开机器学习的功劳，比如实时识别拍照时文字信息的 Live Text、通过分析照片信息生成相应歌曲推荐的相册回忆功能等等。

与此同时，苹果进一步强调，所有这些机器学习都是发生在设备端，这意味着几点：其一，由于计算发生在本地，能够最大限度保护用户隐私；其二，本地计算能力的高低与苹果 A 系列处理的性能息息相关，新设备与旧设备的计算效率也会出现差异，这也会成为促使用户更换设备的一大动力；其三，相比于其他通过云端实现的 AI 能力，苹果采用本地计算能够极大降低延迟，带来更快的人机交互。

字节跳动的云计算野心

关于字节做云的消息已经有一段时间了，上月「晚点」给出了新的进展：字节跳动的云计算业务火山引擎将于今年 9 月到 10 月发布 IaaS 服务，正式进入公有云市场。

目前除了「晚点」的这个消息，市场还没有其他声音，字节跳动也未正面回应该消息。

字节跳动做云的最大「受害者」是阿里云，上一季度阿里云季度营收放缓，根本原因在于其海外市场失去了 Tiktok，正如阿里巴巴 CEO 张勇在财报分析师会议上所强调的，该客户不再与阿里云合作，是一个非产品层面的决策。与此同时，字节国内的众多业务都使用了阿里云的产品，FT 此前曾援引分析师的预测，阿里云未来将出现「字节跳动风险」，也就

是因为失去字节业务后的营收放缓甚至下滑的风险。

其次，发力云计算成为字节跳动的另一个重要增长赛道，这是字节跳动上市故事里的关键「情节」，目前该公司的云计算业务主要在 SaaS 和 PaaS 层面，但只有正式踏入 IaaS，才可以称之为「云计算玩家」。

但云计算是一个非常「笨」而且慢的生意，一方面是前期巨大的资金投入，参照 AWS、阿里云的盈利速度；另一方面则是客户服务与运营成本，这也是一个与字节 2C 生意完全不同的商业模式。

另一方面，中国的云计算 IaaS 竞争基本已经结束，阿里云、腾讯云牢牢占据着 40% 左右的份额，更进一步去看，整个 2021 年，包括阿里云、腾讯云和华为云都在聚焦政企大市场，并已进入到「[拼刺刀](#)」的关键时期，留给字节以及其他中小公司的市场份额还有多少，目前来看还是一个未知数。

趋势·报告

当一个 AI 系统获得诺贝尔医学奖

「What if」是《经济学人》杂志网站上的一个年度畅想类栏目，最新一篇的[主题](#)很有趣：如果人工智能系统获得诺贝尔奖怎么办？

文章给出了一个未来情节：一个名为 yulya 的淋巴瘤诊断系统在过去 18 个月里拯救了 400 多万人的生命，从而终结了自 2020 年新冠疫情以来最大的公共健康危机，但当诺奖委员会决定将医学奖授予该系统的时候，斯德哥尔摩爆发了来自人类的抗议。

这是一个类似科幻小说的情节，合理或不合理并不在这篇文章的讨论范围之外，但其中有几个值得思考的角度。

首先，当人类社会越来越依靠机器学习系统的时候，人机关系的变化方向会更微妙，一方面，从业者了解机器学习的能力与局限性，另一方面，绝大多数公众无法面对一个可以媲美自身智能水平的「非生物体」，这种复杂情感在巨大的认知鸿沟面前会变得越发不可调和。

但就像文章中的一个细节所指，如果未来诺奖委员会的成员们都是机器学习系统使用者，他们会为机器学习系统「正名」，这也意味着，随着机器学习使用者们掌握一个领域的话语权，人机关系会重新改写。

其次，一个机器学习系统的成功，到底是机器的成功还是其背后团队的成功？从运行机制上说，机器学习或深度学习本质上还是一个「黑盒子」，没有人会清楚解释，当大量数据交给一个机器学习系统之后，它会给出怎样的结果，这也引出更深层次的问题：如果机器

学习系统的输出导致一系列问题，那么谁来承担这个责任？

这是一个接下来需要不断探讨的领域，[世卫组织](#)上月针对现阶段医疗场景中的 AI 应用，提出了几个原则，或许可以作为一个指引，包括：

- 保证人的自主权
- 促进人类福祉、安全与公共利益
- 确保透明度、可解释性与可理解性
- 培养责任与问责机制
- 确保包容与公平性
- 促进可持续性的 AI

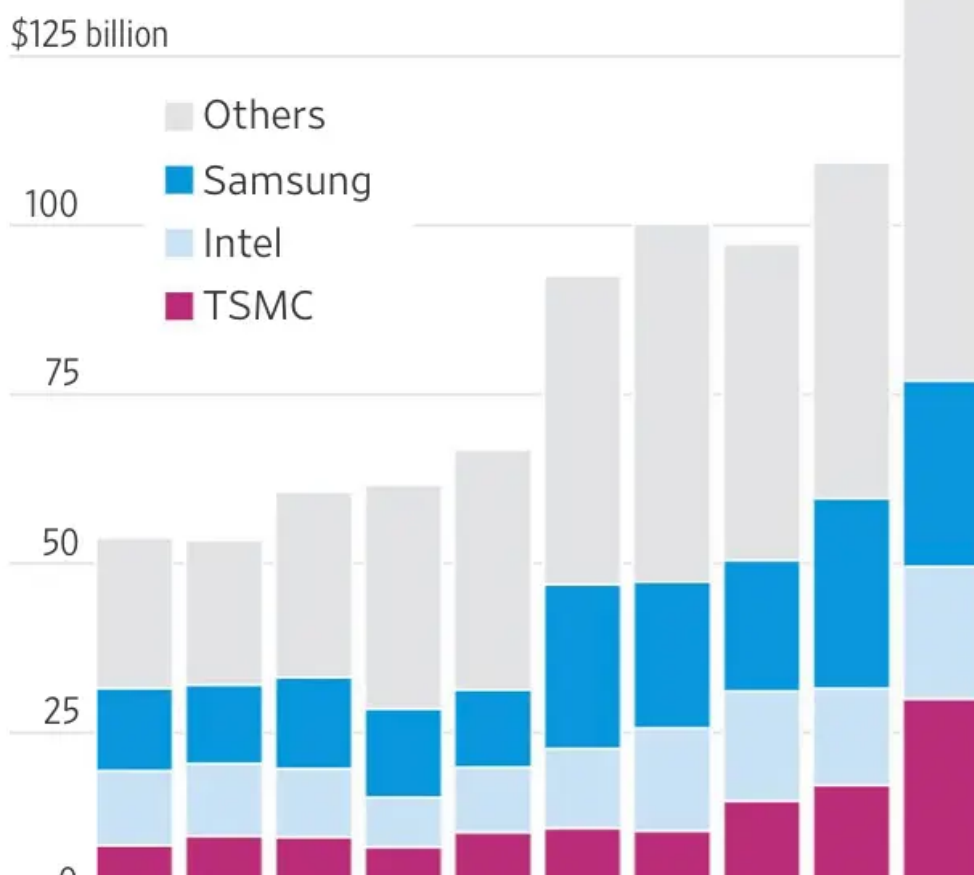
第三，也需要看到一个事实：机器学习或深度学习，不过是迈向强人工智能的一个方向，而不是唯一途径。

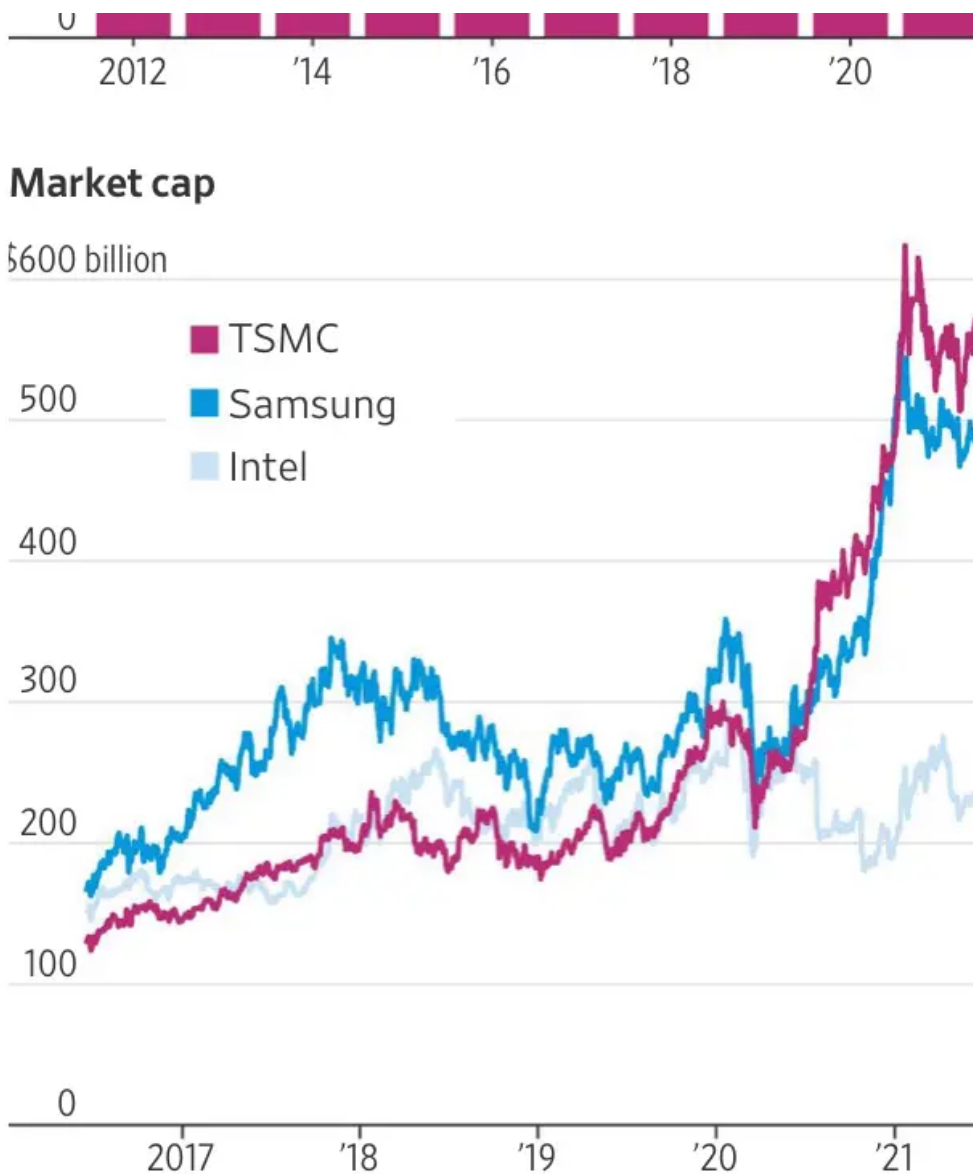
这意味着，或许在未来十年或二十年里，我们可以看到其他人工智能的建造方法，就像 2012 年之后，机器学习/深度学习迅速占领媒体版面一样，我们也会在某个时间段看到另一种「XX 学习」，它可能比机器学习更强大、更节能也更透明，从而成为又一个潮流。

只有一个台积电

[WSJ](#) 提出了一个非常尖锐的议题：芯片正在变得如此重要，但整个世界仅仅依靠一个位于台湾的台积电，是不是太脆弱了？

Semiconductor company capital expenditures





Sources: VLSI Research (capital expenditures); FactSet (market cap)

文章很长，我个人认为上面这张图很重要。

中国超大规模语言模型

以 GPT-3 为代表的超大规模语言模型正在成为新的潮流，上月，来自北京的智源人工智能研究院发布新版超大规模智能模型「悟道 2.0」，这是该机构 3 月份发布 1.0 版本后的新升级，根据[官方新闻稿](#)，悟空 2.0 模型参数规模达到 1.75 万亿，是 GPT-3 的 10 倍，也是目前中国首个、全球最大的万亿级模型。

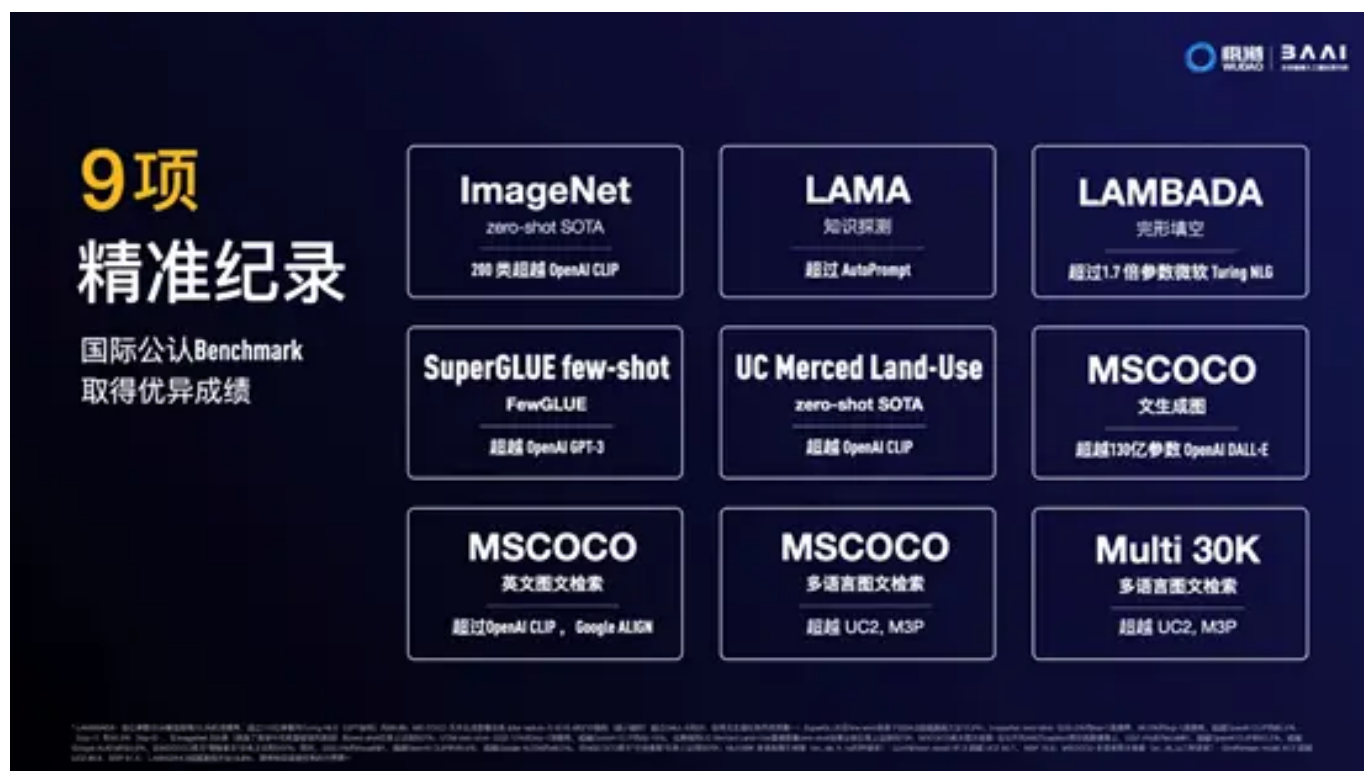
悟道2.0 全球最大

1.75 万亿

10 倍于 GPT-3 参数量

悟空 2.0 模型的训练数据量高达 4.9 TB，包括 1.3TB 中文文本数据、2.5TB 中文图文数据、1.2TB 英文文本数据。

如此规模的参数以及庞大的训练数据，其在多项测试中的成绩也不错，下图是官方披露的数字。



公开资料显示，北京智源人工智能研究院成立于 2018 年 11 月，成立的背景是「落实『北京智源行动计划』的重要举措.....由北京市科委和海淀区政府于2018年11月推动成立的新型研发机构」。

该机构设置理事会、管理团队、学术委员会等多个机构，其中理事会成员包括原微软亚太研发集团 CTO 张宏江、百度 CTO 王海峰、小米技术委员会主席崔宝秋、原华为诺亚方舟

实验室主任李航（现已加入字节跳动）等多位企业级高管，而学术委员会成员包括中科院院士张钹、美国加州大学伯克利分校计算机教授 Michael Jordan 等。

从上述人员配置可以看出，这是有关部门最喜欢的「产学研」架构，目标也是贯通人工智能从实验室到产业落地的方方面面。

回到悟道 2.0 超大规模模型的话题上，一如 GPT-3 一样，如何将如此大规模的语言模型融入到各个产业，是一个不亚于模型研发的难题，官方消息是，目前已经与新华社、小米、美团等 22 家科技公司签约。更进一步去看，相比于 OpenAI 所面临的产业与舆论环境，悟道 2.0 以及智源人工智能研究院无需担心资金问题，也不用关心语言模型中的偏见争议。

韩国也在近期披露了自己的「GPT-3」模型，韩国最大搜索公司 Naver 发布了一个名为 HyperCLOVA 的语言模型，拥有 2040 亿参数，采用韩语作为训练数据。

未来一段时间，不同国家、不同语言的超大规模语言模型的「竞赛」还在持续，这种算力、数据层面的「肉搏」能否带来人工智能的真正突破还无从知晓。

AI 伦理与机器学习的信任链条

皮尤的一份报告显示 AI 伦理发展的「缓慢进度」。这份报告调查了 602 位科技公司高管，其调查结果包括：

- 68% 的受访者表示，关注公共福利的 AI 伦理不会在 2030 年前普及；
- 相比于 AI 伦理，受访者的态度显示当下 AI 的焦点是利润与社会控制；

你可以在[这里](#)获取这份 127 页的报告。

过去几年，随着大量机器学习数据集、开源工具的出现，机器学习的门槛被大大降低，研究者、开发者很多时候无需进行重复工作，直接使用开源社区上的工具即可。

从这个角度看，机器学习已然成为一种基于信任的领域，研究者/开发者「默认」自己所采用的公开数据集或开源工具是安全且值得信任的，这也构成一个威胁：如果有人在这些数据集或开源工具中恶意注入带有偏见的数据或代码该怎么办？

你可以在[这里](#)查看 CSET 的这份报告，通过拆解机器学习不同流程中的威胁，提供了几个解决方案。

国家层面的 AI 投资

欧美国家正在加大对 AI 领域的投资。

根据美国国防部新的预算草案，从下个财年（2022 年 10 月起），五角大楼希望在 AI 领域投资 8.74 亿美元，支撑该机构 600 多个 AI 相关项目，进一步增强美军的作战能力，这笔

预算金额要比今年增加 50%。

与此同时，[美联社](#)的报道指出，美国参议院通过一项前沿技术投资法案，这项法案的核心是向美国商务部拨款 500 亿美元，旨在推动包括 AI、半导体在内的前沿技术研发，从而应对来自中国的竞争。

北约也在上月成立一个科技孵化中心，投资 AI 在军事领域的应用，[FT](#) 称，该计划将组建一个由各个成员国自愿出资的投资基金，通过这个基金，投资一些 AI 创业公司，鼓励他们将 AI 技术和产品应用到军事领域。