

智能科技 跨界相变

— 2024 数字科技前沿应用趋势

序

过去一年，我们见证了数字科技的加速度。每个人都身处变革巨浪之中，既对大模型的突破进展无比兴奋，也对未来充满了无限憧憬。

新年伊始，眺望未来2-3年的科技趋势。我们正驶向一个由连接衍生交互、由计算催生智能的时代。

高性能计算、量子计算、云计算和边缘计算这“四大计算”融汇贯通，正催生全新的计算范式。

通用人工智能渐行渐近，大型模型走向多模态，AI智能体可能成为下一代平台，“AI科学家”有望加速问世。全球已达共识，AI治理将引领我们踏上更智慧、更安全的未来。

机器人演进加速，灵巧手让人形机器人更敏捷、更像人。AI与生命科学的交融，将帮助我们洞察更多的生命奥秘。

我们会进入一个3D、全真在场的新世界。脑机接口不仅在医疗实现突破，正拓展至更多的互动场景。或许不久的将来，我们将亲自见证意识上载从科幻成为现实。

手机卫星电话、垂直起降飞机将改变我们的应急方式和出行模式。交通网、信息网、能源网的融合，将推动智能、绿色的能源变革。

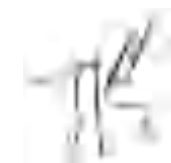
未来已来，一个充满韧性和重塑的全新时代即将揭开序幕。让我们拥抱变革，共同谱写人类与科技和谐共生的新篇章！

专家推荐

在这个日新月异的变革时代，我们有幸见证并参与了数字科技的创新突破。特别是近年来，以大模型为代表的AI技术，让我们看到了通用人工智能的曙光，有望大幅加速技术演进的步伐。

腾讯的2024数字科技趋势，从连接、交互、计算和智能四个维度，对100多项未来技术和重点方向给出了趋势性判断。从星地直连的卫星互联网，到垂直起降飞机的未来交通网，再到能源、信息和交通的多网协同，未来网络连接的广度和深度都迎来无限可能。交互方式正在发生变革，数字交互引擎不仅让虚拟世界更真实，也让真实世界更丰富，未来的全新3D视界呼之欲出，脑机接口的新进展也带来了更大的想象空间。计算无疑是各类智能的底层基础支撑，也成为当前竞争的焦点领域。而智能的升级，不仅会给机器人注入具身智能，还给微观世界的基因计算带来新突破。

预测未来的最好方式，就是创造未来。相信在行业界的共同努力下，这些前瞻洞见会更快到来，从趋势变成现实。



丁汉

中国科学院院士

专家推荐

去年以来，以生成式AI为代表的人工智能技术发展激荡人心，正在引发智能的元革命。可以遇见的未来，智力将会成为一种基础设施服务，新的IaaS呼之欲出。类比摩尔定律，未来可能会出现新的智能定律，对信息世界、物理世界和生命世界带来全方位的影响。

未来十年，AI会无所不在。首先，现有的各类应用可以用AI重新做一遍，新的交互变革、体验创新将会带来更新的智能硬件、更多的智能服务，孕育出比历次工业革命都巨大的产业机会。其次，AI会成为各领域的底层操作系统，AI+机器人的具身智能，AI+生命科学的基因计算，AI+未来出行的自动驾驶汽车和垂直起降飞机，甚至AI+脑机接口的硅基和碳基结合的新生命体，都会一步步成为现实。最后，AI能力的提升也伴随着风险的扩大，我们要充分发挥技术发明的智慧和把握应用方向的智慧，让AI科技更善良、更有创意，朝着强化人、成就人的方向可持续发展。



张亚勤

中国工程院院士、清华大学智能产业研究院（AIR）院长

数字科技星图

DIGITAL TECHNOLOGY STAR CHART

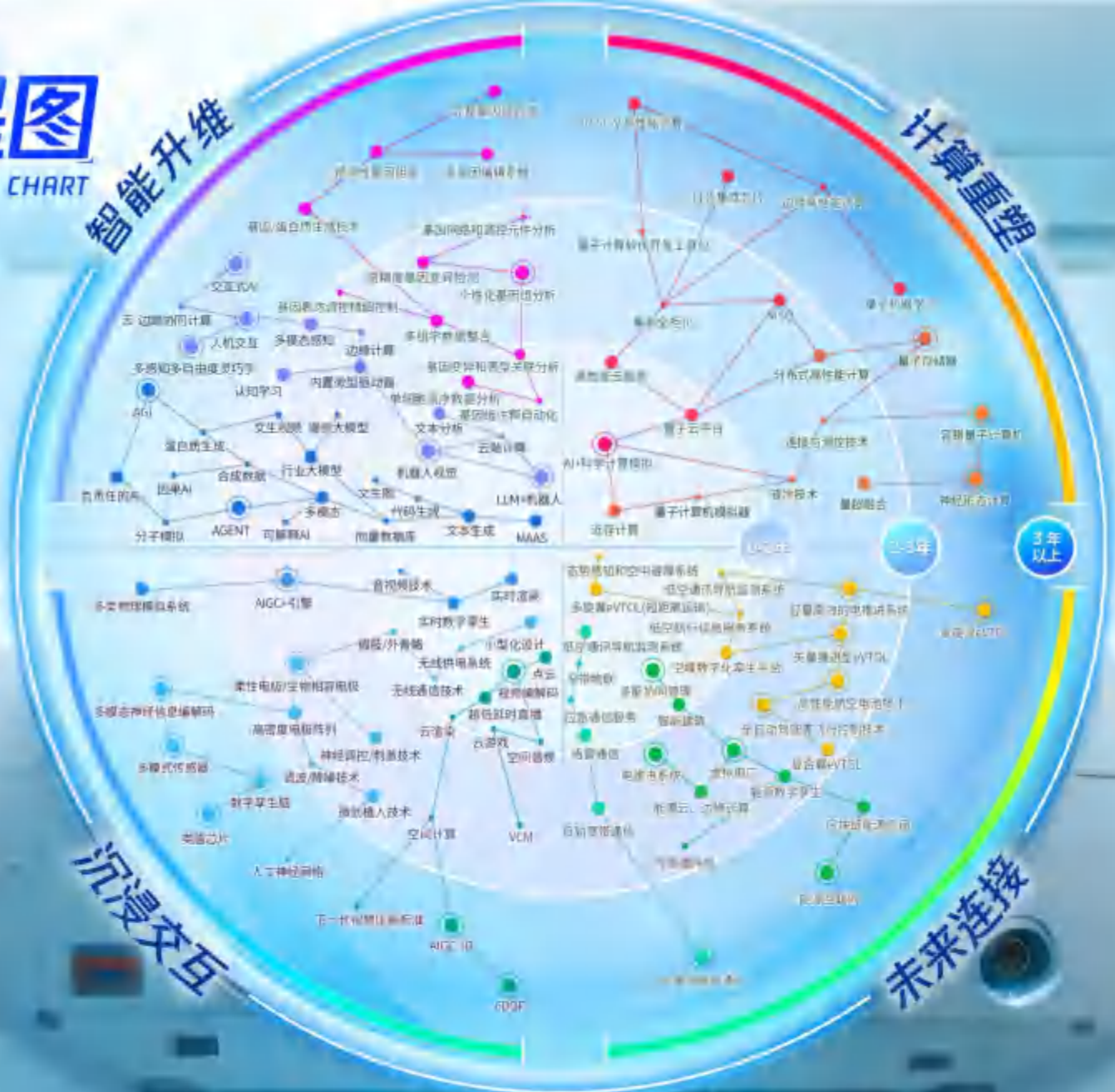
Tencent 腾讯

智能升维

计算重塑

沉浸交互

未来连接



计算重塑

高性能计算的“四算解构”

智能升维

多模态智能体、人形机器人、AI+基因计算

沉浸交互

数字交互引擎、脑机接口、沉浸式媒体

未来连接

星地直连通信、eVTOL、量子电厂

技术影响力

影响力中 影响力大 影响力极大

技术成熟度

3-5年 3年以上

2027

2026

2025

2024

2023

2022

2021



目录

一、计算重塑

高性能计算的“四算聚变”

二、智能升维

多模态智能体加速AGI进程

AI加速人形机器人“手、脑”进化

AI+基因计算解读生命密码

三、沉浸交互

数字交互引擎激发超级数字场景

脑机接口从医疗突破迈向交互革命

沉浸式媒体催生3D在场

四、未来连接

星地直连通信推动泛在网络覆盖

eVTOL加速空中出行奔赴新时代

多能流实时协同重塑虚拟电厂

高性能计算的 “四算聚变”

本章节与行业机构「光子盒」联合研究推出



AI生成

高性能计算的“四算聚变”

- 今年，全球各地高性能计算集群迎来向2.0架构（CPU+GPU）的升级潮，高性能计算集群、量子计算、云计算和边缘计算的“四算融合”也成为高算 3.0 演进的新趋势，衍生新一轮科技探索。高算相关研究机构和企业纷纷加大了在计算单元、存储、网络互联、软件中间件、算法等关键技术上的科研投入以适应新形势，并努力寻找穿越“内存墙”的有效路径。
- 未来几年，高性能算力应用将爆发，以人工智能和科学计算模拟为代表的算法、软件、以及相关的研究成果和记录将迎来一轮刷新。加之可持续计算的加大投入，高性能计算技术应用发展呈现快演进、重效能的新形势。



01



量超云边

高性能计算3.0将以高性能集群为基础，融合量子计算、云原生和边缘高性能，形成并行与分布式一体的计算架构。

02



破墙而出

处理器、存储和网络等关键硬件技术，正在不断进化、重组甚至是革新，试图突破现有的性能瓶颈，其中存算一体技术为“后冯诺依曼时代”提供新的发展路径。

03



模拟一切

计算机模拟，因为算力的提升、软件工具的迭代、以及与人工智能的结合，将成为大语言模型之后、迭代最快的应用发展路径。

04



可持续计算

持续高企的耗能，使得“计算能效”成为可持续计算的重要指标，高性能计算各个层面为“降耗提效”加紧探索。

发展阶段与大事件



至今
→



未来
→



1.0: CPU为核心计算单元

2.0: CPU+GPU成主流

3.0: CPU+GPU+QPU 异构计算集群

图片均来自互联网公开信息

2023 chronicle of events 大事记

纠错量子技术实现“盈亏平衡点”

2月，Google Quantum AI团队在《自然》杂志上发表名为《通过扩展表面码逻辑量子比特来抑制量子错误》的论文，证明将多个量子比特分组组合成为一个逻辑量子比特的纠错方法可以提供更低的容错率，进而证明量子纠错达到“盈亏平衡点”，量子计算机将“越纠越对”。这是量子计算发展的重要里程碑，为实现通用计算所需的逻辑错误率指出了新的途径。

盘古气象模型投入欧洲气象预测

7月，欧洲中期天气预报中心 (ECMWF) 发布了4月和7月进行的对比测试报告，将华为盘古气象模型于欧洲数值模型进行了对比。报告显示，盘古气象模型在一系列精度指标上展现出优势，在处理气象学家关心的极端天气预报方面表现出色。其1小时-7天的预测精度均高于传统算法 operational IFS，同时预测速度提升10000倍。相关论文也发表在《自然》杂志。

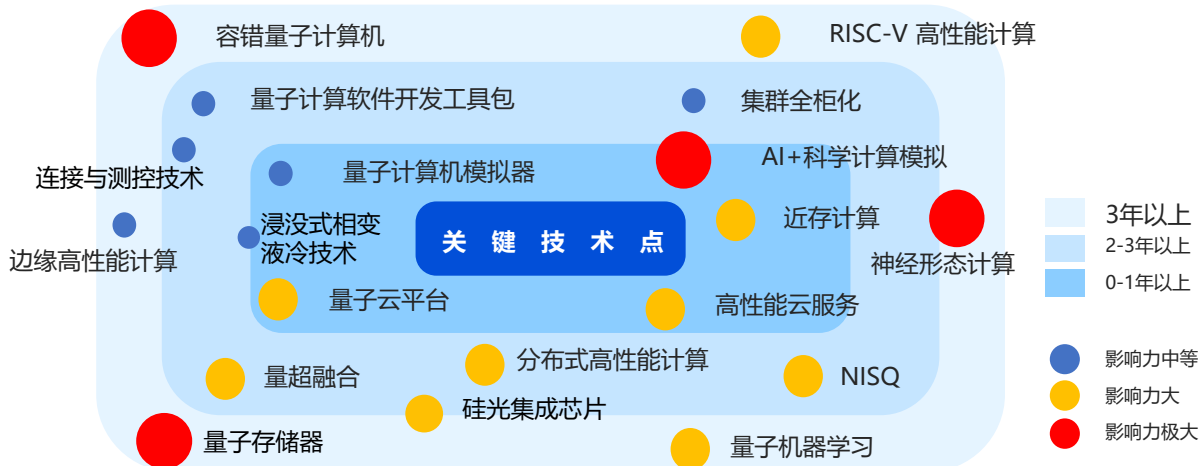
首个RISC-V高性能服务器集群交付

9月，算能与山大共同研发了面向高密度计算和数据中心场景的RISC-V融合服务器集群方案正式交付，这是RISC-V在数据中心的首次商业落地，标志着RISC-V正式迈入高性能计算领域。该集群共有3072核，采用48颗算能科技SG2042 RISC-V SoC 64核CPU芯片。

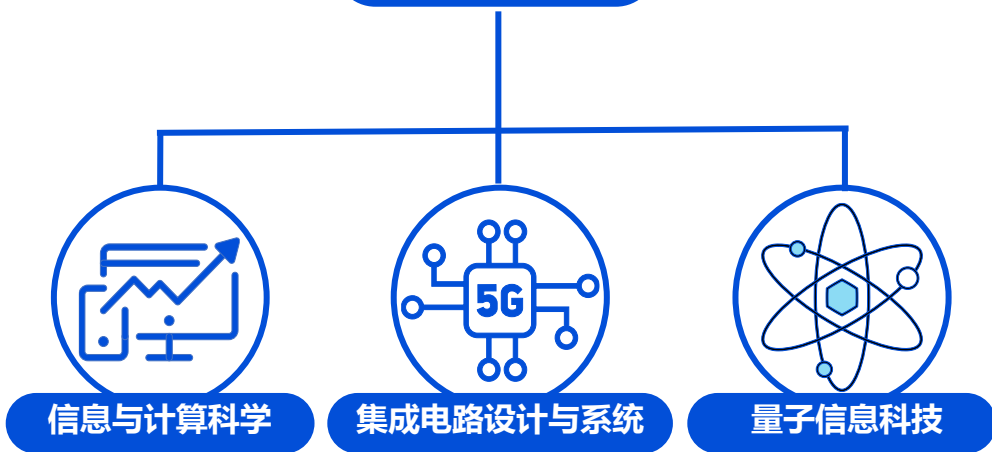
最新E级高性能计算机Jupiter开建

10月4日，欧洲高性能计算联合组织EuroHPC启动了欧洲第一台百亿亿级高性能计算机JUPITER的建设。该系统专为科学和工业领域的最大可能模拟和人工智能应用而设计，预计2024年推出。JUPITER采用了当前最先进的高性能计算技术方案，包括欧洲自研CPU Rhea、英伟达GH200、液冷机柜等，此外系统还预留了量子计算和神经形态计算的扩展规划。

技术热点

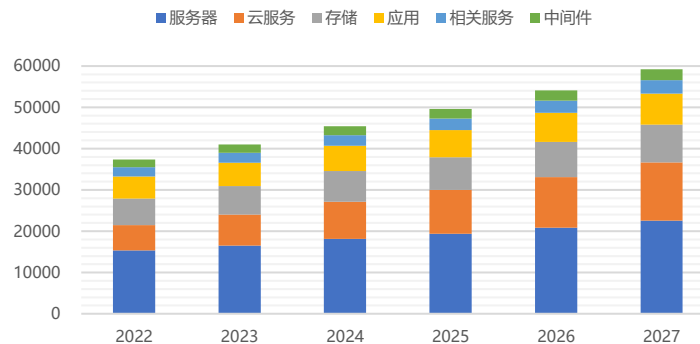


关键学科



市场趋势

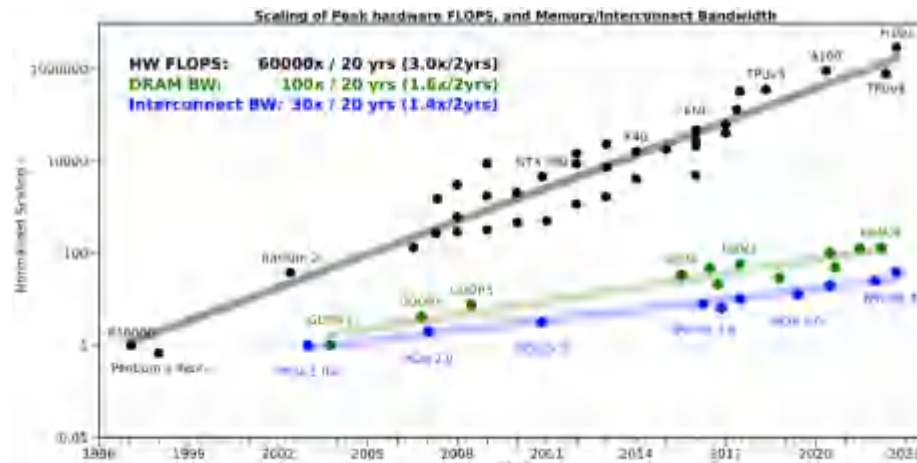
高性能计算市场趋势 (单位: 百万美元)



未来 5 年，高性能计算传统市场的 CAGR 将维持在 8%，而高算云服务的 CAGR 将超过 18%。合并市场总量预计将在 2027 年达到 592 亿美元。

数据来源: https://hyperionresearch.com/wp-content/uploads/2023/11/E-Joseph_HPC-Market-Update-Introduction_Hyperion-Research-SC23-Briefing.pdf

热议问题



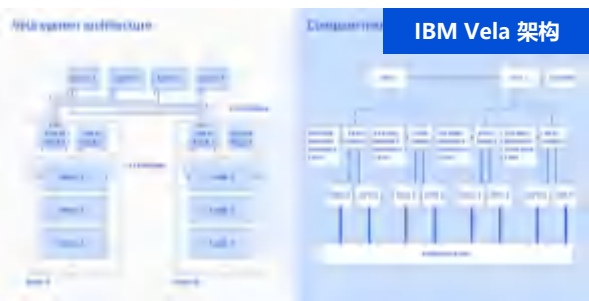
冯诺依曼架构下，限制高性能算力扩展的“内存墙”越发明显
<https://medium.com/riselab/ai-and-memory-wall-2cb4265cb0b8>

趋势要点1: 高算集群、量子计算、云计算和边缘计算“四算”融合，衍生新技术路径

- 随着高性能计算机群不断引入web 服务和容器化等云计算技术，以及云计算平台开始提供分布式高性能虚拟机服务，两者技术栈开始迅速同化；
- 传统高性能计算与量子计算的融合，成为未来高性能算力跃升的高潜路径，世界各大研究机构纷纷开展试验；
- 量子计算系统的扩展引入了集群的思路，量子云平台将成为量子计算软件开发和应用探索的重要基础设施；
- 高性能、低功耗的芯片技术正在培育高性能边缘计算的基础，云边协同的分布式架构将成为高性能计算未来的典型架构之一。

云超同化

高性能算力的充分利用，云计算技术的引入是必要选择。现有高算集群云原生化脚步进一步加快，如IBM 云原生高算Vela，以及 AWS 和英伟达联合的Project Ceiba；



另一方面云计算平台对于并行+分布式的高性能计算方案实践，也展现了进一步释放现有算力价值的可能。比如谷歌云提出“算力多切片训练”方案，以实现超出一一般高算集群负载的超大规模 AI 训练。

量超融合

量超融合成为行业共识：通过将计算任务在量子计算机和高性能集群之间进行分解和调配，实现量超协同，在大幅节约资源的情况下，双向发挥量子计算机和超级计算机各自优势。预计连接两者的量子测控系统将成为量超融合“三步走”的关键系统之一。



2023年11月，芬兰VTT技术研究中心开放量子计算机，并通过LUMI超级计算机进行访问

量子集群出现、云服务普及

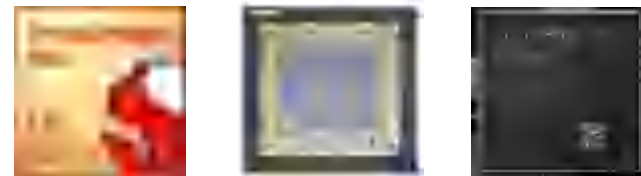
量子计算集群出现：IBM 公布可扩展量子计算系统 IBM Quantum System Two，引入模块化架构和计算机集群设计，可让系统容纳 10000 个量子比特。



量子云服务普及：这有助于上层软件和应用的高效研发。全球已有超过20家机构推出该服务，机构从整机公司（如IBM），发展到云服务商（如AWS），再到科研院所（如北京量子院）。

高性能计算走向云边协同

高性能边缘计算（HPEC）成长中：目前HPEC单点算力可达千万亿次每秒，自动驾驶和机器人是典型应用场景。除传统计算平台外，神经形态芯片、光电计算芯片等新技术的成熟也在不断提升边缘算力和能效。



可以运行文生图大模型的骁龙 8Gen3 很快将面世。

清华大学研发的 ACCEL 光电计算芯片在视觉 AI 任务中，比同用途常规芯片快三千余倍，能效提升四百万余倍。

IBM神经形态芯片 NorrthPole 比相同制程的 CPU 芯片能效提升 25 倍。将在自动驾驶等场景试点。

云边协同：为获得实时计算决策以及更高的信息安全环境，将部分数据处理、模型训练以及推理等工作，从数据中心/云平台迁移至云边协同架构下，是必要的技术路径。

趋势要点2：高性能芯片多元发展，QPU 发展由数量转向质量

- 既谷歌和亚马逊，微软、Meta 今年先后发布自研高性能芯片；而老牌芯片公司也相继补齐 CPU 和 GPU 产品矩阵；同时，ARM和RISC-V芯片在高性能领域的应用探索也在增加，高性能芯片性能之争加剧；
- 冯诺依曼架构的“内存墙”越发明显，业界纷纷采用存算一体技术，提升现有芯片性能的同时，研发神经形态计算芯片，探索“破墙”之路；
- 量子比特制备的各条技术路线均有进展，竞争更加激烈；而芯片互联、逻辑量子比特制备等突破，显示量子计算芯片发展开始追求质量和实用。

芯片技术多元化发展加剧

众多科技企业纷纷加入高性能芯片大战的同时，芯片技术的多元化发展进一步加剧。

微软的ARM 架构 CPU Cobalt100

Meta RISC-V 架构 SoC AI 加速芯片

位于葡萄牙的高性能计算集群 Deucalion 将采用富士通的 ARM 架构高性能芯片 A64FX。

算能科技日前向山东大学交付 国内首台 RISC-V 服务器集群，该集群拥有 48 颗 RISC-V 高性能芯片 SG042。

存算一体，突破“内存墙”

英伟达、Intel、AMD 等纷纷采用 HBM和 LPDDR 等内存技术实现近存计算，内存带宽突破5TB/s，算力再提升。



AMD MI300X 采用 192GB HBM3e 内存



英伟达 Grace CPU 采用 960GB LPDDR5X 内存

神经形态计算 (Neuromorphic Computing)

是“后冯诺依曼时代”突破“内存墙”的重要计算技术路径。受人类大脑原理的启发，神经形态计算芯片通过存内计算 (Compute in Memory) 方式，实现高算力的同时实现超低功耗。目前该类芯片尚处研发早期，Intel, IBM, 中科院和清华等发布阶段成果。

量子比特制备路径之争将更激烈

多条技术路径并进：超导、离子阱、光量子、冷原子等路径均有进展，竞争将更激烈。



九章三号首次突破255个可控光量子数目。



Quantinuum推出32量子比特GHZ状态的新一代H2离子阱量子计算机。



Intel 推出名为Tunnel Falls、含有12个硅自旋量子比特的量子芯片。



AWS (亚马逊) 发布超导量子芯片，提出“主动+被动”纠错方案，专注提升纠错能力。

QPU 发展从数量到质量



千级量子计算机达成：IBM 发布了1121量子比特的 Condor 处理器；Atom Computing将在24年推出全球首台超过1000量子比特的中性原子量子计算机。

芯片互联出现：IBM 发布133比特具备互联结构的 Heron 处理器。这是国际首次完成芯片互联的结构，模块化量子计算机时代即将到来。



逻辑量子比特制备突破：哈佛大学、QuEra、MIT和NIST/马里兰州大学联合实现在48个逻辑量子位上执行复杂纠错量子算法。

趋势要点3: 云原生技术将释放高性能算力潜能, 科学计算模拟应用大量增加

- 云原生技术的加快普及, 以及虚拟化和容器化技术与传统计算变成平台的结合, 将加快AI 计算和传统科学计算应用的大规模迁移, 高性能算力的潜力将进一步得到释放;
- 人工智能大模型的研发、科学计算应用软件的迁移, 以及科学计算建模算法与人工智能技术的结合, 将比预计的快许多, 科学家和科研团队的生产力将迅速提高, 科学研究将迎来快速进步的新形势。

云原生技术加快普及

云平台容器服务将加快对高性能计算服务的支持, 并进一步兼容CUDA 等并行计算平台, 为用户提供高度自动化的高性能计算工作环境, 大大提高研发和实验效率。

云平台	HPC云服务	容器服务	计算平台
谷歌云	H3, A3	GKE	CUDA
	TPU v5e		OpenCL
亚马逊云	EC2 UltraCluster	EKS	OpenACC
微软云	NC H100 V5	AKS	OpenMP
英伟达云	DGX Cloud	NCT	SYCL
IBM云	VPC	KS	ROCm
腾讯云	THCC	TKE	oneAPI
阿里云	SCC	ACK	
百度云	CHPC	CCE	
华为云	GACS	CCE	

+

科学模拟应用和成就将迎来大爆发

算力骤增, 算法升级: 传统高算集群正在迅速的向CPU+ GPU的计算架构升级, 如芬兰的 LUMI; 同时, 全球新建高算集群也如雨后春笋, 如德国的JUPITER 和沙特的 Condor Galaxy; 而云超同化使各地高性能算力更易访问; 此外, **许多传统模拟数值算法经 AI 技术优化后, 性能大幅提升。**

需求强烈: 各国科研团队、企业和研究机构纷纷基于高性能算力展开各项科学计算模拟实验。



量子电路模拟



核聚变模拟



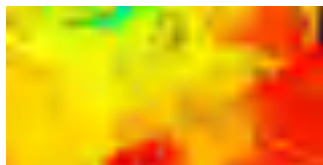
血流和癌细胞模拟



飞机材料腐蚀模拟



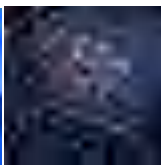
大涡模拟



气象模拟与预测



地理空间模拟



暗物质模拟

更多高算集群尝试 LLM 训练

传统高算集群通过加装 GPU 集群、结合云平台高性能计算服务得到升级后, 将具备更强的 AI 加速计算能力, 进而可以支持人工智能大语言模型的训练。

- 智谱 AI 的 ChatGLM3支持在神威高算上运行;
- 芬兰将基于 LUMI 创建项目 LumiLLM, 研制多语言模型, 全面覆盖欧洲官方语言;
- 美国阿贡国家实验室开始基于 Aurora 创建 ScienceGPT, 预计参数数量达到 1 万亿, 将为广泛的科学研究提供帮助。



Aurora 高算集群于今年完成安装, 理论峰值算力达 2E。

趋势要点4：业界将加大可持续高性能计算的践行力度

- 可持续计算，是数字经济和双碳目标背景下，高性能计算技术演进历程中的路标和灯塔。拥有可持续高性能计算的技术，也将成为重要的竞争力；
- 环境可持续的同时，计算能效将成为评估高性能计算技术先进性的重要指标，存算架构、冷却技术、计算机软硬件和供需匹配3条路径将并行展开探索；
- 机器学习、神经网络等人工智能技术与传统科学计算建模的结合，将有助于大大提升建模效率和模拟精度，进而提高能效；

践行可持续的高性能计算

以TOP500 排名第一的Frontier 高算集群为例，其算力可达 1.6EFlop/s，功率可达 20 兆瓦，相当于近 1万户家庭用电水平。

据统计，数据中心的耗电量约占全球耗电量的 1%，我国 2020 年占比是2.7%。虽然当前尚无准确的数据中心能耗预测，但急剧增加的人工智能大模型训练等高耗能计算应用，势必会使这个数字继续增加。

可持续计算，通常指在设计、建造和使用计算机技术（包括硬件、系统、软件）的过程中，力求实现**最大的能源效率和对环境影响的最小化**。

世界经济论坛发布《2023 十大新兴技术》报告，“可持续计算”位列其中。Intel、IBM、英伟达等先进计算企业也在不遗余力的践行和推动可持续计算。

计算能效将更受关注

包括Green500、CCF、以及业界机构与企业，均采用“**浮点运算次数/每瓦特 (Flops/Watt)**”作为能源效率评估的核心指标。

2023 年 11 月 Green500 排名第一的是位于美国的 Henri 高算集群，能效达到 65.4GFlops/Watt。

更合理的算力供需匹配

更先进冷却技术

更高能效的计算软硬件

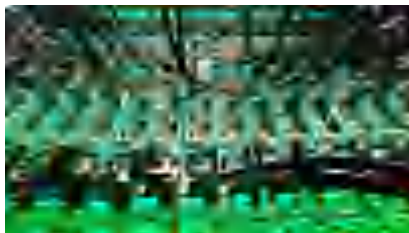
提高计算能效的3个途径

追求“量子效用”

IBM 提出“当量子计算机能够以超出暴力经典计算方法的规模执行可靠的计算，从而为计算问题提供精确的解决方案时，即能获得量子效用”。同时建议基于 100 多个量子位来探索能获得量子效用的应用案例。

浸没式液冷加快普及

传统风冷所需能耗可占数据中心总能耗的 20-30%。液冷，特别是浸没式液冷能耗相比风冷低 30%以上。同时，利用环境水冷却的水下数据中心也在实验中。



英特尔将 24 台装有至强处理器的服务器，置于一个充满合成且非导电油的槽中冷却。

高效软硬件持续探索

高带宽内存：H200 采用高带宽内存 HBM3e，能耗降至 H100 的一半；

超低功耗计算单元：神经形态计算具有超低功耗下实现高算力的潜力，值得持续探索。



清华大学研发出全球首枚忆阻器存算一体芯片，能耗仅为同类AI用途ASIC的 3%



IBM 研发出基于相变存储器、用于深度学习推理的存算一体芯片 AIMC，在 8 位向量计算任务中，能效可达 9.76 TFlops/w

AI+科学模拟：深势科技团队将机器学习技术与科学模拟相结合，复现了今年戈登贝尔奖获奖实验，而相比获奖方案，复现实验所需资源减少四个数量级，模拟速度提升三个数量级。

**多模态智能体
加速 AGI 进程**



多模态AI智能体的生成未来

- 生成式AI推动技术迈入了通用AI的门槛，从理解到生成，从感知到决策，人工智能的能力进一步提升。加上多模态、Agent、以及具身智能等方向的持续探索，AI有望完成“感知—决策—行动”的闭环。
- 迈入AGI门槛后，AI将有望成为通用目的技术，进而成为新生产力，给全球经济、社会带来全方位的巨大影响。



01



多模态、Agent和端侧

从文本到图像、音视频等更多维度，多模态将推动AI应用的广度；Agent为AI带来“手和脚”，大幅拓展应用深度。端侧大模型成为各大硬件厂商重点，有望带来新生态。

02



AI for Science

生成式AI对科学研究的助力日益凸显，“AI科学家”有望来临。

03



价值对齐

RLHF和宪法性AI等多种技术和治理方式，持续强化与人类价值的协调，对齐也成为大模型的重要竞争力之一。

发展阶段、学科背景、热点方向

弱人工智能

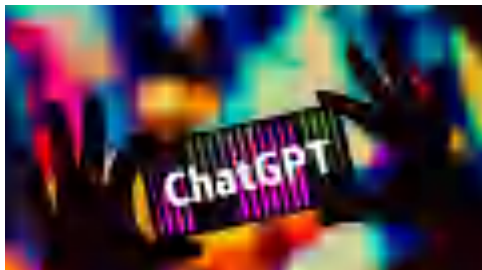


GPT4迈进了AGI的门槛



在特定领域协助或主动完成某些任务。如自动驾驶、医疗对话等领域强于人类。

通用人工智能



能否进化到下一阶段，存在争议



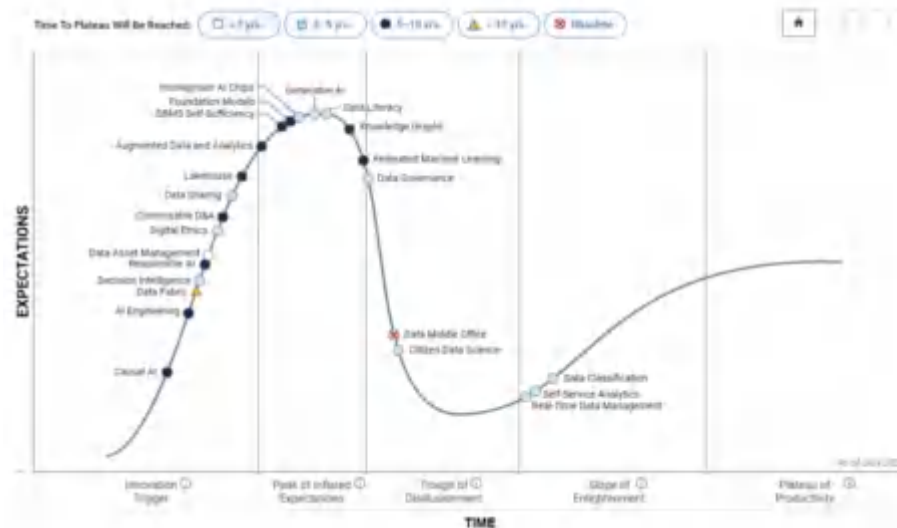
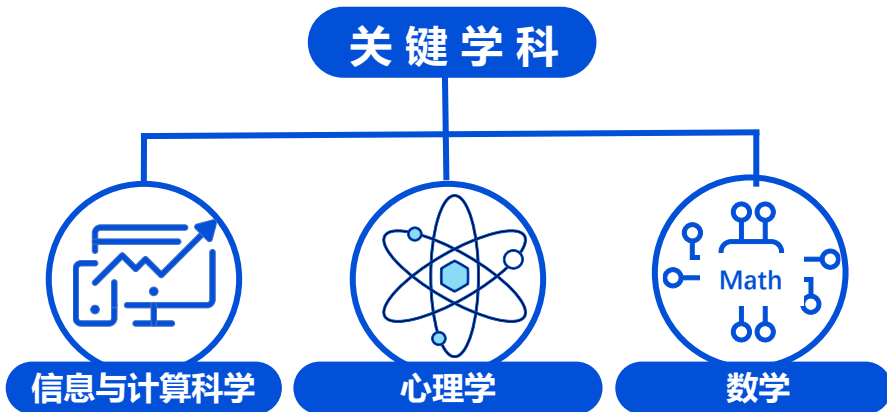
具有可泛化的任务执行能力。可通过图灵测试，让人类误以为是人类，或通过大学考试。

强人工智能



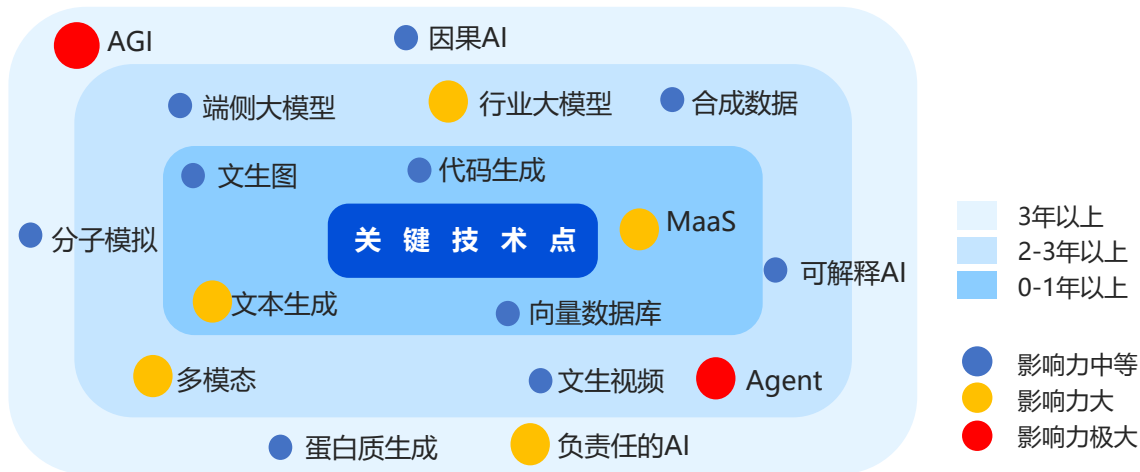
明显比人更有智慧，或解决问题能力明显超越人类，帮助解决危害人类的问题。

关键学科

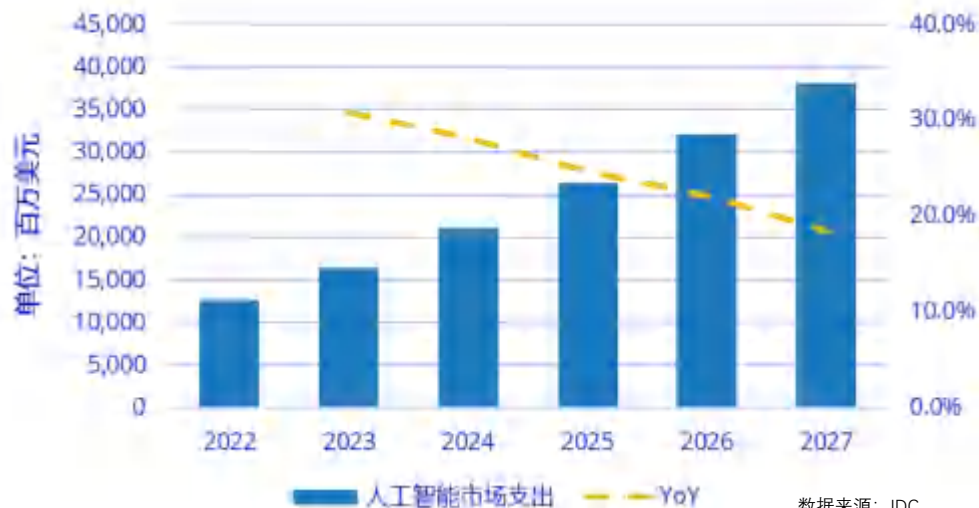


Gartner: 生成式AI是当前炒作顶点

技术预测



行业预测



2023 chronicle of events 大事记

GPT4发布

3月, 在ChatGPT发布4个月后, GPT4正式发布, 带来新的多模态识图能力演示, 回答的准确性相比GPT3.5大幅提升了40%。其人类考试的成绩较GPT3.5大幅提升, 达到Top 10%水平。
11月6日, Open AI开发者日发布GPT-4 Turbo, 允许用户上传资料, 生成自定义的智能助理。同时开放应用商店, 打造类似APP Store的应用生态。

12家中国AI大模型首批备案通过

8月31日起, 12家大模型陆续通过《生成式人工智能服务管理暂行办法》备案。北京5家, 百度“文心一言”、百川智能“百川”大模型、抖音“云雀”、中科院“紫东太初”、智谱“智谱清言”; 上海3家, 商汤“商量 SenseChat”、MiniMax“ABAB 大模型”、上海人工智能实验室“书生通用大模型”; 安徽1家, 讯飞“星火”; 浙江1家, 阿里“通义千问”; (5) 天津1家, 360智脑; (6) 广东1家, 腾讯“混元”。11月, 第二批大模型备案通过11家。至此, 国内已形成一批初见水平的AI创业企业, 如 minimax、智谱、百川、月之暗面 (moonshot)。

全球高度关注AI治理

11月1日, 由英国倡议发起的首届“全球AI安全峰会”在伦敦布莱切利园开幕, 布莱切利庄园是二战期间盟军破译密码的主要地点。会议首日, 美国、中国、欧盟和大约20个国家共同签署了《布莱切利宣言》, 重点在于识别人工智能产生的共同关注风险, 建立对风险的科学认知, 并制定跨国风险缓解政策。

趋势要点1: 多模态解析世界的本来面貌, 并实现“三生万物”

多模态是人类世界的本来样貌, AGI的发展趋势一定是朝向多模态。技术将从单一的文本、图像、视频(2D和3D), 再到声、光、电, 甚至分子、原子等各类模态, 而且具备跨模态迁移的特性。通过设计不同模态对应的解码器, 可以实现多模态之间的切换。未来理想的框架是: “多模态的对齐和融合 + 统一的编码器和解码器”, 可以更好地解决多模态的理解和生成任务。

从LLM走向LMM

多模态大模型百花齐放:

Blip2、InstructBlip、LLaVA、VisualGLM、MiniGPT5、VideoLLaMA、ImageBind、Meta-Transformer、GPT4V、Gemini、Pika、Runway、混元、文心一言、ChatGLM、MiniMax等。

多模态具有更突出的优势:

- 带来了更灵活和友好的用户交互。
- 是一个更完善的任务求解器, 可以使用更多样化的任务类型。

主要的实现方法:

- 多模态指令、多模态上下文学习、多模态思维、LLM 辅助视觉等

从专业单反到“傻瓜相机”

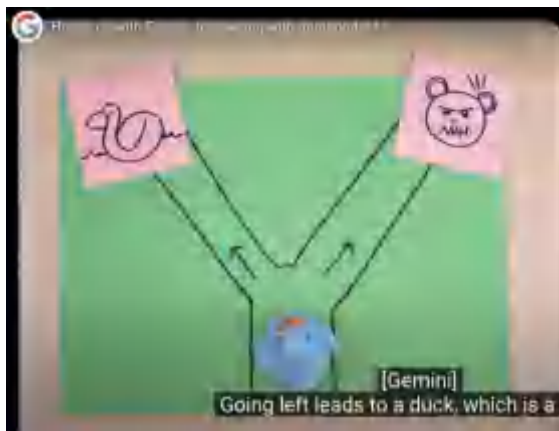
将LLM(大语言模型)和文生图相结合, 让用户可以不必再学习复杂的Promot技巧, 直接用日常语言提示, 即可生成专业级的图片。这种方式, 巧妙地利用了LLM的理解能力, 进一步提升了文生图的效果。



腾讯混元生成图片“醉后不知天在水, 满船清梦压星河”

图读世界

多模态大模型可以实现对图像和视频的理解, 包括数量、类型、空间关系、人物、地点、事件、时间序列, 以及图像中隐含的信息(如搞笑图、医学诊断等)。



Google Gemini的演示

超级“变形金刚”

北大联合腾讯提出LanguageBind的多模态预训练框架。用语言作为与其它模态之间的纽带, 冻结语言编码器, 然后用对比学习方法, 将各个模态映射到一个共享的特征空间, 实现多模态数据的语义对齐。



由香港中文大学多媒体实验室和上海人工智能实验室 OpenGVLab联合开发的Meta-Transformer可以应用于许多应用领域, 包括3D识别、夜间安全、天气预报等。

趋势要点2: 从大脑到Agent, 大模型从CoPilot副驾, 走向主驾驶

- Agent是指能够独立思考、自主行动并可以与环境交互的软件程序或机器人等实体。Agent包含三步: PPA, 即感知(Perception)--规划(Planning)---行动(Action)。
- 人工智能之父、图灵奖得主马文·明斯基 (Marvin Minsky) 在1986年出版了一本里程碑式的著作《思维的社会》(The Society of Mind), 试图解读人类思维这个复杂的过程。Minsky认为社会中的某些个体经过协商之后可求得问题的解, 这些个体就是Agent。Agent应具有社会交互性和智能性。
- 比尔盖茨撰文, 表示AI Agent将是下一个平台, 人工智能即将彻底改变人们使用计算机的方式并颠覆软件行业。在不久的将来, 任何上网的人都将能够拥有由人工智能驱动的个人助手, 远远超越今天的技术水平。

Agent发展如火如荼



Figure 1: Illustration of the growth trend on the field of LLM-based autonomous agents.

《A Survey on Large Language Model based Autonomous Agents》

LLM+Memory+Planning+Tool use

Agent是主动释放LLM潜能的关键, 让大模型从“有脑”到“有手有脚”

调用更便捷

在大模型的加持下, 根据会话自动调用相关的插件

开发门槛更低

以自然语言的方式完成插件、各类应用的开发和部署

AI时代的新入口

每个人将能拥有由AI驱动的个人助手, 无需再根据需求下载不同的APP, 把需求告知代理即可

对话式编程

微软AutoGen发布短短两周内, 星标量从 390 增到 10K, 并在 Discord 上吸引了 5000 多名成员。



AutoGen框架允许多个 LLM 智能体通过聊天来解决任务。LLM 智能体可以扮演各种角色, 如程序员、设计师, 或者是各种角色的组合, 对话过程就把任务解决了。

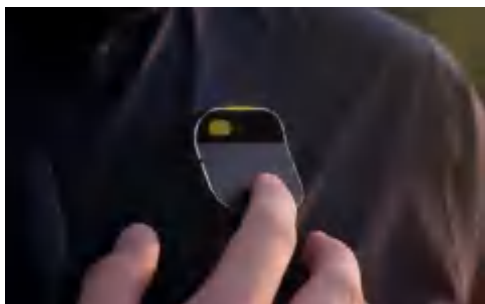
趋势要点3: 端侧大模型加速部署, 或将成为未来交互新入口

- 大模型正在向端侧转移, AI 推理将在手机、PC、耳机、音箱、XR、汽车, 以及其它可穿戴式新型终端上运行。目前, 一些手机已经在利用本地 AI 支持如暗光拍摄、降噪和人脸解锁等功能。未来借助端侧大模型, 并结合向量化后的各类个人数据, 用户可以跟手机进行更流畅的交互, 实现各种原生操作和功能。
- 端侧大模型应用的三种可能: (1) 端侧原生集成AI模型, 类似siri, 帮助用户调用其他软件, 从而可能成为硬件新入口 (2) 将大模型作为独立app, 例如MIT一位教授将开源模型集成在手机端做一个独立app; (3) 将大模型接入即时通讯软件作为chatbot, 例如What 's App已经集成了Meta AI。

新型终端



新一代 Ray-Ban Meta AI 互动



AI PIN等原生大模型终端试水

智能车



FSD V12或重启中国计划; 自研 DOJO超算架构提升大模型能力



联想推出AI PC

手机终端



谷歌Pixel 8 Pro 发布了 AI Core 应用首个更新, Gemini发布端侧 Nano 版

手机端侧大模型

苹果	AjaxGPT, 性能超GPT3.5, PC上2000亿参数
小米	13亿和60亿参数, 面向手机、汽车和机器人
VIVO	130亿蓝心大模型实现端侧跑通, 开源7B大模型
OPPO	安第斯 (AndesGPT), 10亿参数
荣耀	魔方, 70亿参数



高通骁龙8 Gen3 支持终端侧运行100亿参数的模型

端侧大模型的主要优势:

1. 本地数据处理效率更高
2. 节省云端服务器带宽和算力成本
3. 对用户数据更好的隐私保护
4. 开启更多交互新方式、新体验

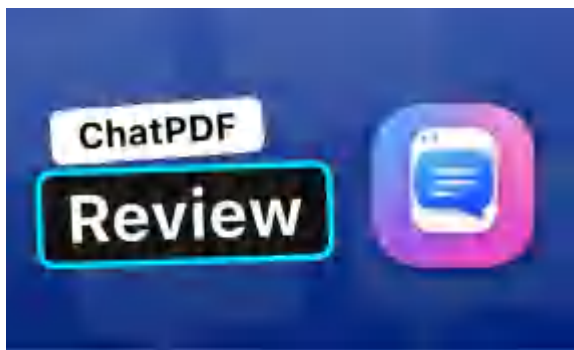
趋势要点4: AI4S助力科学探索, 贯穿科研全流程

- 在AI技术的加持下, 科学研究的效率大大增加。通过理论探索、设计实验、分析数据等方向为科学发现提供动力。AI与各个科学领域结合后, 正在发生一场充满潜力和挑战的科技革命。

AI文献阅读与问题提出 → 数据优化处理 → 数据分析挖掘 → 科研大模型

文献阅读及问题提出

对现象和数据观察提出某种猜想; 对既有文献的梳理, 在此基础上提出新的研究问题, 帮助科学家发现新问题。

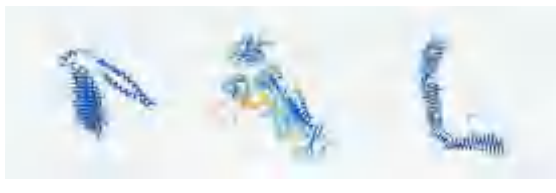


- 科研神器ChatPDF可以用十秒解析上万字的研究文献, 分析文献结构、定位内容, 提高阅读效率。

数据优化和处理

AI可以通过大量数据和复杂的数据分析, 帮助提取有用信息和填补数据缺失, 从而提高数据质量和利用效率。降低对数据的依赖。

数据分析和挖掘



AlphaFold已预测出超过100万个物种的2.14亿个蛋白质结构, 几乎涵盖了地球上所有已知蛋白质。AI可以帮助科学家预测蛋白质的结构, 从而优化药物研发过程。

AI成为生命科学新引擎

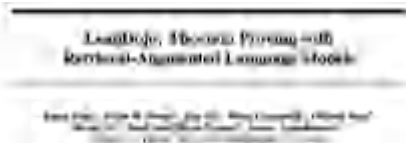
生成式AI成为基因科学的新引擎: 生成式AI被用于预测基因序列、发现新的药物靶点以及设计新型的生物材料。

- 基因数据分析
- 蛋白质表达及测量
- 基因序列分析

生成式AI与科研

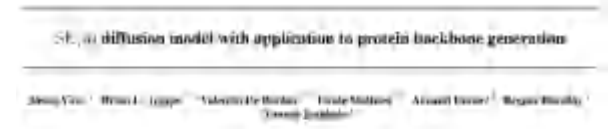
菲尔兹奖得主, 数学家陶哲轩认为, 预计如果使用得当, 到2026年, AI将成为数学研究和许多其他领域值得信赖的合著者。

一篇用GPT证明数学定理的论文引发巨大关注



自然科学领域大模型

DARWIN, 是一个为物理、化学和材料科学应用而精心设计的专业化大语言模型(LLM), 在多个科学任务中取得了最先进的结果, 旨在通过人工智能驱动的自动化来增强和加快探索发现的过程。



麻省理工学院(MIT)的研究人员开发出一种扩散模型——FrameDiff, 该模型能够生成自然界中不存在的新型蛋白质结构。该模型可以生成多达500个氨基酸序列的蛋白质主链, 且无需依赖于预训练蛋白质结构预测网络。

趋势要点5.1: 价值对齐是大模型的必由之路, 将成AI产品的核心竞争力

- 随着大模型, 尤其是LLM成为人机交互的新界面, 并涌现出强大的推理和内容生成能力, 大模型的价值对齐日益成为一个关键问题, 即让大模型的能力和行为跟人类(使用者)的价值、伦理原则和真实意图相一致, 确保人类与人工智能协作过程中的信任与安全。
- 业界和研究人员已在探索实现大模型价值对齐的技术和治理措施, 诸如人类反馈的强化学习(RLHF)、宪法AI(Constitutional AI), 可扩展监督(如AI监督)、模型可解释性等, 以构建更加安全可靠且有用的大模型。未来对齐技术的发展, 需要更好将人类监督和AI监督结合起来。

RLHF是当前比较有效的对齐技术

RLHF包括初始模型训练、收集人类反馈、强化学习、迭代过程等几个步骤, 其核心思路是要求人类训练员对模型输出内容的适当性进行评估, 并基于收集的人类反馈为强化学习构建奖励信号。



AI对齐的多种技术和治理措施

训练数据干预

- 构建价值对齐的专门数据集。
- 对训练数据进行记录, 以识别是否存在代表性或多样化不足的问题。
- 对训练数据进行人工或自动化筛选、检测以识别、消除有害偏见。

可解释、可理解的大模型

- 为了实现AI价值对齐, 人们需要理解人工智能如何作出决策。
- 例如OpenAI利用GPT-4来针对其大语言模型GPT-2的神经网络行为自动化地撰写解释并对其解释打分; 有研究人员则从机制解释性 (mechanistic interpretability) 的角度来应对AI对齐问题。

红队测试

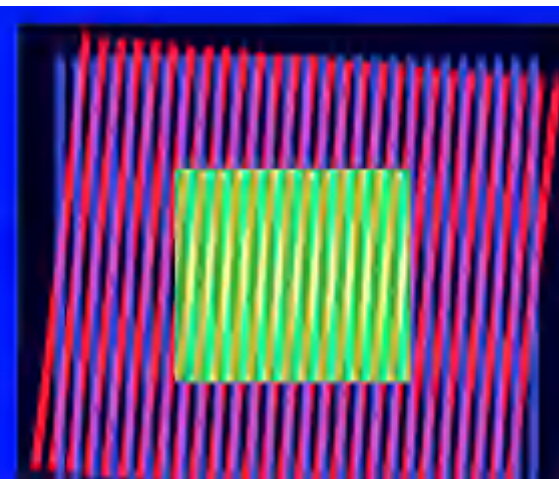
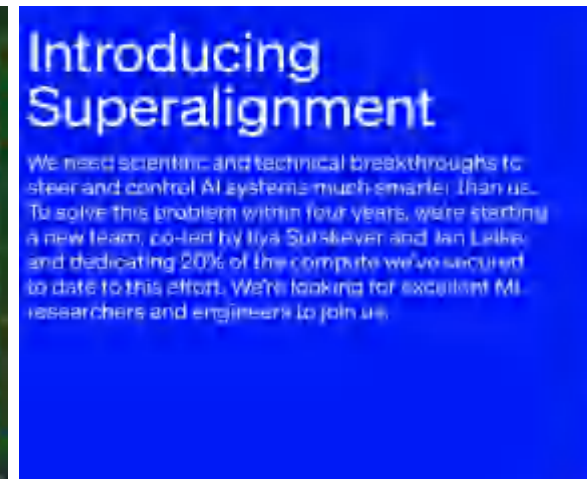
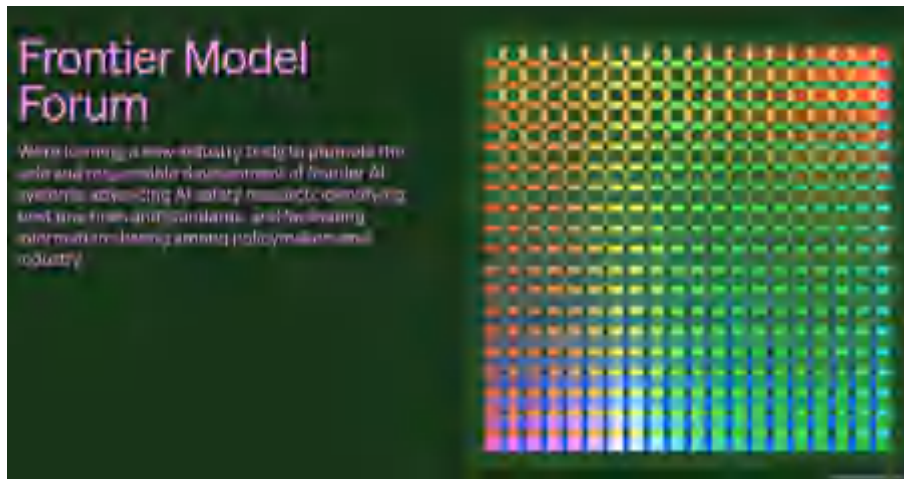
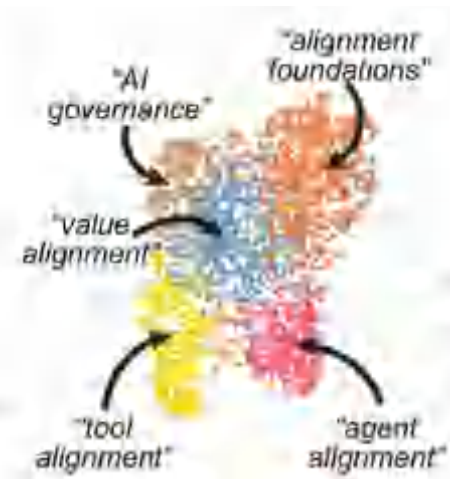
- 在模型发布之前邀请内部或外部的专业人员, 对模型发起各种对抗攻击, 以发现潜在问题并予以解决。

基于检索或外部工具的对齐

- 将要对齐的价值观保存为一个外部模块, 大模型回复时需要检索出要对齐的价值观并作出合适的回复。
- 通过外部工具(搜索引擎、代码编译器、计算器等)辅助回复的生成, 以确保生成内容的真实性。

趋势要点5.2: 以价值对齐确保人工智能的未来, 更好实现科技向善

- **凝聚行业共识**, 推动形成**大模型价值对齐的技术和伦理指南**, 总结推广最佳实践。
- **鼓励开放研究**, 支持关于AI安全性、公平性、可解释性、价值对齐和其他伦理议题的研究, 并鼓励研究者公开其研究结果, 促进全球社区协作。
- **将人类监督和AI监督更好地结合**, 当下的AI价值对齐工作还面临一个关键问题, 即在人类的智能基本保持不变的前提下, 随着AI能力持续提升, 人类自己对前沿AI模型的有效监督将变得越来越困难。因此, 为确保AI安全, 需要使我们监控、理解、设计AI模型的能力与模型本身的复杂性同步发展。例如, OpenAI成立超级对齐团队, 核心是希望探索利用AI来帮助人类解决比人类强大的AI系统的价值对齐问题。
- **基于自然反馈对齐的探索**: 这是更天然的强化学习方式, 如传言Open AI的Q*, 一方面突破了人类数据的限制, 可以自己生产海量训练数据, 另一方面, 模型具备了自主学习和自我改进的能力。





AI加速人形机器人

“手、脑”进化

AI加速人形机器人“手、脑”进化

- 根据工信部印发《人形机器人创新发展指导意见》，人形机器人集成人工智能、高端制造、新材料等先进技术，有望成为继计算机、智能手机、新能源汽车后的颠覆性产品，将变革人类生产生活方式，重塑全球产业发展格局。特别是在关键技术突破方面，打造人形机器人“大脑”和“小脑”、突破“肢体”关键技术、健全技术创新体系。
- 2023年以来，基于视觉-语言的大模型嵌入机器人本体，加速思考（“感知脑”）和执行（“灵巧手”）突破，推动机器人迈向智能化带来新的可能。在思考能力层面，大模型的嵌入极大提升机器人感知环境、分解任务、规划流程以及与环境交互的能力。同时，云边结合的分布式计算平台发展，强化了机器人的训练和分析决策速率。在执行层面，以灵巧手为代表的关键技术，进一步强化了人形机器人末端执行应用能力，集中体现在微操作、近操作等环节。
- 随着视听触多模态、端侧算力、运动控制技术的进步，人形机器人技术将加速迭代，走向更柔性、更智能、更灵巧。



01



大模型

大模型的嵌入极大提升机器人分解任务、规划子任务，与环境交互的能力

02



多训练平台

云边结合的分布式计算平台发展，强化了机器人的训练和分析决策速率

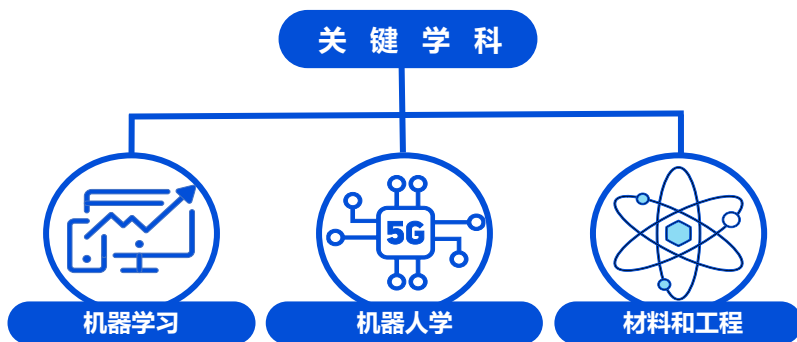
03



灵巧操作

多感知、多自由度功能融合、微型驱动器产业化，强化了机器人的微操作、近操作能力

关键学科、行业趋势、典型事件



人形机器人市场规模2028年有望增长到138亿美元



- Markets and Markets对人形机器人市场预测, 市场规模将从2023年的18亿美元增长到2028年的138亿美元, 其复合年增长率可达50.2%。
- 高盛预测在最理想的情景下(产品设计、用例、技术、可负担性和公共接受度等障碍被克服), 2035年人形机器人市场或将达到1540亿美元。

(资料来源: Markets and Markets, 2023)

2023

chronicle of events

大事记

0预训练完成复杂指令

7月, 李飞飞团队发布最新成果VoxPoser具身智能技术, 可以使机器人直接听懂人类的自然语言指令并完成复杂任务, 无需额外的数据和训练。

特斯拉擎天柱再次升级

人形机器人Optimus快速迭代, 商业化进程持续推进。特斯拉擎天柱的力控能力、抓取复杂物体的能力、利用FSD技术的视觉感知和处理能力。

多家科技公司布局具身智能

OpenAI投资了人形机器人公司1X, 英伟达CEO黄仁勋也公开唱多“具身智能”。微软基于ChatGPT 自然语言理解和推理能力, 生成控制机器人的相关代码。腾讯通过将预训练 AI 模型和强化学习技术应用到机器人控制领域, 使得Max能够在复杂环境中展现出栩栩如生的运动行为, 并学会障碍追逐比赛的智能策略。

国家政策支持创新体系建立

工业和信息化部关于印发《人形机器人创新发展指导意见》的通知, 部署到2025年, 人形机器人创新体系初步建立, “大脑、小脑、肢体”等一批关键技术取得突破。

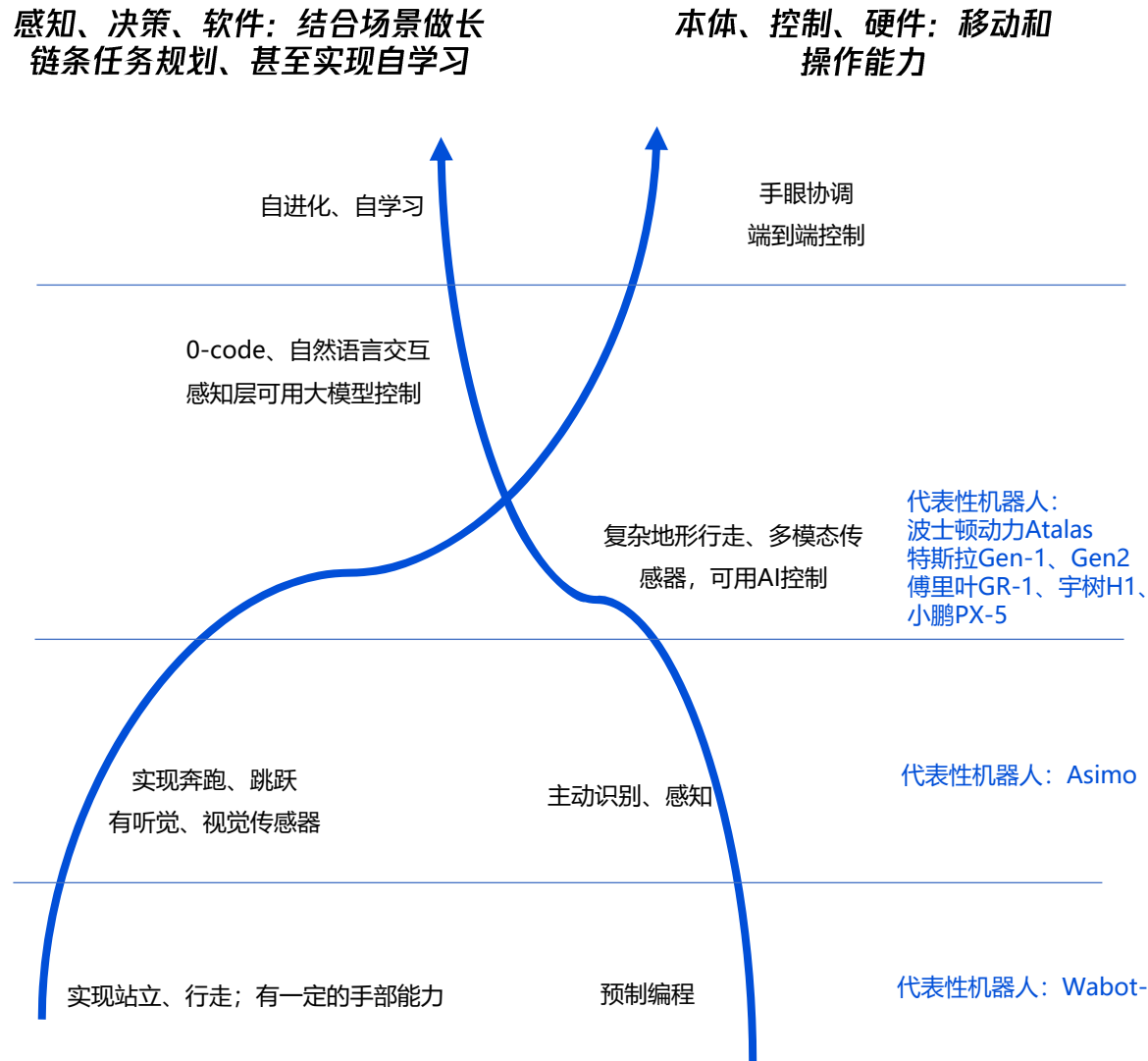
关键技术点及成熟度



1848年，图灵在“机器智能”的文章中，首次提到“嵌入式智能（肉体智能）”（embodied-intelligence）和“非嵌入式智能（无肉体智能）”（disembodied-intelligence），这被公认是“具身智能”概念的首次提出；值得注意的是，1950年这两个核心概念进一步发展为“智力”和“体力”，进一步将AI和机器本体做了区分。

从Wabot-1开始，机器人的“硬件”和“软件”技术开始了各自发展到相互融合的螺旋上升，研发主体从AI和机器人本体两条路线逐渐汇聚。当前，以特斯拉Tesla-bot为代表的一系列人形、AI机器人颇受瞩目。具身能力是这一阶段机器人的技术前沿布局方向。

目前技术还在发散期，未来3-5年将进一步收敛



趋势要点1：大模型在文本、自然语言和视觉领域取得重要突破，提升机器人的理解能力

- 英伟达创始人黄仁勋在ITF World 2023大会上表示，人工智能的下一次浪潮将是具身智能，即能理解、推理以及与真实物理世界互动的智能系统。在Chatgpt出现之前，大模型与具身智能领域的结合更多出现在感知层面，在海量数据集上预训练的视觉模型只是作为一种更好的表征提取器来提升机器人在场景中的感知能力，而具体的规划与动作执行依然需要大量的具身场景数据进行训练。AI大模型，以及后续的多模态，可以从语音、视觉感知、决策、控制等多方面为机器人更好进行学习训练和进化。
- 当前，算法的难点在于对随机性的自适应能力和尝试规划等问题。大模型目前可以完成离散动作的决策和选取，把通用任务进一步描述分解成个别动作，未来将持续发展算法的泛化能力，向端到端大模型发展。

李飞飞和斯坦福： Vox Poser实现具身智能技术突破

过往机器人的控制模式是预设轨迹，导致机器人的控制与行动比较受局限。而李飞飞团队在2023年发布的VoxPoser系统，将LLM（大语言模型）和VLM（视觉语言模型）接入机器人，前者用来理解人类指令并生成交互代码，实现与后者的交互，而VLM进行规划路径，生成操作指示地图3D Value Map。

两者能力结合从而实现通过自然语言指令与机器人交互，可将复杂指令转化为具体行动规划，而无需预设数据和提前训练，同时，系统也具备很强的抗干扰能力，可以在遇到干扰因素时快速重新规划。

VoxPoser还产生了四个涌现能力：（1）评估物理特性；（2）行为常识推理；（3）精度矫正；（4）基于视觉的多步操作；



借助大模型，机器人获得应急能力，可应对突发状况



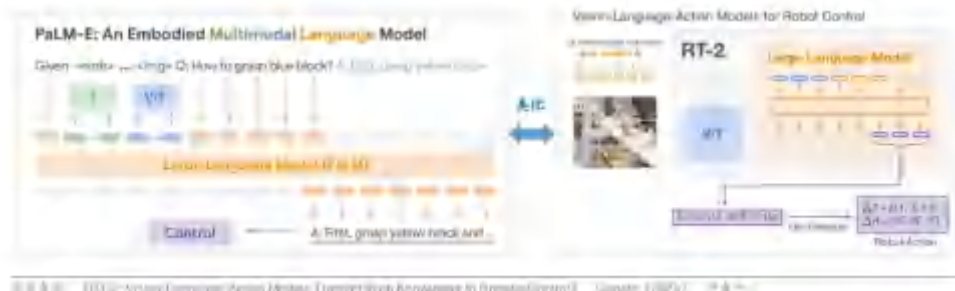
Voxposer技术原理示意：LLM和VLM的双向交互

谷歌：从语言模型到多模态模型综合布局

PaLM-E和RT-2采用了类似Token转化思路。

PaLM-E的技术前身是谷歌的“PaLM”大语言模型，模型参数规模达到5620亿，集成用于控制机器人的视觉与语言，特征是无需重新训练即可执行各种任务。通过添加感知信息以及机器人控制，实现了PaLM模型的具像化，使之能跟真实物理世界互动。PaLM-E具有“正迁移”的表现，这意味着它可以从一项任务学到的知识和技能迁移到另一项任务，较单任务执行机器人性能获得显著提升。

Robotics Transformer 2 (RT-2) 是由Google DeepMind突出的一个用于控制机器人的视觉-语言-动作 (VLA) 的AI模型。RT-2使用经过精调的LLM来输出运动控制命令，可移植性训练数据中未明确包含的任务，并在新出现技能评估中将基线模型的表现提升了3倍。



英伟达：将动作Token映射到手臂上

英伟达VIMA支持文本、视觉、语音等多模态作为机器人的任务输入；并使用预先训练的 T5 模型对多模态提示进行编码。在零样本泛化训练中，VIMA的任务成功率比之前的最优方法提高了最多2.9倍。



趋势要点2：云边结合的分布式算力平台发展强化了训练速率和数据质量，降低产业化门槛

- 由于机器人是 AI、算力、IoT、底层硬件等综合技术的集大成者，所以“具身智能”的进化速度取决于硬件和软件双轮驱动。当下，机器人技术受到训练数据短缺的限制。不同于大语言模型的数据集可以直接从互联网端获取，机器人的数据都是动态的数据集，需要在仿真环境中，获得动态模拟数据，也依赖于和物理环境的基础和交互，这些都需要一定的时间。
- 当前，机器人训练数据有两类，包括真实数据和合成数据。例如，谷歌RT-1模型是基于17个月收集了13万台机器人的13万条真实数据，而英伟达也在通过少量示范+大规模自动生成的技术路线，为机器人“脑”的发展，创造“数据粮仓”。
- 未来，通过云-边-端融合的机器人系统和架构，例如云端运行超大模型Nvidia A100 GPU，边缘运行小模型。让机器人达到数百万千万级水平，从而降低价格成本，实现大规模商用。

谷歌云：平台推动机器人自主决策水平

Google 专门开发协议，将模型部署在多TPU 云服务实现实时推理。实验结果表明，基于现成的VLM 模型PaLM-E 和 PaLI-X，RT-2 取得了良好的泛化和涌现性能。

英伟达：开发MiniGen可自动生成数据集

英伟达与得克萨斯大学奥斯汀分校合作发布了一片最新论文，介绍了一个名为“MimicGen”的系统，只需少量人类示范，便能自动生成大规模的机器人训练数据集。



特斯拉：人型机器人采用自动驾驶视觉神经网络算法和Dojo计算平台

特斯拉计划利用Dojo对海量的视频数据进行无监督学习，以加速特斯拉的Autopilot和完全自动驾驶（FSD）系统的迭代，同时为特斯拉的人形机器人Optimus提供算力支持。

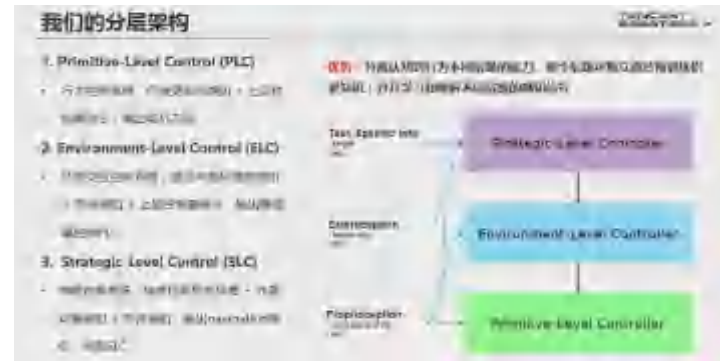
Dojo将加速人形机器人的开发，其能够为机器人的神经网络训练提供算力支持，更快速度地处理海量数据，有望推动机器人加速落地。



腾讯：通过游戏仿真引擎和真实狗动捕数据，搭建分层预训练模型

腾讯Robotics X机器人实验室和腾讯游戏合作，通过游戏技术推动了仿真引擎的准确性和高效性，并在游戏制作和研发过程中积累了多元的动捕素材。这些技术和数据对实验人员基于物理仿真的智能体训练，以及真实世界机器人策略部署，起到了帮助作用。

通过引入预训练模型和强化学习技术，可让机器狗分阶段进行学习，有效的将不同阶段的技能、知识积累并存储下来，让机器人在解决新的复杂任务时，不必重新学习，而是可以复用已经学会的姿态、环境感知、策略规划多个层面的知识，并“举一反三”。



趋势要点3：多感知、多自由度功能融合的灵巧手强化了人形机器人的微操作能力

- 机器人迈向通用智能的道路上，还需要突破工程方面的难题才能发挥更大生成价值，带来颠覆性应用落地。目前论文和一些demo展示的机器人与大语言模型结合偏重交互问题，而机器人仍需解决复杂环境的通用移动能力和高精度操作能力。
- 软件方面，强化学习控制算法按智能体(agent)是否理解环境与自身的动态模型可分为无模型强化学习算法与基于模型的强化学习算法。
- 硬件方面，灵巧手作为机器人实现操作的终端工具十分重要。空心杯电机、触觉感知传感器的发展，提升了人形机器人末端执行器的灵敏度和鲁棒性。

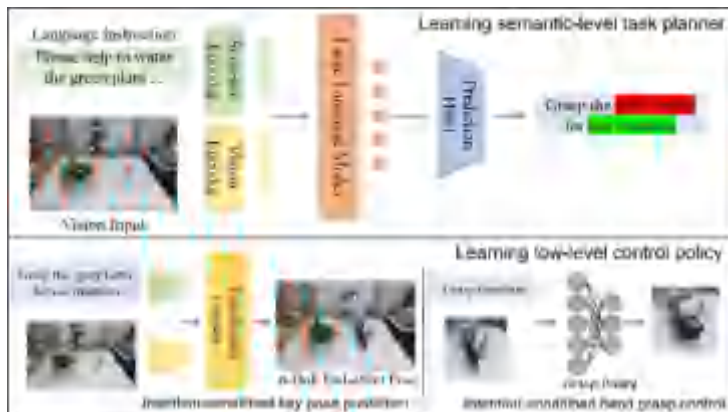
空心杯电机与灵巧手关节适配性高，特斯拉 Optimus 选用其作为动力源，带动产业热炒

内置式多指灵巧手是灵巧手发展的主要技术路线。特斯拉人形机器人擎天柱灵巧手有6个执行器，可实现11个自由度，其中空心杯电机为核心部件。



中国科学院自动化研究所提出基于视觉-语言大模型推理的灵巧手类学习框架

中国科学院自动化研究所王鹏研究员研究组研发的Casia Hand系列灵巧手及类人灵巧操作机器人在2023世界机器人大会上首次展出，课题组提出了一系列基于灵巧手的类人灵巧操作学习模型和算法，并在实际机器人上进行了验证。包括：基于视觉-语言大模型推理的灵巧手类人功能性工具操作学习框架。



腾讯研发拥有8个自由度、带触觉感知的灵巧手，可以完成调酒等复杂动作

腾讯Robotics X实验室公布最新机器人研究进展，首次展示在灵巧操作领域的成果，推出自研机器人灵巧手 TRX-Hand和机械臂 TRX-Arm。其中，灵巧手TRX-Hand拥有像人手一样灵活的操作能力，可适应不同场景，灵活规划动作，自主完成“操作”。



AI+基因计算

解读生命密码



AI加速驱动基因测序、基因编辑、基因合成发展

- 基因是生命的最基本元素。基因计算指使用计算方法来分析和理解基因信息的过程。
- 人工智能、大数据、云计算等数字技术正在广泛应用于基因检测、分析、预测、调控以及生物合成等方面，通过大算力的支持，充分利用多模态的海量基因数据，可以帮助我们解答生命科学的重要问题。
- AI与基因计算的结合，开辟了巨大的科学和应用前景，但同时也存在基因隐私保护、基因编辑伦理等问题。



01



生物育种

AI+生物技术加速育种4.0时代到来

02



生命健康

AI+基因计算助力实现个性化健康预测

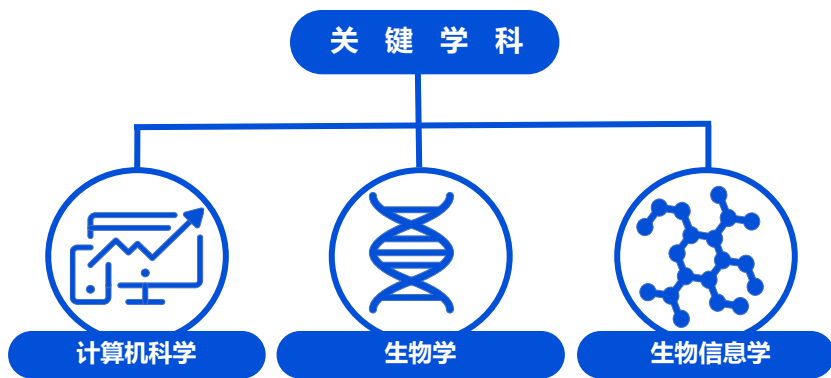
03



生物制药

AI+基因工程重塑分子药物设计和研发规则

学科背景、热点方向



根据BCC Research统计，全球消费级基因检测市场在2026年将达到377.2亿美元，2021-2026年亚太地区复合增速达 20.9%

根据Statista统计，2023年至2030年，全球基因编辑行业市场规模将以22.3%的复合增速进行增长，2023年市场规模约为66.19亿美元，到2030年，预计达到360.61亿美元

市场研究公司Global Market Insights Inc.发布报告预测，到2032年，基因组学行业中的人工智能市场规模将达到125亿美元，复合年增长率为39.2%

热点方向

基因数据处理技术

存储和管理海量生物数据的技术，支撑生物数据的快速查询和分析

基因表达分析技术

对基因组中的基因表达程度进行量化，帮助了解基因功能、调控机制及基因变异影响等

基因网络分析技术

利用复杂网络理论研究基因调控网络、蛋白质相互作用等，支持了解生物系统运作机制

基因智能设计技术

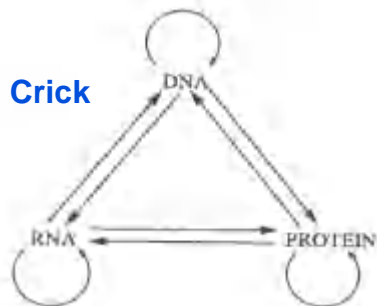
利用AI算法和大数据分析优化基因组设计和编辑过程，创造具有特定性状或功能的生物体

发展阶段

前基因组时代

生物数据库建立、检索工具开发及DNA/蛋白质序列分析

Francis H. C. Crick
提出中心法则



Frederick Sanger等
提出第一代DNA测序
技术



1958-1980s

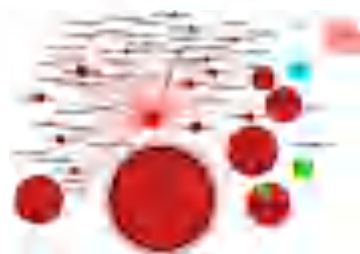
基因组时代

基因组测序和识别、网络数据库系统建立和交互界面开发

人类基因组计划,
完成92%基因组



第二代测序技术
(高通量测序)



1990s-2010s

后基因组时代

大规模基因组分析、多模态数据聚合处理、单细胞测序技术、基因编辑、基因疗法等

《科学》公
布迄今最完
整人类基因
组测序结果



基因
编辑
未来
应用
前景



AI分析海量
多模态基因
数据



2010s-

技术预测



2023 chronicle of events 大事记

基因组编辑

2023年1月，诺贝尔化学奖得主Jennifer A. Doudna教授团队在《科学》杂志发表文章《CRISPR technology: A decade of genome editing is only the beginning》；4月，中科院开发了PrimeRoot系统，在水稻和玉米中实现了长达11.1Kb的大片段DNA的高效精准定点插入，相关成果发表于《自然-生物技术》

基因预测模型

2023年9月，Google DeepMind团队基于AlphaFold开发了新的人工智能模型AlphaMissense，可用于判别人类的基因突变是否会导致疾病，该研究进展「Accurate proteome-wide missense variant effect prediction with AlphaMissense」发布在《Science》；11月，DeepMind宣布，AlphaFold预测范围从蛋白质扩展至DNA、RNA等

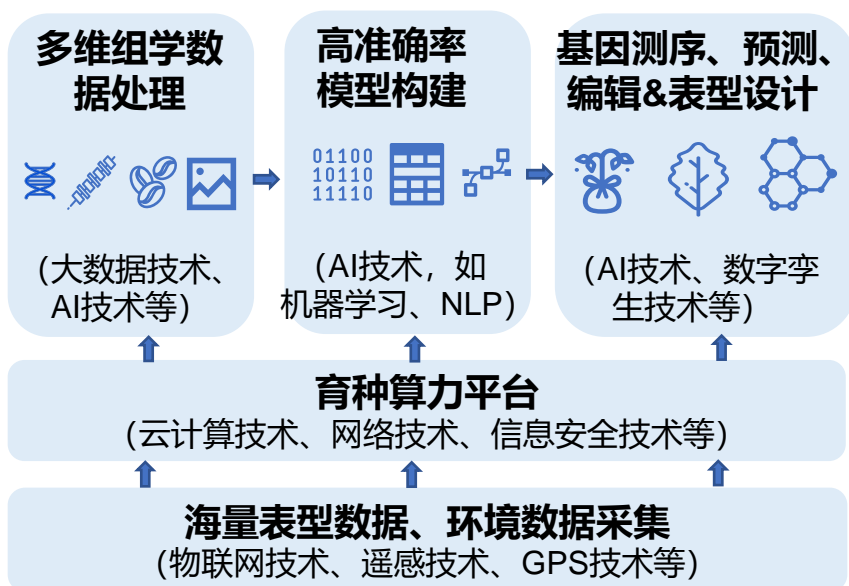
生物医药对话大模型

2023年9月，清华大学智能产业研究院、水木分子发布新一代对话式药物研发助手ChatDD (Drug Design) 和全球首个千亿参数多模态生物医药对话大模型ChatDD-FM 100B

趋势要点1: AI+生物技术加速育种4.0时代到来

- 基于AI+生物技术，进行标准化多维大数据关联分析和设计，利用基因编辑、合成生物等前沿技术，精准调节控制目标性状的基因和网络，选育新品种
- AI大模型强大的数据处理和模式识别能力，将用于支持对遗传信息的深度挖掘、表型特征的高精度预测以及育种过程的智能化决策
- 分子育种正在产生大量的表型和基因型数据，对海量基因测序数据、蛋白质结构数据、图像数据等数据的处理成为关键

育种4.0中的数字技术



表观预测：提升表观合成的精准度

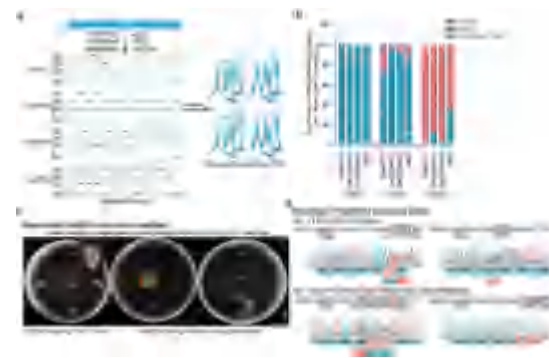
通过整合多组学数据，构建智能模型，预测并优化人工设计的合成表观回路，大幅度提升表观合成的精准度，推动合成表观技术在作物改良上的应用。中国农科院生物所提出利用合成表观技术设计和创制智能作物（SMART Crop）的途径和路线图。



来源：中国农科院生物所

基因编辑：提高作物产量和性能

2023年4月，农业农村部发布《2023年农业用基因编辑生物安全证书(生产应用)批准清单》。万建民院士团队开发了新型的多重正交碱基编辑器MoBE和随机化多sgRNA组装技术，获得了新的抗除草剂水稻突变植株。



来源：南京农业大学官网

趋势要点2: AI+基因计算助力实现个性化健康预测

- AI与基因计算在个体健康管理、疾病预防、诊断和治疗等方面的应用将持续深化。例如，结合基因组信息和AI算法，构建高度精准的疾病预测模型，预测个体在未来发生特定健康问题的风险，从而实现早期干预
- 通过AI，可实现复杂数据处理、解决数据噪声问题，也可高效的对基因与功能之间的复杂调控关系进行识别、挖掘
- 随着单细胞组学、单细胞多组学和空间组学突破，基因测序技术正在从多细胞/组织层级向单细胞层级进化

基因测序：单细胞注释模型

scBERT:在单细胞测序领域首次引入BERT范式。在预训练阶段通过大规模训练数据学习到通用的基因和表达编码，提升模型泛化能力、在跨平台数据集的鲁棒性；基于少量数据微调，降低对精标注数据依赖，提升模型准确性。



腾讯AI Lab scBERT模型

基因成像：依托空间组学数据库

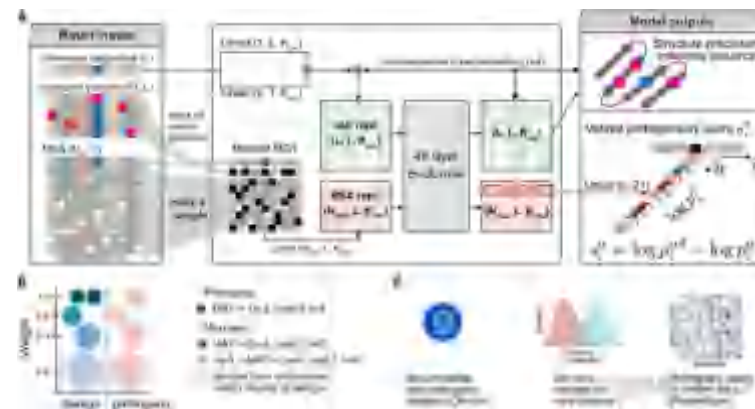
从空间角度出发，深入研究基因及其相互作用。空间组学技术借助各类成像或者测序技术，可以同时测量细胞的基因表达谱和空间位置，得到细胞微环境里其他细胞的基因表达和细胞间的相互作用。例如，基因交互式可视化、脑组织空间结构等。



腾讯AI Lab 空间多组学数据库 (SODB)

基因预测：分析基因突变的致病性

在基因突变如何影响人类健康方面，AlphaMissense模型对全部7100万种可能的基因错义突变进行了分类，将其中89%的变异分类为“可能致病”和“可能良性”。相比之下，人类专家目前的成绩是0.1%。



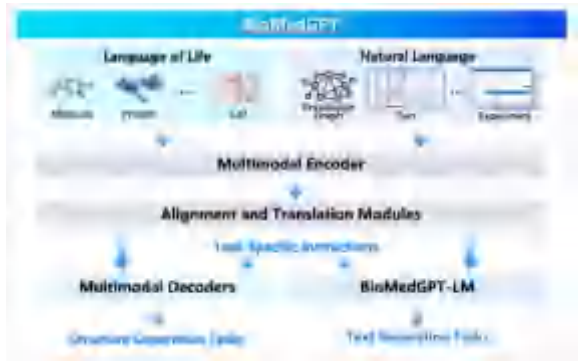
来源：谷歌DeepMind

趋势要点3: AI+基因工程重塑分子药物设计和研发规则

- AI技术在药物研发中，主要用于海量文献信息分析整合、发掘药物靶点、化合物高通量虚拟筛选、全新分子设计/优化、分子 ADMET 成药性预测、分子逆合成分析、耐药性预测等场景
- 通过量子化学计算、AI技术与生物医药技术的结合，开发实用化、产业化的合成生物学技术，可提升企业创新药研发实力

多模态数据统一表示

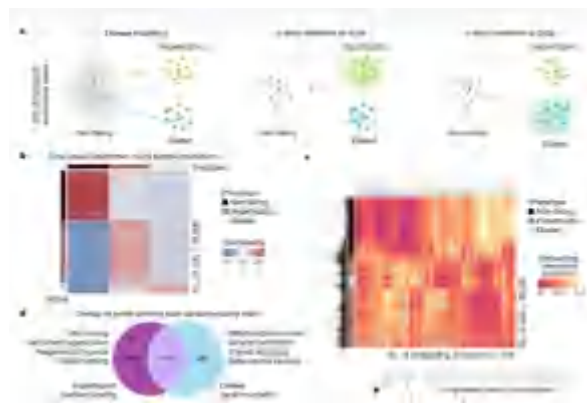
多模态生物医药领域的基础模型——BioMedGPT，旨在将生物世界中的分子、文本和知识进行统一表示学习，以提高各项下游任务的能力。BioMedGPT在数据层面整合了基因、分子、细胞、蛋白、文献、专利、知识库等多源异构的数据。



来源：清华AIR&水木分子

基因分析发掘药物靶点

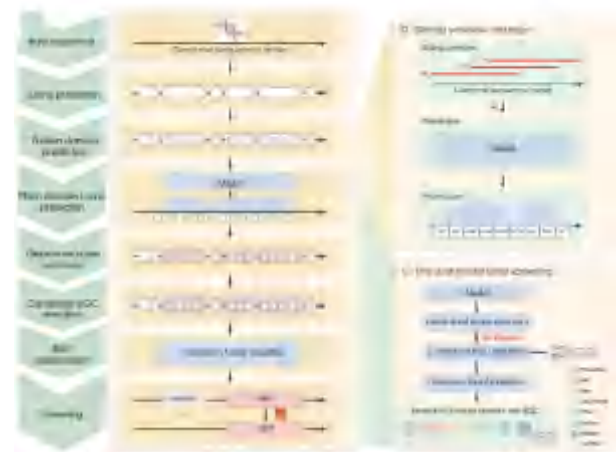
AI绘制基因互作网络，加快疾病治疗靶点发现。美国丹娜-法伯癌症研究所基于人体基因表达数据集，预训练了一个基于迁移学习的AI模型——Geneformer。对于罕见病，可能仅需少量的数据就能推测出这些罕见疾病的病理机制。



来源：《Nature》期刊, Transfer learning enables predictions in network biology

AI助力生物合成基因簇挖掘

采用深度学习方法对基因组数据进行分析，可发现、筛选并鉴定出潜在的、具有新颖结构的活性化合物相关的生物合成基因簇。



腾讯量子实验室模型流程图



数字交互引擎

激发“超级数字场景”

数字交互引擎激发“超级数字场景”

- 数字交互引擎是在文化创意场景下产生、伴随数字文化产业升级而不断实现技术迭代一类工具集，集成了物理模拟、3D建模、实时渲染等多种前沿技术，是文化科技融合的典型产物。数字交互引擎主要由图形模块、仿真模块、实时渲染等模块构成，它以软件代码包形式创造虚拟场景，动态呈现其外观变化，支持其与物理世界进行实时交互。
- 在发展前期，数字交互引擎主要应用于游戏场景，在行业场景下被称为“游戏引擎”；在服务游戏产业高效构建虚拟世界、与现实世界高质量交互的过程中，不断实现技术迭代、提升跨平台通用能力，逐步成为跨行业、跨场景应用的数字交互引擎。当前，数字交互引擎已经应用于文旅、汽车、工业等多元领域，成为构建实时虚拟世界、实现虚实交互的关键工具集。



01



打造构建“超数”
的高效工具集

数字交互引擎与AIGC互相驱动，打造构建超级数字场景的高效工具集

02



有望成为大众化
3D内容工具

大众应用方面，数字交互引擎或将走向UGC内容生产与交互工具

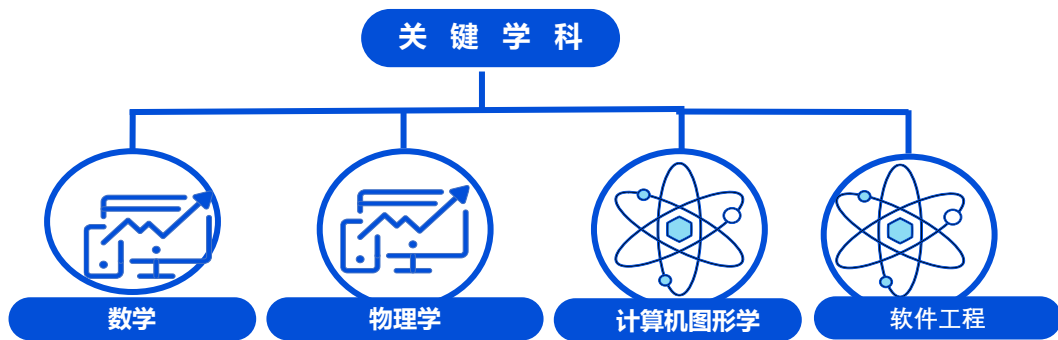
03



助推行业数字
孪生走向实时

行业应用方面，数字交互引擎推动各行业数字孪生走向实时性

学科背景、关键点、行业影响力



关键点和方向:

- 1、实现高性能渲染、呈现高度拟真视觉效果
- 2、工具链不断成熟、使用门槛降低
- 3、数字交互引擎与AI加速融合互驱
- 4、助力数字孪生、广泛链接实体行业

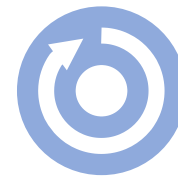
数字交互引擎是关系到我国科技竞争力的重要技术吗? *



92%认同

92%的受访专家认为，数字交互引擎是关系到我国科技竞争力的重要技术

数字交互引擎是否代表游戏产业原生的科技力量? *



受访专家普遍认同数字交互引擎代表着游戏产业原生的科技力量

您认为使用数字交互引擎赋能城市数字孪生能够带来什么样的提升? *

82%

提升数字孪生项目实施效率

82%的受访专家认为使用数字交互引擎能够提升数字孪生项目实施效率

73%

优化实时交互能力，画面可以实时响应

73%的受访专家认为使用数字交互引擎能够优化实时交互能力，画面可以实时响应

发展阶段

数字交互引擎基于开发数字文化创意的需求而产生，不断实现技术迭代、提升跨平台通用能力

诞生起源：为简化数字文化创意开发流程，数字交互引擎应运而生 >>>

20世纪90年代初

早期游戏开发者构建通用代码的开发工具，如知名游戏制作工具RPG Maker，已具有**数字交互引擎的雏形**

1993年

“3D游戏之父”卡马克在开发游戏《毁灭战士（DOOM）》的过程中开发了Id Tech引擎，**被广泛认为是应用于游戏行业的第一款数字交互引擎**

1996年

Id Software公司推出《雷神之锤（Quake）》，Quake引擎被认为是游戏行业的第一款**3D数字交互引擎**

技术迭代：技术不断迭代升级，强化实时光影渲染、物理仿真等关键能力，国产自研引擎兴起 >>>

1998年

引擎能力逐步完整

初代虚幻引擎（Unreal Engine）推出，将渲染、碰撞检测、AI、图形、网络和文件系统集成为一个完整的引擎

2006年左右

算力和表现力快速提升

Xbox360、PS3等主机平台发售，更强的硬件架构带来算力和游戏表现力的快速提升
第一批基于UE3的主机游戏和PC游戏发布，将游戏表现力提升到新高度
艺电（EA）、育碧（Ubisoft）等大型游戏厂商的自研引擎在各自的游戏类型上表现更好，不断提升游戏表现力

2010年左右

国产自研引擎起步发展

国产游戏厂商陆续研发自研引擎，逐步推出QuickSilverX、START云游戏引擎、NeoX、Messiah等引擎

2020年左右

支持实时光线追踪

2018年英伟达（Nvidia）发布支持实时光线追踪的RTX系列显卡
2022 UnrealEngine5正式发布，核心的实时全局光照技术以及Nanite支持海量多边形技术
全球头部厂商的自研引擎也纷纷实现了基于光线追踪的实时动态全局光照技术，标志着引擎进入实时动态全局光照时代

跨界赋能：
适配主流平台，构建跨行业通用能力>>>

2005年左右

逐步适配多种高性能硬件平台

2005年初代Unity引擎发布，数字交互引擎的适配平台从PC、主机拓展至Mac

2008年左右

移动端兼容性优化

随着移动互联网时代到来，Unity率先支持iOS、安卓等平台，数字交互引擎开启“移动端跃迁”

2018年左右

逐步开启跨行业应用

Unity支持的研发平台超过20余个，主流商业引擎涉足汽车、建筑设计、动画等领域，应用不断超越游戏场景

2022-2023年

加速拓展行业能力与应用场景

国产自研引擎的跨行业应用进入加速期，助力文物保护、智能驾驶、智慧工厂等多元领域

技术预测



2023 chronicle of events 大事记

5月

英伟达发布Avatar Cloud Engine for Games (ACE for Games)，可高效生成智能化的游戏角色

6月

Unity发布Muse和Sentis，用生成式AI加速实时3D应用创建，简化游戏项目集成 AI 的过程

6月

南航翔翼与腾讯游戏共同打造国产自研全动模拟机视景软件系统，国产引擎促进我国民航工业核心自研技术发展

6月

科技部将“三维数字交互引擎关键技术研发与应用生态构建”纳入国家重点研发计划

趋势要点1: 数字交互引擎与AIGC互相驱动, 打造构建超级数字场景的高效工具集

随着以数字交互引擎为代表的游戏科技进入游戏之外的更多领域, 游戏将迎来新的产业扩容, 成为不断创造新价值与新可能的“超级数字场景”。游戏作为前沿科技的“试验场”, 在深度跨界、产业扩容的过程中, 也将推动数字交互引擎进一步与多种前沿技术广泛连接, 其中数字交互引擎与AIGC的加速融合将成为重要趋势, 两者的互相驱动, 将打造构建超级数字场景的高效工具集、辐射多元社会场景。

01

数字交互引擎及其生成的游戏场景, 是AIGC爆发、迭代背后的重要推动力: 首先, 游戏是AI的重要训练场景, 促进AI决策更加智能; 同时, 数字交互引擎结合PCG (程序化内容生成) 技术, 已大量应用于3D内容制作, 可助力AIGC提升3D内容生成能力; 未来, 数字交互引擎及其创建的3D数字资产也将为AI大模型训练提供重要支持。

02

数字交互引擎融入更多AI能力, 正在加速数字文化产业的工业化进程: 当前AIGC已渗透进游戏制作的多个环节, 包括动画、对话和语音、美术、3D资产及场景等, 简化内容开发流程, 提升影视、游戏、广告等行业的智能化、工业化水平。

03

数字交互引擎与AIGC的结合, 将打造构建超级数字场景的高效工具集。随着数字交互引擎的应用不断拓展, 游戏正在成为一个不断创造新价值和新可能的“超级数字场景”。未来, 作为底层工具的数字交互引擎与AIGC相结合, 将加速打造更多的“超级数字场景”, 在感知与认知的双重维度上, 实现人类生产与生活模式的全面跃迁。

英伟达的Omniverse平台集成了数字交互引擎、AI等多种技术能力, 不仅能支持高效开发场景逼真、物理精确的游戏, 也能广泛助力工业制造、自动驾驶等领域的虚拟仿真与协作。

趋势要点2: 大众应用方面, 数字交互引擎或走向UGC形态内容工具

随着数字交互引擎与AIGC的融合加深、工具门槛进一步降低, 数字交互引擎有望以更低的使用门槛为大众提供3D内容的生产力。正如短视频变革了今天的社会信息传播方式, 未来, 数字交互引擎也有望作为大众化的3D内容生产与交互工具, 支持人们创造自身在虚拟空间中的形象及资产等, 在数实相生世界里实现生产与生活。

Roblox等游戏展现了“UGC+游戏+社交”的可能形态, 面向大众的游戏创作工具已初现雏形; 同时, UE、Unity等主流商业引擎集成大量AI工具, 不断借助AI技术提升高质量角色创作、大规模场景生成的效率, 降低开发者的使用门槛与创作难度。

Unity



Unity引擎中集成大量AI制作工具

Machine Learning Agents平台

更智能、更大规模地打造出各类风格的数字资产, 还可以让这些数字资产以更接近真实世界的方式与真人实现交互

通过深度强化学习和模仿学习相结合的方式, 让游戏开发者能够轻松地训练用于2D、3D和VR/AR游戏的智能Agent

数字交互引擎要走向大众化3D内容工具, 还需要从技术、生态、商业模式等维度积累资源、构建能力

🔧 技术

高质量的画面效果及实时算力支撑

- 提升关键模块技术水平: 如实时渲染、物理模拟等
- 云游戏提供强大算力支撑: 平台算力云化, 降低用户的接入成本

🌐 生态

丰富的开发者与多元的内容体系

- 降低准入门槛: 提供更完善素材和UGC工具, 实现低门槛、高效率的创作
- 构建用户社区: 构建社区和丰富的生活、娱乐、社交功能, 构建开发者生态

💰 商业

良好可持续的盈利模式

- 多元化商业模式: 探索多元商业模式, 推动创作者变现, 聚集内容开发者

趋势要点3：行业应用方面，数字交互引擎推动各行业数字孪生走向实时性

数字交互引擎凭借其在数据可视化、实时渲染、友好交互等方面的能力，已成为众多行业数字孪生重要的构建及运行平台，并支撑各行各业的数字孪生应用层面提升实时性。未来，随着数字交互引擎集成更多AI能力，将进一步提升数字孪生构建的效率与智能化水平，助力数字孪生技术以更智能化的形式服务于各行各业，促进各领域提高研发生产效率、降低产业创新风险。

感知：实时可视化

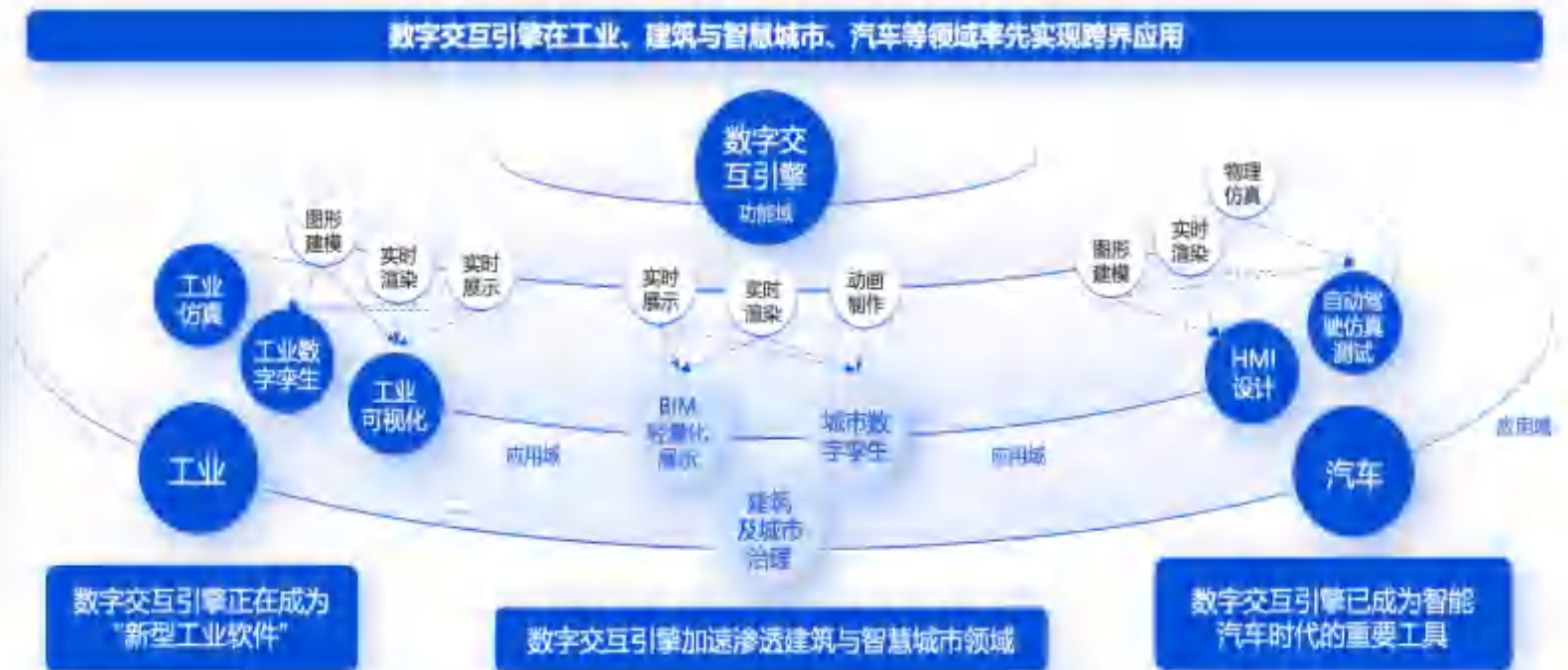
支持数据的实时感知接入并进行可视化展现

展示：实时渲染

能够实现3D模型的实时渲染和展示

交互：实时调用

支持用户对孪生体的快速查看、调用和修改





脑机接口 从医疗突破迈向 交互革命

脑机接口

- 脑机接口 (BCI, brain-computer interface)：在人或动物脑与外部设备间建立的直接连接通路，实现大脑与外部设备的直接交互。根据传感器或电极植入部位不同，主要分为非侵入式 (Non-invasive) 和侵入式 (Invasive) 两大类。
- 经过百年的近现代技术发展，脑机接口已形成了一系列基本的技术研究和应用范式，但由于对人脑认识的局限性，目前脑机接口仍处于发展早期。不过近年在数字技术尤其AI不断突破的加持下，脑机接口也呈现出加速发展的趋势，预计医疗、军事、教育、混合现实交互、类脑智能等领域将发挥重点带动作用，全球商用市场将以17%以上的年增长到2023年突破60亿美元。
- 从长远看脑机接口的意义更为重大，是人类应对人工智能威胁、弱化老龄化社会冲击、探索人类本质等重大问题，构建人机和谐社会的重要路径之一。



01



医疗突破

脑机接口将加速脑科学发展，助力医疗领域神经系统疾病监测与治疗突破

02



交互革命

脑机接口将与混合现实等技术深度融合，推动新一代人机交互模式变革

03

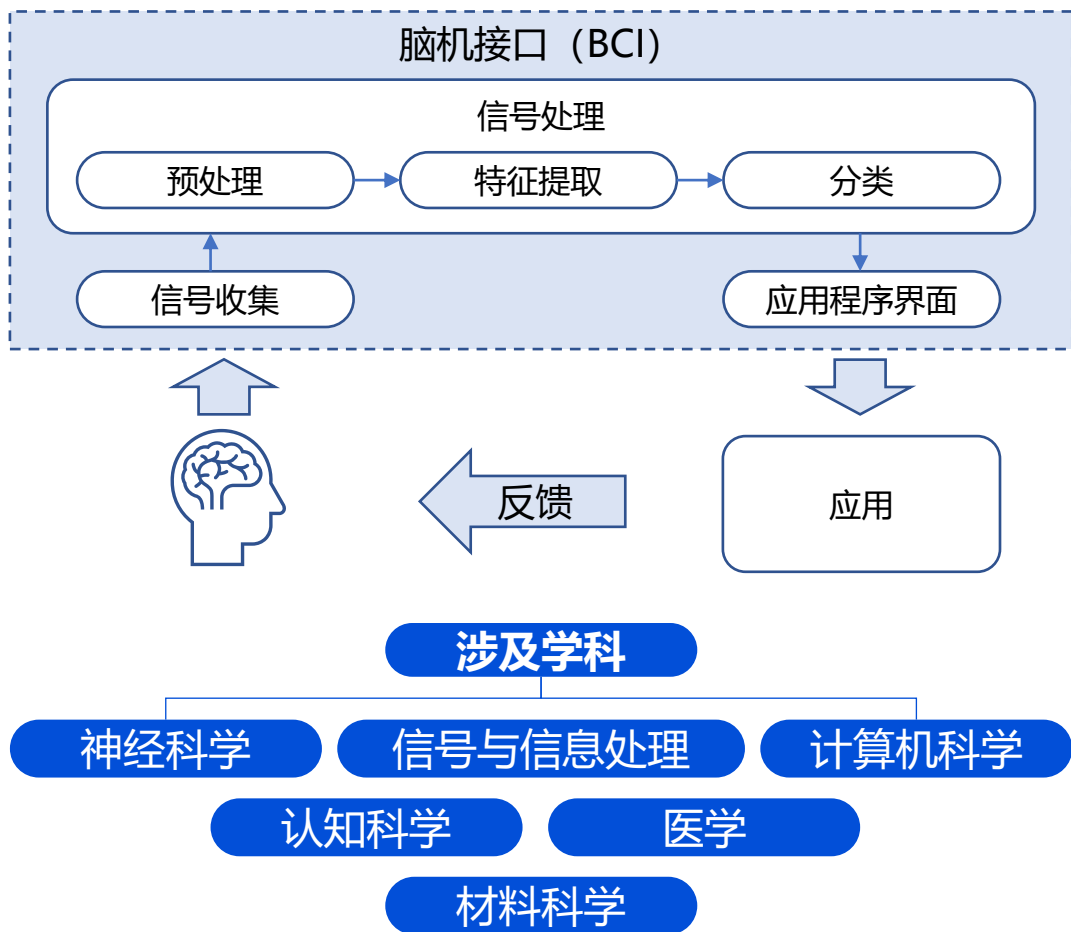


人机共生

脑机接口与人工智能相辅相成，成为促进人机和谐共生的重要路径之一

脑机接口主要涉及6大学科交叉，目前经历5个早期发展阶段

脑机接口的基本工作原理



起源：1780s

- 1780年：意大利Luigi Galvani发现静电能让死青蛙腿肌肉抽搐，称动物电流
- 1875年：英国Richard Conton发现非人类动物大脑中的电信号
- **1924年：德国Hans Berger开发了脑电图（EEG）首次记录人脑中的电信号**

科学设想：1930s

- 1938年：美国Herbert Jasper构思绘制了从脑电波中解码出语言的示意图
- **1973年：美国Jacques Vidal论文首次提出“脑机通信”的概念**
- 1984年：加拿大William Gibson在科幻《神经漫游者》描绘脑机接口设想

科研论证：1960s

- 1969年：德国Eberhard Fetz研究验证猴子能学会用意念控制仪表的指针
- 1988年：美国Farwell和Donchin创建P300拼写器，后成为重要的范式之一
- 1997年：美国Richard Norman开发了犹他电极阵列，后广泛应用

实用探索：1990s

- 1984年：美国FDA批准人工耳蜗临床中用于成人
- 1997年：美国FDA批准脑深部电刺激(DBS)用于治疗原发性震颤和帕金森病
- 2005年：美国Cyberkinetics公司BrainGate实现人体脑机接口控制机械臂

新技术融合：2000s

- 2001年：美国ASU发表具有生物活性的基于聚酰亚胺的柔性皮质内电极阵列
- 2014年：IBM发布TrueNorth类脑芯片，推动更高效、低能耗的类脑计算发展
- 2019年：美国Neuralink公司发布具备1024个通道的高速宽带脑机接口系统

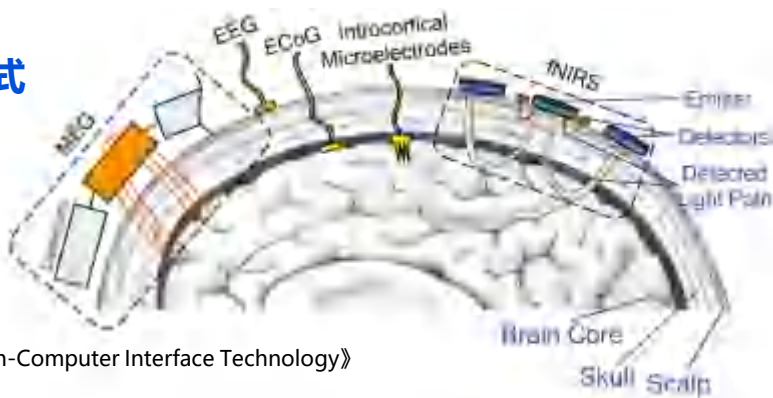
技术突破近期主要在信号采集和AI算法，远期关键在新材料和脑科学

关键技术蓝图



主要的传感器安装模式

- 脑磁图MEG
- 脑电图EEG
- 脑皮层电图ECoG
- 皮质内微电极 (IM)
- 功能近红外光谱fNIRS



来源：《State-of-the-Art on Brain-Computer Interface Technology》

主要性能趋势

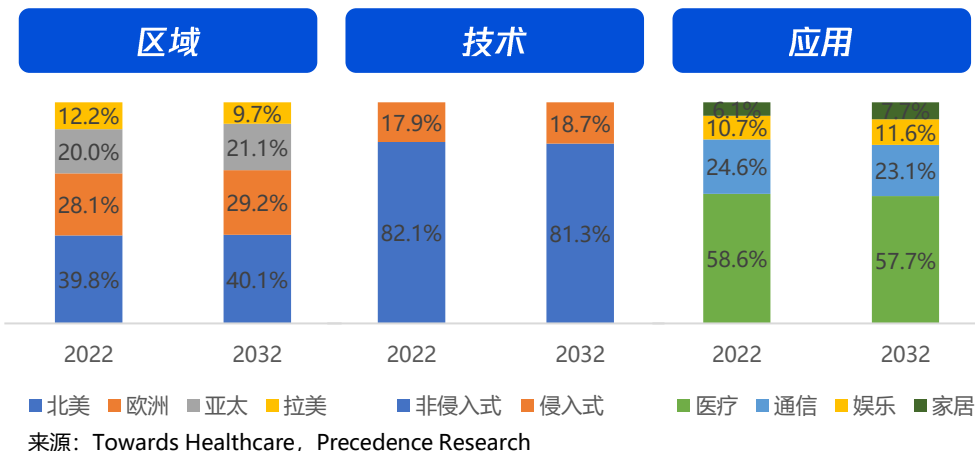
技术指标	侵入式BCI 技术目标	侵入式BCI 技术要点	非侵入式BCI 技术目标	非侵入式BCI 技术要点	共性 技术要点
采样率 (Sampling rate)	每秒 100,000Hz	高密度电极阵列, 新型生物相容性电极	1000Hz以上	高密度通道	新算法, 减噪技术
时延 (Time Delay)		无线通信, 新算法	10毫秒以内	无线通信, 新算法	新算法, 无线通信
空间分辨率 (Spatial Resolution)	毫米级或细胞级	高密度电极阵列, 微创植入技术, 新型生物相容性电极	-	高密度通道	-
控制自由度 (Degree of Freedom)	对数十个乃至上百个独立控制通道的解码	高密度电极阵列, 神经调控技术	-	高密度通道	新算法, 个性化优化
信息传输速率 (Information Transfer Rate)	每分钟百字量级	无线通信, 高密度电极阵列	每分钟60字以上	无线通信, 高密度通道	新算法, 减噪技术, 无线通信
准确率 (Accuracy)	95%以上	新算法, 神经调控技术, 个性化优化	90%以上	新算法, 减噪技术	新算法, 减噪技术, 个性化优化
使用寿命 (Functional Lifetime)	5-10年以上	新型生物相容性电极, 无线供电系统, 小型化设计	依赖于设备设计和材料性价比	新型生物相容性材料, 如导电聚合物水凝胶	新型生物相容性材料

脑机接口尚处市场早期，空间有限，但应用场景需求和价值潜力很大

市场空间预测 (亿美元)

研究机构	2023E	2030E	增长率
Precedence	23.5	64.2	16.7%
Gran View	20.4	63.2	17.5%
Zion	21	61	16.5%
Allied	22	54.6	13.9%
Data Horizon	23.6	75	17.8%
平均值	22	64	17%

市场细分预测



关键驱动因素

- **人工智能加速迭代与威胁:** AI大模型加速迭代, 在越来越多任务中展现出对人类脑力的超越性, 这种颠覆性威胁下, 脑机接口是未来人与AI共生的重要出路之一
- **神经相关疾病患病率攀升:** 1) 老龄化加深, 伴随癫痫、阿尔茨海默症和帕金森症等增多; 2) 神经发育障碍问题越来越受重视, 如自闭症、注意力缺陷、学习障碍等; 3) 残疾人对辅助技术需求强
- **现代军事人机协作需求加强:** 在无人机、自主武器等大量应用的情况下, 脑机接口可以提高军队人机协作能力, 同时改善训练和医疗效果

2023 chronicle of events 大事记

Neuralink获批首次人体临床试验

2023年5月, 马斯克的脑机接口公司Neuralink首次获得美国FDA批准, 开始人体临床试验开放招募, 可能率先针对颈脊髓损伤或肌萎缩侧索硬化症 (ALS) 导致四肢瘫痪的患者。Neuralink的核心并非底层技术突破, 而是强大的工程能力, 实现了高通量、微型化脑机接口装置的机器人精准微创植入, 有望加速脑机接口产品化、商业化的进程。

瑞士脑机科研团队让瘫痪男子再行走

2023年5月, 瑞士洛桑联邦理工大学等机构通过在人体内植入脑-脊接口 (BSI), 成功让因脊髓受损而下半身瘫痪的患者恢复自主行走能力。研究团队在《自然》杂志公布的研究成果显示, 在患者头部和脊髓周围植入传导装置, 能够实现在大脑和脊髓间重建“数字桥梁”。此项技术展现了脑机接口技术应用对残障人群的巨大价值。

苹果Vision Pro上市暗藏脑机接口

2023年6月, 苹果正式发布头显产品Vision Pro, 宣布进入空间计算时代, 其卓越的混合现实 (XR) 体验引发全球热议。除手动、声控等交互外, Vision Pro还实现了流畅的眼动交互, 通过人眼与大脑连接, 创造了一种简略的脑机接口 (BCI) 体验, 如监测瞳孔变化预测用户行为从而实时重建UI。这打开了未来XR和BCI融合创新的探索空间。

趋势要点1: 脑机接口加速脑科学研究, 助力医疗领域神经系统疾病监测及诊疗突破

- **医疗领域仍将是牵引脑机接口发展的重点领域。**脑机接口与医疗的结合应用展现出了广阔的前景, 不仅能为瘫痪和神经退行性疾病患者提供重要的沟通和控制解决方案, 还有潜力推动神经科学和临床神经医学的研究进展
- 未来脑机接口将重点向疾病治疗、人体增强等多样化的应用场景深化, 潜在应用功能包括恢复 (如神经和精神疾病治疗)、替换 (如控制神经假体)、增强 (如增强感官体验)、补充和改善等

神经发育障碍和退行

针对神经发育障碍、退行导致的各类疾病 (如癫痫、阿尔茨海默症、孤独症), 更精准的监测、分析、干预等提升诊疗效果



癫痫: 2021年浙大实现国首例款自研闭环神经刺激系统 Epilure™ 治疗癫痫临床植入



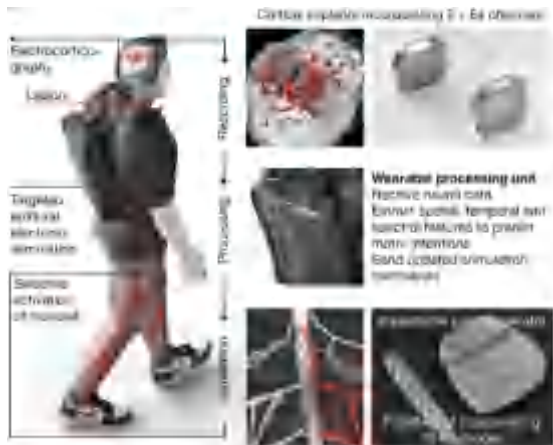
帕金森: 2022年宣武医院完成国内首创闭环神经刺激器治疗帕金森病植入



孤独症: BrainCo推出脑机接口社交沟通训练系统支持孤独症儿童治疗

肢体缺陷及运动障碍

一是重复刺激患者大脑, 实现运动相关神经元部分连接或功能修复; 二是获取患者运动意图, 实现对外部设备控制完成行动



瑞士EPFL脑-脊接口让瘫痪者恢复自主行走能力

感觉缺陷和损伤

解码患者自身感觉信息, 实现感觉补偿甚至修复, 已在听觉、视觉、触觉等感觉缺陷和损伤诊疗中发挥积极作用

BCM通过动态电流电极, 成功在受试者脑海中呈现指定图像 (Cell)



Stanford、UCSF的读脑设备让失语者重新说话 (Nature)



心理健康与精神疾病

通过脑电信号分析和调控人的情绪和认知状态, 辅助诊疗抑郁症、焦虑症、精神分裂症等, 并通过神经反馈训练促进患者康复



抑郁症: Inner Cosmos推出目前最小、侵入性最少的植入物“数字药丸”并开展临床实验



Muse推出基于EEG的冥想和睡眠头带帮助用户日常精神和睡眠状态调整优化

趋势要点2: 脑机接口与混合现实深化融合, 将推动新一代人机交互模式变革

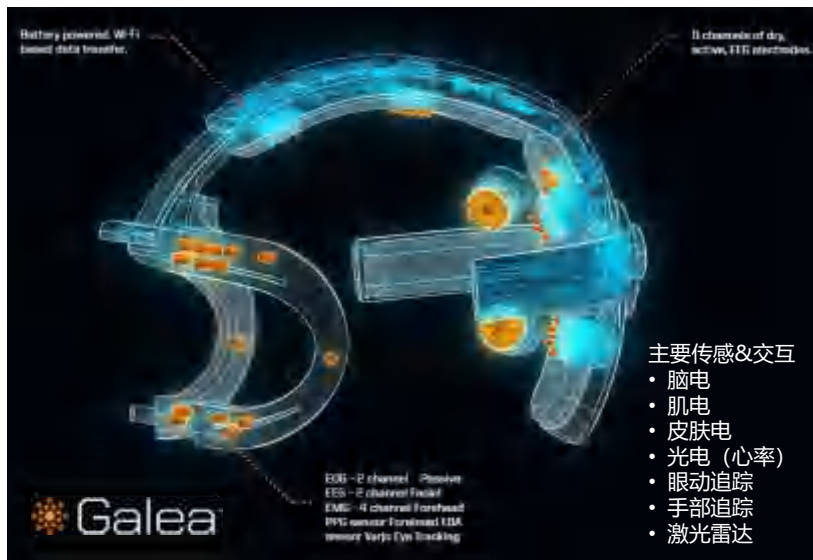
- 人机交互模式的演进需求将驱动脑机接口与其他技术的融合发展。通过脑机接口, 我们可以实现直接从大脑获取信息和指令, 这将与虚拟现实 (VR)、增强现实 (AR) 和元宇宙等新兴技术有效融合, 提供更为直接和自然的交互方式
- 未来脑机接口与混合现实的结合, 特别在商业及生活服务方面的适用领域会十分广泛, 如娱乐、社交、身份识别、疲劳干预、个性化学习等。这类脑机技术多采用非植入式, 更加安全便捷、更易被大众接受, 真正形成消费级应用才能加速普及

脑机接口将促进人机自然交互全面升级



【例】Meta2022年发布肌电图+AI腕带交互装置

【例】OpenBCI与Varjo合作开发的Galea BCI系统, 整合包含多种传感器的非侵入式脑机接口系统和混合现实系统, 创造出了能进行多重体验交互的新型软硬件平台



娱乐: 更沉浸



游戏主播用脑机头戴玩3A游戏

生活: 更便捷



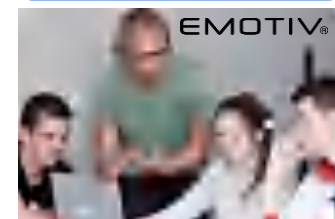
基于脑机接口的智能家居系统研究

工作: 更高效



脑机头戴对司机疲劳监测与警报

学习: 更专注



脑电图耳机+神经科学APP教学

趋势要点3: 脑机接口与人工智能相辅相成, 成为促进人机和谐共生的重要路径之一

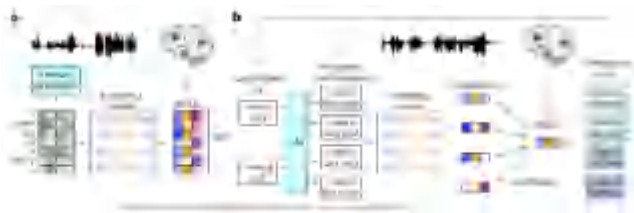
- 应对人工智能的潜在威胁提高了脑机接口发展的必要性。脑机接口有望架设人脑与人工智能的直接、高速连接, 一方面能增强人脑信息处理和交互能力, 避免被人工智能直接替代; 另一方面能促进人工智能向人脑机制深入学习演进, 走向更可信可控、低耗高效的发展方向
- 类脑计算是脑机接口与人工智能结合的重点发展方向之一, 主要包括神经科学研究 (特别是对大脑信息处理基本原理)、类脑计算器件 (硬件)、类脑学习与处理算法 (软件), 长期可能促成新的类脑计算结构体系、赛博格 (人机融合体)、脑联网等的发展突破



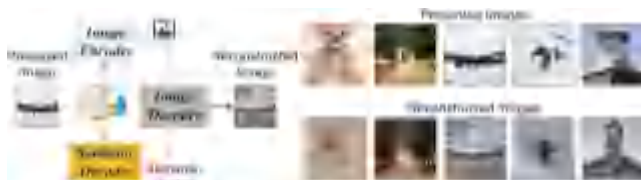
“人类必须成为机器人才能避免AI的统治……这主要与带宽有关, 即你的大脑和你的数字版本之间的连接速度” ——马斯克

大脑信息解码与重建

AI技术的加速突破 (如LLM、DM等), 能支持更高效处理脑机接口采集的大量脑信号, 提升对大脑文字、影像等信息解码与重建效率, 促进大脑信息处理机制等的基础研究, 提升未来“脑-脑通信”的可能性



来源: 《Semantic reconstruction of continuous language from non-invasive brain recordings》



来源: 《Improving visual image reconstruction from human brain activity using latent diffusion models via multiple decoded inputs》

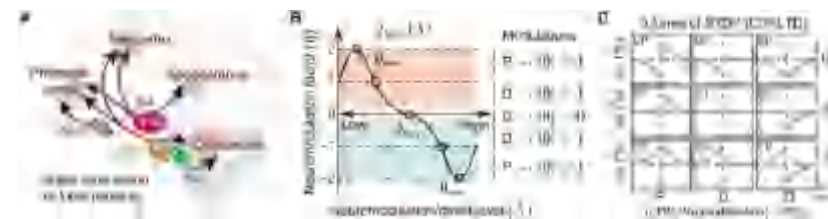
德克萨斯大学奥斯汀分校研究团队实现, 基于GPT-1和大脑fMRI数据训练的模型, 信息重建准确率:

- 感知语音 (听录音): 72-82%
- 想象语言 (想故事): 41-74%
- 无声电影 (看电影): 21-45%

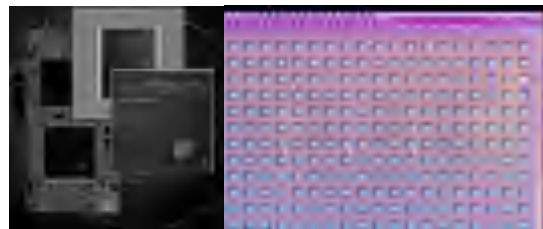
大阪大学研究团队实现, 基于扩散模型, 重建通过fMRI获得的人脑活动图像, 有效降低计算成本同时保留高生成性能

类脑计算软硬件

脑机接口促进脑数据采集和脑科学研究, 能反过来支持类脑计算的框架、算法、芯片等技术创新, 跳出冯-诺伊曼计算结构局限, 推动更低功耗、高效率、可信可控的AI发展, 降低AI风险、促进向人机和谐共生方向发展



中国科学院自动化研究所团队, 创新一种基于神经调制依赖可塑性的新型类脑学习方法NACA, 通过全局多巴胺的弥散支持与输入信号同步, 甚至先于输入信号的正向信息传播, 再加上选择性的对STDP的调整, 使得NACA表现出明显的快速收敛和缓解灾难性遗忘优势



IBM在其类脑芯片TrueNorth的基础上酝酿用于神经推理的新芯片原型NorthPole, 融合了类脑计算和硅优化计算之间、计算和内存之间、硬件和软件之间的界限, 相比传统CPU能以极其节能的方式快速执行计算任务, 大幅提升能效25倍

**沉浸式媒体
催生3D在场**



沉浸式、3D化的新视界

- 未来视频向沉浸式体验、高效生成演进、从面向人眼视觉到面向机器视觉、从消费级到产业端。
- 体验、内容生成、标准和网络协议是多媒体的关键技术。伴随着设备从平面到XR的升级、AI技术的不断进步，以及产业互联网对视频的需求提升，多媒体体验将更加沉浸。同时，AIGC技术将高效生成文字、图片和视频。深度学习视频编解码也将进一步压缩，提升效率。此外，在应用层和网络层的优化，将进一步推动多媒体赋能产业互联网。



01



沉浸式

沉浸式技术日趋成熟，模型成为视频的信息载体

02



多媒体 + AI

AIGC已经商用，大模型在垂直领域不断强化多媒体能力

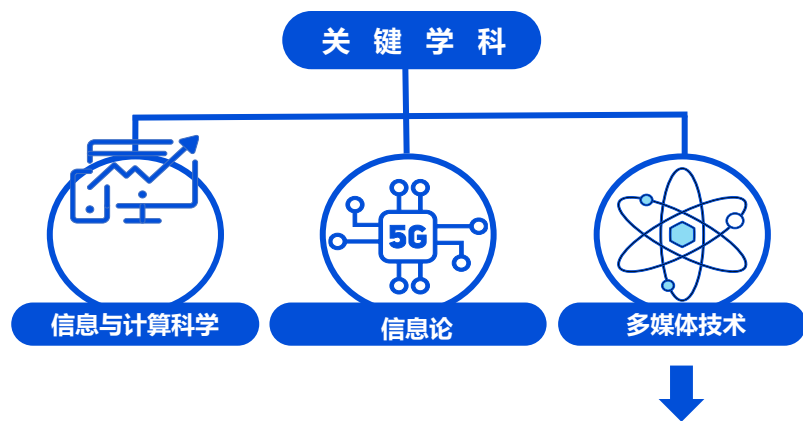
03



协议

底层协议根据场景需求优化完善，推动多媒体赋能产业互联网

学科背景



多媒体压缩、多媒体处理、多媒体理解、机器视觉

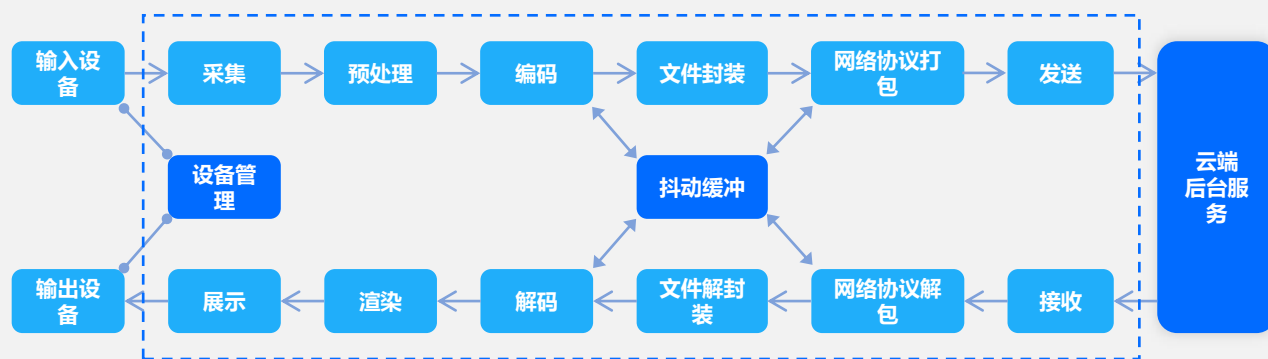
2021.6-2023.6网络视频（含短视频）用户规模及使用率



来源: CNIC 中国互联网络发展状况统计调查

2023.5

音频数据处理过程



技术关键点和指标

生成编辑: 硬件采集

视频压缩编码: 编解码标准

视频传输: 网络传输协议、时延、Qos、Qoe

视频的呈现: 设备、交互、清晰度、HRD

发展阶段

时间	阶段	分辨率	清晰度	标准	设备	聚焦方向	交互	领域
1990年代	电视阶段	352×240 (NTSC制式) 352×288 (PAL制式)		MPEG-2 电视传输渠道	电视 广播		单向交互	
21世纪初	PC流媒体阶段	720p、1080p (标清和高清)	(全高清, Full HD)	H.264编码	PC	实现流媒体	互动快进、双向弱交互	面向C端, 链接人到人的
2010-2020		1080p、4K	(UHD, 超高清)	H.265 (HEVC) AV1	PC手机	更高压缩比和分辨率	基于移动设备、移动多媒体的双向交互	
2020年-今	移动互联网阶段	4K、8K	HDR (10比特)	H266 VVC	手机、XR	更高压缩比和更多样的视频形态	低时延、实时互动、3DoF、6DoF强交互	面向产业, 链接人和物、设备的



故宫 x 腾讯——沉浸式数字体验展

未来发展方向

- 1、如何提高qoe和qos: 进一步降低时延、提高压缩比。
- 2、更高效的内容生成和呈现: AIGC、HDR技术
- 3、更多样内容、互动内容、新媒体: 沉浸式交互和体验
- 4、更深入产业, 助力产业互联网

技术预测



2023 chronicle of events 大事记

空间计算

9月，苹果发布会介绍，用户可利用iPhone 15 Pro 超广角和主摄一起来拍摄带三维感的视频，并在Apple Vision Pro 中重现，这也是首款可以实现空间视频拍摄的智能手机。

AIGC技术日趋成熟

Stable Diffusion、Midjourney、Runway、Pika等专业模型，以及GPT4V、Gemini等多模态模型先后发布，文生图、文生视频技术日益成熟。

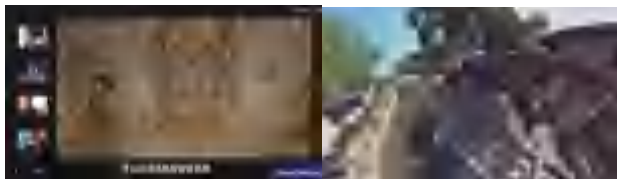
裸眼3D

1月，腾讯申请“全息裸眼3D设备”专利获批。9月，腾讯会议展示裸眼3D效果，同时具备双目视差和移动视差。观看者除了能体验到出屏的立体感，同时也可以通过左右移动，看到不同视角的立体内容，更真实，沉浸感更强。

趋势要点1: 沉浸式技术日趋成熟, 模型成为视频的信息载体

视频向更加沉浸式, 双向交互方向演进, 3DoF视频已经商用落地, 在微信视频号、快手等平台可以体验。裸眼3D产品已经产品化, 谷歌、腾讯均发布了裸眼3D的产品。6DoF视频和体积视频目前还在演进和探索中, 目前点云的技术路线相对比较成熟, 有望实现突破。未来伴随硬件的进一步成熟和6DoF技术演进, 更多的UGC 3D沉浸式内容将会出现, 同时模型文件将成为多媒体的重要载体

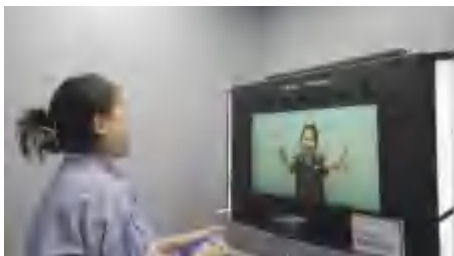
3DoF产品商用落地



敦煌研究院·VR360
远程会诊系统

一部手机游云
南·VR360游白沙细月

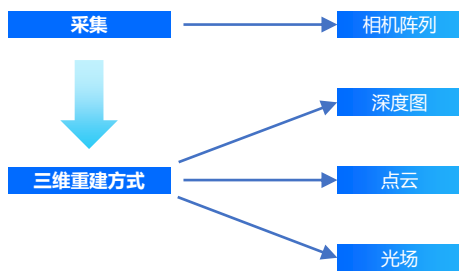
裸眼3D已经产品化



腾讯会议裸眼3D

6DoF视频探索中

三维重建技术是6DoF视频的核心, 目前采集端用相机阵列进行采集, 三维重建有深度图、点云和光场等不同方式, 其中深度图的方式较为成熟。



体积视频还在演化

以极高的精度和细节捕获空间中深度、色彩等信息, 生成3D模型素材, 实现更为沉浸的体验。全球体积视频市场从 2022 年的 17.9 亿美元增长到 2023 年的 23.1 亿美元, 复合年增长率 (CAGR) 为 28.7%。

硬件推动

- 高通xr芯片
- 苹果头显
- 裸眼3D屏幕日益成熟

标准推动

- H265成熟: mv-hevc、3d-hevc,
- H266推动沉浸视频演进: 针对全景视频、3D视频提出方案
- 6dof 点云标准: PCC

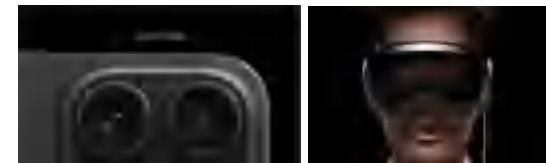
大规模分发链路支持

云上直播和点播转码产品化

- 多媒体容器格式支持
- 分发协议优化
- 4k、8k、转码

UGC 3D内容

随着Vision Pro、新款Iphone的推出, 用户未来可以直接采集和显示3D图片和视频。



模型文件成为重要载体

模型文件成为载体, 视频和应用的界限将模糊化。Nerf是未来重要方向。



趋势要点2: 大模型在垂直领域不断强化多媒体内容能力, 进一步提升视频压缩能力

- AIGC技术在多媒体的文字、图片生成等方面已经商用, 未来聚焦生成更加稳定的视频和3D内容。同时在垂直领域如数字人、超分、老片修复等方向不断强化多媒体的能力
- 随着更多样化的块划分方法和编码模式的不断涌现, 以及更复杂的预测和变换技术的引入, 传统视频编码算法的复杂度不断提高, 完全依赖传统编码框架技术来提高视频编码压缩效率变得越来越困难。深度学习技术在图像分类、目标检测等计算机视觉任务上已取得了巨大的成功, 近几年, 深度学习技术为图像/视频编码框架定义了全新的结构范式, 实现了图像和视频编码器性能的显著提升, 这为图像/视频编码领域带来了新的研究机遇。

AI生成视频、3D是主要探索方向

1. 技术进步在急剧加速; 2. 3D生成很有可能在2年内可用; 3. 创新多出现在多领域交叉领域

文本生成文本	文本生成图片	生成动画	生成音频	生成视频	生成3D模型
人类平均水平	人类专业水平	可用	可用	探索	不可用
ChatGPT LMDA	Midjourney Stable Diffusion	WhiskerVideo DeepMotion	Audius	Image Pictory	Magic3D Dressify

AI 在垂直领域不断强化多媒体能力

数字人: Talking Face Generatio, 利用Gan大模型, 数字人动作更自然



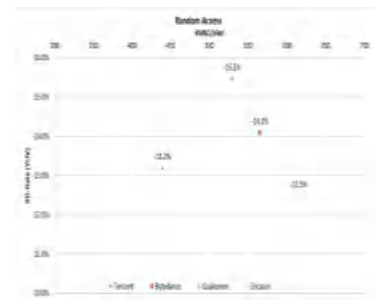
在超分、老片修复等方向: 利用大模型利用解决视频稳定性, 简化人工手动优化



AI增强传统编码能力

在传统编码框架中引入深度神经网络来进一步提升压缩性能。深度学习策略实现对传统编码标准中的块划分、预测模式等大量待搜索对象快速判决, 从而缓解编码端搜索压力、降低计算复杂度。

JVET 开展的神经网络视频编码 (NNVC) 探索实验和增强压缩视频编码 (ECM) 探索实验, 相比VVC/H.266标准参考软件VTM, 在RA和 AI配置下, Y、Cb、Cr三个通道BD-rate分别节省: {-21.17%, -32.29%, -33.05%} 和{-11.06%, -22.62%, -24.13%}, 具备演进成为下一代视频编码标准的技术潜力。



端到端神经网络视频编码技术: 打破传统编码框架, 完全使用深度学习实现编解码流程。基于神经网络视频编码, 利用海量数据集进行神经网络训练, 学习去除视频压缩失真任务中的先验知识。例如jpeg-ai的主要动机是基于深度神经网络的工具在图像编码、计算机视觉和图像处理任务中的出色性能。JPEG AI旨在开发一种图像编码标准, 以满足云存储、视觉监控、自动驾驶车辆和设备、图像采集存储和管理、视觉数据实时监控和媒体分发等广泛应用的需求。

趋势要点3: 底层协议根据场景需求优化完善, 推动多媒体赋能产业互联网

多媒体技术在产业互联网中的应用, 将根据不同场景进行优化, 从而更好地满足产业互联网的需求。在产业互联网中, 网络协议需要根据场景需求进行优化以提高传输效率。传统的直播传输的播放模型存在诸多问题, 如缓存固定、传输可靠性过高、无法区分视频帧优先级等。针对这些问题, 在消费互联网中的WebRTC通信模型基础上, 进行优化, 在网络时延、QoE和可靠性等进行优化, 对场景需求进行最优匹配。

在C端低延时的演进路线

基于TCP 协议

2005-RTMP协议

2009-HTTP Live Streaming(HLS)协议

2019-LL-HLS(Low-Latency HLS)

采用chunk编码, 将延时降低到3秒左右

基于UDP协议

1998-RTSP 协议 over UDP

2017-Secure Reliable Transport(SRT)协议

2011-webrtc 协议

时延和c
端QoS
时延和稳定

优化

信令改造, 利用miniSDP和0-RTT的结合, 大幅减少信令耗时、提升信令交互成功, 进而降低首帧耗时和提升开播成功率。
音视频改造, 让WebRTC支持AAC, H.265, 附加前向纠错, 抗50%以上丢包。还引入了B帧, 增强了画质, 同时大幅减少了码率。
传输改造, 采样柔性分级丢帧的传输策略来渐进式降低码率, 以适应弱网情况。支持P2P分发网络, 能够将看同一视频流的用户群就近地组织成网络, 相互分享传输。

快直播的延时可降低到800ms以内, 并同时兼顾延时、卡顿和首帧耗时, 综合QoS远超传统直播

优化

信令改造: 提升信令链路对网络异常抵抗能力, 减少网络异常恢复时间, 提升视频应用稳定性
音视频改造, 优化相机采集、视频渲染和视频编解码耗时, 从音视频处理层面减少端到端画面延迟。
传输改造, 以减少视频传输延迟为目标, 适当平衡抗丢包和抗网络波动能力, 减少视频延迟; 并引入多网传输策略, 减少单一网络依赖, 提升传输稳定性;

画面延时可降低到100ms以内, 兼顾抗弱网能力

2B2C: 快直播

电商领域, 随时发起的低延时秒杀活动赋予了电商直播更强的互动娱乐性, 拉动了用户观看时长以及交易额增长

企业体育直播, 多机位镜头结合超低延时直播技术应用, 能够无缝切换现场多视角镜头, 现场沉浸式体验更佳。

To B

工业远程作业, 基于<100ms的极低视频延迟, 实现远程流畅工程作业, 降低作业危险和现场往返时长, 提升作业效率。

星地直连通信推动 泛在网络覆盖

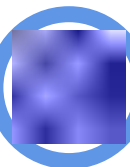


星地直连通信推动泛在网络覆盖

- 实现Anytime Anywhere Anyone Anything 一直以来就是通信技术发展的目标，尽管目前地面蜂窝网络已经实现了70%人口的覆盖，但是仅有20%的地表面积被网络覆盖，大部分地区还处于失联状态。基于卫星的无线通信技术可以以较低的成本实现广域覆盖，并不受地质灾害、安装环境的影响，实现宽窄带通信服务
- 星地直连需要通过商业航天和通信技术结合来实现。受益于商业航天近些年的快速发展，火箭发射成本得到了大幅度的降低，卫星的研制成本和性能也有了大幅度的提升，可以满足星地直连高通量卫星研制和大规模星座建设和发射的需求。基于地面5G衍生的非地表网络协议实现了产业链融合和赋能，并有望实现星地通信一致化的体验。
- 未来1-3年，随着卫星研制能力的提升和星地融合通信协议的进步，基于平板天线的宽带卫星通信系统和棉线微型设备的窄带星地融合通信系统有望进入大规模普及阶段，基于宽带的手机直连卫星通信技术也将得到快速的发展。



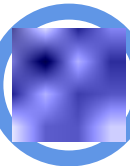
01



卫星研制“工业化”

通过使用工业级COTS器件、批量生产和可回收技术，卫星的研制和发射成本大幅度下降，大规模星座部署成本可控，服务能力大幅度提升

02



通信标准“一体化”

3GPP基于5G开展非地表网络NTN技术研究，R17版本完成面向高轨弯管卫星的窄带物联网标准研究，实现星地通信技术一体，实现产业链融合

03

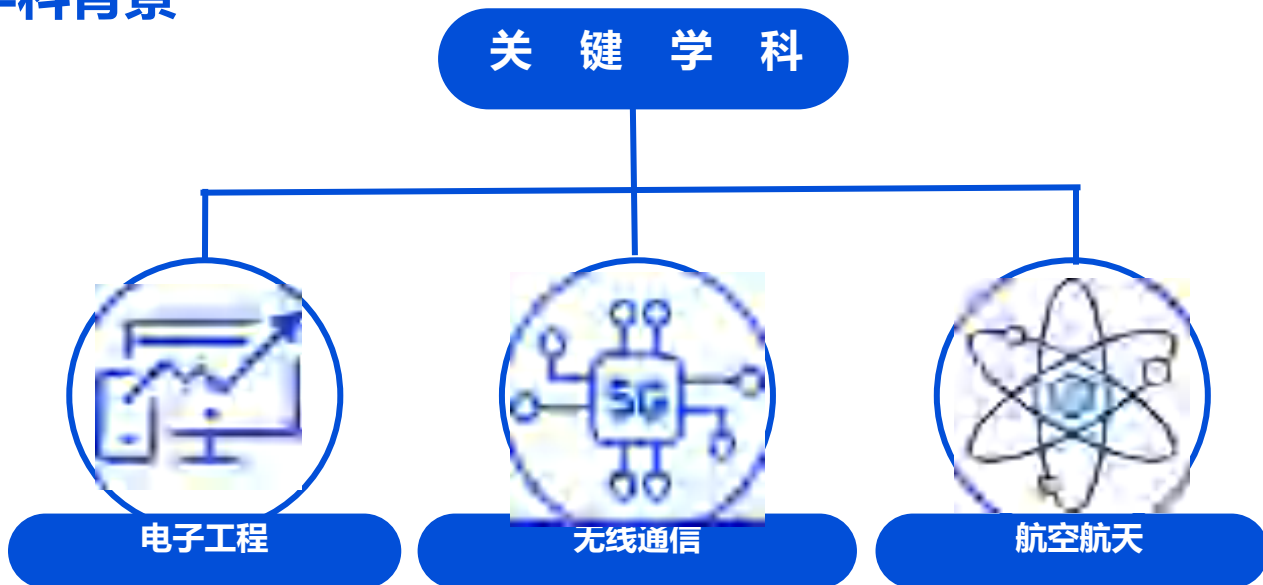


卫星通信“大众化”

卫星通信系统正在拓展到更泛在、更商业、更大众的应用领域，必然催生大量的软硬件需求。包括即时通信和应急通信产品的新需求，以及泛在物联网的爆发等。卫星网络安全也进入公众视野。

» 发展阶段、学科背景、热点方向

学科背景

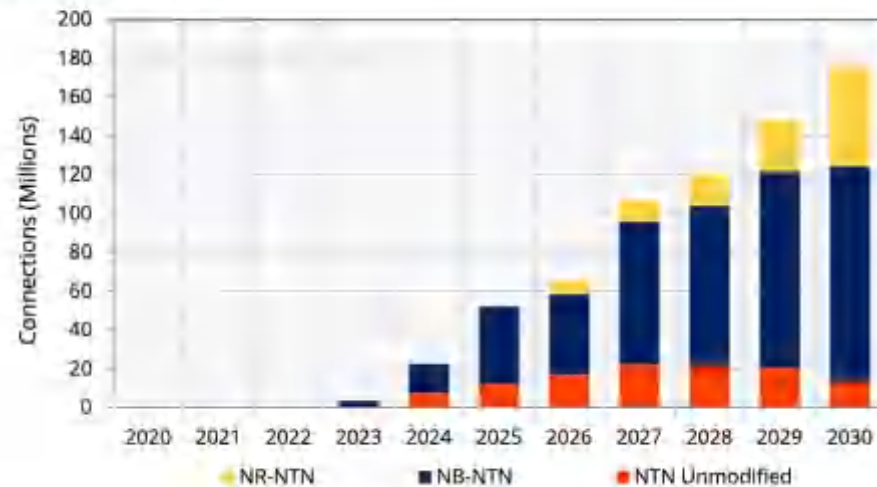


技术关键点和方向

业务需求、星座规模、服务能力和带宽之间相互影响、相互促进。卫星成本降低和通信技术演进推进手机直连卫星应用的加速发展；手机直连卫星的发展促进应用需求的多样化，增加了对业务多样性和带宽的需求，又进一步对星座建设和通信技术演进提出要求；卫星数量越多，卫星的制造成本降低又进一步推动产业的发展。

行业预测

Chart 1: NTN Mobile Connections by Service Segment
World Markets: 2020 to 2030
(Source: ABI Research)



根据ABI研究预测，随着NTN技术的发展，未来手机直连卫星会进入到快速增长阶段。到2030年全球会有1.7亿台NTN移动终端设备，总共带来163亿美元的收入，从2022年到2030年间的复合年均增长率可以达到76%。

技术预测



时间点	协议	芯片	设备	天地融合
1990年代	以铱星系统为代表的私有协议	专用芯片	专用通信设备	不融合
21世纪初	借鉴地面蜂窝通信协议研发专用卫星通信协议 如全球星的通信协议是基于CDMA系统的	专用芯片	专用设备	不融合
2022年	借鉴地面蜂窝通信协议研发专用卫星通信协议 如全球星的通信协议是基于CDMA系统的	专用芯片	在通用设备中植入独立的专用卫星通信芯片	业务融合
2022年-至今	基于3GPP NTN协议实现和地面蜂窝网络的兼容	通用芯片	在通用设备中植入独立通用的卫星通信芯片	协议兼容
未来	星地融合的卫星通信协议	通用双模芯片	标准商用手机	协议融合

2023 chronicle of events 大事记

星链2.0 mini版本入轨

2023年2月27日，星链公司发射星链2.0 mini版本卫星，重量800kg，支持S频段，计划支持T-mobile手机上星应急通信，开启低轨手机直连星座组网；2024年初，发射了6颗支持存量手机直连的星链卫星。

华为Mate60 Pro发布

2023年8月，华为发布Mate 60 Pro手机，支持通过天通卫星的语音通信功能，标志着商用手机进入卫星通信时代。2024年初，荣耀Magic6系列旗舰新品，号称实现了体积最小、信号最稳、且最省电的手机卫星通信体验。

基于NTN的窄带通信测试成功

2023年，多家运营商和私营公司成功开展了基于3GPP IoT NTN的星地直连通信，验证了基于地面蜂窝的通信协议可以支持使用现有卫星进行通信

趋势要点1: 卫星研制“工业化”，卫星研制成本和周期快速下降

- 近年来，商业航天实现突破性进展，通过使用工业级COTS器件、批量生产和可回收技术，卫星的研制和发射成本大幅度下降，大规模星座部署成本可控，服务能力大幅度提升。

工业级器件大规模使用

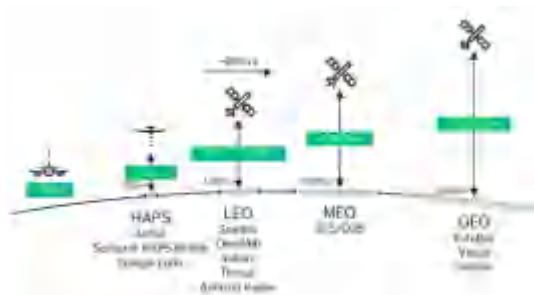
基于工业级器件降低卫星研制成本：大量的商业航天公司使用低成本高性能的工业级器件替代传统的宇航级器件研制低轨卫星，并通过多模冗余设计解决空间辐噪导致的失效问题，大幅度降低了卫星的研制成本和时间。

火箭可回收技术

SpaceX通过高推重比发动机、多级并发、动态调度等技术实现火箭发动机的可回收和重复使用，降低火箭的发射成本，并进一步结合星箭一体化技术增加单个火箭运载卫星的数量，从而大幅度降低卫星的部署成本。



2017年3月，猎鹰9号火箭B1021成为世界上首个两次被成功回收的轨道级助推器



ABI: 各种卫星的轨道高度示意

卫星批量化生产

批量化生产卫星降低卫星的测试成本和研制周期：通过卫星工厂来批量化生产标准卫星，可以降低卫星的研发和测试成本，并减少卫星的研制周期，为短时间内大规模星座部署提供保障。

趋势要点2：通信标准“一体化”，为产业链融合提供有效支撑

- 通信标准“一体化”：3GPP基于5G开展非地表网络NTN技术研究，R17版本完成面向高轨弯管卫星的窄带物联网标准研究，R18推进基于宽带的5G NR NTN通信协议的研发工作，实现星地通信技术一体，实现产业链融合。

3GPP确定技术演进标准

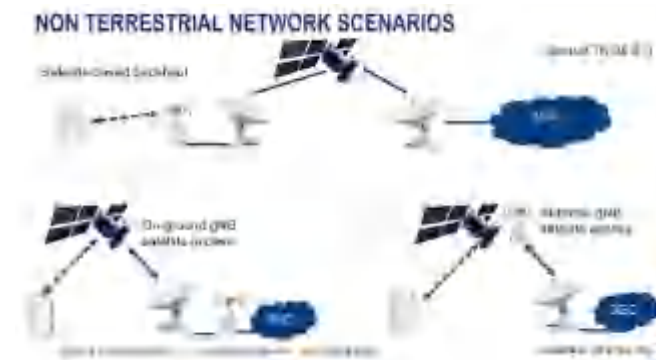
3GPP R17标准正式发布，通信国际标准化组织3GPP R17版本发布，支持面向卫星物联网的IoT NTN标准，标准完成了基于高轨弯管卫星的NB-IoT和eMTC系统的支持，后续会逐渐演进到低轨和NR NTN；

ITU划分手机直连频段

国际电联ITU正式确定NTN网络的可用频段：ITU划定N255和N256频段可用于部署基于NTN的卫星窄带物联网系统，L频段的34MHz FDD带宽被定义为N255，S频段的30MHz定义为N256频段；

3GPP IoT NTN在轨测试成功

基于R17 IoT NTN的在轨测试成功：多个运营商和卫星物联网公司开展了基于高轨卫星和R17 IoT NTN标准的在轨测试，并获得成功，标志着基于高轨卫星IoT NTN的通信协议已经初步实现了技术闭环



<https://www.connectivity.technology/2020/05/r-technical-explainer-on-3gpp-5g-non.html>

趋势要点3: 卫星通信“大众化”

- 卫星通信系统正在从专业和特殊行业应用拓展到更泛在、更商业、更大众的应用领域，必然催生大量的软硬件需求。其中一个重点趋势是对于宽窄带结合的即时通信和应急通信产品的新需求，由此带来对语音和视频压缩技术的新需求。另一个重点趋势是泛在物联网的爆发。
- 近年来，卫星网络安全攻击事件激增。例如，卫星提供商Viasat的KA-SAT网络遭遇“多方面”攻击，导致乌克兰与欧洲部分区域KA-SAT卫星宽带用户服务全部中断，被忽视的卫星通信系统网络安全进入公众视野。

商用手机具备卫星应急通信功能

2022年末，华为和苹果相继发布手机支持应急通信短报文功能；
2023年10月，华为发布Mate 60 Pro手机支持天通卫星的语音通话功能。



低轨手机直连开展技术验证

Starlink发射2.0mini版本卫星，相比1.0版卫星，体积提升了4倍，重量增加了3倍，通信能力提升了4倍，未来支持S频段的T-Mobile的手机直连；AST Mobile开始部署基于低轨的超大阵列天线以支持手机直连功能，并开展了多次在轨测试。



IoT NTN进入部署阶段，网络安全引起关注

2023年MWC展上，鹏鹄物宇等单位联合发布了基于IoT NTN的卫星物联网终端设备和模组，设备体积小，功耗低，有效降低了卫星物联网的使用门槛，为后续大规模普及提供铺垫。
腾讯安全天马实验室聚焦前沿卫星互联网络及新基建下各类产业应用场景的安全研究，率先原创发现多个影响范围广泛的通用型高危漏洞，为卫星电话、地图测绘、广播电视、气象监测等应用领域提供安全保障。



*eVTOL*加速空中
出行奔赴新时代



» eVTOL加速空中出行奔赴新时代

- 随着城市化进程加速，交通拥堵和环境污染问题日益突出，人类社会对绿色、高效的交通方式的需求日益迫切。在电动飞机技术、新能源技术和数字技术的发展和共同引领下，电动垂直起降飞行器（electric vertical take-off and landing, eVTOL）逐渐从概念走向现实，将彻底改变以地面交通为主导的出行模式，加速空中出行奔赴新时代。
- eVTOL以电池为能源，采用分布式电推进系统，不仅大幅降低飞行噪音，提高操作安全性，还能实现垂直起降，无需跑道。这使其成为一种具备应用潜力的绿色、智能空中交通工具。近期，我国工信部等四部门共同发布《绿色航空制造业发展纲要（2023-2035年）》明确提出了到2025年实现eVTOL试点运行。**在政策的大力支持下，空中出行有望在2025年后逐步成为公众的交通选择。**



01



电动化

在全尺寸原型机试飞测试中，多旋翼、复合翼、矢量推进构型的技术逐步得到提升，其中矢量推进构型成为业界主流方案。eVTOL的动力装置电推进系统也向更高效、更轻量级的方向发展。

02



长续航

宁德时代研发能量密度高达500Wh/kg的锂电池技术，锂电的里程碑式突破有望推动eVTOL续航里程大幅提升，为eVTOL执行城际/区域空中交通奠定技术基础。

03



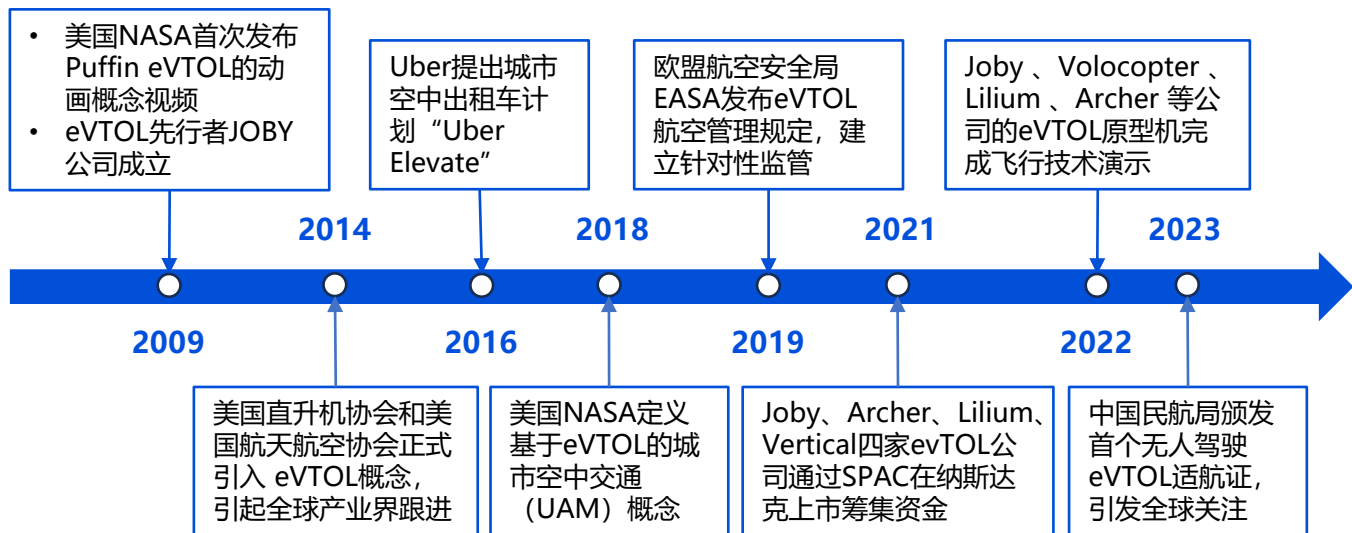
智能化

长期来看无人驾驶eVTOL是降低运营成本的重要方向，短期内有人驾驶是主流选择。在相对复杂的低空环境中实现无人驾驶，数字化智能化的空中交通管理系统成为必不可少的低空新基建。

eVTOL的发展历程、学科背景、应用场景

eVTOL从概念诞生到如今成为新兴行业，得到了航空航天企业、运输行业、政府、军方和学术界等机构的大力推广和支持。

eVTOL的发展历程



eVTOL的相关学科

- 空气动力学
- 飞行力学
- 结构力学
- 仿生学
- 材料学
- 计算机
- 控制理论

eVTOL的应用场景

旅游观光、机场摆渡、城市空中出租车、城际通航、区域客运、商务飞行等	城市末端配送、紧急医疗服务、城际空中快递、乡村物流
载人客运	载物货运
空中巡逻、应急处突、交通执法、缉毒缉私、空中监控、警力机动、军需运输、救援搜索、空中通讯等	城市管理、消防灭火、应急救援、环保监测、疫情防控、自然保护、社区治理、农林植保
国防安防	公共服务

技术预测



技术成熟度

- 3年以上
- 2-3年以上
- 0-1年以上

- 影响力中等
- 影响力大
- 影响力极大

备注：受限于篇幅，以上仅列出部分eVTOL的关键技术。

行业趋势

载人客运是eVTOL的发展方向和核心场景，未来市场规模预计超万亿美元。

eVTOL将率先在载货物流、城市服务、消防救灾等场景启动运营，待各项技术成熟政策完善和市场接受，才会迈入大范围载人客运时代。许多公司已经提前构想和研究空中交通的“共享出行”运营模式，摩根士丹利预测城市客运和区域客运等潜在市场规模在2040年将超过1万亿美元。

eVTOL续航里程持续提升，带动空中交通转向服务城际和区域出行。

受限于电池能量密度不足，电动航空最初定位是城市空中交通场景。随着电池技术发展和续航里程提升，城际和区域出行的商业模式将为运营商带来更高的单位经济效益，为客户有效节省时间；同时所需的飞行频率和机队密度较低，对城市带来的影响更小。

eVTOL主机厂的适航取证工作有序推进，eVTOL产业生态体系加速构建。

随着各主要民航监管部门对多个型号和构型的eVTOL的审查工作不断推进，头部OEM有望在2025-2026年集中获得适航证。在此背景下，主机厂、零部件供应商、基础设施建设商、服务提供商以及领域外玩家（如互联网、物流、汽车企业）都在提前布局eVTOL市场。

2023 chronicle of events 大事记

多家厂商完成全尺寸eVTOL的飞行测试

2023年，Joby, Archer, Lilium, Vertical等头部eVTOL公司已完成首架全尺寸原型机的飞行测试，实现了垂直起降、高速巡航、悬停飞行过渡到前飞等各项主要飞行任务，展示了eVTOL技术的可行性和未来的商业化潜力。

电动飞机高能量密度电池技术取得突破

2023年4月，宁德时代发布了其研发成功单体能量密度500Wh/kg的凝聚态电池，用于民用电动载人飞机项目的合作开发，被航空界认为是极大提升电动飞机续航里程的关键性技术。

全球民航部门加速完善eVTOL监管体系

2023年，欧盟EASA在原有适航认证基础以外还从监管角度定义了飞行器运行要求、电池电量储备要求、飞行员执照、空域整合以及合规要求等。美国FAA和中国CAAC也在积极制定eVTOL的监管框架并发布相关文件。

全球首个无人驾驶eVTOL适航证颁发

2023年10月，中国民航局为亿航智能EH216-S型号载人无人驾驶航空器颁发型号合格证，也是全球首例面向无人驾驶eVTOL的型号合格证。亿航智能随后也举行了EH216-S型号的商业载人首飞演示。

趋势要点1: 矢量推进型eVTOL成为市场认可的主流技术路线

- 在eVTOL商业化的不同技术路线中，相较于多旋翼构型，矢量推进型和复合翼构型eVTOL在航程、巡航速度和载重比方面优势明显，具有较好的有效载荷、最大起飞重量和运营经济性，因此城际运输等空中交通商业场景中应用推广显示出更大的潜力。
- eVTOL普遍采用分布式电推进系统（DEP, Distributed Electric Propulsion）作为动力装置，有利于提升飞行器的气动效率、载运能力、环保性和鲁棒性等，是一种航空领域的颠覆性技术。

eVTOL整机研发项目的主要技术路线形成

eVTOL基于推进动力方式形成多种主要技术路线

技术路线	示意图	工作原理	相关企业
多旋翼		通过多个（通常多于4个）固定螺旋桨实现起降和巡航动作	Volocopter 亿航智能 空中客车
复合翼 (升力+巡航)		升力与巡航用的螺旋桨是独立的，分别实现垂直起降和巡航	Volocopter Vertical 亿航智能 霍尼韦尔
倾转旋翼		通过倾转不同螺旋桨或机翼方向实现飞行姿态控制与起降	Joby Archer 时的科技
涵道矢量推进		通过改变涵道推力方向，实现不同场景下的垂直起降于巡航	Lilium 德塔航空

备注：后两类飞行器可通过改变螺旋桨/机翼/涵道方向实现飞行器的起降和巡航，又称为矢量推进型。

矢量推进型和复合翼型的航程和效率优势较为明显

构型的优势取决于定位的商业场景及其对应航程距离

多旋翼	多旋翼构型实现技术路线简单，但是气动效率较低，有效载荷和航程相对有限。目前电池能量密度能支撑约30分钟飞行，用于短距离运输场景，无法支持中长途载客运输。
复合翼	“升力+巡航”复合翼构型可以有效地提升巡航效率、航程和安全性，但是垂直起降和平飞携带的无效结构重量过多，限制飞行器的有效载荷和飞行速度。
倾转旋翼	倾转旋翼构型重量轻、推力大、效率高，在速度和航程上具有优势，但开发难度、成本和风险高。
涵道矢量推进	倾转涵道风扇构型的突出优势是在巡航阶段所消耗的能量更小，但是悬停的效率较低、消耗能量高，更适合长距离飞行场景。同时较其他机型更容易降低噪音和振动。

全球超900个eVTOL研发项目中，矢量推进构型是受青睐



来源：美国垂直飞行协会，截止至2023年11月

电推进系统向更高效、轻量和安全的方向发展

- eVTOL采用分布式电推进系统，利用电力驱动多个推进器作为动力装置，能有效提升飞行器气动效率、载运能力、环保性和鲁棒性等，是一种航空领域的颠覆性技术。
- 电推进系统下一步发展的方向是更高效、更轻量、更安全。**具体包括但不限于：高压平台的应用和推广普及；通过提高电压，降低线束直径；进一步提高功率元器件IGBT的效率；轴向磁通电机取代径向磁通电机，降低电机重量；提高电源模块和驱动模块的效率。

eVTOL使用电推进系统具有多方面优势

效率	<ul style="list-style-type: none"> DEP推进效率达95%-97%，比先进的涡扇发动机高出20%。 电机功率重量比可达其他发动机的6倍。
安全	<ul style="list-style-type: none"> 功率不因高海拔或炎热天气而衰减。 多个电机的冗余设计提升安全性。
可持续	<ul style="list-style-type: none"> 纯电飞机运行全过程零排放。 大幅度降低社区噪声水平15dB以上。

趋势要点2: 高性能电池技术突破提升eVTOL续航能力

- 从经济性的角度出发, 绝大多数主机厂都在正式生产的eVTOL机型中采用更为成熟稳定、能量密度和功率密度相对较高的锂电池, 以实现航程距离较长的空中交通, 另一方面也满足适航认证过程中对于飞行器安全性的要求, 加速适航取证后的量产制造。
- 未来随着氢气储存技术和氢燃料电池的稳定性得到显著提升, 以氢能源为动力的低空飞行器也有望得到应用推广。

锂电池成为eVTOL商用的首选能源

头部厂商选择现阶段具备比较优势的锂电池

- 航空业绿色转型的能源选择有锂电池、氢燃料电池、电子燃料(e-Fuels)等。在头部厂商的目标航程范围内, 锂电池的技术更加稳定成熟, 且能源效率高、单位成本低(见下表), 且目前可通过与新能源汽车电池生产商合作进行量产。

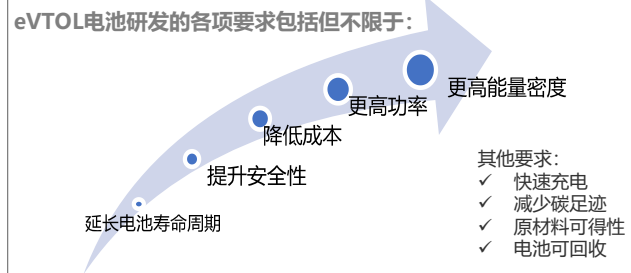
指标	电池 Batteries	绿氢 e-Hydrogen	电子燃料 e-Fuels
一级能量效率	73%	22%	13%
电力价格	~0.36美元/kWH		
成本/千瓦时轴功率	~0.5美元/kWH	~1.7美元/kWH	~2.8美元/kWH
飞行航程	1100 (2040年) -2000km (2050年)	最长达~ 3400km	最长达~ 1600km

来源: Lilium Battery Webinar, 2023年11月

电池是eVTOL续航提升的主要驱动力

电池需满足安全、高能量、可量产等条件

- 在满足航空级别安全要求的条件下, 提高电池的能量密度(Wh/kg)和功率密度(W/kg), 且以业界可接受的成本实现量产, 是电池研发的重要方向。



当前电池性能尚不足支持eVTOL长距离飞行

- 使用能量密度300Wh/kg左右的电芯, 估算不同构型的eVTOL的单次充电的最大航程: 多旋翼构型为30-40km, 复合翼构型为250-280km, 倾转旋翼构型为250km, 涵道矢量推进构型为175km。

高能量密度电池正在跨越商业化门槛

航空级别的高能量密度电池商业化渐近

- 高能量密度电池技术被航空界认为是极大提升电动飞机续航里程的关键性技术。
- 2023年4月, 宁德时代发布了单体能量密度500Wh/kg的凝聚态电池, 用于电动飞机研发。对于部分头部企业而言, 电池能量密度达到500Wh/kg, 有潜力支持现有机型飞行400-500km。

超高能量密度电池技术的研发厂商举例

公司(国家)	电池能量密度	商用时间
宁德时代(中国)	500Wh/kg的凝聚态电池	2023年发布, 未公布商用时间
Amprion Technologies(美国)	400Wh/kg的硅阳极锂离子电池	2024年商用
Lilium With IonBox(德国)	330Wh/kg的硅阳极锂离子电池	已经用在测试机型中

来源: 宁德时代, Amprion, Lilium

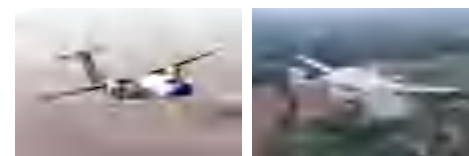
氢能源在航空的应用潜力巨大

氢能应用潜力大但受限于总成本高

- 长期来看, 氢燃料电池的能量密度最高可达锂电池的上百倍, 然而, 氢燃料电池受成本重量和储运及潜在安全风险等制约, 短期内应用空间有限。

氢能飞机商业化进度慢于电动飞机

- 德国H2FLY今年发布了一款双体式、四座位、单发的液态氢飞机HY4, 并完成了有人驾驶的飞行演示。
- 美国Universal Hydrogen今年试飞氢能源飞机, 获得FAA颁发的特殊适航证。
- 空中客车公司计划在2035年之前推出氢能飞机。



Universal Hydrogen 氢能飞机

H2FLY氢能飞机

趋势要点3: 智能化技术加速无人驾驶空中交通愿景实现

- 得益于智能驾驶技术发展与政策支持等有利因素，eVTOL飞行器正逐步从传统的有人驾驶模式过渡到更高效的无人驾驶模式，呈现出“软件定义飞行器”的趋势。
- 数字技术逐步在低空空域管理中应用和推广，从支持空域数字化到助力监管部门在复杂和密集的低空空域中实现有效监管，有力促进和保障低空经济发展。

全自动无人驾驶eVTOL是未来低空出行的演进方向

数字技术进步前提下，“软件定义飞行器”成为可能

- 未来eVTOL的竞争力关键既包括飞行器的设计和性能，也包括决策AI为核心的智能驾驶软件技术，同时还需要配备高效的数字化空中交通管理系统。

无人驾驶的核心是满足民航管理部门对安全性的要求

- 无人驾驶不仅是技术实现的问题，还需要满足民航管理部门对航空安全性的要求。为了使eVTOL符合适航认证对于安全性的要求以及更容易被乘客接受，主机厂会对早期机型配备飞行员或无人驾驶安全员，但长期看实现自主飞行是大幅度降低运行成本的方向之一。



Lilium Jet 是配备飞行员的eVTOL



亿航EH216-S是无人驾驶eVTOL

低空基础设施建设助力高效安全空中交通管理

低空空域交通管理面临高度复杂性和不确定性

- eVTOL未来在低空运行的流量大于现有通航直升机，但数量级会少于城市内的网约车。eVTOL运行调度的环境是三维，还要考虑到天气等因素对低空空域的影响，环境的复杂性、不确定性较高。

低空基础设施可支撑高效安全的空中交通管理



低成本、高精度、高可靠性的通讯、导航、监视系统，可以实时获取航空器的信息，降低航空器间隔，提高空域流量和安全性。



基于传统气象雷达以及激光雷达等新一代传感器收集的气象数据，结合AI大模型等技术，为低空飞行带来更精准、网格化的气象服务。



基于云计算、边缘计算、深度学习和人工智能的管控和调度系统，为eVTOL飞行提供决策支持，比如实时航路规划、起降场选择。

数字化平台应用加速低空空域开放和利用的进程

我国正在加强和规范空域资源管理、促进空域有效利用

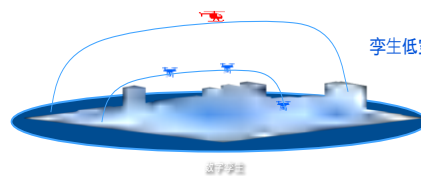
- 空域是国家重要战略资源。2023年11月，国家空管委发布《中华人民共和国空域管理条例（征求意见稿）》，提出实现空域资源科学精细配置、分级分类管理、动态灵活使用。

数字化平台应用加速空域数字化

- 集成物联感知、空间构建、时空计算、逼真渲染、仿真推演、数字孪生等技术的数字化平台，为空域管理部门进行民用场景的空域设计、航道规划、模拟测试提供数字化工具。

空域建模

- 基于地理信息数据和BIM空间数据融合，及场景自定义编辑，实现大规模、高精度三维建模

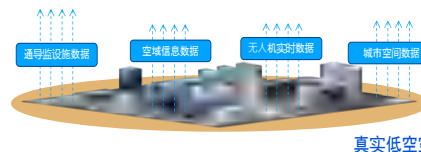


孪生低空空域 时空计算

- 基于实时四维时空计算，实现无人机飞行计划和飞行航迹的时空冲突校验，保障飞行活动安全有序

数据融合

- 无人机、空域、CNS设施实时数据融合，低空空域运行态势实时感知



仿真模拟

- 依托低空空域数字孪生体，构建面向行业的算法仿真训练能力，加速产业协同发展

腾讯数字孪生技术在低空空域的应用



多能流实时协同

重塑虚拟电厂

多能流实时协同重塑虚拟电厂

- 随着新能源技术与信息技术的发展与成熟，虚拟电厂将成为双碳背景下关键环节能源结构转型的重要解决方案。当前，未来电网的源、荷、储三端正在发生重大变化；源端：波动的清洁能源将大规模、高比例地接入电网；负荷端，大量用户或将迎来参与发电、储能与电网响应的“新身份”；储能端，大量电化学储能技术的发展，以及氢储能技术的研发，正在大幅减少能量的存储与运输成本。
- 以往虚拟电厂更多应用在较为狭窄的应用场景，接下来，我们应该将其上升到城市甚至是城市间的尺度。城市运行中有三类调节性负荷，分别是工业、算力和交通。随着工业革新、大模型算力、新能源汽车等领域的发展，其电力需求仍在不断攀升，是危也是机，让数字化聚合起来的“电厂”，可以通过承担多网耦合、协同的工作，将上述几类用电大户，转变为可调节的资源，从而有效应对由于能源结构转型所带来的电网压力，保障新能源下新型电力系统的顺利转型。

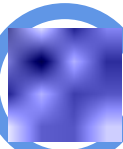
01



工业-电力

互联网协同炼钢厂，以钢铁工艺流程中可调节负荷为初步虚拟电厂试点，并逐步囊括多种可调节资源

02



算力-电力

多个IDCs通过光纤可以进行数据负荷传输，并与其之间传输数据的光纤网共同构成算力网络实现协同

03

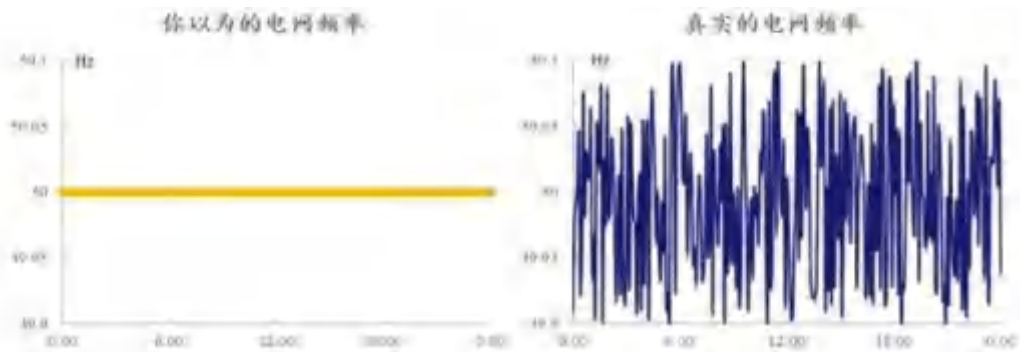


交通-电力

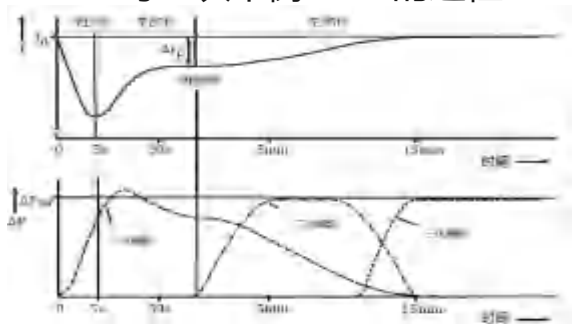
电动汽车(EV)将电力网络与交通网络这两个复杂的人造系统紧密耦合，交通网将成为强大的调控资源

虚拟电厂：以数字化聚合起来的“电厂，由于大电网新能源占比不断提升，虚拟电厂需要承担多网耦合、协同的工作

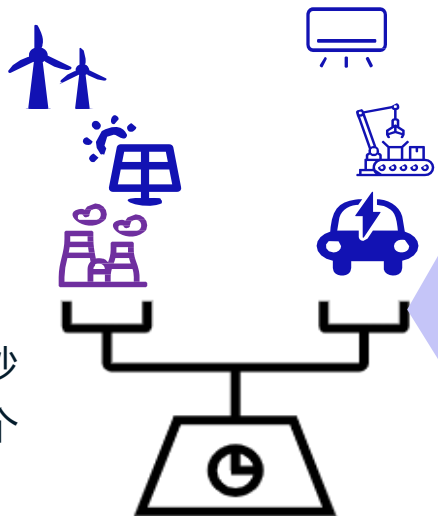
挑战



每一次平衡delta的过程



- 整个南方电网的响应时间只有几秒
- 电力的供需平衡在总量和时间两个维度上都有挑战



机遇

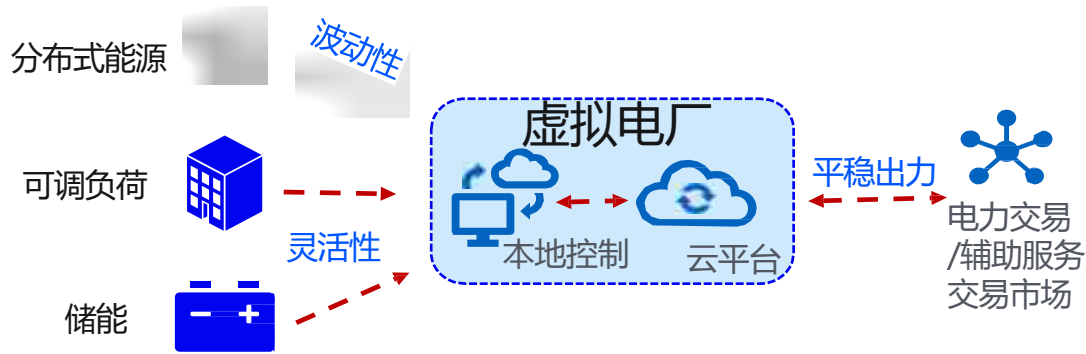
引导自备电厂、传统高载能工业负荷、工商业可中断负荷、电动汽车充电网络、虚拟电厂等参与系统调节，省级电网基本具备5%以上的尖峰负荷响应能力

——2021/10 国务院《2030年前碳达峰行动方案》

- 收益：浙江2021~ 小时、分钟、秒级响应，支付4元/KWh收益
- 社会价值：10MW灵活性可等效每年减排9000吨，节省燃气电厂备用费8000万元

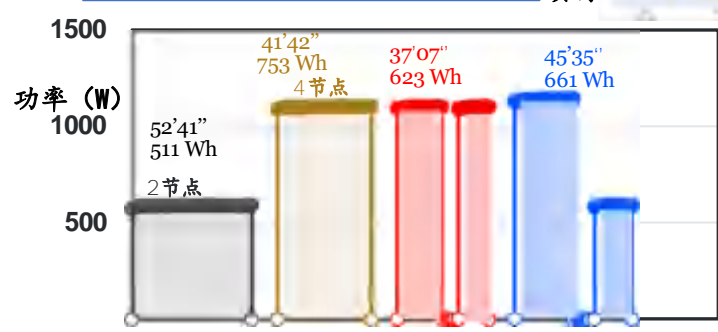
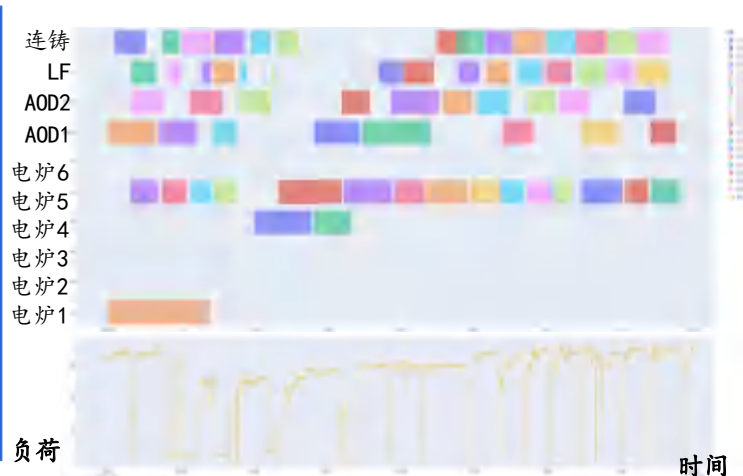


虚拟电厂：以数字化聚合起来的“电厂”



新型虚拟电厂须从系统工程的角度，寻找城市运行当中核心三类调节性负荷并进行整合，建立长期虚拟电厂管理机制

工业-电力协同
生产节奏调整，短流程
钢铁“华容道”



算力-电力协同
腾讯IDC
离线业务/断点续算/可
续渲染

交通-电力协同
智能充电、V2G等

新型电力系统的储能需求
与C端互动，平移伸缩电动车充
电速度，聚合为“大电池”

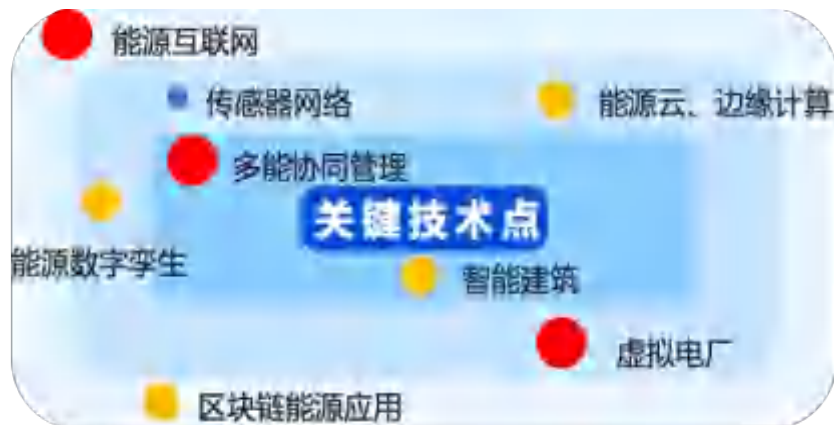


智能、有序、高效的数字化控制手段成为当下能源变革的先决条件

可再生能源消纳以及电力电子化是实现碳中和的关键，稳定电力供应是实打实的刚性需求。随着近年来新能源行业发展，电力电子资源的不断增加，国家需要有更高效的数字化控制手段

-相关学科与技术方向:

-2023年大事记电力协同应对手段:



蔚来参与全国最大规模V2G需求响应

8月23日，车网互动验证中心（简称e-Park）的V2G充放电系统需求响应试验在无锡正式启动。该系统是目前国内规模最大的V2G充放电系统，为新型电力系统建设提供了有力支撑。

山东电力交易现“负电价”运营

“五一”期间，山东电力现货市场实时交易电价波动剧烈，区间为1047.51元/兆瓦时至-80元/兆瓦时（约为1.05元/度至-0.085元/度），多次出现负电价。

中广核新能源虚拟电厂参与响应

2023年中，中广核新能源深圳虚拟电厂成为首批满足并网接入要求的标准化虚拟电厂，并成功参与首轮精准响应，响应容量和响应精度位于虚拟电厂运营商前列。

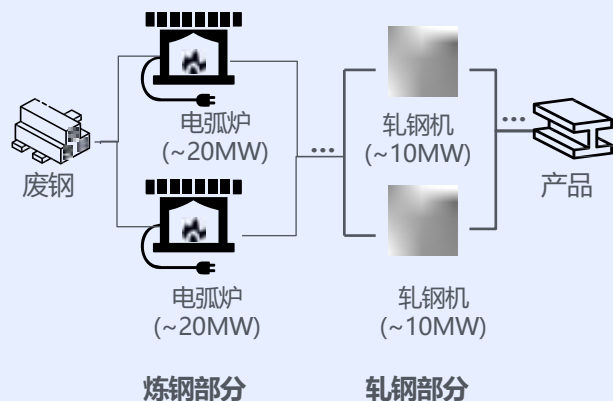
趋势要点1: 规模巨大、可调性强的工业可调节负荷可提供规模灵活性

腾讯以钢铁工艺流程中可调节负荷为初步虚拟电厂试点，并逐步囊括多种可调节资源；初步盘查，可通过调整电炉的生产节奏和功率以获得负荷灵活性；可根据不同优化目标对当日生产计划进行调整

-容量：典型短流程电炉炼钢生产线可提供5-20MW灵活性：

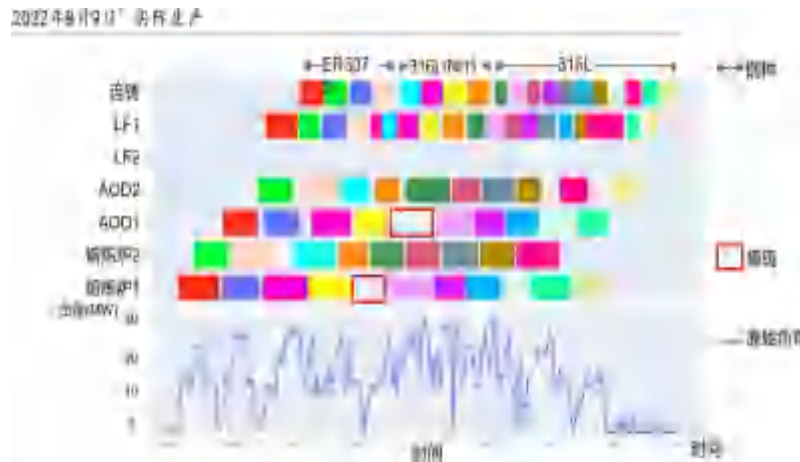
设备：在保证生产情况下调节产出；组织：需要平衡设备运行和人力排产；电网交互控制：电网需要按不同时间尺度通知、控制负荷方

典型短流程电炉炼钢生产线可提供5-20MW灵活性



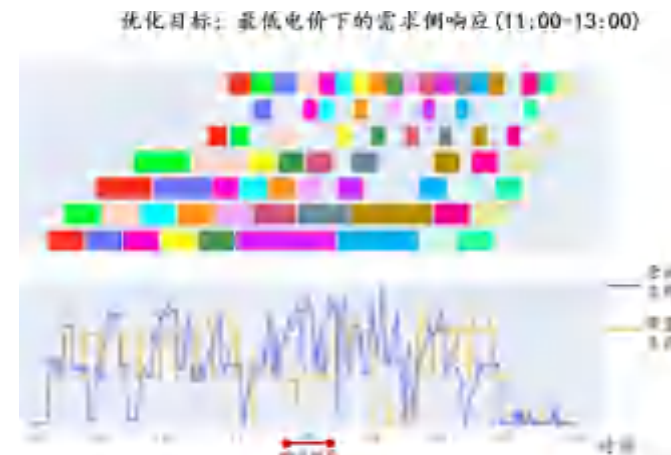
-方法：调整电炉的生产节奏和功率以获得负荷灵活性：

在响应时段降低电炉功率，延长电炉时间，或设备错峰，获得灵活性，目标在不损失总产量的要求下：1、降低整体用电费用；2、降低负荷、获得补助



-策略：可根据不同优化目标对当日生产计划进行调整模型与目标策略：

可以设定不同的生产目标，如：最快生产时间、最快生产时间+模铸、最低电价、最大化峰谷错开、最低电价下的需求侧响应（未来引导的响应方式）

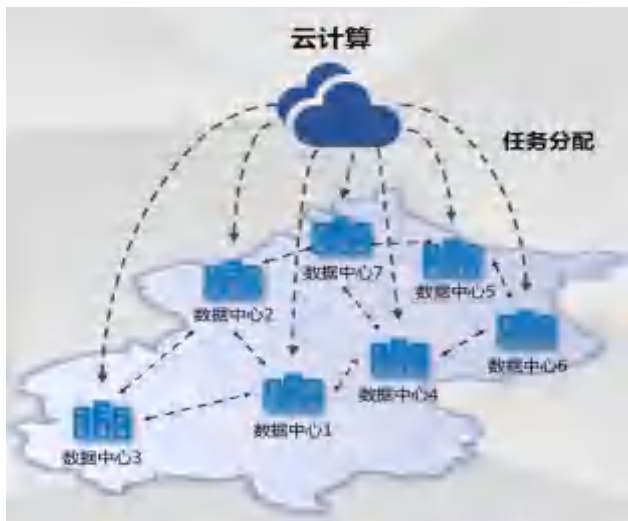


趋势要点2：数字基础设施的能耗优化，数据中心任务移动改变负荷

根据与其他数据中心是否互联，数据中心有两类：独立的数据中心（如传统的机房）；互联网数据中心（IDC）：多个IDCs通过光纤可以进行数据负荷传输，并与其之间传输数据的光纤网共同构成数据网络（i.e., 算力网络）

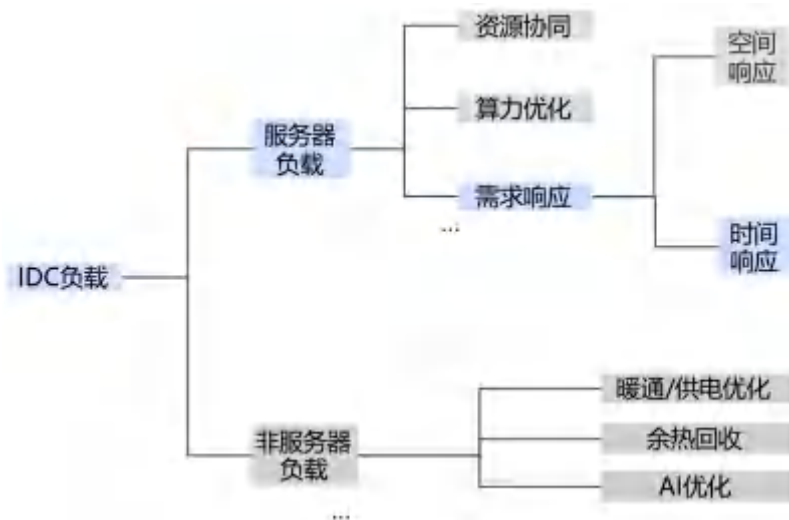
-转移：IDC负荷通过算力网络转移数据负荷，从而转移电力：

东西电力资源差异大，云计算场景下IDC负荷通过算力网络转移数据负荷，AI海量计算时代尤其重要。



-策略：数据中心腾讯自身应用场景需求响应特性测试：

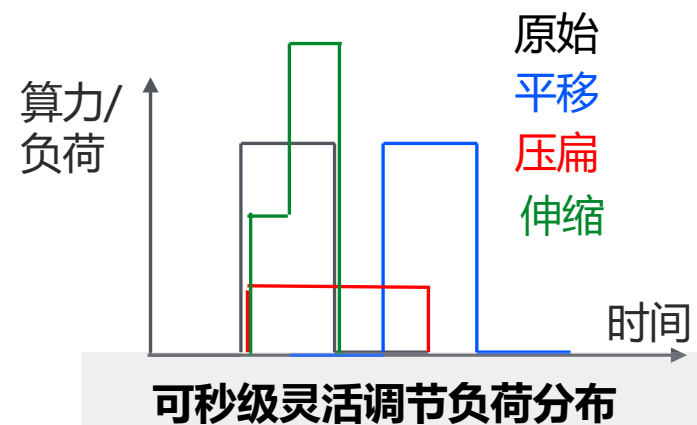
在不影响任务表现的情况下，调节服务器功率，关注响应速度、响应深度、响应时长、响应精度



-时效：探索数据中心的服务器的负荷灵活性策略，灵活调节：

对实时性不敏感的计算任务进行扩缩容和“断点续算”，快速改变负荷分布

包括科学计算、视频渲染等独立/耦合并行任务



趋势要点3：与C端互动，平移伸缩电动车充电速度，聚合为“大电池”

电动汽车(EV)将电力网络与交通网络这两个复杂的人造系统紧密耦合；EV 的充放电与出行由驾乘用户决定，形成以EV 为核心的信息-物理-社群系统(Cyber-Physical-Social System, CPSS)，带来新的问题与调控潜力

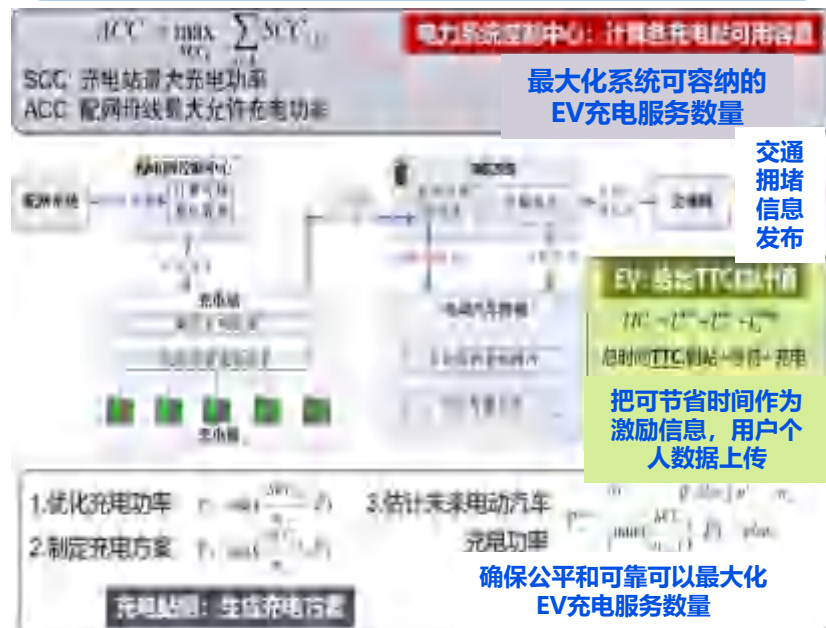
-时空协同：充电负荷通过交通网络转移，改变电网潮流分布：

电力-交通网络耦合的定义：充电负荷通过交通网络转移，改变电网潮流分布，新能源车时代尤其重要。



充电负荷时空分布（上海）

-方案：将电动汽车引导到对交通系统与电力系统都“最佳”位置充电：



-试点：基于价格调控的电动营运车辆调度是接下来重要的尝试方向：

目标：出行需求/充电负荷的供需不平衡？
可通过各类价格对营运车辆进行适当引导
出行价格可用于对车流密度与供需关系调控
充电价格可用于对充电需求时空分布的调控

出行服务供需不平衡



充电供需不平衡



顾问

郭凯天 腾讯集团高级副总裁

策划

司晓 腾讯研究院院长

指导专家

(名单按姓氏的首字母顺序排列)

李强

腾讯集团副总裁
政企业务总裁

李航

腾讯集团市场与
公关部总经理

李学朝

腾讯智慧交通
副总裁

刘杉

腾讯云副总裁
腾讯多媒体实验室负责人
腾讯杰出科学家

刘威

腾讯杰出科学家

刘颖

腾讯云副总裁

刘煜宏

腾讯数据平台部
副总经理

商世东

腾讯云副总裁
腾讯天籁实验室负责人

王迪

腾讯机器学习平台部
副总经理

岳淼

腾讯集团市场与
公关部副总经理

张军

腾讯集团市场与
公关部总经理

张胜誉

腾讯量子实验室负责人
腾讯杰出科学家

张正友

腾讯首席科学家、腾讯AI Lab
及腾讯Robotics X 实验室主任

赵岩

腾讯星海实验室负责人
腾讯杰出科学家

特别鸣谢

(名单按姓氏的首字母顺序排列)

安柏霖	陈创前	陈俊文	陈磊	陈立翰	陈梦珂	陳銘良	陈强	陈晓岸	陈煜东	代凌燕	窦淼磊	冯宇彦
付涛	高峰	谷晓峰	郭庆来	韩磊	郝少刚	贺辉群	侯文祯	胡仁成	胡轶强	黄婷婷	纪多斌	焦典
井天增	孔倩	李力	李芮	李伟	李忻	李阳阳	李源	李志成	黎国龙	梁聪慧	廖芳莉	林海川
刘超	刘金松	刘晶	刘兆瑞	罗璇	吕奇瑞	毛峻岭	孟祥芝	倪瑞捷	聂再清	潘兴德	庞宏启	戚元艷
茹炳晟	宋丹丹	唐敏豪	滕超	王焕超姚	王军	王宇凡	吴昊	吴羽	肖玮	谢良	熊依雪	徐佳康
杨卿	杨雪	杨子翊	杨钟灵	建华	姚麟	叶佳宾	曾耿华	曾钦杵	邹弘宇	张策	张川川	张建青
张露文	张强	张陶	张翔	张意晗	张云飞	赵军	郑兴源	郑一聪	郑宇			

主编

刘琼

执行编委

王强 刘莫闲 陈守双 周政华

主笔

高性能计算：刘莫闲

多模态智能体：王强、曹建峰、徐思彦

人形机器人：徐一平、王焕超

AI+基因计算：袁媛

数字交互引擎：陈孟、李南

脑机接口：吴朋阳

沉浸式媒体：宋扬、王强

星地直连通信：王鹏

eVTOL：陈楚仪、张雪琴

虚拟电厂：李瑞龙

联合出品

腾讯星海实验室、腾讯云异构计算产品中心、腾讯量子实验室、腾讯数据平台部、腾讯机器学习平台部、腾讯云智能数智人，腾讯Robotics X、腾讯AI Lab、腾讯北极光工作室、腾讯互娱研发效能部引擎技术中心、腾讯多媒体实验室、腾讯会议天籁实验室、腾讯 SSV 应急开放平台、腾讯安全天马实验室、腾讯智慧交通、腾讯SSV碳中和实验室、腾讯院士专家工作站、腾讯集团公共事务部、腾讯集团市场与公关部、腾讯新闻、腾讯研究院

Tencent 腾讯