

“巧妇难为无米之炊”

- 米 —— 数据
- 巧妇 —— 算法
- 锅 —— 算力



画一幅画，主题：巧妇难为无米之炊，图中要有灶台，巧妇侧身，“灶台”用土灶，“巧妇”手持铁铲，灶台上有一口大铁锅，锅里啥都没有



2018年诺贝尔经济学奖获得者William D. Nordhaus在《计算过程》一文中提出：“算力是设备根据内部状态的改变，每秒可处理的信息数据量”。

海量数据处理和各种数字化应用都离不开基于算力的加工和计算

基础算力

由基于CPU芯片的服务器提供的算力，主要用于基础通用计算，如通用的商用服务器，是基础算力的主要硬件设施。

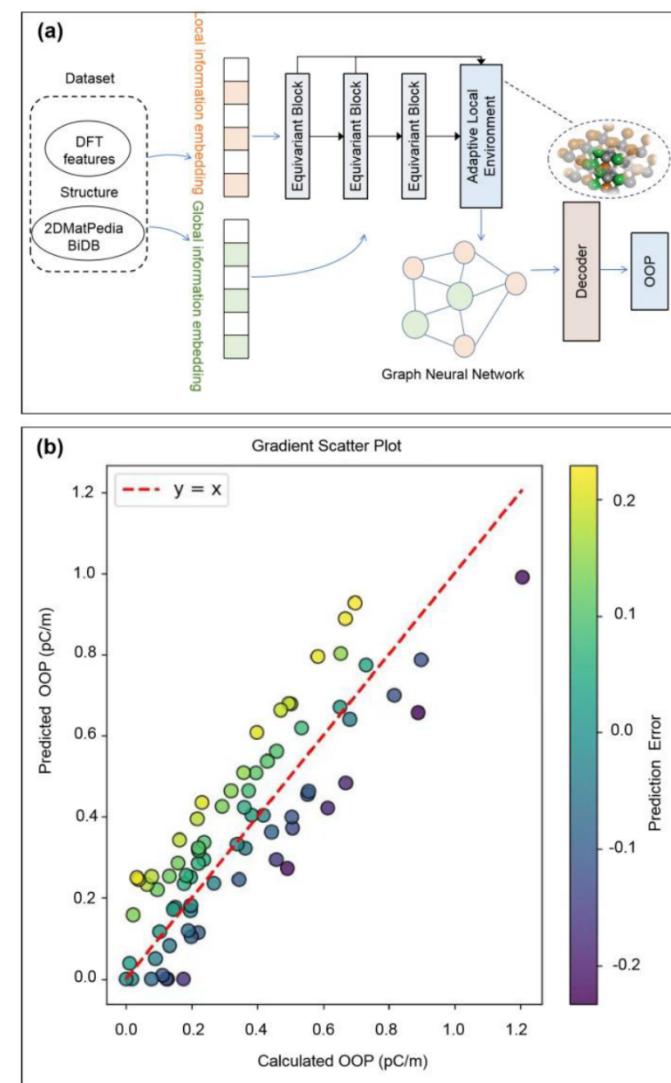
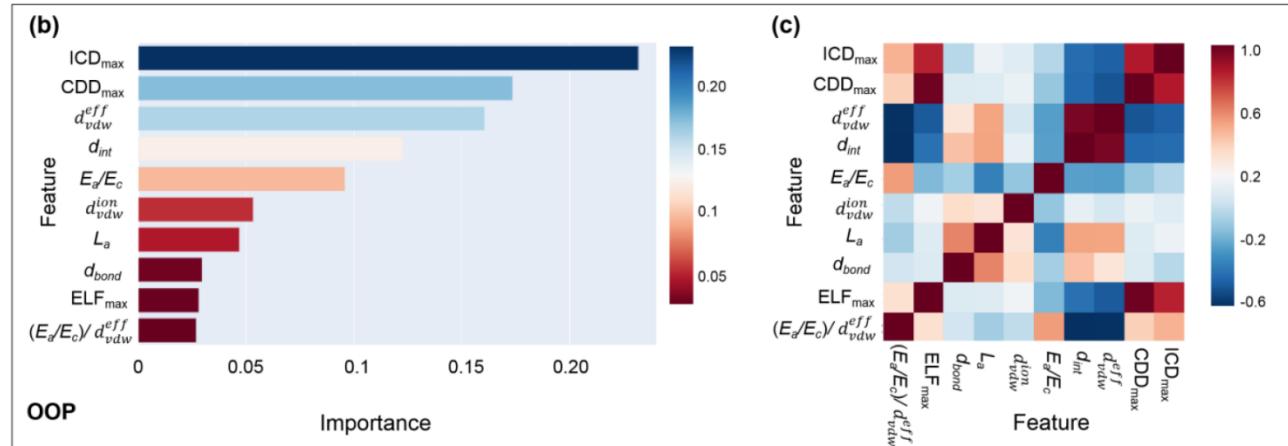
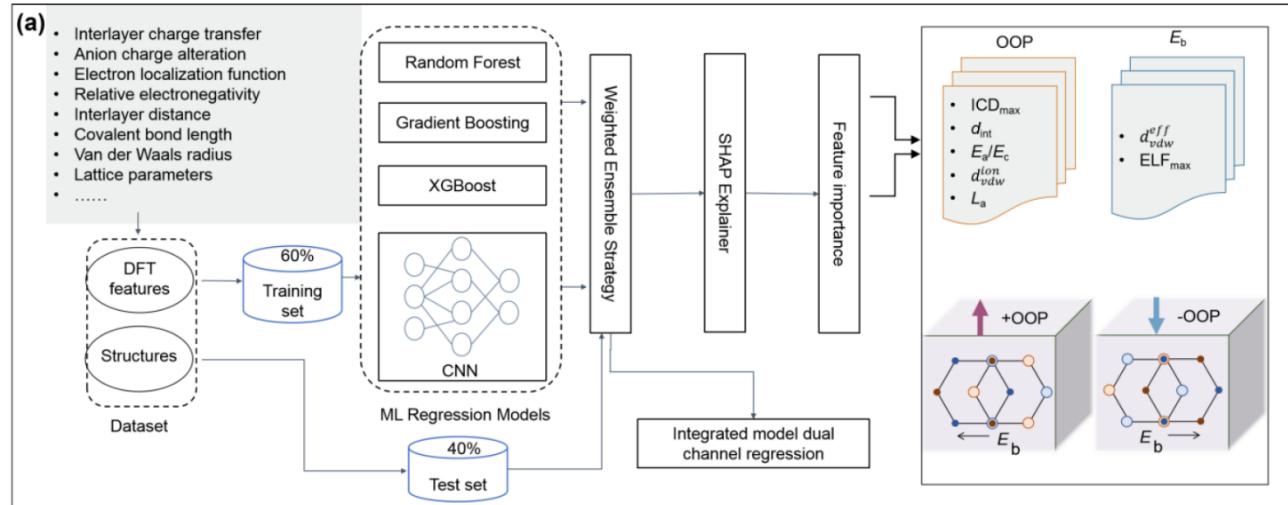
超算算力

由超级计算机提供的算力，主要用于高精度科学领域的计算，比如天气预报、药物筛选、基因分析、仿真模拟等等。

智能算力

由基于GPU、FPGA、ASIC等AI芯片的加速计算硬件提供的算力，用于AI模型的训练和推理计算，比如语音、图像和视频的处理。

例：材料铁电性特征提取与强度预测



- 机器学习-影响材料铁电性的主要物理化学特征
- 机器学习-预测铁电性的强度

Wang et al. Physical Review B 111(2025) 094106

各类型的算力都是不可或缺的

- 不同的计算场景需要不同类型的算力
- 完成一项科研任务可能需要多种算力



画一幅画，这幅画中要包含几种不同类型的“锅”，
例如：炒菜用的铁锅、炖汤用的砂锅、煎牛排用的平底锅

数据越来越多、算力相对不足

第二届中国数据大会 | 中国科学院院士周志华：从算力资源供给出发 提高算力利用率

2025-12-15 00:54 · 天眼新闻

“机器学习技术应用依赖算法、数据、算力三大核心要素。”12月13日，在第二届CCF中国数据大会开幕式现场，中国科学院院士、南京大学教授周志华以“考虑算力资源供给影响的机器学习理论框架”为主题作主旨报告。



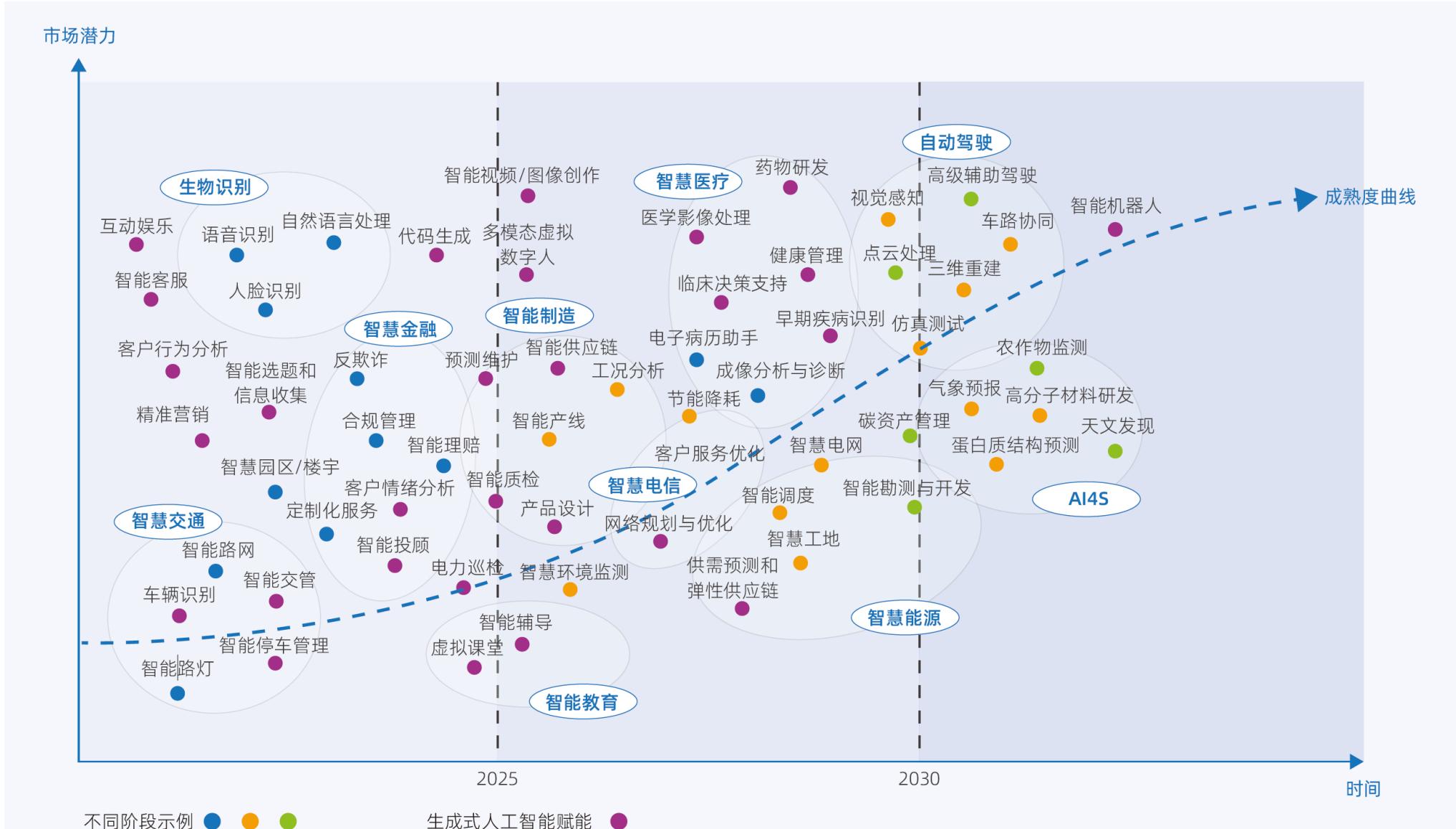
中国科学院院士、南京大学教授周志华

“算力已成为决定学习效果的关键变量。”周志华强调，同样的算法、同样的数据，不同的算力供应可能导致不同的学习结果，算力充足时模型可充分学习数据特征，不足时即便数据再丰富也只能通过抽样处理，导致数据资源浪费。

从应用层面看

- 科研范式变革催生算力的新需求

中国人工智能场景发展



来源：IDC, 2025

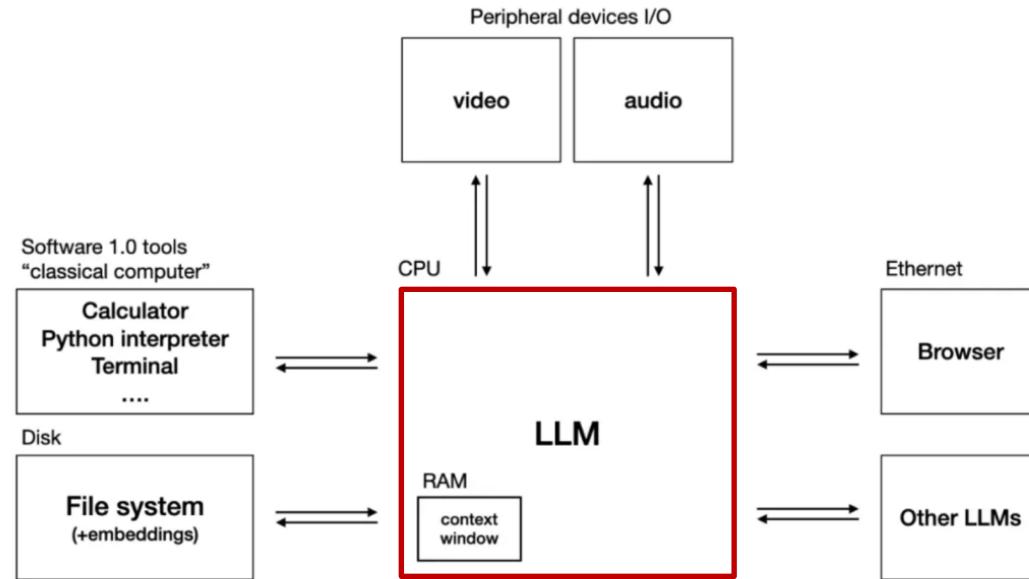
从应用层面看

- 科研范式变革催生算力的新需求

从计算范式看

- 大语言模型成为计算系统的核心

大语言模型成为计算系统的核心



图片来源: Andrey Karpathy

- 大语言模型置于计算系统的核心位置，类似于传统操作系统中 CPU 的角色
- 通过函数调用机制连接并控制外围设备(如视频、音频)、传统软件工具(如计算器、Python 解释器)、存储系统以及网络资源(浏览器、其他 LLM)

计算范式转变

传统计算机需要精确的指令序列，而 LLM 可以理解模糊的人类意图并将其转换为具体操作

抽象层次提升

如同 CPU 让程序员不必关心底层电路细节，LLM 让用户不必关心具体的程序实现细节

人机交互优化

Agent 替代人完成作步骤，普通用户也能完成复杂的计算任务

从应用层面看

- 科研范式变革催生算力的新需求

从计算范式看

- 大语言模型成为计算系统的核心

从外部环境看

- 超算 / 智算中心的建设如火如荼

国家超级计算中心（科技部批准成立的数据计算机构）



天津超算中心 (2009)

- “天河一号”
- “天河三号”



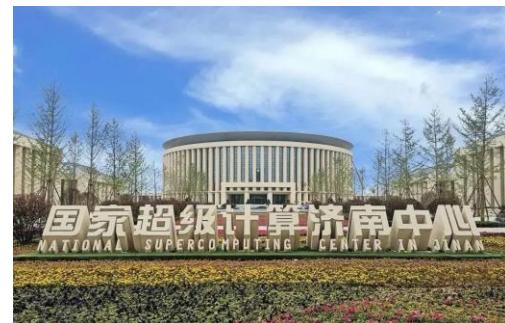
深圳超算中心 (2010)

- 资源利用率在国家级超算中心中首屈一指



长沙超算中心 (2010)

- 中西部地区的首家国家超级计算中心



济南超算中心 (2011)

- 首台采用自主处理器的千万亿次超级计算机“神威蓝光”



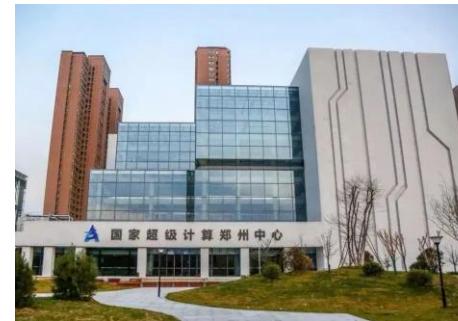
广州超算中心 (2013)

- “天河二号”



无锡超算中心 (2016)

- 全球首台峰值运算超过10万亿次/秒的超级计算机——“神威·太湖之光”



郑州超算中心 (2019)

- “嵩山”超级计算机



昆山超算中心 (2020)



成都超算中心 (2020)

- 最高运算速度达到10亿亿次/秒，进入全球前十



西安超算中心 (2020)



1

更充沛且不浪费的算力

课题组算力 << 校级算力

2

更便宜且可持续的算力

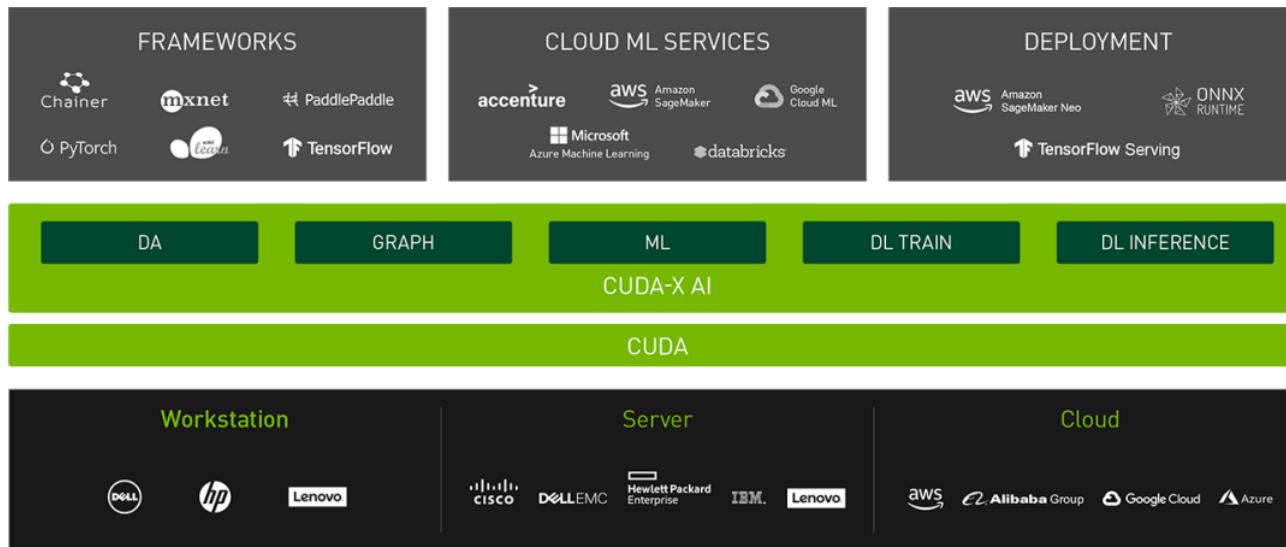
校级算力 << 国家级超算中心

3

更安全且更易用的算力

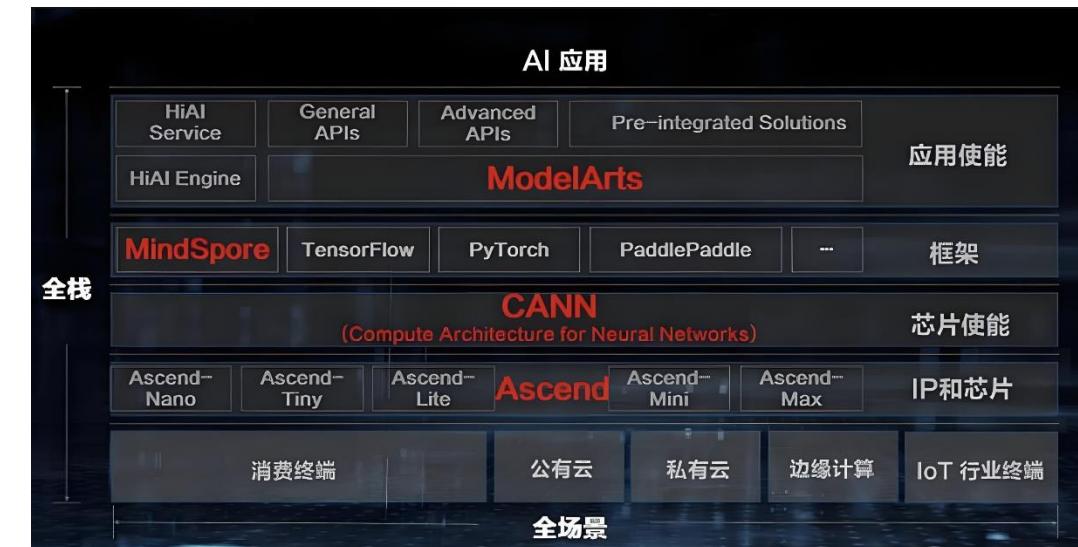
算力本地化、数据不出校

CUDA vs CANN



CUDA (Compute Unified Device Architecture, 统一计算设备架构)

支持开发者利用GPU(图形处理器)的强大并行计算能力高效处理各种复杂的计算密集型任务的一种并行计算平台和编程模型。



CANN (Compute Architecture for Neural Networks, 神经网络异构计算架构)

把AI训练框架(PyTorch、TensorFlow、MindSpore等)和昇腾芯片连接起来，使开发者不用关心芯片细节就能调用底层算力。

算力服务三种模式



01

“买一辆车”

裸金属服务

02

“租一辆车”

虚拟机服务

03

“打一辆车”

API服务

国内高校早期侧重于超算(用于高精度科学领域计算)算力建设

中国科大

The screenshot shows the homepage of the Chinese Academy of Sciences Supercomputing Center. The header features the university's name in English and Chinese, along with a logo. Below the header is a banner with the text "推动科学计算 培养人才" (Promote scientific computing, cultivate talents). The main navigation menu includes: 首页 (Home), 系统平台 (System Platform), 新闻公告 (News Announcements), 业界动态 (Industry Dynamics), 培训信息 (Training Information), 业务服务 (Business Services), 成果展示 (Achievement Display), 运行监控 (Operation Monitoring), 用户申请 (User Application), 资料手册 (Manuals), and 联系方式 (Contact Information). On the left, there are sections for "中国科大超级计算中心简介" (Introduction to the Chinese Academy of Sciences Supercomputing Center), "新闻公告" (News Announcements), "业界动态" (Industry Dynamics), "培训信息" (Training Information), and "计算软件" (Computing Software). The "新闻公告" section lists several news items with dates from 2024-02-28 to 2024-09-12. The "业界动态" section lists items from 2024-07-10 to 2020-06-20. The "培训信息" section lists items from 2024-04-01 to 2020-09-07. The "计算软件" section lists various software packages. A sidebar on the right shows a diagram of a dual-path power distribution system with numbered nodes 1 through 6. At the bottom, there is a copyright notice: Copyright 2009 中国科学技术大学超级计算中心 All Rights Reserved.

瀚海22超级计算系统

含2个管理节点、2个用户登录节点、1个可视化节点、25个8卡GPU计算节点（单节点64颗Intel Xeon Scale 8358 CPU核（2.6GHz, 48MB L3 Cache）、1TB内存、8颗NVIDIA A100 Tensor Core GPU（80GB显存、SXM4、600GB/s NVLink卡间互联）、10个3TB内存4路Intel CPU计算节点（单节点96颗Intel Xeon Scale 6348H CPU核（2.3GHz, 132MB L3 Cache）、3TB内存）、80个双路Intel CPU计算节点（单节点64颗Intel Xeon Scale 8358 CPU核（2.6GHz, 96MB L3 Cache）、256GB内存）、66个双路AMD CPU计算节点（单节点192颗AMD EPYC 9654 96核CPU核（2.4GHz, 768MB L3 Cache）、768GB内存）、11PB可用容量高性能存储，采用Mellanox HDR 200Gbps InfiniBand高速互联。纯计算节点共20352颗CPU核及200颗A100 GPU，总双精度浮点计算能力：3.06PFLOPS（千万亿次/秒，CPU：1.12PFLOPS，GPU：1.94PFLOPS）；Tensor Float 32(TF32)AI算力：62.4PFLOPS。

瀚海20超级计算系统

瀚海20超级计算系统采用Mellanox HDR 100Gbps高速互联，具有Intel Xeon Scale 6248、华为鲲鹏920 5250等不同类型CPU及NVIDIA Tesla V100 GPU和华为Atlas 300 AI卡等协处理器，共计2个管理节点、2个用户登录节点、720个普通CPU计算节点（采用高效节能的板级液冷技术）、10个双V100 GPU计算节点、8个2TB Intel AEP大内存节点、20个华为鲲鹏CPU计算节点构成（其中10个各含6颗华为Atlas 300 AI加速卡），计算节点共30480颗CPU核心和20块NVIDIA V100 GPU卡，总双精度浮点计算能力：2.51PFlops（千万亿次/秒，CPU：2.37PFlops，GPU：0.14PFlops），Atlas计算能力：3840 TOPS INT8 + 15360T FLOPS FP16。

曙光TC4600百万亿次超级计算系统

曙光TC4600百万亿次超级计算系统（519万亿次/秒），兼为安徽省教育科研网高性能计算平台，也为入网用户提供高性能计算服务。此套系统采用100Gbps和56Gbps高速互联，具有Intel Xeon E5 2600 v3、E5 2600 v4、E7 8800 v4、E3 1240 v5等不同类型CPU及Intel Xeon Phi和NVIDIA Tesla K80 GPU等处理器，共计1个管理节点、2个用户登录节点、7个存储节点、2个Lustre LNet路由节点及506个计算节点构成。计算节点共12200颗CPU核心，512颗Intel Xeon Phi融核（MIC）KNL核心和39936颗NVIDIA CUDA核心，总双精度峰值计算能力为每秒519万亿次（CPU：482.82万亿次/秒，GPU：14.96万亿次/秒，Intel Xeon Phi融核：21.28万亿次/秒）。

国内高校早期侧重于超算(用于高精度科学领域计算)算力建设



上海交大

上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

网络信息中心
Network & Information Center

首页 / 中心概况 / 部门领导

分管校领导

党委常委、副校长 管海兵

中心领导

书记 沈宏兴 hxshen@sjtu.edu.cn
主任 薛广涛 gt_xue@sjtu.edu.cn
副书记 周子豪 zhzhou@sjtu.edu.cn
副主任 姜开达 kaida@sjtu.edu.cn
副主任 林新华 james@sjtu.edu.cn

部门领导

数据业务部主任 章思宇 dfxbb@sjtu.edu.cn
计算业务部主任 罗 萍 xluo@sjtu.edu.cn
综合业务部主任 尹守婷 yst_1992@sjtu.edu.cn
网络业务部主任 谢 锐 sherry@sjtu.edu.cn

高性能计算

国内高校一流

高性能计算能力由思源一号 (CPU部分) 、π2.0和ARM平台提供，总计达到7 PFLOPS，聚合存储能力达到30PB。思源一号是目前国内高校第一的高性能计算集群，共 938个CPU节点，CPU采用Intel Xeon ICX Platinum 8358。π2.0集群是国内最早使用Intel Cascade Lake CPU构建的超算，拥有656台计算节点。ARM超算平台是国内首台基于ARM处理器的校级超算集群，一共有100个计算节点，单节点配备128核 (2.6 GHz)。

- 思源一号 (CPU部分)
CPU: 60032 核 算力: 5 PFLOPS
- π2.0集群
CPU: 26240 核 算力: 2 PFLOPS
- ARM集群
CPU: 12800 核 算力: 133 TFLOPS

人工智能计算

国内高校领先

人工智能计算能力由思源一号 (GPU部分) 和人工智能计算平台提供，张量计算能力达到 44 PFLOPS。思源一号GPU 采用 NVIDIA HGX A100 4-GPU，共 23 个节点。人工智能计算平台由8台NVIDIA DGX-2服务器提供，每台DGX-2配置16块NVIDIA Tesla V100 GPU加速卡。

- 思源一号 (GPU部分) : 92 块 A100 GPU
张量计算能力: 28 PFLOPS
- 人工智能计算平台: 128 块 V100 GPU
张量计算能力: 16 PFLOPS

复旦大学



2023年6月国内高校最大的云上科研智算平台CFFF上线

- CFFF全称Computing for the Future at Fudan，国内高校最大的云上科研智算平台。CFFF平台由面向多学科融合创新的AI for Science智能计算集群“切问”一号和面向高精尖研究的专用高性能计算集群“近思”一号两部分组成。
- CFFF拥有**1200张A100、120张A800**，整体算力达28Pflops，可实现千卡规模的并行计算，45亿参数大模型一天训完。
- CFFF目标：建成一批具有世界级影响力的科学大模型：生命科学大模型、材料科学大模型、大气科学大模型、集成电路大模型等。

国内高校新一轮建设重点都为智算(用于AI模型的训练和推理计算)



同济大学

2023年2月22日，同济大学信息办发布《科学计算设备采购项目》招标公告，建设校级科学计算平台，采购GPU节点231台，CPU节点591台，其中189台八卡A800GPU节点共**1512张A800**，智算规模与算力为全国高校一流。

The screenshot shows the Tongji University Procurement and Bidding Management Office website. The main content area displays the title '网络管理部 科学计算设备、科学计算中心机房硬件设备 采购项目成交...' (Procurement Project Result Announcement), the date '发布时间: 2024-07-05' (Release Date: July 5, 2024), and the number of views '浏览次数: 2494次' (View Count: 2494 times). Below this, there is a table summarizing the procurement details:

项目名称:	科学计算设备、科学计算中心机房硬件设备
采购单位:	网络管理部
采购单位联系人及方式:	许维胜 65983045
成交标的:	1、名称 科学计算设备 数量 1 单价 296468000.00 2、名称 科学计算中心机房硬件设备 数量 1 单价 10000000.00
成交单位:	赛尔网络有限公司 (911101087226182167)
成交金额:	142573000.00 人民币
合同签订日期:	2024-07-05

同济大学信息办科学计算设备采购项目公开招标公告

2023年02月22日 上海 招标公告 附件

发布时间	2023-02-22	项目编号	点击查看
招标预算	30646.8万元	资质要求	点击查看
HOT 采购规模走势	点击查看	HOT 招标方式	点击查看
HOT 潜在报名单位	点击查看	HOT 潜在中标人	点击查看
招标单位	点击查看	联系人: 查看	相关项目: 查看 累计招标金额: 查看
代理机构	点击查看	联系人: 查看	相关项目: 查看 合作业主数量: 查看
招标状态	【已截止】		
招标进展	<ul style="list-style-type: none">中标结果: 同济大学信息办科学计算设备采购项目中标公告招标变更: 同济大学信息办科学计算设备采购项目答疑、澄清公告招标公告: 同济大学信息办科学计算设备采购项目公开招标公告	2023-04-02	2023-02-24
附件	附表-科学计算设备详细技术参数.pdf	下载	预览

网络管理部 科学计算设备、科学计算中心机房硬件设备 采购项目成交...

发布时间: 2024-07-05 浏览次数: 2494次

项目名称:	科学计算设备、科学计算中心机房硬件设备
采购单位:	网络管理部
采购单位联系人及方式:	许维胜 65983045
成交标的:	1、名称 科学计算设备 数量 1 单价 296468000.00 2、名称 科学计算中心机房硬件设备 数量 1 单价 10000000.00
成交单位:	赛尔网络有限公司 (911101087226182167)
成交金额:	142573000.00 人民币
合同签订日期:	2024-07-05

全国第一个以
高校为建设主
体的**国产化**
人工智能算力
平台



The screenshot shows the "Service Guide" section of the Southeast University Network and Information Center website. The sidebar lists categories like Campus Network Services, National Chip Computing Power Services, Campus Information Construction, and System Introduction. The main content area discusses the university's self-developed AI computing power platform.

平台介绍
当前位置：首页 -> 服务指南 -> 国产算力服务 -> 平台介绍

东南大学昇腾算力平台是全国第一个以高校为建设主体的全国产化人工智能算力平台，可为广大师生提供优质、自主可控的国产化AI算力。平台目前覆盖了大模型服务、科研、教学等多个领域，用户覆盖本校30余个学院及行政单位，具备易用性、高可控、轻量化等优势，让更多非计算机大类师生也能快速上手。同时，将有具有丰富AI开发经验的原厂工程师为师生提供点对点指导服务，帮助每个课题组高效利用国产化算力。

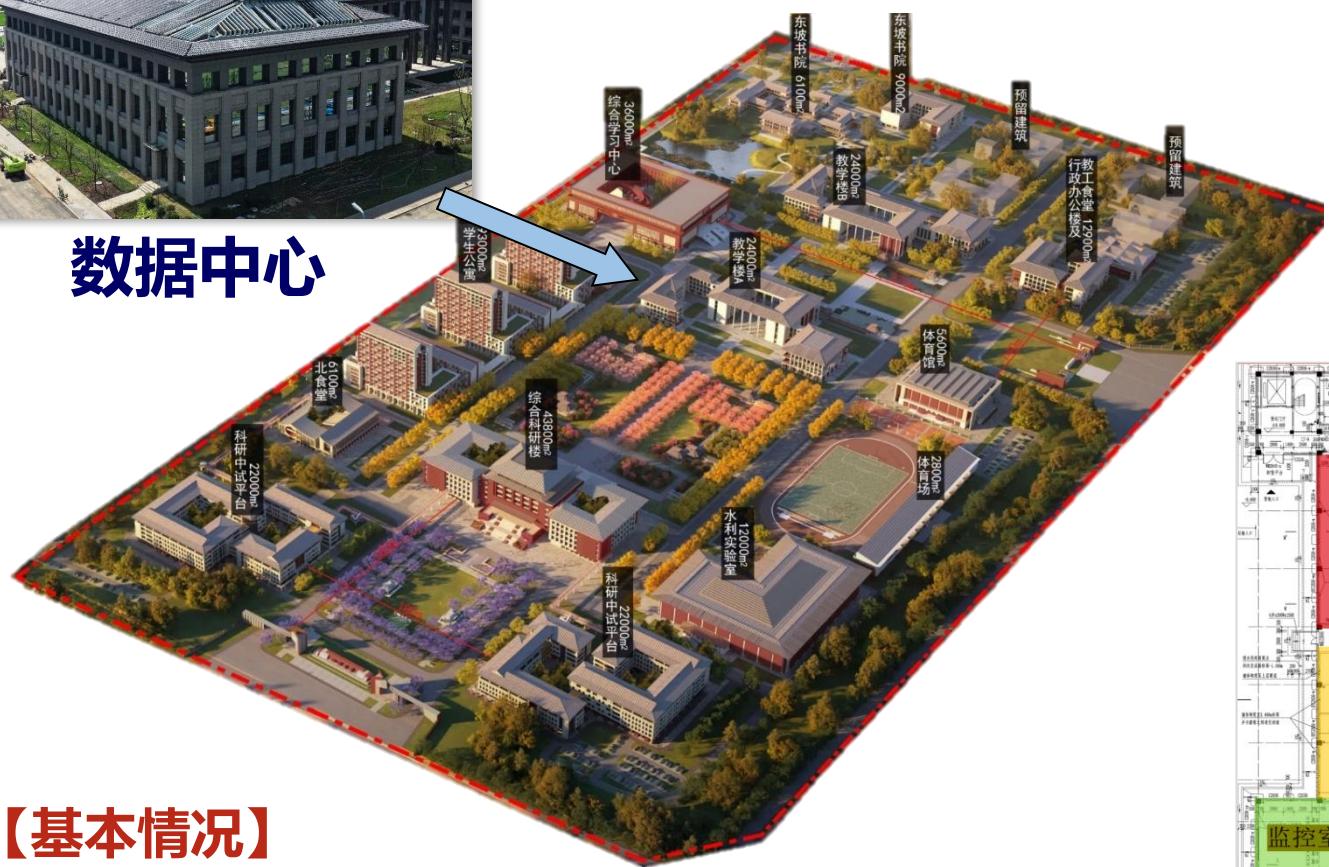
资源概况
当前位置：首页 -> 服务指南 -> 国产算力服务 -> 资源概况

东南大学现有昇腾AI服务器共45台，其中昇腾910训练服务器34台，算力规模85P；昇腾300I DUO推理服务器8台，算力规模4.5P；昇腾300I Pro推理服务器3台，算力规模0.8P。

四川大学新一代数据中心



数据中心

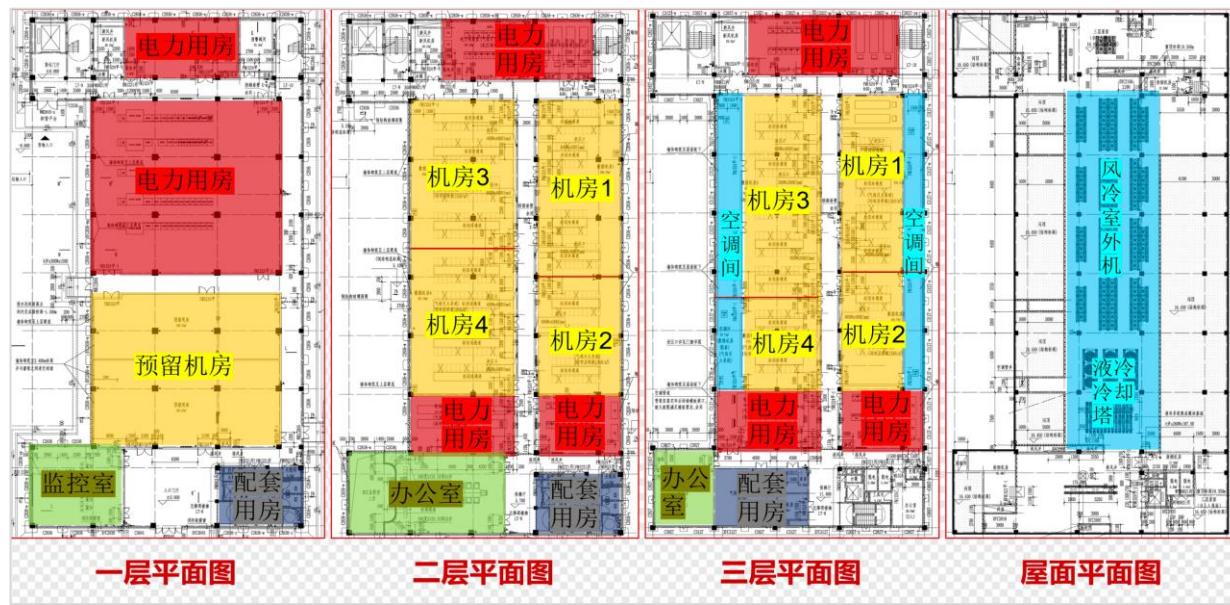


【基本情况】

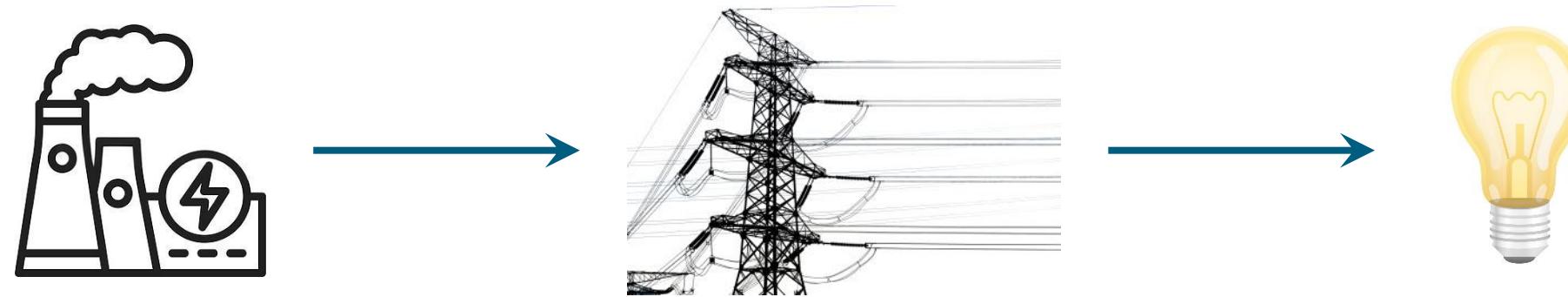
- 建筑面积7000m², 设计PUE≤1.3, 设计机柜总数 498 架
- 除未配置柴油发电机组及冷却塔外, 整体建设标准参照国标 (GB50174) A级标准

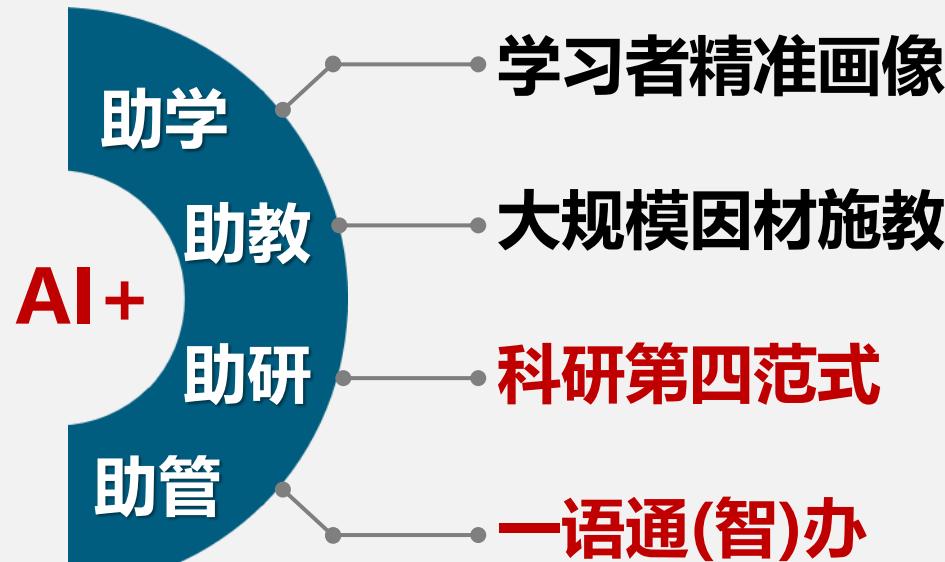
【功能布局】

- 一层规划基础设施及监控中心
- 二层规划托管、云资源机房及电力配套
- 三层规划算力机房及电力配套, 屋顶放置室外机和冷却装置等



信息世界三要素：计算、存储、传输





智能体：AI时代的应用程序

教学数据：完善大川学堂功能，汇聚全量教学资源

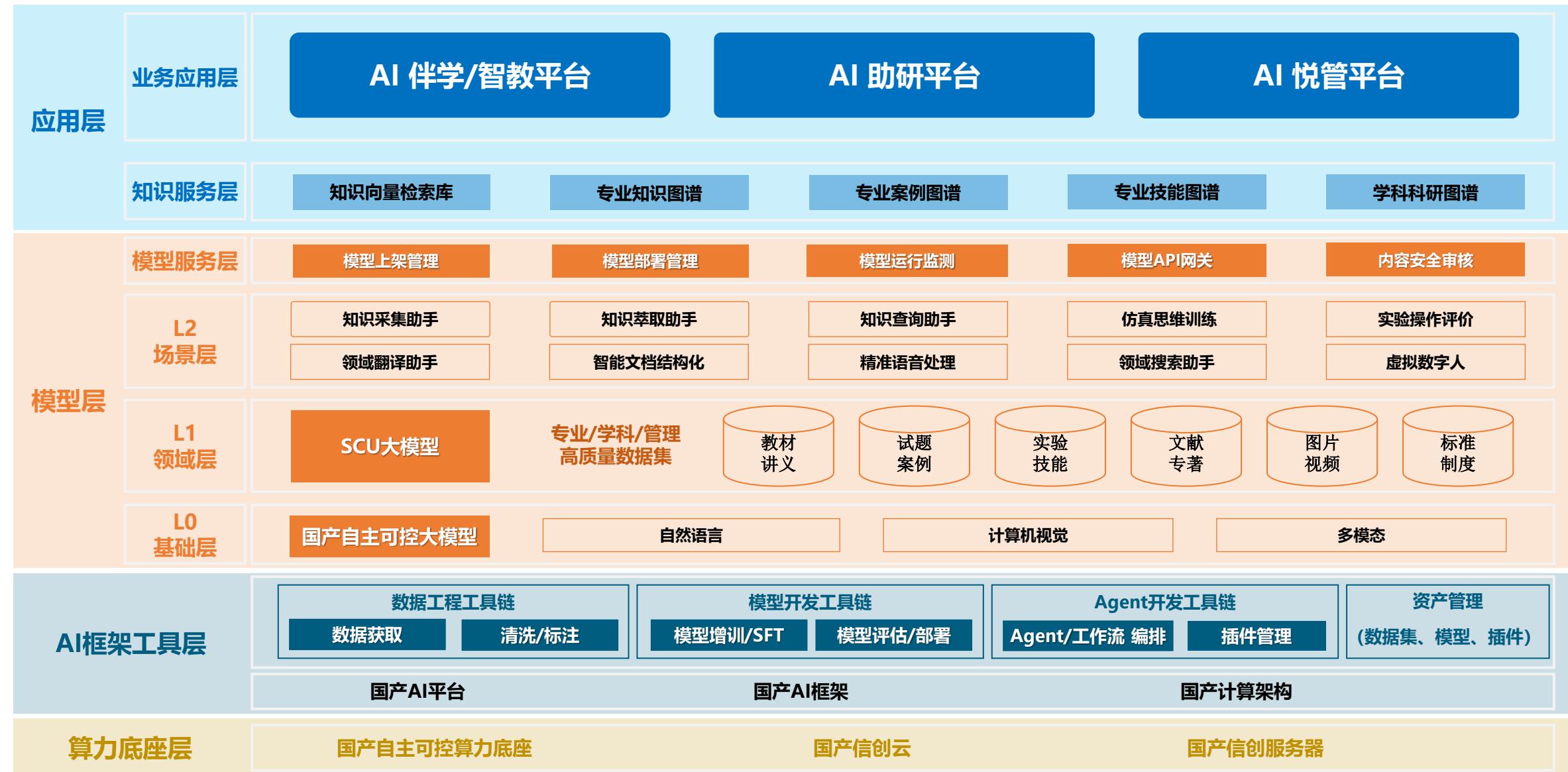
科研数据：提供知识提取服务，构建高质量知识库

管理数据：汇交业务系统数据，精准师生个人画像

算力基础：超算算力 + 智算算力 + 调度平台

智慧校园进阶：一卡通办 → 一网通办 → 一语通办

AI 赋能人才培养、科学研究、管理服务的总体技术架构



“十五五”教育数字化关键字



谢谢！请批评指正