智能计算中心规划建设指南

国家信息中心信息化和产业发展部

二零二零年十二月

智能计算中心定义



概念界定

智能计算中心(简称智算中心)是基于最新人工智能理论,采用领先的人工智能计算架构,提供人工智能应用所需算力服务、数据服务和算法服务的公共算力新型基础设施,通过算力的生产、聚合、调度和释放,高效支撑数据开放共享、智能生态建设、产业创新聚集,有力促进AI产业化、产业AI化及政府治理智能化



A 核心技术AI化

- ・ 基于深度学习、强化学习等创新AI技术
- · 重点围绕生产算力、聚合算力、调度算力、释放算力四大关键环节提升AI算力

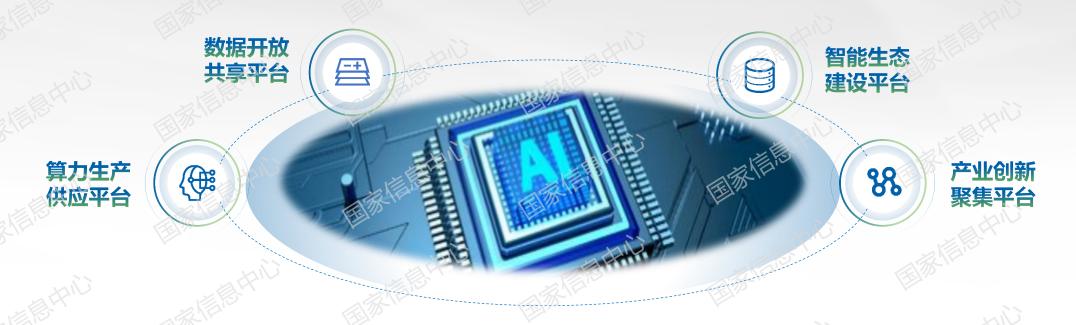
B 输出产品AI化

- 面向政府、企业等输出包括AI数据库、AI模型、AI开放平台等在内的多种AI产品
- 基于AI产品矩阵可以通过打包或定制化服务等方式助力AI产业化、产业AI化和政府治理智能化

C 服务应用AI化

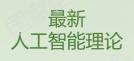
- 基于**先进的AI算力基础平台**、AI算力调度平台和AI算法模型,打造人工智能开放服务平台
- 汇聚并赋能行业AI应用,助力行业智慧应用高效化开发,加速行业和产业AI化

智能计算中心主要内涵及功能定位





智能计算中心技术架构



深度学习 自监督学习 强化学习 自动机器学习 跨媒体多模态

领先 人工智能计算架构

> AI芯片 AI服务器 高速互联 深度学习框架 资源调度

服务 数据服务 算力服务 算法服务 平台 算力生产供应平台 数据开放共享平台 智能生态建设平台 产业创新聚集平台 CNN RNN GNN GAN 模型文件 数据 处理库 高效模型 高效模型 高吞吐 训练框架 优化方法 推理引擎 释放算力 图像 开发接口 机器 学习 深度学习 强化 学习 自监督 学习 自然语言处理 自动机器学习平台 在线服务 AI场景 AI算法 AI工具 AI服务 智 算 CPU GPU 算力服务管理 细粒度切分 调度算 数据处理服务 中 FPGA 共享超分 开发训练服务 算力标准化 模型推理服务 负载均衡 容器化 算力调度 作 业 | 网络虚拟化 随需扩展 Al计算集群 聚合算 网络卸载 环节 弹性 可伸缩扩展 智能管理 高带宽 低延迟 软件定义存储 力 智能存储 算力聚合 算力集群 HHHHH AI训练服务器 CPU FPGA 软件生态 生产算 **Benchmark** Al训练算力 AI推理算力 先进 多样 成熟丰富 强大 高效 易用 力 算力芯片 算力机组 算力生态 算力输出

AI产业化

识别检测 语音交互 AI芯片 自动驾驶 机器人

产业AI化

智能制造 医疗影像 无人商店 智能客服 智慧物流 智慧农林

政府治理智能化

智慧交通 应急管理 防洪减灾 环境保护 地理测绘

...

建设意义和应用需求

一、加速AI产业化创新发展



高性能识别检测平台往往需要支持处理10万+的物体识别能力,具备分钟级别的模型生成响应速度,具备模型持续优化升级的能力,需要支持诸如深度分离卷积、残差密集网络(RDN)和隐式三维朝向学习等最新机器学习算法,同时支持各类算法的灵活配置以满足不同实际应用场景的需求



语音交互算法的训练语料数据量将很快突破百万小时,训练数据包含大量不同用户的口音数据、多领域歧义语料数据和具备复杂语法规则的数据。随着大规模语料样本数据的不断积累,需要更好利用智算中心具备的大数据挖掘技术、深度学习算法,构建更加高级的语音模型,为客户提供流式计算、在线计算、大规模离线计算等多种形式的服务



在为各类用户主体提供多元化AI算力服务的同时, AI芯片产业的快速发展必将带动智能计算中心的建设部 署



自动驾驶行业具备典型的海量数据、高并发、实时处理等计算属性,呈现出高度多元主体协同、智能识别和感知、系统模拟仿真计算、决策分析和预测预警等特性。该产业的高度集成化需要智能计算中提供先进的AI技术支撑、快速推动新产品的研发、测试和应用



机器人产业的发展迫切需要智能计算中心平台提供全链条AI技术支持,包括AI芯片、视觉识别、语音识别、机械臂和导航技术等,根据不同场景的不同诉求,提供"硬件+软件+服务"的全流程快速定制化专用机器人产品,大幅缩短产业链流程。

建设意义和应用需求

二、驱动产业AI化转型升级



传统的视觉检测系统难以应对越来越复杂的检测 场景以及更高的产品质量检测需求,工业质检车间的自动化和智能化越来越成为行业大势,这也为AI赋能工业 质检提供了落地的土壤



智能客服实际上是集合人工智能学、计算机科学、语言学等多门学科的综合应用,需要让机器主动去认知和学习,不断强化行为模式,提高思考能力,从而更加灵活地完成各项工作任务。智算中心能够为NLP(自然语言处理)、NLU(自然语言理解)、ASR(自动语音识别技术)、TTS(语音合成技术)等核心技术提供算力支持



医疗影像智能诊断应用需要基于大量的数据进行 知识挖掘和深度学习训练,获取具有实际医疗应用价值 的模型,这对数据的存储与计算都提出了相当高的硬件 需求,因此,需要智算中心提供强大的算力支持和一体 化开发环境



智慧物流产业的发展需要应用先进的AI算法、软硬一体的机器人产品与操控系统和整仓集成的完整生态。通过AI技术从图像和视频数据里获得高层次、可理解的信息,智能区分人和物,并能在50%变化率动态场景里稳健安全运行



无人商店的应用落地需要实现精准识别客户身份商品标签、智能计算、支付等全流程智能购物,同时需要基于大量的销售数据、行为轨迹、消费习惯等大数据挖掘与深度学习分析。智算中心强大的算力支持和人工智能技术应用是无人商店未来持续拓展的必要基础支撑



农业数据整体呈现规模庞大、类型多样、实时数据更新频繁等特征,涉及农林渔牧副各产业的多种数据针对农林渔牧副产业的不同特点,个性化定制智能解决方案变得尤为重要。需要有配套的智算中心作为支撑,实现农林感知数据的大规模存储,借助强大的算力支持农林数据的高效计算、训练和推演,最终实现快速和智能化的农林生产决策

智能计算中心的建设原则



总体规划 政企协同

加强区域智能计算中心的 统筹规划,探索开放共赢 的新型建设运营模式



需求牵引 先进适用

聚焦人工智能 领先技术,开展典型示范 和场景应用



夯实基础 培育生态

发挥对数字经济产业、 传统产业数字化的叠加 效应、乘数效应



保障安全 创新发展

保障网络安全和数据安全, 创造新供给, 激发新需求, 培育新动能

建设原则

智能计算中心的建设要点









以全面提升 AI计算力为核心

以促进数据 开放共享为基础

以培育区域 智能生态圈为根本

着力推动 AI产业创新聚集

- 据 OpenAl 统计,自2012 年以来随着深度学习模型的演进,模型计算所需计算量已经增长30万倍, Al模型构建所需算力呈现出阶跃式增长特点
- IDC在《2020H1中国加速计算市场调查报告》指出,预计到2024年AI加速计算市场规模将达到2019年的2.6倍

智算中心

- ✓ 免费或低成本提供很多开源的人工智能算法及其代码
- ✓ 以强大算力驱动AI模型对数据进行深度加工,源源不断产生各种智慧计算服务
- ✓ 通过平台开放接口的方式将行业领军企业的算法能力、数据资源及运营服务等**输出给**IT 基础薄弱的企业



01

降低AI应用成本 提高算力效率

- 算法创新、数据和可以用于训练的算力被认为是推动人工智能发展的三个关键要素
- 未来随着智算中心大规模投入使用,以AI算力、数据为基础、以AI模型生成和多场景应用的AI技术生态链将逐渐形成

智算中心

- ✓ 实现从模型开发、训练、部署、测试、发布的"流水式生产"与一站式交付,加速算力产业链的形成
- ✓ 加快AI算法公司和行业用户快速对接,带动一批新兴的专业化大数据企业崛起,驱动一批创新型企业加速成长
- ✓ 带动新能源、新材料等高新技术产业发展,进而构建起以AI算力为核心的人工智能产业 新生态体系



加速AI生态对接 推动创新产业聚集

- 智算中心是智慧时代经济社会运行所必须的关键公共算力新型基础设施
- 数据中心是产业转型升级、新旧动能转化的必要支撑,但现有数据中心算力服务能力已经难以有效满足实践需求

智算中心

- ✓ 凭借其强大的AI算力能够为不同用户提供算法产品与服务,解决用户面临的业务痛点, 实现用AI**为传统行业用户赋能**
- ✓ 使得智慧计算像水电一样成为城市的基本公共服务,进而帮助城市中小企业、创新型企业和传统企业降低企业AI技术研发、应用和部署成本,增强企业创新和转型发展能力



促进产业转型升级 加速新旧动能转化

- 智算中心作为"新基建"数字基础设施的主力军,其所承载的AI算力将是驱动政府治理智能化发展的核心动力
- 丰富的应用场景为智算中心提供了大量AI训练数据,海量训练数据对于数据中心算力提出更高要求



加强政府治理提升公共服务能力

智算中心

- ✓ 政府作为现代政府治理主体,能够在智算中心强大算力支撑下开展精细化、智能化政府治理,智算中心将在智慧交通、应急管理、防洪减灾、环境保护、地理测绘等场景中有着极大应用空间
- ✓ 智算中心作为未来AI算力生产供应中心,其算力能够充分满足未来政府治理和公共服务数据训练要求



改善经济运行格局 激发经济发展活力

- 人工智能计算需求未来将占据80%以上的计算需求,智能计算成为未来经济的主要增长点之一
- 到2035年,人工智能将帮助各国显著扭转经济增速近年来的下滑趋势
- 智能计算产业作为数字经济的重要载体,将作为一支新兴力量,带动全球数字经济走向繁荣

智算中心

- ✓ 埃森哲研究,通过有效应用智能计算,中国**经济增长率有望上升至**7.9%,增长额高达7.1万亿美元。
- ✓ 麦肯锡的预测,人工智能将每年为中国经济增长贡献0.8至1.4个百分点

智能计算中心的建设、运营模式



建设运营总体思路

- 采用政企合作建设运营的框架,高效整合政企资源。
- 注重激发市场创新活力、提高建设运营效率。
- 发挥智算中心作为创新载体的公益属性,促进成果转化应用,带动新兴产业发展与产业链上下游高效协同。

模式名称	优势	劣势	适用范围
政府投资建设模式	建设运维过程更加安全可控政府对项目决策和执行效率高	财政负担全部投资,压力大 资金使用效率可能不高	涉及国家安全或重大公共利益的项目 (如政务领域、基础数据库、信息安全平台设施 建设等项目)
企业建设 运营模式	财政资金压力较小 大企业平台化和整体性优势,有利 于加快区域间协同发展。	政府控制力较弱 存在大企业绑架风险	不涉及国家安全,具有公共服务特征和一定公益性和盈利空间的项目 (如通信网络、社区、文旅等项目)
政府购买服务模式	财政短期资金压力较小 资金使用效率高,政府策略灵活	有效服务市场供给不足 不利于长期稳定协作	基础性、公共性、公益性的共性产品和服务项目 (如云计算服务、通用软件、城市管理等项目)
政府和社会资本合作模式	同时发挥政府和市场的优势,既安全有序,又有市场化效率 提升项目建设运营专业化程度 减小财政资金压力和投资风险	机制设计复杂,存在制度性风险 融资成本较高。	属于公共服务领域,市场化参与程度较高、国家安全和公共利益方面潜在风险较低、需求长期稳定、使用者付费比例高于10%(原则上属于能源、交通运输、市政公用、农业、林业、水务、环境保护、保障性安居工程、教育、科技、文化、体育、医疗卫生、养老、旅游等行业领域的项目)



智能计算中心的建设、运营模式



投资主体

政府主导模式,政府作为投资主体加快推进智能计算中心落地

- 智能计算中心的公共属性
- 智能计算中心的自主性



承建方主体

政府主导下的政企合作模式,由企业具体承建智能计算中心

· 行业头部企业/国有资本控股企业



运营主体

运营主体类型应更加多元,运营模式也更加灵活。

- 更宜采取"投-建-运"一体化模式
- 可委托外包给政府出资控股或国有资本控股的专业运营公司







智能计算中心的建设、运营模式



基于政府独立投资的建设模式

政府直接投资和管理,需要对项目的建设周期、存在风险、成本收益等有很高的把控能力。优势在于政府直接主导智能计算中心建设,可以起到示范效应,建设完成后**智能计算中心所有权归政府所有。**



基于**特殊项目公司** (SPV) 的建设运营模式

政府与企业共同出资成立智能计算中心建设运营项目公司,政府与企业在合作框架协议下按比例出资建设智能计算中心,**政府应对项目公司具有控股权。**



基于"国家-地方-企业"共建的建设运营模式

地方政府会同相关企业共同申请与智能计算中心相关的国家级试点示范项目工程,以项目工程为载体,集国家、地方和企业三方力量,由国家和地方按比例出资建设智能计算中心。

智能计算中心的实施路径



需求导向, 明确建设思路

- 一、需明确建设需求
- 二、根据需求制定设计方案、确定承建方与建设模式
- · 具备一定人口规模和经济实力的中等以上城市
- 重点发展人工智能产业的城市



改建结合, 优化建设方式

- 一、建设专用智算中心
- 对于产业智能化发展需求迫切、传统数据中心规模较小的地区
- 二、数据中心智能化改造
- · 硬件方面:采用CPU+GPU芯片的方案
- · 软件方面: 建立面向人工智能应用的通用算法和模型平台



软硬兼施,提升服务能力

- 一、采用先进硬件芯片,强化算力支撑
- 采用新型人工智能芯片和新型计算方法相结合提供更 快的计算能力
- 二、建立通用软件平台,加强算法保障
- 建立面向人工智能应用需求的通用算法和模型平台



先进适用,统一建设标准

- 一、面向智算中心长期应用
- 坚持实用、经济的建设原则
- 二、面向智算中心的多领域应用
- 坚持资源共享、弹性扩展的建设原则



规范推进,加强项目管理

- 一、确保合规推进
- 结合本地规划、科学论证、审批报备
- 二、开展全流程项目管理
- 工程咨询→设计→监理→验收测试→成效评价



丰富场景, 打造产业生态

- 一、深度挖掘市场需求,丰富智算中心应用场景
- 智能制造、智慧医疗、智慧农业等行业智能应用场景
- · 智算中心可有效提升城市科学治理能力和政府治理水平
- 二、产学研用结合, 打造人工智能产业生态

版权声明

本报告由国家信息中心信息化和产业发展部联合浪潮共同研究发布。

本报告中所涉及的图片、表格及文字内容的版权归国家信息中心和浪潮共同所有。其中部分数据在标注有来源的情况下,版权归属原数据公司所有。本报告取得的数据来源于公开资料,如有涉及版权纠纷问题,请及时联络我们。

任何机构、个人在引用本报告数据或转载有关报告内容时,需标注来源。违反上述声明者,将追究其相关法律责任。

版本号: 2020.12

全文下载



