

人工智能赋能新质生产力，国产 AI 产业迎来新机遇

--7月人工智能月报

核心观点:

- 7月板块震荡微跌，板块调整接近尾声。**人工智能板块指数（884201.WI）7月涨跌幅为0.52%，同期上证综指下跌0.97%，沪深300下跌0.57%，创业板指数上涨0.28%，7月板块整体继续维持弱势震荡调整。我们认为，人工智能板块年初受到催化迎来新一轮上涨，3月底进入震荡调整周期，调整时间超过4个月，估值泡沫逐渐消化，有望逐渐迎来估值修复。
- 人工智能驱动发展新质生产力，抓紧打造自主可控的产业链供应链。**7月21日，《中共中央关于进一步全面深化改革推进中国式现代化的决定》发布，报告中提出要健全因地制宜发展新质生产力体制机制；健全促进实体经济和数字经济深度融合制度；健全提升产业链供应链韧性和安全水平制度。我们认为，新质生产力的核心在于创新，人工智能是最契合新质生产力核心的技术，或将成为驱动新质生产力发展的中坚力量及重要引擎，赋能传统行业并衍生出多元化场景，促进实体经济与数字经济深度融合，推动产业数字化高质量发展。同时报告中提出要“抓紧打造自主可控的产业链供应链”，为全会报告中唯一采用“抓紧”表述，将自主可控提升至前所未有的高度，未来我国将人工智能产业链硬件芯片、大模型、软件生态等细分领域将加速实现自主可控，国产人工智能产业链打开广阔市场空间，建议持续关注产业链投资机会。
- 车路云一体化助力“大模型”上车，萝卜快跑火爆出圈。**7月3日，工信部等五部门正式公布将在北京、上海、广州、深圳、重庆、沈阳等20个城市开展智能网联汽车“车路云”一体化应用试点，随后百度旗下的萝卜快跑无人驾驶出租车快速走红，百度今年5月公布的第一季度财报显示，萝卜快跑的自动驾驶订单约82.6万单，同比增长25%，截至2024年4月19日，萝卜快跑累计订单超过600万单。我们认为，车路云一体化有望推动我国智慧城市、智能交通系统、大数据平台等建设，加速更高级别自动驾驶落地，车路云一体化顶层设计与政策法规推动我国智能网联车发展，是我国智能网联车产业化进程“里程碑”时刻。
- 深化财税体制改革，数字化成为财税体制改革重要抓手。**三中全会提出深化财税体制改革，健全有利于高质量发展、社会公平、市场统一的税收制度，优化税制结构，根据艾瑞咨询预测，2025年中国财税数字化市场规模将达702亿元。我们认为，财税体制改革的重要抓手之一是构建财税数字化，打造财税数字化生态体系是关键。通过信息化技术、数据要素等生产要素将赋予财税体制改革广阔空间，未来发展方向应该在财税数字化生态体系上，通过生态体系企业可以整合自身资源同时为客户提供一站式服务，财税数字化建设迎来加速发展。
- 最强开源大模型 Llama 3.1 横空出世，OpenAI 开战轻量化小模型。**美国当地时间7月23日，Meta推出最新的开源大模型系列 Llama 3.1，总共有8B、70B、405B三个参数规模，三个版本上下文窗口长度从8k扩展16倍到128k，并且同时支持8种语言，其中 Llama3.1-405B 采用超过15万亿个 token 记性训练，使用1.6万张英伟达 H100 芯片训练；美国时间7月18日，OpenAI 正

计算机行业

推荐

维持评级

分析师

吴砚靖

☎：(8610) 66568589

✉：wuyanqing@chinastock.com.cn

分析师证书编码：S0130519070001

鲁佩

☎：(021) 20257809

✉：lupei_yj@chinastock.com.cn

分析师证书编码：S0130521060001

研究助理

胡天昊

☎：(8610) 80927637

✉：hutianhao_yj@chinastock.com.cn

行业数据

2024-7-31



资料来源：Wind，中国银河证券研究院

式发布了多模态小模型 GPT-4o mini, 比 OpenAI 当前的最先进小模型更强大、价格更低。我们认为, 一方面, Meta 一直坚持大模型开源, 并表示 Llama 3.1-405B 多项基准测试超过 OpenAI 的 GPT-4o, 将成为行业的重要里程碑, 未来将有更多开发者加入到开源大模型以及生态的使用和建设, 开源大模型将成为未来人工智能发展的方向。另一方面, OpenAI 推出 GPT-4o mini 向大模型“轻量化”进军, 将更加适用端侧 AI, 扩宽端侧 AI 场景。

- **投资建议:** 建议关注 1、国产算力产业链及生态伙伴; 2、算力基础设施; 3、逐步向国产的自主 AI 算力为底座迁移的 MAAS 及下游应用端; 4、AI PC 产业链整机及代工、结构件、开发套件厂商; 5、自动驾驶产业链及车路云一体化厂商。建议关注: 科大讯飞、海康威视、大华股份、中科曙光、工业富联、网宿科技、海光信息、龙芯中科、华勤技术、润泽科技、曙光数创、中科创达、锐明技术、万集科技、千方科技、金溢科技、税友股份、中国软件、万兴科技、国能日新、神州数码、金山办公、深桑达 A、超图软件、萤石网络、彩讯股份、同花顺、财富趋势、嘉和美康、上海钢联、柏楚电子、中控技术、诺瓦星云等。
- **风险提示:** 技术研发进度不及预期风险; 供应链风险; 政策推进不及预期风险; 消费需求不及预期风险; 行业竞争加剧风险。

目录

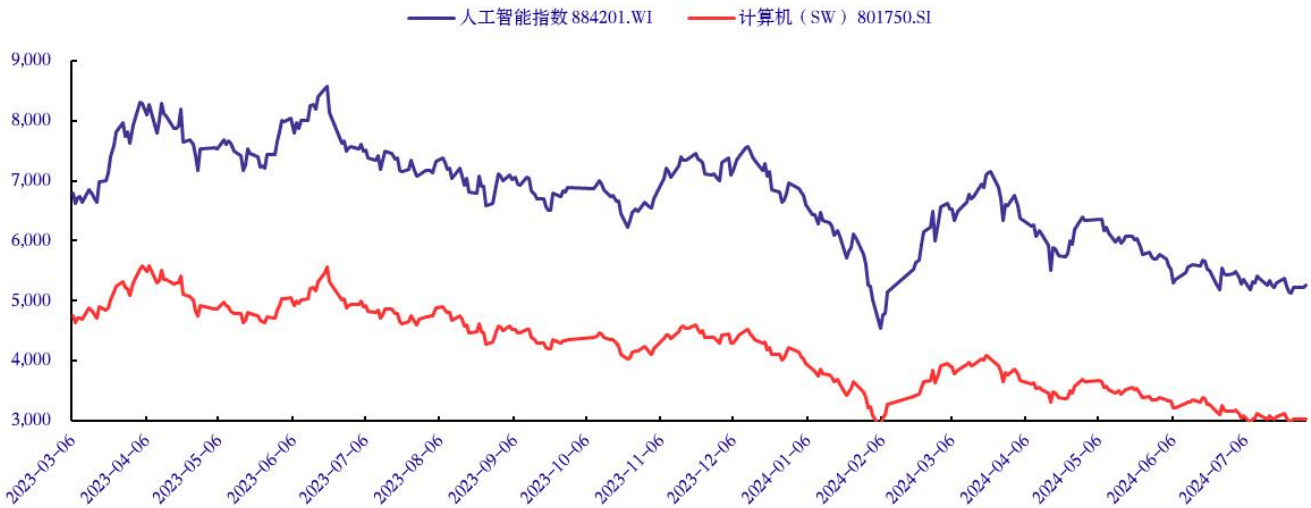
一、市场行情回顾.....	2
(一) 整体行情.....	2
(二) 代表企业.....	2
(三) 板块估值.....	3
二、人工智能产业政策动态.....	5
(一) 数据要素、数据交易所最新新闻及政策.....	5
(二) 算法端：国内外巨头大模型动态.....	8
(三) 算力端：AI 服务器、AI 芯片最新动态.....	10
三、前沿行业动态.....	11
(一) 前沿技术动态.....	11
1. PNN 或可从根本上改变 AI 系统高能耗、低吞吐率、高延迟等问题.....	11
2. FlashAttention-3 发布，H100 的 FLOP 利用率提升至 75%.....	12
3. 谷歌 DeepMind 新算法 JEST 训练提效 13 倍，能耗暴降 10 倍.....	13
(二) 前沿政策动态.....	14
四、前沿企业动态.....	14
(一) 前沿产品动态.....	14
1. 英伟达联合 Mistral AI 发布了可在本地硬件上高效运行的 to B 端 AI 大语言模型.....	14
2. 蚕丛机器人在 2024 WAIC 发布智能双模充电桩（FlashHub）.....	15
3. 昆仑万维等发布通用 Agent 新框架——Cradle.....	16
4. OpenAI、Mistral AI & 英伟达、HuggingFace 竞逐小模型，苹果 DCLM 全开源.....	17
5. 全球首个芯片设计开源大模型 SemiKong，革新半导体工艺和制造技术.....	19
6. 中国农业大学发布“神农大模型 2.0”：覆盖育种、种植、养殖、农业遥感及气象.....	19
(二) 投融资事件.....	21
五、投资建议.....	21
六、风险提示.....	22

一、市场行情回顾

(一) 整体行情

A 股人工智能指数（884201.WI）截至 7 月末收盘价为 5257.12，月涨跌幅为 0.52%。计算机行业指数（801750.SI）截至 7 月末收盘价为 3037.17，月涨跌幅为-0.30%。

图 1：6 月人工智能指数表现



资料来源：wind,中国银河证券研究院

(二) 代表企业

A 股 Wind 人工智能指数（884201.WI）截至 7 月 31 日总市值 13070.32 亿，含成分股 71 支，权重等分。上市板分布为主板 14 支，创业板 30 支，科创板 8 支，中小板 19 支。

表 1：7 月份人工智能板块涨幅前十

股票代码	股票简称	7 月 31 日收盘价（元）	月涨跌幅	相对计算机
688256.SH	寒武纪-U	264.06	32.91%	33.21%
300613.SZ	富瀚微	38.83	17.99%	18.29%
300222.SZ	科大智能	6.3	17.54%	17.84%
300245.SZ	天玑科技	6.14	14.34%	14.64%
300078.SZ	思创医惠	2.45	8.89%	9.19%
600756.SH	浪潮软件	10.84	7.33%	7.63%
300250.SZ	初灵信息	11.63	6.70%	7.00%
001339.SZ	智微智能	29.36	6.34%	6.64%
603019.SH	中科曙光	43.9	5.78%	6.08%
600728.SH	佳都科技	3.92	5.66%	5.96%

资料来源：Wind,中国银河证券研究院

表 2：7 月份人工智能板块跌幅前十

股票代码	股票简称	7 月 31 日收盘价（元）	月涨跌幅	相对计算机
002368.SZ	太极股份	16.74	-25.11%	-24.81%

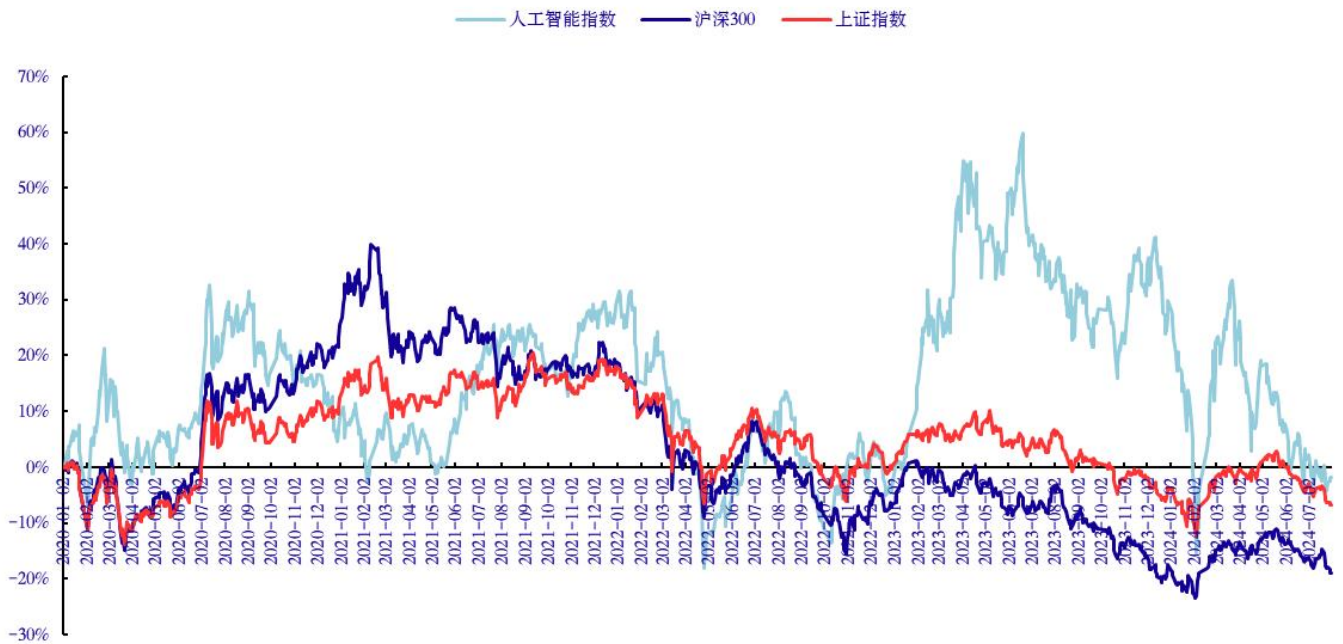
002123.SZ	梦网科技	6.35	-9.67%	-9.37%
301025.SZ	读客文化	9.08	-9.11%	-8.81%
002298.SZ	中电兴发	4.03	-8.82%	-8.52%
300624.SZ	万兴科技	48.58	-8.46%	-8.16%
002230.SZ	科大讯飞	39.72	-7.29%	-6.99%
300781.SZ	因赛集团	48.69	-6.69%	-6.39%
300418.SZ	昆仑万维	30.14	-6.51%	-6.21%
603533.SH	掌阅科技	17.54	-6.05%	-5.75%
002415.SZ	海康威视	29.07	-5.95%	-5.65%

资料来源: Wind,中国银河证券研究院

(三) 板块估值

人工智能指数(884201.WI)重要成分股过去三年整体营业收入复合增长率 13.33%, 净利润复合增长率 -8.33%, 截至 7 月 31 日平均估值 PE(TTM) 62.33 倍, PS(TTM) 3.09 倍。

图 2: 7 月人工智能板块市场表现



资料来源: wind,中国银河证券研究院

表 3: 人工智能主题基金一览

基金代码	基金简称(官方)	基金类型	基金规模(亿元)	7月31日收盘价 (元)	近1月回报 (%)	近3月回报 (%)	近6月回报 (%)
001986	前海开源人工智能	契约型开放式	6.64	1.24	-3.29	-3.21	10.47
005729	南方人工智能主题	契约型开放式	3.73	1.85	-0.59	-3.83	13.80
005844	东方人工智能主题 A	契约型开放式	5.01	0.82	1.62	-3.10	12.00
005962	宝盈人工智能 A	契约型开放式	4.76	1.98	-2.22	-5.93	21.52
005963	宝盈人工智能 C	契约型开放式	1.99	1.89	-2.29	-6.13	21.03
006281	万家人工智能 A	契约型开放式	13.92	2.03	-3.52	-5.48	31.42
008020	华富中证人工智能产业 ETF 联接 A	契约型开放式	1.90	0.65	-4.41	-10.57	14.76
008021	华富中证人工智能产业 ETF 联接 C	契约型开放式	1.44	0.64	-4.42	-10.63	14.61

请务必阅读正文最后的中国银河证券股份有限公司免责声明。

008585	华夏中证人工智能主题ETF联接A	契约型开放式	5.31	0.66	-3.93	-9.04	14.40
008586	华夏中证人工智能主题ETF联接C	契约型开放式	5.18	0.65	-3.97	-9.12	14.20
009239	融通中证人工智能主题C	契约型开放式	1.07	1.10	-4.03	-9.03	14.18
011832	西部利得中证人工智能A	契约型开放式	1.25	0.70	-3.61	-8.97	12.36
011833	西部利得中证人工智能C	契约型开放式	0.88	0.69	-3.64	-9.07	12.13
011839	天弘中证人工智能主题A	契约型开放式	1.63	0.72	-3.94	-8.90	14.56
011840	天弘中证人工智能主题C	契约型开放式	5.92	0.72	-3.96	-8.95	14.44
012733	易方达中证人工智能主题ETF联接A	契约型开放式	4.03	0.82	-3.94	-8.97	14.38
012734	易方达中证人工智能主题ETF联接C	契约型开放式	5.45	0.82	-3.94	-9.00	14.32
014162	万家人工智能C	契约型开放式	16.25	1.99	-3.58	-5.67	30.90
014630	汇添富中证人工智能主题联接A	契约型开放式	0.06	1.02			
014631	汇添富中证人工智能主题联接C	契约型开放式	0.06	1.02			
017811	东方人工智能主题C	契约型开放式	21.16	0.82	1.59	-3.19	11.79
159702	汇添富中证人工智能ETF	契约型开放式	0.13	0.79			
159819	易方达中证人工智能ETF	契约型开放式	55.86	0.69	-4.16	-9.50	15.12
161631	融通中证人工智能主题A	契约型开放式	5.18	1.12	-4.00	-8.93	14.40
512930	平安中证人工智能ETF	契约型开放式	7.99	0.99	-4.12	-9.57	14.51
515070	华夏中证人工智能ETF	契约型开放式	22.15	0.88	-4.15	-9.55	14.98
515980	华富中证人工智能产业ETF	契约型开放式	15.36	0.73	-4.59	-11.29	15.29
517800	方正富邦中证沪港深人工智能50ETF	契约型开放式	1.00	0.53	-4.71	-8.47	17.58

资料来源: Wind, 中国银河证券研究院

表 4: 国内人工智能上市公司近况一览 (截至 2024.7.31)

证券代码	证券简称	2023年 营收增速 (%)	2023年 净利润增速 (%)	2024Q1 营收增速 (%)	2024Q1 净利润增速 (%)	总市值 (亿 元)	市盈率 PE (TTM)	市销率 PS (TTM)	月涨跌幅 (%)	今年以来 涨跌幅 (%)
000977.SZ	浪潮信息	-5.41	-15.65	85.32	73.88	548.52	29.19	0.74	2.45	12.62
002230.SZ	科大讯飞	4.41	45.18	26.27	-265.24	918.22	221.40	4.50	-7.29	-14.15
002236.SZ	大华股份	5.41	257.05	2.75	18.92	494.17	6.65	1.53	-2.98	-16.93
002362.SZ	汉王科技	3.56	4.83	10.08	48.89	39.55	-	2.67	-2.82	-36.85
002405.SZ	四维图新	-6.72	-183.61	0.04	-54.03	167.63	-	5.37	13.34	-20.79
002415.SZ	海康威视	7.42	8.50	9.98	14.06	2,712.41	19.09	2.98	-5.95	-13.99
300229.SZ	拓尔思	-13.84	-74.56	-15.35	-15.14	103.87	309.81	13.86	-2.76	-22.37
300474.SZ	景嘉微	-38.19	-103.82	66.27	83.68	302.28	254.40	39.96	4.24	-6.63
601360.SH	三六零	-4.89	77.02	-12.53	39.98	534.47	-	6.07	-2.60	-15.93
603019.SH	中科曙光	10.34	14.21	7.86	-14.07	642.37	34.77	4.42	5.78	11.64
688088.SH	虹软科技	26.07	65.91	13.94	21.43	109.32	115.72	15.78	-5.58	-33.11
688169.SH	石头科技	30.55	72.36	58.69	95.22	414.00	18.44	4.44	-19.86	11.20
688207.SH	格灵深瞳	-25.84	-523.61	-54.51	-2,893.67	30.92	-	13.71	4.10	-43.03
688256.SH	寒武纪-U	-2.70	33.86	-65.91	12.77	1,102.34	-	167.08	32.91	95.66
688787.SH	海天瑞声	-35.33	-299.14	41.04	95.34	31.44	-	17.29	-5.05	-27.73
688793.SH	倍轻松	42.30	71.77	28.59	956.13	28.83	-	2.15	15.66	-0.33
002410.SZ	广联达	-0.42	-84.53	-1.08	-88.62	181.14	304.47	2.82	13.57	-36.18
688327.SH	云从科技-UW	19.33	28.51	19.99	-9.27	111.26	-	17.46	-5.13	-36.36
688343.SH	云天励飞-U	-7.36	14.01	78.57	-52.69	104.59	-	19.04	3.15	-40.20
688246.SH	嘉和美康	-3.04	-75.34	14.00	48.96	28.13	52.77	3.96	-0.29	-39.26

请务必阅读正文最后的中国银河证券股份有限公司免责声明。

603893.SH	瑞芯微	5.17	-67.81	64.89	468.09	258.89	117.19	11.03	4.63	-2.35
300033.SZ	同花顺	0.14	-17.88	1.47	-15.04	574.75	41.53	16.08	3.10	-30.72
300496.SZ	中科创达	-3.73	-44.51	1.01	-52.27	208.38	53.63	3.97	-0.46	-43.01
688111.SH	金山办公	17.27	17.05	16.54	37.61	974.26	68.73	20.60	-7.41	-33.14
688475.SH	萤石网络	12.39	72.78	14.61	37.38	227.35	38.08	4.55	-15.36	-9.13
300634.SZ	彩讯股份	25.18	44.92	10.64	32.23	75.61	21.74	4.92	0.36	-16.69
300624.SZ	万兴科技	25.49	45.99	0.90	10.23	93.92	103.21	6.33	-8.46	-28.10
301162.SZ	国能日新	26.89	27.15	25.75	20.57	38.18	44.29	8.00	-12.49	-25.77
688188.SH	柏楚电子	56.61	49.56	40.31	47.24	300.97	38.08	19.85	-20.62	-18.29

资料来源: Wind,中国银河证券研究院

表 5: 境外上市人工智能企业近况一览 (截至 2024.7.31)

证券代码	证券简称	今年中报营业收入 (亿元)	今年中报营业收入同比增长率 (%)	今年中报净利润 (亿元)	净利润近 1 年增长率 (%)	总市值 (亿美元)	市盈率 PE (LYR)	市销率 PS (LYR)	月涨跌幅 (%)	今年以来涨跌幅 (%)
TSLA.O	特斯拉	468.01	-3.02	26.07	19.44	7111.91	47.42	7.35	12.50	-10.41
QCOM.O	高通	193.25	3.13	50.93	-44.09	1863.05	25.76	5.20	-16.19	16.49
NVDA.O	英伟达	206.99	38.07	82.32	205.17	25517.58	85.74	41.89	-16.04	109.49
MSFT.O	微软	1185.37	15.23	441.61	21.17	31435.81	35.67	12.82	-5.38	12.88
META.O	脸书				68.53	11748.97	30.05	8.71	-8.14	31.13
GOOGL.O	谷歌	1652.81	14.47	472.81	23.05	20962.70	28.41	6.82	-6.51	22.05
BIDU.O	百度				168.75	311.17	10.85	1.64	2.61	-25.48
AAPL.O	苹果	2103.28	-0.78	575.52	-2.81	33550.97	34.59	8.75	3.88	13.94
9988.HK	阿里巴巴	4589.46	11.20	620.89	28.53	14954.06	16.94	1.44	9.65	4.46
2158.HK	医渡科技	3.56	-24.85	-0.76	74.43	38.00	-17.67	4.27	-6.28	-29.25
0700.HK	腾讯控股				-38.79	33760.56	26.55	5.02	-2.74	24.43
0020.HK	商汤-W				-6.54	425.22	-5.98	11.31	-8.33	4.31

资料来源: wind,中国银河证券研究院

二、人工智能产业政策动态

(一) 数据要素、数据交易所最新新闻及政策

表 6: 数据要素最新新闻及政策

日期	具体内容
7.23	<p>河南省出台“数据要素×”行动实施方案</p> <p>河南省政府印发了《扎实推进 2024 年下半年经济稳进向好若干措施》，其中明确提出要做大做强数字经济，并配套出台了一系列具体措施，旨在通过数据要素的市场化运作和算力基础设施的建设，推动全省数字经济高质量发展。为了充分发挥数据要素在经济发展中的关键作用，河南省出台了“数据要素×”行动实施方案，并制定了数据要素市场培育行动方案。这些方案旨在通过遴选 30 个以上数据要素开发利用的解决方案，推动数据资源的有效整合与利用。同时，河南省还大力支持郑州创建国家级全域数字化转型试点城市，通过示范引领，带动全省数字化转型步伐的加快。在平台经济领域，河南省同样不遗余力地推动高质量发展。为此，该省出台了支持平台经济高质量发展的若干措施，并建立了平台经济创新发展联席会议制度，以加强跨部门协作，共同推进平台经济的规范健康发展。此外，河南省还制定了省重点培育企业和重点引进企业的两个清单，通过政策扶持和资金引导，培育一批具有核心竞争力的平台经济企业。在算力基础设施建设方面，河南省加大了投入力度。研究制定了全省算力基础设施建设规划，明确了未来一段时期</p>

	<p>内算力基础设施建设的目标和任务。其中，国家超算互联网核心节点工程和航空港区智算中心的建设被列为重点项目，旨在提升全省的算力供给能力和服务水平。同时，河南省还启动了省级数据标注基地建设试点，为数据资源的标注、分类和共享提供有力支撑。在基础设施建设方面，河南省全年计划新建 5G 基站 2.9 万个，数据中心标准机架达到 20 万架。</p>
7.22	<p>国新办举行“推动高质量发展”系列主题新闻发布会</p> <p>7月22日下午，国新办举行“推动高质量发展”系列主题新闻发布会，国家数据局局长刘烈宏表示“深入贯彻落实党的二十届三中全会精神，我们将坚持以数据要素市场化配置改革为主线，完善数据要素市场制度和规则，培育全国一体化数据市场，促进数据要素开发利用。”国家数据局将抓好数据产权、流通交易、收益分配、安全治理等政策制定，加快构建适应数据要素特征、符合市场规律、契合发展需要的基础制度。国家数据局数字科技和基础设施建设司司长杜巍介绍，下一步，国家数据局将会同有关部门和地方，加快推动全国一体化算力网建设，包括推动多源异构算力协同发展，促进各类新增算力向国家枢纽节点集聚；推动国家枢纽节点和需求地之间 400G/800G 高带宽全光连接，引导电信运营商提升“公共传输通道”效能，推进算网深度融合；推动行业数据和算力协同，实现数据可信流通，提升数据处理能力和治理水平；推动算力电力融合发展，强化枢纽节点与非枢纽节点的协同联动，支持绿电资源丰富的中西部非枢纽节点融入全国一体化算力网建设；推动提升全国一体化算力网安全防护能力，促进高质量发展和高水平安全良性互动。</p>
7.20	<p>引育数商企业 300 家以上 济南市将打造数据要素枢纽型城市</p> <p>济南市政府办公厅印发《济南市推动数据要素市场化配置改革加快数字经济发展行动方案（2024-2025 年）》，济南市将打造数据要素枢纽型城市。鼓励和支持企业探索开展数据产品加工生产与交易、数据基础设施建设、大型数据集建设、行业模型开发等业务，2024 年引育数商企业 300 家以上，2025 年新增数商企业超过 500 家。济南市将建立济南市数据要素流通服务中心和数据要素流通服务平台，贯通数据要素服务链条，构建数据资源确权、登记、合规、入表、评估、交易等全方位服务生态，推动场内、场外数据交易协同发展。2024 年服务企业 500 家以上，打造数据产品 300 个以上，面向企业全面开展数据入表指导服务，完成 30 家左右重点企业的数据入表工作，2025 年实现济南市数据流通交易体系全面定型，打造数据要素枢纽型城市。2024 年年底，智能算力规模达到 3000PFlops，占比达到全市总算力的 60%；2025 年年底，智能算力规模达到 5000PFlops，占比达到全市总算力的 70%。接入全国头部云服务厂商 5 家以上，服务具有全国影响力的通用及行业大模型企业 20 家以上，形成“智算+云服务+AI”的算力服务生态。</p>
7.20	<p>中国联通发布数据要素开放合作计划 携手多方共绘数字未来新篇章</p> <p>2024 中国联通合作伙伴大会“人工智能赋能数据要素论坛”分论坛在上海市世博中心成功举办。论坛上隆重举行了中国联通数据要素开放合作计划启动仪式与可信数据资源空间联合运营发布仪式，标志着中国联通在推动数据要素市场化配置、加速数字经济高质量发展方面迈出了坚实的一步，以联通数据生态为核心的合作计划正式开启。公安部三所、工信部电子五所、湖北省数字产业发展集团、上海数交所、广州数交所、西部数据交易有限公司、山东数据交易有限公司、上海数商协会、腾讯云、上海理工大学、广东数字政府研究院、等政府、企业、科研院校参加中国联通数据要素开放合作计划启动仪式。广州数据交易所、杭州临平大数据经济有限公司、安徽大数据产业发展有限公司及云从科技、亚信科技等共同参加了可信数据资源空间联合运营发布仪式。</p>
7.18	<p>杭州数据要素运营中心启动 打造数据可信操作空间</p> <p>7月18日，杭州数据要素运营中心(中国数谷·临平)(下称“杭州数据要素运营中心”)启动仪式在杭州市临平区举办。此为浙江省内目前唯一落地的、具备多元数据开发利用能力的可信操作空间。杭州数据要素运营中心汇聚信息化项目研发、安全运营、模型训练、数据加工、接洽展示等五大核心模块，逐步形成集“承接、运营、招商、展示、示范”五大功能于一体的新型杭州数据要素运营体系。</p>
7.17	<p>南京市江宁区成立全市首个数据要素型企业集聚产业联盟</p> <p>南京市江宁区数据产业生态联盟成立大会暨联盟企业座谈会召开，江宁区数据产业生态联盟揭牌成立。该联盟为南京市首个数据要素型企业集聚的产业联盟，成员单位达 40 余家。</p> <p>数据产业是当今数字经济的核心，是数字化转型和智能化升级的重要支撑。近年来，江宁区坚持科技创新</p>

	<p>引领，推进国家、省级和市级数据要素重点项目落地，吸引一批权威的数据要素型企业集聚。统计显示，江宁区目前已集聚数字经济核心产业规上重点企业 454 家、国家重点实验室 7 家，全区人才资源总量超 40 万人。</p>
7.12	<p>四部门：加快推动数字贸易创新发展，适应贸易数字化发展需要，探索将更多数据要素纳入承保评审服务</p> <p>《意见》提出，大力发展服务贸易和数字贸易。指导地方商务主管部门加快推动服务贸易、数字贸易创新发展，加快旅行、运输等重点领域外向型发展和服务外包转型升级，支持进口研发设计、节能环保等优质服务。充分发挥服务贸易创新发展引导基金作用，进一步带动社会资本投资。适应贸易数字化发展需要，积极开展知识产权、商标权、科技成果等无形资产质押融资服务，探索将更多数据要素纳入承保评审服务。依托国家服务贸易和数字贸易创新平台建设，推动银行保险机构与平台内企业高效对接。</p>
7.10	<p>广州：拟立法推动数据要素纳入国民经济和社会发展的统计核算体系</p> <p>7月10日，《广州市数据条例（草案二次审议稿·征求意见稿）》（简称《条例》）正在广州市人大常委会官网公开征求意见，拟于2024年7月进行第二次审议。《条例》明确，建立数据交易场所、数据商及第三方专业服务机构等多方参与的数据要素市场，规范引导数据流通交易，推动数据要素高效、有序流通。</p> <p>《条例》还提出，广州市统计部门应当推动数据要素纳入国民经济和社会发展的统计核算体系。广州市人民政府及统计等有关部门应当探索建立数据要素配置的统计核算指标体系和评估评价指南，评价各行政区、功能区、行业领域内数据对经济社会发展的贡献。《条例》提出，促进数据和实体经济深度融合，鼓励各类市场主体加快制造业、商贸业、服务业等产业数字化转型，实现以数据要素创新驱动数字经济高质量发展。同时，促进数据和公共服务深度融合，提高公共卫生、医疗、教育、养老、就业、文娱、体育、旅游等民生领域的数字化水平。《条例》还提出，南沙深化数据开发合作。南沙区人民政府应当完善数据基础设施，推动南沙（粤港澳）数据服务试验区建设，加快数据要素市场培育，加强与港澳数据交流合作，促进数据高效、有序流动。同时，培育算力、数据算法、数据加工、数据服务等数据类核心产业，打造数据产业集群，建设国际数据合作产业集聚发展区。</p>
7.6	<p>激活数据要素 X 筑基新质生产力——2024WAIC 产业数据要素流通与应用论坛成功举办</p> <p>7月6日，主题为“激活数据要素 X 筑基新质生产力”—2024WAIC 产业数据要素流通与应用论坛于上海世博中心 430 会议室成功举办。本次论坛由世界人工智能大会指导，中国发展战略学研究会数字经济战略专委会主办，中国发展战略学研究会数字经济战略专委会、上海零数科技有限公司联合承办。作为 2024 WAIC 重要分论坛之一，论坛邀请了上海市人民政府办公厅、上海市数据局、上海市城市运行管理中心等相关省市领导，以及中国工程院院士邬贺铨，中国发展战略学研究会副理事长李淑敏，中国电子信息产业集团有限公司副总经理陆志鹏，国家信息中心信息化和产业发展部主任单志广，中国科学院大学经济与管理学院教授孙毅，广电运通集团副总经理关键伟，赛迪网总裁庄健等专家代表齐聚一堂，深入探讨数据要素市场建设思路，共商数据要素产业发展未来。</p>

资料来源：Wind、中国经营报、央广网、新华日报、中国日报网、新华网、光明网、新京报，中国银河证券研究院整理

表 7：数据交易所最新新闻及政策

日期	具体内容
7.23	<p>上海数据交易所上线公共数据专区</p> <p>为提高公共数据开发利用效率，汇聚高质量供给，更好地赋能产业行业应用。上海数据交易所方面介绍，该所目前已经甄选 300 多个涉及公共数据的数据产品集中展示在公共数据专区，推动相关政府机关及企业开展公共数据资产管理，引入可承接市场拓展和技术支撑的服务商，构建公共数据开发利用的市场化运营机制，为数据产品流通交易保驾护航。</p>
7.18	<p>上海数据交易所碳板块正式发布上线，聚焦数据要素与绿色低碳重点场景</p> <p>上海数据交易所碳板块正式发布上线，聚焦绿色低碳领域的数据流通交易，汇聚碳排放数据，重点拓展绿色金融和碳排放管理等核心场景，打通绿色低碳领域的数据流通交易链路，为推动我国绿色化、低碳化、智能化转型提供有力支撑。截至今年 7 月 15 日，全国碳市场碳配额累计成交量已达 4.65 亿吨，累计成交额近 270 亿元。随着全国碳排放权交易市场的不断成熟，碳数据的重要性和价值日益凸显。上海数据交易</p>

	所此次推出的碳板块，正是基于这一背景，旨在汇聚碳排放数据，拓展绿色金融和碳排放管理等核心场景，为相关企业和机构提供更加高效、便捷的数据服务。
7.10	<p>河南：出台“数据要素×”行动实施方案 研究制定全省算力基础设施建设规划</p> <p>7月10日晚，上海数据交易所与欧洲知名金融数据交易平台 Neudata 宣布正式启动“全球数商招募计划”，并同步在线上举办首场全球数商企业沟通会。本次沟通会以“国际数据供应商在上海数据交易所的机会”为主题，吸引了来自全球各地 100 多家涉及人工智能、大宗商品、媒体与情报分析、投资管理、金融服务等诸多行业及领域的企业参与，共同探讨中国数据要素市场的广阔前景与合作机遇。Neudata 方面表示，依托与上海数交所的战略合作，双方将持续推动全球企业开展数据跨境流动业务，进一步促进全球数据流通交易。</p>

资料来源：Wind、中国新闻网、贵阳发布、新华网、光明网、新京报，中国银河证券研究院整理

（二）算法端：国内外巨头大模型动态

表 8：国内人工智能大模型动态

时间	模型	主要内容
2024.7.30	生数科技 Vidu	7月30日，生数科技自研视频大模型 Vidu 全球上线。Vidu 四月底首次亮相，此次上线开放了文生视频、图生视频两大核心功能，提供 4s 和 8s 两种时长选择，分辨率最高达 1080P。效果方面，Vidu 在延续高动态性、高逼真度、高一致性等基础上，新增了角色一致性、动漫风格、文字与特效画面生成等能力。目前，Vidu 实测推理速度为生成一段 4 秒片段需要 30 秒。
2024.7.28	“谛听”大模型	7月28日，“谛听”地震波大模型在四川成都发布，该大模型由国家超级计算成都中心、中国地震局地球物理研究所以及清华大学联合开发，是首个亿级参数量的地震波大模型。成都超算中心运营管理有限公司董事长郭黎介绍，2023年9月，国家超级计算成都中心与中国地震局地球物理研究所合作共建了“地震大模型创新应用联合实验室”，与清华大学、中国科学院地质与地球物理研究所合作启动了“谛听”地震波大模型的训练。“谛听”数据集是国内首个，也是目前国内外最大规模、样本类型和标注最为全面的地震学专业 AI 训练数据集之一，半年多时间里，实验室研究人员利用“谛听”数据集获得了首个亿级参数量的地震波大模型。
2024.7.18	讯飞星火 Spark Pro-128K	讯飞星火推出了 Spark Pro-128K 大模型，具备超长文本处理能力，能够显著提升 AI 大模型在实际应用中的价值。长文本处理能力是 AI 大模型发展的重要方向，对于金融、医疗、科研等行业尤为重要，能够快速理解和生成大量信息，提高工作效率。Spark Pro-128K 大模型的 API 调用已经面向开发者开放，价格低至 0.21 元/万 tokens。该模型不仅能处理超长文本，还具备内容溯源功能，当提供问题时，模型不仅能给出答案，还能指出答案的参考来源，提高回答的可信度。这一功能支持中英文，并且模型能够直接理解英文内容，无需翻译。此外，讯飞星火还提供了丰富的大模型 API 矩阵，包括 Ultra、Lite、Pro、Max 等不同版本的通用大模型，以及智能 PPT 生成、角色扮演、图片生成、图片理解等定制化 API，满足不同企业的个性化业务需求。开发者可以通过实名认证免费领取 Spark Pro-128k 版本的 500 万 tokens，以体验和利用这一大模型的能力。
2024.7.12	粉笔头	7月12日，粉笔正式推出其自主研发的首个专注于职教行业的垂域大模型，并将于8月1日上线粉笔 AI 老师“粉笔头”，为学员提供个性化、智能化的辅导服务。在垂域大模型的基础上，粉笔结合 10 年来的教研积累、独有数据搭建 RAG 系统，采用分治法，加强大模型对问题的理解，将复杂的问题进行拆分，赋予 AI 老师几乎媲美真人的答疑对话能力，协助老师更好的服务学员。
2024.7.12	上体体育大模型	7月12日，国内首个体育大模型——上体体育大模型近日发布。大模型由上海体育大学与百度合作成立的上体—百度飞桨智慧体育技术创新中心研发，具有数据专、算法强、算力优、

		应用广等特点。上体育大模型包含体育文献、动作识别与技战术分析、多模态三个垂直大模型。其中，体育文献大模型通过学习国内外体育文献资料，能对体育问题进行专业系统解答；动作识别与技战术分析大模型能自动解析体育训练的视频与图像，有效输出人体姿态的解析结果及距离、速度、高度、角速度等量化指标，有助于深入分析生物力学；多模态大模型能有效支持学科交叉融合研究，其数据分析结果可用于比赛视频 AI 解说及个性化课程生成。据悉，大模型研发团队正在服务跳水、游泳、田径、体操、蹦床、攀岩等多支国家队的日常训练和巴黎奥运会备战工作，通过在训练场馆部署人工智能智算设备和应用系统，助力运动员提升训练质量，提高在国际赛场的竞争力。
2024.7.5	PathOrchestra	7月5日，在2024世界人工智能大会期间，商汤医疗联合中华医学会病理学分会王哲副主委团队、清华大学何永红教授团队共同发布国内首个病理大模型 PathOrchestra。该模型基于国内规模最大的数字病理图像数据集训练，并实现全球最广泛的临床任务赋能，完成了病理 AI 领域从“单模专病”到“一模多病”的跨越式突破。商汤医疗已将该成果深度嵌入业内领先的数智化病理科整体解决方案，助力医院“一站式”建设大模型驱动的数字智慧病理科，开启病理诊断“新纪元”。

资料来源：Wind、证券时报网、太平洋科技、CSDN，中国银河证券研究院整理

表 9：国外人工智能大模型动态

时间	企业名称	主要内容
2024.7.16	微软	<p>微软计划为 Excel、谷歌 Sheets 等表格应用程序，开发全新的 AI 语言大模型——SpreadsheetLLM。</p> <p>该模型旨在理解和处理复杂的电子表格，它可以理解复杂表格中数据的结构化性质以及表格中的引用和公式，并提供基于电子表格数据的智能见解和建议，还可以自动化处理数据清理、格式化等繁琐耗时的任务。微软将该模型的研究论文发表在了 arXiv 上。SpreadsheetLLM 在企业中的应用可以帮助员工提高生产效率，解锁新的机会、推动创新，并最终为客户和利益相关者创造更多价值。目前，微软还没有公布何时向公众发布 SpreadsheetLLM 的消息。</p>
2024.7.18	OpenAI	<p>OpenAI 踏入小模型战场，发布 GPT-4o Mini</p> <p>美国当地时间 7 月 18 号，OpenAI 正式发布了 GPT-4o Mini，这是其最新的小型人工智能模型（小模型）。该公司表示，GPT-4o Mini 比 OpenAI 当前的最先进小模型更强大、价格更低。该模型将从即日起向开发者和 ChatGPT 网站及移动应用的用户开放，企业用户将在下周获得访问权限。OpenAI 称，GPT-4o Mini 在涉及文本和视觉的推理任务上优于行业领先的小模型。知名的小模型有 Anthropic 的 Claude 3 Haiku 和谷歌 Gemini Flash。它们很适合处理高频次的简单任务，在实际应用场景中，用户的操作可能经常要反复调用这些模型。对于使用 OpenAI 应用程序编程接口（API，Application Programming Interface）的开发者，它的定价为每百万输入 token 15 美分，每百万输出 token 60 美分。这个价格比今年初发布的 GPT-3.5 Turbo 版本便宜 60% 以上。</p>
2024.7.23	Meta	<p>Llama 3.1: Meta 迄今为止规模最大、质量最高的开源模型</p> <p>当地时间 7 月 23 日，Meta 正式发布 Llama 3.1。在官方博客中，Meta 称该模型是目前“最强大的开源基础模型”，并表示：“直到今天，开源大模型在功能和性能方面大多落后于封闭模型。现在，我们正在迎来一个开源引领的新时代。”作为开源领域的常青树，Llama 系列一直在引领开源大模型的脚步，Meta 也被奉为与 OpenAI 齐名的巨头，分别代表着开源与闭源的技术走向。发布 Llama 3.1 的同时，Meta 创始人、CEO 扎克伯格也发布长文强调了开源的意义，他表示，开源是 AI 积极发展的必要条件，“今天我们正迈出下一步——使开源 AI 成为行业标准。”Llama 3.1 包含 8B、70B 和 405B 三种参数规模，适用于多种场景，包括多语言代理、复杂推理和编码助手等。该系列模型上下文窗口增加到了 128K，扩大 16 倍；支持多种语言，包括英语、德语、法语、意大利语、葡萄牙语、印地语、西班牙语和泰语等；</p>

		提升了工具使用能力，支持搜索和 Wolfram Alpha 的数学推理；拥有更宽松的许可，允许使用模型输出改进其他 LLMs。评测结果显示，Llama 3.1 405B 可与 GPT-4o、Claude 3.5 Sonnet 和 Gemini Ultra 等业界头部模型媲美，这也是 Meta 迄今为止最强大的模型。Meta 指出，最新一代的 Llama 将激发新的应用程序和建模范式，包括利用合成数据生成来提升和训练更小的模型，以及模型蒸馏——这是一种在开源领域从未有过的能力。
2024.7.25	Mistral AI	法国初创 AI 公司 Mistral AI 出新一代旗舰模型 Mistral Large 2。 该模型拥有 1230 亿参数，128k 上下文窗口，可以在单个节点上以高吞吐量运行，支持包括中文在内的数 10 种语言、80 多种编程语言，在多项基准测试中接近 GPT-4o 和 Claude 3.5 Sonnet，预训练版本的多任务语言理解准确率高达 84%，击败 Llama 3.1 70B。除了性能上的提升，Mistral AI 还投入大量成本增强该模型的推理能力，训练的重点放在减少模型产生“幻觉”的概率，通过微调模型，使其在响应时更加谨慎和敏锐，提升模型输出的精准度和可信度。

资料来源：Wind、智通财经、华尔街见闻、IT之家、澎湃新闻，中国银河证券研究院整理

（三）算力端：AI 服务器、AI 芯片最新动态

表 10：最新 AI 服务器、AI 芯片动态

日期	具体内容
7.26	亚马逊研发新 AI 芯片，成本仅为英伟达一半 亚马逊在位于得克萨斯州奥斯汀的一家芯片实验室里，对一款受到严密保护的新型服务器设计进行了测试。亚马逊云业务 AWS 旗下安纳普尔纳实验室（AnnapurnaLabs）的工程总监拉米·辛诺（RamiSinno）表示参观实验室时表示，该服务器配备了亚马逊的人工智能（AI）芯片，与市场领导者英伟达的芯片竞争。亚马逊正在开发自己的处理器，以限制对英伟达昂贵芯片的依赖，该芯片为其亚马逊网络服务的部分人工智能云业务提供动力。AWS 计算和网络副总裁戴维·布朗（DavidBrown）表示：“因此，在某些情况下，性能可提高 40% 甚至 50%，因此成本应该是使用英伟达芯片的一半。”
7.20	Tenstorrent 推出基于 RISC-V 架构的高性能 AI 芯片 7 月 20 日消息，由传奇芯片架构师 JimKeller 领导的 AI 芯片新创公司 Tenstorrent 近日通过官网宣布，推出新一代基于 RISC-V 架构的高性能 AI 芯片 Wormholen150，以及基于该芯片的 PCIe 卡 Wormholen300 和面向软件开发人员的 TT-LoudBox 和 TT-QuietBox 工作站。Tenstorrent 首席执行官 JimKeller 表示：“将更多产品交到开发人员手中总是有益的。使用我们的 Wormhole 卡发布开发系统有助于开发人员扩大规模并开发多芯片 AI 软件。除了这次发布之外，我们很高兴我们的第二代 Blackhole 的流片和通电进展非常顺利。”据介绍，Wormholen150 采用 Tenstorrent 自研的灵活、可扩展的 WormholeTensix 人工智能计算内核，拥有 72 核心和 108MBSRAM，主频为 1GHz，运行功率高达 160W，可以提供 262TeraFLOPs(FP8)的算力。单芯片卡配备了 12GBGDDR6 内存，带宽为 288GB/s。
7.19	OpenAI 与博通携手探索新 AI 芯片研发 7 月 19 日，在人工智能领域持续深耕的 OpenAI，近期传出与博通（Broadcom）等芯片设计商进行深入谈判的消息，旨在共同研发全新的人工智能芯片。这一举措不仅标志着 OpenAI 在硬件自主化道路上迈出了重要一步，也体现了其对于提升全球半导体基础设施的宏伟愿景。据知情人士透露，OpenAI 目前开发的 ChatGPT、GPT-4、DALL-E3 等先进人工智能模型，在运算过程中高度依赖昂贵的图形处理单元（GPU）。为了摆脱对外部供应商的过度依赖，并降低高昂的硬件成本，OpenAI 正积极探索自主制造 AI 芯片的可能性。这一战略转型不仅将增强其技术独立性，还有望在未来的人工智能竞赛中占据更有利的位置。
7.10	三星电子首次涉足 2 纳米 AI 芯片 将与日本 AI 企业合作开发 2024 年 7 月 10 日，记者从三星电子获悉，该公司将与日本人工智能企业合作开发一款基于 2 纳米工艺的前沿人工智能（AI）加速芯片，芯片应用于服务器中以提供算力，满足生成式 AI 对高性能计算的需求，这意味着该公司将开始涉足人工智能计算芯片的开发。日本人工智能企业 Preferred Networks 是三星电子的合

作方，也是这款芯片的使用方。Preferred Networks 总部位于东京，长期开发 AI 的软件和硬件技术，为制造、交通、医疗、娱乐和教育等行业提供 AI 解决方案和产品。三星电子表示，生成式 AI 已经成为全球关注的焦点，然而，现有芯片的计算架构存在一定限制，很难满足训练 AI 技术所需的高速处理和计算能力。公司曾开发 AI 用存储芯片，目前开始在 AI 算力芯片领域做相关的开发。

资料来源：wind、华尔街见闻、IT之家、澎湃新闻，中国银河证券研究院整理

三、前沿行业动态

（一）前沿技术动态

1. PNN 或可从根本上改变 AI 系统高能耗、低吞吐率、高延迟等问题

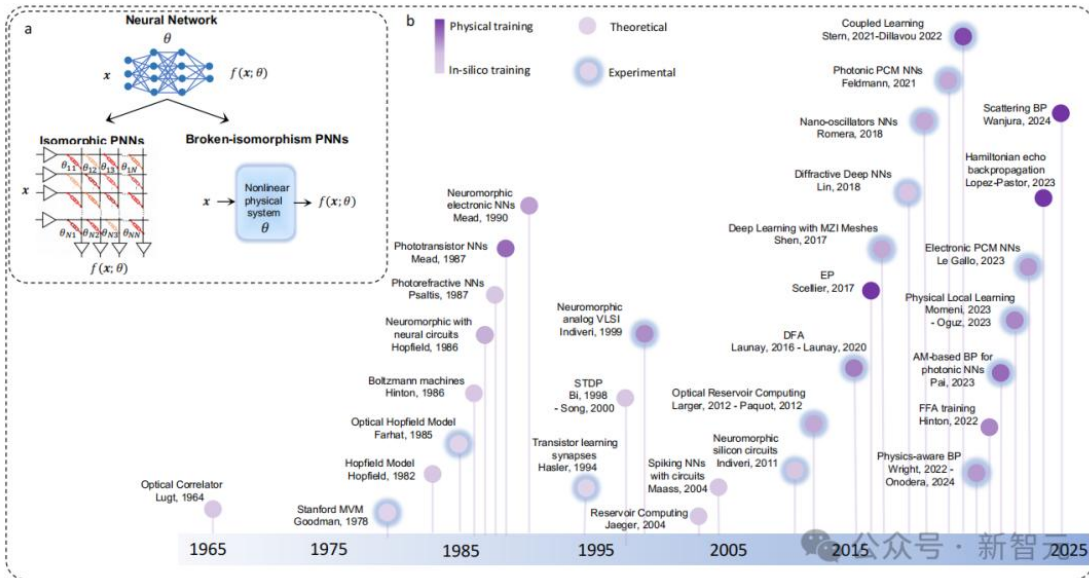
PNN 训练的综述性论文《Training of Physical Neural Networks》提出，PNN 可以从根本上改变 AI 系统的可能性和实用性，实现前所未有的模型规模。论文的合著者云集了众多顶尖机构，包括微软研究院、DeepMind、剑桥、耶鲁、康奈尔、斯坦福、普朗克研究所、EPFL、UCLA 等。

当前典型的 AI 系统存在高能耗、低吞吐率、高延迟等问题，其中的核心矛盾在于内存和处理单元的分离，而且两者之间的数据传输速度较低。由于看到了这些性能限制，研究人员对可替代当前 AI 系统的其他计算平台重新产生了兴趣，例如光学、光子学和模拟电子学。论文将这些非常规计算平台统称为 PNN，指代利用物理系统的属性执行计算的一种类神经网络，与当前深度学习使用的人工神经网络（artificial neural network，ANN）相区分。和 ANN 类似，PNN 系统同样使用可训练权重 θ 处理输入数据 x ，但不同之处在于，系统中至少有一部分是模拟的而非数字的。这意味着部分或者全部的输入/输出数据被连续编码为物理参数，权重也可以是物理参数，以期在性能和效率上超越数字硬件。

从是否模仿数字神经网络的角度，PNN 可以被分为两类：同构（isomorphic）PNN，和破坏同构（broken-isomorphism）PNN。前者会设计出严格的、逐个操作的数学同构的硬件来执行数学变换，比如用于矩阵-向量乘法的忆阻器交叉开关。后者则会彻底打破数学同构的想法，直接训练硬件的物理变换。但这种方法的复杂之处在于，我们不知道通用计算或通用函数逼近需要哪些特征。

相比数字方法，破坏同构 PNN 的计算效率更高，从而为速度更快、更有扩展性、能量效率更高的机器学习方法开辟了道路。

图 3: PNN 训练方法发展时间表

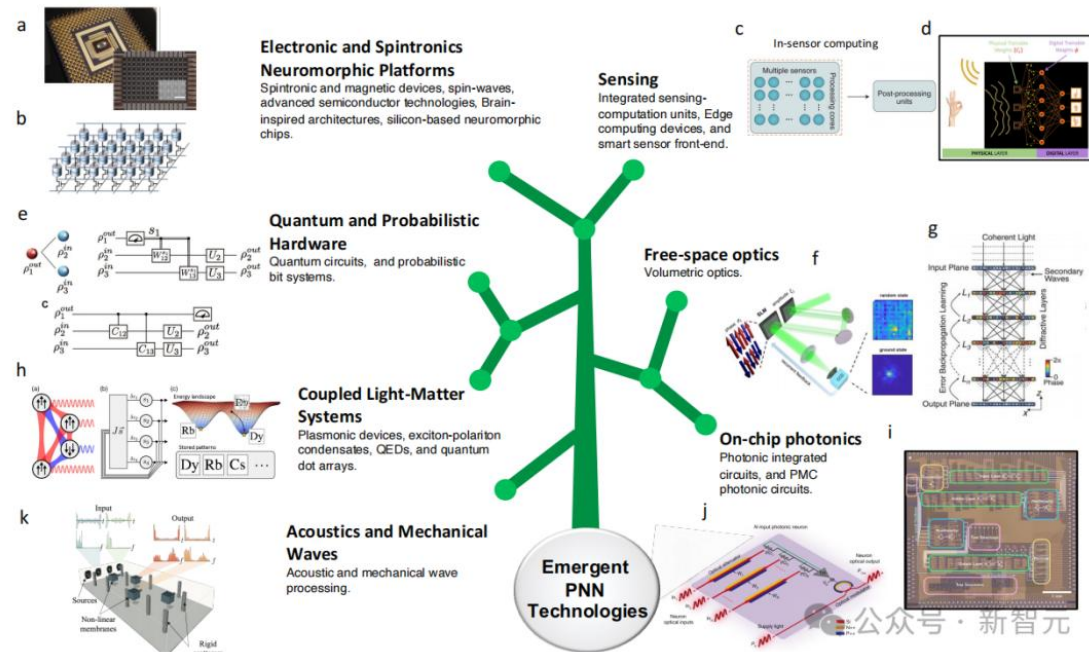


资料来源: 《Training of Physical Neural Networks》(Momeni et Al., 2024), 新智元, 中国银河证券研究院

研究人员发现, 如果 PNN 硬件设计得当, 其不同的底层物理特性可能会使其表现出与数字电子系统不同的能量缩放行为。这意味着, 在模型规模足够大的情况下, PNN 可能实现比数字系统更高的效率。尽管模拟硬件有许多间接成本, 例如数模转换成本。PNN 的光学点积能量缩放优势可能转化为与人工智能模型推理类似的缩放优势, 因为大多数模型主要由点积组成。

在 PNN 方面, 量子计算、概率计算、光子计算、光物质计算和混合计算都是大有可为的发展方向。

图 4: 新兴的 PNN 技术



资料来源: 《Training of Physical Neural Networks》(Momeni et Al., 2024), 新智元, 中国银河证券研究院

2. FlashAttention-3 发布, H100 的 FLOP 利用率提升至 75%

请务必阅读正文最后的中国银河证券股份有限公司免责声明。

对 Transformer 架构来说，注意力机制既是核心优势，也是重要瓶颈。其理论计算量是序列长度的二次方，因此拖慢了计算速度，阻碍了在 LLM 中的长上下文应用。FlashAttention 通过减少内存读写次数，开创了一种在 GPU 上加速注意力机制的方法，现在大多数库都使用它来加速 Transformer 的训练和推理。

23 年 7 月，FlashAttention-2 发布，相比第一代实现了 2 倍的速度提升，比 PyTorch 上的标准注意力操作快 5~9 倍，达到 A100 上理论最大 FLOPS 的 50~73%，实际训练速度可达 225 TFLOPS（模型 FLOPs 利用率为 72%）。FlashAttention（以及 FlashAttention-2）使得大语言模型的上下文长度在过去两年中大幅增加，从 2-4K（如 GPT-3、OPT）扩展到 128K（如 GPT-4），甚至达到 1M（如 Llama 3、Gemini 1.5 Pro）。然而，FlashAttention-2 尚未运用到硬件中的最新功能，在 H100 上仅实现了理论最大 FLOPS 35% 的利用率。

2024/7/12，FlashAttention-3 归来，专攻 H100 GPU，只能在 H100 或 H800 上运行，不支持其他 GPU 型号。FlashAttention-3 的改进将带来以下变化：

更高效的 GPU 利用率：新技术使 H100 GPU 的利用率从之前的 35% 提升到 75%。这使得 LLM 的训练和运行速度显著提高，达到了之前版本的 1.5~2 倍。

更好的低精度性能：FlashAttention-3 在保持准确性的同时，可以使用 FP8 这样的较低精度。这不仅加快了处理速度，还能减少内存使用，从而为运行大规模 AI 操作的客户节省成本并提高效率。

在 LLMs 中使用更长上下文的能力：通过加速注意力机制，FlashAttention-3 使 AI 模型能够更高效地处理更长的文本。这意味着应用程序可以理解 and 生成更长、更复杂的内容，而不会影响速度。

图 5：FlashAttention-3 论文以及 GitHub 上用于 Beta 测试的源代码



FlashAttention-3:
Fast and Accurate Attention with Asynchrony and Low-precision

Jay Shah¹, Ganesh Bikshandi¹, Ying Zhang², Vijay Thakkar^{3,4}, Pradeep Ramani³, and Tri Dao^{5,6}

¹Colfax Research ²Meta ³NVIDIA ⁴Georgia Tech ⁵Princeton University ⁶Together AI 新智元
{jayshah,ganesh}@colfax-intl.com, yingz@meta.com, {vithakkar,praman}@nvidia.com, tri@tridao.me

FlashAttention-3 beta release

FlashAttention-3 is optimized for Hopper GPUs (e.g. H100).

Blogpost: <https://tridao.me/blog/2024/flash3/>

Paper: <https://tridao.me/publications/flash3/flash3.pdf>

公众号 · 新智元

注：论文地址：<https://tridao.me/publications/flash3/flash3.pdf>；项目地址：<https://github.com/Dao-AILab/flash-attention>
资料来源：新智元，中国银河证券研究院

3. 谷歌 DeepMind 新算法 JEST 训练提效 13 倍，能耗暴降 10 倍

ChatGPT 一天用掉超 50 万度电，相当于 1.7 万个美国家庭的用电量。国际能源署（IEA）预测，从 2022 年到 2026 年，数据中心的用电量将翻一番。随着 AI 计算需求的膨胀，还需要用水来冷却计算系统。研究称，微软用水量从 2021 年到 22 年飙升了 34%，ChatGPT 每处理 5-50 个提示就会消耗接近半升水。

最近，谷歌 DeepMind 研究团队提出了一种加快 AI 训练的新方法——多模态对比学习与联合示例选择 (JEST)，大大减少了所需的计算资源和时间。JEST 以 13 倍更少的迭代次数，以及 10 倍更少的计算量，超越了最先进的模型。

预训练的参考模型，已经学习了什么样的数据是有“优质的”或“有用的”。然后通过模型，来引导数据选择那些精心筛选过的小型数据集。这一发现揭示了，数据筛选水平可以作为评判 Scaling Law 的一个新维度。

最近的研究表明，在下游任务未知时，静态数据集的过滤会限制模型性能。这篇论文的结果则表明，相比单独选择样本的方法，在线构建 batch 能提高预训练的效率。无论是使用 JEST 参考模型对数据集进行预评分，还是通过可学习性评分来根据模型需求进行动态调整，都可以成为通用基础数据集的更有效率的替代方案。

但该方法也有局限性。虽然 JEST 同时实现了性能增益和训练成本降低，但依旧依赖于小型、精心管理的参考数据集，它指定了未经管理的更大数据集中优先考虑的分布。因此，未来的工作可以探索一种方法，从指定的下游任务中如何推断出参考数据集的组成和分布。

(二) 前沿政策动态

表 11: 相关政策法规

中央/地方	发布机构	时间	文件名称	主要内容
中央	工业和信息化部，中央网络安全和信息化委员会办公室，国家发展和改革委员会，国家标准化管理委员会	2024/7/2	关于印发国家人工智能产业综合标准化体系建设指南(2024版)的通知	提出：到 2026 年，标准与产业科技创新的联动水平持续提升，新制定国家标准和行业标准 50 项以上，引领人工智能产业高质量发展的标准体系加快形成。开展标准宣贯和实施推广的企业超过 1000 家，标准服务企业创新发展的成效更加凸显。参与制定国际标准 20 项以上，促进人工智能产业全球化发展。重点方向为基础共性标准(人工智能术语、参考架构、测试评估、管理、可持续等)、基础支撑标准(基础数据服务、智能芯片、智能传感器、计算设备、算力中心、系统软件、开发框架、软硬件协同等)、关键技术标准(机器学习、知识图谱、大模型、自然语言处理、智能语音、计算机视觉、生物特征识别、人机混合增强智能、智能体、群体智能、跨媒体智能、具身智能等)、智能产品与服务标准(智能机器人、智能运载工具、智能移动终端、数字人、智能服务等)、赋能新型工业化标准、行业应用标准、安全/治理标准等

资料来源：工业和信息化部等政府网站，中国银河证券研究院

其他产学研合作动态：

7 月 4 日，由长三角国创中心发起的“长三角人工智能+产业创新联合体”正式宣布成立。未来，联合体将聚焦人工智能技术在各行业的融合与应用，推动人工智能企业与各行各业龙头企业的交流互动，加速人工智能技术与产业的融合创新与成果转化，共同推动“人工智能+”产业创新发展。

7 月 11 日，国内首个算力互联公共平台发布。平台将对全国范围内的算力资源进行标识注册和测试，通过平台可以查询全国范围内的算力资源和相关算力调度服务，为各行各业提供真实、可信的算力支持，加速推动算力互联互通。

四、前沿企业动态

(一) 前沿产品动态

1. 英伟达联合 Mistral AI 发布了可在本地硬件上高效运行的 to B 端 AI 大语言模型

IT 之家 7 月 19 日消息，英伟达联合法国初创公司 Mistral AI，发布了全新的 Mistral-NeMo AI 大语言模型，拥有 120 亿个参数，上下文窗口(AI 模型一次能够处理的最大 Token 数量)为 12.8 万个 token。Mistral-NeMo AI 大模型主要面向企业环境，让企业不需要使用大量云资源的情况下，实施请务必阅读正文最后的中国银河证券股份有限公司免责声明。

人工智能解决方案。通过提供一种可在本地硬件上高效运行的模式，英伟达和 Mistral AI 正在解决阻碍许多企业广泛采用人工智能的问题，如数据隐私、延迟以及与基于云的解决方案相关的高成本。此举有可能创造公平的竞争环境，让资源有限的小型企业也能利用人工智能功能，而这些功能以前只有拥有大量 IT 预算的大型企业才能使用。该模型以 Apache 2.0 许可发布，允许商业使用，这将加速其在企业环境中的应用。

2. 蚕丛机器人在 2024 WAIC 发布智能双模充电桩（FlashHub）

7月7日，上海蚕丛机器人科技有限公司在2024世界人工智能大会（简称2024 WAIC），隆重发布智能双模充电桩（FlashHub）等创新产品。

作为 FlashX 移动充电场景的补充，FlashHub 拥有无需依赖电容，布置灵活，不受电源位置限制，大幅简化施工难度，可降低高达 90% 的传统充电桩施工布线成本等多种优势，适用于大型停车场或者社区环境下，以极低的成本和迅速的部署满足停车场对于充电设施静态覆盖的要求，实现插枪即走、无感充电的服务。

FlashHub 分为直流快充的 FlashHub F4 和交流慢充的 FlashHub S12 两种型号。

直流快充的 FlashHub F4 输出电压为 200-1000V，工作温度 -20℃ 至 +50℃，系统最大功率 120kW，单枪最大功率 120kW，最多覆盖车位 8 个。

交流慢充的 FlashHub S12 输出电压为 220V，工作温度 -20℃ 至 +50℃，系统最大功率 105kW，单枪最大功率 7kW，最多覆盖车位 24 个。

图 6：蚕丛机器人 FlashHub 示例图



资料来源：中国机器人网，中国银河证券研究院

24 年以来，蚕丛机器人已完成多处海外商业化试运营，下一步按照国家“一带一路”的倡议，在海外能源转型需求大的国家和地区完成部署。在国内一线城市，明年其目标是做成 5000 个一线城市大型停车场的移动充电服务，形成规模部署，达成网络效应。

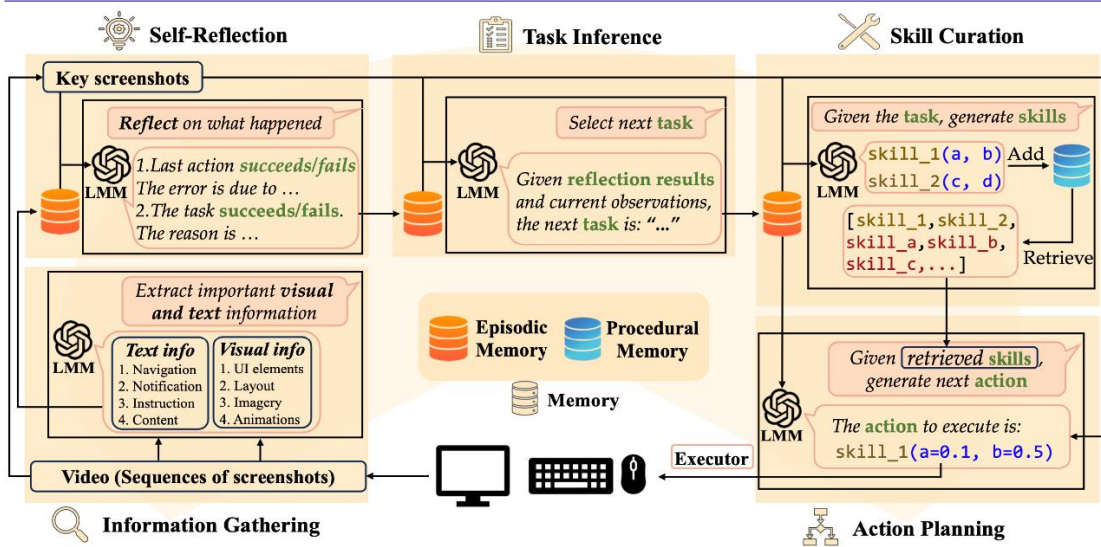
发布会上，蚕丛机器人分别与上海张江智能电力发展有限公司、上海宝华企业集团有限公司、招商蛇口商管公司华东区域公司等签署了战略合作框架协议。

3. 昆仑万维等发布通用 Agent 新框架——Cradle

近日，昆仑万维携手北京智源人工智能研究院、新加坡南洋理工大学、北京大学等顶尖名校机构，联合提出了迄今为止第一个既能玩多种商业游戏又能操作各种软件应用的 AI 框架——Cradle。在这个全新的通用计算机控制框架加持下，AI Agent 无需训练便能像人一样直接控制键盘鼠标，不依赖任何内部 API，实现任意开源软件交互。

Cradle 一共由 6 个模块组成：信息收集、自我反思、任务推断、技能管理、行动规划，以及记忆模块。Cradle 高度的通用性，来源于其对和电脑交互过程中的原始输入输出的合理封装和抽象。以从屏幕中显示的视频图像作为输入，提取其中的文本和视觉信息进行决策，并且输出最底层的操作系统中控制键盘和鼠标的信号去和电脑交互，使其可以不依赖于任何假设和任何内部 API 进行交互。

图 7: Cradle 对电脑交互原始输入输出的合理封装和抽象



资料来源：新智元，中国银河证券研究院

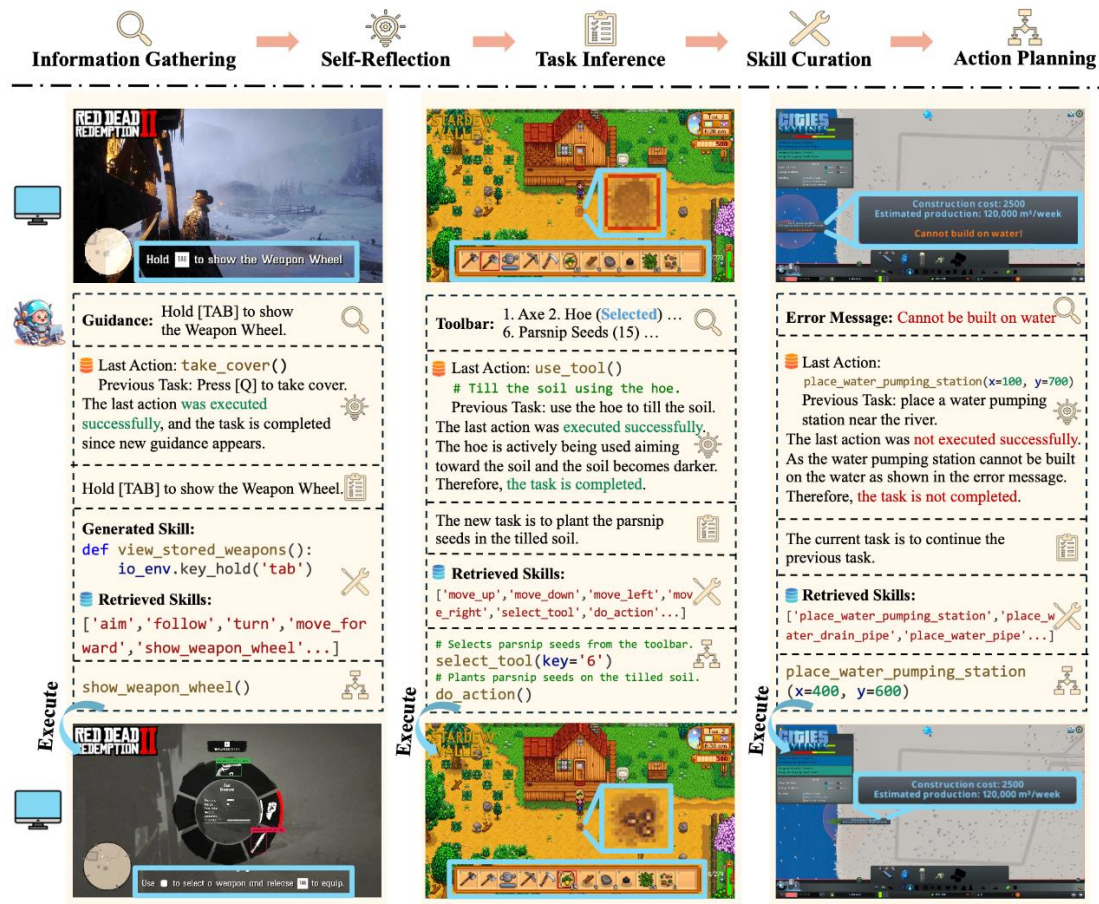
同时，Cradle 强大的决策推理模块让其得以自发和软件进行交互并且完成任务，这个过程可以被简单地总结为：反思过去，总结现在，规划未来。

反思过去： Cradle 使用执行过往动作过程的视频作为输入，分别提取出其中关键的文本和视觉信息，通过反思来判断上一步动作是否执行成功任务是否完成以及如何改进。

总结现在： 反思完之后，Cradle 需要总结当前情况，并且以此为根据来决定是否更换任务目标或是修改任务内容。

规划未来： 最后 Cradle 会根据当前任务以及现状生成或者更新自身的技能，并且从已经学会的技能中检索一部分和当前任务相关的技能作为备选，然后从中选取合适的技能实例化为动作去执行。

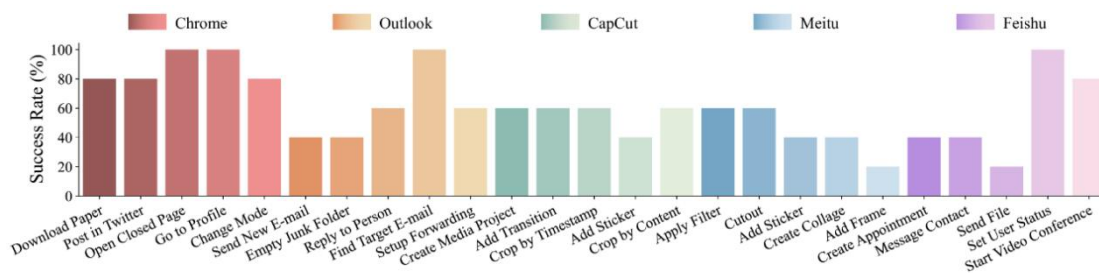
图 8: Cradle 强大的决策推理模块



资料来源: 新智元, 中国银河证券研究院

此外, Cradle 还可实现论文下载、发邮件、修图、剪视频、发送飞书文件等。

图 9: Cradle 在日常软件应用的各项任务执行上表现不俗



注: 论文地址: <https://arxiv.org/abs/2403.03186>; 项目主页: <https://baai-agents.github.io/Cradle>; 代码链接: <https://github.com/BAAI-Agents/Cradle>

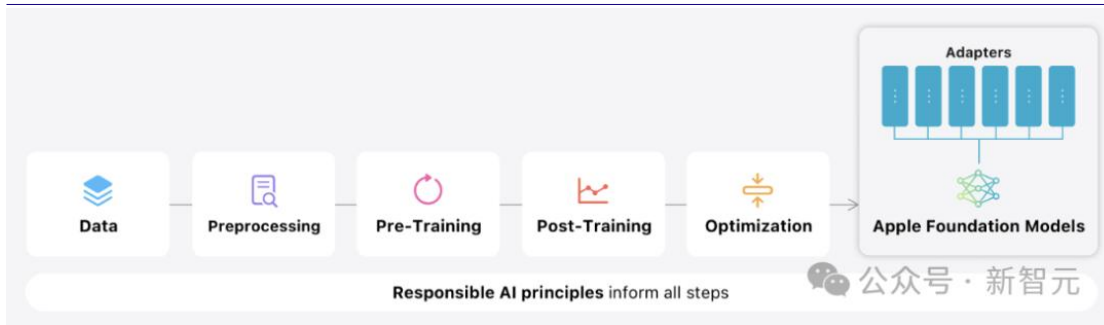
资料来源: 新智元, 中国银河证券研究院

4. OpenAI、Mistral AI & 英伟达、HuggingFace 竞逐小模型, 苹果 DCLM 全开源

2024 年 7 月, OpenAI 发布 GPT-4o mini, Mistral AI 联手英伟达发布 Mistral NeMo, HuggingFace 发布 SmoLLM, 小模型赛道愈发拥挤。无独有偶, 在 WWDC 2024 上, 苹果重新定义了 AI——Apple Intelligence。这是一款深度集成到 iOS 18、iPadOS 18 和 macOS Sequoia 中的个人智能系统。苹果 AI 并非追求越大越好, 而是更强调 AI 模型的定制化。大厂们开始追逐的小模型, 优势在于成本低、速度快、

更专业，通常只使用少量数据训练，为特定任务而设计。大模型变小，再扩大规模，可能是未来发展的趋势之一。Andrej Karpathy 认为我们需要模型先变得更大，再变得更小，因为我们需要“巨兽”将数据重构、塑造为理想的合成形式，逐渐得到“完美的训练集”，再喂给小模型。

图 10: 苹果基础模型概览



资料来源: 新智元, 中国银河证券研究院

此外，作为苹果在 AI 领域的研究成果之一，DCLM(即 DataComp-LM, Data Competition for Language Models)旨在通过高质量的数据集来提高大型语言模型(LLM)的性能。DCLM 项目 6 月在 Hugging Face 上发布了一个开源的 70 亿参数模型，称为 DCLM-7B 1。这个模型的性能已经超越了 Mistral-7B，并正在逼近其他领先的开源模型，包括 Llama 3 和 Gemma 1。

DCLM 项目的核心是来自 Common Crawl 的 240T 新候选词库，通过固定训练代码，鼓励研究人员提出新的训练集来进行创新。研究团队发现基于模型的过滤，即由机器学习模型从较大的数据集中自动过滤和选择高质量数据，可能是构建高质量训练集的关键。

DCLM-7B 模型使用基于 OpenLM 框架的预训练方案，在 MMLU 基准上 5-shot 准确率达到 64%，与 Mistral-7B-v0.3 (63%) 和 Llama 3 8B (66%) 相媲美，并且在 53 个自然语言理解任务上的平均表现也可与 Mistral-7B-v0.3、Llama 3 8B 相媲美，而所需计算量仅为 Llama 3 8B 的 1/6。

值得注意的是，DCLM 项目不仅开源了模型权重，还开源了训练代码和预训练数据集，这使得它被描述为“真正开源的最佳模型”，而 Llama、Gemma、Phi 等一系列“开源”模型都是只放权重、不公布数据。此外，DCLM 项目还发布了 1.4B 版本的模型，并且在 Apache 2.0 许可下发布，允许商业使用、分发和修改。

图 11: 2024 年全球发布的 5 个领先的小模型

5 Leading Small Language Models of 2024



datasciencedojo
— data science for everyone —

资料来源: 新智元, 中国银河证券研究院

请务必阅读正文最后的中国银河证券股份有限公司免责声明。

5.全球首个芯片设计开源大模型 SemiKong，革新半导体工艺和制造技术

2024年7月10日，在 Semicon West 2024 大会上，Aitomatic 发布了首个 SemiKong 半导体行业设计的新模型，将革新半导体工艺和制造技术。SemiKong 是由 Aitomatic 与 FPT Software 合作开发，在处理行业特定任务时，表现优于通用大模型，如 GPT 和 Llama 3。据报道，该模型在计算速度、功耗控制、设计精度等多项关键指标上，均展现出了卓越的性能表现，达到了业界领先水平。

SemiKong 的出现，可以逐渐降低半导体生产成本，未来几年内消费者可能会看到更强大的智能手机、笔记本电脑和智能家居设备以更低的价格上市。Aitomatic 表示，SemiKong 有望在未来五年内，重塑价值 5000 亿美元的半导体行业。下一个更强大的 SemiKong 版本计划在今年 12 月推出，首个针对工艺特定模型预计在 9 月发布。

6.中国农业大学发布“神农大模型 2.0”：覆盖育种、种植、养殖、农业遥感及气象

第八届智慧农业创新发展国际会议于 7 月 13 日在北京举办，中国农业大学在会议上发布了“神农大模型 2.0”。相比 1.0 版本，“神农大模型 2.0”在图像、声音、视频、文件等多模态交互及智能化推理方面获得提升，使大模型能够覆盖育种、种植、养殖、农业遥感及气象等多个农业应用场景。“神农大模型 2.0”含有四个专门化的农业专业大模型：

“神农·固芯”育种大模型：结合多源高维数据推理分析，优化育种流程，减少盲目性与随机性，为种子筛选与适应性评估提供科学依据，保障粮食安全

“神农·筑基”种植大模型：涵盖水肥管控、病虫害识别、作物生长建模等功能，结合物联网技术，实现种植环境的智能化监测与精准调控，提升农业种植效率

“神农·强牧”养殖大模型：提供养殖环境监测、疫病干预及合理喂养方案，推动养殖业向智能化、科学化转型

“神农·问穹”遥感气象大模型：专注于气象与遥感监测，通过精准预警减少自然灾害损失，保障农业生产安全

图 12：神农大模型 2.0 功能升级与场景应用



资料来源：中国农业大学新闻网，中国银河证券研究院

(二) 投融资事件

表 12: 7 月企业投融资事件

融资方	赛道	简介	融资日期	融资轮次	融资金额	本轮投资方
鲸鱼机器人	教育机器人	一家专注青少年人工智能与教育机器人产业的高科技企业，拥有覆盖 3—22 岁全年龄段，适合家庭、学校、培训机构等全场景应用的全球最广的教育机器人产品线，包括移动机器人、积木机器人、飞行机器人、仿生机器人等 4 大系列、400 多款教育机器人，每年向全球 22 个国家与地区出口 50 万套产品。	2024/7/26	B 轮	超 1 亿人民币	上海久事集团、陆家嘴集团、上文投、五星体育传媒等多个国资平台共同设立的上海体育产业基金
百川智能	AI 大模型	百川智能成立于 2023 年 4 月 10 日，由前搜狗公司 CEO 王小川创立。公司核心团队由来自搜狗、百度、华为、微软、字节、腾讯等知名科技公司的 AI 顶尖人才组成。百川智能成立不到 100 天，便发布了 Baichuan-7B、Baichuan-13B 两款开源可免费商用的中文大模型。	2024/7/25	A 轮	50 亿人民币	阿里、小米、腾讯、亚投资本、中金等头部大厂和市场化投资机构，和北京市人工智能产业投资基金、上海人工智能产业投资基金、深创投等国资背景产业投资基金。
橡鸮科技	机器人	成立于 2021 年，是一家服务型机器人研发商，专注于面向酒店、餐饮等生活服务场景，利用机器人自主移动算法、多任务并发调度、云端“智能化”管理等技术，为用户提供酒店自助终端、配送服务机器人等类型产品。	2024/7/22	战略融资	近 2 亿人民币	京东
感图科技	计算机视觉	成立于 2018 年，是一家网络计算机视觉技术及产品研发商，主要业务是将计算机视觉技术应用于精密外观检测场景。目前感图科技主要聚焦于 SMT 半导体和精密机械制造领域，为头部手机零配件供应商和汽车关键安全类铸件制造商提供 AI 表面检测一体机及 AI 检测视觉系统，同时面向部分终端客户提供完整的产品解决方案。	2024/7/19	C2 轮	数亿人民币	博华资本，小橡创投，金华市东富投资合伙企业
元石科技	AI 大模型	成立于 2022 年，旗下产品：问小白，主要进行多模态大模型的研发，基于文本、图像、视频、音频等多种模态数据进行学习训练的人工智能算法。	2024/7/19	Pre-A 轮	3200 万美元	Redpoint Ventures，经纬创投，宿华
蚂蚁工场	AI-工业	成立于 2016 年，是集互联网+、数字敏捷智造、供应链服务于一体的工业互联网平台，旨在面向自动化装备行业，构建海量工业零部件线上交易、智能制造、数字服务的现代制造服务生态体系。	2024/7/18	C++ 轮	超 1 亿人民币	北京国管，亦庄国投
蚌壳智能	AI 大模型	一家基础大语言模型开发商，拥有适合国内企业快速部署的通用大模型、代码大模型、多模态图文大模型、多模态语音大模型等多个基础大语言模型。	2024/7/18	A 轮	1.2 亿人民币	天阳科技
元鼎智能	家用清洁机器人	成立于 2015 年，是一家室外智能清洁产品研发商，专注于泳池清洁机器人、割草机器人等其他室外智能清洁产品的创新研发和全球品牌销售。	2024/7/17	B 轮&B+ 轮	超 6000 万美元	B 轮：招银国际（CMBI）领投，复星锐正、蜂巧资本跟投，老股东 XVC 加码；B+ 轮：由全球头部新能源企业领投，老股东增持
逐际动力	人形机器人	成立于 2022 年，总部位于深圳，主要产品包括全尺寸人形机器人、四轮足机器人、双足机器人等，落地应用于智能制造、工业巡检、物流配送、家庭服务等领域。	2024/7/15	A 轮	近 1 亿人民币	阿里巴巴、招商局创投、上汽集团旗下尚颀资本领投，原始股东峰瑞资本、绿洲资本和明势资本跟投
芯盟科技	AI 芯片、存储芯片	成立于 2018 年，是一家异构集成芯片研发及生产商，公司拥有异构集成技术方案所需的系统方案（SOH™ IP 和设计服务）、芯片配套（VHSM™ 及其他辅助芯片）和集成制造（3D 和 2.5D 异构集成系统硬件制造实现）相关能力，助力并引导客户芯片满足大算力、高带宽、低功耗等场景需求。	2024/7/11	B++ 轮	数十亿人民币	产业方领投，光谷产投、普华资本、谢诺辰途、招银国际等机构跟投
硅基流动	深度学习	成立于 2023 年 8 月，是一家专注于 AI Infra（AI 基础设施）领域的创业公司。硅基流动的产品已在多个行业中广泛应用，并且被业内顶级云厂商、芯片公司和大模型公司相继认证为最高等级合作伙伴。	2024/7/4	天使+ 轮	近 1 亿人民币	由某知名产业方领投，跟投资方包括智谱 AI、360 和水木清华校友基金等企业及机构，老股东耀途资本继续超额跟进，华兴资本担任独家财务顾问

资料来源：来觅研究院，36 氪，腾讯新闻，证券时报，企查查，IT 桔子，中国银河证券研究院

五、投资建议

建议关注 1、国产算力产业链及生态伙伴；2、算力基础设施；3、逐步向国产的自主 AI 算力为底座迁移的 MAAS 及下游应用端；4、AI PC 产业链整机及代工、结构件、开发套件厂商；5、自动驾驶产业链及车路云一体化厂商。建议关注：科大讯飞、海康威视、大华股份、中科曙光、工业富联、网宿科技、海光信息、龙芯中科、华勤技术、润泽科技、曙光数创、中科创达、锐明技术、万集科技、千方科技、金溢科技、税友股份、中国软件、万兴科技、国能日新、神州数码、金山办公、深桑达 A、超图软件、萤石网络、彩讯股份、同花顺、财富趋势、嘉和美康、上海钢联、柏楚电子、中控技术、诺瓦星云等。

请务必阅读正文最后的中国银河证券股份有限公司免责声明。

六、风险提示

技术研发进度不及预期风险；供应链风险；政策推进不及预期风险；消费需求不及预期风险；行业竞争加剧风险。

图表目录

图 1: 6 月人工智能指数表现	2
图 2: 7 月人工智能板块市场表现	3
图 3: PNN 训练方法发展时间表	12
图 4: 新兴的 PNN 技术	12
图 5: FlashAttention-3 论文以及 GitHub 上用于 Beta 测试的源代码	13
图 6: 蚕丛机器人 FlashHub 示例图	15
图 7: Cradle 对电脑交互原始输入输出的合理封装和抽象	16
图 8: Cradle 强大的决策推理模块	17
图 9: Cradle 在日常软件应用的各项任务执行上表现不俗	17
图 10: 苹果基础模型概览	18
图 11: 2024 年全球发布的 5 个领先的小模型	18
图 12: 神农大模型 2.0 功能升级与场景应用	20

表格目录

表 1: 7 月份人工智能板块涨幅前十	2
表 2: 7 月份人工智能板块跌幅前十	2
表 3: 人工智能主题基金一览	3
表 4: 国内人工智能上市公司近况一览 (截至 2024.7.31)	4
表 5: 境外上市人工智能企业近况一览 (截至 2024.7.31)	5
表 6: 数据要素最新新闻及政策	5
表 7: 数据交易所最新新闻及政策	7
表 8: 国内人工智能大模型动态	8
表 9: 国外人工智能大模型动态	9
表 10: 最新 AI 服务器、AI 芯片动态	10
表 11: 相关政策法规	14
表 12: 7 月企业投融资事件	21

分析师承诺及简介

本人承诺以勤勉的执业态度，独立、客观地出具本报告，本报告清晰准确地反映本人的研究观点。本人薪酬的任何部分过去不曾与、现在不与、未来也将不会与本报告的具体推荐或观点直接或间接相关。

吴砚靖 TMT/科创板研究负责人

北京大学软件项目管理硕士，10年证券分析从业经验，历任中银国际证券首席分析师，国内大型知名PE机构研究部执行总经理。具备一二级市场经验，长期专注科技公司研究。

鲁佩 机械行业首席分析师

伦敦政治经济学院经济学硕士，证券从业9年，2021年加入中国银河证券研究院。曾获新财富最佳分析师、IAMAC最受欢迎卖方分析师、万得金牌分析师、中证报最佳分析师、Choice最佳分析师、金翼奖等。

免责声明

本报告由中国银河证券股份有限公司（以下简称银河证券）向其客户提供。银河证券无需因接收人收到本报告而视其为客户。若您并非银河证券客户中的专业投资者，为保证服务质量、控制投资风险、应首先联系银河证券机构销售部门或客户经理，完成投资者适当性匹配，并充分了解该项服务的性质、特点、使用的注意事项以及若不当使用可能带来的风险或损失。

本报告所载的全部内容只提供给客户做参考之用，并不构成对客户投资咨询建议，并非作为买卖、认购证券或其它金融工具的邀请或保证。客户不应单纯依靠本报告而取代自我独立判断。银河证券认为本报告资料来源是可靠的，所载内容及观点客观公正，但不担保其准确性或完整性。本报告所载内容反映的是银河证券在最初发表本报告日期当日的判断，银河证券可发出其它与本报告所载内容不一致或有不同结论的报告，但银河证券没有义务和责任去及时更新本报告涉及的内容并通知客户。银河证券不对因客户使用本报告而导致的损失负任何责任。

本报告可能附带其它网站的地址或超级链接，对于可能涉及的银河证券网站以外的地址或超级链接，银河证券不对其内容负责。链接网站的内容不构成本报告的任何部分，客户需自行承担浏览这些网站的费用或风险。

银河证券在法律允许的情况下可参与、投资或持有本报告涉及的证券或进行证券交易，或向本报告涉及的公司提供或争取提供包括投资银行业务在内的服务或业务支持。银河证券可能与本报告涉及的公司之间存在业务关系，并无需事先或在获得业务关系后通知客户。

银河证券已具备中国证监会批复的证券投资咨询业务资格。除非另有说明，所有本报告的版权属于银河证券。未经银河证券书面授权许可，任何机构或个人不得以任何形式转发、转载、翻版或传播本报告。特提醒公众投资者慎重使用未经授权刊载或者转发的本公司证券研究报告。

本报告版权归银河证券所有并保留最终解释权。

评级标准

评级标准	评级	说明
评级标准为报告发布日后的6到12个月行业指数（或公司股价）相对市场表现，其中：A股市场以沪深300指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准，北交所市场以北证50指数为基准，香港市场以摩根士丹利中国指数为基准。	行业评级	推荐：相对基准指数涨幅10%以上
		中性：相对基准指数涨幅在-5%~10%之间
		回避：相对基准指数跌幅5%以上
	公司评级	推荐：相对基准指数涨幅20%以上
		谨慎推荐：相对基准指数涨幅在5%~20%之间
		中性：相对基准指数涨幅在-5%~5%之间
	回避：相对基准指数跌幅5%以上	

联系

中国银河证券股份有限公司研究院

深圳市福田区金田路3088号中洲大厦20层

上海浦东新区富城路99号震旦大厦31层

北京市丰台区西营街8号院1号楼青海金融大厦

公司网址：www.chinastock.com.cn

机构请致电：

深广地区：程曦 0755-83471683 chengxi_yj@chinastock.com.cn
 苏一耘 0755-83479312 suyiyun_yj@chinastock.com.cn
 上海地区：陆韵如 021-60387901 luyunru_yj@chinastock.com.cn
 李洋洋 021-20252671 liyangyang_yj@chinastock.com.cn
 北京地区：田薇 010-80927721 tianwei@chinastock.com.cn
 褚颖 010-80927755 chuying_yj@chinastock.com.cn