

英伟达业绩超预期，首批 11 家大模型 获批开放

计算机行业

推荐

维持评级

--8月人工智能行业月报

核心观点:

- **8月人工智能板块小幅调整，跑赢多数板块。**人工智能板块行业指数（884201.WI）8月涨跌幅为-1.25%。2023年年初至今上证综指上涨0.99%，沪深300下跌2.75%，创业板指数下跌10.41%，人工智能板块上涨37.36%。8月受到基本面和情绪面的影响，大部分行业出现了不同程度的回调，整体来看人工智能板块跌幅较小。回顾近几月板块连续出现回调，我们认为伴随产业的高速发展及政策的持续催化下，板块调整逐渐结束，有望迎来新一轮上涨周期。
- **数据要素市场催化不断，相关产业落地有望提速。**2023年8月21日，财政部在官方网站发布了“关于印发《企业数据资源相关会计处理暂行规定》的通知”（以下简称《暂行规定》），明确了数据资产的适用范围、会计处理方法、列示及披露原则等。自2024年1月1日起施行，这意味着数据要素明年元旦起将正式计入资产负债表，入表意味着数据完成了从自然资源到经济资产的跨越，作为数字经济时代的第一生产要素，数据有望成为政企报表及财政等收入的重要支撑。
- **英伟达业绩超预期，AI芯片需求持续火爆。**2023年8月23日英伟达交出了营收翻番，净利润激增8倍的超预期答卷，期内实现收入135.1亿美元，同比上涨101%，高于预期的111.9亿美元；净利润61.9亿美元，同比上涨843%。目前，随着人工智能技术在各个领域的广泛应用，AI芯片市场需求持续增长。根据相关数据显示，2023年全球AI芯片市场规模达到了1500亿美元，预计到2028年将达到2500亿美元。我们认为，大模型带来算力升级需求，相关产业将持续受益，有望迎来爆发式增长。
- **11款大模型获批上线，加速应用场景落地。**8月31日消息，据多家媒体报道，我国首批11款大模型产品将陆续通过《生成式人工智能服务管理暂行办法》（以下简称《暂行办法》）备案，可正式上线面向公众提供服务。分别是：百度（文心一言）、智谱AI（智谱清言）、百川智能（百川大模型）、抖音（云雀）、中科院（紫东太初）、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）、科大讯飞（星火大模型）以及华为、腾讯的大模型产品。
- **投资建议：**持续看好整个板块产业趋势，当下时点，继续看好AI+多场景应用，我们看好AI+金融、AI+医疗、AI+教育、AI+法律、AI+新能源、AI+政务（包括媒体类舆情类）、AI+办公等细分赛道的优质公司。重点推荐：同花顺、嘉和美康、彩讯股份、拓尔思、萤石网络、国能日新、上海钢联、深桑达A、中科曙光、中科星图、海康威视、金山办公、柏楚电子、通达海等公司。
- **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。

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

行业数据

2023/8/31



资料来源：wind, 中国银河证券研究院

目录

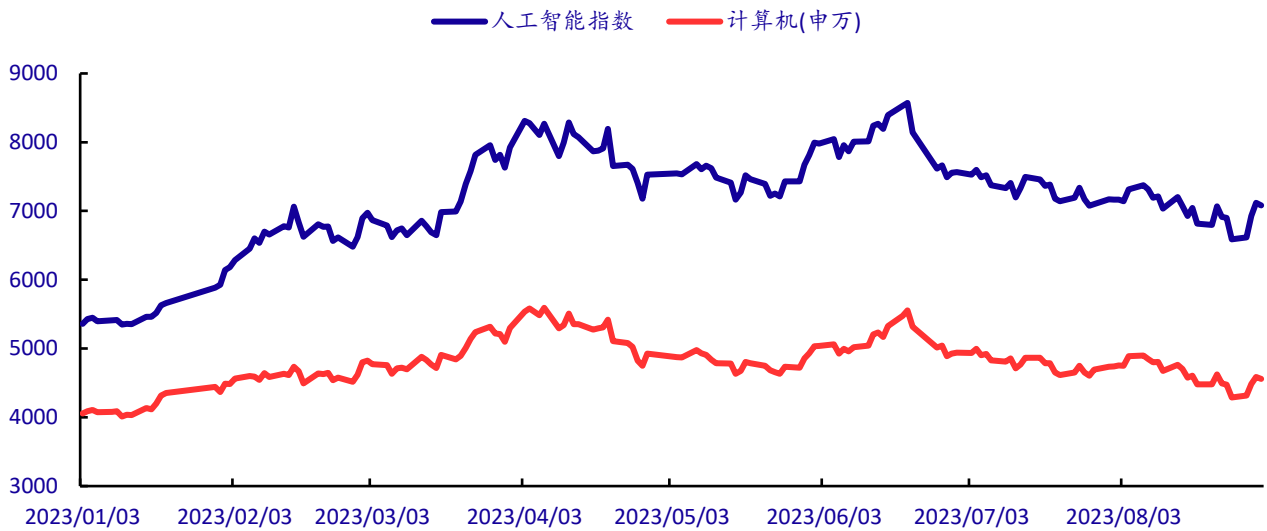
一、市场行情回顾.....	3
(一) 整体行情.....	3
(二) 代表企业.....	3
(三) 板块估值.....	4
二、人工智能产业政策动态.....	6
(一) 数据要素、数据交易所最新新闻及政策.....	6
(二) 算法端：国内外巨头大模型动态.....	8
(三) 算力端：AI 服务器、AI 芯片最新动态.....	10
三、人工智能前沿行业动态.....	11
(一) 前沿技术动态.....	11
1、GLM 技术团队发布长文本理解评测数据集.....	11
2、网易互娱 AI Lab 图像生成论文被 AI 顶会接收.....	12
3、AI 通过自监督训练重建生物样品的微观图像.....	13
4、阿里达摩院发布癌症分析通用模型.....	13
5、DeepMind 发布提高翻译质量的算法 ReST.....	14
6、研究称 ChatGPT 满足一些指标能证明其有意识.....	15
7、SuperCLUE 发布中文大模型 8 月榜单.....	15
(二) 前沿政策动态.....	16
四、前沿企业动态.....	17
(一) 前沿产品动态.....	17
1、多领域机器人广泛应用.....	17
2、全球多家机构致力打造 ChatGPT 式大模型产品.....	20
3、OpenAI 旗下 ChatGPT 功能优化.....	27
4、全球芯片研发取得新进展，与 AI 发展息息相关.....	29
5、无人机应用.....	32
6、京东方和视涯科技打入苹果头显供应链，为“Vision Pro”提供屏幕.....	33
7、讯飞星火 V2.0 突破代码能力，一个指令生成贪吃蛇游戏.....	34
8、亚马逊首将 LLM 落地电商，拟推出 AI 详情页文案创作工具.....	35
(二) 投融资事件.....	36
五、风险提示.....	37

一、市场行情回顾

(一) 整体行情

A股人工智能指数(884201.WI)截至8月末收盘价为7081.79,月涨跌幅为-1.25%。计算机行业指数(801750.SI)截至8月末收盘价为4557.13,月涨跌幅为-3.72%。人工智能指数价格一直高于计算机行业指数,但8月跌幅小于计算机行业指数。

图1: 8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8月31日总市值15698.06亿,含成分股66支,权重等分。上市板分布为主板17支,创业板28支,科创板6支,中小板15支。

表1: 8月份人工智能板块涨幅前十

股票代码	股票简称	8月31日收盘价	月涨跌幅	相对计算机
301316.SZ	慧博云通	40.28	75.13%	78.85%
002235.SZ	安妮股份	8.07	24.73%	28.45%
300465.SZ	高伟达	11.91	24.32%	28.04%
600633.SH	浙数文化	17.01	19.12%	22.84%
300245.SZ	天玑科技	10.13	12.18%	15.90%
002152.SZ	广电运通	12.55	9.04%	12.76%
300400.SZ	劲拓股份	15.91	4.88%	8.60%
300047.SZ	天源迪科	7.80	4.28%	8.00%
002073.SZ	软控股份	7.09	3.20%	6.92%
688228.SH	开普云	46.15	2.51%	6.23%

资料来源: wind, 中国银河证券研究院

表 2: 8 月份人工智能板块跌幅前十

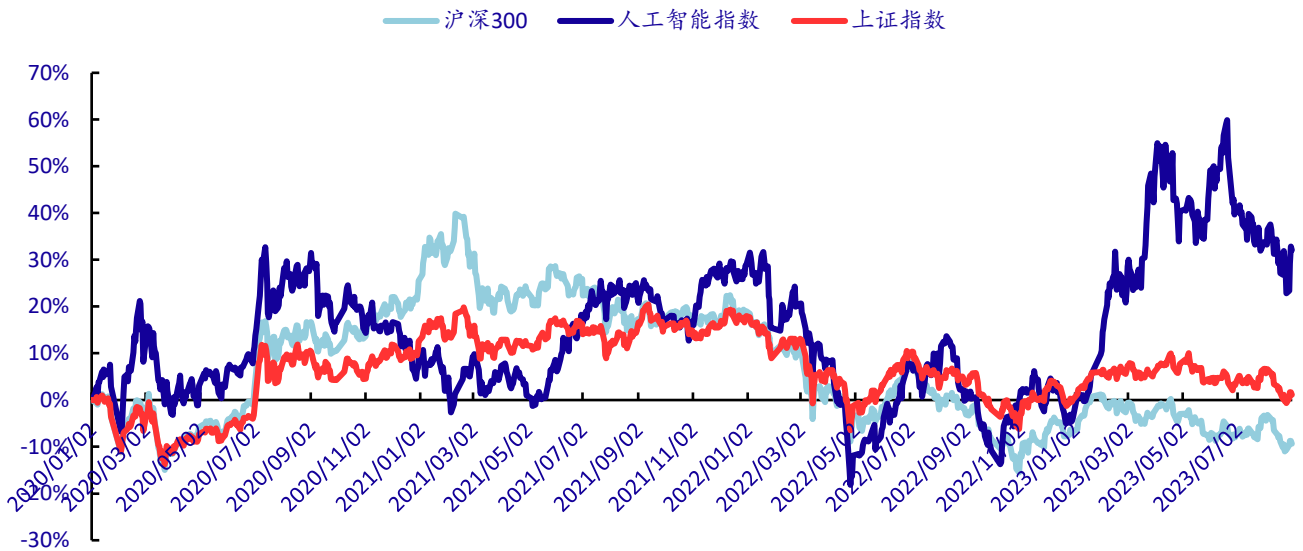
股票代码	股票简称	8 月 31 日收盘价	月涨跌幅	相对计算机
600536.SH	中国软件	43.20	-17.00%	-13.28%
000977.SZ	浪潮信息	39.13	-14.40%	-10.68%
002230.SZ	科大讯飞	54.83	-12.80%	-9.08%
600570.SH	恒生电子	36.00	-12.45%	-8.73%
002368.SZ	太极股份	36.88	-12.15%	-8.43%
603893.SH	瑞芯微	65.48	-11.99%	-8.27%
300223.SZ	北京君正	73.84	-11.90%	-8.18%
300613.SZ	富瀚微	49.43	-10.82%	-7.10%
000681.SZ	视觉中国	15.68	-9.99%	-6.27%
300458.SZ	全志科技	24.25	-9.92%	-6.20%

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 11.33%, 净利润复合增长率 3.97%, 截至 8 月 31 日平均估值 PE(ttm)90.27 倍, PS (ttm) 3.99 倍。

图 2: 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 [单位]亿元	最新收盘日 [单位]元	近 1 月回报 [单位]%	近 3 月回报 [单位]%	近 6 月回报 [单位]%
001986	前海开源人工智能	契约型开放式	8.97	1.61	-7.17	-3.49	-6.19
005729	南方人工智能主题	契约型开放式	4.61	2.24	-4.54	-5.52	4.42
005844	东方人工智能主题 A	契约型开放式	6.30	1.03	3.36	-9.98	14.45
005962	宝盈人工智能 A	契约型开放式	6.03	2.22	-2.54	-3.17	-0.67
005963	宝盈人工智能 C	契约型开放式	2.37	2.13	-2.61	-3.37	-1.07

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。



006281	万家人工智能 A	契约型开放式	15.70	2.12	-6.54	-16.61	8.58
008020	华富中证人工智能产业 ETF 联接 A	契约型开放式	0.96	0.78	-6.80	-10.96	-0.38
008021	华富中证人工智能产业 ETF 联接 C	契约型开放式	1.13	0.77	-6.83	-11.03	-0.53
008585	华夏中证人工智能主题 ETF 联接 A	契约型开放式	4.92	0.78	-6.09	-9.71	1.29
008586	华夏中证人工智能主题 ETF 联接 C	契约型开放式	3.19	0.77	-6.11	-9.77	1.15
009239	融通中证人工智能主题 C	契约型开放式	0.91	1.28	-5.94	-9.44	0.93
011832	西部利得中证人工智能 A	契约型开放式	1.49	0.85	-4.23	-8.95	-0.93
011833	西部利得中证人工智能 C	契约型开放式	1.00	0.85	-4.27	-9.04	-1.12
011839	天弘中证人工智能主题 A	契约型开放式	1.03	0.85	-5.84	-9.09	2.46
011840	天弘中证人工智能主题 C	契约型开放式	3.95	0.84	-5.85	-9.12	2.36
012733	易方达中证人工智能主题 ETF 联接 A	契约型开放式	3.26	0.97	-6.17	-9.52	1.55
012734	易方达中证人工智能主题 ETF 联接 C	契约型开放式	2.79	0.97	-6.17	-9.54	1.50
014162	万家人工智能 C	契约型开放式	12.54	2.09	-6.61	-16.77	8.17
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	19.95	1.03	3.32	-10.08	14.24
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	38.84	0.81	-6.51	-10.04	1.41
161631	融通中证人工智能主题 A	契约型开放式	5.51	1.30	-5.92	-9.35	1.13
512930	平安中证人工智能 ETF	契约型开放式	3.04	1.23	-6.52	-9.72	1.50
515070	华夏中证人工智能 ETF	契约型开放式	14.89	1.04	-6.44	-10.02	1.44
515980	华富中证人工智能产业 ETF	契约型开放式	6.88	0.89	-7.34	-11.60	-1.16
517800	方正富邦中证沪港深人工智能 50ETF	契约型开放式	0.75	0.67	-8.05	-7.89	-4.76

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

证券代码	证券简称	营业收入同比	营业利润同比	2023 中报	2023 中报	总市值	市盈率	市销率	月涨跌幅	今年以来
		增长率	增长率	营收增速	净利润增速					
		[单位]%	[单位]%	[单位]%	[单位]%	[单位]亿元	PE(TTM)	PS(TTM)	[单位]%	[单位]%
000977.SZ	浪潮信息	3.70	0.13	-28.85	-66.67	576.05	39.69	0.97	-14.40	82.31
002230.SZ	科大讯飞	2.77	-79.81	-2.26	-93.87	1,269.69	355.95	6.81	-12.80	67.26
002236.SZ	大华股份	-6.91	-34.33	3.88	32.20	710.95	25.57	2.29	-1.69	93.02
002362.SZ	汉王科技	-13.19	-257.57	4.71	-24.03	59.40	-41.88	4.16	-3.91	56.47
002405.SZ	四维图新	9.37	-512.76	9.71	-376.58	244.91	-41.09	7.04	-8.12	-6.53
002415.SZ	海康威视	2.14	-19.98	0.84	-6.21	3,291.84	26.51	3.94	-2.86	3.77
300229.SZ	拓尔思	-11.81	-49.08	4.58	-5.27	163.51	133.25	17.64	-7.43	75.67
300474.SZ	景嘉微	5.56	-6.09	-36.55	-106.14	339.61	216.90	35.56	-4.44	36.26
601360.SH	三六零	-12.54	-264.55	-6.64	43.43	798.85	-39.23	8.68	-5.81	70.95
603019.SH	中科曙光	15.44	34.85	7.33	22.78	595.53	37.07	4.45	-9.38	84.60
688088.SH	虹软科技	-7.22	-57.60	29.73	-9.61	159.15	305.44	26.11	2.35	74.60
688169.SH	石头科技	13.56	-15.88	15.41	19.93	385.28	29.49	5.44	6.35	66.43
688207.SH	格灵深瞳	20.47	131.10	34.35	115.28	56.92	120.36	14.45	-0.50	33.77
688256.SH	寒武纪-U	1.11	-60.67	-33.37	12.87	661.55	-56.11	98.49	-3.18	191.06
688787.SH	海天瑞声	27.32	-35.65	-35.13	-188.76	49.26	-682.91	22.13	-4.98	92.23
688793.SH	倍轻松	-24.69	-239.06	30.70	30.39	29.10	-25.46	2.81	-22.16	-3.72
002410.SZ	广联达	17.80	44.53	10.71	-36.54	408.56	49.94	5.97	-24.02	-42.34

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。

688327.SH	云从科技-UW	-51.06	-38.53	-58.16	5.32	166.95	-19.69	55.85	-6.94	47.22
688343.SH	云天励飞-U	-3.44	-14.91	-20.90	8.36	191.91	-44.87	37.82	-4.52	-48.09
688246.SH	嘉和美康	9.97	1.98	15.89	2.96	48.05	67.99	6.35	5.85	18.49
603893.SH	瑞芯微	-25.34	-57.55	-31.34	-90.89	273.66	547.97	16.68	-11.99	-4.50
300033.SZ	同花顺	1.40	-12.15	5.83	-5.29	899.67	54.02	24.72	-11.46	72.15
300496.SZ	中科创达	31.96	17.89	0.32	-9.25	367.08	48.05	6.73	-11.38	-19.92
688111.SH	金山办公	18.44	6.71	21.25	12.58	1,821.95	152.19	42.71	-2.81	49.44

资料来源: wind, 中国银河证券研究院

表 5: 海外人工智能上市公司近况一览

证券代码	证券简称	营业收入 [单位]亿元	营业收入 (同比增长率) [单位]%	净利润 [单位]亿元	净利润 (近1年增长率) [单位]%	总市值 [单位]亿元	市盈率 PE(TTM)	市销率 PS(TTM)	月涨跌幅 [单位]%	今年以来 涨跌幅 [单位]%
0020.HK	商汤-W	38.09	-18.97	-60.45	64.73	522.12	-8.01	12.25	-15.68	-29.73
2158.HK	医渡科技	12.37	42.69	-7.62	83.03	41.80	-5.83	2.74	-29.86	-34.92
BIDU.O	百度	1236.75	-0.66	75.59	-26.08	499.30	22.77	2.81	-8.44	24.87
GOOGL.O	谷歌	2828.36	9.78	599.72	-21.12	17169.68	28.17	6.07	2.60	54.34
MSFT.O	微软	1982.70	17.96	727.38	18.10	24351.79	33.65	12.28	-2.22	37.60
NVDA.O	英伟达	269.14	61.40	97.52	0.83	12190.69	118.06	45.30	5.62	237.82
TSLA.O	特斯拉	814.62	51.35	125.56	127.51	8191.44	67.17	10.06	-3.50	109.51

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
8.31	<p>全国首个电力数据要素创新实验室成立</p> <p>8月31日, 全国首个电力数据要素创新实验室正式启动。据悉, 该实验室由深圳数据交易所、南方电网数字平台科技(广东)有限公司合作成立, 以推动电力数据要素流通为目标, 研究解决数据要素确权授权、合规管理体系、市场体系构建的重点和难点问题, 旨在以电力数据产品为试点实现市场生态。</p>
8.30	<p>上海数交所与院士团队共筑数据安全底座</p> <p>8月30日, 上海数据交易所(以下简称上海数交所)与中国工程院院士邬江兴领衔的复旦大学大数据研究院签署合作协议, 共同构建覆盖数据安全流通全要素、全过程的可信框架, 以提供低成本、高效率、可信赖的数据流通环境。据介绍, 此次合作构建的可信交付框架秉承开放、可控原则, 将达到物理域、信息域、认知域的可信共识, 为安全、高效、开放、协同的生态建设, 以及异构空间互联互通提供关键支撑。此框架还可以整合上海数交所的一链多系统等数据流通新型基础设施, 将可信交付融入业务场景, 提供可信保障。</p>
8.21	<p>财政部印发《企业数据资源相关会计处理暂行规定》通知</p> <p>为规范企业数据资源相关会计处理, 强化相关会计信息披露, 8月21日, 财政部制定印发了《企业数据资源相关会计处理暂行规定》(财会〔2023〕11号, 以下简称《暂行规定》), 自2024年1</p>

	<p>月 1 日起施行。制定《暂行规定》将有助于进一步推动和规范数据相关企业执行会计准则，准确反映数据相关业务和经济实质。同时，也将为持续深化相关会计问题研究积累中国经验，有助于在国际会计准则相关研究制订等工作中更好发出中国声音。制定《暂行规定》，进一步强化数据资源相关信息披露，将有助于为有关监管部门完善数字经济治理体系、加强宏观管理提供会计信息支撑，也为投资者等报表使用者了解企业数据资源价值、提升决策效率提供有用信息。</p>
8.21	<p>中国首家数据流量生态园数据要素流通交易规模破 34 亿元</p> <p>山西数据流量生态园(以下简称“生态园”)，作为中国首家以数据流量为特色的数字经济产业园区，截至 8 月 21 日，生态园数据要素流通交易规模突破 34 亿元。截至目前，生态园数据要素流通交易规模达 34 亿元，拉动入园企业营收超 195 亿元，带动就业 8000 余人，吸引 580 余家数字经济企业入驻，包括上市公司 13 家、“专精特新”企业 23 家，“四上”企业 29 家，形成以数字基建为底层支撑的 16 大数字应用场景。下一步，生态园将按照“前轻后重、数实融合”的发展理念，不断强化“全链路一站式的数字化转型服务商”定位，重点打造以先进制造业为支撑的数字智造 1 大主导产业，培育人工智能、算力算法、数据流量等多种业态，形成“1+X”产业生态体系，持续用数字技术为传统产业赋能赋智赋值，努力将园区建设成为中部地区数字经济发展新高地。</p>
8.18	<p>全国首个能源数据要素 OID 登记子节点正式运营</p> <p>贵州电网有限责任公司获得贵阳大数据交易所授权，正式开展全国首个能源数据要素 OID (对象标识符) 登记子节点运营管理，将能源数据要素在数据商、数据资源、数据资产、数据中介、数据交易、数据信托等领域实行登记服务。通过对能源数据要素进行统一的标识和命名，不仅可有效提高数据要素的可识别性和可管理性，为数据要素的组织和应用提供更好的基础和支持，还可高效助力各分省能源数据中心建设及应用。据了解，能源数据要素 OID 登记子节点一方面将为能源数据要素的确权提供可靠依据，通过对能源数据要素的来源、提供者、权利人等关键信息的记录，确保数据权属明确、来源可追溯，为后续数据利用和交易提供有效保障；另一方面，将大力促进能源数据的流通和合理利用，通过及时记录能源数据的使用期限、使用次数、使用限制等重要信息，实现对能源数据要素流通交易全生命周期的管理。</p>
8.15	<p>上海市数据要素市场改革:从数据走向繁荣</p> <p>8 月 15 日，上海市人民政府办公厅印发了《立足数字经济新赛道推动数据要素产业创新发展行动方案(2023-2025 年)》(以下简称《方案》)。方案阐述了数据要素产业发展的具体目标和策略，旨在依据党中央和国务院的总体战略，以构建强大的数字国家和数字中国为目标，推进数据要素产业的创新发展。该计划的基础原则是国家数据基础设施和培育数据要素市场，并将这些要求与城市的数字化转型需求相结合，旨在推动数据要素产业的演进，催生数字红利的释放，提升数字经济的质量。这一行动方案采取多层次方法，作为实现这些战略目标的关键蓝图。</p>
8.8	<p>工信部: 加快培育数据要素市场进一步夯实数字经济发展基础</p> <p>工信部将联合有关部门，围绕三方面做好相关工作。一是认真贯彻落实《“十四五”数字经济发展规划》，推动构建数据基础制度，加快培育数据要素市场，持续完善数字经济发展顶层设计，发挥好数字经济部际联席会议机制的作用，推动数字经济发展各项重大任务落实落细。二是深入贯彻落实《“十四五”智能制造发展规划》，实施工业互联网创新发展工程，持续深化工业互联网融合应用，进一步夯实数字经济发展基础。三是加强部门协同合作，继续落实好支持软件产业发展及工业基础装备改造的各项税费优惠政策，加强政策跟踪分析评估，持续优化税费管理服务措施。</p>

资料来源: wind, 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
8.25	深圳国际仲裁院与深圳交易集团、深圳数据交易所签署战略合作协议

	<p>8月25日,深圳交易集团有限公司(深圳公共资源交易中心)(以下简称“深交易”)、深交所所属企业深圳数据交易所(以下简称“深数所”)在前海国际仲裁大厦分别与深圳国际仲裁院(以下简称“深国仲”)签署合作协议。深国仲党组书记、院长刘晓春,深交易党委书记、董事长陈戈出席签约仪式。本次战略合作协议的签署是深交易、深数所与深国仲深化合作、携手前行的全新起点。未来,三方将共享优势资源、凝聚发展合力,共筑高质量发展高地。</p>
8.23	<p>杭州数据交易所正式揭牌! 华数获得全国首张“数字证书”</p> <p>8月23日下午,“中国数谷”2023杭州峰会(夏季)在高新区(滨江)举行。会上,杭州数据交易所(以下简称“杭数所”)正式揭牌,为数据交易双方发出全国首批“数据合规流通数字证书(以下简称‘数字证书’)”意味着,“中国数谷”在探索构建全流程数据合规监管和数据流通交易制度方面迈出了重要一步。会上,“中国数谷”领导小组还发布了《“数谷”要素改革——数商企业认定标准及引育政策方向》,未来,将持续围绕繁荣数据产业、培育壮大数据市场主体的目标,开展三类数商认定、评价工作,并出台相应引育政策。</p>
8.21	<p>上海数据交易所发布“数据要素市场繁荣计划”</p> <p>2023年8月21日,上海数据交易所发布“数据要素市场繁荣计划”(下称“繁荣计划”),将面向全国范围内的数据要素市场相关企业,推出关于开展促进数据要素流通专项激励资金的方案,拟设立一亿元人民币的专项激励资金,以繁荣数商生态、活跃场内交易,充分激发数据要素市场活力。界面新闻从上海数据交易所获悉,“繁荣计划”涵盖三大方面的激励举措,一是首次登记挂牌补贴,数据产品供方企业首次在上海数交所首次登记并挂牌的,按照每家不超过10万元的额度给予一次性补贴;二是数据产品交易示范补贴,对符合一定交易合同数量且达成相应交易额的数据产品供需方企业给予相应激励资金;三是优质数商培育补贴,鼓励数商更多、更好地为场内交易主体进行多元服务,活跃场内交易。</p>
8.5	<p>人民数据与多家数据交易所签署合作协议,正在积极打造云、数、链三大数据平台</p> <p>8月5日,人民网·人民数据分别与贵阳大数据交易所、浙江大数据交易中心、杭州国际数字交易中心、武汉长江大数据交易中心、海南国际知识产权交易所等多家数据交易所签署合作协议,携手各方打造稳定、可靠、安全的大数据产品和服务,共同促进当地数字经济高质量发展。人民网·人民数据正在积极打造云、数、链三大数据平台。其中云平台作为数据核心资源,承担数据存储功能,致力于打造数据云生态;数据确权流通平台是有效进行数据资产管理的平台,致力于数据确权、授权、流通交易三个方向;人民链是数据服务的底层技术支持,用于服务大数据的“存、管、用”。</p>

资料来源:wind,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
2023.8.31	百度“文心一言”	8月31日,百度宣布其全新一代大语言模型“文心一言”正式向公众开放,同时开放的还包括国内首款 AI 原生对话应用“文心一言 APP”,以及具备全新生成式 AI 能力的百度搜索“AI 伙伴”。数据统计显示,文心一言开放首日,百度搜索有超 3 亿次需求由生成式智能引擎解决;百度搜索“AI 伙伴”当日访问用户数突破 400 万。
2023.8.31	智谱清言	8月31日,北京智谱华章科技有限公司(以下简称“智谱 AI”)正式上线首款生成式 AI 助手智谱清言。该助手基于智谱 AI 自主研发的中英双语对话模型 ChatGLM2,用户可通过苹果商店 AppStore、安卓主流应用市场(华为、OPPO、vivo、应用宝及小米)下载或搜索微信小程序进行体验。

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。

2023.8.31	商汤“日日新”	8月31日,商汤科技日日新大模型旗下自然语言应用“商量 SenseChat”正式面向广大用户开放服务。商汤“日日新 SenseNova”大模型体系和生成式 AI 产品系列,包括商量、秒画、如影、琼宇、格物,分别对应着自然语言交互、AI 文生图、数字人、3D 大场景重建、3D 小物体生成这五个主流的生成式 AI 应用。未来,商汤将依托强大的基模型,以及在算力、数据、算法基础上的积累和丰富 Know-how,持续升级商汤日日新大模型体系下的多种生成式 AI 产品,满足各类用户的需求。
2023.8.31	紫东太初	“紫东太初”全模态大模型是8月31日首批通过《生成式人工智能服务管理暂行办法》备案,可面向全社会开放服务的8个大模型之一,是国内少有的千亿级参数规模的大模型。下一步,“紫东太初”大模型将围绕武汉优势产业,打造更多面向 B 端和 C 端的服务,比如智能汽车、智能制造、智慧医疗、智慧政务、智慧教育等,利用当地的人才优势和场景优势,快速实现优势产业的数字化升级。
2023.8.17	抖音云雀大模型	8月17日消息,字节跳动旗下 LLM 人工智能机器人“豆包”现已开始小范围邀请测试,用户可通过手机号、抖音或者 AppleID 登录。“豆包”是字节跳动公司基于云雀模型开发的 AI 工具,提供聊天机器人、写作助手以及英语学习助手等功能,它可以回答各种问题并进行对话,帮助人们获取信息,支持网页 Web 平台,iOS 以及安卓平台,但 iOS 需要使用 TestFlight 安装。
2023.8.15	讯飞星火认知大模型 v2.0	8月15日,科大讯飞发布了讯飞星火认知大模型 V2.0 版本,重点升级了代码能力和多模态交互能力。科大讯飞董事长刘庆峰表示,代码能力是支撑认知大模型智慧的关键维度,而多模态能力则是实现通用人工智能的必经之路。针对企业用户对于人工智能的需求,此次科大讯飞与华为联合发布了讯飞星火一体机,为企业提供国产自主创新平台上的自主、安全、可控的私有化部署大模型。
2023.8.8	百川大模型	8月8日,百川智能在北京宣布新一代大模型 Baichuan-53B 正式上线。百川智能表示,在文本创作的创意性、风格模仿和实用性上,Baichuan-53B 都可以做到足够好,大部分任务上都能给出不错的回应。在 Baichuan-53B 上,搜狗强调大模型和搜索进行了很高程度的融合,希望能通过这种机制给未来的搜索模型打下了基础。
2023.8.8	广汽 AI 大模型平台	8月8日,广汽正式推出 AI 大模型技术的最新研发成果——广汽 AI 大模型平台,据悉,该平台是汽车行业首创的大模型平台技术,将在近期搭载于高端智能轿跑昊铂 GT 亮相。广汽 AI 大模型平台聚合了多种 AI 大模型,并与广汽智能网联底层能力深度融合,不仅能调用通用大模型的能力,还能基于智能汽车的应用场景构建专用模型,实现全场景智能体验的跃迁。此外,广汽 AI 大模型平台以海量的用车和研发数据集为基础,将结构化数据汇聚到 AI 中台,通过中台训练形成多种特定场景模型;平台可以根据应用需要,灵活调用多个模型,从而获得最优的推理结果,让应用服务精准、高效运行。
2023.8.3	阿里云通义大模型	8月3日,阿里云开源通义千问 70 亿参数模型,包括通用模型 Qwen-7B 和对话模型 Qwen-7B-Chat,两款模型均已上线魔搭社区,开源、免费、可商用。此举让阿里云成为国内首个加入大模型开源行列的大型科技企业。

资料来源:wind,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
2023.8.28	Meta	Llama 代码生成模型能力直追 GPT-4 研究团队在 Phind 内部数据集上对 CodeLlama-34B 和 CodeLlama-34B-Python 进行微调之后发现,这两款模型微调之后在 HumanEval 测试中的通过率均已超过 GPT-4 在 3 月份的成绩。本次发布的 Code Llama 是在 Llama 2 的基础上,通过特定的代码数据集进一步训练而来,支持 C++、Java、Python、PHP、Typescript (Javascript)、C# 和 Bash

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。

		等众多流行语言，依然是开源可商用。Code Llama 对编程专家和初学者都非常有用，无论是用专业的编程语言还是用自然语言（如普通话）来描述编程需求，Code Llama 都能理解，并很好地生成相应的代码或与代码相关的解释。这极大降低了开发门槛和效率。
2023.8.22	Snap	Snap 推出 AI 大模型 Text2Cinmagraph Snap 近日联手卡内基梅隆大学共同发表论文，公布一款 AI 大模型 Text2Cinmagraph。据悉，Text2Cinmagraph 可以将图片转换成动画，还可以根据文本内容来调整动画的细节。论文表示，利用现有的自然图像和视频数据集，Text2Cinmagraph 可以准确地分割现实图像，并根据语义信息预测合理地运动。
2023.8.16	谷歌	谷歌最强大模型 Gemini 被曝今年秋季发布 谷歌数百人团队正在为将于今年秋季发布的大型 AI 模型 Gemini 而努力，这是谷歌迄今构建的风险最高的产品之一。据悉，Gemini 将 GPT-4 等大型语言模型的文本能力与文生图能力相结合。谷歌员工还讨论了使用 Gemini 提供分析图表或创建带有文本描述的图形、使用文本或语音命令控制软件等功能。
2023.8.9	IBM	IBM 宣布旗下 AI 平台将托管 Meta 大语言模型 8月9日，IBM 表示，将在自己的企业人工智能平台上托管 Meta Platform 的人工智能语言模型 Meta Llama 2。IBM 的 watsonx 平台可以帮助企业将人工智能整合到其工作流程中，为一些客户提供了 Meta 发布的 Llama 2 的早期访问权。IBM 表示，Meta 的开源人工智能模型发布后，还将推出更多软件，如人工智能调整工作室、情况说明书和其他生成式人工智能模型。

资料来源：wind，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
8.30	谷歌推出第五代 AI 芯片：训练和运行 AI 模型的速度提高 5 倍 当地时间 8 月 29 日，谷歌在旧金山的年度云会议 Google Cloud Next 上发布了新的人工智能芯片，即第五代定制张量处理器（TPU）芯片 TPU v5e，用于大模型训练和推理。与上一代芯片相比，TPU v5e 每一美元的训练性能提高 2 倍，每一美元的推理性能提高 2.5 倍。
8.25	AI 服务器需求增加推动，二季度全球 DRAM 销售额环比增长 20% 从机构最新的数据来看，生成式人工智能聊天机器人及大型语言模型对半导体领域的拉动，已在 DRAM 上体现，二季度全球 DRAM 的销售额环比就有大幅增加。研究机构的数据显示，在今年二季度，全球 DRAM 的销售额达到了 114.3 亿美元，同比增长 20.4%。机构在报告中还表示，二季度全球 DRAM 的销售额环比大幅增加，是由于 AI 服务器需求增加，推升了对高带宽存储器的需求，三大 DRAM 厂商的出货量因此也大幅增加。除了 DRAM，生成式人工智能聊天机器人及大型语言模型的训练和应用，也大幅拉升了对高性能 GPU 的需求，主要供应商英伟达在截至 7 月 30 日的 2024 财年第二财季的营收，就同比大增 101%，环比也大增 88%，达到了 135.07 亿美元，净利润超过 60 亿美元，同比增长超过 400%。
8.23	英伟达 Q2 净利润暴涨八倍，AI 芯片每年需求增长超 20% 英伟达公布了截至 2023 年 7 月 30 日的第二季度财报。受益于 AI 技术的快速发展，英伟达业绩暴涨，远超市场预期。英伟达 CEO 黄仁勋表示，一个新的计算机时代已经开始，全球公司都在从通用计算向加速计算和生成式 AI 转型，采用生成式 AI 的竞赛已经开始。净利润上，按照 GAAP（一般公认会计原则）算约为 61.9 亿美元，同比暴涨 843%，近乎 8.5 倍。非 GAAP 净利润为 67.4

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。

	<p>亿美元，同比暴涨 422%。第二季度的业绩暴增主要是由于 AI 需求快速增长所带来的数据中心业务的膨胀，财报显示，英伟达二季度数据中心营收 103.2 亿美元，创下历史纪录，与上年同期相比增长 171%，与上一财季相比增长 141%，而分析师预期为 79.8 亿美元。用四个字来形容英伟达此次的营收表现，那就是远超预期。</p>
8.23	<p>能效达传统芯片 14 倍，IBM 开发出新 AI 芯片</p> <p>8 月 23 日，《自然》发表的研究报道了一种能效为传统数字计算机芯片 14 倍的人工智能（AI）模拟芯片。这一由 IBM 研究实验室开发的芯片在语音识别上的效率超过了通用处理器。该技术或能突破当前 AI 开发中因算力性能不足和效率不高而遇到的瓶颈。研究团队此次开发了一个 14 纳米的模拟芯片，在 34 个模块（tile）中含有 3500 万个相变化内存单元。对于更大的 Librispeech 模型来说，该芯片能达到每秒每瓦 12.4 万亿次运算，系统性能估计最高能达到传统通用处理器的 14 倍。研究团队总结道，该研究在小模型和大模型中同时验证了模拟 AI 技术的性能和效率，有望成为数字系统的商业可行的替代选择。</p>
8.22	<p>全球领先系统制造商推出 NVIDIA AI 就绪型服务器，为企业大幅提升生成式 AI 性能</p> <p>2023 年 8 月 22 日 — NVIDIA 今日宣布，全球领先的系统制造商将推出 AI 就绪型服务器，其支持同于今日发布的 VMware Private AI Foundation with NVIDIA，帮助企业使用自有业务数据来定制和部署生成式 AI 应用。NVIDIA AI 就绪型服务器将采用 NVIDIA L40S GPU、NVIDIA BlueField -3 DPU 和 NVIDIA AI Enterprise 软件使企业能够微调生成式 AI 基础模型并部署生成式 AI 应用，比如智能聊天机器人、搜索和摘要工具等。这些服务器还提供由 NVIDIA 加速的基础设施和软件，以支持 VMware Private AI Foundation with NVIDIA。</p>
8.18	<p>联想集团发布两款全新 AI 服务器</p> <p>8 月 18 日消息，今日，联想集团在 2023 中国算力大会上发布两款全新 AI 服务器——联想问天 WA7780 G3 AI 大模型训练服务器和联想问天 WA5480 G3 AI 训推一体服务器。至此，联想集团 AI 服务器已实现对 AI 大模型、超大模型训练，AI 中小模型训练、推理和 AI 推理的全域覆盖，参数规模从十亿级到千亿级。</p>
8.8	<p>英伟达发布 AI 芯片 GH200，黄仁勋称“它会疯狂推理”</p> <p>英伟达当地时间周二在洛杉矶举行的 SIGGRAPH 大会上宣布了一项人工智能芯片的新配置，英伟达创始人 CEO 黄仁勋表示，新的芯片能加速生成式人工智能应用程序，并降低大模型的运行成本，让数据中心规模化。这款最新发布的人工智能芯片 GH200 使用与英伟达当前最高端的 AI 芯片 H100 相同的 GPU，但配备 141GB 的内存以及 72 核的 ARM 芯片，H100 的内存为 80GB。英伟达副总裁 Ian Buck 在一场媒体发布会上表示，新版本的芯片增加了高带宽内存的数量，这种设计能够驱动更大的 AI 模型运行。GH200 经过优化，可以执行 AI 推理功能，从而有效地为类似 ChatGPT 等生成式 AI 应用程序提供支持。</p>

资料来源：wind，中国银河证券研究院整理

三、人工智能前沿行业动态

（一）前沿技术动态

1、GLM 技术团队发布长文本理解评测数据集

近日，GLM 技术团队发布专门针对模型长文本理解能力的评测数据集 LongBench。该数据集包含了 13 个英文任务、5 个中文任务和 2 个代码任务。多数任务的平均长度在 5k-15k 之间，共包含约 4500 条测试数据。LongBench 包含单文档 QA、多文档 QA、摘要、Few-shot 学习、代码补全和合成任务等六大类任务 20 个不同子任务。

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。

利用该评测数据集，GLM 技术团队分别对 GPT-3.5-Turbo-16k、Llama2-7B-chat-4k、LongChat-7B-16k、XGen-7B-8k、InternLM-7B-8k、ChatGLM2-6B、ChatGLM2-6B-32k* 等 7 个支持长文本的模型的性能。其中 ChatGLM2-6B-32K 在 ChatGLM2-6B 的基础上进一步强化了对于长文本的理解能力，能够更好的处理最多 32K 长度的上下文。结果如下所示：

图 3: LongBench 测评结果

英文榜单

	Avg	单文档 QA	多文档 QA	摘要	Few-shot学习	代码补全	合成任务
GPT-3.5-Turbo-16k	45.5	39.8	38.7	26.5	76.0	54.5	37.8
Llama2-7B-chat-4k	29.0	24.8	21.4	23.9	50.5	47.3	5.9
LongChat-7B-16k	33.7	29.3	16.1	25.8	59.9	57.0	14.2
XGen-7B-8k	28.7	24.5	20.4	24.8	58.7	38.0	5.6
InternLM-7B-8k	24.7	17.1	20.8	13.3	52.7	39.7	4.7
ChatGLM2-6B	26.0	23.1	15.0	22.9	46.1	46.1	2.7
ChatGLM2-6B-32k	42.7	32.8	34.0	28.6	68.1	52.7	39.8

中文榜单

	Avg	单文档 QA	多文档 QA	摘要	Few-shot学习	代码补全	合成任务
GPT-3.5-Turbo-16k	44.5	61.2	28.7	16.0	29.2	54.5	77.5
Llama2-7B-chat-4k	13.5	11.6	1.9	0.2	19.8	47.3	0.5
LongChat-7B-16k	23.7	26.6	19.1	14.0	20.8	57.0	4.8
XGen-7B-8k	14.5	14.2	9.1	1.5	20.0	38.0	4.2
InternLM-7B-8k	18.6	33.3	8.9	13.0	15.5	39.7	0.9
ChatGLM2-6B	22.5	33.0	15.2	14.6	20.5	46.1	5.5

资料来源: GLM, 中国银河证券研究院整理

2、网易互娱 AI Lab 图像生成论文被 AI 顶会接收

近日，网易互娱 AI Lab 的一篇关于图像生成的学术论文《General Image-to-Image Translation with One-Shot Image Guidance》被国际计算机视觉顶级学术会议 ICCV 2023 接收。网易互娱 AI Lab 提出了一种通用的基于单样本图像引导的图像编辑技术 VCT (visual concept translator, 视觉概念转换器)，给单张参考图像，经过少量的优化步骤，即可把图中的物体或风格迁移至源图像，同时不改变源图像的整体结构。

VCT 相较于以往的方案有以下优势：(1)应用泛化性: VCT 不需要大量的数据进行训练，且生成质量和泛化性更好。其基于反演的思路，以在开放世界数据预训练好的高质量文生图模型为基础，实际应用时，只需要一张输入图和一张参考图就可以完成较好的图片编辑效果。(2)视觉准确性: 相较于近期文字编辑图像的方案，VCT 利用图片进行参考引导。图片参考相比于文字描述，可以更加准确地实现对图片的编辑。(3)不需要额外信息: 相较于近期的一些需要添加额外控制信息（如遮罩图或深度图）等方案来进行引导控制的方案，VCT 直接从源图像和参考图像学习结构信息和语义信息来进行融合生成，不需要额外的限制。

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。

图 4: 视觉概念转换器图像编辑演示

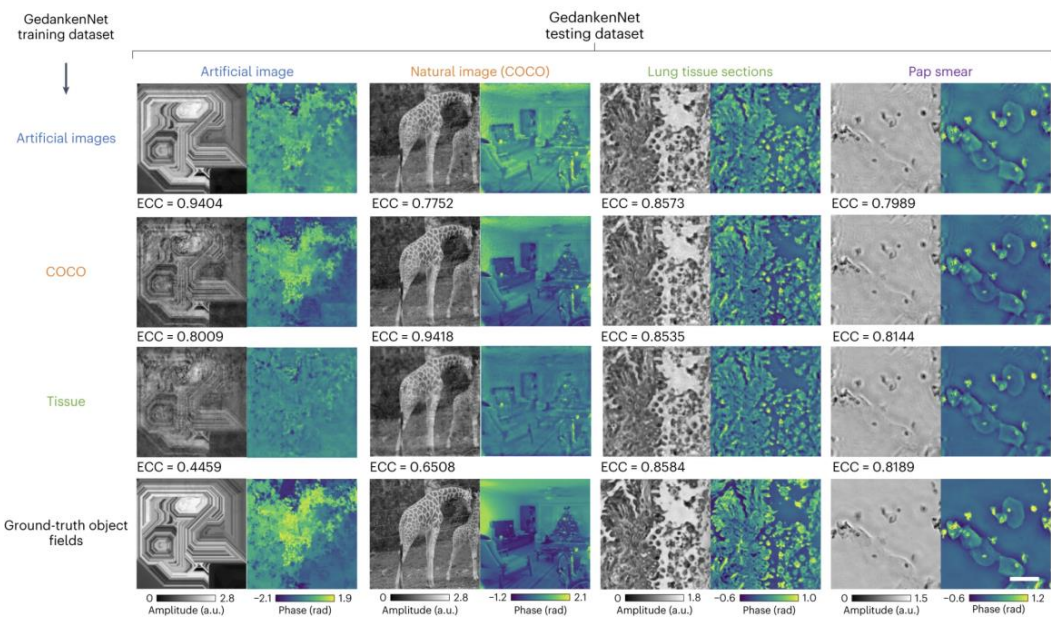


资料来源: ICCV, 中国银河证券研究院整理

3、AI 通过自监督训练重建生物样品的微观图像

来自加州大学洛杉矶分校的研究人员 8 月 7 日在国际学术顶级期刊 Nature 的子刊 Nature Machine Intelligence 上发表一篇新论文, 提出了一个自监督 AI 模型 GedankenNet。该模型无需标注或实验训练数据, 具有更好的泛化能力, 可从物理定律和思想实验中学习。经过自监督训练, GedankenNet 成功在全息图中重建了从未见过的生物样品的微观图像。

图 5: 不同训练数据集上 GedankenNet 的性能表现



资料来源: Nature Machine Intelligence, 中国银河证券研究院整理

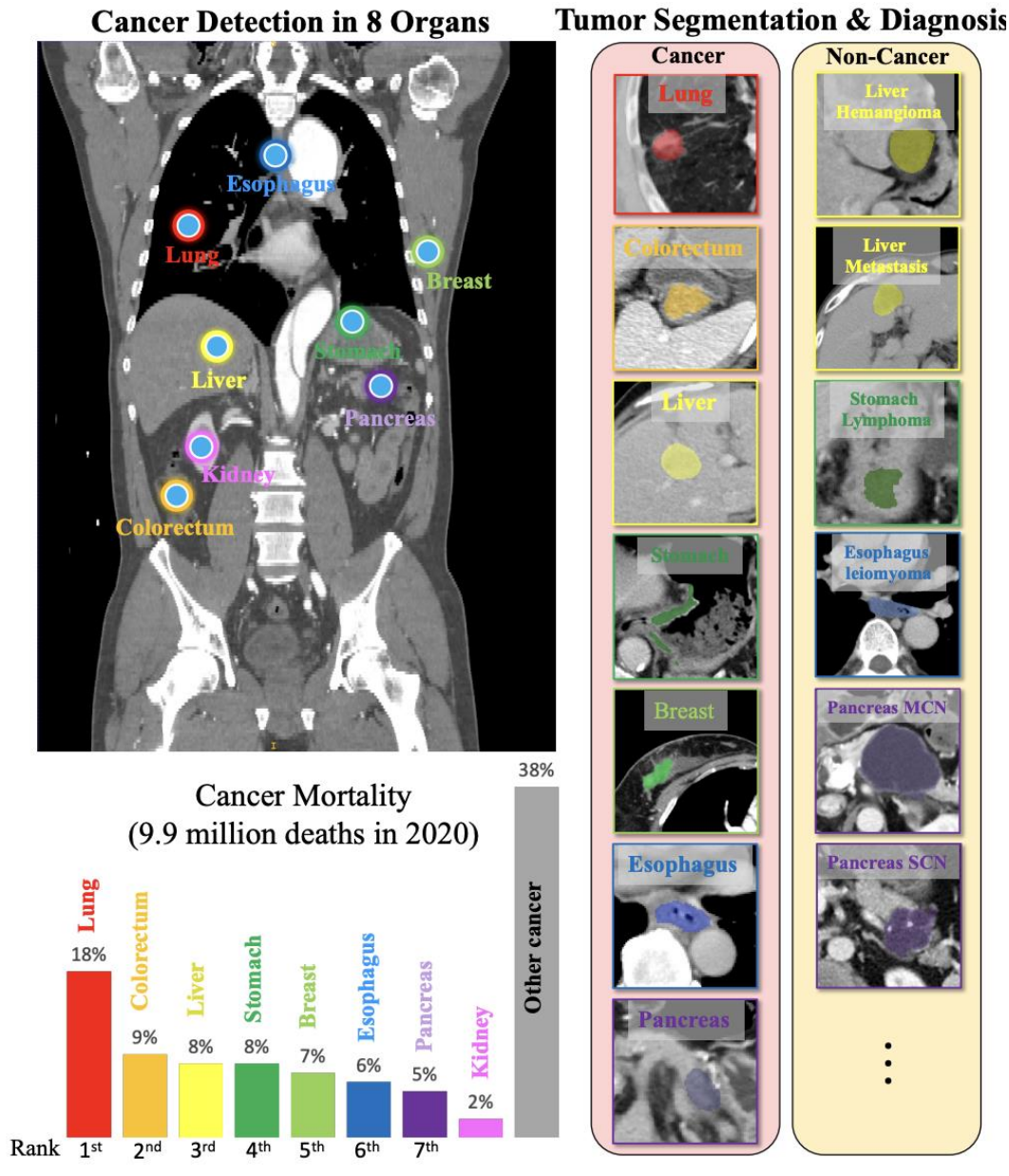
4、阿里达摩院发布癌症分析通用模型

阿里达摩院近日发布多癌影像分析通用模型, 可检测、分割和诊断八种主要的高发致死癌症, 据称有助于实现多癌统一诊断, 降低漏诊概率。

针对市面上医疗 AI 模型出现的高假阳性、漏检等问题，达摩院医疗 AI 团队联合中山大学肿瘤防治中心、四川省肿瘤医院、浙大附属第一医院、盛京医院、广东省人民医院等单位，提出了一个统一的多癌影像分析通用模型（cancerUniT），以 Mask Transformer 语义分割为基础，解决多种肿瘤图像此前难以统一检测、分割和诊断的问题，适用于八种主流的高发高致死癌症（肺、结肠、肝、胃、乳腺、食管、胰腺、肾）以及相关器官中的肿瘤子类型。

根据官方数据，在一组 631 名患者的对比测试中，该模型肿瘤检测、分割和诊断任务的性能均优于 8 个特定器官的单模型组合，检测任务的平均敏感性达到 93%，平均特异性达到 82%。

图 6: 通过 CT 扫描对八个主要器官进行癌症和非癌症检测、分割和诊断



资料来源: arXiv, 中国银河证券研究院整理

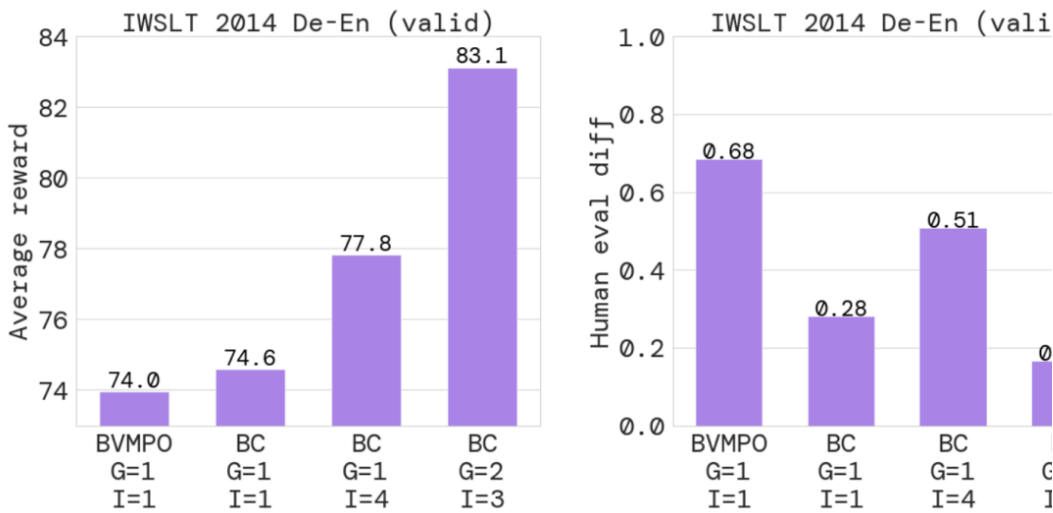
5、DeepMind 发布提高翻译质量的算法 ReST

据机器之心报道，谷歌 DeepMind 于 8 月 21 日发布论文，提出一种使 LLM 与人类偏好对齐更简单的算法 ReST (Reinforced Self-Training)。不同于 RLHF (基于人类反馈的强化学习)

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。

使用人类反馈改进语言模型，ReST 通过生成和使用离线数据进行训练，从而使得 LLM 与人类偏好保持一致。研究团队表示，虽然 ReST 可用于所有生成任务，但本文的重点是机器翻译。研究者在机器翻译任务上进行了实验，测试基准包括 IWSLT 2014、WMT 2020 和 Web Domain。结果表明，ReST 可以极大地提高翻译质量。

图 7: 基于学习奖励和人类评估的绩效比较



资料来源: arXiv, 中国银河证券研究院整理

6、研究称 ChatGPT 满足一些指标能证明其有意识

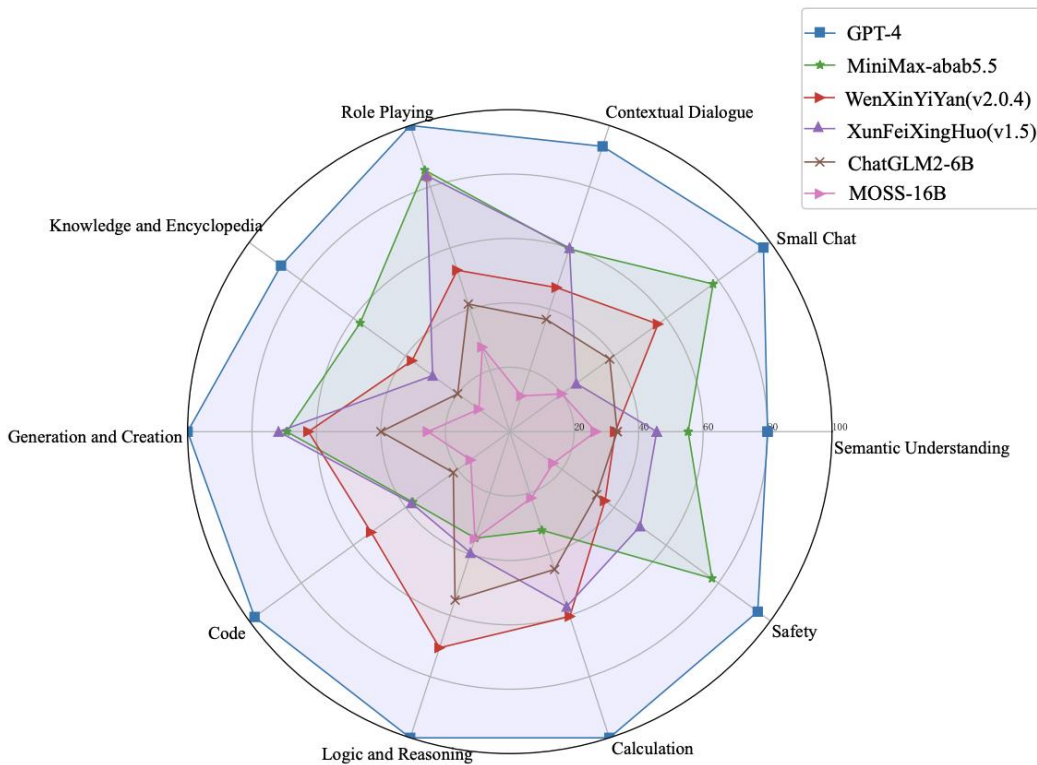
据国际顶刊 Nature 新发布的文章，由 19 名神经科学家、哲学家和计算机科学家组成的小组制定了一份标准清单，用以测试人工智能系统是否已经具有意识。该报告评估了 ChatGPT 等大型语言模型，发现此类系统可以说具有与全局工作空间理论相关的一些意识指标。然而，最终，这项工作并不表明任何现有的人工智能系统都是意识的有力候选者。

为了制定标准，作者假设意识与系统如何处理信息有关，且通过大脑扫描和其他技术对人类和动物进行研究的基于神经科学的意识理论可以应用于人工智能。基于这些假设，研究小组选择了其中六种理论，并从中提取了一系列意识指标。其中之一是全局工作空间理论，声称人类和其他动物使用许多专门的系统（也称为模块）来执行认知任务，例如看和听。作者表示，这篇论文远未达到如何评估人工智能系统意识的最终结论，他们希望其他研究人员帮助完善他们的方法。

7、SuperCLUE 发布中文大模型 8 月榜单

据 CLUE 中文语言理解测评基准微信公众号发文，SuperCLUE 发布中文通用大模型综合性评测基准 2023 年 8 月榜单。本次评测选取了目前国内外最具代表性的 16 个通用大语言模型，8 月评测数据集为全新的 3337 道测试题，榜单由总排行榜、OPEN 多轮开放问题排行榜、OPT 三大能力客观题排行榜、十大基础能力排行榜、开源排行榜五个部分组成。其中，百川智能的 Baichuan-13B-Chat(V2)于总排行榜、OPEN 排行榜、基础能力排行榜、开源排行榜四个部分排名第一，OPT 排行榜第一位是 MiniMax 的 MiniMax-abab5 模型。

图 8: 所选 LLM 在 OPEN 上的能力分布雷达图。



资料来源: arXiv, 中国银河证券研究院整理

(二) 前沿政策动态

本月, 国家各部门、成都市、北京市、深圳市相继公布人工智能产业相关条例, 填补法律法规空白, 支持产业发展。

表 11: 相关政策法规

时间	文件	主要内容
2023.08.22	《新产业标准化领航工程实施方案(2023—2035年)》	提出聚焦新一代信息技术、新能源、民用航空等8大新兴产业, 以及元宇宙、人形机器人、生成式人工智能等9大未来产业, 统筹推进标准的研究、制定、实施和国际化
2023.08.16	《关于促进民营经济做大做强若干措施》	深圳将在新能源汽车、人工智能、新型储能等新兴领域大力培育一批民营领军企业, 培育一批国家级、省级中小企业特色产业集群; 支持民营企业纳入国家、省民营企业百强名单, 对入选国家级、省级的专精特新民营企业给予奖励
2023.08.16	《北京市促进机器人产业创新发展的若干措施》	提升机器人关键技术创新能力, 组织实施机器人产业“筑基”工程, 发布产业关键技术攻关清单, 围绕机器人操作系统、高性能专用芯片和伺服电机、减速器、控制器、传感器等关键零部件, 以及人工智能、多模态大模型等相关技术
2023.08.15	《人工智能法示范法 1.0(专家建议稿)》	建议采用负面清单管理制度对人工智能进行风险管理, 负面清单内的研发、提供活动采取许可机制, 实行事前监管; 负面清单外的研发、提供活动则采取备案机制, 事后监管机制; 创设性提出设置国家人工智能主管机关, 负责统筹人工智能发展和管理
2023.08.02	《关于实施促进民营经济发 展近期若干举措》	其中, 提出支持民营企业参与重大科技攻关, 牵头承担工业软件、云计算、人工智能、工业互联网、基因和细胞医疗、新型储能等领域的攻关任务

资料来源: 政府网站, 中国银河证券研究院整理

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。

四、前沿企业动态

(一) 前沿产品动态

1、多领域机器人广泛应用

German Bionic 推出医护人员专用外骨骼 Apogee+, 可协助人员举起 30 千克重物。曾推出 Cray X 的德国仿生学公司 German Bionics 在今年年初公布了旗下新款外骨骼装备 Apogee, IT 之家早前曾报道, Apogee 是 German Bionic 旗下最轻的外骨骼装备, 每次抬起动作可以抵消用户下背部最高约 30kg 的负荷, 并提供主动行走辅助以减少疲劳。German Bionic 日前基于 Apogee 外骨骼, 设计了 Apogee+, 官方表示, Apogee+ 是第一款专为医护人员设计的外骨骼, 它可以减轻体力劳动的负担, 并考虑到医院的实际情况, 进行了专用设计。

German Bionic 声称, 医疗保健工作经常需要长距离行走、长时间站立和搬运重物, 护士和医生所承受的体力负荷可能是很多人没有考虑到的, 因此该外骨骼着重于减轻医护人员的体力负荷。据悉, Apogee+ 的防尘和防水性能达到 IP54 标准、能够协助医护人员举起重达 30 千克的重物 (和原版 Apogee 外骨骼一致)。

图 9: Apogee 外骨骼演示



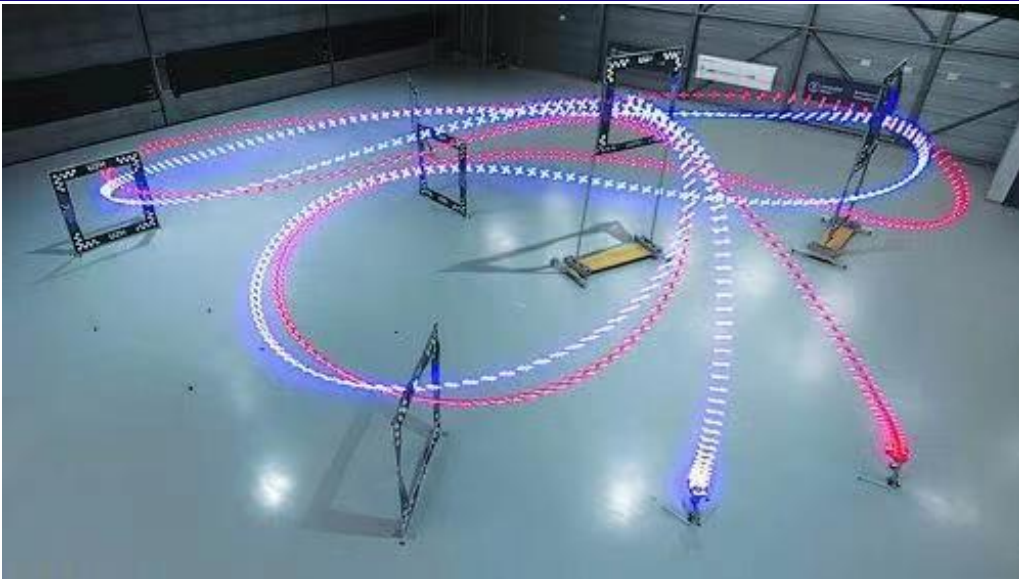
资料来源: German Bionic, 中国银河证券研究院整理

AI 操控无人机能力超越人类冠军。《自然》30 日发表的一篇文章报道了一个名为 Swift 的人工智能 (AI) 系统, 该系统驾驶无人机的能力可在真实世界中一对一冠军赛里战胜人类对手。研究结果标志着移动机器人学和机器智能的一个里程碑, 并可启发在其他物理系统中的部署, 如自动驾驶的地面车辆、飞行器和个人机器人。

深度强化学习系统已在各类博弈中战胜过人类, 包括国际象棋、雅达利游戏、《星际争霸》和 GT 赛车。不过, 这些大部分都是模拟和棋盘游戏环境, 而不是物理世界的竞赛。在第一人称视角的无人机竞速中, 职业选手需驾驶高速飞行器穿过三维赛道。每位飞行员通过机载摄像

机的视频直播从无人机的视角观察周围环境。自动驾驶无人机要达到职业飞行员的水准很难，因为机器人需要在其物理限制下飞行，同时只能根据机载传感器估算其速度和方位。

图 10: 无人机竞速



资料来源: Nature, 中国银河证券研究院整理

江西抚州首台诉讼智能机器人“上岗”。8月23日，一位身穿黑色T恤的中年男子走进江西省抚州市乐安县人民法院诉讼服务中心，一台诉讼智能机器人迎了上来。经过几分钟交流后，该男子在机器人的引领下到服务“一号改革工程”窗口办理了立案手续。这是抚州市首台诉讼智能机器人。自今年7月在乐安“上岗”以来，累计接待访客134人次，解答法律咨询176条，检索案例和法律条文54件。“除了立案指引，这台机器人还有解答法律咨询、检索法律条文、查询类案判例、提供文书模板、指导证据收集和诉讼费用计算等十余项功能。”乐安县法院诉讼服务中心庭长晏诚介绍说，该诉讼智能机器人操作起来非常简单，可根据自己的需求通过语音对话或者点击屏幕选择相关的业务，并根据其语音提示与显示界面进行下一步操作，整个过程便捷流畅，点击即看。

图 11: “乐安”诉讼智能机器人

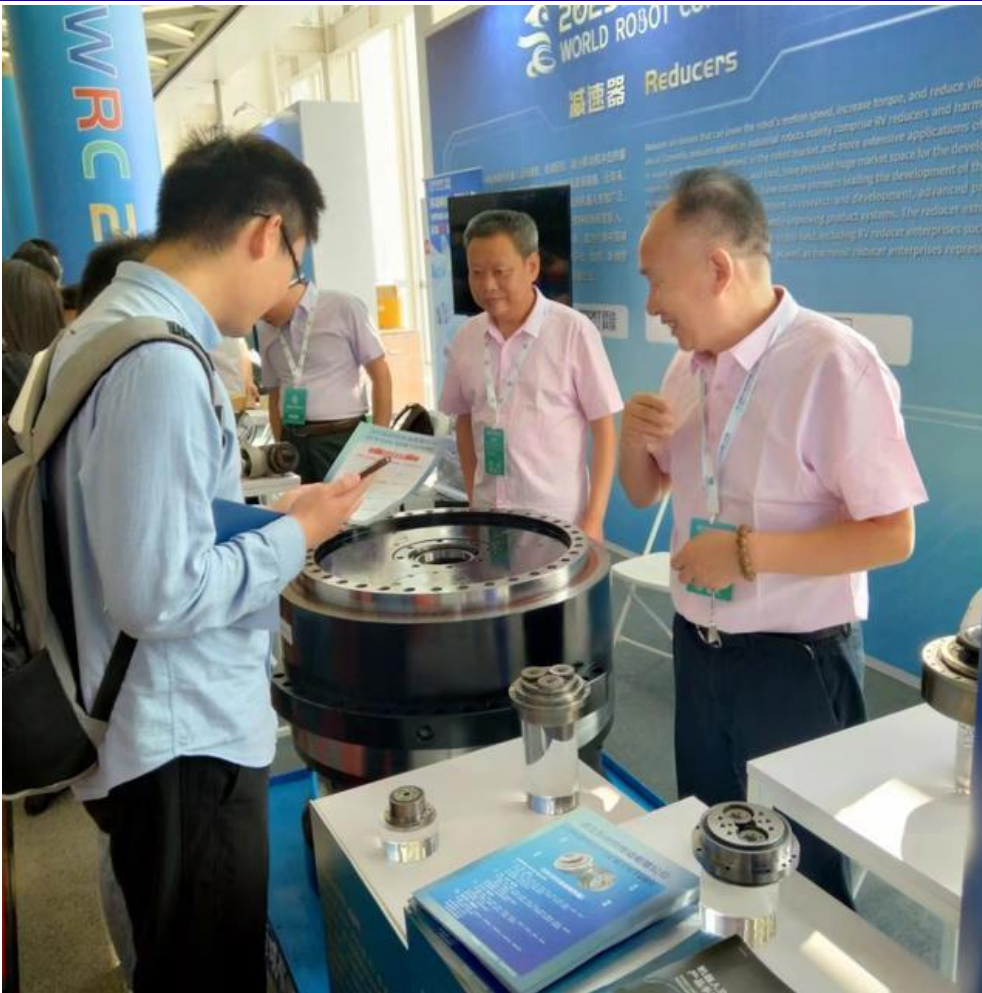


资料来源: 法制日报, 中国银河证券研究院整理

三星电子开始制定人形机器人发展战略。据业内人士透露，三星电子最近开始制定进军机器人市场的战略，由设备体验（DX）部门的规划团队主导该项目。分析师表示，此举针对的是特斯拉、谷歌和微软正在开发的由人工智能驱动的人形机器人领域。为此三星还推迟了用于医疗保健的可穿戴机器人的上线。

武汉自研全球最强机器人关节亮相。8月16日，2023世界机器人大会在北京开幕，包含论坛、博览会、机器人大赛等活动。武汉精华减速机有限公司（以下简称“武汉精华”）携全球载重量最大的RV减速机亮相博览会。大会现场，一台直径1.1米、高1.3米的“圆柱型黑大个”吸引了众多国内外参展商围观。武汉精华董事长余运清在一旁自豪地介绍：“这台RV减速机自重2吨，但载重量高达45吨。我们检索了国内外所有公开信息，这台RV减速机是目前全球已知承载重量最大的RV减速机。”“这台RV减速机能让其承载的设备精准运行，在直径20米的空间里，其控制机械臂运行的误差不会超过10毫米。”余运清介绍，这款减速机是国际合作项目，将应用在国际热核聚变实验堆装置中，维护“人造太阳”——大型核聚变反应堆的运行。

图 12: 向机器人行业上下游厂商介绍全球载重量最大的 RV 减速机情况



资料来源：长江日报，中国银河证券研究院整理

擎朗智能公布多款新品机器人，预计四季度上市。在日前举行的世界机器人大会上，擎朗智能正式公布了多功能配送机器人 t10 和全新的清洁服务机器人产品 c30，预计将于今年第四季度正式上市。据了解，多功能配送机器人 t10 集“配送高手”“全景感知”“多模态交互”“广告营销”四大特性于一身，其具备的开放式大容量托盘、ai 相机检测、大幅升级的运动控制性能和

全能底盘，不仅避震强化，运行可靠，而且有着极限 59cm 的狭窄过道通过能力；同时，依托先进传感器融合，t10 大幅提升了避障、定位、自主路径规划、安全配送流程，并且能够引入更高性能的人机交互；23 寸高清大屏配置，引入了云端广告营销运营系统，实现引流和全能配送二合一。另一款超能清洁机器人 c30 面向酒店场景，具有四大基础清洁功能，单次清洁面积不低于 1500 平米，以专业清洁、便携易用、智能灵巧、云端管理为主要特性。

图 13: 多功能配送机器人 t10



资料来源：擎朗智能，中国银河证券研究院整理

全国首例，青大附院成功完成机器人辅助单髁翻修术。近日，青大附院骨科医院副院长、关节外科主任张海宁教授团队成功完成全国首例 MAKO 机器人辅助下单髁关节翻修术。机器人辅助关节置换手术是当前关节外科领域发展的热点和方向，也代表了智能化关节修复重建领域的最高水准。近年来，青大附院关节外科团队成功开展了包括 MAKO 机器人在内的 8 种各类型机器人辅助的关节置换手术 700 余例，位居国内前列、省内领先水平，积累了丰富的临床经验。在此基础上，张海宁教授团队大胆创新，将机器人应用拓展到单髁翻修这种难度较高的手术当中，并取得了良好的临床效果。

2、全球多家机构致力打造 ChatGPT 式大模型产品

谷歌助手将引入大型语言模型技术。根据 Axios 近期看到的一封发给员工的谷歌内部电子邮件，谷歌计划彻底重塑谷歌助手（Google Assistant），看到了“探索由最新大型语言模型

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。

技术驱动的超级助手的巨大机会”，将用最新的类似于 ChatGPT 和 Bard 聊天机器人的生成式 AI 技术。谷歌表示，改进谷歌助手的一些工作已经开始，首先是该产品的移动版本，正在重组谷歌助手团队，并将进行少量裁员。Axios 获悉，此举将涉及取消谷歌助手数千名员工中的数十个工作岗位。

培生集团将引入生成式 AI 学习工具。据路透社 7 月 31 日报道，全球知名的英国教育集团培生集团将在即将到来的学年将生成式 AI 学习工具引入其在线高等教育平台，试图减轻投资者对 AI 对数字教育服务威胁的担忧。在周一公布的财务业绩中表示，ChatGPT 式的工具将为其在线教科书和评估平台 Pearson+和 Mastering 中的某些课程为学生提供实时个性化支持。培生集团 CEO 安迪·伯德（Andy Bird）称，培生仍处于生成式 AI 开发之旅的早期阶段，但表示“生成式 AI 对公司来说是真正的长期积极因素。”这些工具不仅会为学生提供“找到答案的捷径”，还会通过将 AI 功能与出版商的专有知识产权和数据集相结合，进一步个性化学习体验。

吉利将发布首个汽车行业全场景 AI 大模型。据 IT 之家报道，在吉利汽车月初举行的内部大会上，CEO 淦家阅透露，吉利汽车已具备全栈自研的 AI 大模型技术，将于下半年对外发布。淦家阅介绍称，该模型将是“全球首个汽车行业全场景大模型”，也是“应用场景最丰富”的 AI 大模型，包括绘画、音乐、语言等功能。该大模型也具备自动驾驶 AI 大模型能力，号称实现十亿级别、百亿级别、千亿级别数据大模型全覆盖。

亚马逊组建新团队研发 AI 大模型。据 Business Insider 近日报道，亚马逊 CEO 安迪·贾西（Andy Jassy）发布一封内部电子邮件，任命 Alexa 高级副总裁兼首席科学家罗希特·普拉萨德（Rohit Prasad）作为他的直接下属，负责领导一个新成立的“中心团队”，将集中资源来构建亚马逊最广泛的大型语言模型。贾西将其称作亚马逊“最雄心勃勃”的大型语言模型。

百度千帆大模型平台接入 Llama 2 等 33 个模型。百度智能云“千帆大模型平台”升级了三大功能。百度智能云 AI 与大数据平台总经理忻舟表示，目前，千帆大模型平台已经全面接入 Llama 2 全系列、ChatGLM2、RWKV、MPT 等 33 个大模型，成为国内拥有大模型最多的平台。接入的模型经过千帆平台二次性能增强，模型推理成本可降低 50%。同时，千帆平台上线了国内最全的预置 Prompt 模板，模板数量多达 103 个，覆盖对话、游戏、编程、写作十余个场景。第三，千帆是国内首家提供插件服务的大模型平台，本次升级再次发布多款全新插件。

图 14：2023 百度云智大会



资料来源：百度，中国银河证券研究院整理

Uber 正在开发 AI 聊天机器人。据彭博社 8 月 1 日报道，Uber 正在开发一款由 AI 驱动的聊天机器人，以集成到其应用程序中，加入众多使用该语言工具来改善客户服务、营销和其他自动化任务的公司行列。Uber CEO 达拉·科罗斯瓦西（Dara Khosrowshahi）周二在接受彭博科技节目采访时强调了 Uber 目前在其业务中使用 AI 的方式，但没有提供有关聊天机器人能够做什么的具体细节。

腾讯混元大模型开始应用内测。据 36 氪报道，腾讯自研的“腾讯混元大模型”近日已经进入应用内测阶段，多位腾讯员工透露，近期收到了内部邮件邀请，可以通过网页或者小程序体验混元大模型，腾讯内部多个业务也已经接入混元大模型进行测试。在具体的功能点上，除了较为常见的聊天对话，腾讯混元大模型设置了一个名为“灵感发现”的栏目，主打内容创作。根据指引，员工可以让 AI 生成小红书种草文案、短视频脚本、社群营销文案以及朋友圈文案等；还支持不同风格的“绘画”模式，比如赛博朋克风、像素插画风、马赛克风格等。

中石油大学与阿里云共建勘探开发基础大模型。今天，中国石油大学（北京）（以下简称“中石大”）与阿里云在 2023 中国油气人工智能科技大会上签署了战略合作协议。基于中石大在油气科学领域的积累与阿里云在云计算、大模型等领域的技术优势，双方将为能源行业联合孵化场景化解决方案，通过“云+AI”助力石油石化领域由数字化向智能化转型。结合中石大在勘探开发领域的科研成果和经验沉淀，双方将联合打造勘探开发领域的行业模型算法库，助力缩短勘探时间，提升作业效率。同时，依托阿里云通义千问大模型，共建勘探开发领域基础大模型。

伦敦证交所正与微软等合作开发 AI 模型。伦敦证券交易所（LSE）正与微软和几家银行合作，创建定制的生成式 AI 模型，此举表明金融服务业正寻求在不暴露专有数据的情况下利用这项技术。伦敦证券交易所正与微软合作创建“定制的大型语言模型”。“我们正在与客户进行对话，讨论他们是否有机会使用我们现有的大量数据，以安全的专有方式将这些数据与他们的数据融合在一起。”伦敦证券交易所首席执行官 David Schwimmer 称，银行对创建自己的生成式 AI 模型感兴趣，因为他们希望确保自己的数据不会被用到任何其他大型语言模型。

中国气象局研讨 AI 气象预报大模型建设。据中国气象局微信公众号发文，8 月 3 日，中国气象局围绕 AI 气象预报大模型建设展开研讨，分析当前大模型发展面临的形势，进一步明确建设方向、目标和实现路径。局党组书记、局长陈振林强调，要深刻理解建设大模型对助力气象高质量发展的重要战略意义，坚定信心，狠抓落实，有序推动大模型建设。会议指出，当前 AI 技术应用在气象预报领域已呈现出快速发展趋势。建设 AI 气象预报大模型，有利于开辟气象高质量发展新赛道，有利于为实现气象科技能力现代化注入新动能，有利于加快推动气象业务技术体制改革，气象部门要提高政治站位，深刻理解建设大模型对助力气象高质量发展的重要战略意义，加快发展 AI 气象应用。

韩国互联网巨头 Kakao 将推 AI 大模型。据站长之家报道，Kakao 计划将 KakaoTalk 打造成一个综合的人工智能平台，结合通信、社交网络和商业功能。他们计划吸引超过 1000 万日活跃用户，占韩国人口的五分之一。为了实现这一目标，Kakao 将发布一个超大型生成式 AI 模型 KoGPT2.0，该模型号称与 Facebook 的 GPT 模型相媲美。这将为用户提供更高质量的智能助手和交流工具。

同花顺：目前公司大模型还处于备案阶段。根据同花顺 8 月 3 日发布的投资者关系活动记录表，目前公司的大模型还处于备案阶段，具体应用上线还需监管部门审批通过。未来大模型将优先应用于金融行业，同时将深入运用到公司现有的产品和服务体中，深度赋能 i 财财、同花顺 APP、iFind、爱基金以及面向 B 端客户等应用场景。同花顺较早高强度深耕 AI 技术在

金融领域的运用，因此在金融领域的垂类模型中，公司在数据、技术、业务场景、客户理解等方面拥有深厚积累，具备一定的优势。

字节对话类 AI 产品 Grace 已开启内测。据 Tech 星球 8 月 5 日报道，字节跳动代号 Grace 的对话类 AI 项目已开启测试，测试界面曝光，测试网站已搭建好，或为后期官网。Tech 星球向字节跳动方面进行求证，字节跳动相关负责人回复称：“目前产品还不成熟，还在内测阶段，与国外的领先模型相比，还有较大差距。”据悉，Grace 的测试网站为 gracebot.cn，跳转该网址后，可以看到网站标题显示“Grace-你的 AI 伙伴”等字样，目前支持邮箱、Gmail、飞书账号登录，但需得到授权。登录成功后，即可跳转至 Grace 的聊天对话界面。除了文生文对话形式外，Grace 还支持文生图。

智源最强开源语义向量模型 BGE 上线。据智源研究院微信公众号近日发文，为加快解决大模型的制约问题，近日，智源发布最强开源可商用中英文语义向量模型 BGE (BAAI General Embedding)，在中英文语义检索精度与整体语义表征能力均超越了社区所有同类模型，如 OpenAI 的 text embedding 002 等。此外，BGE 保持了同等参数量级模型中的最小向量维度，使用成本更低。将 LangChain 与智源 BGE 结合，可以轻松定制本地知识问答助手，而不需要花较高的成本训练垂类大模型。本次 BGE 模型相关代码均开源于 FlagOpen 飞智大模型技术开源体系旗下 FlagEmbedding 项目，一个聚焦于 Embedding 技术和模型的新版块。智源研究院将持续向学术及产业界开源更为完整的大模型全栈技术。

图 15: 智源研究院 FlagOpne 飞智大模型技术开源体系



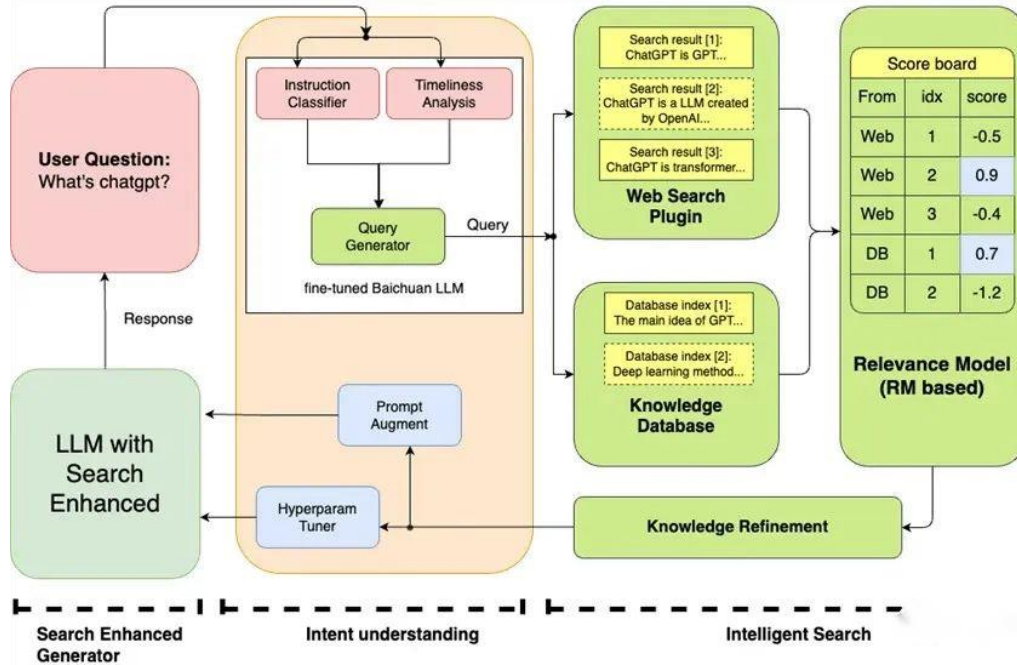
资料来源: 智源研究院, 中国银河证券研究院

元象开源百亿参数通用大模型。据 36 氪报道，AI 驱动的 3D 内容生产与消费一站式平台元象 XVERSE 宣布开源其百亿参数的高性能通用大模型 XVERSE-13B，可免费商用。据悉，基于 XVERSE-13B 生成的服务包括文本生成、自动化写作、数据分析、知识问答等方面。元象 XVERSE 于 2021 年在深圳成立，2022 年 3 月完成 A 轮、A+ 轮共计 1.2 亿美元融资，投资方包括腾讯、高榕资本、五源资本、高瓴创投、红杉中国、淡马锡、CPE 源峰等。

百川智能发布 530 亿参数的闭源大模型。近日，百川智能发布 530 亿参数规模的闭源大模型 Baichuan-53B，这是百川智能发布的第三个大模型，主要服务 B 端行业，预计下个月将会开放 API 等相关组件。百川智能 4 月 10 日成立后，6 月 15 日发布了 70 亿参数规模开源

模型 Baichuan-7B，7月11日发布了130亿参数规模大模型 Baichuan-13B，到今天，Baichuan-53B 已经是其发布的第三个模型。百川智能创始人、CEO 王小川透露，季度末，百川智能将发布千亿参数规模的模型对标 GPT。

图 16: Baichuan-53B 的搜索增强系统图示



资料来源: 百川智能, 中国银河证券研究院

广汽推出 AI 大模型平台。《科创板日报》8 日讯，广汽近日正式推出 AI 大模型技术的最新研发成果——广汽 AI 大模型平台，该平台号称“汽车行业首创的大模型平台技术”，将在近期搭载于高端智能轿跑昊铂 GT 亮相。广汽 AI 大模型平台聚合了多种 AI 大模型，能调用通用大模型的能力、基于智能汽车的应用场景构建专用模型。

百度智能云联合金蝶推出“金蝶云·苍穹 GPT”。8 月 8 日，金蝶面向全球推出新一代企业管理产品——金蝶云·苍穹 GPT，定位为最懂管理的企业级大模型平台。本次发布的金蝶云·苍穹 GPT 基于百度智能云千帆大模型平台打造，千帆平台为金蝶云·苍穹 GPT 及其各个领域应用提供了大模型推理、微调以及算力资源服务，落地了财务分析、人才推荐、供应链风险分析、自然语言生成报表、流程以及低代码页面等高价值应用场景。

Stability AI 推出生成式 AI 编程产品 StableCode。8 月 8 日，Stability AI 宣布推出其首个用于编码的大型语言模型生成式 AI 产品 StableCode。该产品旨在帮助程序员进行日常工作，同时也为准备将技能提升的新开发人员提供了一个很好的学习工具。StableCode 通过使用三种不同的模型来帮助开发人员编程，提高效率，分别是基础模型、用于解决复杂编程任务的指令模型、为用户提供单行和多行自动补全建议的长上下文窗口模型。

小米大模型首次曝光。小米大模型 MiLM-6B 近日现身 C-Eval、CMMLU 大模型评测榜单。截至当前，MiLM-6B 在 C-Eval 总榜单排名第十、同参数量级排名第一，在 CMMLU 中文向大模型排名第一。得分方面，在 C-Eval 评估中，MiLM-6B 的平均分为 60.2。

图 17: CMMLU 中文向大模型排名

#	Model	Creator	Submission Date	Avg ▾	Avg(Hard)	STEM	Social Science	Humanities	Others
0	ChatGLM2	Tsinghua & Zhipu.AI	2023/6/25	71.1	50	64.4	81.6	73.7	71.3
1	GPT-4*	OpenAI	2023/5/15	68.7	54.9	67.1	77.6	64.5	67.8
2	AILMe-100B v2	APUS	2023/7/25	67.7	55.3	65.4	72.3	71.2	64
3	SageGPT-V0.2	4Paradigm	2023/7/25	66.6	61.1	67.9	76.6	66.9	54.9
4	SenseChat	SenseTime	2023/6/20	66.1	45.1	58	78.4	67.2	68.8
5	亦免	北京容联易通信息技术有限公司	2023/8/8	64.1	43.2	58.5	76.6	66.9	60.3
6	InternLM	SenseTime & Shanghai AI Laboratory (equal contribution)	2023/6/1	62.7	46	58.1	76.7	64.6	56.4
7	ChatGLM2-12B	Tsinghua & Zhipu.AI	2023/7/26	61.6	42	55.4	73.7	64.2	59.4
8	UniGPT	Unisound	2023/7/26	60.3	46.4	57.7	69.3	58	59
9	MiLM-6B	Xiaomi	2023/8/9	60.2	42	54.5	71.7	62.7	57.7
10	Qwen-7B	Alibaba Cloud	2023/7/29	59.6	41	52.8	74.1	63.1	55.2
11	BatGPT-15b-sirius-v2	SJTU & WHU	2023/8/4	57.4	36.9	50.5	72.1	60.7	53.3
12	Instruct-DLM-v2	DeepLang AI	2023/7/2	56.8	37.4	50.3	71.1	59.1	53.4
13	GS-LLM-Alpha	共生矩阵科技(深圳)有限公司	2023/7/26	55.6	35.8	47	69.7	61.8	52
14	DFM2.0	AI-Speech & SJTU	2023/7/10	55.4	38.3	47.5	64.6	58.7	58.2
15	XVERSE-13B	XVERSE Technology	2023/8/6	54.7	33.5	45.6	66.2	58.3	56.9
16	HITsz-Lychee-Base-11B-V0.1	HITsz (哈工大深圳)	2023/8/6	54.7	44	50.8	61.3	57	53.8
17	ChatGPT*	OpenAI	2023/5/15	54.4	41.4	52.9	61.8	50.9	53.6
18	Claude-v1.3*	Anthropic	2023/5/15	54.2	39	51.9	61.7	52.1	53.7
19	TeleChat-E	China Telecom Corporation Ltd.	2023/7/4	54.2	41.5	51.1	63.1	53.8	52.3
20	CPM	ModelBest	2023/7/5	54.1	37.5	47.2	62.7	58.4	54.8
21	Baichuan-13B	Baichuan	2023/7/9	53.6	36.7	47	66.8	57.3	49.8
22	DLM-v2	DeepLang AI	2023/7/2	53.5	35.3	47	64.7	56.4	52.1
23	InternLM-7B	Shanghai AI Laboratory & SenseTime	2023/7/5	52.8	37.1	48	67.4	55.4	45.8
24	ChatGLM2-6B	Tsinghua & Zhipu.AI	2023/6/24	51.7	37.1	48.6	60.5	51.3	49.8
25	llamav2_cn_bairong_chater	百融云创	2023/7/28	50.5	35.5	47.4	60.6	50.5	47.1
26	EduChat	ECNU	2023/7/18	49.3	33.1	43.5	59.3	53.7	46.6
27	AndesLM-13B	AndesLM	2023/6/18	46	29.7	38.1	61	51	41.9
28	Claude-instant-v1.0*	Anthropic	2023/5/15	45.9	35.5	43.1	53.8	44.2	45.4
29	MiLM-1.3B	Xiaomi	2023/8/9	45.8	29.9	38	59.5	49.8	43.4

资料来源: 智源研究院, 中国银河证券研究院

英伟达发布 CALMAI 模型 可模拟 50 亿个人体动作。《科创板日报》11 日讯, 英伟达近日和以色列理工学院、巴伊兰大学和西蒙弗雷泽大学合作, 发布了一篇关于 CALMAI 模型的技术论文。英伟达表示 CALM 的全称是条件对抗潜在模型 (Conditional Adversarial Latent Models), 用于训练定制虚拟角色。英伟达表示, 在真实世界训练 10 天, 相当于在模拟世界里训练 10 年时间。CALMAI 模型在训练之后, 可以模拟 50 亿个人体动作, 涵盖行走、站立、坐姿、跑步、用剑战斗等人类动作。

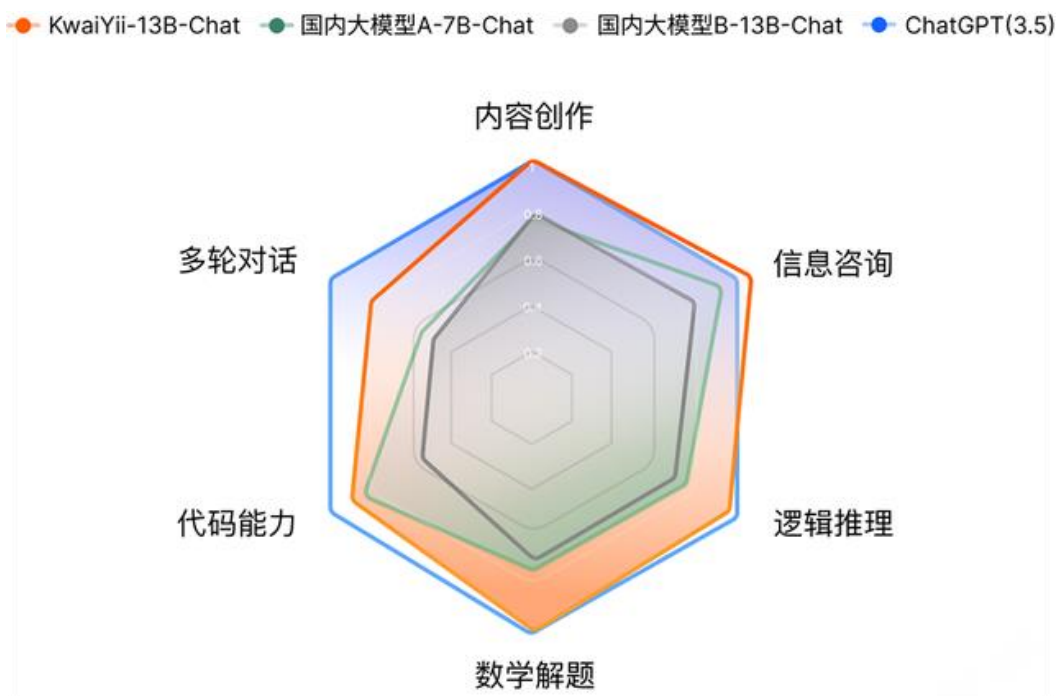
讯飞正式发布星火认知大模型 V2.0。8 月 15 日, 科大讯飞正式发布了星火认知大模型 V2.0, 重磅升级两大能力: 代码能力和多模态能力。代码能力方面, 讯飞星火 2.0 进行 5 个维度的升级, 包括代码生成、代码补齐、代码纠错、代码解释、单元测试生成。多模态能力方面, 讯飞星火 2.0 在图像描述、图像问答、识图创作、文图生成、虚拟人合成等方面全新升级。讯飞星火还发布了多项能力应用新成果, 包括星火语伴 2.0、星火教师助手、星火汽车助理等等。讯飞大模型正在行业和企业中落地, 此次科大讯飞联合华为推出了星火一体机, 采用 5 种定制优化模式 10+ 可选场景包, 支持训练和推理。

中国中医科学院与阿里云共创中医药行业大模型。据阿里云微信公众号发文，中国中医科学院与阿里云近日在北京签署合作协议，双方将共同启动数智中医创新联合实验室，并联合打造中医药行业大模型，推动中医药行业数智化。依据合作协议，双方将共同建设、运营“数智中医创新联合实验室”，面向行业联合打造中医药行业大模型——岐黄助手，围绕中医药数字化、产业化，共同搭建、推广、运营涵盖中医药服务业、中医药产业的一站式数字化云服务平台。

普华永道为员工提供私人生成式 AI 工具。据 CIO Dive 8 月 15 日报道，普华永道近期正在试用一款内部生成式 AI 工具 ChatPwC，可作为员工的对话式人工智能助手。该工具采用了 OpenAI 的技术，并根据相关数据进行了微调。目前，约有 1000 名员工可以使用该工具，公司计划在今年夏天重点扩大使用范围。

快手推出“快意”AI 大模型。据 IT 之家报道，近日，快手自研大语言模型“快意”(KwaiYii)开启内测。据其官方 GitHub 主页称，快意的 13B 规模系列模型 KwaiYii-13B 中，预训练模型 KwaiYii-13B-Base 在多个知名大模型排行榜上位居前列，对话模型 KwaiYii-13B-Chat 支持内容创作、信息咨询、数学逻辑、代码编写、多轮对话等任务。人工评估结果表明，KwaiYii-13B-Chat 在内容创作、信息咨询和数学解题上接近 ChatGPT 3.5 版本同等水平。

图 18: 多个 AI 大模型的人工评估结果



资料来源: AI 前沿观察者, 中国银河证券研究院

清华系团队开源多模态生物医药大模型。据清华大学智能产业研究院微信公众号 8 月 18 日发文，清华大学智能产业研究院 (AIR) 携手大模型初创公司水木分子开源可商用多模态生物医药百亿参数大模型 BioMedGPT-10B。据介绍，该大模型可用于提升药物研发各个环节的效率，包括新药立项评估、药物设计和优化、临床试验设计、适应症拓展等。该模型在生物医药领域的问答能力号称“比肩人类专家水平”，在多个生物医药问答基准数据集上实现了 SOTA (state-of-the-art model, 目前最先进的模型)，已成功通过了美国医师资格考试。据悉，该模型已于 8 月 14 日上线开源。此外，该团队一同上线了 BioMedGPT-LM-7B，号称是“首个可商用、生物医药专用的 Llama2 大模型”。

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。

阿里云等联合发布教育垂直大模型智海-三乐。8月21日，中国人工智能学会等联合主办的 Engineering 大讲堂暨“智行中国”第五期系列论坛在浙江大学举行。论坛上，浙江大学联合高等教育出版社、阿里云等发布了基于通义千问 7B (70 亿参数) 模型训练的智海-三乐教育垂直大模型，该模型已在阿里云灵积平台 (DashScope) 上线，对外提供 API 服务；智慧司法技术总师系统、浙江大学、上海交通大学与阿里云等联合发布了行业首个《法律大模型评估指标和测评方法 (征求意见稿)》，涵盖法律大模型能力体系、评估指标、测评方法、典型应用场景等内容，适用于法律大模型系统测试的设计和 implement 等任务场景，旨在推动法律大模型的研发、评测和应用的规范化。

澜舟科技推出 400 亿参数通用大模型孟子 GPT。8月27日，国内大模型创企澜舟科技正式发布参数量 400 亿的孟子 GPT-40B 通用大模型、参数量 70 亿的孟子 GPT-7B 金融大模型，并推出基于孟子 GPT 大模型打造的会议内容分析平台澜舟智会。孟子 GPT-40B 通用大模型现已开启邀测，中文能力更加突出，兼顾多语言能力，语料来自网页、百科、社交媒体、新闻以及澜舟科技与合作伙伴的一些高质量中文开源数据集。

图 19: 孟子 GPT 简介图示



资料来源：澜舟科技，中国银河证券研究院整理

航旅纵横发布首个民航大模型“千穰”。据新浪财经报道，8月25日，中航信移动科技有限公司航旅纵横团队发布了首个民航领域垂直大模型“千穰”。该模型将垂直领域海量民航数据与通用的视觉、语言、多模态和计算等大模型相融合，具备 10 亿、100 亿、300 亿等多种参数规模，可以用于旅客满意度分析、行业智能客服、多模态点评感知等服务。目前，“千穰”已在航旅纵横 APP 应用登陆，旅客可通过与其对话获取信息服务。

3、OpenAI 旗下 ChatGPT 功能优化

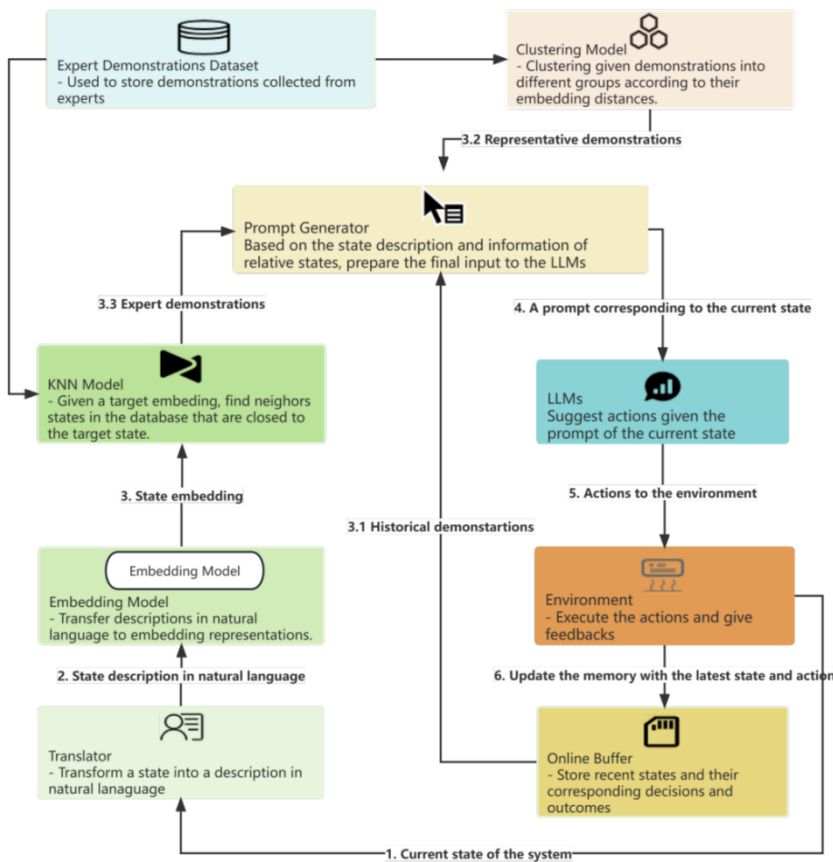
ChatGPT 进入主流科学搜索引擎。据 Nature 8月2日报道，AI 驱动的对话式聊天机器人即将进入更多主流科学搜索引擎。8月1日，荷兰出版巨头 Elsevier 为其 Scopus 数据库的部分用户发布了由 ChatGPT 支持的 AI 界面。英国公司 Digital Science 亦宣布对其 Dimensions 数据库进行 AI 大语言模型助手的封闭试验。美国科睿唯安公司称在努力将大语言模型纳入其 Web of Science 数据库。

ChatGPT 更新提示词示例等新功能。据品玩报道，OpenAI 今天宣布，从下周开始，将会推出一系列小更新来改善 ChatGPT 的使用体验。更新之后，ChatGPT 将会增加提示词示例，可以帮助新用户了解如何使用 ChatGPT。还有新的回复建议功能，可以根据聊天内容给用户提出对话建议。另外更新之后 Plus 用户将会默认使用 GPT-4 版本进行对话，而不是目前默认的 GPT-3.5 版本。此外，ChatGPT 还将支持上传多个文件，用户可以通过代码解释器来实现这个功能。更新之后的 ChatGPT 还会加入登录状态保持和键盘快捷键功能。

OpenAI 在中国申请注册“GPT-5”商标。据 IT 之家报道，国家知识产权局商标局官网显示，欧爱运营有限责任公司（OPENAI OPCO, LLC）已于 7 月下旬申请注册了两枚“GPT-5”商标，国际分类分别为 9 类和 42 类（科学仪器、设计研究），商标状态均为申请中。此前美国专利商标局（USPTO）信息显示，OpenAI 已于 7 月 18 日申请注册“GPT-5”的商标。从商标信息来看，GPT-5 将提供文本生成、自然语言理解、语音转录、翻译、分析等功能。

微软新研究支持 GPT-4 对空调进行工业控制。据机器之心报道，最近，微软亚洲研究院提出可以将 LLM 用于工业控制，而且仅需少量示例样本就能达成优于传统强化学习方法的效果。该研究尝试使用 GPT-4 来控制空气调节系统（HVAC），得到了相当积极的结果。研究团队设计了一种机制来从专家演示和历史交互挑选示例，还设计了一种可将目标、指示、演示和当前状态转换为 prompt 的 prompt 生成器。然后，再使用生成的 prompt，通过 LLM 来给出控制。

图 20: 利用 GPT-4 控制 HVAC 系统流程



资料来源: arXiv, 中国银河证券研究院

OpenAI 宣布使用 GPT-4 进行内容审核。OpenAI 近日宣布使用 GPT-4 进行内容策略开发和内容审核决策，实现更一致的标注，为策略细化提供更快反馈循环，并减少人类的参与。使用 GPT-4 的内容审核系统可以更快地迭代策略更改，开发和定制内容策略的过程从几个月缩短到几个小时。GPT-4 还能够解释长内容策略文档中的规则和细微差别，并立即适应策略更新，从而实现更一致的标注。或者，为了大规模处理大量数据，OpenAI 可使用 GPT-4 的预测来微调一个更小的模型。OpenAI 相信，这为数字平台的未来提供了一个更积极的愿景，AI 可以根据特定平台的策略帮助调节在线流量。任何拥有 OpenAI API 访问权限的人都能用这种方法来创建自己的 AI 辅助审核系统。

OpenAI 推出 GPT-3.5 Turbo 微调功能。OpenAI 于当地时间 8 月 22 日宣布推出 GPT-3.5 Turbo 微调功能和 API 更新，并将于今年秋天推出 GPT-4 微调版本。开发人员可以自定义更适合其所需用例的模型，并大规模运行这些自定义模型。早期测试表明，GPT-3.5 Turbo 的微调版本在某些细分任务上可以匹配甚至超越 GPT-4。OpenAI 称，传入与传出微调 API 的数据不会被用来训练其他模型。此外，OpenAI 更新了 GPT-3 模型，推出 babbage-002 和 davinci-002 作为 GPT-3 基础模型 ada、babbage、curie 和 davinci 的替代品，这些模型将于 2024 年 1 月 4 日关闭。

ChatGPT 自定义指令功能全面开放。OpenAI 于 8 月 22 日在其社交媒体 X 上宣布 ChatGPT 自定义指令功能向欧盟和英国开放，该功能可以让用户对机器人设置一些特定的规则和条件，来控制机器人的行为和输出。此前，OpenAI 曾于 8 月 10 日宣布该功能面向除欧盟和英国外所有用户免费开放。

OpenAI 推出企业版 ChatGPT。8 月 29 日，OpenAI 于官方博客发布了企业版 ChatGPT，该版本提供企业级安全和隐私，包括静态和传输中数据加密、符合 SOC 2 标准和承诺不会将数据用于训练模型；适合大规模部署的功能，包括具有批量会员管理功能的管理控制台、单点登录、域名验证和分析仪表盘；并且是迄今为止最强大的版本，提供无限制的 GPT-4 访问、速度提升 2 倍、无限制高级数据分析、可用于 4 倍输入和文件的 32k Tokens 上下文窗口、可共享的聊天模板和免费的 API 点数。其早期客户包括普华永道、Canva 可画和雅诗兰黛等行业巨头。OpenAI 称，目前正在开发更多功能，后续将更新自定义功能、为小型团队提供自助式商务服务、更强大的数据分析和浏览版本、更多垂直领域解决方案等。

4、全球芯片研发取得新进展，与 AI 发展息息相关

(1) 国内动态

国芯科技：重点布局边缘计算芯片和 AI 芯片。国芯科技在互动平台回复投资者问题时表示，边缘计算芯片和 AI 芯片是公司重点布局的方向，公司在边缘计算芯片方面已开展了二代芯片的研发和产业化应用。国芯科技在 AI 芯片的布局包括：（1）在云 AI 芯片方面，围绕 AI 服务器应用，开发系列化芯片产品，目前已有芯片产品包括 Raid 控制芯片和超高速加解密处理芯片，未来将继续开展 AI 服务器芯片组的开发和应用；（2）在端 AI 芯片方面，已在生物特征识别领域推出了包括轻量级 AI（卷积协处理器）和安全处理的 SoC 芯片，实现指纹和人脸识别应用，未来将在现有基础上继续发展生物特征识别领域的高性能端 AI 芯片；（3）正在开发 RISC-V CPU 内核系列，目前正在开展将 AI 引擎加入到 RISC-V CPU 内核中，形成具有 AI 引擎的 CPU 内核系列；（4）正在为客户 AI 芯片提供定制服务，目前已有多个 AI 芯片定制服务的在手订单；（5）将继续和合作伙伴合作，联合开展 GPU 和 NPU 等技术的研发和应用。

鸿海向 AI 芯片创企耐能投资约 1400 万美元。据界面新闻报道，鸿海 8 月 3 日公告，子公司 Icreate Investments 取得 AI 芯片初创公司耐能智慧（Kneron）股份，投资金额约 1400 万美元（约合新台币 4.26 亿元）。鸿海表示，此次现金增资为长期投资。耐能主攻神经网络处理器及各种图像识别功能，先前耐能募集资金，研发下一代边缘计算 AI 芯片，或可支持鸿海的电动车开放平台。

RISC-V 芯片创企进迭时空获数亿元 A 轮融资。近日，杭州 RISC-V 芯片创企进迭时空已完成数亿人民币 A 轮融资，本轮融资由联想创投、君联资本领投，北京国管顺禧基金、经纬创投、耀途资本、万物资本、真格基金、海阔天空创投、趣合投资跟投，凡卓资本担任独家财务顾问。本轮融资将主要用于下一代高性能 RISC-V CPU 产品的研发及市场进一步拓展，加速 RISC-V 生态迭代。

耐能推出 KL730 芯片 支持轻量级 GPT 大语言模型。据耐能官方公众号近日发文，AI 初创公司耐能近日宣布推出 KL730 芯片。该芯片以实现 AI 功能为目的，集成了车规级神经处理单元（NPU）和图像信号处理器（ISP），可无缝接入图像、视频、音频和毫米波等数字信号，支持多类行业的 AI 应用开发。据称，KL730 能效比相较于耐能过往的芯片提升了 3 至 4 倍，且比主要行业同等产品提高了 150%~200%，提供每秒 0.35 - 4 tera 有效计算能力，扩展了支持最先进的轻量级 GPT 大语言模型（如 nanoGPT）的能力。据称，KL730 定位改变 AIoT 领域的安全性，从而使用户能够部分或完全离线地在终端运行 GPT 模型。该芯片配合耐能的私有安全边缘 AI 网络 Kneo，让 AI 运行在用户的边缘设备上，从而让他们更好地保护数据隐私。

图 21: KL730 性能表现测试



资料来源：耐能，中国银河证券研究院

字节发布自研芯片，联合英伟达开源多媒体框架。8 月 22 日，火山引擎于“视象新生”主题火山引擎视频云&AIGC 大会上发布了自研视频编解码芯片，可以根据视频场景，提供不同的算力支持；携手英伟达发布多媒体处理框架 BMF（Babit MultiMedia Framework），并全面开源，包含 BMF 框架层整体开源、9 个开箱即用案例、20+API 调用范例，旨在帮助企业 and 开发者将各类视频的原子能力与方案做到优势互补、协同工作和快速落地。

昆仑芯首批加入 AI 芯片适配基地建设。在 8 月 19 日的 2023 中国算力大会上，宁夏发布全国首个由政府主导建设的 AI 芯片适配基地。大会期间，昆仑芯携两代芯片系列产品亮相，成为首批加入 AI 芯片适配基地建设的头部企业。昆仑芯前身为百度智能芯片及架构部，于 2021 年 4 月完成独立融资，首轮估值约 130 亿元，专注打造拥有通用性、易用性和高性能的通用 AI 芯片。目前，昆仑芯已实现两代通用 AI 芯片系列产品的量产及落地应用，在互联网、智慧工业、智慧交通、智慧金融等领域均有规模部署。

(2) 国际动态

报告称英伟达 AI GPU 产能在 2024 年增加 2.5 倍。《科创板日报》7 日讯，研究机构 Aletheia 报告指出，预估 AI 服务器市场规模将在 2024 年翻倍、2025 年达到 1350 亿美元，是 2022 年规模的 4.5 倍。其中，GPU 在 AI 服务器中成本占比最高，Aletheia 认为，GPU 市场到 2025 年将较 2023 年成长 3 倍。英伟达为了满足 AI 服务器市场庞大的需求，将着手大幅增加 AI GPU 产能，预估英伟达 AI GPU 产能在 2024 年增加 2.5 倍；公司数据中心业务营收在 2023-2026 年的年复合成长率将达 85%。

图 22: Nvidia H100 架构图



资料来源: Nvidia, 中国银河证券研究院

英伟达 H100 在 eBay 网站上涨到 4.5 万美元。据第一财经报道，目前 GPU 价格仍在上涨，作为人工智能基础设施的“硬通货”，海外已经有创业企业开始利用 GPU 进行抵押融资。这种新兴的融资途径凸显了此类硬件在资本密集型人工智能“军备竞赛”中的价值。查询 eBay 网站发现，H100 的售价已经高达 4.5 万美元，这较今年 4 月份 4 万美元的价格涨幅超过 10%，而且货源较上半年也显著减少。英伟达内部人士告诉第一财经记者，电商渠道的价格不受英伟达管控，“属私人在网上转售”。

三星加入 AI 芯片竞赛，将为 Groq 生产 4 纳米加速器芯片。据界面新闻报道，美国 AI 芯片公司 Groq 于 8 月 14 日晚宣布，三星晶圆代工部门（三星 Foundry）泰勒工厂将为其生产 4 纳米制程的 AI 加速器芯片。Groq 创始人兼 CEO Jonathan Ross 表示，将通过三星 4 纳米晶圆代工工艺实现 Groq 的技术进步。Groq 由谷歌前员工 Jonathan Ross 在 2016 年创立，主要针对

云计算市场开发 AI 芯片。Groq 主要成员曾参与研发谷歌的张量处理单元（Tensor Processing Unit, TPU），公司成立后推出首个云端推理芯片 GroqChip，可以实现 16 个芯片间互连并配置 230 MB SRAM，算力可达 750 TOPS。2021 年 4 月，Groq 获 D1 Capital、老虎环球基金等机构 3 亿美元融资。

沙特阿拉伯采购至少 3000 块 H100 芯片。据外媒报道，沙特阿拉伯和阿拉伯联合酋长国正在采购高性能英伟达芯片，以加入全球 AI 军备竞赛。据知情人士透露，沙特阿拉伯采购了至少 3000 块 H100 芯片，总价值约 1.2 亿美元；阿联酋获得了数千个英伟达芯片的使用权，并开发了自己的开源大语言模型 Falcon。

5、无人机应用

浙江机器人产业集团发布垂起固定翼无人机-机器人自动换电机巢。8 月 25 日，全国首台垂起固定翼无人机-机器人自动换电机巢新品发布会暨“科创中国·宁波”无人机产业趋势分享会在余姚市机器人小镇成功举行。在无人机技术快速发展和广泛应用的今天，此次全国首台垂起固定翼无人机-机器人自动换电机巢的正式发布吸引了行业广泛关注的目光。该产品由浙江机器人产业集团联合极臻智能、远度科技共同开发，集成人工智能、机器视觉和自动化控制等领域领先技术，实现了高效、智能的无人机换电过程。其创新研发的国内首创蛙跳式巡检模式，多旋翼无人机动态协同算法、固定翼机械臂换电技术，不仅大幅提高了无人机飞行时间、作业半径和巡检效率，而且满足应急管理长时间巡检需求，突破应急响应新高度，真正实现无人机自动值守、巡检、回收，是一次重要的创新突破。

图 23：“科创中国·宁波”无人机产业趋势分享会



资料来源：浙江机器人产业集团，中国银河证券研究院

河北首次开展超高压输电线路“无人机+机器人”带电作业。“作业环境安全，启动智能机器人自主带电作业。”8 月 22 日，在河北石家庄，国网河北超高压公司无人机作业班成员郝顺杰操作一架大疆 T-50 无人机，将输电线路智能机器人吊装并放置在距地面 30 米的 500 千伏桂北 II 线导线上。机器人在导线上平稳行走，并移动至指定作业位置，仅用时 30 分钟就自主完成了缺陷处理。这是河北首次采用“无人机+机器人”的方式在 500 千伏超高压输电线路开展带电作业，采取这种科技手段，进一步提高了带电作业效率。“过去以人工带电作业方式处理这样的销钉缺陷，需要 5 名不同专业的人员进行协作，时间大约需要 1 小时 20 分钟，而采用

‘无人机+机器人’来操作，作业时间缩短 60%以上，还减少了人工登塔作业的安全风险。”参与机器人研发的国网河北超高压公司带电作业班班长刘洪吉说。

国内首家大型无人机考试中心落户北京定陵机场。近年来，无人机在社会生产生活各个方面发挥着重要作用，随着无人机相关法规政策的出台，无人机发展事业将步入正规化轨道。日前，中国民用航空局大型无人机考试中心在北京定陵机场运行启动，这标志着国内首家大型无人机考试中心正式落户定陵机场，将对无人机培训、考评的规范化发展起到重要的推动作用，并进入有法可依的新阶段。

IEC 首个电力无人机国际标准立项。科技日报记者 8 月 4 日从中国电科院获悉，由该院牵头发起的 IEC（国际电工委员会）63439-2-1“架空电力线路巡检用无人机通用技术要求”正式获批立项。该标准是电力无人机领域首项 IEC 标准，对于全球范围内开展电力无人机方面的国际标准化工作具有重要里程碑意义。通过无人机执行自主巡检，可高频次、近距离、多角度自动拍摄电力线路杆塔及通道高清图像，及时排除缺陷隐患，保障线路安全稳定运行。此前，世界三大国际标准组织（ITU、IEC、ISO）在架空电力线路无人机巡检领域尚处空白。此项 IEC 标准将对架空电力线路巡检用无人机装备（产品）的抗电磁干扰性能、飞行性能、巡检功能等技术要求进行规定，为相关无人机装备（产品）的设计研发、生产制造、应用维护提供统一规范。该标准将对世界各国提升无人机装备质量，推动电网运维数字化转型和智能化升级等发挥重要指导作用。

6、京东方和视涯科技打入苹果头显供应链，为“Vision Pro”提供屏幕

据 The Information 报道，苹果公司目前正在测试京东方和视涯科技两家中国供应商制造的 Micro-OLED 屏幕，以便将其纳入其 MR 头显的未来型号中，实现成本降低、产量提高。

近年来，中国政府一直对发展 Micro-OLED 等先进显示技术给予支持，广州等主要制造中心的地方政府纷纷出台激励措施，鼓励本地企业开发此类技术。京东方财报显示，该公司去年获得了 54.59 亿的政府补贴，是前一年所获得补贴的两倍多。在当地政府的扶持和企业自身发展下，京东方于 2016 年进入苹果供应链，成为 MacBook 和 iPad 的液晶面板制造商，帮助苹果减少对三星技术的依赖。2020 年，京东方开始为 iPhone 12 供应更高端的 OLED 面板。但熟悉京东方的人士说，京东方生产的面板质量一直落后于三星和 LG 这两家主要的 iPhone 显示屏供应商。

而视涯公司成立于 2016 年，专门投资 2.75 亿美元用于生产 Micro-OLED 显示屏，目标是到 2020 年每年生产 2000 万块显示屏。The Elec 曾在 6 月份报道，视涯公司希望向苹果供应显示屏。

苹果公司如今测试京东方和视涯两家中国供应商制造的 Micro-OLED 屏幕，此举表明，苹果公司首席执行官 Tim Cook 做到了他在 2016 年许下的承诺——努力改善两国政府关系，帮助提升中国的技术实力。苹果公司使用更多中国供应商的举动还表明，中国帮助国内制造企业在新兴技术领域发展的经济政策使它从中获益。

Vision Pro 的显示屏采用了被称为 Micro-OLED 的先进技术，这项技术可以为头显周围环境提供超清晰的图像。Micro-OLED 的生产难度很大，因为它们是直接在硅晶圆上而不是玻璃上制造的，所以它们可以做得特别小、特别薄，但像素密度仍然很高。分析师认为，它们是 Vision Pro 中最昂贵的部件之一。

苹果公司在提升产量、降低成本与面临商业与声誉风险的多种考量中，目前正在测试京东方和视涯两家中国供应商制造的 Micro-OLED 屏幕，与中国供应商的合作未来可能进一步深化

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。

增强。此举对于苹果公司与中国供应商双方而言，具有成本效益、生产能力与速度、技术创新和研发合作等优势，对于苹果公司的全球竞争优势和市场地位以及国内先进制造商未来成长具有重要意义。

图 24: Apple Vision Pro



资料来源：智东西，中国银河证券研究院

7、讯飞星火 V2.0 突破代码能力，一个指令生成贪吃蛇游戏

8月15日，讯飞星火大模型迎来年内第二次重磅升级，V2.0来了！星火大模型 V2.0 主要升级点体现在代码能力、多模态交互上，科大讯飞董事长刘庆峰谈道，大模型的代码能力是连接数字世界的关键支撑，多模态是让通用人工智能赋能工业的各种具体场景，未来进入家庭的最重要核心能力。

星火大模型的应用落地也正在稳步向前，5月6日星火大模型发布以来，科大讯飞人工智能开放平台新增了70万个开发者团队；讯飞星火小助手6月9日发布至今，已经新增了7800多个用户自己创建的小助手；讯飞学习机5月销量同比上涨了超100%，6月上涨超200%……这一过程中，用户体验又再反向提升大模型的能力，使得星火大模型的语言理解、逻辑推理的测试得分不断提高。可以看出，通用人工智能时代的发展，正在让各行各业的人都能释放自己的想象力，并通过大模型的能力将其快速实现，每个人都拥有个人AI助手的愿景正越来越接近。

到了最重磅的代码能力展示环节，刘庆峰谈道，代码能力是认知智能大模型智慧提升的关键维度，在提升这一能力时可以同步提升长文本理解、逻辑推理、任务规划能力，以及帮助用户执行复杂任务、连接数字世界，还能降低数字经济的创业门槛。星火能快速将视频片段转化成GIF动图，并且动图也会相应的缩小文件大小，方便用户发送分享。星火生成经典的贪吃蛇小游戏，包含72行代码，并且每一段都给出了相关的注释。

基于OpenAI构建的代码能力公开测试集 HumanEval 和认知智能全国重点实验室构建的代码实用场景测试集 iflyCT-py，星火在python语言下，代码生成和代码补齐能力超过 ChatGPT，其它能力正在快速赶超 ChatGPT。

图 25: 星火代码大模型能力提升显著



资料来源: 科大讯飞, 中国银河证券研究院

8、亚马逊首将 LLM 落地电商, 拟推出 AI 详情页文案创作工具

8月10日消息, 据 The Information 报道, 亚马逊正在为其电商卖家用户推出一款 AI 工具, 用于为商品详情页面 (Listing) 撰写文案。这将是亚马逊将大语言模型 (LLM) 整合到电子商务业务中的首批实例之一。上周, 亚马逊 CEO 安迪·贾西 (Andy Jassy) 曾在季度财报电话会议上谈到, AI 是其公司未来几年计划的核心, 包括电子商务在内的多个业务部门都在开展生成式 AI 项目。

亚马逊此次推出的新功能会在卖家上架新商品时提示描述其产品的关键词。根据亚马逊卖家顾问 Jantzen Russell 向 The Information 分享的截图显示, 该工具会在卖家撰写商品详情页面时为其生成可供选择的标题、商品描述和商品要点等要素。亚马逊发言人称, 公司目前正在测试新的 AI 解决方案, 让卖家以最小的工作量创建能够吸引潜在客户注意力的商品详情页面。关于亚马逊在该商品描述工具中具体使用了哪些 LLM 的问题, 该公司发言人没有回答。

在亚马逊之前, 已有其他电商公司更早推出了 AI 辅助商家功能。电商软件公司 Shopify 曾在今年 2 月推出了一款由 ChatGPT 驱动的商品描述工具, 又在上个月推出了一款 AI 助手, 可以辅助卖家自动完成一些简单的工作。Jungle Scout 和 CommerceIQ 等初创公司也开始提供 AI 功能, 为亚马逊商品详情页面生成文案。

国内电商巨头淘宝天猫在 5 月发起“商家 AI 生态伙伴计划”, 计划开放 7 大商家经营场景, 其中包括向商家提供店铺名称、LOGO、详情页文案和商品主图等智能化功能, 帮助商家高效运营店铺。

亚马逊正积极推出一款 AI 工具, 为其电商卖家提供商品详情页面文案。但该工具目前仍处于内测阶段, 只向少数卖家提供, 存在一定的报错率。亚马逊也提醒卖家在使用时注意检查,

避免违反其产品刊登准则。作为全球最大的电商平台，亚马逊这一行为可能会促进智能电商规模化落地。

图 26: 亚马逊商品详情页面的七个组成要素



资料来源: Amazon, 中国银河证券研究院

(二) 投融资事件

表 12: 8 月企业投融资事件

企业	简介	融资情况	本轮投资方
陪跑兔科技	专注于人工智能应用开发的科技公司, 致力于为中小微企业提供先进、高效的智能化解决方案, 助力企业迈向智能化时代	完成 3000 万人民币天使轮融资	星柚资本
了然视觉	为客户提供汽车零部件生产过程提供智能化的解决方案, 如智能化的大部件识别、抓取, 不同零部件的缺陷、式样检测等	完成未透露金额 Pre-A 轮融资	金鹏佳投资
进进时空	致力于高性能 RISC-V 处理器核、高性能芯片、核心软件系统的创新与研发, 为行业提供高性能 RISC-V 处理器和软件系统。	完成数亿人民币 A 轮融资	联想创投, 君联资本等
澜码科技	基于大语言模型的驱动, 致力于重塑人机互动的模型, 实现智能驱动与决策。	完成数千万人民币 A 轮融资	IDG 资本, 联新资本, Atom Capital, 数联投资

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。

生数科技	积极打造世界领先的多模态大模型，探索生成式AI在艺术设计、游戏制作、影视后期、内容社交等场景的商业赋能。	完成数千万人民币天使+轮融资	锦秋基金
达奇月泉	仿生智能产品开发商，经营范围包括假肢、矫正器、人工智能软件、智能机器人等。	完成千万级人民币天使轮融资	中关村发展启航产业投资基金
粒界科技	致力于通过自研的实时图形视觉技术，实现智慧城市、工业生产和影视动漫实现数据采集、可视化和互动等问题。	完成千万级美元 B++轮融资	比亚迪
智子引擎	进行图文匹配、图文检索、图像描述生成、文本描述生成等多种任务的训练，根据输入的图片或文字，“元乘象 ChatImg”可以进行智能聊天、讲故事、写广告等工作。	完成数千万人民币天使+轮融资	中金资本，某头部投资机构，软通高科

资料来源: Wind, 36 氪, 中投网, 中国机器人网, 中国银河证券研究院

五、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。

插图目录

图 1: 8 月人工智能指数表现.....	3
图 2: 人工智能板块市场表现.....	4
图 3: LongBench 测评结果.....	12
图 4: 视觉概念转换器图像编辑演示.....	13
图 5: 不同训练数据集上 GedankenNet 的性能表现.....	13
图 6: 通过 CT 扫描对八个主要器官进行癌症和非癌症检测、分割和诊断.....	14
图 7: 基于学习奖励和人类评估的绩效比较.....	15
图 8: 所选 LLM 在 OPEN 上的能力分布雷达图。.....	16
图 9: Apogee 外骨骼演示.....	17
图 10: 无人机竞速.....	18
图 11: “乐安”诉讼智能机器人.....	18
图 12: 向机器人行业上下游厂商介绍全球载重量最大的 RV 减速机情况.....	19
图 13: 多功能配送机器人 t10.....	20
图 14: 2023 百度云智大会.....	21
图 15: 智源研究院 FlagOpne 飞智大模型技术开源体系.....	23
图 16: Baichuan-53B 的搜索增强系统图示.....	24
图 17: CMMLU 中文向大模型排名.....	25
图 18: 多个 AI 大模型的人工评估结果.....	26
图 19: 孟子 GPT 简介图示.....	27
图 20: 利用 GPT-4 控制 HVAC 系统流程.....	28
图 21: KL730 性能表现测试.....	30
图 22: Nvidia H100 架构图.....	31
图 23: “科创中国·宁波”无人机产业趋势分享会.....	32
图 24: Apple Vision Pro.....	34
图 25: 星火代码大模型能力提升显著.....	35
图 26: 亚马逊商品详情页面的七个组成要素.....	36

表格目录

表 1: 8 月份人工智能板块涨幅前十.....	3
表 2: 8 月份人工智能板块跌幅前十.....	4
表 3: 人工智能主题基金一览.....	4
表 4: 国内人工智能上市公司近况一览.....	5
表 5: 海外人工智能上市公司近况一览.....	6
表 6: 数据要素最新新闻及政策.....	6
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.....	16
表 12: 8 月企业投融资事件.....	36

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。

评级标准

行业评级体系

未来6-12个月，行业指数相对于基准指数（沪深300指数）

推荐：预计超越基准指数平均回报20%及以上。

谨慎推荐：预计超越基准指数平均回报。

中性：预计与基准指数平均回报相当。

回避：预计低于基准指数。

公司评级体系

未来6-12个月，公司股价相对于基准指数（沪深300指数）

推荐：预计超越基准指数平均回报20%及以上。

谨慎推荐：预计超越基准指数平均回报。

中性：预计与基准指数平均回报相当。

回避：预计低于基准指数。

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。

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	liyongyang_yj@chinastock.com.cn
北京地区：	田薇	010-80927721	tianwei@chinastock.com.cn
	唐嫚羚	010-80927722	tangmanling_bj@chinastock.com.cn