



商汤集团股份有限公司

2023中期业绩简报

坚持原创，让AI引领人类进步

2023.08



画面由商汤秒画生成

01 概览

-

2023 年上半年，大模型引领 AI 行业变革



“日日新”大模型体系
研发突破



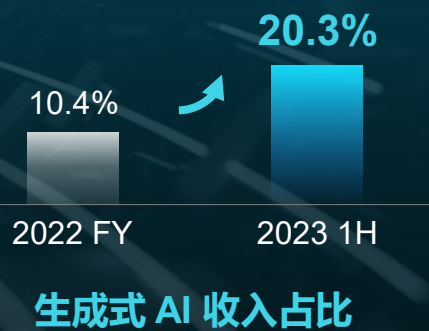
生成式 AI
产品族



生成式 AI 业务
三位数强劲增长

14.3 亿元

2023 1H 集团营业收入
(同比 +1.3%)



23.9 亿元 (↓7%)

2023 1H 经调整后净亏损²
(同比收窄)

备注：
1. 截至2023年6月30日数据
2. 优先股及其他金融负债的公允价值亏损、以股份为基础的薪酬开支、对优先股股东的股份支付费用和按公允价值计量且变动计入损益的金融资产的公允价值利得/(亏损)

前瞻布局 SenseCore 大装置，为大模型提供丰沛算力

全国最大
上海临港AIDC

~30,000 GPU¹
(vs. 2023年3月: ↑ 11%)

~6,000 P¹
(vs. 2023年3月: ↑ ~20%)
峰值算力不断扩充

算力节点拓展
广州、重庆、深圳、福建



备注：
1. 截至2023年8月28日数据

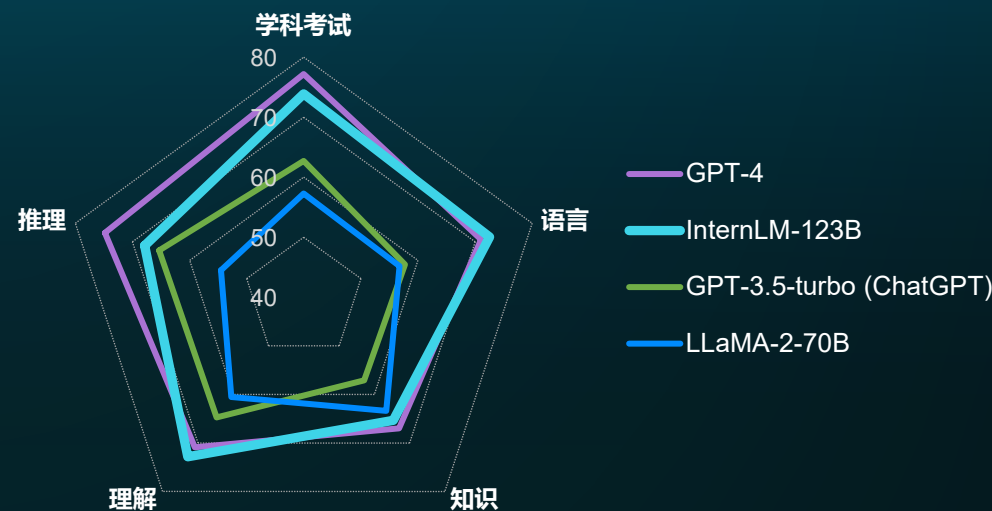
日日新大模型体系下，InternLM-123B模型完成训练，性能测试排名全球第二



在全球51个知名评测集测试中排名全球第二

排名	评测集名称	InternLM-104B	InternLM-123B	ChatGPT	GPT-4	LLaMA-2-70B
综合考试	MMLU	64.3	72.9	69.1	83.0	69.8
	C-Eval	54.4	67.5	52.5	69.9	50.1
	AGIEval	37.1	57.8	39.9	55.1	40.0
	ARC-c	76.6	90.2	83.7	93.6	78.3
	ARC-e	85.5	91.9	88.9	95.4	85.9
知识问答	CommonSenseQA	70.5	88.5	80.2	88.3	78.3
	NaturalQuestions	29.6	36.8	27.0	40.4	34.2
阅读理解	C3	62.8	95.5	85.6	95.1	79.0
	CMRC	40.0	62.9	40.1	42.6	43.7
	RACE(Middle)	83.4	94.9	85.6	93.7	81.6
	RACE(High)	79.2	91.7	80.3	91.0	79.9
	LAMBADA	77.7	87.0	57.5	65.5	78.9
推理	WinoGrande	68.7	86.4	58.8	78.9	69.8
	StoryCloze	78.6	99.8	97.2	99.5	89.7
	HellaSwag	81.3	91.8	79.5	91.4	82.3
	StrategyQA	73.0	81.7	66.1	79.7	74.7
	PIQA	82.2	86.3	81.7	89.2	82.5
	SIQA	62.6	82.4	72.4	70.1	64.8
	BigBench-Hard	59.5	71.1	70.1	86.7	64.9
数学	GSM8K	52.7	76.4	78.2	91.4	63.3

与GPT-4、GPT-3.5-turbo、LLaMA-2-70B的比较



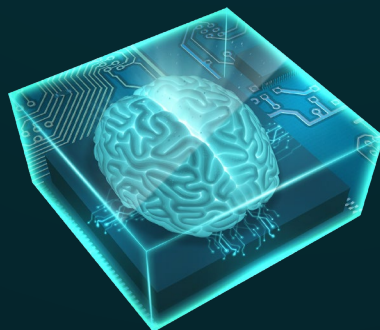
生成的内容更加准确可靠

具备自主反思和修正错误能力

代码解释器和插件调用能力升级

可灵活搭建AI智能体应用

超越 GPT-4 之路 —— 23 年内研发规划蓝图



千亿参数 基模型训练

训练know-how迁移

稳定大算力支持

高价值数据输入

算法“配方”实验

百亿
参数模型 × 100+
训练次数

十亿
参数模型 × 1000+
训练次数

数据采集和清洗

业内
领先
总数据储备

10万亿+
Token
预计年底
高质量数据储备

超大规模GPU集群



以大量高质量数据训练出最具价值的基模型



超越 GPT-4 之路 —— 23 年内研发规划蓝图

104B

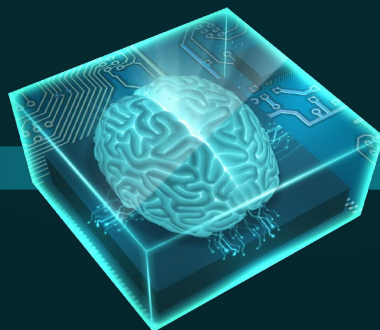
参数

2023 Q2

123B

参数

2023 Q3



每季度
迭代进化

更强大的
基模型

训练know-how迁移

千亿参数
基模型训练

稳定大算力支持

算法“配方”实验

百亿
参数模型

x

100+
训练次数

十亿
参数模型

x

1000+
训练次数

500人精英团队

+

大量高质量数据

+

10,000卡算力

高价值数据输入

数据采集和清洗

业内
领先

总数据储备





10万亿+
Token

预计年底
高质量数据储备

超大规模GPU集群



开源InternLM-7B —— 全球性能最好的轻量级基模型

	 InternLM-7B	 LLaMA2-7B	 LLaMA2-13B	 ChatGLM2-6B
MMLU ¹	52.0	46.8	55.0	46.0
AGIEval ¹	43.7	21.8	30.9	39.0
C-Eval ¹	54.1	32.5	41.4	52.6
SuperClue-Opt ²	53.9	未上榜	34.3	53.8
FlagEval ³	0.618	0.463	0.495	0.362

数据来源:

1: OpenCompass大语言模型评测榜单 (202308)

2: SuperClue-Opt开源模型榜单 (202307)

3: 智源研究院有监督微调榜单 (202308)

即将开源性能更强基模型InternLM-20B

更优秀的
工具调用能力

...

模型大小适中
运行成本较低

...

适用AI智能体
等应用构建

丰富的开源体系 长久的开源历史

OpenMM Lab

2018开源
Github星标 87,000+

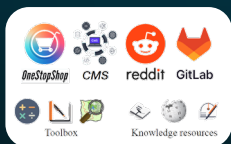
OpenCV Lab OpenDI Lab

- 决策智能
- 大语言模型
- 数据平台
- 高性能训练和推理框架
- AI智能体框架

 商汤 |  大装置

“生产力”大模型，将成为“AI操作系统”和“流量入口”

“聊天模型”无法实现的应用层技术



工具调用

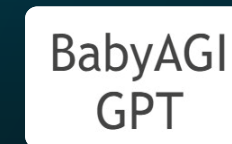


智能体交互



机器人控制

海外只有GPT-4，可支撑“生产力”应用



“..... are driven by GPT-4.....”

能准确调用外部工具的大模型，才谈得上“生产力”

大模型调用API的错误率	GPT-4	LLaMA	StarCoder	CodeGen
错误率	19%	61%	68%	93%
- API选择错误	0%	22%	22%	30%
- API参数设定错误	14%	32%	23%	63%
- 代码不可执行	5%	7%	23%	0%

数据来源：《On the Tool Manipulation Capability of Open-source Large Language Models》

“AI操作系统”和“流量入口”

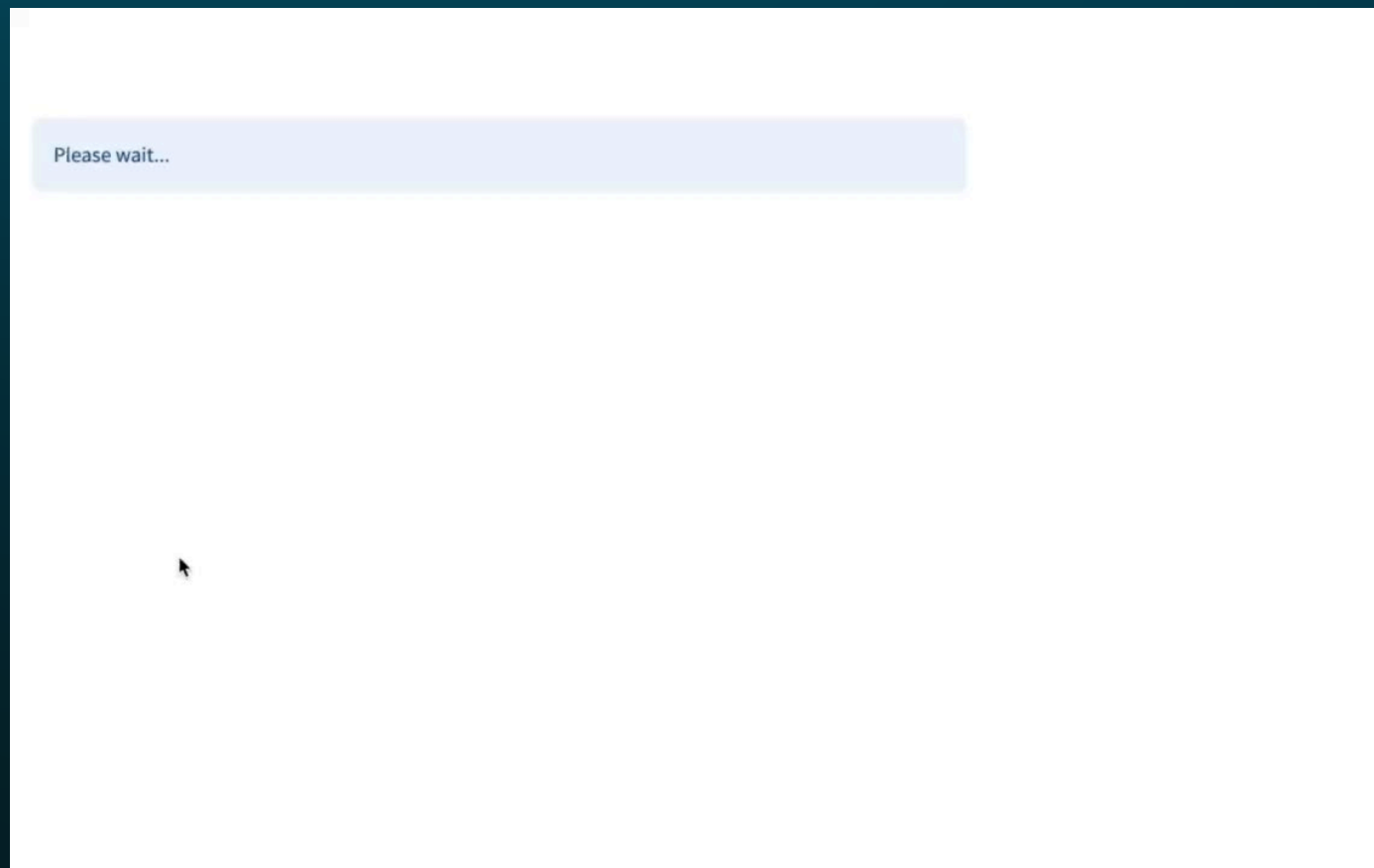


v3.0

- API函数调用 (function calling)
- 代码解释器 (code Interpreter)
- 第三方插件 (plug-in)
- 其他外部工具调用

问题:

通过全国人口数据表格和中国地图,
画出全国各省市的人口分布图



迭代频率“日日新”：半年内秒画完成两次大版本升级

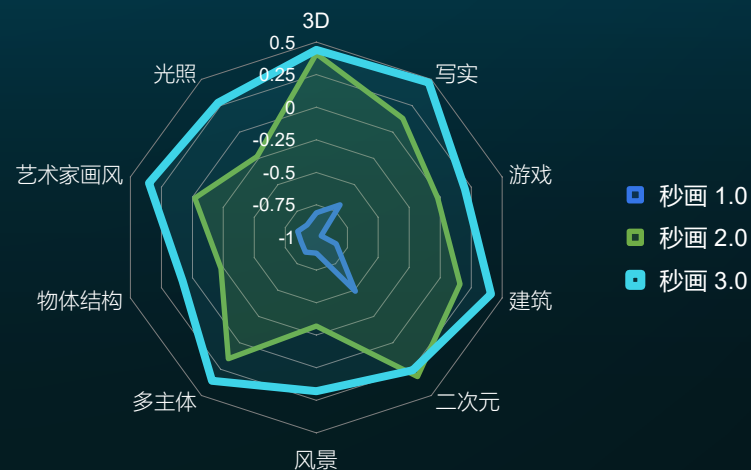
投入2,000 GPU 训练，国内最好、全球前三的文生图模型



秒画3.0 在COCO benchmark
超过 Imagen和 DALL·E 2



1.0版本到3.0版本各个场景能力显著提升



迭代频率“日日新”：半年内秒画完成两次大版本升级



秒画1.0



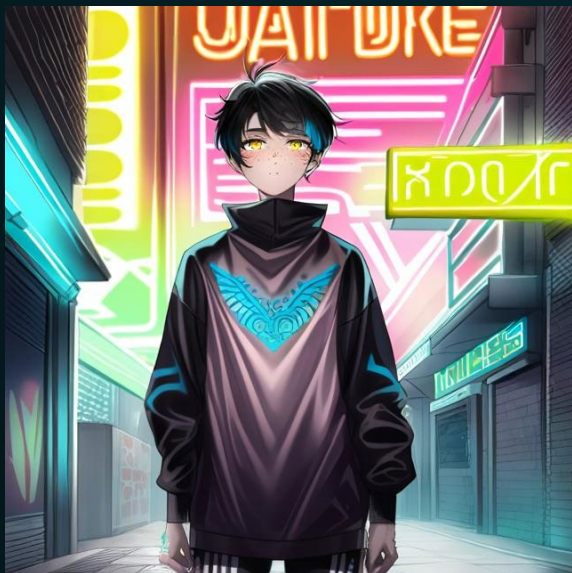
秒画2.0



秒画3.0

提示词：宇航员在太空中冲浪，非常详细，高质量，数字绘画，戏剧性的灯光，artstation 上的流行趋势，杰作，插图

迭代频率“日日新”：半年内秒画完成两次大版本升级



秒画1.0



秒画2.0



秒画3.0

提示词：一位20岁的女子，留着凌乱的黑色短发，方下巴，长着雀斑，穿着街头服装，站在霓虹灯招牌、蓝色和粉色霓虹灯的小巷里

相同提示词：

“中国龙，蓝色摆件，珠宝风格”

相同提示词：

“中国龙，蓝色摆件，珠宝风格”

秒画1.0及秒画2.0：无描述词自动优化功能

“中国龙，蓝色摆件，珠宝风格”



秒画 1.0



秒画 2.0

相同提示词：

“中国龙，蓝色摆件，珠宝风格”

秒画1.0及秒画2.0：无描述词自动优化功能

“中国龙，蓝色摆件，珠宝风格”



秒画 1.0



秒画 2.0

秒画3.0：新增描述词自动优化功能

“想象一下一个中国龙的3D渲染图，具有极其精细和详细的图案。龙站在一个蓝色玛瑙海上，增添了深度和神秘感。龙的身上装饰着精美的珠宝，反射着光线，为场景增添了奢华的触感。”



秒画 3.0

大模型研发成果高效产出

全球首个

高分辨率，对核心大气变量超过10天有效预报

风乌气象
预报模型
4月发布



书生2.5
多模态模型
3月开源

全球首个

解译类别最全的
遥感大模型

SenseEarth3.0
遥感模型
5月发布



GITM
通才智能体
5月发布

国内首个&最佳

国内首次性能表现超越
ChatGPT (GPT-3.5)

书生·浦语等
系列大模型
6月发布



UniAD自动驾驶
多模态模型
6月获奖

国内最佳

最接近全球SOTA
国内性能最强

日日新
sensenova
7月更新



全球最佳

全球开源模型中
ImageNet准确度最高、
规模最大的模型

国内最佳

最接近全球SOTA
国内性能最强

全球最佳

完全解锁《我的世界》整
体科技树，超越OpenAI、
DeepMind

全球首个&最佳

全球首个感知决策一体化
方案，CVPR最佳论文

通用
人工智能

视觉感知

语言理解



内容生成

推理决策

商汤的通用人工智能之路

大模型训练计算量示意

10²⁶
10²⁵
10²⁴
10²³
10²²
10²¹
10²⁰
10¹⁹
10¹⁸
10¹⁷
10¹⁶

注：
1. 示意计算量 = 模型参数量 x 训练数据量 (tokens)
2. 数据来自公开媒体预估



前瞻性布局
大模型“训推一体”
大装置平台

2019年至今

基模型能力率先
突破GPT-3.5

Model	MMLU	AGIEval	AGIEval (GK)	C-Eval	Gaokao
GPT-4	86.4	56.4	58.8	68.7	-
SenseChat 2.0	68.6	49.9	58.5	66.1	-
InternLM	67.2	49.2	57.3	62.7	65.6
ChatGPT	67.3	42.9	45.8	54.4	51.3
LLaMA-65B	63.5	34.0	32.7	38.8	19.0
GLM-130B	44.8	34.2	38.1	40.3	22.4

2023年1-6月

基模型能力
跻身全球TOP3

2023年7-12月

成为领先的大模型
应用生态平台

“AI操作系统”和
“流量入口”

2024年

向通用人工智能
更高标准
持续迭代



2025年以后

积极支持行业模型、插件和应用产品生态布局



备注:

1. 《Generative AI could raise global GDP by 7%》，高盛研究
2. 《The economic potential of generative AI》，麦肯锡报告

日日新大模型应用矩阵



自然语言交互

出色的语言理解

秒解复杂问题

提供定制化建议

多种大小规模系列



AI 文生图

摄影级图像质量

快速生成多风格图片

用户可自定义模型

自动优化提示词



数字人生成

快速、低成本生成

电影级清晰度

流畅的合成语音

高拟真表情
生动的生成动作



3D大场景重建

厘米级精度重建

高效率实时渲染

实景编辑实现

丰富的资产特性



3D小物品重建

细致还原物体

支持多样纹理材质

实时高逼真渲染

降本增效高质量



智慧商业



智慧生活



智能汽车



智慧城市

日日新大模型应用矩阵



自然语言交互

应用行业：

医疗、金融、代码编程、汽车等



AI 文生图

应用行业：

电商、广告、建筑设计等



数字人生成

应用行业：

视频制作、直播、客服等



3D大场景重建

应用行业：

影视创作、建筑、园区管理等



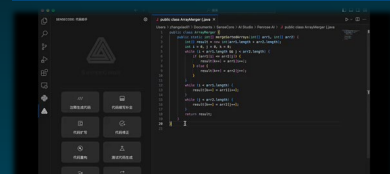
3D小物品重建

应用行业：

商品营销、展览、拍卖等



医疗 - 「商汤·大医」



代码编程 - 「商汤·AI代码助手」翻译



AI 虚拟试穿



AI 广告生成



直播数字人



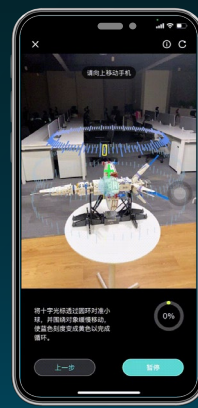
广电主持数字人



数字孪生应用平台



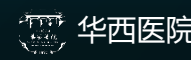
数字孪生城市规划

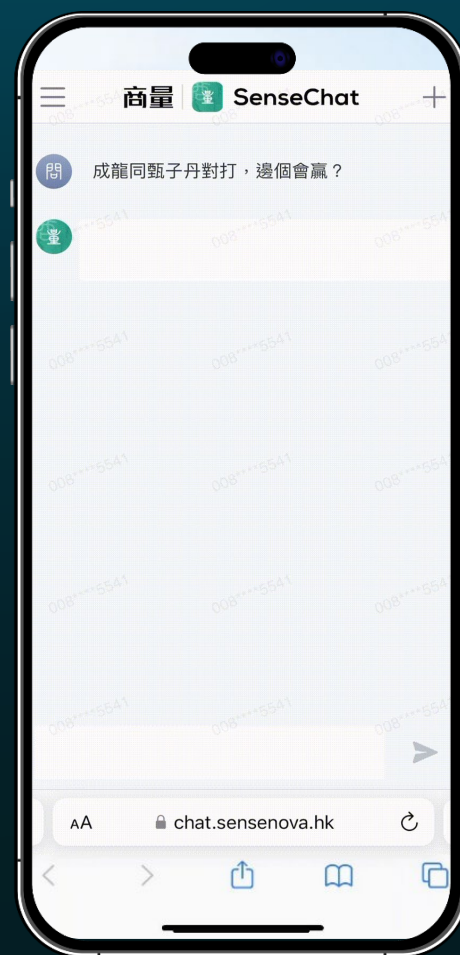
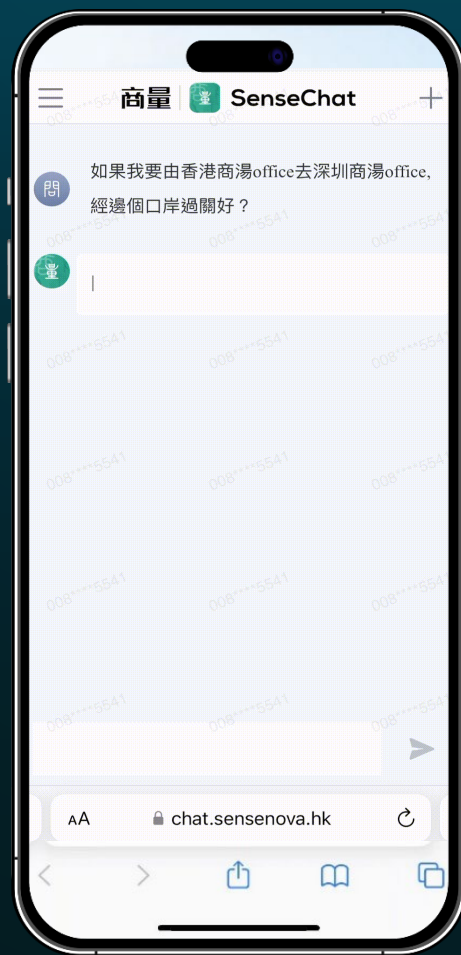


物品采集



珠宝集团3D数字化





- 1 本土內容理解
- 2 口語化問題理解
- 3 口語化回答
- 4 地道文化認識

02

业务回顾

-

大模型体系推动全线业务焕新升级



生成式AI商业化率先体现

20%

生成式AI相关收入占比

领先的大模型能力和稀缺的大装置资源
引领各板块客户进入 AI 2.0 时代

智慧商业

收入占比: 60%
FY2022占比: 38%

行业大模型广泛布局

10+

定制行业大模型

+60%

单客户收入提升

智慧生活

收入占比: 22%
FY2022占比: 25%

生成式AI日常应用铺天盖地

+48%

客户数增长

数亿

终端、企业与
个人用户触达

智能汽车

收入占比: 6%
FY2022占比: 8%

从研发转化为规模化量产

+5.7x

量产收入增长

+29%

单车毛利提升

智慧城市

收入占比: 13%
FY2022占比: 29%

战略性转型, 聚焦高质量业务

64%

国内一二线城市
收入占比

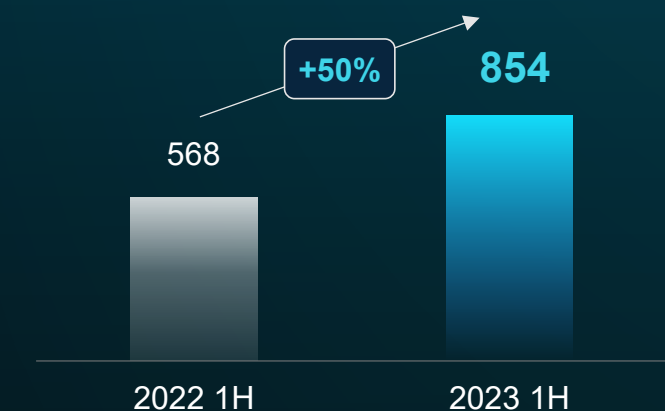
优化

现金流

备注:
1. 截至2023年6月30日数据

大模型时代带动企业智能化转型， 收入大幅上涨

单位：人民币百万元



481

客户数量

+60%

单客户收入同比增长

与灯塔客户合作， 广泛布局行业大模型

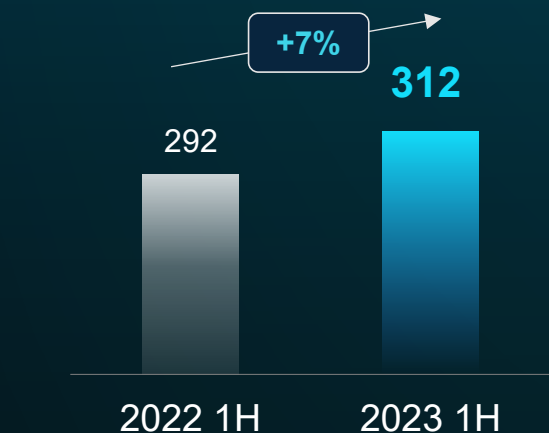
能源 | 金融 | 地产 | 人力资源
传媒 | 工业制造 | 咨询 |

...



生成式AI产品广泛落地，客户数量大幅增加

单位：人民币百万元



收入占比 21% → 22%

48%
客户数量同比增长¹

SenseMARS

生成式AI产品与能力赋能上百个客户
触达数亿终端、中小企业与个人用户



与电商、媒体、社区平台打造行业大模型



- 小样本、批量化数字人制作
- 服务电商、直播、银行客服等场景



3D大场景重建
实现超100平方公里
4K高精度建模



3D小物品重建
已与香港珠宝集团
实现合作

智能终端

三位一体业务战略
探索大模型能力手机端应用



传感器

多款交付中
收入大幅增长



ISP芯片

首款芯片量产
研发中低端版本



SDK

2亿+
新增赋能设备



“手机端AI助理”

“基于自然语言交互的
AI操作系统”



智慧
医疗

“大医”助手
提升医院诊疗效率



家庭互动
机器人

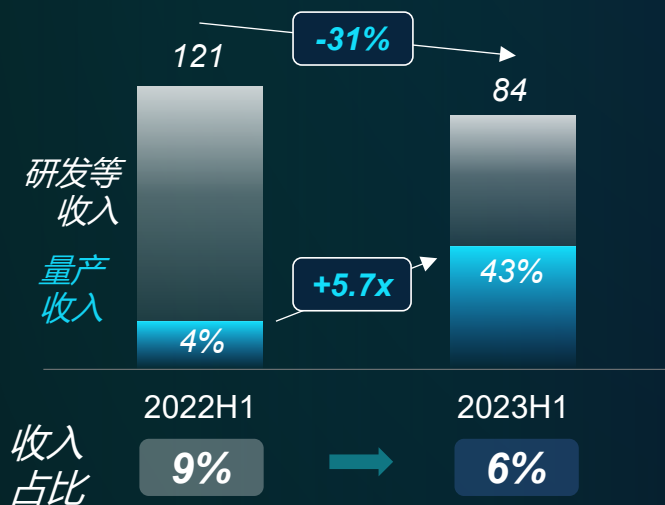
围棋版元萝卜
下棋机器人面市

备注：

1. 2023年上半年客户数量 = 270

量产收入快速增长，单车价值提升

单位：人民币百万元



39万+

新增量产交付数量

500万+

新增定点数量

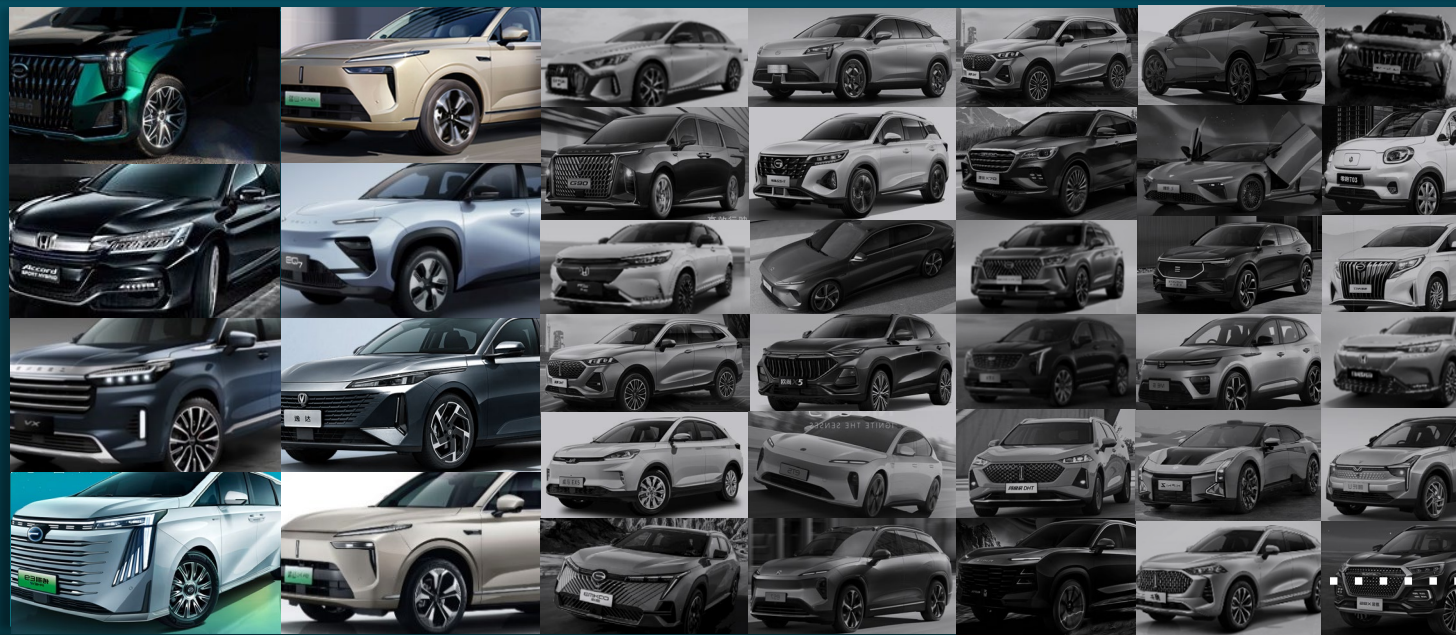
30+

累计车企客户

160+

累计定点车型

已量产交付的合作车企与车型



BYD 比亚迪汽车



GEELY



CHERY



ZEEKR



长城汽车 Great Wall Motors



长安汽车



NIO



广汽集团



HiPhi



JETOUR



HONDA



NISSAN



LEAPMOTOR



北汽集团 FIC Group



上汽



WELTMEISTER 威马汽车



Cadillac

NO.1

座舱AI软件
市场份额¹

50+



累计量产
交付车型

+29%



单车毛
利增长

智驾算法行业领先，推动大模型与智驾结合

BEV感知

- 国内率先完成BEV感知量产交付 
- 获2022年Waymo挑战赛冠军 

UniAD

- 业内首个感知决策一体化的自动驾驶通用大模型 
- 获评CVPR 2023最佳论文 
- 多个关键指标超越SOTA结果



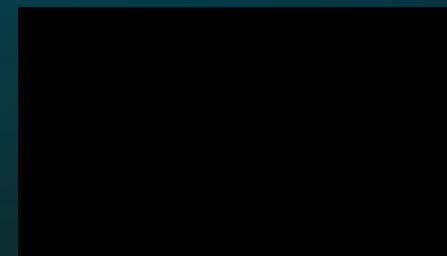
多模态大模型

- 预研自动驾驶多模态大模型
- 提升感知决策能力，推动感知决策全融合

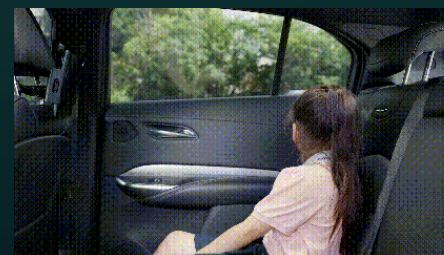
大模型产品首发，打造车舱流量入口



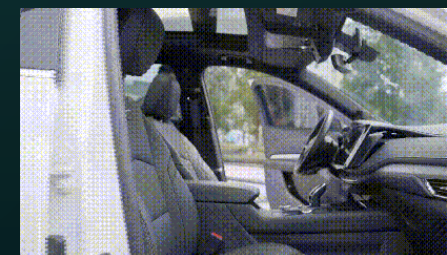
健康问诊



旅游规划



儿童伴读



AR说明书

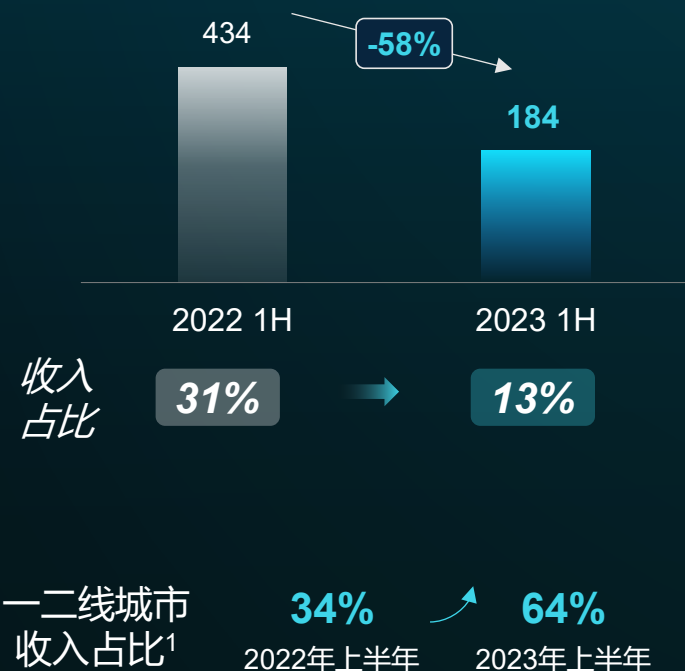
基于大装置+大模型，提供数据闭环训练服务

合作车企



城市智能化升级投入减缓 坚定推进业务战略性转型

单位：人民币百万元



战略转型初现成果

聚焦精品客户
提升收入质量
改善现金流

大模型突破持续提供 城市数字化转型新动能

新基建 投入



助力城市运营区域智算中心

数字 政务



数字办事员提高办事大厅与市民热线的回应效率，创建城市服务新范式

备注：

1. 指智慧城市总体收入中来自中国一二线城市的收入比例

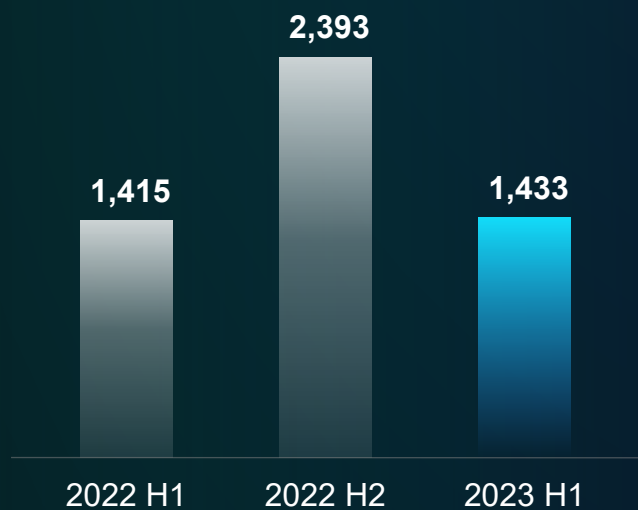
03

财务回顾

-

营业收入

(人民币百万元)



(14.3%)

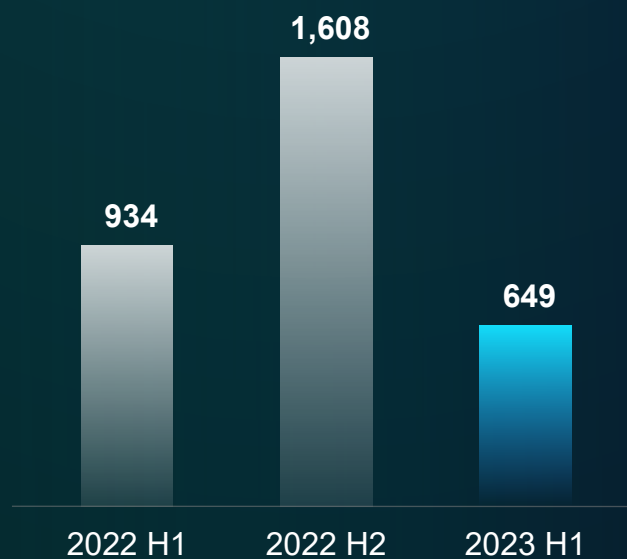
(21.5%)

1.3%

年增长率

毛利

(人民币百万元)



66.0%

67.2%

45.3%

毛利率

调整后净亏损⁽¹⁾

(人民币百万元)



(181.2%)

(90.8%)

(167.0%)

调整后净亏损率

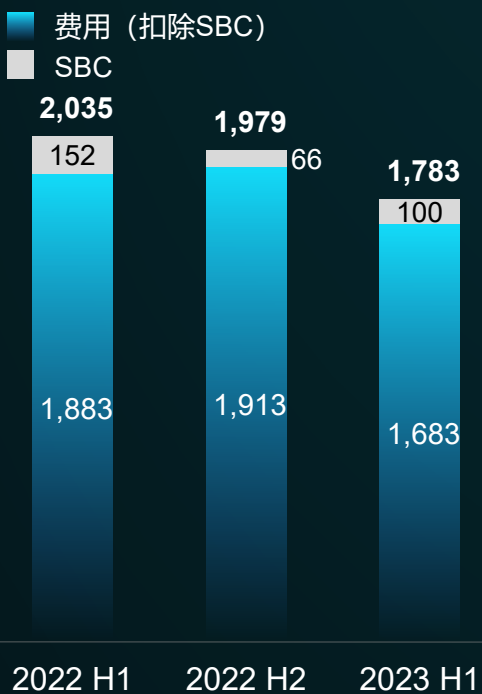
备注:

1. 调整项: 优先股及其他金融负债的公允价值亏损、以股份为基础的薪酬开支、按公允价值计量且变动计入损益的金融资产的公允价值利得/(亏损)

运营支出

研发费用

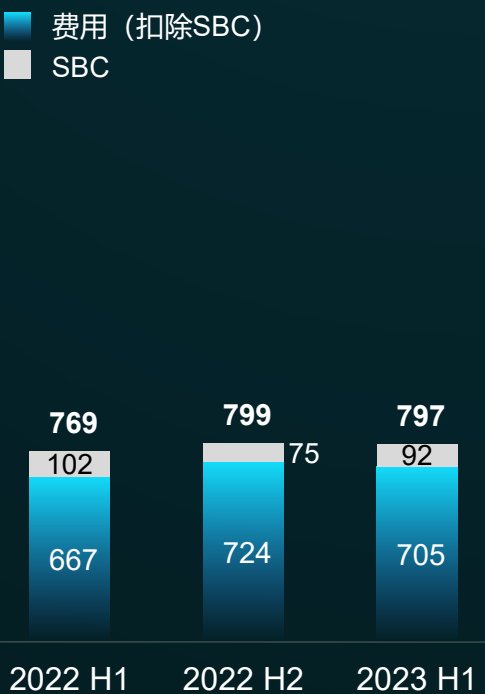
(人民币百万元)



年增长率 (扣除SBC)
 44.1% 9.0% (10.6%)

管理费用

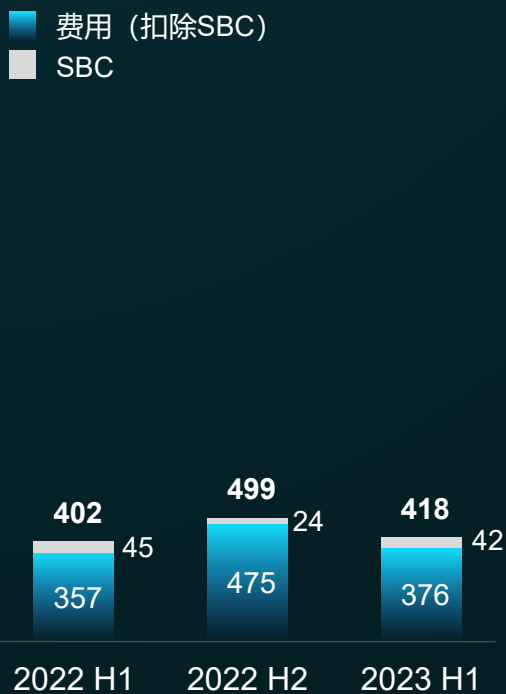
(人民币百万元)



年增长率 (扣除SBC)
 30.3% (9.9%) 5.8%

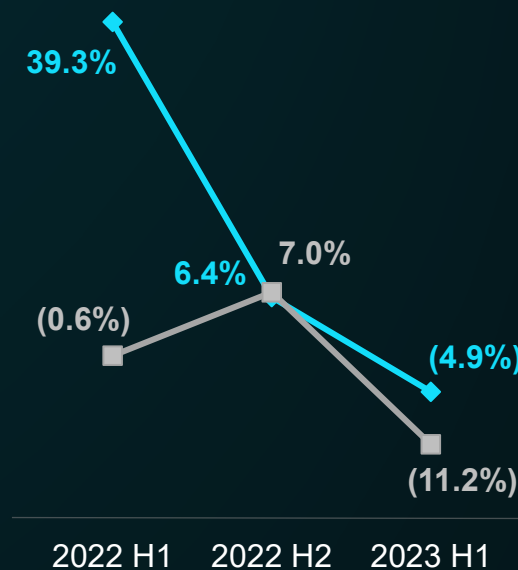
销售费用

(人民币百万元)



年增长率 (扣除SBC)
 33.5% 29.7% 5.4%

运营支出⁽¹⁾年增长率 (扣除SBC)

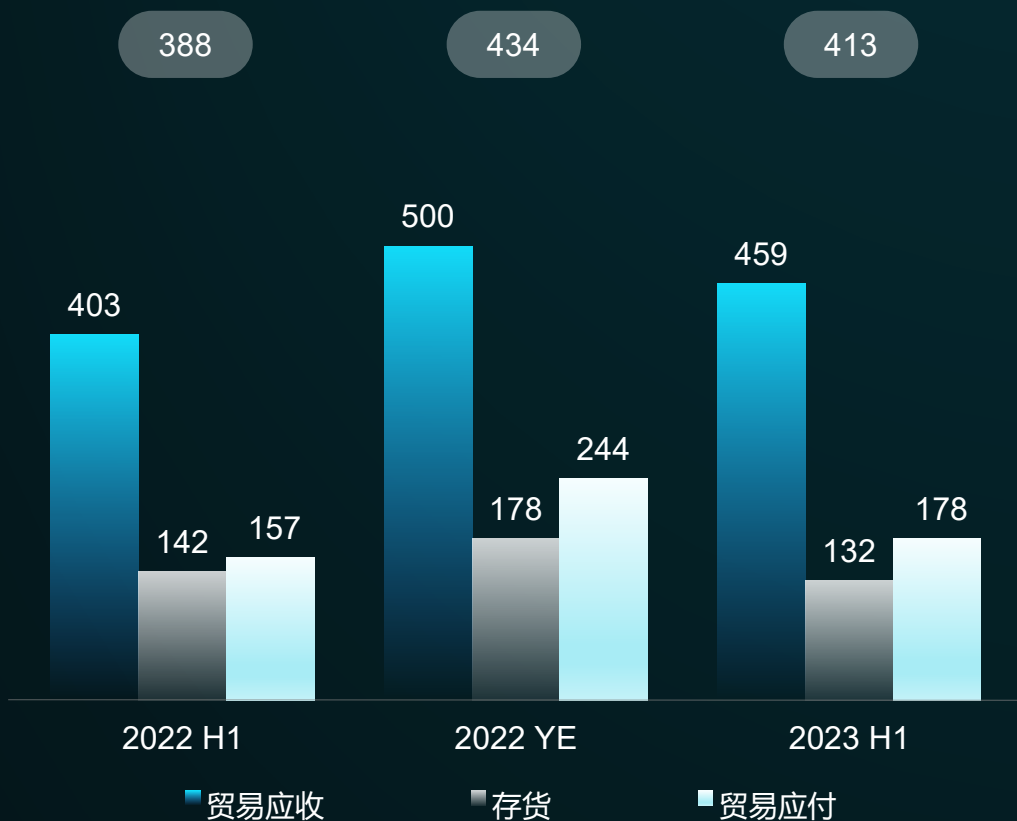


— 年增长率 (扣除SBC)
 — 环比增长率 (扣除SBC)

备注：
 1. 运营支出 = 研发费用 + 管理费用 + 销售费用

周转天数 – 以期末值计算⁽¹⁾

现金循环周期 Cash Conversion Cycle



减值拨备占贸易应收额比率 (资产负债表)



金融资产减值亏损净额 (利润表)

(人民币百万元)



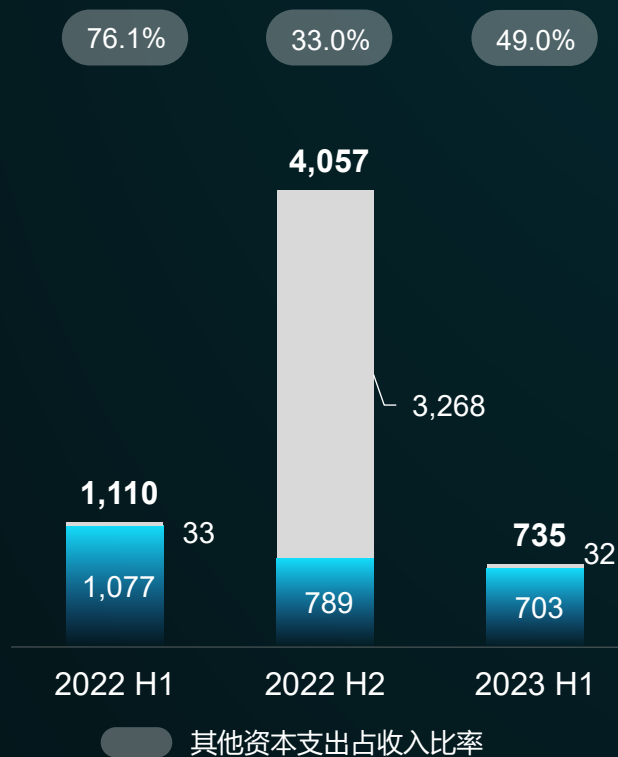
备注:

1. 以期末值计算的周转天数等同于期末值除以过去12个月的收入/成本, 并乘以365天 (期间长度为一年)

资本支出⁽¹⁾

(人民币百万元)

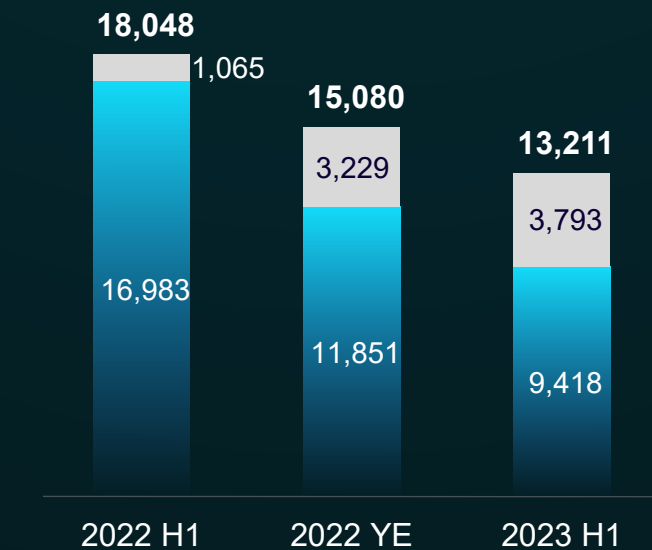
- 其他资本支出⁽¹⁾
- 办公楼层相关资本支出



总现金储备⁽²⁾

(人民币百万元)

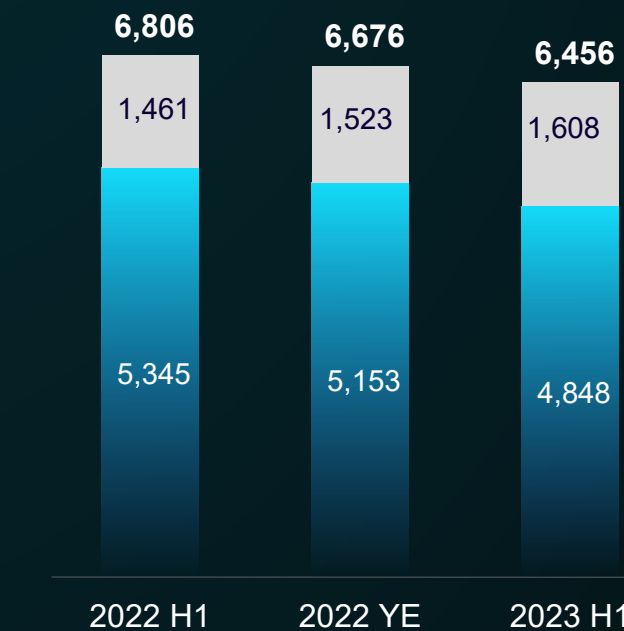
- 净现金
- 总借款负债



股权及债券投资余额⁽³⁾

(人民币百万元)

- 股权投资余额
- 债券投资余额



备注:

1. 定义为购买物业、厂房及设备 + 购买无形资产 + 购买土地使用权的现金流出; 其他资本支出定义为资本支出-办公楼层相关资本支出
2. 定义为现金和现金等价物 + 定期存款 + 结构性存款 + 受限制现金
3. 定义为按公允价值计量且变动计入损益的金融资产 (非流动资产)

Q & A

-



THANKS!

坚持原创，让 AI 引领人类进步

2023.08



画面由商汤秒画生成