

合合信息 (688615.SH)

智能文字识别领军，B/C端双轮驱动成长

合合信息：智能文字识别龙头，AI和大数据技术赋能B/C端。1)公司是智能文字识别领军企业，基于其AI及大数据技术积累，已经形成了覆盖B端与C端全面产品矩阵。其中，C端业务主要是扫描全能王、名片全能王、启信宝3款核心产品，是公司当前最主要的收入来源；公司B端业务为面向企业客户提供以智能文字识别、商业大数据为核心的技术服务，截至2024年年报，《财富》杂志2024年8月发布的世界500强公司名单中，公司B端客户已覆盖超过130家。2)公司整体股权结构稳定集中，截至2025H1公司实控人镇立新持有公司股权24%，公司前十大股东持有公司股权约61%。管理层技术出身，具备AI领域深厚技术基础。3)公司正处于成长期，收入端与利润端皆稳步增长。2022-2024年员工人数持续提升，体现出公司经营效率和管理质量的持续改善。

智能文字识别市场和商业大数据市场处于快速成长期，全球市场空间广阔。1)全球OCR市场预计到2032年将达到519.2亿美元，在2026E-2032E年期间，复合年增长率将达到15.24%。AI技术快速成熟+80%企业文档有望走向自动化，将驱动OCR技术转向智能文字识别，进一步打开市场空间。2)2024年国内商业大数据服务行业市场规模已超600亿元，同比增长20.76%，市场规模广阔且增长迅速。由于该行业具备较高的技术壁垒+资质壁垒，市场竞争格局相对集中。C端市场呈现出三足鼎立局面，“一信两查”已占领用户心智；B端市场集中度略低于C端，碎片化需求带来更多玩家生存空间。

竞争壁垒在哪里？C端深入用B端智通最后一公里，B+C端协同发展筑定扎实内功。1)公司深耕智能文字识别技术，准确率与实用性领先全球，核心APP在全球拥有庞大的用户基础，截至2025H1月活合计超过1.8亿。2023年至2025H1公司用户数和付费率逐年提升，印证公司产品市场的认可度与用户心智持续提升。2)公司深耕细作B端业务已近十年，打通落地的最后一公里，针对不同行业场景和客户痛点提出有针对性的解决方案，能够有效帮助客户降本增效，提供实际业务价值（例如伟巴斯特项目中帮助客户整体作业效率提升500%~1000%）。3)公司兼具有成长潜力的C端业务、稳定盈利的B端业务，C端/B端业务在研发、营销投入方面具备协同性，技术算法模块可研发复用，C端优良口碑有利于B端业务开展，可降低B端营销成本。

投资建议：首次覆盖，给予“买入”评级。基于对公司基本面、商业模式以及行业趋势的判断与分析，我们预计公司2025E/26E/27E有望实现总营业收入18.3/23.0/28.5亿元，同比增速达到28%/25%/24%；实现归母净利润4.87/6.10/7.55亿元，同比增长21.5%/25.4%/23.8%。截至2026/1/31，公司PS(2026E)低于可比公司平均水平，具备估值性价比，首次覆盖给予“买入”评级。

风险提示：AI技术迭代不及预期的风险，海外业务发展不及预期的风险，宏观经济波动的风险。

财务指标	2023A	2024A	2025E	2026E	2027E
营业收入(百万元)	1,187	1,438	1,834	2,300	2,847
增长率 yoy (%)	20.0	21.2	27.5	25.4	23.8
归母净利润(百万元)	323	401	487	610	755
增长率 yoy (%)	13.9	23.9	21.5	25.4	23.8
EPS 最新摊薄(元/股)	2.31	2.86	3.48	4.36	5.40
净资产收益率(%)	32.4	15.1	15.7	17.2	18.8
P/E(倍)	116.0	93.6	77.0	61.4	49.6
P/B(倍)	37.5	14.1	12.1	10.6	9.3

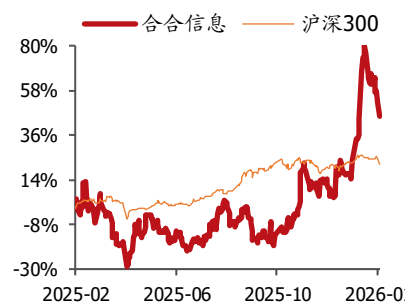
资料来源：Wind，国盛证券研究所 注：股价为2026年02月02日收盘价

买入(首次)

股票信息

行业	软件开发
02月02日收盘价(元)	267.73
总市值(百万元)	37,482.20
总股本(百万股)	140.00
其中自由流通股(%)	69.05
30日日均成交量(百万股)	4.38

股价走势



作者

分析师	孙行臻
执业证书编号	S0680526010001
邮箱	sunxingzhen1@gszq.com
分析师	陈泽青
执业证书编号	S0680523100001
邮箱	chenzeqing3655@gszq.com

相关研究

财务报表和主要财务比率
资产负债表 (百万元)

会计年度	2023A	2024A	2025E	2026E	2027E
流动资产	1608	3201	3876	4607	5410
现金	353	1495	1733	1999	2299
应收票据及应收账款	90	105	135	169	210
其他应收款	7	7	9	11	14
预付账款	12	14	17	21	26
存货	2	4	4	5	7
其他流动资产	1144	1577	1978	2400	2854
非流动资产	189	288	305	294	271
长期投资	0	0	0	0	0
固定资产	54	142	151	154	152
无形资产	3	48	58	58	53
其他非流动资产	132	97	96	81	67
资产总计	1797	3488	4181	4900	5681
流动负债	747	794	1051	1313	1621
短期借款	0	0	0	0	0
应付票据及应付账款	56	63	78	97	120
其他流动负债	691	731	974	1216	1501
非流动负债	52	37	38	38	38
长期借款	0	0	0	0	0
其他非流动负债	52	37	38	38	38
负债合计	799	831	1089	1351	1658
少数股东权益	0	0	0	0	0
股本	75	100	140	160	160
资本公积	208	1453	1463	1473	1493
留存收益	715	1115	1502	1929	2382
归属母公司股东权益	998	2658	3092	3550	4023
负债和股东权益	1797	3488	4181	4900	5681

现金流量表 (百万元)

会计年度	2023A	2024A	2025E	2026E	2027E
经营活动现金流	451	456	777	913	1101
净利润	323	401	487	610	755
折旧摊销	34	73	80	86	93
财务费用	-3	-4	0	0	0
投资损失	-1	-2	-2	-2	-3
营运资金变动	124	21	222	217	255
其他经营现金流	-25	-32	-10	1	1
投资活动现金流	-635	-662	-485	-494	-518
资本支出	-76	-174	-86	-76	-71
长期投资	-566	-503	-399	-419	-449
其他投资现金流	7	15	0	1	2
筹资活动现金流	-21	1260	-51	-153	-282
短期借款	0	0	0	0	0
长期借款	0	0	0	0	0
普通股增加	0	25	40	20	0
资本公积增加	0	1245	10	10	20
其他筹资现金流	-21	-10	-101	-183	-302
现金净增加额	-200	1059	238	266	300

利润表 (百万元)

会计年度	2023A	2024A	2025E	2026E	2027E
营业收入	1187	1438	1834	2300	2847
营业成本	187	226	283	353	437
营业税金及附加	7	10	11	14	17
营业费用	335	429	532	667	826
管理费用	76	84	110	138	171
研发费用	323	390	499	625	774
财务费用	-7	-14	-15	-17	-20
资产减值损失	0	0	0	0	0
其他收益	39	87	92	115	142
公允价值变动收益	20	24	0	0	0
投资净收益	1	2	2	2	3
资产处置收益	0	0	0	0	0
营业利润	323	423	508	636	788
营业外收入	0	0	0	0	0
营业外支出	1	3	1	1	1
利润总额	322	421	507	636	787
所得税	-1	20	20	25	31
净利润	323	401	487	610	755
少数股东损益	0	0	0	0	0
归属母公司净利润	323	401	487	610	755
EBITDA	328	455	572	705	860
EPS (元/股)	2.31	2.86	3.48	4.36	5.40

主要财务比率

会计年度	2023A	2024A	2025E	2026E	2027E
成长能力					
营业收入(%)	20.0	21.2	27.5	25.4	23.8
营业利润(%)	15.2	31.0	20.0	25.4	23.8
归属母公司净利润(%)	13.9	23.9	21.5	25.4	23.8
获利能力					
毛利率(%)	84.3	84.3	84.6	84.6	84.7
净利率(%)	27.2	27.9	26.5	26.5	26.5
ROE(%)	32.4	15.1	15.7	17.2	18.8
ROIC(%)	27.8	13.5	15.0	16.5	18.1
偿债能力					
资产负债率(%)	44.5	23.8	26.0	27.6	29.2
净负债比率(%)	-28.9	-54.3	-54.4	-54.9	-55.9
流动比率	2.2	4.0	3.7	3.5	3.3
速动比率	2.1	4.0	3.6	3.5	3.3
营运能力					
总资产周转率	0.8	0.5	0.5	0.5	0.5
应收账款周转率	13.0	14.7	15.3	15.1	15.0
应付账款周转率	3.7	3.8	4.0	4.0	4.0
每股指标 (元)					
每股收益(最新摊薄)	2.31	2.86	3.48	4.36	5.40
每股经营现金流(最新摊薄)	3.22	3.25	5.55	6.52	7.86
每股净资产(最新摊薄)	7.13	18.98	22.09	25.35	28.73
估值比率					
P/E	116.0	93.6	77.0	61.4	49.6
P/B	37.5	14.1	12.1	10.6	9.3
EV/EBITDA	-0.9	41.3	62.6	50.4	41.0

资料来源: Wind, 国盛证券研究所 注: 股价为 2026 年 02 月 02 日收盘价

内容目录

1.智能文字识别龙头，B/C 端双轮驱动成长	4
2.智能文字识别市场和商业大数据市场处于快速成长期	12
2.1 智能文字识别市场空间广阔，可解决传统 OCR 的缺点	12
2.2 商业大数据市场方兴未艾，数据价值深度发掘赋能下游行业	14
3.竞争壁垒：C 端深入用户心智，B 端打通最后一公里，B+C 协同发展沉淀扎实内功	18
3.1 公司深耕智能文字识别技术，准确率与实用性领先全球	18
3.2 C 端：已建立用户心智与广泛用户基础，高粘性+高迁移成本构成护城河	20
3.3 B 端：深耕优质客户+深度行业场景 know-how，打通企业级 AI 应用最后一公里	21
3.4 B+C 联动：B 端与 C 端业务形成协同，联动发展提升研发与销售效率	23
4.盈利预测	25
5.风险提示	28

图表目录

图表 1: 公司主营业务构架图	4
图表 2: 公司产品矩阵	5
图表 3: 公司各业务板块商业模式梳理	6
图表 4: 公司前十大股东明细（截至 2025H1）	7
图表 5: 公司营业总收入及同比增速	7
图表 6: 公司归母净利润及同比增速	8
图表 7: 2025H1 公司各大业务板块结构占比	8
图表 8: 公司各业务板块收入情况（单位：亿元）	9
图表 9: 公司各主要业务板块收入同比增速情况	9
图表 10: 公司国内外收入结构占比	9
图表 11: 公司销售毛利率与销售净利率情况	10
图表 12: 公司人均创收与人均创利情况	10
图表 13: 公司销售/管理/财务费用率情况	11
图表 14: 公司研发投入占营业收入比例	11
图表 15: 2024-2032E 年全球智能文字识别服务市场规模统计	12
图表 16: 2017-2027E 年中国智能文字识别服务市场规模统计	12
图表 17: 2022-2027E 中国智能文字识别细分市场年均复合增速	13
图表 18: 智能文字识别技术与传统 OCR 技术的区别	13
图表 19: 智能文字识别技术的功能	14
图表 20: 国内商业大数据服务行业市场规模及增速	15
图表 21: 商业大数据服务商业模式	16
图表 22: 商业大数据产业链图谱	17
图表 23: 公司智能文字识别技术能解决现实场景下的众多技术挑战	19
图表 24: 公司主要 C 端产品月活用户数及同比增速	20
图表 25: 公司累计付费用户数及同比增速	20
图表 26: 公司 C 端 APP 付费率逐步上行	21
图表 27: 合合信息可为保险行业提供的服务及解决方案	22
图表 28: 合合信息可为地产行业提供的服务及解决方案	22
图表 29: 合合信息合作大量 B 端优质客户	23
图表 30: 2025E-2027E 公司盈利预测	26
图表 31: 2025E-2027E 公司毛利率预测	27
图表 32: 公司销售/管理/研发费用率情况	27
图表 33: 同业可比公司估值情况（截至 2026/1/31）	28

1. 智能文字识别龙头，B/C端双轮驱动成长

公司是智能文字识别龙头，以 AI 和大数据技术同时赋能 B 端与 C 端。合合信息基于自主研发的领先的智能文字识别及商业大数据核心技术，为全球 C 端用户和多元行业 B 端客户提供数字化、智能化的产品及服务。公司 C 端业务主要为面向全球个人用户的 APP 产品，包括扫描全能王、名片全能王、启信宝 3 款核心产品；公司 B 端业务为面向企业客户提供以智能文字识别、商业大数据为核心的服务。公司的 C 端产品覆盖了全球百余个国家和地区的亿级用户，B 端服务覆盖了近 30 个行业的企业客户。截至 2024 年年报，《财富》杂志 2024 年 8 月发布的世界 500 强公司名单中，公司 B 端客户已覆盖超过 130 家。

图表1: 公司主营业务构架图



资料来源：公司招股书，国盛证券研究所

公司主要产品矩阵包括：

● 智能文字识别：

- 1) C 端 APP：扫描全能王、名片全能王
- 2) B 端技术基础服务
- 3) B 端标准化服务：名片全能王企业版 SaaS 服务
- 4) B 端场景化解决方案

● 商业大数据：

- 1) C 端 APP：启信宝 APP
- 2) B 端基础数据服务
- 3) B 端标准化服务：启信宝企业版 SaaS 服务
- 4) B 端场景化解决方案

图表2: 公司产品矩阵

所处领域	核心技术	产品/服务矩阵	业务内容
智能文字识别	智能图像处理、复杂场景文字识别、NLP 技术	C 端 APP	扫描全能王 APP: 智能文字扫描及识别软件, 可将复杂场景下的文档转变为扫描仪效果的 PDF 文件或图片, 并进一步识别为文本, 实现文档资产的便捷管理 名片全能王 APP: 智能名片及人脉管理软件, 可将复杂场景下的纸质名片转变为电子联系人, 实现名片资产的便捷管理
		B 端基础技术服务	证照、银行卡、名片、通用文本、表格等智能文字识别的基础技术服务
		B 端标准化服务	名片全能王企业版 SaaS 软件服务
		B 端场景化解决方案	财报机器人、票据机器人、合同机器人、智能配置机器人等与业务场景紧密结合的文档智能识别机器人 智能文字识别 AI 训练平台
商业大数据	大数据挖掘、知识图谱、NLP 技术	C 端 APP	启信宝 APP: 企业商业信息查询平台, 汇集境内 2.3 亿家企业等组织机构的超过 1,000 亿条实时动态数据, 并提供多种数据挖掘分析报告
		B 端基础数据服务	数据包服务 数据 API 服务
		B 端标准化服务	启信宝企业版 SaaS 软件服务
		B 端场景化解决方案	企业知识图谱解决方案: 包括金融风险知识图谱解决方案、供应链大数据风控平台、企业信息实时分析大屏 启信宝商业全景数据库

资料来源: 公司招股书, 国盛证券研究所

商业模式上, 公司 C 端业务以订阅制为主, 提供稳健现金流, B 端则通过技术服务收费。公司 C 端、B 端业务采取不同的商业模式: 1) C 端 APP: 扫描全能王、名片全能王、启信宝, 都以年度、月度 VIP 会员付费为主, 采取订阅制, 可以为公司带来稳健的现金流。2) B 端: 通过技术服务收费, 具体收费模式可分为基础数据服务、标准化服务、场景化解决方案三种模式。

图表3: 公司各业务板块商业模式梳理

业务分类	产品/服务矩阵	业务内容	主要盈利模式	销售模式
智能文字识别	C端APP	扫描全能王APP、名片全能王APP	(1) 月度、年度等各类型的VIP会员付费 (2) 其他按次付费产品	用户直接在公司App或Web端购买相关产品
	B端基础技术服务	证照、银行卡、名片等智能文字识别的基础技术服务	软件永久一次性授权费+升级更新及技术支持服务费, 或者按期间收取软件使用费	主要为直销, 经销为辅
	B端标准化服务	名片全能王企业版SaaS软件服务	按期间收取软件使用费	主要为直销, 经销为辅
	B端场景化解决方案	智能文字识别AI机器人、智能文字识别AI训练平台	软件一次性授权费+升级更新及技术支持服务费	主要为直销, 经销为辅
商业大数据	C端APP	启信宝APP	(1) 月度、年度等各类型的VIP会员付费 (2) 付费报告等非会员类产品	用户直接在公司App或Web端购买相关产品
	B端基础数据服务	数据包服务数据API服务	数据包: 一次性收费数据 API: 按期间内的调用次数或调用量收费	主要为直销, 经销为辅
	B端标准化服务	启信宝企业版SaaS软件服务	按期间收取软件使用费	主要为直销, 经销为辅
	B端场景化解决方案	企业知识图谱解决方案 启信宝商业全景数据库服务	企业知识图谱解决方案: 软件永久一次性授权费+升级更新及技术支持服务费 数据库: 一次性数据库落地费用+数据更新服务费	主要为直销, 经销为辅
互联网广告推广	互联网广告推广	基于C端APP, 为B端各类广告客户提供广告营销服务	根据不同位置、时长、出现方式等以及广告客户不同的需求来确定结算方式和收入	直销与第三方平台合作推广模式为主, 经销为辅
手机厂商技术授权	手机厂商技术授权	将名片识别、银行卡识别等技术模块授权给手机厂商、集成在其手机产品中	按集成该技术的手机出厂量或者按照合同约定的授权期间向公司支付费用	主要为直销

资料来源: 公司招股书, 国盛证券研究所

公司股权结构集中稳定, 管理层技术出身, 具备AI领域深厚技术基础。截至2025H1公司实控人镇立新持有公司股权24%, 公司前十大股东持有公司股权约61%, 公司整体股权结构稳定集中。实控人镇立新与二股东罗希平都毕业于中国科学院自动化研究所, 为模式识别与智能系统博士, 具备深厚AI领域技术基础。

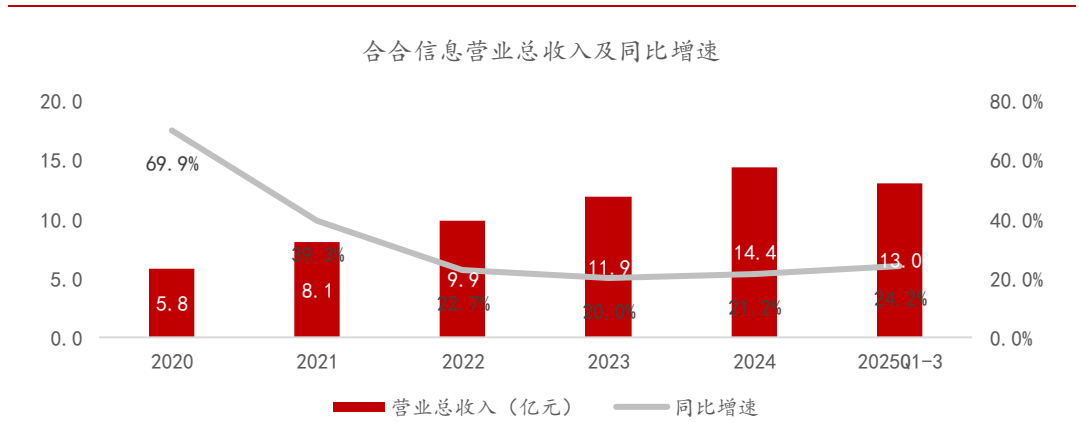
图表4: 公司前十大股东明细 (截至 2025H1)

序号	股东名称	持股数量(股)	占总股本比例(%)
1	镇立新	33,862,920	24.19
2	罗希平	7,177,800	5.13
3	东方富海(芜湖)股权投资基金(有限合伙)	7,094,745	5.07
4	常州鼎仕投资合伙企业(有限合伙)	6,749,295	4.82
5	陈青山	6,574,995	4.70
6	经纬(杭州)创业投资合伙企业(有限合伙)	5,496,750	3.93
7	宁波梅山保税港区启安股权投资合伙企业(有限合伙)	5,297,565	3.78
8	上海卉新投资中心(有限合伙)	4,769,285	3.41
9	上海目一然投资中心(有限合伙)	4,049,641	2.89
10	龙腾	3,694,110	2.64
	合计	84,767,106	60.56

资料来源: wind, 国盛证券研究所

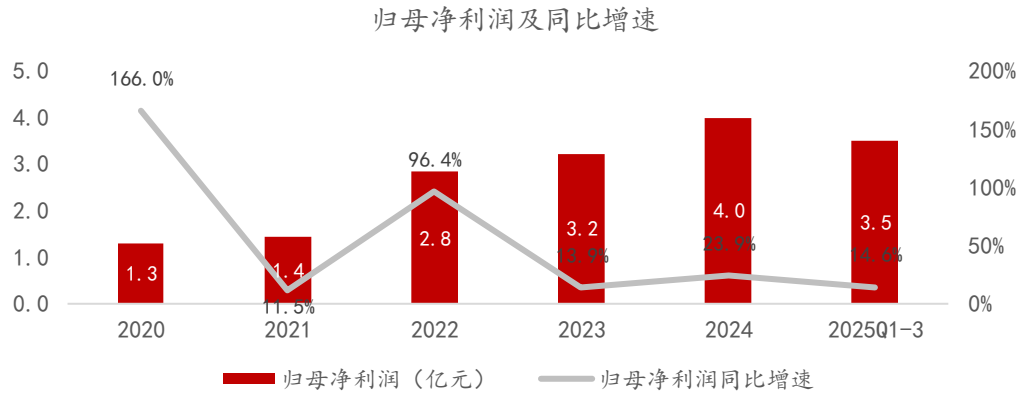
收入端、利润端快速增长, 公司正处于成长拓展期。2020 年至 2025Q3, 公司收入端、利润端都呈现出快速增长态势。2025Q1-3, 公司营业总收入达到 13 亿元, 同比增长 24.2%; 实现归母净利润 3.5 亿元, 同比增长 14.6%。

图表5: 公司营业总收入及同比增速



资料来源: wind, 国盛证券研究所

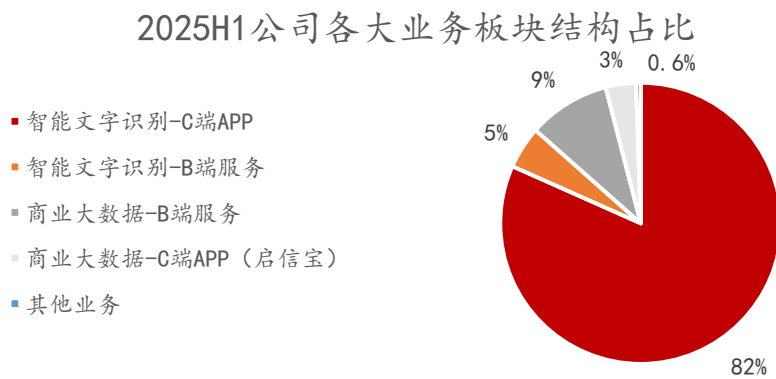
图表6: 公司归母净利润及同比增速



资料来源: wind, 国盛证券研究所

智能文字识别 C 端业务贡献公司收入的 82%，是公司最主要的收入来源。公司业务版图可大体分为智能文字识别 C 端/B 端、商业大数据 C 端/B 端四大业务板块，其中智能文字识别 C 端 APP 是最大的业务部门，2025H1 贡献公司整体收入的 82%，是公司最重要的收入来源。

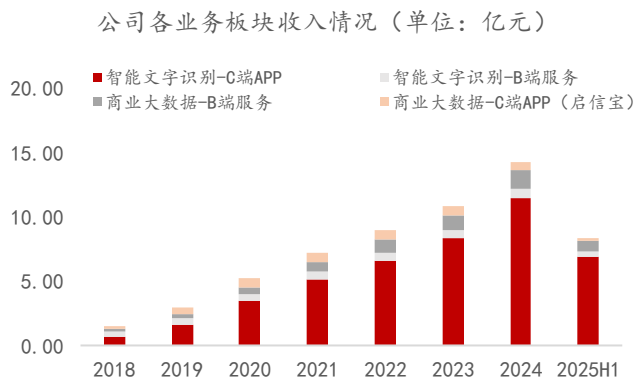
图表7: 2025H1 公司各大业务板块结构占比



资料来源: wind, 国盛证券研究所

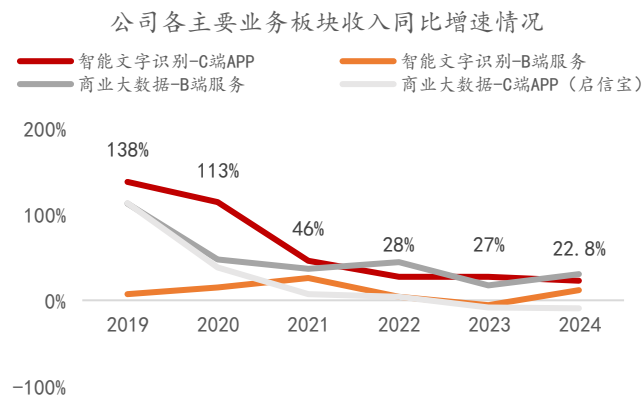
智能文字识别 C 端和商业大数据 B 端业务引领公司成长。自 2021 年以来，智能文字识别 C 端业务和商业大数据 B 端业务就是公司增速最快的业务板块，引领公司整体成长。2024 年，智能文字识别 C 端 APP 业务增速达到 22.8%，商业大数据 B 端业务增速则达到 30%。

图表8: 公司各业务板块收入情况 (单位: 亿元)



资料来源: wind, 国盛证券研究所

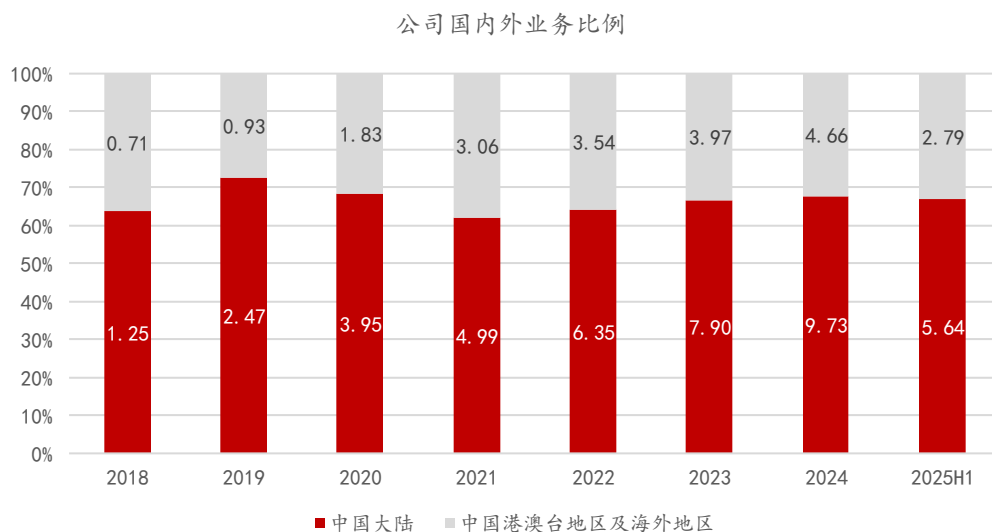
图表9: 公司各主要业务板块收入同比增速情况



资料来源: wind, 国盛证券研究所

海外收入占比达到 **33%**, 海外市场深耕细作已初见成效, 具备较大商业化潜力。截至 2025H1, 公司中国港澳台及海外地区业务占比已经达到总收入的 **33%**, 成为重要的收入来源。根据公司官方投资者交流纪要内容披露, 公司在海外运营中将推行差异化战略: 聚焦不同地区用户的个性化需求, 针对性开发产品功能、提升用户体验。短期内不急于推进全面商业化变现, 而是更注重沉淀用户的长期价值, 待用户使用习惯养成且黏性达到一定程度后, 再将国内经过验证的投放与转化经验复用于海外市场, 从而实现商业变现与用户体验的平衡。从当前海外业务收入来看, 还有较大的商业化潜力。

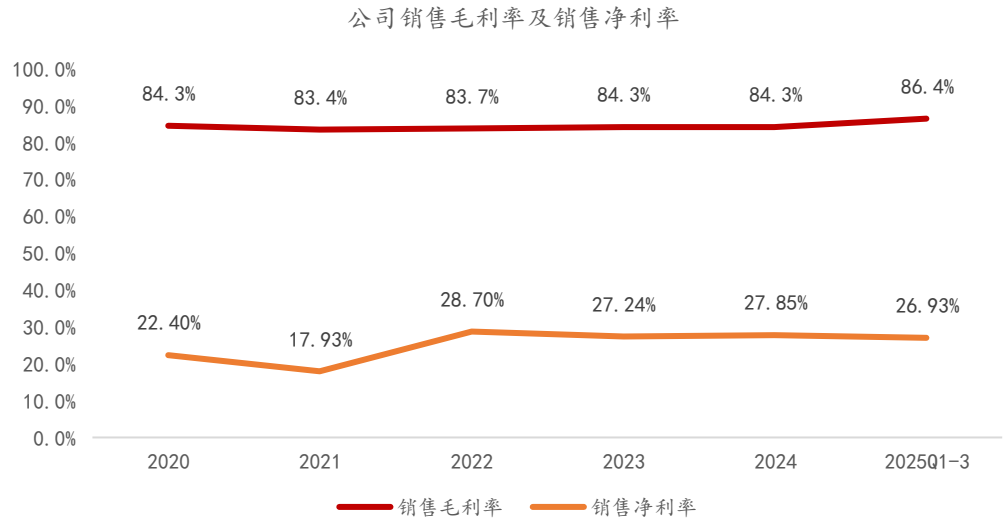
图表10: 公司国内外收入结构占比



资料来源: wind, 国盛证券研究所

毛利率常年维持在 **80%**以上, 净利率保持相对平稳, 体现出公司商业模式所带来的盈利能力优越性。公司最重要的收入来源为智能文字识别 C 端业务, 其商业模式为订阅制, 天然具备高毛利的特性。2020 年至 2025Q3, 公司毛利率常年稳定维持在 **80%**以上, 而净利率保持相对平稳的状态, 体现出公司优越商业模式所带来的盈利能力。

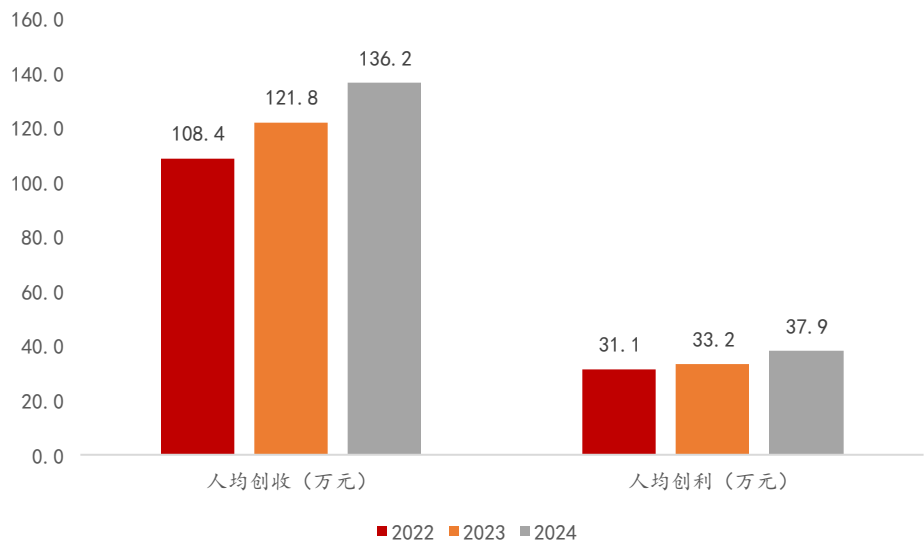
图表11: 公司销售毛利率与销售净利率情况



资料来源: wind, 国盛证券研究所

公司管理效率持续提升, 经营健康, 员工人效持续提升。2022-2024年, 公司人均创收显著提升, 从2022年的108万元提升至2024年的136万元; 人均创利从2022年的31万元提升至2024年的38万元, 体现出公司的经营效率提升与经营质量的健康。

图表12: 公司人均创收与人均创利情况

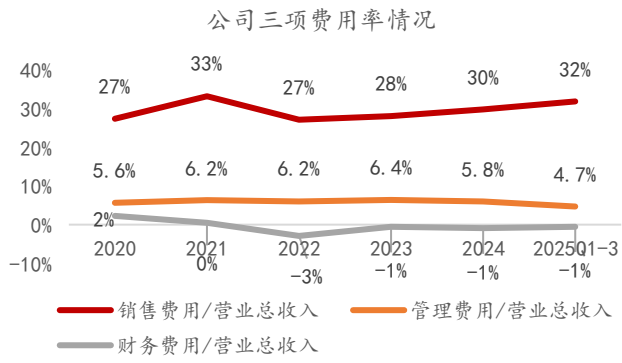


资料来源: wind, 国盛证券研究所

销售费用率稳中微升, 管理费用率及财务费用率基本维持稳定。自2020年至2025Q3, 公司三项费用率波动不大, 销售费用率稳中有升, 来自于公司市场拓展的客观需求, 管理费用率和财务费用率则大体维持稳定, 体现公司经营质量的稳健。2025Q1-3, 公司销售/管理/财务费用率分别达到32%/4.7%/-1%。

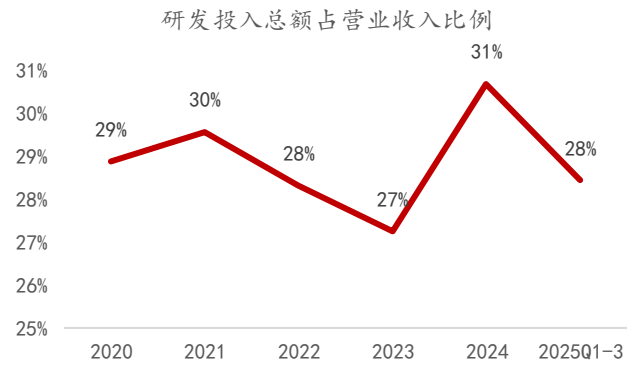
常年坚定投入研发, 为公司打下坚实的技术基础。作为AI及大数据相关企业, 公司研发投入常年高企, 自2020年至2025Q3, 研发费用率一直在27%-31%之间波动, 持续坚定的研发投入为公司的技术壁垒打下坚实基础。

图表13: 公司销售/管理/财务费用率情况



资料来源: wind, 国盛证券研究所

图表14: 公司研发投入占营业收入比例



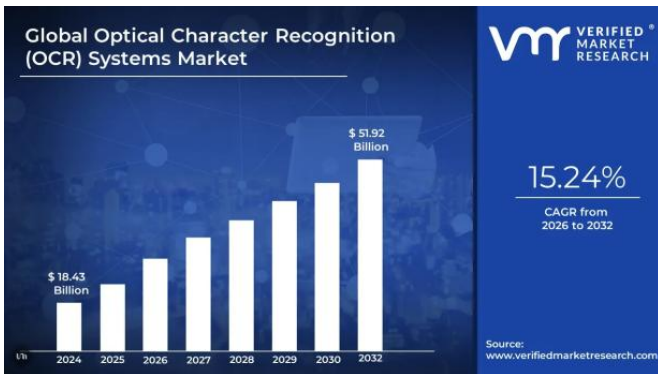
资料来源: wind, 国盛证券研究所

2.智能文字识别市场和商业大数据市场处于快速成长期

2.1 智能文字识别市场空间广阔，可解决传统 OCR 的缺点

2024 年全球 OCR 市场规模已达到 184 亿美元，国内 OCR 市场规模达 105 亿元，处于快速成长期。1) 根据 Verified market research 数据，全球光学字符识别 (OCR) 系统市场规模在 2024 年达到 184.3 亿美元，预计到 2032 年将达到 519.2 亿美元，在 2026E-2032E 年期间，复合年增长率为 15.24%。预计到 2026 年初，该市场已成熟为以 AI 为先的基础设施，OCR 不再是一个独立的工具，而是成为智能文档处理 (IDP) 的主要入口，使纸质信息能够无缝集成到企业 ERP 和 CRM 系统中。2) 根据智研咨询数据，2024 年中国智能文字识别市场规模已从 2017 年的 6.1 亿元增长至 105.3 亿元，预计 2027 年国内智能文字识别市场规模有望达到 170 亿元。从细分市场看，伴随下游个人用户和企业客户对文档资产数字化与智能化升级需求的驱动下，国内智能文字识别各类型细分市场均将保持较高增速。预计 2022-2027 年我国智能文字识别中基础技术服务、标准化产品 (C 端)、标准化产品 (B 端)、场景化解决方案 CAGR 分别为 16.5%、22.4%、30%、31.6%。

图表15: 2024-2032E 年全球智能文字识别服务市场规模统计



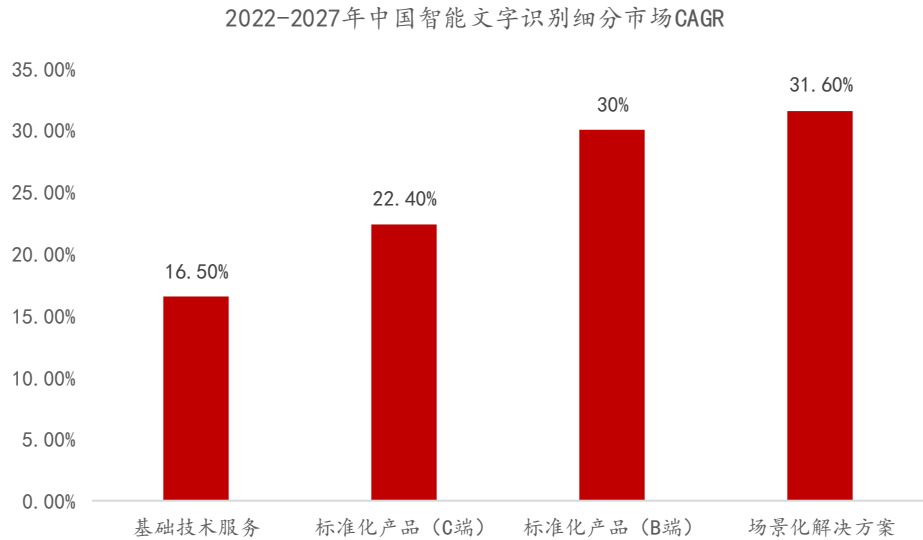
资料来源: verified market research, 国盛证券研究所

图表16: 2017-2027E 年中国智能文字识别服务市场规模统计



资料来源: 智研咨询, 国盛证券研究所

图表17: 2022-2027E 中国智能文字识别细分市场平均复合增速



资料来源: 智研咨询, 国盛证券研究所

智能文字识别强于传统 OCR 技术, 适用于更复杂多样的应用场景。智能文字识别融合了智能图像处理、基于深度学习的复杂场景文字识别、自然语言处理 (NLP) 等多项 AI 技术, 具备更多认知与理解能力, 可适应多语言、多版式、多样式等复杂场景, 并可应用到多个商业化场景中, 例如票据分类、证照票据结构化、合同关键信息抽取、智能审核等。传统 OCR 技术针对特定场景的文字图像进行建模, 受限于手工设计特征的单一表达能力和复杂的处理流程, 一旦跳出当前场景, 遇到复杂场景, 模型就会失效。而基于深度学习、NLP 的智能文字识别技术可解决传统 OCR 的缺点, 灵活应对自然场景下多形态文本以及手写文本。

图表18: 智能文字识别技术与传统 OCR 技术的区别

项目	传统 OCR 技术的特点	结合深度学习、NLP 的智能文字识别技术的特点
底层算法	需要假设模型, 人工特征提取, 无法做到完全的数据驱动, 依靠统计模式识别	基于深度学习算法, 无需假设模型, 完全的数据驱动与 NLP 技术相融合, 增强了对内容的“理解”能力
识别精度	识别精度不高, 仅对标准印刷文字的识别准确率较高	识别精度和适应的场景大幅度的提升
场景适应性、抗干扰性、稳定性	对输入的图像质量要求非常高, 无法处理成像复杂、文字复杂、场景复杂的图像	通过深度学习技术并与智能图像技术相结合, 自动地对曲形、褶皱、模糊、噪点、光线变化、形变、色彩等各种复杂场景下的文字图像进行建模和处理, 实现高性能的复杂场景文字识别性能

资料来源: 公司招股书, 国盛证券研究所

图表19: 智能文字识别技术的功能



资料来源: 公司招股书, 国盛证券研究所

AI 技术快速发展, 86% 企业文档有望走向自动化, 驱动 OCR 技术转向智能文字识别。

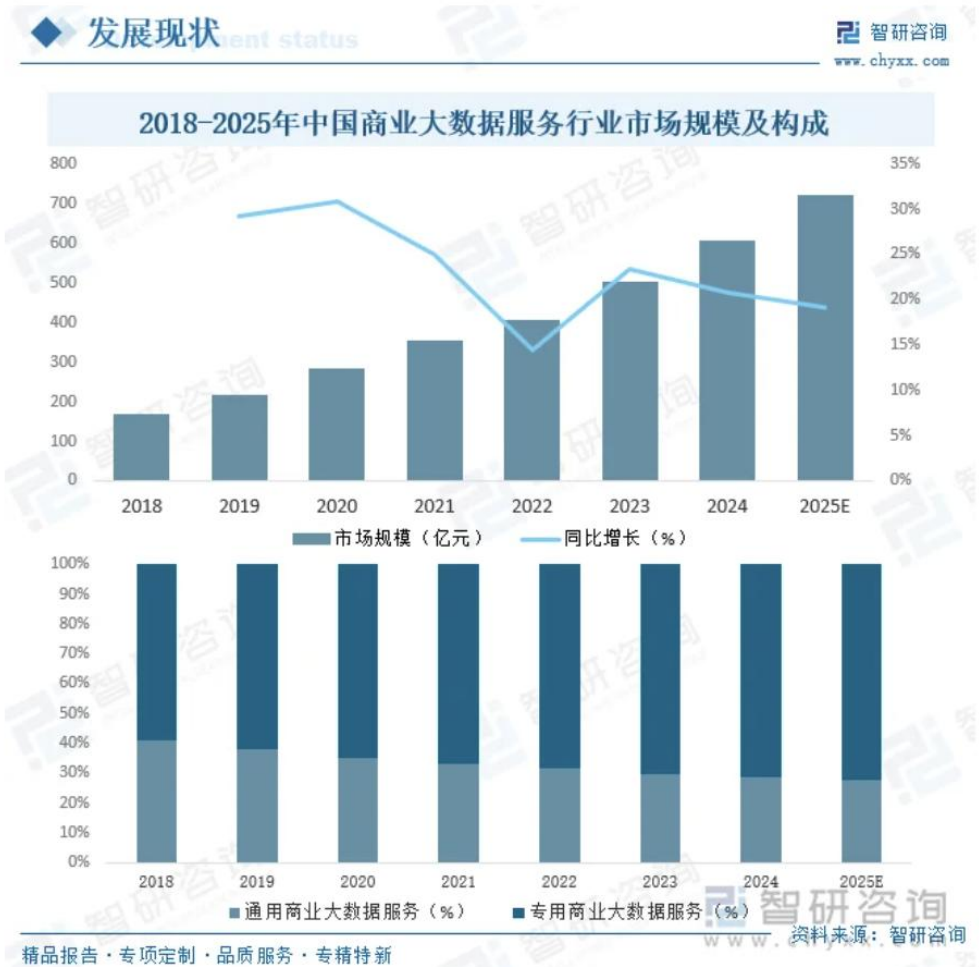
从技术上讲, OCR 正从传统的矩阵匹配快速转向智能字符识别 (ICR) 和基于神经网络的提取。这些先进的系统能够以超过 99% 的准确率识别复杂的手写体和多语言文本。这种增长根本上是受超自动化趋势的驱动, 预计 80% 的全球企业将在今年年底前采用文档自动化, 以缓解不断上升的劳动力成本并消除手动数据输入错误。

2.2 商业大数据市场方兴未艾, 数据价值深度发掘赋能下游行业

商业大数据服务属于新兴的大数据服务的一种, 通过深度挖掘数据价值赋能下游, 行业发展方兴未艾。商业大数据围绕企业的各类原始数据 (包括但不限于: 基本信息、股权、司法涉诉、信用、董监高、产业链、舆情等), 利用大数据挖掘、知识图谱等技术, 通过“数据、信息、知识、智能” 4 个层次, 挖掘商业数据背后蕴藏的价值, 赋能各行各业。

2024 年国内商业大数据服务行业市场规模已超 600 亿元, 市场规模广阔且增长迅速。根据智研咨询数据统计, 2024 年我国商业大数据服务行业市场规模达 605 亿元, 同比增长 20.76%, 其中, 通用商业大数据服务约占 29%, 专用商业大数据服务约占 71%。

图表20: 国内商业大数据服务行业市场规模及增速



资料来源: 智研咨询, 国盛证券研究所

按照下游客户类型及交付形式, 中国商业大数据服务可分为 **C端 APP**、**B端基础数据服务** (如 **API**、**数据包**)、**B端标准化服务** (如**数据终端**、**分析决策工具**、**SaaS 软件**等) 和 **B端场景化解决方案**四种形态。

图表 21: 商业大数据服务商业模式

分类	定义及形态	下游客户	未来发展分析
C 端 APP	商业数据查询平台 APP	主要为 C 端个人消费者	随着社会整体信用意识的提升, 个人用户在商业谈判、产品营销、求职、投资、采购等多种场景均对查询企业商业资料与信用数据具有需求, 商业数据查询 APP 的用户群体逐渐扩大, 覆盖行业包括金融、律师、媒体、工业、零售等, 预计商业数据查询 APP 市场将持续增长
B 端基础数据服务	以 API 接口或数据包形式将商业大数据提供给客户	下游 B 端企业客户类型多样	大数据已经渗透到各行各业, 成为重要的生产要素, 企业客户需要合规、维度完整、实时更新、稳定可靠的数据接口, 以满足其基于商业大数据的业务管理、监控、开发等需求, 预计 B 端基础数据服务将不断增长
B 端标准化服务	面向 B 端的标准化 SaaS (软件即服务) 产品, 提供征信报告、评级报告等标准化服务, 交付方式轻量化	主要为对标准化 SaaS 产品和服务有需求的企业, 例如中型企业或大型企业的部分部门	商业大数据可在智能营销、风险控制方面赋能中小型企业, 但中小型企业预算有限、IT 基础薄弱, 因此价格适合、即插即用的 SaaS 服务可满足中小型企业的需要, 预计商业大数据 SaaS 服务市场将快速增长
B 端场景化解决方案	根据客户场景化需求开发行业知识图谱、数据中台建设等大数据解决方案	主要为具有场景化需求、付费能力较强的大型企业	大型企业存在复杂的业务部门, 各个部门数据标准未统一、数据尚未打通, 需要贴合业务流程、符合其风控、营销、供应链管理、分析决策等方面的场景化解决方案。大型企业由于预算相对充足, 在国家政策的有利推动下, 对企业数字化转型的投入更具有付费意愿。预计商业大数据场景化解决方案市场将保持高速增长

资料来源: 公司招股书, 国盛证券研究所

商业大数据行业具备较高的技术壁垒+资质壁垒, 市场竞争格局相对集中。

1) 技术壁垒: 由于大数据行业的大数据挖掘、知识图谱等技术壁垒较高, 且数据采集投入较高, 并且商业大数据企业如果涉足企业征信领域, 需要树立公正、可信赖、具有知名度的品牌形象, 因此商业大数据市场竞争格局相对集中。由于数据具有“一次采集、多次应用、多领域应用”的特点, 首次采集数据的成本较高, 但使用数据的成本呈现边际成本递减的特点, 这一特点也助于头部企业强者恒强。

2) 资质壁垒: 截至 2024 年末, 中国人民银行各分支行公示的已备案企业征信机构共 154 家, 2022-2024 年各年末同比上一年末净增加的备案企业征信机构数量分别为 6 家、9 家、5 家, 新增备案企业数量较少, 每年亦有部分备案企业注销。在已备案的企业征信机构中, 已经开发出成熟的大数据产品的机构更是少数。因此, 企业征信备案构成行业一项较高的进入壁垒。

其中, 按 B 端/C 端细分市场来看:

- **C 端市场: 已成三足鼎立局面, “一信两查” 已占领用户心智。** 商业大数据 C 端 APP 市场集中度较高, 国内已形成三足鼎立的局面, 头部三家分别为启信宝、企查查、天眼查 (简称为“一信两查”), “一信两查” 已具备先发优势, 凭借各自的数据与服务特色, 积累了大量用户, 并形成品牌效应。“一信两查” 利用大数据挖掘技术, 突破原有征信数据的局限, 带来更丰富、深入的分析维度, 补充、完善了征信系统企业的覆盖率与信息的维度。近几年我国政府愈加重视第三方征信在社会信用体系建设的重要作用, 随着社会整体信用意识的提升, 个人用户在商业谈判、产品营销、求职、投资、采购等多种场景均对查询企业商业资料与信用数据具有需求, 商业数据查询 APP 的用户群体逐渐扩大, 覆盖行业包括金融、律师、媒体、工业、零售等, 为头部商业大数据企业带来发展机遇。
- **B 端市场: 集中度略低于 C 端市场, 碎片化需求带来更多通用/垂类玩家生存空间。** 商业大数据 B 端企业服务市场的集中度相对低于 C 端市场, 参与者包括本土大数据企业以及跨国企业的国内子公司。本土商业大数据 2B 服务商主要由“一信两查”、中数智汇以及一些垂直领域大数据服务商。

图表22: 商业大数据产业链图谱



资料来源: 智研咨询, 国盛证券研究所

3.竞争壁垒：C端深入用户心智，B端打通最后一公里，B+C协同发展沉淀扎实内功

3.1 公司深耕智能文字识别技术，准确率与实用性领先全球

➤ 智能文字识别方面：

公司多年深耕智能文字识别技术，识别准确率全球领先。公司自2006年成立起就将文字识别、图像处理作为核心研发方向，在智能文字识别领域已拥有超过10年的研发经验及技术积累，识别准确率稳居全球领先水平。在2019年的国际文档分析识别大会(ICDAR)中，公司研发团队取得表格识别竞赛的冠军，并与华南理工大学团队联合取得票据识别竞赛的冠军；在2020年国际模式识别旗舰会议(ICPR)中，公司研发团队在信息图表识别竞赛的7大任务14个子任务中与华南理工大学、联想研究院团队联合获得11项第一。

有效解决现实场景中的畸变、背景干扰等识别类技术挑战，具备更强可落地性。公司通过建模实现对复杂图像的精准预处理，例如对模糊图像进行去模糊化和图像增强，对形变图像进行还原矫正，对图像中因光照不均产生的阴影问题进行处理，对不同角度拍摄的图像进行定位和透视变换矫正。同时，公司通过对文本的类型、文档的整体拓扑结构信息、多个文本条目关联前后语义的联想等进行理解与分析，实现关键信息的定位和抽取和结构化输出，同时对已识别的字符进行纠错，例如错字、前后文漏字和符号错误等情况，提升识别精准度。

名片全能王字符识别率超过99%，扫描全能王APP对手写体文档字符平均识别率也高达97%。根据中国合格评定国家认可委员会(China National Accreditation Service for Conformity Assessment 英文缩写为：CNAS)认可的检测实验室(NO.CNASL8192)2020年出具的一系列《软件测试报告》：(1)名片全能王APP针对常规的多语言名片的字符识别率平均值为99.43%，名片识别对象包括简体中文、繁体中文、英文、日文、韩文及西欧语言；(2)扫描全能王APP针对常规的印刷体文档字符平均识别率为99.77%，手写体文档字符平均识别率为97.00%。

图表23: 公司智能文字识别技术能解决现实场景下的众多技术挑战



资料来源：公司招股书，国盛证券研究所

➤ 商业大数据方面：

积累海量数据，积极深度挖掘数据价值。2015 年，公司开始布局大数据产业的商业大数据细分领域，利用前期积累的深度学习、NLP 等 AI 技术的积淀，拓展知识图谱及大数据挖掘技术。公司内部的大数据底层平台快速成熟化，2018 年，公司自主研发的企业知识图谱解决方案开始进行商业落地；2019 年，商业大数据量已突破 700 亿条，汇集境内 2.1 亿家企业等组织机构；2020 年，商业大数据量已突破 1,000 亿条，汇集境内 2.3 亿家企业等组织机构。

公司在商业大数据领域的核心技术为知识图谱、大数据挖掘、NLP 技术，技术先进性主要体现在：

- (1) 多源异构的超大规模动态知识图谱构建技术：公司通过知识图谱、NLP 技术，将企业的工商、司法、信用、舆情等多种多源异构信息中的实体进行对齐和融合，并进一步通过 NLP 技术挖掘出不同实体之间的显性和隐性关系，构建出上亿的节点、数十亿条关系的超大规模商业知识图谱，并在此基础上，利用相关技术将实时更新的多源异构数据动态的融合到知识图谱中，实现知识图谱的动态更新。该技术也同时解决了企业知识图谱构建过程中的同名节点的实体消歧问题。
- (2) 超大规模知识图谱推理与挖掘技术：公司通过超大规模知识图谱推理与挖掘技术，对商业知识图谱进行推理与挖掘，获取更多新的商业关联关系。该技术还可以与客户的私有数据进行联合建模，进一步拓展业务场景的边界。例如对企业的股权信息与客户的担保信息等数据进行联合挖掘分析，自动挖掘出集团关系、一致行动人、实际控制人、受益人，集团内企业循环担保圈等企业数据产品。该技术在大规模知识图谱的循环圈检测任务中，将传统方法长达数月的计算时间和数 T (1T=1024G) 的存储量，缩减至 2 小时的计算量和 30G 的存储量。

3.2 C 端：已建立用户心智与广泛用户基础，高粘性+高迁移成本构成护城河

公司 3 款核心 APP 在全球拥有庞大的用户基础，2025H1 月活合计超过 1.8 亿。

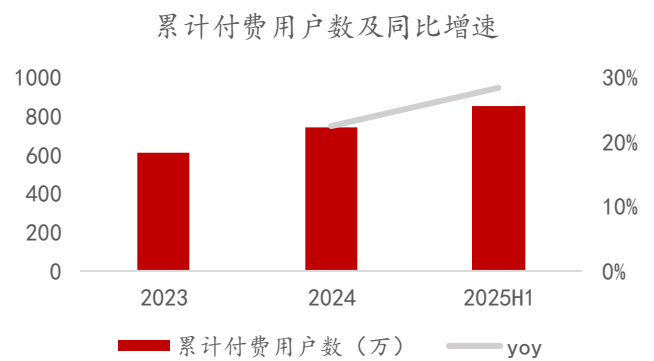
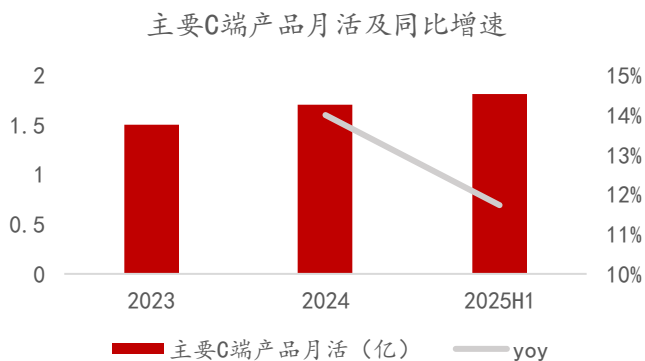
截至 2021 年 3 月末，公司扫描全能王、名片全能王、启信宝 APP 在 App Store 与 Google Play 应用市场的全球用户累计首次下载量合计超过 6 亿，2021 年 3 月的月活合计约 1.2 亿。截至公司 2025 年半年报，公司 C 端 APP 月活用户数超过 1.8 亿。巨大的 C 端用户基数为公司奠定了领先的行业地位。

品牌知名度+产品口碑已经建立，用户心智深入人心。公司 APP 品牌知名度与用户口碑俱佳，尤其是名片全能王、扫描全能王在 2009 年、2010 年上线，在品牌影响力方面具有约十年的深厚积累：早在 2010 年，名片全能王即被纽约时报（The New York Times）称为“为低效的名片管理带来有效的秩序”；在 2013 年，扫描全能王被时代周刊（Time）评选为“iPhone 必备的 50 个应用之一”、被华盛顿邮报（The Washington Post）报道为“手掌扫描仪”；在 2013 年，名片全能王被 BBC 报道为“更智能化的名片管理”；名片全能王在 2012-2016 年被上海市软件协会连续评为“年度优秀软件产品”；在 2017 年，扫描全能王被 App Annie 评为“一带一路最佳 APP”。根据 App Store 数据，扫描全能王付费版已被 App Store 精品推荐超过 660 次；名片全能王已被 App Store 精品推荐 440 多次。2020 年，启信宝获得上海大数据联盟颁发的“数字赋能先锋奖”。从 App Store 的搜索关键词来看，截至 2021 年 8 月初，“扫描全能王”的搜索热门程度高于“扫描”或者“扫描文档”，说明在用户心智中扫描全能王一定程度上已成为扫描的代名词。

广泛可延展的应用场景带动用户群体的不断拓宽，付费率持续上行。自 2023 年至 2025H1，根据公司官方财报可知，公司 C 端 APP 月活用户数、累计付费用户数持续上行，快速增长，且公司付费率逐年提升，至 2025H1 付费率已经上升至 4.71%，体现出公司产品市场认可度的持续提升。

图表24：公司主要 C 端产品月活用户数及同比增速

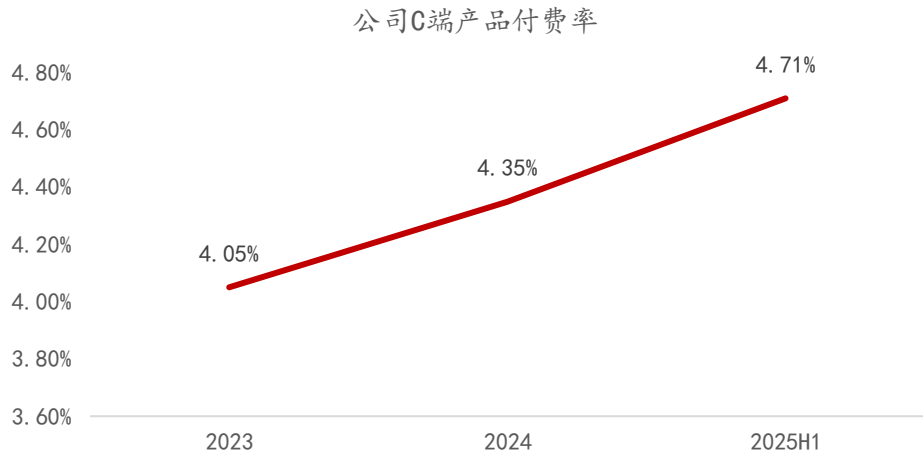
图表25：公司累计付费用户数及同比增速



资料来源：公司官方财报，国盛证券研究所

资料来源：公司官方财报，国盛证券研究所

图表26: 公司 C 端 APP 付费率逐步上行



资料来源: 公司官方财报, 国盛证券研究所

我们认为, 公司 C 端 APP 已经建立起用户心智与广泛用户基础, 高粘性+高迁移成本构成高竞争壁垒。从商业模式看, 用户在扫描全能王、名片全能王上持续积累了大量文档资产或名片资产, 用户粘性强、迁移成本高, 因此用户留存率高; 同时用户通过扫描全能王的文件分享、通过名片全能王交换电子名片时形成自然的网状裂变, 为产品推广带来低成本的扩张, 因此获客成本较低。从应用场景看, 用户可将扫描全能王作为文档资产管理工具, 将名片全能王作为人脉资产管理工具, 两款 APP 切实为用户带来无纸化移动办公的效率提升, 其中尤其是扫描全能王的应用场景尤为广泛, 不仅财会、行政、销售等商务场景对其具有应用需求, 教师、学生、家庭等生活场景也对证件、试卷、手写文件、报税单等文档具有电子化管理的需求, 其应用场景的可延展性强, 用户群体不断拓宽。因此扫描全能王、名片全能王自 2009、2010 年上线至今, 仍保持较高的发展速度, 具有较长的产品生命周期, 未来发展潜力较大。

3.3 B 端: 深耕优质客户+深度行业场景 know-how, 打通企业级 AI 应用最后一公里

公司深耕细作 B 端业务已近十年, 针对不同行业场景和客户痛点提出有针对性的解决方案。公司早在 2016 年即成立企业级智能解决方案事业部, 已经在 B 端业务方面探索出了成熟的路径, 培养了一支协作力度高的研发+咨询+销售+客户服务团队, 服务质量与交付效率均较高。公司销售团队对客户所在行业的真实业务场景、业务逻辑、产业链上下游等有充分的理解, 围绕客户的痛点, 充分调动公司内部如研发、实施、售后服务等各类资源。在售前阶段, 深入了解客户使用需求和现有痛点, 和研发团队及时沟通并推荐和客户需求匹配的产品, 对于需要进行场景化开发的产品, 公司的销售团队会对客户业务的使用场景进行多轮沟通, 并将调研结果及时反馈给研发人员, 研发团队进行产品策划、设计、开发、测试, 贴合客户业务场景做算法优化, 从易用性、稳定性提升客户体验, 从而在业务场景中创造价值。在售中和售后阶段, 公司会通过线上咨询、远程支持、现场服务等形式, 为客户做出准确的故障诊断, 提供专业、快速的解决方案。例如:

- 保险行业: 可提供启信宝商业大数据、证照机器人、票据机器人等。
- 地产行业: 可提供知识图谱、启信宝商业大数据、证照/合同/票据/表格机器人等。

图表27: 合合信息可为保险行业提供的服务及解决方案

保险业务流程	对公司服务的潜在需求	公司服务提供的价值
营销环节	接入启信宝的商业大数据	发现客户，将目标企业客户标签、筛选，对风险企业进行初步过滤
投保环节	证照机器人	对客户证照进行自动识别、信息录入
核保环节	接入启信宝的商业大数据	对客户进行全面的尽职调查按照中国人民银行的要求进行反洗钱查询
理赔环节	证照机器人、票据机器人	对证照、票据进行分类、录入、比对、解析将重复性高、骗保嫌疑高的资料与历史数据比对，查验出骗保嫌疑的理赔资料，为保险公司节约理赔支出

资料来源: 公司招股书, 国盛证券研究所

图表28: 合合信息可为地产行业提供的服务及解决方案

地产业务流程	对公司服务的潜在需求	公司服务提供的价值
营销环节	知识图谱	将企业的工商、司法、经营、风险等数据与地产相关的土地信息、城乡规划、工程信息、房屋信息、物业信息、行政处罚等数据相融合，为房地产企业构建全面精确的地产项目画像，增大地产项目交易成功率对公司承建的地产项目进行风险监测，当出现负面信息时及时处理，减少消息蔓延范围
招商环节	接入启信宝的商业大数据	通过启信宝的产业综合数据库，通过筛选模型，从中挑选出有投资布局该区域可能性的企业形成短名单，推进招商产业化、智能化、精准化
招标环节	接入启信宝的商业大数据	利用启信宝商业大数据将供应商进行筛选，进行关联关系排查，避免围标、串标、陪标等问题，对准入的供应商做全方位的资格审查，在招标后持续对供应商进行舆情、经营风险监控
审核环节	证照/合同/票据/表格机器人	地产行业具有较多的审批环节，涉及大量证照、合同、票据、表格，公司 AI 机器人可对各种版式的文档进行电子化的录入、核对

资料来源: 公司招股书, 国盛证券研究所

有效帮助客户降本增效、改善获客与风控，为下游客户提供实际业务价值。公司的智能文字识别服务、商业大数据服务帮助客户切实解决了降本增效、改善风控、高效获客的业务痛点，提供了较高的技术附加值。以合合信息官方披露的伟巴斯特客户案例为例，合合信息基于 OCR 技术+AI 算法，助力伟巴斯特解决供应商发票效率及合规风控难题：

- **伟巴斯特业务痛点在于传统发票处理耗费大量人工，效率较低。**伟巴斯特集团总部位于德国，是一家拥有百年历史的汽车零部件供应商。2001 年，伟巴斯特（中国）成立并迅速成为中国汽车天窗市场的领导者，在中国拥有 11 家工厂，年产值达 120 亿。此次伟巴斯特是首次与合合信息合作，实现中西联合解决复杂场景下的应付自动化项目，该项目也是伟巴斯特首个财务共享中心的创新项目。伟巴斯特每月约需处理 1.5 万张供应商采购发票，传统作业不仅耗时耗力，且各类成本较高，影响业务流转效率，还隐含少量难以察觉的财税风险。痛点如下：

- 发票处理量大：约 1.5 万-2 万张/月
 - 国内供应商发票需要人工处理
 - 对中国税法的合规校验效率低
 - 国内票与海外票需要同时管理
- **采用合合信息的票据智能解决方案，整体作业效率提升 500%~1000%。**合合信息票据智能解决方案，赋能企业发票自动接收、数据录入、安全稽核、ERP 三单匹配等场景。支持复杂单证/多语言票据识别，多项自动化核验，支撑不同系统应付流程。从供应商信息平台、海内外票据识别、合规稽查、数据转换、ERP 订单匹配等维度，帮助企业实现票据扫描识别输入、真伪验证、稽核检查、凭证归档，从而更好的建立数字中台以及决策分析。借助合合票据机器人+SAP ERP 系统，可助力伟巴斯特实现票据扫描、识别、验真、合规自动检查、发票与订单数据匹配、数字化数据、影像留存全流程财税管理，数字化的发票与订单管理让财税处理中的每个环节都可调阅可追溯，**整体作业效率提升 500%~1000%。**

我们认为，公司 B 端业务深耕优质客户，多元行业场景 know-how 使得公司得以打通企业级应用的最后一公里，为用户有效降本增效，有望建立较深的业务护城河。截至 2024 年，公司 B 端服务已覆盖了银行、证券、保险、政府、物流、制造、零售等行业的众多头部客户，《财富》杂志 2024 年 8 月发布的世界 500 强公司名单中，公司客户已覆盖超过 130 家。

图表29: 合合信息合作大量 B 端优质客户

共创共赢的合作伙伴



资料来源：合合信息官网，国盛证券研究所

3.4 B+C 联动：B 端与 C 端业务形成协同，联动发展提升研发与销售效率

公司兼具有爆发性成长潜力的 C 端业务、稳定盈利的 B 端业务，C 端与 B 端业务在研发、营销投入方面具备较高协同性，具体表现在：

1) **研发投入方面：技术算法模块可复用。**名片全能王、扫描全能王 APP 以智能文字识别技术模块为核心，启信宝 APP 以大数据技术为核心，在算法打磨、实际应用场景的落地方面均形成深厚积累，为后续发展 B 端企业服务、将 AI 技术在不同行业场景上实现真正的商业落地奠定了技术基础。B 端业务是由 C 端的 AI 核心技术输出，将 AI 技术在不同行业场景上实现真正的商业落地。另一方面，B 端已有近 30 类行业解决方案，在不同行业场景下的深度理解，以深度服务 B 端客户的过程中，可以反哺 C 端垂直场景的开发，例如：利用银行业的 B 端客户的风险门户开发经验，不断优化启信宝 APP 的风险信息模块。

2) 营销投入方面：C 端优良口碑有利于 B 端业务开展，可降低 B 端营销成本。C 端用户是口碑的重要传播者，C 端 APP 的忠诚用户可能是 B 端客户的关键决策者，其在使用 C 端 APP 时通过优质且流畅的用户体验已经对公司的技术能力、服务质量有了良好的印象，例如公司在启信宝企业版 SaaS 软件客户开发的过程中常常会遇到公司启信宝 APP 的忠实用户，而启信宝 APP 与企业版在底层技术、功能设计方面具有一定相似之处，因此 B 端客户可以迅速理解启信宝企业版的技术优势和服务价值；再例如，在拓展智能文字识别的 B 端服务项目时，B 端客户的关键决策者可能已经是扫描全能王或名片全能王的忠实用户，通过扫描全能王或名片全能王的使用感受已对公司的 AI 技术落地能力和服务价值拥有了深度的体验。因此公司 C 端产品巨大的用户量和良好的口碑可以促进公司 B 端业务的客户拓展，降低 B 端业务的营销成本。

我们认为，公司基于深度学习算法、智能文字识别、NLP、知识图谱等前沿技术，布局 C 端与 B 端丰富的产品矩阵，未来公司 C 端与 B 端业务联动的想象空间广阔，公司可进一步提升不同业务协同性，同时继续推出更具创新性的智能商业产品和解决方案，从而进一步打开市场空间。

4. 盈利预测

根据以上对公司业务的分析与理解，我们对公司 **2025E-2027E** 核心业务经营数据预测如下：

- 1) 智能文字识别-C 端 APP: 当前该业务正往海外拓展, 具备增长加速可能。考虑到公司历史上 2022-2023 年间该业务增速达到过 27-28%, 2024 年同比增速约为 23%, 在此基础上加速, 因此我们预计该业务 2025E/26E/27E 收入同比增速分别达到 30%/27%/25%。
- 2) 智能文字识别-B 端服务: 该业务公司具备技术积累, 而 B 端市场相对 C 端较为稳健, 因此有望在未来数年维持良性平稳的成长, 我们预计该业务 2025E/26E/27E 收入同比增速分别达到 15%/15%/15%。
- 3) 商业大数据-B 端服务: 该业务正处于快速成长期, 2024 年增速已达到 30%。考虑到公司成长的持续性, 我们预计该业务 2025E/26E/27E 收入同比增速分别达到 30%/27%/25%。
- 4) 商业大数据-C 端 APP: 考虑到过去两年该业务趋于收缩的态势, 我们预计该业务未来数年有望维持平稳, 预计该业务 2025E/26E/27E 收入同比增速分别达到 0%/0%/0%。
- 5) 其他主营业务: 体量较小, 预计未来数年维持平稳, 2025E/26E/27E 收入同比增速分别达到 0%/0%/0%。

图表30: 2025E-2027E 公司盈利预测

收入 (亿元)	2022A	2023A	2024A	2025E	2026E	2027E
智能识别系统	7.25	9.00	12.16	15.70	19.83	24.69
智能文字识别-C端 APP	6.54	8.33	11.41	14.83	18.84	23.55
智能文字识别-B端服务	0.71	0.67	0.75	0.86	0.99	1.14
商业大数据类	1.74	1.85	2.14	2.59	3.12	3.74
商业大数据-B端服务	0.98	1.15	1.50	1.95	2.48	3.10
商业大数据-C端 APP (启信宝)	0.76	0.70	0.64	0.64	0.64	0.64
互联网广告推广	0.82	0.96	-	-	-	-
技术授权	0.05	0.04	-	-	-	-
其他业务	0.02	0.01	0.05	0.05	0.05	0.05
总收入	9.9	11.9	14.4	18.3	23.0	28.5
收入同比增速						
智能识别系统	25%	24%	35%	29%	26%	24%
智能文字识别-C端 APP	28%	27%	23%	30%	27%	
智能文字识别-B端服务	4%	-6%	11%	15%	15%	15%
大数据类	23%	6%	15%	21%	20%	20%
商业大数据-B端服务	44%	17%	30%	30%	27%	25%
商业大数据-C端 APP (启信宝)	4%	-8%	-9%	0%	0%	0%
其他主营业务	N/A	N/A	N/A	0%	0%	0%
总收入同比增速	23%	20%	21%	28%	25%	24%

资料来源: wind, 国盛证券研究所

注: 2024 年起, 为更清楚体现收入与产品的相关性, 保证产品收入分类与公司未来发展方向更加一致, 充分体现出公司智能文字识别与商业大数据、C 端与 B 端产品服务的特性, 公司 2024 年主营业务收入分类口径将 APP 互联网广告收入合并至 APP 收入中进行披露, 将收入占比较低的手机厂商技术授权合并至其他。因此“其他主营业务”的同比增速一栏中, 2022-2023 年数据填写为 N/A (不适用)。

我们对公司各项分业务毛利率预测如下:

- 6) 智能文字识别-C端 APP: 该业务商业模式为订阅制, 因此有望长期维持高企的毛利率, 我们预计该业务 2025E/26E/27E 毛利率分别达到 88%/88%/88%。
- 7) 智能文字识别-B端服务: 由于该业务提供的是面向 B 端的服务, 因此该业务毛利率预计将略低于智能文字识别 C 端 app 业务, 我们预计该业务 2025E/26E/27E 毛利率有望稳定在 62.5%/62.5%/62.5%。
- 8) 商业大数据-B端服务: 由于该业务提供的是面向 B 端的服务, 因此该业务毛利率预计将略低于商业大数据 C 端 app 业务, 参考 2024 年公司该业务毛利率情况, 我们预计该业务 2025E/26E/27E 毛利率有望稳定在 66%/66%/66%。
- 9) 商业大数据-C端 APP: 该业务商业模式为订阅制, 边际成本极低, 因此有望长期维持高企的毛利率, 我们预计该业务 2025E/26E/27E 毛利率分别达到 91%/91%/91%。
- 10) 其他主营业务: 体量较小, 预计未来数年毛利率有望维持平稳, 2025E/26E/27E 毛利率分别达到 97%/97%/97%。

图表31: 2025E-2027E 公司毛利率预测

毛利率	2024A	2025E	2026E	2027E
智能识别系统	86.2%	86.6%	86.7%	86.8%
智能文字识别-C 端 APP	87.8%	88%	88%	88%
智能文字识别-B 端服务	62.3%	62.50%	%	62.50%
大数据类	73.6%	72.2%	%	70.3%
商业大数据-B 端服务	65.9%	66%	66%	66%
商业大数据-C 端 APP (启信宝)	91.5%	91%	91%	91%
其他主营业务	97.6%	97%	97%	97%
总收入口径毛利率	84.4%	84.6%	84.6%	84.7%

资料来源: 公司官方年报, 国盛证券研究所

注: 公司 2024 年年报调整过一次披露口径, 故毛利率无 24 年前同比口径数字

我们对公司各项费用率预测如下:

- 1) 销售费用率: 公司当前正处于成长阶段, 并扩张海外业务中, 对于销售费用仍有客观需求, 我们预计 2025E/26E/27E 公司销售费用率分别达到 29%/29%/29%。
- 2) 管理费用率: 公司管理稳定, 经营质量健康, 我们预计公司 2025E/26E/27E 管理费用率分别达到 6%/6%/6%。
- 3) 研发费用率: 公司是 AI 及大数据企业, 我们预计公司将坚定持续投入研发, 2025E/26E/27E 研发费用率预计分别达到 27.2%/27.2%/27.2%。

图表32: 公司销售/管理/研发费用率情况

收入 (亿元)	2022A	2023A	2024A	2025E	2026 E	2027 E
销售费用/营业收入	27.19%	28.24%	29.84%	29.00%	29.00%	29.00%
管理费用/营业收入	6.17%	6.41%	5.84%	6.00%	6.00%	6.00%
研发费用/营业收入	28.32%	27.25%	27.15%	27.20%	27.20%	27.20%

资料来源: wind, 国盛证券研究所

投资建议: 首次覆盖, 给予“买入”评级。

- 1) 基于我们对公司基本面、商业模式以及行业趋势的判断与分析, 我们预计公司 2025E/26E/27E 有望实现总营业收入 18.3/23.0/28.5 亿元, 同比增速达到 28%/25%/24%。我们预计公司 2025E/26E/27E 将分别实现归母净利润 4.87/6.10/7.55 亿元, 同比增长 21.5%/25.4%/23.8%。
- 2) 我们选取金山办公、福昕软件、万兴科技和智谱作为可比公司, 其中金山办公与福昕软件为办公软件企业, 其商业模式与业务属性与合合信息相吻合; 万兴科技为多模态 AI 创作工具企业, 其产品具备生产力工具属性, 与合合信息的产品业务属性相似; 智谱为国内 AI 大模型企业, 从人工智能这一技术维度而言, 与合合信息智能文字识别业务属性相匹配, 属于同业可比公司。截至 2026/1/31, 以上可比公司的 wind 一致预期估值, 按平均 PS (2026E) 计算约为 25.5x, 而合合信息仅为 17.4x, 公司 PS 估值低于可比公司平均水平, 具备估值性价比, 首次覆盖, 给予“买入”评级。

图表33: 同业可比公司估值情况 (截至 2026/1/31)

证券代码	证券简称	wind 一致预测 2026E 营业收入 (亿元)	总市值 (亿元)	PS (2026E)
688111.SH	金山办公	71.56	1,560.17	21.80
688095.SH	福昕软件	11.64	97.92	8.41
300624.SZ	万兴科技	18.95	180.48	9.52
2513.HK	智谱	16.00	995.80	62.24
可比公司平均				25.49
688615.SH	合合信息	22.41	390.60	17.43

资料来源: wind, 国盛证券研究所

注: 可比公司盈利预测取自 wind 一致预期

5.风险提示

AI 技术迭代不及预期的风险: 若 AI 技术整体迭代不及预期, 无法达到预想效果, 则公司产品迭代可能不及预期, 可能对公司业务产生不利影响。

海外业务发展不及预期的风险: 若公司海外业务拓展不及预期, 公司收入源头受限, 可能对公司业绩产生不利影响。

宏观经济波动的风险: 若全球宏观经济遭受波动, 下游支出意愿可能收缩, 进而使得公司业绩层面承压。

免责声明

国盛证券股份有限公司（以下简称“本公司”）具有中国证监会许可的证券投资咨询业务资格。本报告仅供本公司的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告的信息均来源于本公司认为可信的公开资料，但本公司及其研究人员对该等信息的准确性及完整性不作任何保证。本报告中的资料、意见及预测仅反映本公司于发布本报告当日的判断，可能会随时调整。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本公司力求报告内容客观、公正，但本报告所载的资料、工具、意见、信息及推测只提供给客户作参考之用，不构成任何投资、法律、会计或税务的最终操作建议，本公司不就报告中的内容对最终操作建议做出任何担保。本报告中所指的投资及服务可能不适合个别客户，不构成客户私人咨询建议。投资者应当充分考虑自身特定状况，并完整理解和使用本报告内容，不应视本报告为做出投资决策的唯一因素。

投资者应注意，在法律许可的情况下，本公司及其本公司的关联机构可能会持有本报告中涉及的公司所发行的证券并进行交易，也可能为这些公司正在提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。

本报告版权归“国盛证券股份有限公司”所有。未经事先本公司书面授权，任何机构或个人不得对本报告进行任何形式的发布、复制。任何机构或个人如引用、刊发本报告，需注明出处为“国盛证券研究所”，且不得对本报告进行有悖原意的删节或修改。

分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的任何观点均精准地反映了我们对标的证券和发行人的个人看法，结论不受任何第三方的授意或影响。我们所得报酬的任何部分无论是在过去、现在及将来均不会与本报告中的具体投资建议或观点有直接或间接联系。

投资评级说明

投资建议的评级标准		评级	说明
评级标准为报告发布日后的6个月内公司股价（或行业指数）相对同期基准指数的相对市场表现。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；香港市场以摩根士丹利中国指数为基准，美股市场以标普500指数或纳斯达克综合指数为基准。	股票评级	买入	相对同期基准指数涨幅在15%以上
		增持	相对同期基准指数涨幅在5%~15%之间
		持有	相对同期基准指数涨幅在-5%~+5%之间
		减持	相对同期基准指数跌幅在5%以上
	行业评级	增持	相对同期基准指数涨幅在10%以上
		中性	相对同期基准指数涨幅在-10%~+10%之间
		减持	相对同期基准指数跌幅在10%以上

国盛证券研究所

北京

地址：北京市东城区永定门西滨河路8号院7楼中海地产广场东塔7层
 邮编：100077
 邮箱：gsresearch@gszq.com

南昌

地址：南昌市红谷滩新区凤凰中大道1115号北京银行大厦
 邮编：330038
 传真：0791-86281485
 邮箱：gsresearch@gszq.com

上海

地址：上海市浦东新区南洋泾路555号陆家嘴金融街区22栋
 邮编：200120
 电话：021-38124100
 邮箱：gsresearch@gszq.com

深圳

地址：深圳市福田区福华三路100号鼎和大厦24楼
 邮编：518033
 邮箱：gsresearch@gszq.com