



关于上海合合信息科技股份有限公司
首次公开发行股票并在科创板上市申请文件的
审核中心意见落实函的回复

保荐机构（主承销商）



北京市朝阳区建国门外大街1号国贸大厦2座27层及28层

上海证券交易所：

贵所《关于上海合合信息科技股份有限公司首次公开发行股票并在科创板上市的审核中心意见落实函》（上证科审〔2023〕414号，以下简称“本次审核中心意见落实函”）已收悉。

根据贵所的要求，上海合合信息科技股份有限公司（以下简称“合合信息”、“发行人”或“公司”）会同中国国际金融股份有限公司（以下简称“保荐机构”）、上海市锦天城律师事务所（以下简称“发行人律师”）、众华会计师事务所（特殊普通合伙）（以下简称“申报会计师”）对本次审核中心意见落实函中所提问题逐项核查，具体回复如下，请予审核。

说 明

如无特别说明，本回复使用的简称与《上海合合信息科技股份有限公司首次公开发行股票并在科创板上市招股说明书》中的释义相同。

| | |
|----------------------|--------|
| 本次审核中心意见落实函所列问题 | 黑体 |
| 对本次审核中心意见落实函所列问题的回复 | 宋体 |
| 对招股说明书的修订、补充及反馈回复的更新 | 楷体（加粗） |

在本回复中，若合计数与各分项数值相加之和在尾数上存在差异，均为四舍五入所致。

目 录

| | |
|--------------------|---|
| 说 明 | 2 |
| 目 录 | 3 |
| 1.关于媒体质疑情况自查 | 4 |

1.关于媒体质疑情况自查

一、请保荐机构自查与发行人本次发行上市相关的媒体质疑情况，并就重大媒体质疑核查并发表意见。

回复：

针对合合信息科创板 IPO 项目，保荐机构、发行人律师、申报会计师持续关注媒体等对本次发行上市相关的报道，并就相关文章涉及的发行人信息披露内容进行了充分核查，核查范围包括自发行人 2021 年 9 月 27 日首次申报至本次审核中心意见落实函回复签署之日与合合信息公开发行上市相关的负面报道，核查数量包括约 50 篇原创报道。其中主要媒体质疑情况如下：

一、媒体质疑总体情况

自发行人于首次申报至本次审核中心意见落实函回复签署之日，除仅对发行人招股说明书相关内容进行简单摘录的报道以及内容相似的重复性报道外，媒体报道的主要质疑关注点包括：1、公司经营资质问题；2、海外经营的政策风险；3、科创属性问题；4、数据安全与个人隐私保护问题；5、启信宝 APP 合规及数据准确性；6、扫描全能王 APP 业务发展可持续性问题；7、名片全能王 APP 业务发展可持续性问题；8、启信宝 APP 业务发展可持续性问题；9、C 端收入真实性；10、私募基金股东的出资人委托持股问题；11、扫描全能王 APP 自动订阅问题；12、本次募集资金规模的必要性；13、股权激励在 2019 年集中支付的合理性；14、2018 年-2020 年前五大客户的占比下降；15、2019 年存在未弥补亏损问题；16、2018 年度至 2021 年 1-3 月期间费用率较高。

二、核查情况

截至本次审核中心意见落实函回复签署之日，针对财经媒体发布质疑报道的上述关注问题，保荐机构、发行人律师、申报会计师的核查情况如下：

（一）公司经营资质问题

1、媒体质疑情况

媒体关注到：上交所要求发行人就未及时取得征信备案和增值电信业务许可、

未履行境外投资备案手续的原因进行说明。

2、核查情况

保荐机构、发行人律师就该问题的核查情况如下：

(1) 发行人已完成未及时取得企业征信机构备案的整改工作

1) 发行人采取的整改措施及结果

发行人及子公司苏州贝尔塔、上海生腾在报告期内存在未取得企业征信机构备案即从事企业征信业务的情形。发行人按照《征信业管理条例》《征信业务管理办法》等法律法规要求采取的具体整改程序、措施及结果如下：

①发行人确定由子公司上海生腾办理企业征信机构备案并统一开展征信类业务：2020年7月24日，人行上海分行发布关于完成对上海生腾的企业征信机构备案的公告。

②发行人C端及B端征信业务全部迁移至上海生腾，并已在法规规定的整改期内整改完毕：企业征信C端业务已于2018年12月完成整改。企业征信B端业务中，新增征信类业务已于2021年6月1日完成整改，存量征信类业务已于2022年6月29日完成整改。

2) 发行人已就企业征信业务的整改方案向主管机关进行了完整汇报。发行人的整改措施、整改程序和整改结果符合监管要求

2022年2月，上海市静安区金融服务办公室组织召开协调工作会议，会上发行人就公司征信业务情况及整改方案向人行上海分行进行了完整汇报，上海市静安区金融服务办公室已于2022年2月11日出具《关于合合信息征信业务相关问题协调工作的情况说明》，对上述情况予以确认。此外，人行上海分行出具情况说明，确认2017年8月1日至2023年1月31日期间，人行上海分行对上海生腾未作出过行政处罚。截至本次审核中心意见落实函回复签署之日，发行人及其子公司未受到征信主管部门的任何处罚。发行人的整改措施、整改程序和整改结果符合监管要求。

(2) 发行人已就上海找贝及上海生腾未及时取得增值电信业务经营资质的情形完成整改工作

1) 发行人采取的整改措施及结果

发行人子公司上海找贝和上海生腾存在未及时取得增值电信业务许可而开展增值电信业务的情形,上海找贝、上海生腾已根据《中华人民共和国电信条例》等法律法规要求积极自行整改。上海找贝已于 2019 年 12 月 16 日取得《增值电信业务经营许可证》,上海生腾已于 2020 年 8 月 17 日取得《增值电信业务经营许可证》。

2) 发行人的整改措施、整改程序和整改结果符合监管要求

上海市静安区金融服务办公室及上海市静安区科学技术委员会已于 2022 年 3 月 21 日出具《情况说明》,经上述部门与上海市通信管理局沟通协调,就上海生腾、上海找贝曾存在未办理《增值电信业务经营许可证》即开展相关业务的情形,确认鉴于上海生腾、上海找贝已自行整改并取得《增值电信业务经营许可证》,未对国家增值电信业务管理造成重大不利影响,因此上述事宜不构成重大违法违规,上海市通信管理局也未因此对合合信息、上海生腾、上海找贝进行任何立案调查或行政处罚。

根据工业和信息化部司局简函《关于为拟上市企业开具合规证明相关问题的复函》(工信管函[2020]1658 号),确认开具“合规证明”不属于电信管理机构《权力清单和责任清单》所列示职责事项,电信业务经营者可通过电信业务综合管理平台(<https://tsm.miit.gov.cn>)自行查询相关信息。经检索电信业务综合管理平台、上海市通信管理局、中国裁判文书网、中国执行信息公开网、信用中国等网站,截至本次审核中心意见落实函回复签署之日,公司及子公司未因上述事宜受到主管部门的任何处罚。

(3) 发行人已完成境外投资审批手续瑕疵的整改工作

1) 发行人已将投资设立 CCI 未履行境外投资备案手续事宜向政府部门进行了完整汇报,并完成整改工作

发行人报告期内曾经的美国子公司 CCI 是发行人境外扫描全能王免费下载

版 APP 对应 Google Play 平台的收款主体，CCI 设立时未履行境内审批程序。发行人已将前述事宜向政府部门进行了汇报，并完成了整改工作。

2021 年 3 月，发行人向上海市商务委提交了《关于上海合合信息科技股份有限公司及其控股子公司境外投资合规性的请示》，就 CCI 相关事宜进行了汇报。2021 年 4 月 20 日，在上海市科学技术委员会的组织协调下，上海市静安区政府等相关政府部门召开了合合信息拟申报科创板上市工作专题会议，并就 CCI 事宜进行了讨论。由于无法为 CCI 补办《企业境外投资证书》及《境外投资项目备案通知书》，公司决定通过报批新设境外企业，将 CCI 原有的业务转移至新设境外企业的方式来解决此法律瑕疵。

对此，发行人已根据主管部门意见及法律法规要求积极自行完成整改：上海临冠的全资子公司上海又冠分别于 2021 年 5 月 26 日、2021 年 6 月 1 日完成新设美国子公司 CamSoft 的相关境外投资境内主管部门审批程序及外汇备案。CCI 已于 2021 年 7 月 23 日将当时的全部业务迁移至 CamSoft。CamSoft 已代替 CCI 成为公司境外扫描全能王免费下载版 APP 对应 Google Play 平台的收款主体，公司后续通过 CamSoft 收取相关款项，经营正常。

2022 年 4 月 19 日，发行人完成了 CCI 的注销手续，美国律师事务所 Squire Patton Boggs (US) LLP 已出具专项法律意见书，对 CCI 注销程序的合法合规性进行了确认。至此，发行人上述不合规的情况已经终止，发行人已完成所有整改工作。

2) 发行人通过将 CCI 的业务转移至 CamSoft 的方式进行整改符合监管要求

上海市静安区金融服务办公室及上海市静安区科学技术委员会已于 2022 年 3 月 21 日出具《情况说明》，经上述部门与上海市商务委、上海市发改委沟通协调，确认合合信息在发现 CCI 存在法律瑕疵后及时进行了整改，CCI 业务已由合合信息子公司 CamSoft 承接，且 CamSoft 在设立时依法履行了相应的境外投资境内主管部门审批程序。上述事宜未对国家境外投资管理造成重大不利影响，因此不构成重大违法违规。

公司已将投资设立 CCI 未履行境外投资备案手续事宜向政府部门进行了汇报。截至本次审核中心意见落实函回复签署之日，发行人未因投资 CCI 未履行境外投资备案手续受到主管部门的行政处罚，CCI 的业务转移、注销进展等事宜符合相关监管要求。

针对公司曾经存在的经营资质问题，保荐机构已在《关于上海合合信息科技股份有限公司首次公开发行股票并在科创板上市申请文件的第二轮审核问询函的回复》（以下简称为“二轮问询函回复”）之“1.关于合规经营”中进行了详尽的说明。发行人实际控制人镇立新已出具承诺，如发行人及其子公司因业务资质方面的法律瑕疵遭受任何处罚或损失，其将就此向发行人或其子公司全额补偿并承担相关费用。

（4）发行人已制定了一系列的内控制度及措施，在新业务开展前进行严格审批

为提升公司业务团队在新业务、新产品开发过程中的合法合规性，公司制定了《新业务及产品开发与设计流程合规规范》，在新业务开展前进行严格审批，主要流程如下：**1）新业务需求评估阶段：**由产品负责人将需求汇报至安全与合规管理部、法务部及董事会办公室，由3大部门共同评估该等新业务是否需要相关资质及合法合规性；**2）新业务上线前的申请阶段：**待需求评估阶段通过后，由产品负责人发起内部申请流程。该等申请流程须经公司安全与合规管理部、法务部、市场部、董事会办公室、财务部、事业部负责人核验申请材料并审批通过。

经核查，保荐机构、发行人律师认为：截至本次审核中心意见落实函回复签署之日，针对发行人曾经存在的未及时取得征信备案和增值电信业务许可、未履行境外投资备案手续的情况，发行人已向相关部门进行了完整汇报，并完成了整改工作。截至本次审核中心意见落实函回复签署之日，发行人不存在因未及时取得征信备案和增值电信业务许可、未履行境外投资备案手续而受到主管部门处罚的情形。发行人已制定了一系列的内控制度及措施，在新业务开展前进行严格审批。前述事宜不构成重大违法违规，不会对发行人业务开展造成重大不利影响，也不构成发行人本次发行上市的障碍。

（二）海外经营的政策风险

1、媒体质疑情况

部分媒体关注公司海外业务面临的印度政府及美国联邦政府的政策风险。

2、核查情况

保荐机构就该问题的核查情况如下：

针对公司境外业务的政策风险，保荐机构核查了印度政府及美国联邦政府的政策变动对发行人过往历史业绩的影响。

2020年6月，印度政府宣布封杀59款中国应用程序，扫描全能王和名片全能王在印度主流应用市场被迫下架。该等事项导致扫描全能王和名片全能王在印度地区产生的收入出现下滑，2020年-2022年收入分别为716.58万元、245.24万元和31.77万元，占公司各期总收入比重分别为1.24%、0.30%和0.03%。

2021年1月5日，美国时任总统特朗普签署行政命令，禁止美国司法管辖范围内的企业及个人与8家中国应用软件进行交易，其中包括扫描全能王，2021年6月11日，该项行政命令被新任美国总统拜登撤销。因美国联邦政府的该项禁令未实际执行且后续被撤销，该等事项对公司报告期内业绩无明显影响。

尽管出现前述政策变动，公司扫描全能王和名片全能王的境外收入在报告期内依然保持增长，2020年-2022年，扫描全能王境外收入分别为14,806.79万元、23,587.13万元和28,266.42万元，名片全能王境外收入分别为966.54万元、1,033.01万元和1,203.43万元。

就发行人海外经营风险，发行人已在招股说明书“第三节 风险因素”之“二、与行业相关的风险”之“（二）海外经营风险”中对政策风险的未来潜在影响进行了特别提示。

经核查，保荐机构认为：报告期内印度政府及美国联邦政府的政策变动对发行人过往历史业绩的影响较小，发行人已在招股书对海外业务的政策风险进行了完善提示。

（三）科创属性问题

1、媒体质疑情况

部分媒体报道质疑发行人的科创属性问题：（1）发行人的核心技术是否符合科创板定位方向，营收主要来源于 C 端业务，是否符合科创成色的要求；（2）文字识别技术是否已发展较为成熟，技术成长空间是否较大；（3）复杂场景的智能文字识别是否是主流市场需求，潜在市场需求是否较大；（4）合合信息已有发明专利中与大数据相关的专利较少；（5）合合信息近三年每年新增已授权专利数量较少，是否与公司连年递增的研发费用投入具有匹配性，专利数量少于腾讯、百度、金山办公是否代表着公司的科创成色不足；（6）2020 年公司研发费用率相比 2019 年出现大幅下降的合理性；（7）服务器及带宽费用划入研发费用的合理性。

2、核查情况

保荐机构、申报会计师就该问题的核查情况如下：

（1）发行人核心技术方向符合科创板定位方向

1) AI 和大数据技术是公司的关键核心技术，属于科创板支持的“硬科技”

公司核心技术为智能文字识别及商业大数据技术：①智能文字识别技术是人工智能（以下简称“AI”）领域的一个重要分支，融合了智能图像处理、复杂场景文字识别、NLP 等 AI 技术，相比传统简单文字识别，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并形成落地的产品或服务，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；②商业大数据技术属于大数据领域的细分领域，融合了大数据挖掘与知识图谱等技术，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。

根据国家发改委在官网发布的对于“硬科技”的名词解释，人工智能和信息技术属于目前硬科技的代表性领域。根据《上海证券交易所科创板企业发行上市申报及推荐暂行规定（2022 年 12 月修订）》：“申报科创板发行上市的发行人，应当属于下列行业领域的高新技术产业和战略性新兴产业：（一）新一代信息技

术领域，主要包括半导体和集成电路、电子信息、下一代信息网络、人工智能、大数据……”。

因此，公司的智能文字识别、商业大数据核心技术属于“硬科技”，属于科创板支持的领域。

2) 公司核心技术方向属于国家战略重点支持的领域

针对发行人核心技术方向是否属于国家战略重点支持的领域，发行人已在《关于上海合合信息科技股份有限公司首次公开发行股票并在科创板上市申请文件的第一轮审核问询函的回复》（以下简称为“一轮问询函回复”）之“1.关于科创属性”之“一、（一）发行人产品及服务与国家科技创新战略和相关产业政策的匹配情况，是否属于国家战略重点支持的领域”中进行了详尽的说明。相关核心内容总结如下：

发行人智能文字识别与商业大数据的产品及服务与国家科技创新战略和相关产业政策匹配，属于国家战略重点支持的领域。

智能文字识别方面：①人工智能已被纳入我国“十三五”与“十四五”纲要，复杂场景文字识别技术已被科技部列入“新一代人工智能共性关键技术”，智能文字识别业务响应了国家“促进人工智能和实体经济深度融合”、“加快数字化发展、建设数字中国”的政策；②智能文字识别技术为国家限制出口的关键技术之一，具有重大的国家战略意义；③公司的智能文字识别技术促进无纸化办公，有助于减少碳排放量，响应了节能减排的国家战略。

商业大数据方面：①大数据已被纳入我国“十三五”与“十四五”纲要，国家政策重视数据要素市场的培育、数据要素价值的利用及释放，公司商业大数据技术对数据要素的价值进行充分挖掘、有效释放；②公司商业大数据业务响应了国家“促进大数据与各行业应用的深度融合”的政策；③公司子公司作为企业征信机构积极参与信用大数据的建设与服务，响应了国家“发展第三方大数据服务产业”、“推进社会信用体系建设”的号召。

3) 公司核心技术处于行业领先水平

针对核心技术的技术水平及技术壁垒问题，发行人已在一轮问询函回复之“1.

关于科创属性”之“一、（二）发行人技术来源，底层技术和通用技术的具体内涵、技术门槛、行业技术水平及成熟度，底层技术系发行人原创、还是在行业通用技术上的改良创新或直接应用”与“1.关于科创属性”之“一、（三）发行人核心技术主要为行业通用技术还是专有技术，技术实现是否依赖第三方硬件，与行业内主要竞争对手的先进性对比情况，‘行业领先’、‘世界领先’相关表述是否客观准确，如否请删除或修改表述”中进行了详尽的说明。相关核心内容总结如下：

通过 10 余年持续的自主研发与技术积累，公司在复杂场景文字识别、智能图像处理、NLP、知识图谱、大数据挖掘等核心技术领域均处于行业领先地位。

①智能文字领域

智能文字识别技术是公司凭借多年技术研发活动中积累的自主研发的算法模型、C 端产品和 B 端服务开发中沉淀的专有技术，而非直接运用行业通用技术。

2023 年 5 月，公司加入工信部下属的中国信通院（工信部直属科研事业单位）牵头的《文档图像篡改检测标准》制定工作，该项标准将确立文档图像篡改检测技术的行业标准、为 AI 视觉图像行业健康发展提供指引，由中国图象图形学学会、中国科学技术大学等业内知名机构联合编制。

2023 年 5 月，公司在中国泰尔实验室（隶属工信部直属科研单位中国信通院的第三方检测实验室，是中国质量认证中心等认证机构的签约实验室）的 OCR 智能化服务评估和智能文档处理系统技术能力模块评估中获得最高评级，包括识别性能、AI 核心能力、数据安全性及可靠性等多维度评估。

在 2023 年国际文档分析识别大会（ICDAR，全球文档图像分析识别领域公认的权威学术会议）的文本图像篡改检测算法竞赛（Competition On Detecting Tampered Text in Images）中，公司研发团队取得文本篡改检测任务（Text Manipulation Detection）的冠军。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。

在 2020 年、2022 年 CNAS（中国合格评定国家认可委员会 China National Accreditation Service for Conformity Assessment）认可的检测实验室出具的一系列

测试报告中，公司智能文字识别技术的识别率行业领先。

公司智能文字识别的技术壁垒主要是根据用户痛点及客户需求去分析处理文档、制定出适应性强的 AI 算法迭代框架和模型。将 AI 技术在资源受限的移动端实现大规模应用的技术门槛较高，需要克服多个技术难关，例如如何在资源受限的移动端设备上设计兼顾性能和效率的轻量级深度学习模型，并克服特定的硬件架构对模型算子不支持的情况，针对特定的硬件架构对模型进行调优，以获得更好的性能和用户体验；同时还要解决用户在不同的使用场景和环境下形成的一系列复杂情况，例如反光、倾斜、背景干扰、弯曲等各种难题。不仅如此，公司也将 AI 技术在 B 端近 30 个行业的业务场景下实现商用级别的广泛应用。

尽管部分媒体报道质疑公司报告期内超过 60% 的收入来自扫描全能王 APP 业务、继而质疑公司的“硬科技”属性，但收入结构特点并不影响公司技术层面满足科创属性认定，扫描全能王 APP 的底层核心技术属于 AI 领域、属于科创板鼓励的“硬科技”方向，APP 作为公司核心技术的应用载体，有利于公司将 AI 技术实现商业化落地，具体分析如下：将 AI 技术落地在 APP 上同时实现全球数亿用户的优质使用体验，在功能设计、使用效率、产品稳定性以及系统安全性等方面的技术门槛高，研发投入高，具备较高的产品附加值，同时 APP 作为标准化程度高的软件，边际成本低，可通过付费订阅服务等方式在全球上百个国家和地区实现快速部署及商业化应用。

综上分析，公司的核心技术属于科创板鼓励的“硬科技”方向，扫描全能王 APP 在全球收入规模大、付费用户多、用户体验佳、市场认可度高，证明了公司在 AI 领域优秀的商业化落地能力。

②商业大数据领域

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：

A.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，可实现高效动态、可实现局部变化数据的检测、高效传输和局部的数据更新融合，使得知识图谱具备自学习、自更新的功能。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。

B.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。针对图谱节点上亿、关系复杂、挖掘和推理的效率低、时间漫长等技术难点，公司采用了子图分割、分布式图计算、图合并等技术，使得推理和挖掘的速度和准确率大幅度提升。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=1,024G）的存储量，缩减至 2 小时的计算量和 30G 的存储量。

C.大数据搜索与分析技术：可在启信宝平台海量商业大数据中进行毫秒级的实时搜索与分析，为用户提供企业信息画像、关联关系分析、风险查询、营销决策分析等大数据搜索与分析服务，是国内领先、实时、动态的商业大数据搜索与分析平台。

商业大数据技术是公司凭借多年技术研发活动中积累的自主研发的算法模型、C 端产品和 B 端服务开发中沉淀的专有技术，而非直接运用行业通用技术。

公司商业大数据的技术壁垒主要是在外部公有数据与 B 端客户私有数据的融合度、超大知识图谱构建的深度与速度、知识图谱与客户所在行业的结合度、知识图谱自动动态更新的及时性和准确度等方面具有竞争优势。另外，公司是中国电子技术标准化研究院牵头成立的标准规范知识图谱联盟理事单位，参与撰写的《标准数字化知识图谱白皮书》于 2021 年 10 月正式发布。

公司的科技成果转化能力强，将 AI 及大数据技术在诸多领域进行商业化落地。凭借领先的自主研发技术、成熟的产品落地能力、优质的用户体验及服务质

量，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¹，B 端服务覆盖了近 30 个行业的企业客户。在 C 端产品方面，截至 2022 年 12 月底，公司扫描全能王、名片全能王、启信宝 3 款 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 8 亿²。截至 2022 年 12 月 31 日，扫描全能王曾在 App Store 上 105 个国家和地区（含中国）的效率类免费应用下载量排行榜位列第一，名片全能王曾在 App Store 上 43 个国家和地区的商务类免费应用下载量排行榜位列第一。2021 年初，公司在 App Annie “2021 Top publisher Awards”中与 Google、Microsoft、Zoom Video Communications 及 Adobe 被评选为全球前五大商业与效率应用企业（Top 5 publishers of Business & Productivity Apps）³。在 B 端业务方面，公司智能文字识别与商业大数据服务已覆盖了银行、证券、保险、政府、物流、制造、地产、零售等行业的众多头部客户，《财富》杂志 2022 年发布的世界 500 强公司名单中，公司客户已覆盖超过 125 家。

公司在研发投入、研发人员数量、收入规模及增长速度方面大幅超过科创属性相关的指标要求：

| 科创属性评价标准一 | 是否符合 | 指标情况 |
|---|------|---|
| 最近三年累计研发投入占最近三年累计营业收入比例 $\geq 5\%$ ，或最近三年累计研发投入金额 $\geq 6,000$ 万元 | 是 | 公司最近三年累计研发投入占最近三年累计营业收入比例为 28.88%，最近三年累计研发投入金额为 68,516.20 万元。 |
| 研发人员占当年员工总数 $\geq 10\%$ | 是 | 截至 2022 年 12 月 31 日，公司全职员工共 912 人，其中研发人员共 499 人，占总人数比重约为 55%。 |
| 应用于公司主营业务的发明专利 ≥ 5 项 | 是 | 截至 2022 年 12 月 31 日，公司发明专利 117 项中应用于主营业务的为 89 项；若将基于同一优先权申请在不同国家/地区获批、内容基本相同或相似的发明专利的视为同族专利，且同族专利仅记为 1 项，则中国境内外专利合计为 79 项，其中应用于主营业务的为 55 项。 |
| 最近三年营业收入复合增长率 $\geq 20\%$ ，或最近一年营业收入金额 ≥ 3 亿 | 是 | 公司最近三年营业收入复合增长率为 30.74%，最近一年营业收入金额为 98,846.18 万元。 |

¹ 注：2022 年 12 月的各渠道月活合计超过 1.3 亿（各渠道月活为直接相加、无法去重，即在多个渠道活跃的同一用户可能被重复计算），其中 APP 渠道的月活合计超过 1.1 亿，小程序、Web 端等其他渠道月活合计约 0.2 亿

² 注：截至 2022 年 12 月末，公司扫描全能王、名片全能王、启信宝 3 款 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 8 亿，其中 2020 年初-2022 年 12 月末的累计新增首次下载量约 3.9 亿，由于部分安卓渠道历史数据不能区分首次与更新下载，所以未统计所有安卓渠道的首次下载量，下同。

³ 注：该排名由 App Annie 根据各 APP 在全球的 iOS 及 Google Play 应用市场 2020 年整体下载量统计得出，下同。

综上所述，经核查，保荐机构认为：发行人核心技术具有行业先进性，发行人具有突出的科技创新能力及科技成果转化能力，符合科创板支持方向、科技创新行业领域和相关指标等科创板定位要求，属于科创板支持的“硬科技”。

（2）智能文字识别技术处于快速发展阶段，远未达到技术成熟阶段

智能文字识别技术融合了智能图像处理、复杂场景文字识别、NLP 等 AI 技术，相比传统简单文字识别，具备更多认知与理解能力。传统简单文字识别已基本达到成熟的阶段，但智能文字识别技术尚处于快速发展阶段，超过 99% 的识别率主要在相对简单标准的印刷文档场景下，在一些实际的复杂应用场景中，比如拍摄质量差、光照不均、纸张褶皱、阴影、版面排版复杂、书写潦草、字体形态各异（含印章/公式/二维码等）等复杂场景下，目前识别率只达到 90% 以上，识别率距离用户满意的水平仍有较大距离，智能文字识别领域仍需攻克诸多技术难题，远未达到技术成熟阶段。

客户的预期是接近 100% 的准确率，几个百分点的识别率的差别对于客户复核识别结果的工作量、产品的体验感的影响巨大。公司在智能文字识别领域深耕超过 15 年，目标是解决 C 端用户及 B 端客户在实际使用中可能遇到的各类痛点问题，使得其在各类复杂场景下的识别的准确率与可信度提升至满意的水平。

同时，公司的智能文字识别技术不仅包括复杂场景文字识别技术，还包括智能图像处理技术与 NLP 技术，相关技术仍处于快速发展阶段，有很多技术创新值得深入探索：

首先，在很多客户的真实应用场景（例如合同、财报自动审核等），不仅需要文字的精准识别，还需要利用 NLP 技术对文档整体进行校对、审核或分析，如何根据客户的需求去分析处理文档、提供更高的技术增值服务还面临大量的技术难点。

另外，公司利用核心技术中的智能图像处理技术，探索了多个 AI 图像安全领域的创新技术方向，相关方向有较大的技术发展与应用空间：

①AI 图像篡改检测技术，对于肉眼无法辨别的伪造痕迹，通过 AI 精准定位篡改痕迹，解决不法分子利用技术手段篡改文本图像的问题，比如从原图中“抠

下”关键要素后移动“粘贴”至另一处的“复制移动”图片篡改手段；

②生成式图像鉴别技术，对于肉眼无法辨别的 AIGC 生成式图片，基于空域与频域关系建模，利用多维度特征来分辨真实图片和生成式图片的细微差异；

③OCR 对抗攻击技术，可对场景文本或者文档内文本进行扰动，为个人或企业的重要文件“上锁”，防止第三方通过文字识别技术读取并保存图像中所有的文字内容，降低个人信息数据泄露的风险，达到保护隐私信息的目的。

综上分析，经核查，保荐机构认为：智能文字识别技术仍处于快速发展阶段，远未达到技术成熟阶段。

(3) 复杂场景的智能文字识别可以解决 C 端和 B 端客户的诸多痛点，具有较大的潜在需求空间

公司核心技术之一“基于深度学习的复杂场景下的文字识别”技术中“复杂场景”指客户在实际的办公/学术/家庭/商务等场景下遇到的任意场景，实际场景中可能碰到的各种各样的痛点问题，公司产品所覆盖的全球不同语言、不同场景的数亿用户会把实际遇到各类痛点放大到极大的数量级。

复杂场景下智能文字识别技术可以为 C 端用户和 B 端客户提供较高的技术增值服务，做到实质性降本增效，因此具有巨大市场需求，具体分析如下：

1) C 端场景

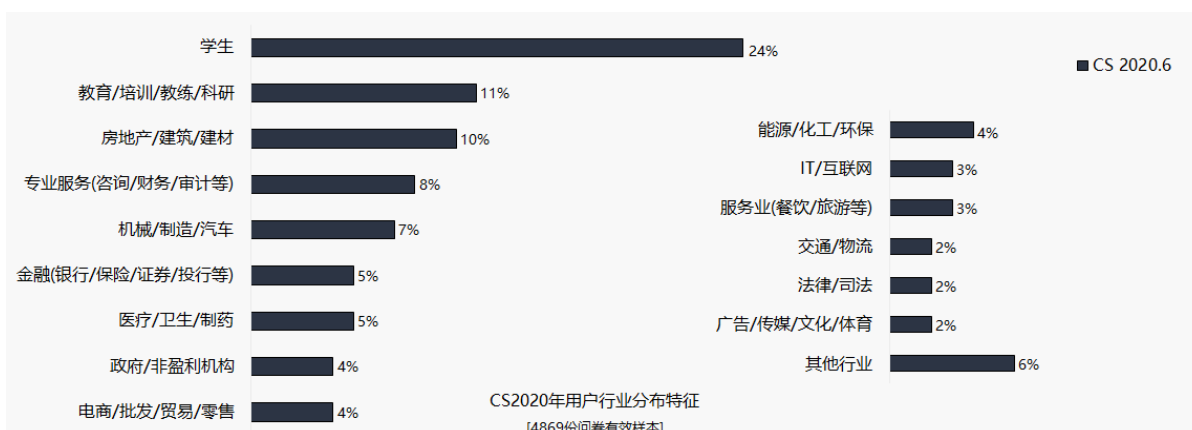
以扫描全能王为例：

| 扫描全能王的代表性功能 | 技术亮点 |
|---------------------------------|---|
| 弯曲及角度矫正 边缘检测 阴影处理 图像增强 | 采用 AI 领域的智能图像处理技术和算法，可对复杂场景下拍摄的倾斜图片进行自动校正，自动检测拍摄图像中的文档四周边缘位置，对边缘进行自动切边，去除多余背景图像，显著提高图像质量，同时去除复杂阴影，大幅提高复杂场景下文字识别的准确率 |
| 图片转 word/ppt/excel | 可将任意文档的照片转换为可编辑的 word/ppt/excel 等多种文件格式，同时做到版面还原，包括还原图片中的表格、数学公式、印章、手写签名等 |
| 多语言处理 | 可以自动辨别语言类型（支持 41 种语言），并支持多角度、多方向的多语言识别 |
| 拍图识别 | 可以实现在任意复杂场景下（比如抖动模糊、反光、远距离低像素、褶皱、低光照噪点等）手机摄像头拍照后进行文字的提取识别 |

| 扫描全能王的代表性功能 | 技术亮点 |
|-------------|------------------------------|
| 表格识别 | 可做到表格线检测、无线表格线检测及表格内段落信息换行提取 |

扫描全能王前述功能的技术基础是公司集合深度学习算法、智能图像处理、自然语言处理（NLP）等多项 AI 技术、研发出可以落地至手机端的轻量级 AI 模型。经公司内部后台统计，以 2023 年 6 月 30 日为例，前述功能单日调用次数的合计数超过 1 亿次，证明用户对公司的核心技术产品具有很强的需求。

使用扫描全能王 APP 的典型用户画像包括老师、学生、销售人员、采购人员、行政人员、财务人员、工程师、医生等多种主流职业，主要应用场景涉及教育、建筑、专业服务、机械、金融、医疗等领域（如下图为公司 2020 年 6 月用户调研所回收的 4,869 份有效问卷的结果分析）。



下表以四类行业的代表性用户来说明 C 端用户使用智能文字识别的应用场景的丰富度：

| 序号 | 用户职业 | 高频使用的文档 | 使用场景 |
|----|------|-------------------------------------|--|
| 1 | 老师 | 论文、报告 | 可从手机拍摄的教学 PPT/论文/报告等文档的图片进行文字提取，作为教学或者个人学习素材 |
| 2 | 工程师 | 开工报告、施工合同、施工材料检验报告、施工蓝图、现场施工图片、采购单等 | 及时备份纸质原件并分享给其他部门，留存采购发票/清单等定期报销，工地现场安装的关键步骤例如管道安装图、压力试验的压力表等，及时扫描处理或进行文字提取 |
| 3 | 银行经理 | 客户保单、证件、贷款合同 | 扫描及识别客户的信息以建立个人客户档案，并及时上传公司系统，部分文件外发时加防盗用水印，注明证件用途 |

| 序号 | 用户职业 | 高频使用的文档 | 使用场景 |
|----|------|----------------------|---|
| 4 | 医生 | 论文、教材、病例诊断报告、学术培训材料等 | 扫描不同病例的诊断结果、修改编辑文档并整理留档，扫描论文、教材等作为教学或学术交流素材 |

综上分析，复杂场景的智能文字识别可以解决 C 端用户的诸多痛点，具有丰富的使用场景和多元化的用户群体，具有较大的潜在需求空间。

2) B 端场景

在 B 端，公司利用智能文字识别服务除具备 C 端产品的功能外，同时结合客户的具体应用场景开发了财报 AI 机器人、票据 AI 机器人、合同 AI 机器人等与实际业务场景紧密结合的 AI 场景化解决方案。

| 序号 | 场景化解决方案 | 亮点 |
|----|-----------|--|
| 1 | 财报 AI 机器人 | 面向银行信贷尽调、融资租赁业务前序审查、供应商准入尽调、供应商风险监控、投资机构投后管理等场景，提供财报智能分析和目标公司尽职调查能力 |
| 2 | 票据 AI 机器人 | 通过不同种类票据（国税税种、地方票种、出租车发票、银行流水、医疗费收据、公路客运发票等），进行智能识别、分类、录入，在企业财务管理、费控管理、对公支付、进项税票、税务分析、财务管理等业务场景中完成智能化的流程改造 |
| 3 | 合同 AI 机器人 | 结构化提取合作主体方关键信息，智能判断合同内是否有公司盖章、法人或代表人签字/手印等，协助工作人员完成自动填表、一致性检查等工作 |

发行人智能文字识别 B 端大大减少企业在实际工作中大量人工录入/核对的重复劳动性工作，相比人工操作的准确度显著提升，降本增效的同时提升了客户满意度，因此具有丰富的应用场景：

| 下游行业 | 行业痛点 | 智能文字识别技术赋能 |
|-----------|--|--|
| 银行、证券等金融业 | <ul style="list-style-type: none"> 大量业务流程仍依赖员工或者用户的人工操作，效率较低，用户体验亟待优化 存在海量未电子化的文档数据，未能打通各业务环节文档数据，实现数据价值 | <ul style="list-style-type: none"> 多终端文档、图像信息的自动采集，批量扫描识别多种非标准化文档 各类表单、文档及工单等文档的智能录入、分类、核验、解析，简化业务流程，降本增效 实现自动核保、远程开户等商业化应用，数字化用户体验，提高用户满意度 |

| 下游行业 | 行业痛点 | 智能文字识别技术赋能 |
|------|---|---|
| 制造业 | <ul style="list-style-type: none"> • 信息化程度较低,大量文档数据未有效管理与利用 • 上下游涉及企业较多,缺乏有效协同运作体系 | <ul style="list-style-type: none"> • 自动识别、录入各类合同、票据等文件 • 自动识别、录入各类识别编码、日期等信息,赋能设备、生产原材料管理等 • 数字化企业运营,上下游协同工作、流程自动化 |
| 物流 | <ul style="list-style-type: none"> • 竞争加剧,利润承压,急需降低成本,提升效率 • 运单版式多样、经手频次高、人工效率低且易出错 | <ul style="list-style-type: none"> • 自动识别不同版式的快递单据,并将运单信息高效录入系统,完成核验 • 自动识别快递信息,进行智能分拣,优化流程 |
| 政府 | <ul style="list-style-type: none"> • 业务处理流程繁杂、所需时间较长,用户体验亟待优化 • 机构间通过纸质文件办公,协同效率较低 | <ul style="list-style-type: none"> • 自动识别、录入各类证件、申请表格等,并与政府机构的核查系统、申请系统等连接,赋能政务自动化转型 • 无纸化办公,提升各机构间协同效率 • 数字化用户体验,提升办事效率,提高用户满意度 |
| 医疗 | <ul style="list-style-type: none"> • 日常运营中产生病历、病程记录等大量纸质表格,信息化程度较低,大量医疗数据未有效利用与价值挖掘 • 浪费大量人力物力进行病历记录,医疗效率亟待提升 | <ul style="list-style-type: none"> • 自动识别、录入病历等纸质文档,对病历内容进行检查、校对,建立电子医疗档案 • 自动识别、分析化验单和收费单等单据,提高诊疗及收费效率,实现数字化运营,利用智能化手段提升医疗效率 |
| 地产 | <ul style="list-style-type: none"> • 业务处理流程繁杂,用户体验较差 • 上下游涉及企业较多,缺乏有效协同运作体系 | <ul style="list-style-type: none"> • 自动识别、录入用户的各类证件,包括房产证、不动产权证等,并自动将信息录入至对应系统,完成核验 • 数字化用户服务,提升办事效率,提高用户满意度 • 数字化企业运营,上下游协同工作,实现流程自动化 |
| 教育 | <ul style="list-style-type: none"> • 传统的教学方式耗费大量的人力和物力,老师、家长及学生亟需便捷的家校沟通平台和高效的智能教辅工具 | <ul style="list-style-type: none"> • 自动识别、录入准考证等信息,电子化管理学生信息 • AI 智能阅卷,为老师/家长减负,提升批改效率 • 拍照搜题,准确识别题目及搜索解析思路 |

推动 B 端市场未来继续快速增长的主要驱动因素是：①由于文档广泛地存在于各行业业务场景，从非标准化文档中提取结构化信息、进行数字化转型在多种行业中均存在需求，而近些年来企业及政府的数字化转型需求快速增强、信息化意识不断提高，信息化与 IT 投入近些年不断提升，且相比发达国家仍有较大上升空间；②我国政府对企业及政府的数字化转型也推出了多项利好政策。例如，2021 年 3 月，“加快数字化发展、建设数字中国”成为我国“十四五纲要”中的重要章节，其提出“加强关键数字技术创新应用、加快推动数字产业化、推进产业数字化转型”的要求，人工智能领域的“智能识别系统”是数字经济重点产业，并提出了“加快建设数字经济、数字社会、数字政府，以数字化转型整体驱

动生产方式、生活方式和治理方式变革”的目标；③应用场景及行业的拓宽驱动行业快速发展，智能文字识别技术可帮助企业与政府实现文档存量数字化、增量电子化，帮助企业与政府降本增效，应用行业从银行、证券、保险等领域快速拓宽至政务、制造、零售等多元化领域。

根据灼识咨询，从中长期可触达市场（Total addressable market）的角度，随着企业对于自动化、智能化转型升级需求以及场景化方案需求的不断提升，中国智能文字识别市场将迎来巨大发展空间，尚未被挖掘的市场规模超过人民币1,000亿元。

综上分析，经核查，保荐机构认为：复杂场景的智能文字识别可以解决 C 端和 B 端客户的诸多痛点，具有较大的潜在需求空间。

（4）大数据相关的已授权发明专利数量较少的原因具有合理性

针对大数据相关的专利数量较少问题，保荐机构核查情况如下：在商业大数据领域，截至 2022 年 12 月 31 日，公司拥有“一种企业集团关系获取方法及系统”发明专利，同时已有 14 项发明专利申请已提交，正在审查中：

| 序号 | 所属国家 | 申请号 | 阶段 | 发明名称 | 申请日 | 对应的核心技术 |
|----|------|----------------|-----|-------------------------|------------|-----------------|
| 1 | 中国 | 201910878683.3 | 审查中 | 一种企业关联关系构建方法及系统 | 2019/9/18 | 超大规模知识图谱构建技术 |
| 2 | 中国 | 202110012787.3 | 审查中 | 一种自动生成查询语句的数据提取方法及装置 | 2021/01/06 | |
| 3 | 中国 | 201910878762.4 | 审查中 | 一种企业受益人运算系统及方法 | 2019/9/18 | 超大规模知识图谱推理与挖掘技术 |
| 4 | 中国 | 201910878748.4 | 审查中 | 一种企业实际控制人运算系统及方法 | 2019/9/18 | |
| 4 | 中国 | 201910878698.X | 审查中 | 一种企业疑似关联关系判定方法及系统 | 2019/9/18 | |
| 5 | 中国 | 201910878735.7 | 审查中 | 一种企业一致行动人运算系统及方法 | 2019/9/18 | |
| 7 | 中国 | 202110532500.X | 审查中 | 基于ANTLR的数据字段血缘分析方法及装置 | 2021/5/17 | |
| 8 | 中国 | 202110532643.0 | 审查中 | 一种合同履行表现量化及短期违约预测的方法及装置 | 2021/5/17 | |

| 序号 | 所属国家 | 申请号 | 阶段 | 发明名称 | 申请日 | 对应的核心技术 |
|----|------|----------------|-----|---------------------------|------------|---------------|
| 9 | 中国 | 202110567097.4 | 审查中 | 一种企业信用评估及信用传导方法及装置 | 2021/5/24 | |
| 10 | 中国 | 202111283821.7 | 审查中 | 一种基于裁判文书的企业司法诉讼关系的构建方法及系统 | 2021/11/1 | |
| 11 | 中国 | 202111440689.6 | 审查中 | 一种知识图谱的扩展更新方法及装置 | 2021/11/30 | |
| 12 | 中国 | 202111444974.5 | 审查中 | 一种企业疑似实际控制人的深度挖掘方法及装置 | 2021/11/30 | |
| 13 | 中国 | 202111601481.8 | 审查中 | 一种企业多层股权的图计算方法及装置 | 2021/12/24 | |
| 14 | 中国 | 202110616330.3 | 审查中 | 一种基于企业关联关系的人名消歧方法及系统 | 2021/06/02 | 基于数据挖掘的人名消歧技术 |

目前公司在商业大数据领域已授权的发明专利较少的原因是，专利的审批时间较长，多项专利申请处于审核过程中，另外专利申请也和公司研发团队的工作重点方向有关。在大数据领域，2015-2017年为公司重点开拓商业大数据业务阶段，未将工作重心在专利申请方面；2018年开始公司基于前期研发经验，从2018年初至2022年12月31日，陆续提交了多项大数据领域的发明专利申请，逐步加强了专利申请力度。

经核查，保荐机构认为：由于发明专利审查的审核周期不确定、公司大数据业务专利申请节奏的原因，公司大数据相关的已授权发明专利数量较少，该等原因具备合理性，公司已有多项大数据相关的发明专利在申请中。

(5) 近三年每年公司已授权发明专利数量变动的原因具有合理性，公司应用于主营业务的已授权发明专利数量大幅超过科创属性要求

针对近三年每年所获得的已授权专利数量情况，保荐机构核查如下：按照授权时间统计，截至2022年12月31日，公司117项已授权发明专利中，2018年之前授权的发明专利共99项，2018年初至2022年12月31日授权的专利共18项。报告期内新授权的发明专利数量增幅低于报告期之前积累的发明专利数量，主要是由于发明专利的审批时间较长，多项发明专利申请处于审核过程中，另外专利申请也和公司研发团队的工作重点方向、专利申请策略有关，对于部分领域

公司认为公布了专利申请文件之后可能过早对外公开了研发核心机密，因此适当放缓了专利申请节奏，公司重点对防守型专利进行优先申请。

截至 2022 年 12 月 31 日，发行人及其子公司共有 34 项与核心技术相关的发明专利申请已提交，主要集中于知识图谱、智能文字识别、智能图像处理、自然语言处理等领域；截至本次审核中心意见落实函回复签署之日，前述 34 项发明专利均在正常申请过程中。

专利申请是保护公司创新成果的手段，专利数量并非反映公司技术先进性的主要因素，不代表公司的所有研发技术成果。公司 C 端及 B 端业务优质的客户体验、不断提升的品牌力以及快速增长的业务规模是公司依靠核心技术驱动、持续进行研发投入带来的优秀成果，是公司技术先进性的重要体现。

以扫描全能王为例，如下表所示，报告期内公司持续进行智能文字识别、图像处理、深度学习相关的底层技术和应用层面的研发项目，主要目的是为了实现核心技术的突破、产品功能的创新、用户体验的升级。相关研发投入也得到了优秀的落地成果，扫描全能王推出“图片转 word/ ppt/ excel”、“拍图识字”、“去摩尔纹”等新功能，实现多项技术创新与升级，报告期内的收入、月活用户数量、付费用户数量、新增付费转化率等关键业务指标保持快速增长趋势；同时公司有选择性的将研发项目的部分技术成果进行发明专利申请，截至 2022 年 12 月 31 日，发行人及其子公司共有 20 项智能文字识别领域的发明专利申请已提交。

| 报告期内扫描全能王相关的研发项目 | 研发的主要内容 |
|--------------------------|---|
| 基于迁移学习和知识图谱技术的智能文本图像理解系统 | 创新研发基于迁移学习的语义理解技术，针对新场景下的样本比较少的文本图像，通过迁移学习模型训练，将已有的通用模型转化为新场景下的模型；创新研发基于知识图谱的语义扩充技术方案。 |
| 基于智能文字识别技术的 AI 云平台 | 结合深度学习技术，对以下几项关键技术进行研发：1) 用于文档图像识别的深度卷积神经网络技术；2) 基于端到端的全连接卷积回归神经网络的文本行识别技术等；3) 基于深度学习的文档全要素检测技术。全方位的提升文字检测识别、文档要素检测、表格检测与识别等方面的性能，并有效的降低文档理解的错误率。 |
| 基于深度学习的电子文档处理系统 | 结合深度学习技术，攻克云端电子文档识别服务并发出现性能瓶颈、PDF 签名功能实时渲染性能开销大等移动端落地的技术难点，研发 PDF 编辑、PDF 转 Word/Excel/PPT、图像转 Word/Excel 等功能。 |

| 报告期内扫描全能王相关的研发项目 | 研发的主要内容 |
|------------------------------|---|
| 基于深度学习的智能扫描系统增强与优化 | 继续优化扫描全能王的底层算法和图像处理能力，进一步扩展扫描全能王的服务能力和服务场景，结合不同场景下用户属性和底层的深度学习算法，为用户提供智能的服务。 |
| 基于深度学习的版面还原技术在移动设备的应用 | 基于深度学习对移动设备扫描的图像内容进行识别，去除图像的背景干扰，保留文字内容和图像内容，并智能还原版式信息，包括还原图片中的表格、数学公式、印章、手写签名等，让用户真正体验到所见即所识别。 |
| 复杂环境下的图像 AI 修复项目 | 目前行业内图像人工修复耗时耗力，通过研发图像修复的人工智能算法，利用智能图像处理技术将模糊破损的照片进行修复、对低分辨率的视频的进行修复，可以大幅提高效率，降低修复成本。 |
| 扫描全能王的 PDF 文档服务扩展矩阵 | 开发扫描全能王的 PDF 文档服务扩展矩阵，为扫描全能王用户在扫描创建文档后的提供签名、水印、加密处理等编辑文档等增值服务，打造优质的 PDF 格式转换效果。 |
| 基于智能文字识别技术的数字化转型平台 | 对各种文档图像的文字检测和识别等关键技术进行升级，重点提高复杂场景下的字符识别率，实现抖动模糊、反光、远距离低像素、褶皱、低光照噪点等复杂场景下手机摄像头拍图后文字自动识别。 |
| 全球企业数字化转型升级的新一代人工智能文档数字化赋能平台 | 针对海外多国语言文本图像，对卷积神经网络技术、表示学习技术、序列学习技术、语言模型技术、命名实体识别等技术进行研发，解决智能文字识别技术在海外多语言情况下落地的难点。 |

部分媒体报道通过横向比较公司与腾讯、百度、金山办公等企业的专利数量来质疑公司的科创成色不足，该等观点不具有合理性，分析如下：

1) 公司应用于公司主营业务的发明专利数量大幅超过科创属性的条件，截至 2022 年 12 月 31 日，公司共有 89 项应用于主营业务的发明专利；若采取同族专利去重的口径⁴，则公司共有 55 项应用于主营业务的发明专利，大幅超过科创属性要求之一“应用于公司主营业务的发明专利 ≥ 5 项”；

2) 如前分析，专利申请是保护公司创新成果的手段，专利数量并非反映公司技术先进性的主要因素。公司主要采取技术秘密保护与专利保护相结合的策略，对于部分领域公司认为公布了专利申请文件之后可能过早对外公开了研发核心机密，另外在业务开展过程中，公司的 C 端用户和 B 端客户并不以发行人所拥

⁴ 将基于同一优先权申请在不同国家/地区获批、内容基本相同或相似的发明专利的视为同族专利，且同族专利仅记为 1 项

有的专利数量作为优先考察因素，因此公司适当放缓了专利申请节奏、采取严格保密措施对自主研发的技术进行保护；

3) 公司的主营业务、业务定位、重点研发方向与腾讯、百度、金山办公等公司存在一定差异：

①腾讯、百度在科技领域的业务范围广泛且多元化，包含云服务器、云存储、智能翻译、人脸识别、语音识别、视频分发加速等领域，文字识别为其中一种领域、但不是其核心业务；

②金山办公与公司的主营业务和业务定位不同，根据金山办公招股说明书，其业务定位为办公软件和服务提供商，主要从事 WPS Office 办公软件产品及服务的设计研发及销售推广。WPS Office 办公软件包含多元化的功能，其核心功能为文档编辑、文档阅读、线上办公协作等，文档扫描与文字识别功能为其中一种，但不是核心功能；

③公司专注智能文字识别领域且持续研发投入超过 15 年，识别精准率行业领先，尽管近几年面临腾讯、百度、金山办公等企业的竞争，公司智能文字识别业务产生的收入自 2019 年的 21,138.16 万元快速增长至 2022 年的 72,489.24 万元，年复合增长率为 51%，快速增长的核心技术业务收入规模可以证明公司优秀的科技成果转化能力，公司技术先进性、科技成果转化能力的分析详见本次审核中心意见落实函回复报告之“二（三）2、（1）、3）公司核心技术处于行业领先水平”。

经核查，保荐机构认为：由于发明专利审查的审核周期不确定、公司的知识产权保护策略等原因，公司近三年每年新增的已授权发明专利数量较少，该等原因具备合理性。公司应用于主营业务的已授权发明专利数量满足科创属性要求，具有突出的技术先进性和科技成果转化能力。

（6）2020 年研发费用率相比 2019 年出现大幅下降主要是由于收入增速较高以及股份支付各期金额差异的影响

单位：万元

| 项目 | 2022 年 | 2021 年度 | 2020 年度 | 2019 年度 |
|---------------|-----------|-----------|-----------|-----------|
| 研发费用 | 27,990.16 | 23,829.24 | 16,696.80 | 21,058.73 |
| 计入研发费用的股份支付费用 | 0 | 0 | 0 | 6,377.18 |

| 项目 | 2022 年 | 2021 年度 | 2020 年度 | 2019 年度 |
|-----------------------|-----------|-----------|-----------|-----------|
| 研发费用（扣除股份支付费用） | 27,990.16 | 23,829.24 | 16,696.80 | 14,681.55 |
| 营业收入 | 98,846.18 | 80,578.16 | 57,824.64 | 34,038.94 |
| 研发费用占营业收入比例 | 28.32% | 29.57% | 28.87% | 61.87% |
| 研发费用（扣除股份支付费用）占营业收入比例 | 28.32% | 29.57% | 28.87% | 43.13% |

如上表所示，尽管 2020 年公司研发费用占收入比重出现大幅下降，但研发费用（扣除股份支付费用）绝对值从 2019 年的 14,681.55 万元提升 14% 至 2020 年的 16,696.80 万元，依然保持稳定增长。2020 年公司研发费用占收入比重出现下滑的主要原因是：1) 计入研发费用的股份支付费用从 2019 年的 6,377.18 万元降低至 2020 年的 0 万元；2) 2019 年公司的收入体量较小，2020 年收入增长速度较高，2020 年相对 2019 年的收入增速达 70%，随着营业收入的快速增长，规模效应逐步体现，因此公司的研发费用率有所下降。

2020 年-2022 年，公司的研发费用占收入比重基本稳定在 28% 左右，研发费用绝对值保持 29% 的年复合增长率。扣除股份支付影响后，发行人与可比公司的研发费用率相比基本合理：

| 可比公司 | | 2022 年度 | 2021 年度 | 2020 年度 | 2019 年度 |
|------|-----------|---------------|---------------|---------------|---------------|
| 公司简称 | 证券代码 | | | | |
| 金山办公 | 688111.SH | 33.41% | 32.98% | 31.44% | 37.91% |
| 福昕软件 | 688095.SH | 30.97% | 24.83% | 16.08% | 15.36% |
| 平均值 | | 32.19% | 28.90% | 23.76% | 26.64% |
| 发行人 | | 28.32% | 29.57% | 28.87% | 43.13% |

资料来源：可比公司定期报告、招股说明书

综上所述，保荐机构、申报会计师认为：发行人 2020 年研发费用率出现大幅下降主要是由于收入增速较高以及股份支付各期金额差异的原因，发行人与可比公司的研发费用率相比基本合理。

(7) 服务器及带宽费用划入研发费用具有合理性

公司将研发部门进行 AI 底层算法与应用模型的训练活动、业务的开发及测试等消耗服务器及带宽的存储和传输等费用计入研发费用。保荐机构、申报会计

师经核查后认为：公司将与研发活动相关的服务器及带宽费用划入研发费用符合费用实质、亦符合相关会计准则，公司已建立较为完善的内控，规范各部门对服务器的领用及消耗，不存在虚增研发费用的情形。

（四）数据安全与个人隐私保护问题

1、媒体质疑情况

部分媒体报道质疑公司曾经存在的向个人供应商采购数据的情形；质疑公司存在销售数据的情形；部分媒体报道质疑公司展业中牵涉到的数据信息安全与合规情况，以及公司涉嫌侵犯用户的数据隐私并违规使用用户信息。

2、核查情况

（1）发行人报告期内不存在向个人供应商采购数据的情形，发行人历史上向个人供应商采购数据的数据来源合法合规

1) 发行人历史上向个人供应商采购数据的事宜履行了内部审批程序，具备业务合理性

2018 年及以前，公司关于数据的自动化采集逻辑是动态采集最新的数据，获取新的数据之后将覆盖原有旧的数据，变更历史未充分留存。但在业务开展过程中，用户仍然存在查询企业变更信息或个体工商户信息的需求。为完善相关数据维度，增强公司整体大数据业务的竞争力，公司选择了具备较强技术背景、且曾经通过自动化采集等手段获取了国家工商公示系统对外公示的工商信息的个人供应商，并与其就数据采集达成了合作。

向该个人供应商进行采购前，公司履行了内部的审批流程，包括确认需求并进行充分评估、内部履行充分的采购流程管理、与个人供应商进行对接合作。公司采购行为按照公司内控制度和流程执行，符合公司关于供应商管理的内控制度和流程的要求。

2) 该等采购的数据来源合法合规，该个人供应商也无需任何资质要求，使用或出售该等数据及该等数据形成的产品合法合规

根据《数据安全法》，国家保护个人与数据相关的权益，鼓励个人合法有效

利用数据；根据《上海市数据条例》，自然人可以合法取得数据，并依法开展数据交易活动。根据公司对上海市经信委和上海数据交易所的访谈，市经信委认为：自然人、法人和非法人组织均可从事数据类服务、数据交易，不存在法律法规禁止相关主体从事数据交易活动的情况。对于自然人而言，在不涉及提供征信服务的情况下，通过公开渠道获取合法的数据后对外销售，无需取得采集、整理、加工或出售数据的资质，只要其从事数据业务的行为符合《数据安全法》《网络安全法》等法律法规的要求即可。个人供应商销售给公司的数据均为公开的基本工商信息，主要包括名称、注册资本、成立时间、工商变更及股东信息等，以及该等信息历史变化情况。

综上，保荐机构、发行人律师认为：公司报告期内不存在向个人供应商采购数据的情形；公司历史上存在向个人供应商采购数据的情形，公司已就该等事宜履行了内部审批程序，具备业务合理性。该个人供应商系技术领域从业人员，与发行人及其客户、其他供应商不存在关联关系或利益安排。根据《数据安全法》、《上海市数据条例》，该等采购的数据来源合法合规，公司不存在违反相关法律法规的情形，该个人供应商不需要获得出售数据的相关资质，也无需任何资质要求；公司使用或出售该等数据及该等数据形成的产品合法合规。

(2)发行人在业务开展过程中不存在将开展 C 端或 B 端业务时获取的用户数据进行销售的情形

发行人存在销售数据的情形，但是发行人销售的是对原始数据加工、处理后的标准化产品，并不是其开展 C 端或 B 业务时获取的用户数据。

公司在开展智能文字识别 C 端业务的过程中，对于用户使用扫描全能王 APP 或名片全能王 APP 扫描、识别的文档、图片或名片等资料本身的内容不做任何的商业化使用，扫描全能王和名片全能王作为高效的办公工具，公司仅靠为用户提供扫描、识别、上传、分享等产品功能，获取用户使用办公工具而支付的会员费。智能文字识别 B 端所获取、使用的各类信息数据为根据协议具体内容、由 B 端企业客户自主提供企业信息，以便公司提供文档扫描与文字识别服务。公司智能文字识别业务为 C 端用户和 B 端客户提供的服务主要为文档扫描与文字、图片识别服务，也不涉及销售或交换信息数据。

公司开展 B 端商业大数据业务、C 端启信宝 APP 的主要数据来源是公司数据中台汇聚的企业数据。公司开展 B 端商业大数据业务不会利用 C 端的用户数据（包括但不限于 C 端用户注册数据、C 端用户上传的文档、名片）。公司对前述数据进行数据挖掘、数据清洗、数据结构化处理等加工后，分别在 C 端形成按期收费的 VIP 会员、付费报告产品，以及在 B 端形成给企业客户提供基础数据服务、标准化服务、场景化解决方案。

就历史上向个人供应商采购数据、公司销售数据相关事宜，发行人已在第一轮问询函回复之“4. 关于数据合规性”之“一、请发行人说明：（1）发行人各项业务及研发分别获取、存储、使用哪些数据，对应的数据来源、数据权属，是否存在销售数据的情形”、“（二）发行人向个人供应商采购数据的主要内容、比例及原因，价格公允性，该等个人是否与发行人及其客户、其他供应商存在关联关系或利益安排”进行了详细说明。

综上，保荐机构、发行人律师认为：公司不涉及将任何非公开数据或个人隐私数据用于销售，也不存在将从供应商采购/互换的以及自动化访问获取、使用的原始数据直接销售给第三方的情形。

（3）数据合规及个人信息保护方面：发行人已建立覆盖数据全生命周期的数据合规及个人信息保护措施并有效执行

1) 内控制度方面：发行人已建立包括数据采集、数据使用、数据访问权限控制、数据导出和数据删除等覆盖整个数据生命周期管理的相关制度，并设计了相应的管理流程，实现全数据生命周期的管理。

2) 公司内部组织架构方面：发行人建立了数据合规与信息安全的组织架构，明确了安全与合规管理委员会在业务合规、网络安全、数据安全和用户个人信息保护等方面的责任以及安全与合规部和各事业部、项目组的职责，并规定网络安全、数据安全和个人信息保护等相关责任机构的汇报层级要求。公司按照“谁主管、谁负责”的原则，根据国家的法律法规，明确数据获取、使用、处理等环节的第一责任人，建立了完整的、多层级、全覆盖的数据合规与信息安全管理组织架构体系。

3) 个人信息保护方面：公司制定了《用户个人信息安全管理规定》，其中规定安全与合规部须定期（至少每年一次）对属于个人信息敏感岗位的相关人员开展个人信息安全的专业化培训和考核，其内容包含个人信息安全职责，以确保相关人员熟练掌握个人信息保护政策和相关规程。被培训人员在培训完成后均须通过培训考核。公司还定期开展数据保护及隐私合规培训。

4) C 端及 B 端业务开展方面：公司在严格遵守隐私协议的情况下获取 App 用户相关信息，用户自主输入数据均在 C 端 APP 的相关隐私政策中明确列示，在用户勾选并授权同意隐私政策后方可采集。C 端均在严格遵守合法、正当、必要和诚信原则的前提下，在获得用户授权的情况下，获取必要的用户信息，不涉及非法采集个人隐私信息的情形；公司与 B 端企业客户在协议中已明确约定 B 端企业客户提供的企业数据、业务资料，以及公司提供的技术内容、形式及要求。公司严格按照协议内容为客户提供相关服务，不存在超出与 B 端企业客户签署协议范围外的情形，与 B 端客户也不存在纠纷或潜在纠纷。

(4) 发行人已对历史上存在的数据合规瑕疵进行了整改

公司已对历史上存在的数据合规瑕疵进行了整改，并对相关 APP 产品的部分版本进行了自查及整改，且根据工信部通报的启信宝 APP 部分版本超范围收集个人信息事宜完成整改工作。启信宝 APP 合规情况详见本次审核中心意见落实函回复：“1. 关于媒体质疑情况自查”之“二、核查情况”之“（五）启信宝 APP 合规及数据准确性”。

(5) 发行人的数据合规管理及个人信息保护已获得相关部门认可，并经第三方专业机构评估

1) 公司的数据合规管理获得主管部门认可：上海市公安局静安分局网络安全保卫支队已通过访谈确认公司不存在因数据采集、购买、保管、使用、销售等事宜违反网络信息安全相关的法律、法规和规范性文件的情形，不存在因网络信息安全方面的违法违规行为而受到行政处罚、或因涉嫌违法违规被立案调查的情形。

2) 发行人已成为相关部门数据安全认证方面的试点单位：发行人已成为信

通院“卓信大数据计划”数据安全能力认证（DSMC）首批企业、数据分类分级首批试点单位、数据安全风险管理认证企业、大数据平台安全认证企业。公司的数据合规建设也作为优秀案例首批入选信通院《2022 企业数据安全优秀案例集》。

3) 发行人已成为相关部门个人信息保护方面的试点单位：发行人参与了中国网络安全产业联盟（CCIA）个人信息保护影响专题工作（PIA）专题工作评估。经由电标院、信通院等单位的法律与技术专家组成的工作组评估后，发行人已成为通过该等评估的首批试点单位。

4) 发行人的主要信息系统已通过定级等级备案：发行人的主要信息系统均通过工信部网络安全防护定级三级备案，并通过公安部国家信息系统安全等级保护三级备案。公司通过《ISO20000&27001 认证（信息技术服务管理体系标准&信息安全管理体认证）》等各项资质认证。

5) 第三方专业机构瑛明、中伦已出具评估报告并确认发行人在所有重大方面符合《网络安全法》等所涉法规的相关规定：公司聘请了专业机构瑛明、中伦进行专项的数据安全合规评估。针对公司境内产品及系统，瑛明、中伦已分别出具《数据安全管理的尽职调查报告》，认为公司在所有重大方面符合《网络安全法》等所涉法规的相关规定。针对公司扫描全能王（境外版）、名片全能王（境外版）等境外产品，瑛明、中伦已确认根据《通用数据保护条例》的相关规定，公司数据保护及隐私管理制度经高阶评估方面没有整改建议。

6) 中国电子技术标准化研究院赛西实验室出具评估报告，对公司数据合规管控给予认可：赛西实验室出具了《数据安全保护能力自评估报告》：合合信息已根据标准要求，建立了完整的覆盖数据全生命周期的数据安全管理制度规范体系；已开展了数据分级分类工作，明确各类数据的敏感程度与安全保护需求，已配备了数据加密、数据脱敏等数据安全保护技术，以保障合合信息各个系统在数据全生命周期的安全；定期对数据安全风险进行检测，对数据接口采取了严格的保护措施，在接口口令认证、数据暴露面、访问权限、高危操作方面暂未发现安全风险。

7) 发行人与主管部门相关单位合作，常态化进行数据安全及合规评估：公

司与工信部直属科研事业单位——中国信息通信研究院、中国电子技术标准化研究院分别签署服务合同，信通院对公司主要产品扫描全能王、启信宝的技术方案进行检测和验证，提供建议咨询服务，电标院对公司数据安全合规进行常态化的指导评估。

公司与上海市通信管理局 2021 年网络安全支撑单位——上海映涵信息科技有限公司、北京智游网安科技有限公司分别签署了服务合同，对公司主要产品扫描全能王、名片全能王、启信宝进行用户隐私权限进行年度测试、对收集使用用户信息进行常态化的合规评估。根据主管部门相关单位的常态化评估建议，公司实时更新并完善内部数据合规管理。

经核查，保荐机构、发行人律师认为：发行人报告期内不存在向个人供应商采购数据的情形，发行人历史上向个人供应商采购数据的数据来源合法合规；发行人在业务开展过程中不存在将开展 C 端或 B 端业务时获取的用户数据进行销售的情形。发行人已对历史上存在的数据合规瑕疵进行了自查及整改，自查及整改结果符合主管部门要求。发行人已建立健全数据合规及个人隐私保护的内部控制制度并有效执行，发行人的数据合规管理及个人信息保护已获得相关部门认可，并经第三方专业机构评估。

（五）启信宝 APP 合规及数据准确性

1、媒体质疑情况

部分媒体报道质疑启信宝 APP 的相关版本于 2021 年 11 月 3 日被工信部通报涉及超范围索取权限、过度收集用户个人信息等问题，报告期内，部分媒体还报道启信宝企业信息展示页面曾存在信息更新不及时导致不准确的情形，比如启信宝产品及市场上其他几款同类 APP 在查询企业信息时显示银行的分行、支行负责人的职位被列成了“法定代表人”，并质疑公司报告期内存在的数据类相关诉讼及纠纷案件。

2、核查情况

（1）发行人已完成启信宝 APP 相关版本的整改工作

2021 年 11 月 3 日，工信部信息通信管理局通报 38 款 APP 超范围索取权限、

过度收集用户个人信息等问题，启信宝 APP 在华为应用市场的 8.1.1.0 版本被指超范围收集个人信息。公司已完成启信宝 APP 相关版本的整改工作，并于 2021 年 12 月 10 日在工信部工作部署而建立的官方检测平台全国 APP 技术检测平台上通过复测、完成整改，符合工信部信息通信管理局的整改要求。

截至本次审核中心意见落实函回复签署之日，发行人未因此受到工信部或其他任何主管部门的行政处罚。发行人也不存在因网络信息安全方面的违法违规行为而受到行政处罚或因涉嫌违法违规而被立案调查的情形。

就上述事宜，发行人已在二轮问询函回复之“1. 关于合规经营”之“(三) 对发行人报告期内所有存在的不合规情形是否构成重大违法违规、是否构成发行上市障碍逐一发表明确核查意见”之“4、启信宝 APP 超范围收集信息”进行了详细说明。

(2) 发行人报告期内涉及的数据类型纠纷案件大部分已审结，发行人涉及的履行义务均已完成

截至本次审核中心意见落实函回复签署之日，除金山兴盛一案（（2021）京 73 民终 3418 号）二审开庭尚未判决、阿鑫实业（（2023）沪 0106 民初 33296 号）及铜林矿业（（2023）沪 0106 民初 33292 号）案件一审未开庭，其他相关案件均已审结，发行人涉及的履行义务均已完成，发行人也不存在因开展数据业务而受到主管部门的行政处罚的情形。

1) 关于金山兴盛一案：发行人已在二轮问询函回复之“2. 关于数据”之“2.3 关于数据管理”之“(二) 发行人数据类纠纷案件的主要情况、解决方式及对发行人的影响，该等案件中发行人存在的问题（如有）及整改措施”对金山兴盛一案进行了详细说明。

2) 关于阿鑫实业、铜林矿业诉上海生腾网络侵权责任纠纷一案：2022 年 12 月 30 日，原告阿鑫实业、铜林矿业诉上海生腾网络侵权责任纠纷案立案。原告主张被告产品公示的原告信息不实，侵害其名誉权并造成损失，诉请判令被告在媒体及产品上公开刊登声明、赔礼道歉，赔偿经济损失、合理费用并承担诉讼费用。2023 年 6 月 26 日，阿鑫实业及铜林矿业撤诉。2023 年 7 月，阿鑫实业、

铜林矿业再次对上海生腾提起诉讼。2023年7月10日，发行人收到上海市静安区人民法院关于阿鑫实业及铜林矿业网络侵权责任纠纷的传票，截至本次审核中心意见落实函回复签署之日，该案一审尚未开庭。

经保荐机构、发行人律师核查：金山兴盛一案系原告以不正当竞争纠纷为由诉苏州贝尔塔、上海生腾及北京百度网讯科技有限公司，实则为数据展示引起的纠纷，根据法院一审判决，苏州贝尔塔和上海生腾在本案中不构成不正当竞争；阿鑫实业及铜林矿业一案，发行人已对案件中所涉纠纷进行了核查并在相关网站上进行了内容调整。上述案件对发行人开展数据业务及发行人的声誉不存在重大不利影响。

(3) 发行人报告期内涉及的数据类型纠纷的原因及核查情况

发行人及其子公司涉及数据类纠纷案件为启信宝产品展示公开信息产生的纠纷，案由主要是“名誉权纠纷”、“侵权责任纠纷”，引起纠纷的主要原因及核查情况如下：

1) 展示相关公开信息引起纠纷：发行人产品展示信息中存在企业相关失信被执行人、诉讼相关案件等从中国裁判文书网、中国执行信息公开网获取的公开信息。部分相对方以侵犯名誉权等为由起诉。根据《征信业管理条例（国务院令 第 631 号）》，启信宝 APP 展示的信息系发行人基于提供企业征信服务的目的而从公开渠道获取的企业数据，启信宝 APP 的运营主体上海生腾作为经人行上海分行备案的企业征信机构，收集并展示相关信息符合法律法规的规定。

2) 信息更新不及时等原因引起的纠纷：发行人通过自动化访问获取企业工商、司法诉讼等公开信息。发行人虽设置了自动化访问定期检查机制，但由于被采集网站实时更新、变化，该种采集方式存在一定滞后性，因此存在信息更新不及时从而引起纠纷。另外，发行人通过自动化访问获取数据时，也存在个别数据采集与实际情况不匹配的情形。

发行人已在二轮问询函回复之“2. 关于数据”之“2.3 关于数据管理”之“(二) 发行人数据类纠纷案件的主要情况、解决方式及对发行人的影响，该等案件中发行人存在的问题（如有）及整改措施”对上述事宜进行了详细说明。

(4) 发行人已制定了一系列措施并有效执行，尽可能确保启信宝展示信息的准确性

保荐机构、发行人律师查询了媒体质疑提及的启信宝中的银行、分行显示信息，页面已正确显示为“负责人”。此外，根据媒体质疑中提及的信息准确性问题，为减少启信宝信息展示更新不及时、个别数据采集与实际情况不匹配的情况，发行人已制定一系列有效措施提高启信宝展示信息的准确性：

1) 制度和内控方面：针对自动化访问技术获取数据的情形，发行人制定了《通过自动化访问获取数据操作规范》等相关的管理制度及流程，对自动化信息采集工作实施严格管理，并对被采集网站做定期检查，确保数据采集符合综合评估；针对外部数据供应商，公司制定了《供应商准入制度》、《数据采购管理规范》等制度文件，严格筛选数据供应商，要求供应商保证数据的更新及时性，并及时替换不符合要求的供应商。此外，针对数据延迟、不准确的情形，发行人还制定了《信息主体异议处理规范流程》，建立了如出现因数据延迟、不准确产生异议时的解决处理机制及内控流程，明确公司对信息主体异议处理的管理要求，规范员工对信息主体异议处理的行为。

2) 业务开展方面：公司不断优化启信宝数据更新的相关功能及用户提示，在启信宝 App 企业信息展示页面显著位置增加了被展示的企业信息的更新时间，并在启信宝网页版企业信息展示页面的“工商信息”一栏下方标注了信息的“更新时间”及“来源”，有助于用户了解该等展示信息的权威性及时效性，用户同样可以通过监控企业动态功能，订阅启信宝推送的企业最新动态，帮助用户及时掌握更具时效性的信息。

3) 技术提升方面：报告期内公司开展基础数据校验、完善内部数据治理体系的工作方法如下：

①建立数据治理及校验规则库，积累业务开展过程中积累的包括数据归类错误/格式错误/数值错误等数据准确性规则，以提升问题数据发现率、提升数据准确性。截至本次审核中心意见落实函回复签署之日，发行人已积累了超过 3,500 条规则，覆盖近 400 张数据表，涵盖了公司商业大数据业务开展所需的主要数据字段和数据维度。

②针对自动化访问技术从互联网获取的公开数据，公司一直在投入研究并优化自动化采集的算法以提升采集效率，从而减少数据更新的滞后。公司利用 AI 技术运用前述数据治理及校验的规则对不同来源融合的数据进行实时校验、核对，保证数据的准确性、时效性。

同时，公司也在持续更新和优化底层技术，以提高数据准确性。公司首先对不同来源、不同结构、不同内容的数据进行清洗处理，并通过实体抽取、关系抽取和属性抽取将企业关联数据融合在一起，形成一个庞大的企业关系知识库，结合全量更新、增量更新和实时更新三种更新方式保证数据的时效性与一致性。公司基于已有的底层数据不断优化数据展示效率，减少数据延迟，做到提供更低延迟、更快响应、更富增值的数据产品与服务。

发行人已在招股说明书“第三节 风险因素”之“一、与发行人相关的风险”之“（二）法律风险”之“5、数据类纠纷相关风险”中进行了特别提示。

经核查，保荐机构、发行人律师认为：启信宝 APP 已通过全国 APP 技术检测平台的复测，超范围收集信息事宜已在规定期限内完成整改，符合工信部信息通信管理局的整改要求，不构成重大违法违规行为；上述诉讼纠纷对发行人开展数据业务及发行人的声誉不构成重大不利影响。发行人不断优化自动化访问获取的规则及算法并进行多重复核校验，持续提升启信宝展示信息的准确性、及时性，并制定了相关纠纷解决机制，规范产生异议情况下的处理流程，降低产生纠纷的风险。前述事项不构成发行人本次发行上市的障碍。

（六）扫描全能王 APP 业务发展可持续性问题的

1、媒体质疑情况

部分媒体报道质疑：（1）扫描全能王 APP 面临激烈竞争，几家互联网大厂已切入了文字识别领域；（2）扫描全能王 APP 年 VIP 用户续费率出现下滑；（3）扫描全能王 APP 新增付费转化率较低，付费用户增长是否进入瓶颈期，未来发展是否具备可持续性。

2、核查情况

保荐机构就该问题的核查情况如下：

(1) 尽管几家互联网大厂早已切入文字识别领域，但公司已建立了较高的竞争壁垒

扫描全能王的用户不仅包括中国境内用户，也包括境外用户。全球范围内业务定位为扫描及文字识别的 APP 产品中知名度较高的产品主要包括微软 2015 年推出的 Office Lens、Adobe 2017 年推出的 Adobe Scan、Google 2017 年推出的 Google Lens。在过去七八年的竞争当中，扫描全能王凭借领先的技术及用户体验，持续保持领先的全球用户规模，扫描全能王的月活在 2021 年即超过 1 亿且持续保证增长趋势，领先于下述竞品的全球月活规模，同时用户评分也保持在领先的水平：

| 文字识别类 APP | 简介 | 上线时间 | 用户数量 (App Store 与 Google play 的月活之和) | 用户体验 (以 App Store 部分地区 2022 年 12 月 31 日的评分为例) |
|---------------------|---|--------|---|---|
| 扫描全能王 CamScanner | 将智能手机变成便携式扫描仪，采集、管理各种文档产品，支持文档图片的智能切边、智能图像增强、票据/证照/表格等多种文档的文字识别功能 | 2010 年 | 2022 年 1-12 月的平均月活为 11,907.4 万 | iOS 中国区评分：4.9，评分个数 367.8 万个 iOS 美国区评分为：4.9，评分个数 104.9 万个 |
| Office lens | 微软集团旗下的扫描与文字识别产品，可将图像转换为 PDF、Word、PowerPoint 或 Excel 文件，将打印或手写文本数字化 | 2015 年 | 2022 年 1-12 月的平均月活为 1,419.7 万 | iOS 中国区评分：4.9，评分个数 2.4 万个 iOS 美国区评分为：4.8，评分个数 7.9 万个 |
| Google Lens | Google 集团旗下开发的一款基于图像识别和 OCR 技术的产品，可以根据图片或拍照识别出文本和物体 | 2017 年 | 2022 年 1-12 月的平均月活为 1,545.5 万 | Google lens 未在 App Store 上架 |
| Adobe Scan | Adobe 集团旗下开发的扫描与文字识别产品，能够将任何纸质文档转换为 Adobe PDF，同时具备 OCR 功能 | 2017 年 | 2022 年 1-12 月的平均月活为 5,110.7 万 | iOS 中国区未上架 iOS 美国区评分为：4.9，评分个数 106.0 万个 |

注 1：APP 简介等信息来自官网等公开信息。

注 2：月活数据来源为 App Annie。

注 3：iOS 没有全球评分，只有分地区评分个数数据，评分满分为 5 分，数据来源为七麦数据。

过去几年中，阿里、百度、腾讯等国内几家互联网大厂将扫描及文字识别功能嵌入其旗下产品中（例如阿里的夸克浏览器、百度的百度网盘、腾讯的 QQ 浏览器等），该等产品的主打功能为浏览器搜索或者文件云存储，扫描及文字识别

不是其主打功能，增加此类功能的原因是为用户提供更多元化服务、满足部分用户的比较基础的扫描与文字识别需求，与扫描全能王的专业定位不同。以 2022 年 12 月中国 App Store 中用户关键词搜索热度为例，“扫描全能王”的搜索热门程度高于“扫描”、“文件扫描”、“文档扫描”、“拍照取字”、“文字识别”、“文字提取”等功能词，说明从用户角度，扫描全能王已成为扫描及文字识别的代名词。

扫描全能王已在用户中建立专业可信的扫描及文字识别产品品牌心智的主要原因是其优质的用户体验，这背后主要是先进的核心技术支持。扫描全能王凭借行业领先的智能文字识别技术，解决了诸多用户痛点和技术难题，例如：①由于文档图像质量差导致“看不清”的问题；②由于版面复杂多样导致的“看不准”的问题；③由于非限定情况下文字识别率低导致的“认不全”的问题；④由于文档类型多样，版式复杂导致的文档图像“难理解”的问题。

在 2020 年、2022 年 CNAS 认可的检测实验室出具的一系列测试报告中，公司智能文字识别技术在各类场景测试中识别率领先于境内外互联网大厂。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。在 2023 年国际文档分析识别大会（ICDAR）的文本图像篡改检测算法竞赛中，公司研发团队取得文本篡改检测任务的冠军。2023 年 5 月，公司在中国信通院旗下中国泰尔实验室的 OCR 智能化服务评估和智能文档处理系统技术能力模块评估中获得最高评级，包括识别性能、AI 核心能力、数据安全性及可靠性等多维度评估。

智能识别文字技术需要长期持续投入较多的资金和尖端技术人才支撑，结合客户实际落地场景的技术难点去打磨 AI 底层算法及应用模型是需要长时间的研发投入、实践改良，难以通过资源投入的方式在短时间进行模仿或者超越。公司在智能文字识别技术领域的研发投入已超过 15 年，因此在各类场景算法模型的打磨与研发积累上具有较高的竞争壁垒。

综上分析，尽管互联网大厂在几年前早已切入文字识别领域，但公司扫描全能王产品凭借领先的智能文字识别技术、专业的品牌定位、优质的产品体验和多年积累的用户规模已经建立了较高的竞争壁垒。

发行人已在招股书“第三节 风险因素”之“二、与行业相关的风险”之“（一）

市场竞争加剧的风险”中补充扫描全能王 APP 业务面临竞争的相关风险提示。

(2) 扫描全能王年 VIP 用户续费率的变化具有合理性

针对扫描全能王 APP 年 VIP 用户续费率变化原因，发行人已在一轮问询函回复之“5.2 关于 C 端业务”之“一、请发行人披露：报告期各期各 C 端产品的付费用户数量及占比、月活用户数量、ARPPU 值、新增付费转化率、老客户续费率等核心业务指标及其统计口径、数据来源，分析指标变动原因和趋势”进行了详尽的说明。相关核心内容总结如下：

| 年份 | 本期年 VIP 到期用户数量 (个) | 本期年 VIP 到期中选择续费的用户数量 (个) | 年 VIP 续费率 |
|--------|--------------------|--------------------------|-----------|
| 2020 年 | 488,726 | 224,448 | 45.93% |
| 2021 年 | 1,232,260 | 542,753 | 44.05% |
| 2022 年 | 2,498,028 | 984,827 | 39.42% |

2020 年-2022 年扫描全能王的年 VIP 续费率呈现下降趋势，主要由于：扫描全能王的运营策略是主推年费 VIP，引导付费用户转化为年 VIP 用户，2019 年开始向部分用户推广年 VIP 订阅第一年价格优惠力度较大（第二年恢复原价）的促销方案。如表格所示，扫描全能王 2020 年、2021 年以及 2022 年各期年 VIP 用户数量分别为 48.87 万、123.23 万、249.80 万（数据口径为上表中当年到期的年 VIP 数量），2020 年-2022 年的年复合增长率达 126%，随着年 VIP 付费用户群体的快速扩大，部分价格敏感的用户在首期 VIP 到期后选择不再续订，导致报告期内的年 VIP 续费率整体有所下降。

尽管 2020-2022 年之间扫描全能王的年 VIP 续费率有所下降，但从绝对金额看，扫描全能王年 VIP 当年到期后续费用户的订单金额从 2020 年的 0.5 亿元快速增长至 2022 年的 2 亿元，依然保持较高的增速。

综上所述，报告期内，扫描全能王年 VIP 用户续费率的变化主要原因是公司运营策略的调整，具有合理性，扫描全能王的续费用户订单金额依然保持较高增速。

(3) 扫描全能王的新增付费转化率、用户群体具有较大增长空间，未来收入增长具有可持续性

总体来说，扫描全能王收入增长将保持可持续性：

首先，报告期内扫描全能王的收入保持 37.77% 的年复合增长率，关键运营数据保持增长趋势：2020-2022 年，扫描全能王的月活用户数量从 9,767.55 万个快速增长到 11,725.74 万个，说明用户的活跃程度持续向好；新增付费转化率从 2020 年的 0.58% 增长到 2022 年的 0.75%，说明用户的付费意愿及付费转化持续提升；当期付费用户从 2020 年的 364.14 万个增长到 2022 年的 508.19 万个，说明付费用户的群体快速扩大。报告期内，扫描全能王的营销推广费较低，因此前述增长的驱动力主要为市场需求、公司自身的技术及产品优化、品牌口碑积累等方面的利好因素，而这些利好因素具有持续性，将驱动收入持续增长：

1) 新用户角度：一方面未来用户群体与使用场景的拓宽将驱动新用户持续增长，另一方面，海外市场存在多个渗透率较低的国家有较大新用户拓展空间，新增付费转化率存在较大提升空间

2020 年-2022 年，扫描全能王的新增付费用户分别为 99.69 万个、107.38 万个、114.56 万个，新增付费转化率分别为 0.58%、0.68%、0.75%，处于持续上升趋势的主要原因是：从公司业务角度，报告期内扫描全能王推出了多个创新功能使得用户使用场景拓宽，优异的用户体验与多年积累的产品口碑使得品牌知名度快速提升；从行业角度，随着用户对个人文档数字化管理的意识增强，在用户群体扩大的同时、用户付费意愿逐步提升。未来前述利好因素将持续支持扫描全能王快速发展，新增付费转化率有较大的提升空间，未来新增付费转化率和付费用户规模的提升主要从以下两个角度：

从用户类型角度，随着消费者对个人数字化资产管理意识和需求不断增强，不仅财会、行政、销售等商务场景需要文档资产管理 APP 实现高效的移动办公，教师、学生、家庭等不同生活场景也对证件、试卷、手写文件等不同文档资产具有电子化管理的的需求，用户群体与使用场景的拓宽将驱动新用户保持持续增长。未来扫描全能王还将进一步研发创新功能，优化不同行业用户群体的体验。

从地域拓展角度，海外一些发展中国家的移动办公 APP 的渗透率与付费率仍有较大的提升空间，如下表所示，扫描全能王在巴西、印度尼西亚、墨西哥等国家的 2022 年新增付费转化率低于 0.5%，大幅低于 1.65% 的中国用户新增付费转化率，同时按照 2022 年付费用户数量占当地移动互联网用户总数量比重计算

渗透率，前述 3 个国家付费用户的渗透率也距离中国有较大差距：

| 代表性国家 | 扫描全能王 2022 年新增付 费转化率 | 扫描全能王 2022 年付费用 户数量（万个） | 当地移动互联网 用户总数量 （百万个） | 扫描全能王 2022 年付费用 户渗透率 |
|-------|----------------------------|-------------------------------|---------------------------|----------------------------|
| 中国 | 1.65% | 348 | 948 | 0.37% |
| 巴西 | 0.44% | 15 | 156 | 0.10% |
| 印度尼西亚 | 0.29% | 14 | 129 | 0.11% |
| 墨西哥 | 0.21% | 5 | 81 | 0.06% |

注 1：当地移动互联网用户总数量的数据来源为市场研究机构 eMarketer；

注 2：扫描全能王 2022 年付费用户渗透率=扫描全能王 2022 年付费用户数量/当地移动互联网用户总数量

未来随着海外新兴国家当地经济水平和数字化水平的发展、付费意愿的增加以及公司本地化运营深入，预计付费率有较大的提升空间，具有一定的收入增长潜力。未来公司将在对海外国家当地政策与市场环境进行深入研究之后，为当地用户提供更符合其使用习惯的产品功能，并进行本地化的产品运营。

报告期内扫描全能王新增付费用户贡献的订单金额从 2020 年的 26,242.63 万元提升到 2022 年的 38,592.36 万元，年复合增长 21%，也印证了扫描全能王在新用户拓展方面的能力。因此基于公司在智能文字识别领域建立的先发优势、积累的品牌影响力、在开发新用户类型及新兴国家用户的研发与运营能力，预计扫描全能王来自新用户的收入增长具有可持续性。

2) 存量用户角度：存量用户粘性强，预计将贡献稳定的收入基本盘

存量用户已经在扫描全能王上持续积累了证件、税单、合同、手写文件等价值较高的文档资产，扫描全能王也可进一步作为用户文档资产管理的平台，因此用户迁移成本高、粘性强。报告期内扫描全能王的续费率基本稳定，其中月 VIP 用户的续费率保持在 75%左右，年 VIP 用户的续费率保持在 40%左右。在用户预收款方面，扫描全能王的合同负债余额从 2020 年末的 12,038.59 万元快速增长至 2022 年末的 28,171.83 万元。因此综上分析，存量用户预计未来将贡献稳定的收入基本盘。

3) 技术角度：智能文字识别技术仍有提升空间，扫描全能王将通过进一步的技术创新提高识别效率与效果

在一些实际的复杂应用场景中，比如拍摄质量差、光照不均、纸张褶皱、阴影、版面排版复杂（比如票据/表格等）、书写潦草（手写）、字体形态各异（含印章/公式/二维码等）等情况下，识别率距离用户满意的水平仍有较大距离。在一些客户的真实业务场景中，容错率要求极高，比如在合同、财报审核等业务场景中，客户的预期是接近 100% 的准确率。因此智能文字识别技术仍有提升空间，需要克服众多技术难点。

通过解决各种场景下的疑难问题，给用户达到甚至超越预期的体验是发行人一直追求的目标。公司计划运用深度学习、NLP 技术进一步提升扫描全能王在复杂场景下的识别效率与效果，提升技术的应用性能，包括多语言场景、多文字方向场景、多样化字体和版面场景，加强对图表识别理解、扫描文档图像分析还原及扫描增强等智能文字识别技术研发，提升文本图像质量、识别准确率，进一步优化用户体验，促使用户为优质体验持续付费，促进收入持续增长。

综上所述，未来扫描全能王的用户群体与使用场景预计将不断拓宽，且扫描全能王计划对核心技术与创新功能进行进一步升级、不断优化用户体验、策略性地拓展海外市场，基于多年积累的用户基数及用户口碑，借助口碑传播的网络效应，预计扫描全能王的未来收入增长具有可持续性。

（七）名片全能王 APP 业务发展可持续性问题

1、媒体质疑情况

部分媒体报道质疑名片全能王 APP 新增付费转化率较低，付费用户增长是否进入瓶颈期，未来发展是否具备可持续性。

2、核查情况

保荐机构就该问题的核查情况如下：

针对名片全能王 APP 新增付费转化率以及未来发展的问題，发行人已在第一轮问询函回复之“5.2 关于 C 端业务”之“一、请发行人披露：报告期各期各 C 端产品的付费用户数量及占比、月活用户数量、ARPPU 值、新增付费转化率、老客户续费率等核心业务指标及其统计口径、数据来源，分析指标变动原因和趋势”以及“二、（二）5、名片全能王 APP 业务的收入具有稳定性与可持续性”

进行了详尽的说明。相关核心内容总结如下：

近几年，微信等社交软件迅速崛起，部分用户使用社交软件添加好友交换联系方式来替代交换纸质名片，导致名片全能王的关键运营指标面临月活下滑的挑战。发行人已在招股书“第三节 风险因素”之“二、与行业相关的风险”之“（一）市场竞争加剧的风险”中补充名片全能王 APP 业务面临竞争的相关风险提示。

应对此挑战，2017 年 5 月开始，公司利用快速崛起的微信小程序渠道，开发便于用户在微信小程序使用的功能，借助微信小程序便于分享、传播的优势，利用小程序打通 APP 用户与微信名片全能王用户的连通。从生命周期角度来看，名片全能王 APP 用户发展已基本进入成熟期，但名片全能王微信小程序等渠道的用户仍具有发展潜力，因此公司对其的业务策略主要为注重付费用户的商业化运营以及小程序电子名片用户的发展，保持收入的稳定性与可持续性。

针对老用户，报告期内名片全能王的老客户续费率相对稳定，其中年 VIP 续费率保持持续上升，自 2020 年的 55.57% 增长至 2022 年的 64.05%。在用户预收款方面，名片全能王的合同负债余额从 2020 年末的 905.31 万元增长至 2022 年末的 1,324.91 万元。用户在名片全能王上会积累电子名片资产，因此对 APP 的使用粘性较高。根据公司对名片全能王月活用户使用年限的分析，2020 年初到 2022 年 12 月间，每月的活跃用户中，使用 2 年以上的活跃用户的比例占比稳定维持在 50% 以上，说明用户粘性保持较好；从付费用户看，2021 年和 2022 年，当月付费用户中，使用 2 年以上的用户占比平均值已超过 70%，说明使用年限越长的用户越倾向于付费。未来名片全能王计划进一步通过付费功能的优化、用户运营优化等策略进一步提升老用户的续费率。

针对新用户，2020 年-2022 年，名片全能王的各期新增付费转化率分别为 0.26%、0.20%、0.23%，各期新增付费用户数分别为 1.17 万个、1.15 万个、1.23 万个。名片全能王新增付费用户数与新增付费转化率基本稳定的原因主要是名片全能王十余年积累的品牌影响力、先发优势、优质的用户体验、产品功能的持续优化。公司在吸引用户付费的产品运营策略方面已经积累了一定的经验。报告期内名片全能王新增付费用户贡献的订单金额从 2020 年的 392.82 万元提升到 2022 年的 463.79 万元。

报告期内，名片全能王的月活受到了微信等社交软件的影响有所下滑。未来用户出于环保、便携度等考虑可能会减少纸质名片的使用，但互换名片作为商务社交中的重要组成部分，用户会互换电子名片作为替代方式。由于社交软件并不实名展示用户的姓名、单位与职位，可信度低于电子名片，而且名片全能王电子名片具有辅助的企业认证、职位认证、实名认证等增加可信度的功能，解决了上述痛点，2020年开始，微信小程序渠道的电子名片用户成为名片全能王新的用户增长渠道，微信小程序月活自2020年的92.66万个增长至121.97万个。未来公司计划加大力度培养用户使用小程序名片全能王电子名片的习惯，挖掘小程序电子名片用户的付费潜力。未来公司计划将电子名片商业模式复制到海外社交软件，优化用户通过名片全能王将名片分享到海外社交软件的路径，开拓基于海外社交软件电子名片用户以及扩展用户付费场景。

从历史数据看，名片全能王的当年付费用户数从2020年的8.24万个增长至2022年的8.43万个，收入由2020年的1,667.80万元增长至2022年的2,080.95万元，2020-2022年的收入年复合增长率为11.70%。报告期内，名片全能王的营销推广费极低，因此付费用户和收入小幅增长的驱动力主要为公司品牌与口碑的持续积累、公司自身的技术提升及产品策略优化、用户粘性较强等方面的因素，而这些驱动因素在中短期内预计不会产生重大不利变化，因此将支持未来收入持续保持稳定。

综上所述，名片全能王APP用户发展已进入成熟期，公司对其的业务重心主要是付费用户的商业化运营以及小程序电子名片用户的发展，预计未来收入具有稳定性与可持续性。

（八）启信宝APP业务发展可持续性问题

1、媒体质疑情况

部分媒体报道质疑启信宝的核心竞争力，启信宝与可比APP企查查、天眼查在产品功能、技术路线、业务模式较为相似，启信宝与企查查、天眼查的用户数量存在差距，启信宝仅在2015年有外部融资，企查查和天眼查进行了多轮外部融资，投放了较多的广告，启信宝广告营销费用超过收入，处于亏损状态，且新增付费转化率逐年下降，未来启信宝业务发展是否具备可持续性。

2、核查情况

保荐机构就该问题的核查情况如下：

针对报告期内启信宝与可比 APP 企查查、天眼查在产品功能、技术路线、业务模式、用户数量、市场地位的比较，发行人已在第一轮问询函回复之“6 关于市场地位”之“一、（四）启信宝与企查查、天眼查在产品功能、技术路线、用户数量、市场地位等方面的对比情况”中进行了详尽的说明。

针对启信宝新增付费转化率下降、启信宝业务处于亏损状态的原因，以及未来启信宝业务发展的可持续性，发行人已在第一轮问询函回复之“5.2 关于 C 端业务”之“二、（二）2、启信宝新增付费转化率逐年下降的原因、影响及业务的收入增长可持续性”中进行了详尽的说明。

相关核心内容总结如下：

（1）启信宝现状分析

商业大数据 C 端 APP 市场集中度较高，主要参与者主要为企查查、天眼查、启信宝与百度旗下的爱企查，其中企查查、天眼查、启信宝（简称为“一信两查”）为头部前三名。商业大数据 C 端 APP 市场存在一定的同质化竞争情况，一信两查的主体数据来源均采取工商司法等公开披露的企业信息，虽然在数据覆盖度、特色数据库、产品付费功能、APP 技术路线方面存在差异化的细节，但从用户感知角度来说，相似度较高，用户切换成本相对较低，因此需要通过持续投入宣传推广的方式提高用户影响力、品牌认知度来保持用户规模和市场份额。

启信宝与其他两家用户数量存在一定差距的原因主要为天眼查与企查查近几年在广告宣传方面投入了较大力度的资源：

根据公开信息查询，天眼查在 2015-2019 年之间进行了四轮私募融资，合计融资金额数亿元；企查查在 2015-2019 年之间进行了六轮私募融资，合计融资金额数亿元。天眼查聘请了演员吴刚作为代言人，企查查聘请了演员林志颖作为代言人。根据 2019 年天眼查以不正当竞争为由起诉企查查的诉讼记录，天眼查首创“查公司，查老板，查关系”广告语，在 2017 年至 2019 年期间在全国各地投入近 2 亿元资金通过地铁广告、微信推送、影视植入等线上、线下方式大范围宣

传和推广天眼查的付费功能。

与之相比，合合信息在营销费用的投入力度有所控制，启信宝未聘请演员作为代言人，主要采取流量推广等方式进行宣传推广，2017-2019年启信宝营销费用为525.93万元、930.29万元、5,925.35万元，与竞品相比较低的营销推广水平部分导致启信宝的月活水平位居行业第三名，且新增付费转化率出现下降。

针对亏损问题，2022年度，公司优化启信宝的投放策略，主动减少了转化率较低的广告投入，启信宝的营销投入下降至2,204.43万元，广告宣传费用同比减少4,180.35万元，2022年启信宝C端业务的净亏损已缩减至944.98万元。

(2) 启信宝未来业务发展可持续性分析

针对启信宝APP未来业务发展可持续性问题分析如下：

1) 首先，行业角度的多重利好因素预计中短期内不会发生重大不利变化

①多重有利政策驱动商业大数据C端APP市场快速发展

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，发布多项支持政策推动社会整体信用意识的提升，促进商业大数据C端APP市场快速发展：2020年4月，《中共中央、国务院关于构建更加完善的要素市场化配置体制机制的意见》正式公布，将“数据”列入五大生产要素范围中，并提出从推进政府数据开放共享、提升社会数据资源价值、加强数据资源整合和安全保护三方面加快培育数据要素市场；2023年2月，中共中央、国务院印发的《数字中国建设整体布局规划》强调“释放商业数据价值潜能”。

②用户场景的拓宽驱动用户规模扩张

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，商业大数据C端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。艾媒咨询数据显示，除了判断企业合作风险（49.6%）、投资参考（45.8%）等商业化用途有较多使用外，求职、理财、消费等出于个人需求而使用的场景也逐渐占据更加重要的地位，2020年中国商业查询用户规模达到4.79亿人，2022年有

望增至 5.89 亿人。根据灼识咨询研究，中国商业信息查询 C 端市场的用户规模正不断扩大，付费率正逐步提升，预估未来长期内的可触达市场（Total addressable market）接近 80 亿元。

③政策严格监管数据合规性，提高行业进入门槛

国家颁布了多项政策严格监管数据合规性，提高了商业大数据 C 端 APP 市场的进入门槛。2021 年 6 月 10 日，全国人民代表大会常务委员会发布了《中华人民共和国数据安全法》，是我国关于数据安全领域的首部基础性法律，为各行业数据安全提供监管依据。2021 年 11 月 1 日，《中华人民共和国个人信息保护法》正式生效。随着主管部门持续健全我国数据合规方面的法律法规，细化数据安全方面的合规要求，健全个人信息保护工作体制机制，规范了行业发展，提高了行业进入门槛，有利于头部规范企业。

2) 启信宝采取差异化竞争策略，更加注重产品自身的竞争力的提升

①将合规与可信度打造为企业品牌力的重要构成

商业大数据行业对数据的真实性、有效性、可靠性极为重视，因此启信宝将数据合规作为经营发展的重要底线，从行业内权威性较高的数据供应商进行数据采购，将合规与可信度打造为企业品牌力的重要构成。启信宝已与人民数据管理（中卫市）有限公司、国家公共信用信息中心、凭安征信等业内权威机构达成长期深度合作，并引入针对不同垂直场景的特色数据，建设数据生态体系。2022 年初，公司入选中国信息通信研究院云计算与大数据研究所牵头发起的“个人信息保护合规审计领航计划”参与单位之一，致力于推动个人信息保护合规审计工作不断发展，提升各行各业的个人信息保护水平。2023 年 3 月，上海生腾成为首批中国信通院授予的“数据安全管理能力认证(DSMC)”企业。

启信宝已建立了包括数据质量标准、数据治理流程标准、数据资产管理标准、数据业务服务标准在内的数据治理标准体系，并落实到启信宝日常数据治理流程中，未来启信宝目标持续建设合规且高价值的数据资产，搭建行业内领先的数据治理体系；

②持续进行技术与产品创新，提升用户体验

开发针对特定垂直场景的差异化增值功能，使得启信宝成为辅助不同行业用户通过数据分析提高运营效率的生产力工具。公司本次首次公开发行募集资金拟投入的项目“商业大数据 C 端产品及 B 端服务研发升级项目”中，计划在启信宝 APP 针对金融、政府、供应链、法律、媒体等垂直场景推出场景化的功能与界面，添加更丰富的功能模块。例如，银行客户经理是启信宝的高付费人群，但目前银行客户经理仍有很多未被满足痛点，比如数据分析报告的自动化生成、数据商机挖掘（根据投资、股东、担保、产业和交易等关系数据进行商机挖掘）等，启信宝计划通过开发差异化的功能，帮助银行客户经理挖掘高价值、易触达的企业客户，并提高尽职调查效率。

加强启信宝 B 端与 C 端业务之间协同，基于 B 端第一线接触不同行业的客户，反哺 C 端垂直行业场景的深度开发，为用户打造更有价值的产品体验。

3) 报告期内启信宝 C 端收入保持小幅稳定增长

从历史数据角度，启信宝 C 端收入从 2020 年的 6,758.80 万元提升到 2022 年 7,635.28 万元。尽管在竞对持续大力度进行广告投放的竞争环境下，启信宝 C 端业务依然保持了稳定的收入贡献。

综上所述，启信宝 C 端业务面临行业方面的诸多利好因素，而且其在数据合规、技术与产品创新等方面进行了差异化竞争壁垒的打造，但行业内的竞争较为激烈、且竞品在营销推广方面持续大力度投入，因此启信宝 C 端收入未来存在波动的可能性，发行人已经在招股书“第三节 风险因素”之“二、与行业相关的风险”之“（一）市场竞争加剧的风险”中补充启信宝 APP 业务面临的竞争及收入下降的潜在风险。

虽然一信两查在 C 端业务竞争趋向同质化，但启信宝在 B 端业务方面拥有差异化的优势，商业大数据 B 端业务相较 C 端业务具有更高的技术壁垒，主要由于：①商业大数据 C 端业务主要提供数据搜索、数据查询的功能，依靠大数据挖掘、分析和搜索技术作为底层技术，而商业大数据 B 端业务不仅需要依靠大数据挖掘、分析和搜索技术，还需要“多源异构的超大规模动态知识图谱构建技术”、“超大规模知识图谱推理与挖掘技术”等相关技术，该等技术的技术壁垒主要是：外部公有数据与 B 端客户私有数据的融合度、超大知识图谱构建的深

度与速度、知识图谱与客户所在行业的结合度、知识图谱自动动态更新的及时性和准确度；②公司商业大数据 B 端业务服务的政府、大型企业客户常常会要求将其本地化的自有数据与公开的工商等数据做融合，因此对数据安全及保密性的要求极高。公司在数据安全、数据加密等方面技术壁垒较高，采用可靠的安全性设计及安全检测流程，将接口及数据传输过程加密，兼顾数据安全性、同步实时性。例如公司的《合合启信宝 FISync 数据同步软件 V1.0》（著作权登记号 2019SR1191785）可以将数据通过加密通道，单向同步至客户的服务器环境，使得可以在保障信息不外泄的条件下，完成客户内部数据与外部大数据的融合，同时数据同步实时性好。该技术已在招商银行、民生银行等多个银行客户项目中使用，满足了银行客户对数据安全及保密性的高要求。

与主要竞争对手的业务以 C 端为重心不同，公司商业大数据业务在 B 端与 C 端同时形成了规模化收入。2020 年-2022 年，商业大数据 B 端业务收入自 5,028.77 万元快速增长至 9,759.56 万元，年复合增长率为 39.31%，高于商业大数据 C 端业务的同期增长率 6.29%，同时 2022 年商业大数据 B 端业务收入绝对值已是 C 端业务收入的 1.3 倍。未来 B 端业务将成为公司商业大数据业务的重要发展驱动力。

（九）C 端收入真实性

1、媒体质疑情况

部分媒体报道对公司的 C 端收入真实性存疑，质疑的点主要包括：（1）收入报告期内快速增长，高于市场增长率；（2）付费用户数量快速增加，且非注册用户付费的数量、金额和占比都在持续增长，但 C 端核查反馈率较低，是否存在刷单情况；（3）应收账款增长与营收增长匹配性，第四季度营收占比较大是否存在合理性；（4）公司的货币资金与交易性金融资产是否真实，是否存在银行存款虚增以配合营收虚增的情况；（5）公司 C 端 APP 来自公司内部后台系统、微信小程序管理后台等多个渠道的月活无法去重是否合理，是否存在虚增月活数量以配合营收虚增的情况。

2、核查情况

保荐机构、申报会计师就该问题的核查情况如下：

(1) C 端收入和利润增长具备真实性和合理性，并已采取多重手段进行核查

C 端收入的主要增长来源于扫描全能王。扫描全能王报告期内收入分别是 33,343.21 万元、49,325.74 万元和 63,287.86 万元，2020 年-2022 年的年复合增长率为 37.77%，占 C 端 APP 业务收入的比例分别为 78.73%、84.28% 和 86.69%，占比较高，因此报告期内 C 端 APP 收入快速增长的原因主要系扫描全能王产品收入快速增长导致，而扫描全能王产品收入快速增长原因主要系报告期内付费用户数量增加和年费订单占比升高所致。

公司 2010 年即率先推出扫描全能王产品，2010-2017 年间在主要应用市场形成用户口碑，积累用户基数，随着产品关键功能与技术的持续优化，以及 2017 年以来用户对个人文档数字化管理的意识增强，用户对扫描全能王产品的付费意愿逐步提升，加之智能手机以及移动支付技术的不断成熟，扫描全能王进行功能的关键优化与技术创新，扫描全能王业务进入快速成长期，推动扫描全能王的各项运营指标在报告期内保持高速增长。

报告期内，公司净利润增长，主要系在收入增长的情况下，业务存在一定的规模效应，此外公司合理控制相关费用支出，如优化了启信宝的广告投入，故净利润也能保持较好的增长趋势。

为充分论证发行人 C 端收入真实性，保荐机构及申报会计师主要采取了如下核查手段：

| 核查程序 | 核查逻辑 | 核查结论 |
|----------------|--|------------------------------------|
| C 端收入增长的驱动因素分析 | 分析 C 端收入增长的驱动因素，既包括移动互联网整体行业分析，说明 C 端收入增长与行业发展趋势相符，也包括基于各 APP 的具体情况，分析其细分行业的红利、功能的关键优化与技术创新、先发优势与龙头地位、以及营销投入对 C 端收入增长的驱动力。 | 各 C 端 APP 在报告期内收入保持快速增长具有商业合理性。 |
| 用户真实性核查 | 结合各 C 端 APP 的收入和用户情况，对其用户进行分层抽样，确保抽取样 | 该部分 C 端用户真实，用户行为特征具有合理性，不存在明显异常情形。 |

| 核查程序 | 核查逻辑 | 核查结论 |
|----------|--|---|
| | 本能够满足全面性和有效性，并对抽取的样本执行电话核查、APP 推送核查。 | |
| IT 审计核查 | 针对报告期内的所有 C 端用户执行 IT 审计程序，包括：信息系统一般控制、信息系统应用控制、业务数据真实性核查，确保核查程序的完整性和有效性。 | 公司的信息系统一般控制及应用控制有效，业务数据真实性核查无明显异常。核查维度包括：（1）用户量核查：区分不同 APP，进行用户增量趋势分析、用户注册时间段分析、注册用户 IP 地址集中度分析、用户行为数据分析、IT 运维成本的匹配分析，确定发行人新增注册用户变动趋势合理，月活用户数与启动次数等变动趋势合理，发行人系统数据与第三方平台记录数据差异较小；（2）订单数据一致性核查：区分不同 APP，进行境内支付方式订单核查（扫描全能王/名片全能王/启信宝）、APP Store 支付方式订单核查（扫描全能王/名片全能王）、Google Play 支付方式订单核查（扫描全能王/名片全能王），确认发行人系统记录的支付订单与第三方在线支付账单基本一致；（3）第三方支付账号与用户 ID 关系核查：区分不同 APP，分析各年度境内支付金额分层合理性、单一账号同时为多个 user_id 支付的情况以及内部人员支付订单金额情况，确认平均支付金额与产品定价相符，不存在异常情形；（4）付费用户及订单数据分析：从 C 端产品各类型付费订单数量、用户的地域分布、购买金额、付费用户的活跃情况、未活跃用户数量等多维度对付费用户及订单量的合理性进行分析。 |
| 资金流水核查 | 针对发行人及其实际控制人、控股股东、董监高、关键岗位人员资金流水进行核查，确保核查范围的完整性，核查标准的有效性，并对存在异常的银行流水采取进一步的补充核查程序。 | 发行人及其实际控制人、控股股东、董事、监事、高级管理人员、销售主管、采购主管、出纳等关键岗位人员个人账户资金流水无异常情形，不存在资金流向用于发行人自充值或刷单的情形。 |
| 其他财务核查维度 | 核查公司 C 端的具体业务情况和会计处理方式，自主获取并核查公司 C 端业务收入的入账依据，获取第三方平台的结算协议、开发者协议等，核查其具体的结算方式；对 Oracle Fusion 收入摊销功能进行了内控测试，确认关键内控点执行且有效，其中检查 Oracle Fusion 系统中的收入摊销公式，确认公式中的摊销 | 公司收入确认政策符合企业会计准则规定，收入确认依据真实、准确、完整，业务系统和财务系统控制有效，符合公司业务实质。 |

| 核查程序 | 核查逻辑 | 核查结论 |
|------|---|------|
| | 逻辑与财务口径摊销规则一致，核查相关收入的完整及准确性； 终端用户抽样核查，按照不同产品、不同付费方式等对用户进行分层抽样，进行电话核查、APP 推送两种方式进行核查； 结合 C 端 APP 业务的收费模式、服务交付方式、服务开通时间、收入确认方法与时点等情况执行截止性测试，核查相关收入确认是否跨期。 | |

(2) 核查反馈率较低具有合理原因并已补充核查程序

以扫描全能王为例，报告期各期支付金额在 500 元以内的账户合计交易金额占已核查用户交易金额比重均在 99% 以上，单个用户交易金额较低，回复问卷的意愿也相对较低，且出于对电信诈骗的警惕以及对个人信息保护意识的不断提高，公众对于陌生在线访谈抵触，因此 C 端用户电话访谈的接听率、APP 推送问卷的回复率普遍较低。

同时，通信主管机构对电话核查亦有所限制。工信部《通信短信息和语音呼叫服务管理规定（征求意见稿）》（2020 年 8 月 31 日）第十六条规定“任何组织或个人未经用户同意或者请求，或者用户明确表示拒绝的，不得向其发送商业性短信息或拨打商业性电话。”该等规定对公司给用户拨打商业性电话、发送商业性短信息的行为作了限制。

为遵守上述规定，避免 C 端用户集中投诉，中介机构在电话核查的过程中，对第一次拨打未接听的，3-5 天后再进行拨打，若仍未接听，不再重复拨打，转而采用 APP 推送问卷的方式进行核查；对第一次拨打后用户接听但明确拒绝配合回答的，不再重复拨打，转而采用 APP 推送问卷核查，而在同等条件下，对用户而言手动填写问卷较口头回复电话更为繁琐，因为问卷填写过程涉及用户打开 APP，点击问卷链接，并逐一回答较多问题后方可提交，故 APP 推送问卷的回复率亦较低。

针对电话核查及 APP 推送核查的所有用户（包含已反馈用户和未反馈用户），保荐机构及申报会计师将其按用户 ID 去重后，对于扫描全能王、名片全能王和启信宝分别从注册时间、注册 IP 集中度、交易金额分布、第三方支付账号与用

户 ID 关系、App Store 和 Google Play 支付方式订单分析、启动次数、启动时间分布及用户反馈异常等方面进行了核查。经核查，不存在用户注册时间异常、用户行为异常的情形。

(3) 应收账款和营业收入变动具备合理性，第四季度收入占比较高符合公司的业务实际情况

1) 应收账款与营业收入具有匹配性

报告期内，发行人应收账款与营业收入的匹配关系如下：

单位：万元

| 项目 | 2022 年度 | 2021 年度 | 2020 年度 |
|-------------|-----------|-----------|-----------|
| 营业收入 | 98,846.18 | 80,578.16 | 57,824.64 |
| 应收账款账面余额 | 10,254.73 | 7,306.11 | 6,383.29 |
| 应收账款占营业收入比例 | 10.37% | 9.07% | 11.04% |

由上表可知，公司各期末应收账款余额与营业收入规模相匹配。报告期内，保荐机构和申报会计师针对应收账款执行了相应的核查程序：

①获取发行人报告期各期末应收账款明细表，分析应收账款变动的原因及应收账款与营业收入规模的关系，判断应收账款变动的合理性；

②对报告期内主要的客户执行函证、走访程序，针对未回函的客户全部执行替代测试，获取并检查相关原始单据；

③获取公司报告期各期末的应收账款期后回款明细表，测试重要客户的收入回款记录，核实公司应收账款余额期后回款情况。

2) 第四季度收入占比较高具有合理性

公司第四季度收入较高主要系由于 B 端收入的影响，C 端收入在第四季度的占比相对平稳。公司 B 端收入在第四季度占比较高，主要原因为：①公司 B 端智能文字识别和商业大数据的主要客户集中在金融、信息科技服务等领域大型企业，上述客户通常会制定年度采购预算，如年末剩余预算较多，会在年末集中进行采购；②各电商平台每年促销活动多集中在年末举行，因此公司 B 端广告推广业务收入在 12 月大幅增加。

C 端 APP 业务方面，2020-2022 年的付费用户数及平均付费金额整体处于上升趋势，因而分季度来看，各个季度占比整体呈逐渐提高的状况。2020-2022 年，公司 C 端业务第四季度收入占全年收入的比例分别为 29.98%、27.38% 和 27.62%。

总体来看，报告期各期，B 端和 C 端收入的第四季度按月收入分布相对平均。

(4) 公司的货币资金与交易性金融资产真实、准确、完整，不存在虚增情况

保荐机构与申报会计师通过执行银行函证、查阅理财产品购买记录与相关凭证、取得第三方平台账单，与企业 C 端业务收入摊销表进行核对，并结合前述其他财务核查手段，确认公司货币资金与交易性金融资产真实、准确、完整，不存在虚增情况。

综上，保荐机构、申报会计师认为：通过采取 IT 审计、资金流水核查、其他财务核查手段等方式，并结合财务指标分析，发行人 C 端收入增长的驱动因素真实合理，C 端收入真实、准确、完整。

(5) 公司 C 端 APP 各渠道月活用户无法去重的原因具有合理性，已对无法去重的原因和各渠道月活进行了充分的披露

公司已在招股说明书中对月活各渠道的统计口径、具体数据以及直接相加计算的过程进行了详细披露，并将各渠道月活无法去重的原因披露在二轮问询函回复之“4.1 关于月活用户”之“(二) 各渠道月活无法去重的原因，将各渠道月活直接相加计算的准确性和合理性、是否存在误导性陈述，在无法去重计算的情况下有无其他替代指标或处理方式”，核心内容总结如下：

截至 2022 年 12 月底，公司扫描全能王、名片全能王、启信宝 3 款 APP 产品 2022 年 12 月的各渠道月活合计超过 1.3 亿（各渠道月活为直接相加、无法去重，即在多个渠道活跃的每一用户可能被重复计算），其中 APP 渠道的月活合计超过 1.1 亿，小程序、Web 端等其他渠道月活合计约 0.2 亿。

公司各渠道无法去重主要是由于对不同渠道的月活进行去重操作的前提是：可以通过用户唯一标识对相同的活跃用户进行识别。然而，公司仅能够从微信小程序管理后台、百度小程序管理后台、百度统计平台获取月活数量的结果，但上

述平台对外提供月活数据时仅提供月活总数，因此公司无法从上述渠道获取带有用户唯一标识的明细数据，导致公司无法通过用户唯一标识对公司统计的各渠道月活进行去重操作，因此将各渠道月活直接相加计算。

保荐机构针对报告期内公司 C 端 APP 各渠道月活数据真实性执行的核查程序包括：1) 对于来源于公司内部后台的大数据系统记录的活跃用户数量，执行了以下程序：对大数据系统执行了信息系统一般控制测试，确认系统运行有效，对大数据系统执行了信息系统应用控制测试，确认系统可以准确记录活跃用户数量；2) 对于来源于第三方渠道记录的活跃用户数量，执行了以下程序：取得外部第三方（微信小程序、友盟等）渠道统计的活跃用户数量截图；3) 对报告期内各渠道月活数据变动的合理性进行分析。

经核查，保荐机构认为：报告期内，公司 C 端 APP 各渠道的月活数据真实，公司 C 端 APP 各渠道月活用户无法去重具有合理性，公司已对无法去重的原因和各渠道月活进行了充分的披露，不存在误导性陈述；公司的 C 端收入真实、准确、完整，不存在虚增月活数量以配合营收虚增的情况。

(十) 私募基金股东的出资人委托持股问题

1、媒体质疑情况

媒体关注到发行人的私募基金股东——常州鼎仕的股东层面存在委托持股情况。相关媒体从中国裁判文书网上发现，上海市徐汇区人民法院于 2021 年 8 月对叶跃民与石仁燕的借款合同纠纷做出了一审民事判决，判决书中提及石仁燕在 2015 年 4 月与叶跃民签订协议，委托叶跃民持有合合信息股权，投资金额为 250 万元。

2、核查情况

保荐机构、发行人律师就该问题的核查情况如下：

叶跃民为常州鼎仕的有限合伙人，根据 2015 年 4 月叶跃民与石仁燕签署的协议，叶跃民受石仁燕的委托代其间接持有合合信息的股权，即石仁燕向叶跃民支付人民币 250 万元用于认购常州鼎仕 250 万元的财产份额，对应当时持有合合信息 0.21% 的股权。因叶跃民涉及民事诉讼，导致其所持有的常州鼎仕财产份额

被司法冻结，截至本次审核中心意见落实函回复签署之日，该等份额尚处于被冻结状态。

2023年3月10日，叶跃民与石仁燕签署解除委托持股关系的《协议书》；根据《协议书》，叶跃民于2023年3月16日向石仁燕支付了股权（份额）转让款，双方委托持股关系解除。

根据对叶跃民的访谈，就上述委托持股、财产份额转让及委托持股解除事宜，叶跃民与石仁燕、常州鼎仕以及常州鼎仕的合伙人之间不存在纠纷或潜在纠纷。

经核查，保荐机构、发行人律师认为：上述委托持股事宜仅涉及常州鼎仕层面的相关财产份额，不涉及发行人直接股东的相关股权。截至本次审核中心意见落实函回复签署之日，叶跃民与石仁燕已解除委托持股关系，且不存在纠纷或潜在纠纷。该等事宜不会对本次发行上市造成法律障碍。

（十一）扫描全能王 APP 自动订阅问题

1、媒体质疑情况

部分媒体报道了扫描全能王自动订阅模式下扣除会员费遭遇用户投诉的情况，并质疑作为美化业务数据的手段。

2、核查情况

保荐机构、发行人律师就该问题的核查情况如下：

（1）报告期初至 2021 年 3 月：扫描全能王 APP 已在支付界面明确提示用户自动订阅相关内容

以连续包年套餐为例，用户点击扫描全能王 APP “升级到高级账户”、高级账户支付界面均被提示“购买连续包年高级账户的账号，会在下一年订阅周期到期前 24 小时自动从微信账号扣费并延长一年高级账户有效期”，界面明确连续包年套餐中首年及次年的高级账户价格。

用户可在扫描全能王 APP 内自主取消自动订阅功能，也可以通过微信、支付宝、APP Store 等原支付渠道、或是点击短信链接的方式取消自动订阅。此外，用户在高级账户有效期到期前 24 小时以前可以随时取消订阅，取消后不再扣费，

若用户不慎被扣费，也可通过联系扫描全能王 APP 人工客服进行退费处理。

(2) 2021 年 3 月至今：公司根据《网络交易监督管理办法》的要求对扫描全能王 APP 自动订阅服务进行了调整

2021 年 3 月 15 日，国家市场监督管理总局公布了《网络交易监督管理办法》。其中，第十八条规定：网络交易经营者采取自动展期、自动续费等方式提供服务的，应当在消费者接受服务前和自动展期、自动续费等日期前五日，以显著方式提请消费者注意，由消费者自主选择；在服务期间内，应当为消费者提供显著、简便的随时取消或者变更的选项，并不得收取不合理费用。

因此，公司对扫描全能王 APP 的自动订阅功能作了进一步调整，在高级账户自动订阅服务截止日期前五日，公司会以手机短信、APP 内消息提醒等方式向用户发送自动订阅提示。用户在选择订阅高级账户时，APP 在订阅选项下方及付款确认界面均展示《扫描全能王自动续费协议》（“《续费协议》”）并在《续费协议》中约定了前述内容。用户在到期前 24 小时以前可以通过 APP 首页“我的”二级菜单“订阅管理”中随时取消订阅，取消后不再扣费。用户确认并同意《续费协议》后，才可以使⽤订阅服务。

经核查，保荐机构、发行人律师认为：扫描全能王 APP 自动订阅及取消方式符合《网络交易监督管理办法》关于自动订阅的相关规定，不存在侵犯用户合法权益情形。截至本次审核中心意见落实函回复签署之日，发行人及其子公司不存在任何因相关 APP 产品自动续费问题而受到主管机关处罚的情形。

发行人已在招股说明书“第三节 风险因素”之“一、与发行人相关的风险”之“（四）经营风险”之“5、公司 C 端产品被消费者投诉的风险”中进行了提示。

针对发行人收入真实性，发行人已在一轮问询函回复之“9.关于收入”之“9.2 关于 C 端收入”之“二、（3）对 C 端用户和收入的真实性发表明确核查意见”进行了核查并发表了明确意见：发行人 C 端用户是真实的，C 端收入不存在跨期情形，C 端收入真实、准确、完整。

（十二）本次募集资金规模的必要性

1、媒体质疑情况

部分媒体报道质疑发行人募投项目的必要性，发行人拟募集资金 149,023.18 万元，其中铺底流动资金合计 34,763.84 万元；报告期末发行人货币资金 55,284.88 万元。

2、核查情况

保荐机构就该问题的核查情况如下：

发行人已在一轮问询函回复之“20.关于募投项目”中对本次募集资金规模的必要性和合理性进行了详尽的说明，相关核心内容总结如下：

（1）公司未来具有较大的资金需求

一方面，公司属于 AI 及大数据科技企业，全球 AI 及大数据的技术和产品呈现日新月异的发展态势，新技术、新产品、新模式的更迭速度较快，公司需紧跟行业尖端技术的发展趋势，加强核心技术优势，结合客户的需求变化，进行技术和产品的持续创新更迭，为保持公司在行业内的技术领先地位，公司持续投入大量资金用于智能文字识别及商业大数据领域的核心技术研发，报告期内发行人的研发费用分别为 16,696.80 万元、23,829.24 万元、27,990.16 万元。其次，为持续开拓新客户、提升市场份额，报告期内公司广告宣传费分别为 9,169.49 万元、16,414.25 万元、15,164.64 万元；销售人员薪酬分别为 5,292.82 万元、8,182.44 万元和 9,767.70 万元。

截至 2022 年 12 月 31 日，发行人货币资金余额为 55,284.88 万元，主要为银行存款和其他货币资金。公司具有存量研发项目后续支出、新增研发项目投入、推广费用支出等业务经营所必需的各类资金需求，资金需求量较大。募投项目“人工智能 C 端产品及 B 端服务研发升级项目”、“商业大数据 C 端产品及 B 端服务研发升级项目”、“人工智能核心技术研发升级项目”、“商业大数据核心技术研发与数据中台建设项目”中涉及的研发投入合计约 7.28 亿元，募投项目中涉及的推广费用支出合计约 3.02 亿元（含销售人员薪资 1.40 亿元，广告营销费用 1.62 亿元），募投项目的建设期为 3 年。

随着公司业务规模的发展壮大，公司后续运营资金需求预计将进一步扩大，公司留存自有货币资金已经无法满足本次募投项目建设的资金需求。

(2) 本次募投项目对公司发展具有重要意义

本次募集资金投资项目符合国家相关的产业规划以及公司未来整体战略发展方向，具有良好的市场发展前景和经济效益。具体如下：

人工智能 C 端产品及 B 端服务研发升级项目是对公司目前主营业务中的智能文字识别 C 端产品和 B 端服务的研发升级及优化，从技术研发升级和营销网络建设两方面加强对智能文字识别业务发展的支撑。

商业大数据 C 端产品及 B 端服务研发升级项目是对公司目前主营业务中的 C 端启信宝 APP 和 B 端商业大数据服务的研发升级及优化，从技术研发升级和营销网络建设两方面加强对商业大数据业务发展的支撑。

人工智能核心技术研发升级项目是对公司目前的智能文字识别方面核心技术的进一步深化、升级和整合，通过更加模型化、平台化的技术开发，支撑公司智能文字识别 C 端产品和 B 端服务等业务发展。

商业大数据核心技术研发与数据中台建设项目包括两方面，一方面，对公司目前的知识图谱、大数据挖掘等核心技术的进一步深化、升级和整合，另一方面，对基础数据、知识数据及智能数据模块进行数据源扩充与功能优化，进一步加强公司的数据中台建设，提升数据覆盖广度与丰富度。

综上所述，募投项目的实施有利于公司人工智能、商业大数据产品和服务功能性的提高、应用场景的丰富，促进公司营收规模的进一步扩大。同时，本次募投项目能够提高公司研发能力、强化公司产品与服务的市场竞争力，对公司的长期发展具有必要性。

经核查，保荐机构认为：发行人本次募集资金规模具有必要性和合理性。

(十三) 股权激励在 2019 年集中支付的合理性

1、媒体质疑情况

部分媒体报道质疑公司于 2019 年在出现 3,000 万元亏损的情况下，仍然集

中实施大额股权激励，此举致使当期亏损达 1.96 亿元，因此公司被怀疑通过发放股权激励进行“突击分红”，一方面让原股东获得最大收益，另一方面起到一定的避税目的。

2、核查情况

保荐机构就该问题的核查情况如下：

公司为建立健全激励机制，使关键管理人员和核心技术人员等员工可以分享到公司经营、发展带来的利益，进一步保持公司的竞争优势，实现公司的持续发展，发行人对其董事、监事、高级管理人员和核心技术人员等员工实施了员工股权激励方案。2014 至 2015 年期间，公司实施第一批股权激励，2017 至 2019 年期间，公司实施第二批股权激励，通过股权激励的方式实现员工利益和公司长远利益的有机结合。

发行人 2019 年股份支付由三部分构成，具体如下：

(1) 2017-2018 年已授予但尚未行权的期权加速行权

2017-2018 年，公司与李仕益等 42 名员工约定，分四期等量授予期权，每期间隔一年。发行人在 2019 年修改期权计划，对已授予的尚未行权的期权立即行权，根据《企业会计准则第 11 号——股份支付》相关规定，尚未行权的期权股份支付成本一次性计入 2019 年度当期损益。2017 年度已授予尚未确认的股份支付成本为 196.15 万元，2018 年度已授予尚未确认的股份支付成本为 969.29 万元，因此发行人在 2017-2018 年已授予的但尚未行权的期权加速行权确认的股份支付金额为 1,165.44 万元。

(2) 在 2019 年期权计划完成前进行第二期、三期授予的期权

根据 2017-2018 年已发放期权证书分四期授予的期权约定，在 2019 年进行第二期或第三期期权授予，根据企业会计准则的规定，在授予日发行人无需进行会计处理；在等待期内的每个资产负债表日，以对可行权权益工具数量的最佳估计数为基础，按权益工具在授予日的公允价值将取得职工或其他方提供的服务计入成本费用。2019 年期权计划修改，发行人已授予的尚未行权的期权均立即行权，根据《企业会计准则第 11 号——股份支付》相关规定，尚未行权的期权股

份支付成本一次性计入 2019 年度当期损益。2019 年确认的股份支付 724 万元。

(3) 2019 年提前授予并加速行权的期权以及 2019 年新增授予立即可行权的期权

2017-2018 年发放期权证书,原约定一部分期权将在 T+2 年和 T+3 年的时点授予,但公司对期权计划进行修改,在 2019 年提前授予并加速行权的期权,本次修订属于股份支付协议缩短等待期的有利修改,应在 2019 年一次性确认股份支付费用。对于 2019 年新增的授予后立即可行权的期权,根据《企业会计准则》的规定,对于授予后立即可行权的换取职工提供服务的权益结算的股份支付,应在授予日按照权益工具的公允价值,将取得的服务计入相关资产成本或当期费用,同时计入资本公积中的股本溢价,并确认为非经常性损益。扣除等待期内离职的激励对象对应的期权取消应调减的股份支付费用,2019 年确认的股份支付费用为 14,710.77 万元。

经核查,保荐机构、申报会计师认为:2019 年以后发行人不存在股权激励的情况。发行人根据业务发展客观需要,为吸引和留住优秀人才,充分调动发行人核心员工的积极性实施股权激励计划,不属于“突击分红”,不存在避税的目的,发行人历史上亦未进行任何分红。发行人严格按照会计准则的要求确认股份支付的费用。

(十四) 2018 年-2020 年前五大客户的占比下降

1、媒体质疑情况

部分媒体报道质疑 2018 年-2020 年之间公司前五大客户的销售收入占比分别为 20.43%、9.06%、6.71%,呈下降趋势,其中尤其是三星的收入占比快速下降,在 2020 年已不在前五大客户名单。

2、核查情况

保荐机构和申报会计师就该问题的核查情况如下:

(1) 2018 年-2020 年前五大客户的占比下降具有合理性

2020 年-2022 年,发行人前五大客户的收入占比分别为 6.71%、8.43%和 7.54%,

占比保持基本稳定。

2018年-2020年发行人前五大客户集中度总体呈下降趋势，主要由于：一方面公司C端收入相比B端增速更快，C端收入占比不断提高。2018年-2020年，公司C端收入占营业收入的比例分别为47.41%、63.93%、73.24%，而C端客户又相对零散，公司的前五大客户皆为B端的中大型企业客户。另一方面，主要的B端客户中，公司对三星的手机厂商技术授权业务销售金额在2018年-2020年期间呈显著下降趋势。

发行人已在招股说明书“第五节 业务和技术”之“三、公司销售情况及主要客户”之“（二）报告期前五大客户情况”中披露了报告期内公司前五大客户的销售收入占比变化的原因。

（2）2018年-2020年公司来自三星收入下降的原因

针对发行人来自三星的收入下降，且三星自2020年已不在前五大客户的原因，发行人已在一轮问询函回复之“5.关于业务”之“5.1 关于B端业务、二、（五）1、报告期内手机技术授权业务大幅萎缩的原因”中进行了详尽的说明，相关核心内容总结如下：

2018年-2020年，发行人来自三星的手机厂商技术授权业务收入逐年降低，分别为1,864.49万元、591.84万元、68.70万元，主要系由于2017年底开始三星逐步开展自主研发文字识别技术，在其手机应用系统中内嵌了基础的文字识别功能，2018年发行人与三星的合同正常到期，未进行续约。2017年底公司观察到手机厂商技术授权业务的变化，在资金和人力资源有限的情况下，公司决定围绕自身核心优势业务，将更多的精力投入到智能文字识别B端业务的技术迭代和客户拓展，降低手机技术授权业务在公司业务矩阵中的重要性水平。2020年至2022年，公司智能文字识别B端业务收入从5,412.40万元增长至7,120.43万元，年复合增长率为15%。

发行人已在招股说明书“第三节 风险因素”之“二、与行业相关的风险”之“（一）市场竞争加剧的风险”中对手机厂商拓展手机自带的扫描功能对发行人业务发展的影响进行风险提示。

综上所述，经核查，保荐机构、申报会计师认为：2018年-2020年公司前五大客户的销售收入占比呈下降趋势、三星在2020年已不在前五大客户名单的原因具有合理性，2020年-2022年公司前五大客户的销售收入占比保持基本稳定。

（十五）2019年存在未弥补亏损问题

1、媒体质疑情况

部分媒体就发行人2019年曾存在的亏损情况质疑其经营可持续性，发行人在2020年与2021年扣非后归属于母公司股东的净利润分别为盈利9,655.71万元、盈利13,917.88万元，2019年公司扣非后归属于母公司股东的净利润为亏损5,195.41万元。

2、核查情况

保荐机构和申报会计师就该问题的核查情况如下：

截至2022年12月31日，公司经审计的合并报表未分配利润为35,401.37万元，公司已不存在累计未弥补亏损。公司报告期内经营业绩持续向好，净利润分别为12,952.93万元、14,448.06万元和28,371.77万元。

经核查，保荐机构、申报会计师认为：报告期内，公司不存在因亏损导致可持续经营能力存疑的问题。

（十六）2018年度至2021年1-3月期间费用率较高

1、媒体质疑情况

部分媒体质疑公司2018年度至2021年1-3月的期间费用率较高，分别为120.94%、139.21%、64.17%和66.34%，2018年度和2019年度尤为偏高，均在130%左右，导致公司2018年度和2019年度分别亏损7,871万元和19,629万元，并进一步导致公司截至2021年3月31日的未分配利润为负。2018年度至2021年1-3月期间费用率较高，短期内持续经营能力是否面临挑战。

2、核查情况

保荐机构和申报会计师就该问题的核查情况如下：

2018-2019 年度，公司期间费用率较高，分别为 120.94%和 139.21%，剔除股份支付影响后分别为 120.29%、92.76%，同行业可比公司金山办公和福昕软件剔除股份支付影响后的期间费用率平均值分别为 70.45%、70.40%。公司的 C 端产品扫描全能王的付费功能主要在 2017 年至 2018 年陆续上线，营业收入逐步实现快速增长，因此 2018-2019 年度剔除股份支付影响后的期间费用率仍偏高主要系营业收入相对较低，分别为 19,603.14 万元和 34,038.94 万元。

2020-2022 年度，公司主营业务收入分别为 57,170.01 万元、80,469.11 万元和 98,613.18 万元，2021 年度与 2022 年度分别同比增长 40.75%和 22.55%。随着公司营业收入的快速增长，尽管公司的期间费用仍保持增长，分别为 37,107.77 万元、56,024.11 万元和 58,081.14 万元，但期间费用率较 2018 年度与 2019 年度已显著下降，分别为 64.17%、69.53%和 58.76%。

2020-2022 年度，公司已实现持续盈利，毛利率较高并保持稳定，分别为 84.33%、83.44%和 83.70%，净利润分别为 12,952.93 万元、14,448.06 万元和 28,371.77 万元。截至 2022 年 12 月 31 日，公司经审计的合并报表未分配利润为 35,401.37 万元，公司不存在累计未弥补亏损。

经核查，保荐机构、申报会计师认为：公司 2018 年度和 2019 年度期间费用率偏高主要系股份支付金额较大、收入基数相对较低影响。2020 年度开始，公司期间费用率保持稳定。报告期内公司营业收入保持持续稳定增长趋势，毛利率较高，盈利情况持续向好，具备稳健的持续经营能力。

三、核查程序及意见

针对媒体质疑情况自查，保荐机构、发行人律师、申报会计师执行了以下主要核查程序：

（一）持续关注有关媒体报道，通过公开网络检索的方式，对媒体关于发行人的报道进行全面完整的检索，全文阅读相关文章，分析是否涉及媒体质疑的情形。

（二）查阅发行人的招股说明书及相关问询函回复意见，核查相关媒体报道中提及的内容。

发行人已在招股说明书“第三节 风险因素”之“三、其他风险”之“(四) 舆情相关风险”中进行了特别提示。

经核查，保荐机构、发行人律师、申报会计师认为：相关媒体报道中不存在针对发行人本次发行上市相关的重大媒体质疑，不会对发行人本次发行上市构成障碍。针对媒体报道中关注的主要事项，发行人已在本回复文件、招股说明书及其他披露文件中进行了充分披露和说明。

（此页无正文，为《关于上海合合信息科技股份有限公司首次公开发行股票并在科创板上市申请文件的审核中心意见落实函的回复》之签章页）

上海合合信息科技股份有限公司

2023年7月24日



发行人董事长声明

本人已认真阅读上海合合信息科技股份有限公司本次审核中心意见落实函回复的全部内容，确认本审核中心意见落实函回复的内容真实、准确、完整，不存在虚假记载、误导性陈述或重大遗漏，并对上述文件的真实性、准确性、完整性、及时性承担相应法律责任。

董事长、法定代表人：


镇立新

上海合合信息科技股份有限公司

2023年 7月 24日

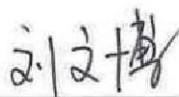


（此页无正文，为《关于上海合合信息科技股份有限公司首次公开发行股票并在科创板上市申请文件的审核中心意见落实函的回复》之签章页）

保荐代表人：



冷小茂



刘文博



保荐人法定代表人声明

本人已认真阅读上海合合信息科技股份有限公司本次审核中心意见落实函回复的全部内容，了解报告涉及问题的核查过程、本公司的内核和风险控制流程，确认本公司按照勤勉尽责原则履行核查程序，审核中心意见落实函回复不存在虚假记载、误导性陈述或重大遗漏，并对上述文件的真实性、准确性、完整性、及时性承担相应法律责任。

法定代表人：_____



沈如军

