



关于北京海天瑞声科技股份有限公司
首次公开发行股票并在科创板上市申请文件
第二轮审核问询函的回复

保荐机构（主承销商）



（深圳市前海深港合作区南山街道桂湾五路 128 号前海深港基金小镇 B7 栋 401）

上海证券交易所：

根据贵所于 2020 年 9 月 15 日出具的上证科审（审核）〔2020〕693 号《关于北京海天瑞声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问询函》（以下简称“问询函”）的要求，华泰联合证券有限责任公司（以下简称“华泰联合”或“保荐机构”）作为北京海天瑞声科技股份有限公司（以下简称“海天瑞声”、“发行人”或“公司”）首次公开发行股票并在科创板上市的保荐机构（主承销商），会同发行人及发行人律师北京市天元律师事务所（以下简称“天元律师”）和申报会计师毕马威华振会计师事务所（特殊普通合伙）（以下简称“毕马威会计师”）等相关各方，本着勤勉尽责、诚实守信的原则，就问询函所提问题逐项进行认真讨论、核查与落实，并逐项进行了回复说明。具体回复内容附后。

说明：

1、如无特别说明，本回复中使用的简称或名词释义与《北京海天瑞声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（申报稿）》（以下简称“招股说明书”）一致。涉及招股说明书补充披露或修改的内容已在招股说明书中以**楷体加粗**方式列示。

2、本回复中若出现总计数尾与所列值和不符的情况，均为四舍五入所致。

3、本回复中涉及的我国、我国经济以及行业的事实、预测和统计，包括本公司的市场份额等信息，来源于一般认为可靠的各种公开信息渠道。本公司从上述来源转载或摘录信息时，已保持了合理的谨慎，但是由于编制方法可能存在潜在偏差，或市场管理存在差异，或基于其它原因，此等信息可能与国内或国外所编制的其他资料不一致。

目 录

1、关于一体化数据处理平台.....	3
2、关于数据服务提供商.....	44
3、关于收入确认政策.....	59
4、关于非直接销售.....	75
5、关于软件产品增值税即征即退.....	83
6、关于收入季节性.....	87
7、关于整体变更时个人所得税事项.....	99
8、关于函证程序.....	102
9、保荐机构在充分核查基础上的总体意见.....	112

1、关于一体化数据处理平台

根据问询回复，发行人业务主要通过一体化数据处理平台开展生产经营。

请发行人说明：（1）一体化数据处理平台各功能模块的具体构成、具体功能和用途以及与发行人生产经营的具体关系、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性；（2）一体化数据处理平台账号注册相关的内部控制程序及有效性，注册账号数量与报告期各期服务采购量和服务人次的匹配关系；（3）是否存在使用客户提供的设备和信息系统开展业务的情形，说明具体的业务内容、涉及的销售收入及其占比。

请保荐机构和申报会计师对上述事项进行核查并发表明确意见，同时说明：（1）注册账号的真实性与变动的合理性、用户行为核查（账号的地域分布与数量、活跃账号数量、单次登录时长与登录时点、登录频率、登录 IP 或 MAC 地址信息等）以及核查的具体范围及比例；（2）信息系统审计的具体范围及完整性，是否存在核查范围受限、历史数据缺失和信息系统缺陷等情形，对信息系统可靠性的核查程序、核查比例、核查过程和核查结论。

答复：

发行人说明事项

（1）一体化数据处理平台各功能模块的具体构成、具体功能和用途以及与发行人生产经营的具体关系、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性；

一、一体化数据处理平台各功能模块的具体构成、具体功能和用途以及与发行人生产经营的具体关系

一体化数据处理平台是发行人的生产系统，嵌入了训练数据生产过程中所需的各类工具、软件模块，主要由发行人的业务部门使用，主要功能包含训练数据生产相关的采集、加工（筛选、转写、各类型标注）、质检等细分环节，具备流程管理、质量管理、数据安全功能。

一体化数据处理平台初始上线时间为2014年9月，上线初期主要集成了各数据处理工具并逐步拓展完善，另在2019年9月陆续上线增加了终端人管理系统及日志管理系统两大功能模块。

平台上线初期至报告期初，发行人对一体化数据处理平台的定位是：可用于生产符合要求的训练数据集，因此着重于打造该平台的生产功能，未对该平台的日志留存等进行强制规定，在2017年未保存生产过程中的日志信息；自2018年起开始保留该平台记载的基本日志信息，如按项目维度的采集/标注的原料数据量（音频文件个数、音频时长、文件大小）等；2019年9月在该平台上线终端人管理系统和日志管理系统后，进一步完善了生产日志信息留存，如生产过程的日志信息、终端人注册信息等。虽然一体化数据处理平台保留的日志信息不能完整覆盖发行人全部报告期，但2017年1月上线运行的业务管理平台（发行人的中台业务信息流转与管理系统）中完整记载留存了采购流程中的关键业务数据，包括每个项目数据服务采购涉及的采集量、标注数据量、标注工作时长、形成的成品数据集信息等，且需要通过发行人内部控制体系中设定的复核机制验证准确后才会进行相应的采购支付，使得发行人与财务报表相关的业务采购内部控制能够健全、有效执行。

经过报告期内持续改进、完善，一体化数据处理平台现由数据处理工具集、终端人管理系统和日志管理系统三大功能模块构成，除可满足训练数据的生产需要外，还可记录采购的数据服务对应的劳务工作量、产出量/数据量等信息，可在验收、决算阶段与发行人业务管理平台生成的财务相关信息进行交叉核对，进一步确保财务信息与生产、业务信息一致、准确。



图：一体化数据处理平台功能模块示意图

一体化数据处理平台各功能模块的具体构成、功能、用途及其与发行人生产经营的关系如下：

（一）模块1：数据处理工具集

该模块是一体化数据处理平台的核心部分，自2014年一体化数据处理平台上线至今持续完善、丰富，集成了发行人自主开发积累的、训练数据生产环节所需的各类数据处理工具。

1、构成：该模块集成了发行人数据采集、加工、质检环节所需的业务操作工具及软件包，包括采集环节使用的录音工具（如桌面录音工具、安卓端及iOS端手机录音工具）、视频采集工具等；加工环节使用的校对工具（如词性校对工具、韵律校对工具、语音与文本一致性校对工具等）、语音切分工具、图片标注工具、视频标注工具；质检工具，如项目内部质量检查工具等。

2、用途：用于训练数据生产流程的采集、加工、质检等各个环节，为其提供所需的软件工具和模块集合。

3、功能：主要功能为训练数据生产及生产相关的流程管理、数据处理、质量管理；可根据项目要求，依据终端人职能、角色为其分配不同的项目访问权限，并相应提供所需的采集/加工/质检等生产工具，并生成训练数据生产过程的日志信息。

发行人自2018年开始对训练数据生产过程中的基本日志信息¹进行保存，以智能语音类项目为例，保存的基本日志信息包括采集/标注人员的注册用户ID、按项目维度的采集/标注的原料数据量（各注册账号用户录制/标注的录音文档对应的音频时长、文件大小、录音语句条数）等。发行人于2019年下半年升级了带有详细日志功能的采集、标注工具，补充记录更多的生产过程信息，如终端人的注册信息、每个终端人的工作量信息等，并新增上线了专门的日志管理系统（具体参见下文“（三）模块3：日志管理系统”介绍），统一管理、存储前述详细日志信息²。

4、与生产经营的关系：服务于训练数据的具体生产过程，与一体化数据处理平台的生产系统定位相符，为发行人业务项目的具体执行、实施提供对应的工具和流程，便于生产过程的数据处理、质量把控，并对生产过程的数据进行记录和保存。

（二）模块2：终端人管理系统

该模块于2019年9月上线，是发行人对于管理人员、公司研发人员、供应商对接人员、终端采集/标注人员等的统一化信息管理模块，是对此前一体化数据处理平台终端人员基本信息保存功能的专门性完善、补充，进一步保证所有人员信息在系统内的完整、真实、可追溯。

1、构成：由终端劳务人员（即“终端人”）管理、终端人员授权文件管理、供应商管理、项目后台管理等子菜单构成，提供标准的人员注册、登录入口。

2、用途：建立统一化、规范化的人员信息库，用于记录终端人、供应商对接人等人员相关信息，并对前述信息进行统一管理；可为终端人提供唯一性标识，使数据处理工具集生成的日志信息能与终端人注册账号对应，实现后续终端人员信息及其生产日志信息的对应、追溯。

3、功能：具体包括终端采集/标注人员等人员的用户注册、注册信息校验、用户登录及工具使用权限管理、项目权限访问控制等；该系统可与数据处理工具集生

¹基本日志信息、详细日志信息具体存储内容、存储情况请参见下文“二、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性/（二）发行人主要运营数据及其保存情况”中具体介绍。

²基本日志信息、详细日志信息具体存储内容、存储情况请参见下文“二、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性/（二）发行人主要运营数据及其保存情况”中具体介绍。

成的日志信息相对接，实现对人员权限、信息、生产数据的统一管理、追溯；可提供终端人及供应商对接人等人员签字确认采集、标注工作量、工作时长情况的入口。

4、与生产经营的关系：服务于生产经营过程中的人员信息管理及生产过程追溯，可对供应商对接人、终端人等进行线上统一管理，同时与数据处理工具集、日志管理系统配合，可记录、留存更为全面的生产过程日志信息³，加强了劳务人员签字确认等外部证据留存以及生产过程数据与财务数据的复核机制，使得数据生产过程更加可追溯，提升了公司整体内部控制水平和经营管理效率。

（三）模块3：日志管理系统

该模块于2019年9月上线，是发行人生产日志的专门管理模块，进一步完善了发行人对于经营生产过程相关日志信息的保存度、留痕度，可做到生产日志信息的集中管理、保存、实时监控。

1、构成：主要包括日志的存储、分析与统计、输入与结果输出等子功能。

2、用途：主要用于全面记录并保存生产过程中的详细日志信息，实时记录劳务工作量等数据处理服务提供情况。

3、功能：针对各采集、标注工具开发了标准的日志记录模块，定义并执行统一的日志格式，可与终端人管理系统、数据处理工具对接，定期收集、定向整理各数据处理工具在服务器上生成的日志，记录终端人的注册、登录信息，以及注册账号参与数据处理的详细日志信息，具体包括被采集人员参与采集的次数、采集事项发生时间、单次采集数据量、累计采集数据量，各标注人员的操作工时、产出量等日志数据和用户行为日志等。可根据以上日志信息，形成后续统计报表、记录统计终端劳务人员提供的工作量情况，对接展示到终端人管理系统模块。

4、与生产经营的关系：该模块全面记录并保留了训练数据生产过程的详细日志信息（终端人注册、登录信息，各终端人的工作量信息、用户行为日志信息等），

³详细日志信息具体存储内容、存储情况请参见下文“二、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性/（二）发行人主要运营数据及其保存情况”中具体介绍。

并生成经营所需的报表，实现训练数据生产可追溯、与财务数据可复核，提升公司整体控制水平和经营管理效率。

二、主要运营数据的保存内容以及保存方式，报告期各期前述运营数据的保存情况及完整性

（一）发行人信息系统基本情况

发行人内部信息系统由三个组成部分构成，包含一体化数据处理平台（前台生产系统）、业务管理平台（中台业务信息流转与管理系统）与用友财务系统（后台记账与报表系统），其中一体化数据处理平台、业务管理平台涉及运营数据的生成或保存。



图：发行人信息系统概览

一体化数据处理平台：初始上线时间为2014年9月，并于2019年9月新增搭载终端人管理系统及日志管理系统两大模块。主要功能请参见上文“一体化数据处理平台各功能模块的具体构成、具体功能和用途以及与发行人生产经营的具体关系”相关介绍。

业务管理平台：于2017年1月上线使用，记载了发行人销售、采购业务相关信息，主要功能包含合同和项目创建、项目执行过程和结算、验收及决算管理，覆盖贯穿项目管理全生命周期，并可根据记录的业务数据生成收入、成本、存货、产品研发费用等业务相关的财务数据。

用友财务系统：于2015年开始上线使用，自上线至2019年4月为用友U8系统，2019年4月至今为用友NC系统。

（二）发行人主要运营数据及其保存情况

发行人生产经营相关的主要运营数据包括生产过程中的基本日志和详细日志信息，终端劳务人员注册信息以及与财务报表相关的核心运营数据等。

各类运营数据的具体保存内容、保存方式、保存该类信息的信息系统及具体的保存期间情况如下：

运营数据类别	保存内容	保存方式	保存数据的信息系统	数据保存覆盖期间
生产过程的日志信息	基本日志信息⁴ ：项目编号、参与项目的终端劳务人员 ID；按项目维度的采集的原料数据量（采集的原料数据文件个数、各数据文件容量大小、采集的各音频/视频文件时长、语音音频涉及的语句条数）；标注的原料数据量（标注的原料数据文件个数、数据文件大小、标注的音频/视频文件时长等） 详细日志信息 ：终端劳务人员的登录 IP、登录设备信息，各项目采集/标注过程的实时记录，终端人执行采集、标注工作的起止时间；终端劳务人员对采集/标注的工作量、工作时长情况的签字确认信息	电子文档形式在发行人服务器保存；同时有备份数据异地存放。	一体化数据处理平台	自 2018 年开始保存基本日志信息；自 2019 年 9 月开始保存基本日志信息+详细日志信息。
终端人注册信息	基本注册信息 ：用户 ID 详细注册信息 ：用户 ID、注册时间、注册 IP、注册设备信息、姓名、身份证号码（脱敏处理）、联系方式信息（脱敏处理）、注册国家等	电子文档形式在发行人服务器保存；同时有备份数据异地存放。	一体化数据处理平台	自 2018 年开始保存基本注册信息；自 2019 年 9 月开始保存基本注册信息+详细注册信息。
与财务报表相关的核心运营数据	项目合同信息、项目管理过程的预算及财务报销信息、项目实施过程的采购确认信息、项目交付验收关键节点过程信息、项目决算管理信息。 以上信息详细记录了一体化数据	电子文档形式在发行人服务器保存；同时有备份数据异地存放。	业务管理平台	自 2017 年起完整保存。

⁴ 以某方言语音识别训练数据项目为例，一体化数据处理平台留存的基本日志信息包含：参与录音及原料数据标注的终端劳务人员的注册账号、采集/标注的录音文件的个数、采集/标注的录音文件大小、采集/标注的录音内容中涉及的语句条数等。

	平台所能够产生的采集数据量、标注数据量、标注工作时长以及成品数据集信息等与财务报表相关的核心运营指标。			
--	---	--	--	--

（三）报告期各期运营数据的保存情况及完整性

报告期各期对运营数据的保存情况一览如下：

期间	是否保存各类运营数据				
	生产过程的日志信息		终端人注册信息		与财务报表相关的完整核心运营数据
	基本日志信息	详细日志信息	基本注册信息	详细注册信息	
2017年					√
2018年1月-2019年9月底前	√		√		√
2019年10月至今	√	√	√	√	√

一体化数据处理平台上线初期，发行人将其定位为能够生产符合质量要求的训练数据工具系统，着重于打造该平台的生产功能，因此未对该平台记录的生产日志进行强制留存，而是对生产日志采取滚动覆盖及定期清理措施，所以该系统中未保留2017年生产日志信息。

自2018年开始，发行人开始留存一体化数据处理平台中记载的与生产过程相关的基本日志信息，包括项目编号、参与项目的终端劳务人员ID、按项目维度的采集原料数据量、标注原料数据量等。

2019年9月，发行人在一体化数据处理平台中上线了终端人管理系统及日志管理系统，在原有基本日志信息留存的基础上增加详细日志信息留存，能够更为全面地记录并保留生产过程中的日志信息和终端劳务人员信息，包括终端劳务人员注册信息、参与的项目编号、登录IP、登录设备信息，各项目采集、标注过程的实时记录，以及终端人执行采集、标注的起止时间等详细日志信息、终端人对服务提供情况和服务量的签字确认信息，使得整个数据生产过程更加可追溯，信息记录及留存更加详细。

除一体化数据处理平台外，发行人2017年1月上线了业务管理平台，并在该平

台完整记载了报告期内与财务报表相关的核心运营指标，包括数据服务采购按项目维度的采集数据量、标注数据量、标注工作时长、形成的成品数据集信息等核心运营数据，能够确保发行人采购信息完整、准确记载。

（四）小结

对于一体化数据处理平台：2017年，发行人未留存生产过程的日志信息，发行人自2018年开始保留一体化数据处理平台中已记载的基本日志信息，并于2019年9月上线日志管理系统，在原有基本日志信息留存的基础上增加详细日志信息留存，能够更为全面地记录并保留生产过程的日志信息；同时，发行人于2019年9月上线终端人管理系统，完整记录终端劳务人员的个人信息、劳务量、签字确认信息等，使得整个数据生产过程更加可追溯，信息记录及留存更加详细。此外，发行人2017年1月上线的业务管理平台完整记载了报告期内与财务报表相关的核心运营指标，能够确保发行人采购信息完整、准确记载。

（2）一体化数据处理平台账号注册相关的内部控制程序及有效性，注册账号数量与报告期各期服务采购量和服务人次的匹配关系；

一、一体化数据处理平台账号注册相关的内部控制程序及有效性

参与采集、加工、质检等训练数据生产环节的终端劳务人员需在一体化数据处理平台注册账号。针对一体化数据处理平台账号注册，发行人采取的内部控制程序如下：

报告期初至2019年9月，终端劳务人员注册账号由一体化数据处理平台生成并分配至终端劳务人员使用。具体流程为：项目经理依据项目的数据处理需求，将项目的采集、加工、质检等任务按人员分配进行合理拆分，一体化数据处理平台根据前述任务拆分情况生成终端劳务人员注册账号，终端采集、标注人员从对接的项目组织人员（即服务公司人员、终端劳务人员的组织人员等）处领取前述账号并登录后，执行该账号对应的后续数据处理任务。其中，对于采集类数据服务，项目组织人员将查验采集人员填写的项目采集信息表，确认其填写的身份信息与其出示的证件一致后才向其发放注册账号。

自2019年9月开始，发行人在一体化数据处理平台上线了终端人管理系统，完善了人员信息的数据留存、注册认证，也完善了账号注册及后续信息保存相关的控制程序，具体如下：

1、注册信息完整性控制：在获得终端人员明确同意及授权后，终端人员填写信息进行注册，平台设置要求终端劳务人员在注册环节填写“电话号码”、“姓名”、“性别”和“生日”等通用的基本必填信息后方能完成注册；同时平台会根据终端劳务人员的国籍等具体属性不同，调整其他需要额外填写的“省份”、“城市”等信息要求。终端人管理系统上线初期，作为系统上线的过渡措施，存在发行人将留存的线下终端人信息导入线上，并要求终端人配合完成必要信息填写以完成注册的情形，此类情形同样要求完整填写必要信息，否则无法完成注册。

2、注册信息真实性控制：终端劳务人员在账号注册过程中，在“电话号码”字段下填写手机号后，一体化数据处理平台将自动发送即时验证码，并要求注册人填入准确的验证码，以确保手机号真实有效；终端劳务人员填写身份证号码后，一体化数据处理平台将根据设定的一般规则对其合理性进行检验，包括长度是否匹配、最后一位校验码是否合理等。终端人管理系统上线初期，作为系统上线的过渡措施，存在发行人将留存的线下终端人信息导入线上，并要求终端人配合完成必要信息填写以完成注册的情形，此类情形同样要求完整填写必要信息并完成验证程序，否则无法完成注册、进行下一步操作。

3、注册账号的权限控制：根据终端劳务人员提供的服务类型（采集、加工、质检等）不同，一体化数据处理平台将为其账号赋予不同权限，并提供与其服务类型对应的工具集供其使用。同时劳务人员仅能查看自己参与服务的项目信息。

4、注册账号信息的后续控制：一方面，一体化数据处理平台对于终端劳务人员注册信息进行有效保护；另一方面，一体化数据处理平台会记录终端人员登录使用的IP地址信息和设备信息，并同步生成终端劳务人员操作行为日志，使得整个数据生产过程更加可追溯，信息记录及留存更加详细。针对账号注册信息，发行人设置的保护措施具体如下：

(1) 数据脱敏和加密传输：发行人对收集的个人信息中的身份证号、联系方式等个人敏感信息等进行去标识化处理，并将去标识化后的个人身份信息与业务数据分开存储；采用加密等安全措施存储和传输数据信息；

(2) 权限、人员设置及内部审批：对被授权访问数据的内部人员设置仅能访问职责所需的数据信息的权限，对安全管理人员、数据操作人员的角色进行分离设置，对数据信息的重要操作设置内部审批流程，如批量修改、拷贝、下载等；

(3) 人员保密协议及培训：发行人与访问、操作数据的员工签署保密协议，对其进行数据安全培训。

综上，报告期初至2019年9月，发行人在终端劳务人员注册账号的申请、权限设立、账号分配等流程管理方面建立了一定的内部控制程序并在报告期内持续改善，2019年9月终端人管理系统上线后，劳务终端人员账号注册信息真实、完整，与账号注册相关的内部控制程序有效执行。

二、注册账号数量与报告期各期服务采购量和服务人次的匹配关系

注册账号数量与各期服务采购量、服务人次/人数的匹配关系如下：

项目	2020年1-9月	2019年度	2018年度
归属于当期执行项目的注册账号数量（万人）①	6.59	6.93	11.75
当期数据处理服务人数（万人）②	6.30	6.91	7.92
当期数据处理服务人次（万人）③	9.01	9.56	11.26
数据服务采购金额（万元）④	5,646.32	6,099.95	7,352.66
注册账号数量与服务人次/人数差异（万人）⑤ （2018年计算方式与2019年、2020年1-9月有所差异，具体参见下文“1、数据口径说明”）	0.29 (⑤=①-②)	0.02 (⑤=①-②)	0.49 (⑤=①-③)
差异率（⑤/①）	4.43%	0.22%	4.14%

注1：归属于当期执行项目的注册账号数量（①）为发行人信息系统记录的、当期提供服务的注册账号数量。

2、当期数据处理服务人数（②）、当期数据处理服务人次（③）分别为各期发行人已结算的数据服务采购对应的服务人数、人次。

3、2017年，发行人未在一体化数据处理平台中保存终端人注册信息，故此处未列示归属于2017年执行项目的注册账号数量。

1、数据口径说明

2018年列示的注册账号数量、差异率计算方式与2019年、2020年1-9月存在差异，具体说明如下：

2018年，发行人一体化数据处理平台记载留存的是终端人基本注册信息，注册账号由一体化数据处理平台根据各项目处理需求生成，并分配给不同的终端劳务人员使用。不同项目的注册账号之间不具备唯一性标识，存在一个终端劳务人员因参与不同项目而被分配多个注册账号的情形，因此该年记录的注册账号数量（归属当年执行项目的）未经去重处理，共计11.75万个。考虑数据可比性，该年差异率使用注册账号数量、当期数据服务采购人次（未去重）计算。

自2019年开始，发行人在内部管理流程中增加了终端人签字确认劳务工作量的环节，该确认环节能够获取终端劳务人员身份识别信息；2019年9月，发行人上线终端人管理系统，作为系统上线的过渡措施，发行人将2019年提供服务的线下终端人信息导入线上，导入过程中已根据终端人身份信息为其分配了唯一识别的账号；发行人上线终端人管理系统后，在终端人注册环节记录身份信息并进行手机号码验证，能够在注册环节为参与多个数据服务采购项目的终端人员提供唯一识别的注册账号。因此，2019年、2020年1-9月，发行人信息系统中记载的终端人注册账号数量均已经过去重处理。考虑数据可比性，该年差异率以注册账号数量、当期数据服务采购人数计算。

2、匹配情况分析说明

（1）注册账号数量与当期服务人次/人数基本匹配

2018年、2019年及2020年1-9月，各期注册账号数量与当期服务人次/人数基本匹配，各期数量略微存在差异，主要原因是：注册账号数量为发行人信息系统记录的、当期提供服务的注册账号数量，当期数据处理服务人数/人次为各期发行人已结算的数据服务采购对应的服务人次/人数，数据服务采购结算时间晚于服务提供时间，使得前述二者数量存在一定差异，各期差异率分别为4.14%、0.22%及**4.43%**，**差异较小**。

综上，各期注册账号数量与当期服务人次/人数基本匹配。

(2) 数据服务采购量与服务人次相匹配

考虑到数据服务采购量涉及智能语音、自然语言、计算机视觉等不同的应用领域，且不同应用领域业务采购量统计单位较多，此处以各期结算的数据服务费金额作为数据服务采购量替代指标进行匹配分析。2018年、2019年及2020年1-9月，各期数据服务采购金额与服务人次同向变动，趋势匹配。

综上，各期数据服务量与当期服务人次匹配。

3、结论

根据（1）中分析，各期注册账号数量与当期服务人次/人数基本匹配；根据（2）中分析，各期数据服务采购量与服务人次相匹配。由（1）、（2）可得，各期注册账号数量与当期数据服务采购量、服务人次匹配。

（3）是否存在使用客户提供的设备和信息系统开展业务的情形，说明具体的业务内容、涉及的销售收入及其占比。

发行人存在少量使用客户提供的设备和信息系统开展业务的情形，具体如下：

情形1-采集：部分情况下，客户会要求发行人使用其提供的设备采集数据。例如：个别客户的算法、模型拟应用于其自研设备（如智能音箱等）上，需进行系统适配，发行人会按照客户要求使用其提供的设备直接采集数据。

情形2-加工：部分客户对数据安全有特殊要求，发行人会应客户要求在其信息系统中直接执行数据标注。

发行人在智能语音、自然语言、计算机视觉各业务领域的训练数据定制服务均涉及少量使用客户提供的设备或信息系统开展业务的情形；但该类情形较少，占报告期营业收入较低。上述各情形的销售收入及占比情况具体列示如下：

单位：万元

项目	2020年1-9月		2019年		2018年		2017年	
	销售收入	占营业收入比	销售收入	占营业收入比	销售收入	占营业收入比	销售收入	占营业收入比

		重		重		重		重
情形1-采集 ⁵	260.45	1.82%	938.88	3.95%	476.50	2.47%	322.94	2.71%
情形2-加工 ⁶	954.62	6.69%	1,084.98	4.57%	747.87	3.88%	65.47	0.55%
使用客户提供的设备和信息系统开展业务的情形合计	1,215.07	8.51%	2,023.86	8.52%	1,224.36	6.36%	388.41	3.26%

注：情形1与情形2不存在重合，即不存在同时使用客户提供的设备及信息系统的情形。

在使用客户提供的设备或信息系统进行数据采集、加工的情形中，发行人依据销售合同（订单）开展采集、加工操作。

对于情形1-采集，发行人使用客户提供的设备采集原料数据，后续使用自有信息系统对前述数据进行加工。采集环节获得的数据即加工环节所需的原料数据，因此发行人信息系统记录有采集数据量、标注数据量信息。一方面，发行人基于自有信息系统记录的采集数据量、标注数据量，以及与供应商约定的标产比换算得到的标注有效工时，向供应商发送验收结算单、进行验收结算。另一方面，发行人也基于前述采集数据量、标注数据量信息，通过向客户项目负责人发送验收申请并得到客户邮件回复确认的方式，与客户进行验收结算。

对于情形2-加工，发行人使用客户提供的信息系统对客户提供的原料数据执行标注，客户通常会在服务期间内为发行人开放查询其信息系统相关信息的端口，发行人可获取客户信息系统记录的标注数据量、相关工作量等信息并在自有信息系统业务管理平台中记录留存。一方面，发行人基于前述标注数据量、相关工作量等信息，依照与供应商约定的标产比换算标注有效工时，向供应商发送验收结算单、进行验收结算。另一方面，发行人基于前述标注数据量、相关工作量等信息，通过向客户项目负责人发送验收申请并得到客户邮件回复确认的方式，与客户进行验收结算。

⁵ 发行人的训练数据定制服务可分为采集+加工服务和纯加工服务两类，仅采集+加工服务类别（发行人根据客户需求设计训练数据集结构、组织原料数据采集、对采集到的原料数据进行加工）涉及使用客户提供的设备开展业务的情形，且该类业务不涉及使用客户的信息系统。此处列示收入为使用客户设备进行采集的采集加工服务收入。

⁶ 发行人的训练数据定制服务可分为采集+加工服务和纯加工服务两类，仅纯加工服务类别（客户提供原料数据，发行人根据客户需求进行加工）涉及使用客户提供的信息系统开展业务的情形，且该类业务不涉及使用客户提供的设备。此处列示收入为使用客户信息系统的纯加工服务收入。

中介机构核查事项

(1) 核查程序

保荐机构、申报会计师主要执行了如下核查程序：

1、查看一体化数据处理平台实际运行情况，查阅该平台各功能模块的系统使用说明，了解一体化数据处理平台的具体功能、用途；访谈该平台的使用人员，了解各功能模块与发行人业务经营之间的关系；

2、访谈发行人信息系统主要负责人，了解发行人信息系统的建设及完善历程、各组成部分及功能定位，取得发行人信息系统导出的运营数据，了解、复核报告期内发行人运营数据的保存方式、保存情况；

3、登录一体化数据处理平台，查看信息系统配置情况，了解账号注册相关的流程、步骤及内部控制程序；针对账号注册过程执行穿行测试，复核验证信息系统可靠性、数据记录准确性；

4、执行信息系统公司层面控制核查（ITEC）、信息系统一般层面控制核查（ITGC）和信息系统应用控制核查（ITAC）程序，对报告期各期间发行人信息系统的内部控制及有效性情况进行核查；执行信息系统数据分析核查（ITDA）程序，对报告期内信息系统留存的业务数据及财务数据进行分析，复核验证数据的真实性、合理性；

5、取得各期注册账号数量、数据处理服务人数数据，分析其匹配关系；

6、访谈发行人主要管理人员，了解发行人使用客户提供的设备和信息系统开展业务的具体情况、业务模式及验收结算情况，查阅相关业务合同，取得相关收入金额统计数据并进行分析。抽查使用客户提供的设备和信息系统开展业务的项目，获取发行人信息系统中记录的采集数据量、标注数据量，与供应商的验收结算文件、业务合同约定、客户信息系统记录的标注数据量、相关工作量及客户验收结算邮件等信息进行交叉比对。

(2) 核查结论

经核查，保荐机构、申报会计师认为：

1、一体化数据处理平台的功能模块设计合理，可有效为发行人业务经营服务。在运营数据保存方面，2017年发行人未留存生产过程的日志信息；发行人自2018年开始保留一体化数据处理平台已记载的基本日志信息，并于2019年9月上线日志管理系统，在原有基本日志信息留存的基础上增加详细日志信息留存，能够更为全面地记录并保留生产过程的日志信息；同时，发行人于2019年9月上线终端人管理系统，完整记录终端劳务人员的个人信息、劳务量、签字确认信息等，使得整个数据生产过程更加可追溯，信息记录及留存更加详细。此外，发行人2017年1月上线的业务管理平台完整记载了报告期内与财务报表相关的核心运营指标，能够确保发行人采购信息完整、准确记载。

2、报告期初至2019年9月，发行人在终端劳务人员注册账号的申请、权限设立、账号分配等流程管理方面建立了一定的内部控制程序并在报告期内持续改善；2019年9月终端人管理系统上线后，劳务终端人员账号注册信息真实、完整，与账号注册相关的内部控制程序有效执行。各期发行人系统注册账号数量与当期数据服务采购量、服务人次相匹配。

3、发行人存在少量使用客户提供的设备和信息系统开展业务的情形，涉及的销售收入金额较小，占对应期间发行人营业收入比重较低。

中介机构说明事项

(1) 注册账号的真实性与变动的合理性、用户行为核查（账号的地域分布与数量、活跃账号数量、单次登录时长与登录时点、登录频率、登录 IP 或 MAC 地址信息等）；

与主要通过互联网提供产品或服务的To C类企业不同，发行人的一体化数据处理平台是内部生产所用的信息系统，核心职能为生产功能，不涉及产品和服务的线上销售、结算及交付。发行人在业务管理平台记载留存了完整的采购信息（采集数据量、标注数据量、标注工作时长、形成的成品数据集信息等），可确保与财务

报表相关的业务采购内部控制健全、有效执行。是否留存终端劳务人员的详细注册信息、生产过程的详细日志信息并不影响发行人与采购相关的财务核算及内部控制。

出于持续提升内部控制水平及经营管理效率，加强数据生产过程的可追溯性、与财务数据的可复核性的考虑，发行人于2019年9月在原有一体化数据处理平台基础上上线了终端人管理系统和日志管理系统，记录留存终端劳务人员的注册账号、工作量确认等信息，同时留存更为详细的生产过程日志信息。

保荐机构、申报会计师（以下合称“中介机构”）针对一体化数据处理平台注册账号的真实性、变动合理性及用户行为开展了核查，确认注册账号真实、数量变动合理，用户行为不存在异常。核查具体情况如下：

一、注册账号真实性、变动合理性核查

2019年9月，发行人在一体化数据处理平台中上线了终端人管理系统、日志管理系统，加强对终端劳务人员的账号注册管理、人员信息留存，以及对生产过程详细日志信息的记录留存。下述核查（一）、（二）、（四）均基于终端人管理系统、日志管理系统上线后记录留存的注册账号及详细日志信息进行；核查（三）基于2018年起记录留存的注册账号及全部日志信息进行。

（一）账号注册内部控制情况核查及穿行测试

中介机构查看了一体化数据处理平台的信息系统配置情况，复核该平台针对账号注册的控制程序⁷情况，并执行穿行测试，对系统可靠性、数据记录准确性加以复核验证：

中介机构使用随机选取的测试设备在一体化数据处理平台上自主注册了终端劳务人员账号（“中介机构测试账号”），并在该平台中创建了测试用项目，以中介机构测试账号参与了数据处理流程。随后，中介机构在一体化数据处理平台中查询了平台记录留存的、与中介机构测试账号相关的终端人注册信息（包括用户ID、

⁷ 详细注册程序可参见本问题“发行人说明事项/（2）一体化数据处理平台账号注册相关的内部控制程序及有效性”中相关表述。

注册时间、注册IP、注册设备、姓名、身份证号、联系方式、国家等)和生产过程日志信息(包括登录IP、登录设备信息、采集/标注的原料数据量、采集/标注的起止时间、工作时长等)。经查验,一体化数据处理平台记录情况与中介机构测试账号实际注册、操作记录相符。

经核查,一体化数据处理平台可准确记录、留存终端人注册信息、生产过程日志信息,与账号注册相关的内部控制程序有效执行。

(二) 注册信息合理性核查

中介机构取得了一体化数据处理平台记录留存的终端人注册信息,包括系统生成的用户ID、注册时间、注册IP、身份证号码、联系方式信息、注册国家等,并对注册相关必填信息的完整性、合理性进行了核查,具体如下:

1、身份必填信息完整性核查

中介机构对终端劳务人员注册一体化数据处理平台账号时的填写的身份信息进行了复核。经核查,各注册账号均按要求填写了所属类别所需的身份必填信息(如证件号码、联系方式等),不存在注册账号身份必填信息缺失的情形。

经核查,发行人系统中不存在注册账号身份必填信息缺失的情况,终端劳务人员注册账号时均填写了所需的身份必填信息,可验证注册账号真实。

2、账号注册信息合理性核查(手机号码合理性核查)

中介机构取得了中国籍终端劳务人员注册账号填写的手机号码信息,对其合理性进行了复核,主要核查手机号码的号段、号码位数是否合理,号码是否含有非数字字符等情形。

经核查,终端人管理系统上线运行初期,因系统验证功能尚在完善,存在4名中国籍终端劳务人员手机号码填写有误的情形,后续发行人已针对该问题进行了系统的升级完善,此后未再出现上述情形。除此之外无其他手机号码不合理情形,终端人注册账号填写的手机号码填写合理,可验证注册账号信息真实。

3、账号注册异常情况核查(同一手机号对应多个注册账号情形核查)

中介机构取得了中国籍终端劳务人员注册账号时填写的手机号码信息，对手机号码的集中情况进行了统计分析，核查是否以同一手机号注册多个账号的异常情况。

经核查，终端人注册账号不存在以同一手机号注册多个账号的异常情形。

综上，除终端人管理系统上线初期因系统功能尚在完善存在4名终端人手机号码填写有误的情形外，发行人不存在注册账号身份必填信息缺失、账号注册信息不合理、账号注册异常的情形，终端劳务人员的注册信息完整、具备合理性。

（三）注册账号变动合理性核查

中介机构取得了2018年-2020年1-9月内各期注册账号数量、数据服务采购金额、服务人数数据，对前述数据匹配性进行了分析，具体分析情况请参见本题“发行人说明事项/（2）一体化数据处理平台账号注册相关的内部控制程序及有效性，注册账号数量与报告期各期服务采购量和服务人次的匹配关系/二、注册账号数量与报告期各期服务采购量和服务人次的匹配关系”。

经核查，发行人系统注册账号数量与当期劳务服务人数、数据服务采购金额同向变动，趋势匹配。

（四）用户行为核查

中介机构取得了一体化数据处理平台中记录留存的生产过程日志信息（包括基本日志信息、详细日志信息），对终端劳务人员行为进行了核查分析，核查分析维度包括账号的地域分布、活跃账号数量、单次登录时长与登录时点、登录频率及登录IP指标，具体核查情况请参见下文“二、用户行为核查”。

经核查，终端劳务人员行为与发行人业务实际开展情况相符，不存在异常。

（五）核查结论

综上，中介机构通过多种核查手段（账号注册内部控制情况核查及穿行测试、注册信息合理性核查、注册账号变动合理性核查、用户行为核查）对一体化数据处理平台注册账号的真实性、变动合理性进行了充分核查，确认注册账号真实、变动合理，用户行为不存在异常。

二、用户行为核查

2019年9月，发行人在一体化数据处理平台中上线了终端人管理系统、日志管理系统，加强对终端劳务人员的账号注册管理、人员信息留存，以及对生产过程详细日志信息的记录留存。本部分用户行为核查均基于终端人管理系统、日志管理系统上线（2019年9月30日）之后记录留存的注册账号及详细日志信息，具体如下：

（一）账号的地域分布和数量情况核查

中介机构取得了2019年终端人管理系统上线后至2020年9月30日期间的注册账号（线上注册部分）⁸清单（共计**6.87**万个）及前述账号的注册IP，按IP地址对应的地域情况统计分析如下：

类别	终端人注册账号数量	占比
前五大分布地区		
北京	7,784	11.32%
河北	7,738	11.26%
广东	5,740	8.35%
河南	5,075	7.38%
山东	3,260	4.74%
前五大分布地区合计	29,597	43.06%
注册账号总数（线上注册部分）	68,735	100.00%

终端人账号的地域分布整体较为分散。前五大分布地区分别为北京、河北、广东、河南、山东，占比分别为**11.32%**、**11.26%**、**8.35%**、**7.38%**和**4.74%**，剩余地区占比均不高于5%，分布较为分散，符合发行人采购业务覆盖地区广的特征。经核查，上述前五大分布地区是发行人经营主体所在地区或智能语音类项目主要的发音采集地区，不存在异常。

（二）活跃账号数量情况核查

⁸ 终端人管理系统上线初期，作为系统上线的过渡措施，存在发行人将留存的线下终端人信息导入线上并要求终端人配合填写必要信息以完成注册的情形，系统将该类账号标记为“线下导入”，区别于线上注册账号。线下导入账号不涉及线上注册IP记录，因此本部分核查仅针对线上注册账号。

一体化数据处理平台是发行人的生产用信息系统，终端劳务人员注册账号并登录该系统是为了执行采集、加工、质检等处理操作。结合前述业务特点，中介机构将注册后曾登录该系统并执行过处理操作的账号定义为活跃账号。

经核查，2019年9月终端人管理系统上线以来至2020年9月30日期间，发行人活跃账号数量（共计**6.79**万个）占注册账号总数⁹（共计**12.41**万个）的比例为**54.67%**。前述活跃账号数量占比不高，主要原因是：2019年9月，发行人上线终端人管理系统，作为系统上线的过渡措施，发行人将2019年提供服务的线下终端人信息导入线上，根据终端人身份信息为其分配了唯一识别的账号（线下导入情形）。线下导入的注册账号中，较多终端劳务人员在前述注册账号分配前已完成其采集、标注任务，完成服务后未再登录系统产生操作行为，与发行人数据服务采购的一次性、临时性特点相符。

除上述线下导入情形以外，中介机构对线上注册账号的活跃情况也进行了核查分析：2019年9月终端人管理系统上线以来至2020年9月30日期间，注册账号总数（线上注册部分）共计**6.87**万个，其中活跃账号共计**5.65**万个，占比**82.13%**。活跃账号比例较高，表明该期间内注册的大部分终端劳务人员均曾使用一体化数据处理平台**进行处理操作，为发行人提供服务**。少数账号尚未有处理操作的原因为：发行人在执行项目较多，依据项目进度，少量终端劳务人员已完成账号注册，但还未**开始**提供数据处理服务。

综上，终端劳务人员注册账号活跃情况合理、不存在异常情况。

（三）单次登录时长与登录时点情况核查

中介机构取得了2019年终端人管理系统上线后至2020年9月30日期间的活跃账号清单（共计**6.79**万个），并取得了前述账号参与处理操作的详细日志信息。考虑到终端劳务人员登录一体化数据处理平台的主要目的是执行处理操作，中介机构

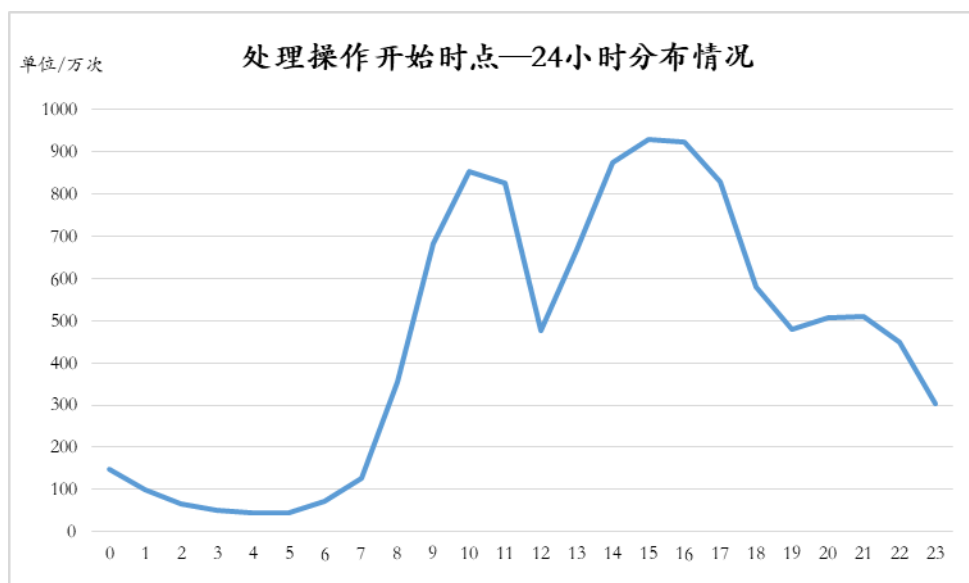
⁹ 终端人管理系统上线初期，作为系统上线的过渡措施，存在发行人将留存的线下终端人信息导入线上并要求终端人配合填写必要信息以完成注册的情形（线下导入情形）。注册账号总数包含前述线下导入情形及线上注册部分。线下导入情形中，截至目前存在部分终端人尚未配合完成信息填写，本部分分析不包含前述注册信息完善中的线下导入账号。

以“处理操作开始时点”作为终端人单次登录时点指标、“处理操作单日工作时长”作为终端人单次登录的时长指标，对终端人的登录时点、单次登录时长情况进行了分析核查分析，

具体如下：

1、登录时点指标（处理操作开始时点）的分析核查

终端劳务人员处理操作开始时点的分布情况如下：

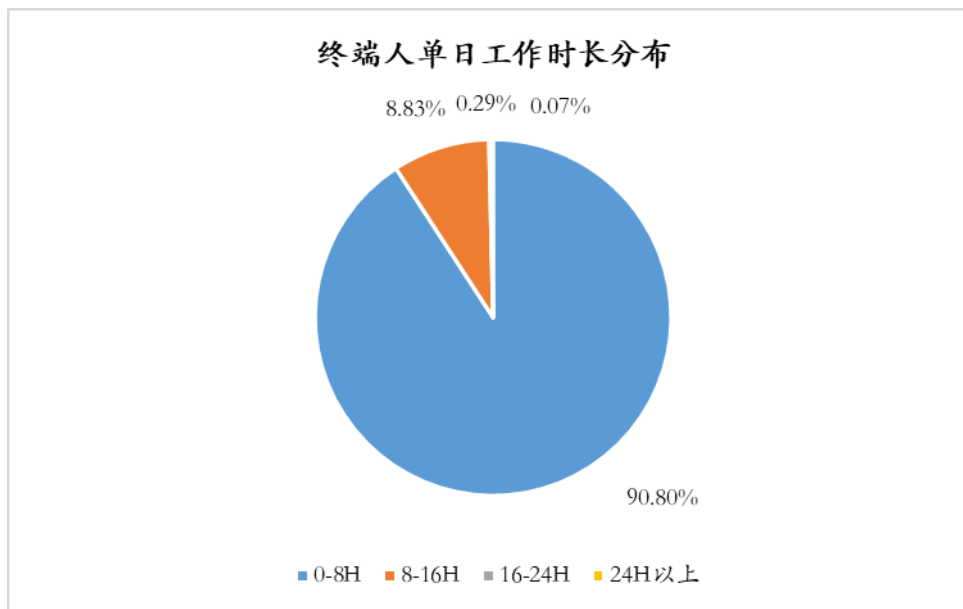


如上图所示，终端劳务人员主要集中在各日8-12时、13至18时之间开始进行处理操作，15时为终端劳务人员开始上线操作的高峰；终端劳务人员在21时至次日4时开始上线操作的情形较少，凌晨4时、5时开始上线操作情形最少。

整体而言，终端劳务人员开始处理操作的时点分布合理，符合普遍的作息规律和工作习惯，不存在异常。

2、单次登录时长（处理操作单日工作时长）指标分析

中介机构对终端劳务人员进行处理操作的单日工作时长分布情况进行了统计，图示如下：



如图所示，终端劳务人员处理操作的单日工作时长主要集中在8小时以内，超过8小时的情形较少，符合普遍的作息规律和工作习惯。

此外，极少数注册账号存在单日工作时长超出16个小时的情况，总共涉及账号**64**个，占对应期间活跃账号总数的比例低于**1.00%**。中介机构对于此类情形执行了进一步分析及核查程序，具体如下：（1）极少数注册账号存在单日工作时长超出24小时的情形（共计**9**个），原因是：个别项目中，存在多名终端劳务人员共享一个注册账号为该项目提供服务的情形，导致该账号记录的累计工作时长超长。为还原各终端人单独工作的时长情况，中介机构将前述账号的工作时长按照MAC地址或项目情况进行了拆分，拆分后各终端人单独工作时长一般为2-12小时，时长合理，不存在异常情况。（2）极少数单日工作时长为16-24小时的账号（共计**55**个）中也存在多名终端劳务人员共享一个注册账号为项目提供服务的情形，导致该账号记录的累计工作时长超过16小时。按照MAC地址或项目情况进行拆分后，除个别账号存在加班情况外，各终端人单独工作时长一般为2-12小时，时长合理，不存在异常情况。

少数终端劳务人员存在单日工作时长在8-16小时的情形，共涉及账号数量**1,888**个，占对应期间活跃账号总数的比例为**2.78%**。除少量账号存在多名终端劳务人员共享一个注册账号为项目提供服务的情形外，终端劳务人员单日工作时长在

8-16小时的情况一般为加班情形。上述终端劳务人员均为标注人员，标注人员相比于采集人员工作时间连续，加班次数较多，符合发行人的实际业务情况，不存在异常情况。同时，针对上述多名终端劳务人员共享账号情形，发行人已对一体化数据处理平台进行了优化，通过改进程序禁止多个设备同时登录同一账号执行处理操作的情形，后续未再出现此类共用情况。

综上所述，终端劳务人员单日工作时长分布合理，符合普遍的作息规律和工作习惯，不存在异常。

（四）登录频率核查

结合发行人业务特点，终端劳务人员在一定期间内登录一体化数据处理平台开展处理操作的天数反映了其在该段时间内的登录频率。

中介机构获取了2019年9月终端人管理系统、日志管理系统上线以来至2020年9月30日期间的活跃账号清单，以上线执行处理操作的天数（以下简称“操作天数”）作为其登录频率指标，对前述用户的登录频率情况分析核查如下：

1、采集环节的终端劳务人员登录频率分析

指标	操作天数： ≤5	操作天数： 5-10	操作天数： >10
有过采集操作的注册账号数量	48,656	173	13
账号分布	99.62%	0.35%	0.03%
平均采集天数	1.20		

有过采集操作的账号的平均采集天数为**1.20**天，整体操作天数以5天内为主（占比**99.62%**）。因发行人的采集任务通常需求人数多、需求多元化（如特定方言、口音、人种要求等），因此多为一次性、临时性交易，同一采集终端人通常只参与一次采集，参与采集的累计时间也较短。上述登录频率的分布情况与发行人实际业务情况相符、不存在异常。

2、标注环节终端劳务人员登录频率分析：

指标	操作天数： ≤20	操作天数： 20-50	操作天数： 50-100	操作天数： >100
----	--------------	----------------	-----------------	---------------

有过标注操作的注册账号数量	15,528	2,794	754	236
账号分布	80.41%	14.47%	3.90%	1.22%
平均标注天数	14.34			

有过标注操作的账号的平均标注天数为**14.34**天，整体操作天数主要集中在20天以内（占比**80.41%**）。前述登录频率分布情况与发行人实际业务中标注人员工作的一般时长规律相符，不存在异常。

综上，发行人终端用户登录频率分布情况与其业务开展的实际情况相符，不存在异常。

（五）登录IP地址信息核查

中介机构取得了终端劳务人员执行采集、标注工作的IP地址数据，对IP地址的集中度（即多个账号登录同一IP地址执行处理**操作**的情况）进行了统计分析核查。具体如下：

核查指标	对应终端劳务人员账号数量 (N) : N=1	对应终端劳务人员账号数量 (N) : 1<N≤10	对应终端劳务人员账号数量 (N) : 10<N≤50	对应终端劳务人员账号数量 (N) : 50<N≤100	对应终端劳务人员账号数量 (N) : N>100
IP地址数量	30,448	5,182	521	48	37
占全量IP地址数量比例	84.03%	14.30%	1.44%	0.13%	0.10%

经统计，终端劳务人员执行处理操作的IP地址较为分散，**84.03%**的IP地址均只记录了1个注册账号登录的情形，不涉及多个账号集中的登录情况。剩余少量IP地址存在多个账号登录并执行处理操作的情形。中介机构对上述存在2个及以上账号在同一IP地址集中登录的情形执行了进一步分析及核查程序，具体如下：

（1）针对账号登录数量为 $1 < N \leq 100$ 的IP地址，中介机构执行了抽查程序（覆盖前述IP地址数量比例为**78%**），核查在相应IP地址登录的终端劳务人员信息及所属项目情况。经核查，由于不同终端劳务人员从属于同一项目，集中为发行人提供服务，从而产生同一个IP地址对应多个终端人账号的情形，不存在异常状况。

（2）针对登录账号数量 $N > 100$ 的IP地址（如下表所示），发行人对各IP地址

情况执行了进一步核查分析：

IP地址	终端人注册 账号个数	IP地址类型	IP地址对应机构 /个人	IP地址对应区域
47. **.***.198	17,425	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
43. **.***.210	3,317	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
124. **.***.178	3,259	发行人所在地IP 地址	发行人	中国北京市
123. **.***.136	2,492	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
123. **.***.130	2,460	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
112. **.***.38	1,500	供应商组织人员 常驻IP地址	北京博睿驰信人 力资源有限公司	中国山东省
47. **.***.115	840	一体化数据处理 平台工具集部署 地址	发行人	中国香港
182. **.***.109	823	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
120. **.***.126	369	终端人员的组织 人员常驻IP地 址	终端人员的组织 人员（即小外包 人员）	中国山东省
125. **.***.148	357	终端人员的组织 人员常驻IP地 址	终端人员的组织 人员	中国浙江省
174. **.***.246	350	终端人员的组织 人员常驻IP地 址	终端人员的组织 人员	加拿大
47. **.***.215	335	一体化数据处理 平台工具集部署 地址	发行人	中国香港
36. **.***.221	307	供应商组织人员 常驻IP地址	北京博睿驰信人 力资源有限公司	中国北京市
61. **.***.109	290	终端人员的组织 人员常驻IP地 址	终端人员的组织 人员	中国湖北省
118. **.***.46	286	终端人员的组织 人员常驻IP地 址	终端人员的组织 人员	日本

1. **.***.120	268	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河南省
61. **.***.108	228	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国湖北省
39. **.***.247	221	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
41. **.***.174	218	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	南非
58. **.***.215	205	供应商组织人员 常驻地IP地址	善世（广东）企 业服务外包有限 公司	中国浙江省
60. **.***.217	188	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河北省
122. **.***.63	170	供应商组织人员 常驻地IP地址	善世（广东）企 业服务外包有限 公司	中国浙江省
67. **.***.197	159	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	加拿大
27. **.***.218	158	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河北省
125. **.***.209	155	供应商组织人员 常驻地IP地址	善世（广东）企 业服务外包有限 公司	中国浙江省
78. **.***.64	153	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	立陶宛
195. **.***.106	147	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	英国
27. **.***.53	145	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河北省
47. **.***.27	144	一体化数据处理 平台工具集部署 地址	发行人	中国北京市
221. **.***.231	142	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国山西省

120. **.***.62	125	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国山东省
115. **.***.203	124	供应商组织人员 常驻地IP地址	善世（广东）企 业服务外包有限 公司	中国浙江省
121. **.***.42	121	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河北省
106. **.***.41	113	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国河北省
112. **.***.85	106	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国山东省
112. **.***.176	104	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国山东省
171. **.***.16	103	终端人员的组织 人员常驻地IP地 址	终端人员的组织 人员	中国湖北省

注：上表中 IP 地址 78. **.***.64 为发行人某立陶宛语音采集项目终端人员的组织人员常驻地 IP 地址，IP 地址 174. **.***.246 为发行人某加拿大美语语音采集项目终端人员的组织人员常驻地 IP 地址，IP 地址 118. **.***.46 为发行人某日语语音采集项目终端人员的组织人员常驻地 IP 地址，IP 地址 41. **.***.174 为发行人某南非祖鲁语语音采集项目终端人员的组织人员常驻地 IP 地址，IP 地址 67. **.***.197 为发行人某加拿大法语语音采集项目终端人员的组织人员常驻地 IP 地址，IP 地址 195. **.***.106 为发行人某香港粤语自由对话语音采集项目终端人员的组织人员（于英国生活的香港居民）常驻地 IP 地址。IP 地址 47. **.***.198、43. **.***.210、123. **.***.136、123. **.***.130、182. **.***.109、47. **.***.115、39. **.***.247、47. **.***.27、47. **.***.215 为发行人一体化数据处理平台工具集部署于各云端服务器的 IP 地址。

经核查，前述账号集中度较高的IP地址主要为发行人部分数据处理工具部署的IP地址、发行人办公地的IP地址、发行人经常合作的供应商或终端人员的组织人员常驻地点的IP地址。登录上述IP地址执行处理操作的账号集中度较高，主要是因为：

- 1、发行人的部分数据处理工具部署在云端服务器上，终端劳务人员远程登录并使用前述工具集时记录的IP地址为统一的云端服务器IP地址；
- 2、部分项目涉及大规模终端人采集，且采集工作在发行人办公地开展，被采集人员使用发行人办公地网络注册账号，登录后进行数据采集；
- 3、发行人长期合作的供应商或终端人员的组织人员在其常驻地统一组织终端人执行处理操作，多名终端人在同一网络环境下登录系统并进行具体操作，体现为同一IP地址登录。上述IP地址的账号分布情况，符合发行人数据服务采购的组织特点。

针对上述登录注册账号数量超过100个的IP地址，中介机构执行了进一步核查程序，具体包括：1、取得前述IP地址对应的登录账号的注册信息（包含姓名、身份证号、联系方式等唯一性标识），并与发行人员工名单核对，确认终端劳务人员数量远大于发行人员工数量，不存在异常；相应注册账号真实、信息有效，不存在脱离实际业务经营虚构账号等异常情况。2、对于供应商、终端人员的组织人员所在地IP地址，抽查核对各项目终端劳务人员的籍贯信息、手机号码归属地与登录IP操作地址是否一致。经核查，终端劳务人员籍贯、手机号码归属地与其登录IP地址一致，不存在异常。

综上，注册账号的IP地址分布情况与发行人业务开展的实际情况相符，不存在异常。

（六）核查结论

综上，中介机构通过多种核查手段（账号的地域分布和数量情况核查、活跃账号数量核查、单次登录时长与登录时点情况核查、登录频率核查、登录IP地址信息核查）对用户行为进行了核查，确认用户行为真实、合理，符合发行人实际业务开展情况或普遍规律，不存在异常。

三、核查范围和比例

核查类型		核查范围		核查比例	
		保荐机构	申报会计师	保荐机构	申报会计师
注册账号真实性核查	账号注册内控核查及穿行测试、注册信息合理性核查	2019年9月终端人管理系统上线至2020年9月30日期间的注册账号（线上注册部分）（共计6.87万个）		覆盖对应期间注册账号总数的58% ¹⁰	覆盖对应期间注册账号总数的55%
注册账号变动的合理性分析		2018年1月1日至2020年9月30日，归属各期执行项目的注册账号		覆盖归属各期执行项目的注册账号数量的100%	
用户行为核查	账号注册地域分布和数量核	2019年9月终端人管理系统上线至2020年9月30日期间的注册账号		覆盖对应期间注册账号总数的55%	

¹⁰ 注册账号真实性核查中，保荐机构与申报会计师核查比例略有差异，主要原因是：终端人管理系统上线初期，有少量线下导入账号未被系统正确标记为“线下导入”。对于此类账号，保荐机构出于谨慎性考虑在核查线上注册账号的同时也对其进行了核查；申报会计师则针对线上注册账号开展核查，未将其纳入核查范围。发行人已针对上述情形对终端人管理系统进行了更新修正，后续未再出现此种情形。

查、活跃账号数量核查	(线上注册部分) (共计 6.87 万个)	
活跃账号数量情况核查	2019年9月终端人管理系统上线至2020年 9月30 日期间的活跃账号(共计 6.79 万个)	覆盖对应期间注册账号总数的 55%
单次登录时长与登录时点情况核查、登录IP地址核查、登录频率核查	2019年9月终端人管理系统上线截至2020年 9月30 日期间的活跃账号(共计 6.79 万个)	覆盖对应期间活跃账号总数的 100%

四、核查结论

综上，中介机构通过多种核查手段对一体化数据处理平台注册账号的真实性、变动合理性以及用户行为进行了充分核查，确认注册账号真实、变动合理，用户行为真实、合理，符合发行人业务实际情况或普遍规律，不存在异常。

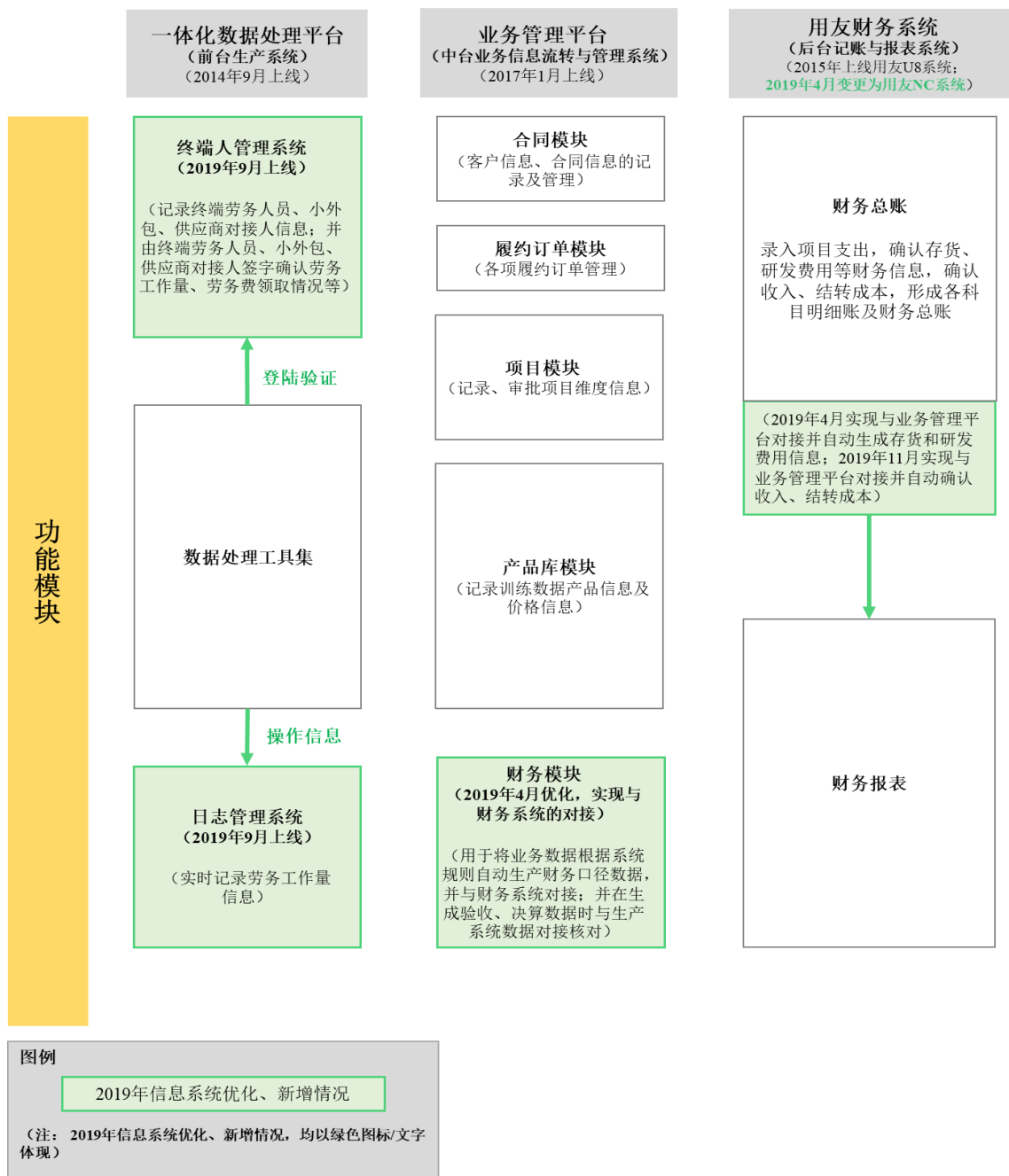
(2) 信息系统审计的具体范围及完整性，是否存在核查范围受限、历史数据缺失和信息系统缺陷等情形，对信息系统可靠性的核查程序、核查比例、核查过程和核查结论。

一、信息系统审计的具体范围及完整性，是否存在核查范围受限、历史数据缺失和信息系统缺陷等情形

(一) 信息系统审计的具体范围及完整性

1、信息系统审计的具体范围

保荐机构(聘请安永(中国)企业咨询有限公司提供信息系统审计专业服务支持)和申报会计师(信息系统审计专家团队提供专业支持)执行了信息系统审计程序，具体范围为发行人的一体化数据处理平台(前台生产系统)、业务管理平台(中台业务信息流转与管理系统)与用友财务系统(后台记账与报表系统)。



图：发行人信息系统构成及功能模块一览

对发行人信息系统基本情况介绍如下：

1、一体化数据处理平台。主要功能包含训练数据集生产相关的流程管理、质量管理、数据安全的管理，嵌入生产所需的各类工具、软件模块，供训练数据生产使用。主要使用部门为公司的业务部门。一体化数据处理平台初始上线时间为2014年9月，后续在2019年9月陆续上线终端人管理系统及日志管理系统。

2、业务管理平台。于2017年1月开始上线使用，记载了发行人销售、采购业务相关信息，主要功能包含合同和项目创建、项目执行过程和结算、验收及决算管理，覆盖贯穿项目管理全生命周期，并可根据平台系统中记录的业务数据生成收入、成本、存货、产品研发费用等业务相关的财务数据。主要使用部门为业务部门及财务部等。

3、用友财务系统。于2015年开始上线使用，自上线至2019年4月为用友U8系统，2019年4月至今为用友NC系统，系统自上线以来未发生重大变更。主要使用部门为财务部。

中介机构对上述信息系统执行了包括信息系统公司层面控制核查（ITEC）、信息系统一般层面控制核查（ITGC）和信息系统应用控制核查（ITAC）在内的信息系统审计程序，对报告期内发行人信息系统相关控制的有效性情况进行了核查。

此外，保荐机构和申报会计师还执行了信息系统数据分析核查（ITDA）程序，对报告期内信息系统留存的业务数据及财务数据的完整性、一致性、准确性进行核查分析。具体分析程序、覆盖期间及核查比例请参见下文“二、对信息系统可靠性的核查程序、核查比例、核查过程和核查结论/（一）保荐机构核查程序、核查比例、核查过程和核查结论及（二）申报会计师核查程序、核查比例、核查过程和核查结论”相关说明。

2、信息系统审计的完整性

发行人的信息系统（一体化数据处理平台、业务管理平台与用友财务系统）为发行人提供了业务支持和财务核算支持，覆盖发行人业务经营管理和财务核算的各个流程环节。保荐机构和申报会计师针对核心业务的各个流程环节，对信息系统的访问控制、程序开发、测试与维护等内部控制及报告期内上述信息系统留存的业务数据和财务数据均执行了信息系统审计程序，信息系统审计具备完整性。

此外，保荐机构和申报会计师对发行人产品和服务的交付方式进行了IT核查。发行人产品和服务的交付方式均为FTP传输（少数情况下应客户要求，发行人在提供FTP传输方式的同时亦提供硬盘交付方式）。发行人使用自有FTP向客户传输数

据，发行人业务人员通过内部邮件方式向IT部门申请所需的FTP地址及账号权限，在对应FTP地址中上传训练数据，并将该FTP地址、账号等信息邮件发送给客户，为客户提供一定期限内数据下载权限。发行人的FTP服务器会自动保存操作日志，日志内容包括用户登录记录、文件上传和下载记录、以及上文件名称和大小等信息。

保荐机构和申报会计师访谈了发行人相关的业务人员，了解FTP传输交付方式相关的业务流程，核查FTP账号权限设立及数据传输流程的内部控制情况，获取了报告期内的FTP日志，并对其进行抽样审阅（2017-2019年各年随机抽查50个项目；2020年前三季度随机抽查50个项目），查看各年抽样项目对应FTP的详细上传、下载日志，及前述项目的客户验收文件、收入确认账务处理情况。经核查，发行人对于FTP传输交付方式设置了相应的内控控制流程，对外交付的训练数据通过FTP进行的上传和下载行为真实合理；客户在通过FTP下载交付的产品/服务后进行验收、回复邮件确认验收完成，发行人在取得客户的验收确认后进行收入确认的会计处理，相关销售交付过程真实、会计处理准确合理，不存在异常情况。

（二）是否存在核查范围受限、历史数据缺失和信息系统缺陷等情形

1、信息系统审计不存在核查范围受限的情形

保荐机构在与发行人签署《保荐协议》时，已与发行人约定，发行人的义务包括按照法律法规的要求和保荐机构的要求提供信息与资料，并对所提供的信息与资料的真实性、准确性和完整性承担责任。申报会计师在与发行人签署《业务约定书》时，已与发行人约定，发行人有责任确保申报会计师不受限制地接触其所认为必要的发行人人员、记录、文件和其他信息，并及时提供审计所需资料。

在各自独立执行信息系统审计的过程中，保荐机构、申报会计师实际的信息系统审计工作并未受到限制。发行人向保荐机构及申报会计师分别开放了业务管理平台和用友财务系统的全部查询权限，保荐机构、申报会计师核查人员独立进入了上述系统，能够查询系统中记录的相关信息；对于一体化数据处理平台，发行人向保荐机构、申报会计师分别提供了登录账号并开放了后台日志信息，核查人员可以不受限制地接触核查过程所必要接触的人员，获取与信息系统审计有关的信息。保荐

机构、申报会计师均已取得了充分、适当的证据资料以支持所出具的信息系统审计意见。

综上，保荐机构、申报会计师的信息系统审计不存在核查范围受限的情形。

2、发行人信息系统中的历史数据情况

发行人的信息系统由一体化数据处理平台、业务管理平台、用友财务系统构成。上述系统构成及功能模块情况参见本题“（一）信息系统审计的具体范围及完整性/1、信息系统审计的具体范围。”

经保荐机构和申报会计师核查，2017年，发行人未留存生产过程的日志信息，发行人自2018年开始保留一体化数据处理平台中已记载的基本日志信息，并于2019年9月上线日志管理系统，在原有基本日志信息留存的基础上增加详细日志信息留存，能够更为全面地记录并保留生产过程的日志信息；同时，发行人于2019年9月上线终端人管理系统，完整记录终端劳务人员的个人信息、劳务量、签字确认信息等，使得整个数据生产过程更加可追溯，信息记录及留存更加详细。此外，发行人2017年1月上线的业务管理平台完整记载了报告期内与财务报表相关的核心运营指标，能够确保发行人采购信息完整、准确记载。

一体化数据处理平台是否完整保存生产过程的详细日志信息，并不影响发行人的业务开展及财务信息的生成和处理。报告期内，发行人在业务管理平台完整保存了与财务报表相关的核心运营数据，在用友财务系统中完整记录和保存了财务数据，发行人与财务报表相关的内部控制制度健全并得到有效执行。

3、发行人信息系统内部控制有效，具备可靠性，不存在重大缺陷

保荐机构和申报会计师执行了信息系统公司层面控制核查（ITEC）、信息系统一般层面控制核查（ITGC）和信息系统应用控制核查（ITAC）程序，对报告期内发行人信息系统（一体化数据处理平台、业务管理平台、用友财务系统信息系统）的可靠性、有效性及发行人业务管理和财务核算的内部控制有效性进行了核查验证。经核查，报告期内发行人对信息系统的内部控制有效，信息系统具备可靠性、不存在重大缺陷。

4、小结

综上所述，保荐机构、申报会计师信息系统审计不存在核查范围受限的情形。经核查，报告期内，发行人的业务管理平台和用友财务系统完整记录、保留了与财务报表相关的核心运营数据和财务数据，不存在历史核心运营数据和财务数据缺失的情况。2017年发行人信息系统中的一体化数据处理平台虽未留存生产日志信息，但并不影响发行人的业务开展及财务信息生成和处理，与财务报表相关的核心运营数据均在业务管理平台中有所记载留存，可支撑财务信息的生成，确保发行人内部控制管理措施的有效实施、财务数据的真实、准确、完整。经核查，报告期内发行人对信息系统的内部控制有效，信息系统具备可靠性、不存在重大缺陷。

二、对信息系统可靠性的核查程序、核查比例、核查过程和核查结论

（一）保荐机构核查程序、核查比例、核查过程和核查结论

保荐机构（聘请安永（中国）企业咨询有限公司提供信息系统审计专业服务支持）组成信息系统核查团队进行了信息系统核查，并出具了《华泰联合证券有限责任公司关于北京海天瑞声科技股份有限公司信息系统专项核查报告》，具体情况如下：

1、核查程序和过程

保荐机构从信息系统公司层面控制、信息系一般层面控制核查、信息系统应用层面控制核查、信息系统数据分析四个层面对发行人信息系统的可靠性进行了核查，具体如下：

（1）信息系统公司层面控制（ITEC）

针对信息系统公司层面控制，保荐机构从公司组织架构、方针政策、制度政策、人员培养、风险管理等多角度进行核查，评估公司的信息系统是否满足企业业务要求，从而甄别公司信息系统管控方针和策略层面是否存在舞弊或错误而导致的重大错误陈述的风险。

信息系统核查团队执行了访谈、资料查阅等程序，对发行人提供的信息技术机

构设置及权责分配、IT战略/计划的制定与实施、IT政策与规章的管理、IT人员培训与考核、IT内部控制与风险管理等方面的材料进行审阅，并与各对接人进行访谈确认。

（2）信息系统一般层面控制核查（ITGC）

针对信息系统一般层面控制，保荐机构对于信息系统变更管理、访问管理、运维管理等从制度、流程、实际执行情况等多方位进行核查，核查公司的信息系统是否在变更、访问和运维方面进行严格完善的控制及管理，从而对公司信息系统内部控制的有效性进行评估。

信息系统核查团队执行了访谈、现场核查、资料查阅等程序，对发行人信息系统变更管理、访问管理、运维管理涉及的制度流程文档等进行审阅，并通过穿行测试和抽样测试的方式核查发行人在实际执行中是否遵循对应制度流程规定。

（3）信息系统应用层面控制核查（ITAC）

针对信息系统应用层面控制，保荐机构对发行人业务流程中依赖于系统功能控制的内容进行核查，检查系统自动控制功能是否有效，相关业务和财务数据是否被准确、完整的记录。

信息系统核查团队执行了现场穿行测试等程序，对发行人业务流程中依赖于系统功能控制的内容，如终端人注册、终端人操作、终端人工作量结算、财务报销等流程进行穿行测试，确认发行人信息系统是否能够正确支持各项业务流程的进行。

（4）信息系统数据分析（ITDA）

信息系统数据分析主要是针对实际业务流程相关数据情况，对发行人业务和财务数据的真实完整性、系统内及系统间的匹配一致性以及业务情况的变化趋势进行分析性核查，从而对公司业务发展情况进行评估。

信息系统核查团队执行了现场数据获取和分析等程序，对发行人业务和财务数据，如终端劳务人员信息表、日志系统终端人工作量信息表、业务管理平台系统合同信息表数据等进行真实完整性、系统内及系统间的匹配一致性以及业务情况的变

化趋势等多角度多维度分析，从而进一步得到分析结果，如是否存在异常情况及其影响程度等。

2、核查比例

对于信息系统公司层面控制（ITEC）、信息系统一般层面控制核查（ITGC）、信息系统应用层面控制核查（ITAC）和信息系统数据分析（ITDA）中的核查，涉及到的核查比例如下表所示：

核查方面		核查范围及比例	
信息系统公司层面控制核查（ITEC）		核查了2017年1月1日至2020年9月30日（报告期）期间，一体化数据处理平台、业务管理平台与用友财务系统的公司层面控制水平	
信息系统一般层面控制核查（ITGC）		核查了2017年1月1日至2020年9月30日（报告期）期间，一体化数据处理平台、业务管理平台与用友财务系统的一般层面控制水平	
信息系统应用层面控制核查（ITAC）		核查了2017年1月1日至2020年9月30日（报告期）期间，一体化数据处理平台、业务管理平台和用友财务系统的应用控制水平	
信息系统数据分析核查（ITDA）	注册账号真实性核查	账号注册内控核查及穿行测试、注册信息合理性核查 对2019年9月终端人管理系统上线至2020年9月30日期间的注册账号（线上注册部分）进行核查，覆盖对应期间注册账号总数的 58%	
	用户行为核查	账号地域分布和数量核查	对2019年9月终端人管理系统上线至2020年9月30日期间的注册账号（线上注册部分）进行分析，覆盖对应期间注册账号总数的 55%
		活跃账号数量核查	对2019年9月终端人管理系统上线至2020年9月30日期间的活跃账号进行分析，活跃账号覆盖同期注册账号总数的 55%
		单次登录时长与登录时点情况核查、登录IP地址核查、登录频率核查	对2019年9月终端人管理系统上线截至2020年9月30日期间的活跃账号的活跃情况进行分析，覆盖同期活跃账号总数的 100% 。
	匹配一致性分析	对2018、2019年、2020年1-9月各期业务管理平台中已确认收入的项目进行了核查，各期收入覆盖比率达到 95%、97%、97% 。	
	趋势分析	对2018、2019年、2020年1-9月各期业务管理平台中已确认收入的项目进行了核查，各期收入覆盖比率达到 95%、97%、97% 。	

3、核查结论

发行人的信息系统内部控制有效、具备可靠性。

（二）申报会计师核查程序、核查比例、核查过程和核查结论

申报会计师利用信息风险管理专家的工作对信息系统可靠性进行专项核查，并

出具了《北京海天瑞声科技股份有限公司信息系统核查报告》，具体情况如下：

1、核查对象及核查内容

基于发行人的业务经营及财务记账的流程，通过访谈、穿行测试等方式，会计师了解并评估了发行人的业务运营流程及支撑的信息系统，确认对发行人信息系统的可靠性和有效性进行核查程序的对象为：一体化数据处理平台（前台生产系统）、业务管理平台（中台业务信息流转与管理系统）与用友财务系统（后台记账与报表系统）。

2、核查程序及核查过程

针对发行人的信息系统，从信息系统一般性控制、信息系统应用控制、计算机辅助程序测试与数据分析三个维度进行穿行测试、控制测试、实质性测试。核查程序及过程的具体细节如下：

（1）信息系统一般性控制（ITGC）

对信息科技治理（含第三方管理）、信息安全管理及系统及数据访问管理、信息系统开发、测试和维护、信息系统运维管理等四个信息系统的企业内部控制管理领域，细分多个控制点，以评估信息系统一般性控制设计和执行的有效性。

在评估过程中，主要采用的程序为现场访谈、观察（含系统查询）和文档审阅，具体如下：

①现场访谈：现场与发行人相关人员进行面对面交流，了解信息科技相关检查领域的管理流程、风险及控制现状。

②现场文档审阅：审阅发行人信息科技管理相关的制度、流程、管理策略等文档，验证相关检查领域控制的设计有效性。检查信息科技管理机制下相关领域和管理流程的审批记录、审阅记录、测试记录、配置信息等，确认相关控制的执行有效性。系统查询账号、权限、参数、日志等，确认相关控制的执行有效性。

③观察：直接观察、查看科技人员在信息系统中执行相关业务的过程，或直接使用发行人授权的账号登录信息系统进行查询、验证。

④现场走访与考察：现场走访发行人信息系统服务器等机器设备所在的机房，对机房的设施及机房门禁等情况进行观察。

针对上述各评估领域，在信息科技一般性控制管理方面未发现发行人存在重大内部控制缺陷。

（2）信息系统应用控制（ITAC）

进行穿行测试，了解并测试发行人的主要业务模式，数据生成和流转的过程及数据存储的情况并形成概要流程图，将每个流程细分了多个步骤与控制，识别系统的自动控制和人工控制，并对识别的系统自动控制进行测试，包括业务经营数据与财务数据传输及财务系统记账自动化、一体化数据处理平台在导入后/标注后/导出后三个不同节点业务操作过程中对文件的保存、终端人注册/登录及操作日志记录、项目报销管理的申请与审批等，以评估发行人信息系统应用控制的有效性。

针对上述各项控制的评估，会计师信息系统审计专家团队通过现场观察发行人业务人员的操作、使用测试账号在一体化数据处理平台进行重新执行等方式、自主提取系统中相关数据进行重新计算。

发行人信息系统的应用控制对业务数据处理、项目管理、财务记账等支撑公司的正常运营与管理流程方面对数据的记录一致、计算准确。同时发行人持续对信息系统进行优化与升级，完善了中台系统业务经营数据与后台系统财务数据传输及财务系统记账自动化、前台系统终端人注册/登录及操作日志记录等方面，系统传输与自动入账完整、准确，日志记录准确。

（3）计算机辅助程序测试与数据分析（CAATs & DA）

申报会计师的信息系统审计专家团队基于对发行人信息系统的业务数据与财务数据一致性核查结果，选取主要项目计算各项指标并进行数据分析，包括项目的设计/导入/交付等环节文件数量一致性、废弃率、单位长度语音标注字节数量、项目标注效率、采集单价、单位语音时长采购成本、项目毛利率等，并进一步分析各项目的指标数据的波动或整体分布情况，以评估发行人业务数据的真实性和合理性。

申报会计师执行了访谈、获取账号权限自主提取数据、重新计算、数据指标分析等程序，对业务和财务数据进行了报告期期间的全量数据测试，保证业务数据和财务数据的一致性；结合财务数据对业务数据进行数据指标分析，并对指标的分布、趋势等情况进行分析，对分析的结果要求发行人提供合理性解释和进一步相关材料支撑。同时，申报会计师结合IT核查的结果，执行了进一步补充审计和核查程序，以确认业务的合理性与真实性。

3、核查比例

针对发行人一体化数据处理平台、业务管理平台与用友财务系统的可靠性，申报会计师的信息系统审计专家团队从信息系统一般性控制、信息系统应用控制和计算机辅助程序测试、数据分析三个维度开展核查，具体的核查比例如下：

核查程序	核查内容		核查范围及核查比例
信息系统一般性控制核查 (ELC&ITGC)	信息科技治理、系统一般性控制		针对报告期各期一体化数据处理平台、业务管理平台与用友财务系统的信息科技治理和系统一般性控制水平进行核查。
信息系统应用控制核查 (ITAC&CAATs)	一体化数据处理平台的数据流转完整性		对2017年1月1日至2020年9月30日期间（报告期）一体化数据处理平台和业务管理平台的系统应用控制水平进行核查。
	业务管理平台对项目管理的控制有效性		
	一体化数据处理平台升级后对终端人登录控制及注册信息、操作日志记录准确性		对2019年9月30日至2020年9月30日期间，一体化数据处理平台对终端人注册、登录及操作环节相关的控制情况及信息记录的准确性进行核查。
	用友财务系统与业务管理平台数据的一致性、用友财务系统相关记账的准确性测试		核对2017年1月1日至2020年9月30日期间（报告期）业务管理平台与用友财务系统的收入、成本相关数据一致性。 核查范围覆盖报告期内完成验收的所有项目，通过计算机辅助程序进行了数据比对。
业务数据和相关财务数据分析 (DA)	注册账号真实性核查	账号注册内控核查及穿行测试、注册信息合理性核查	对2019年9月终端人管理系统上线至2020年9月30日期间的注册账号（线上注册部分）进行核查，覆盖对应期间注册账号总数的 55%
	用户行为核查	账号地域分布和数量核查	对2019年9月终端人管理系统上线至2020年9月30日期间的注册账号（线上注册部分）进行分析，覆盖对应期间注册账号总数的 55%
		活跃账号数量核查	对2019年9月终端人管理系统上线至2020年9月30日期间的活跃账号进行分析，活跃账号覆盖同期注册账号总数的 55%
		单次登录时长与登录时点情况核查、登录IP地址核查、登录频率核查	对2019年9月终端人管理系统上线截至2020年9月30日期间的活跃账号的活跃情况进行分析，覆盖同期活跃账号总数的 100% 。

	项目基础日志与业务财务指标分析	抽样核查了2018年1月1日至2020年9月30日期间确认收入的项目（项目开始及执行的时间段覆盖了2017年1月1日2020年9月30日报告期间），核查项目的收入金额覆盖2018年1月1日-2020年9月30日期间发行人训练数据定制服务和训练数据产品（不含训练数据产品重复销售收入）合计金额的 67% 。
--	-----------------	---

注：ELC 即公司层面信息科技治理，ITGC 即信息系统一般性控制，ITAC 及信息系统应用控制，CAATs 即计算机辅助审计技术，DA 即数据分析

4、核查结论

发行人的信息系统内部控制有效、具备可靠性。

三、总结

1、信息系统审计的具体范围为发行人的一体化数据处理平台（前台生产系统）、业务管理平台（中台业务信息流转与管理系统）与用友财务系统（后台记账与报表系统）。

2、保荐机构和申报会计师针对核心业务的各个流程环节，对信息系统的访问控制、程序开发、测试与维护等内部控制及报告期内上述信息系统留存的业务数据和财务数据均执行了信息系统审计程序，信息系统审计具备完整性。

3、保荐机构、申报会计师信息系统审计不存在核查范围受限的情形。

4、报告期内，发行人的业务管理平台和用友财务系统完整记录、保留了与财务报表相关的核心运营数据和财务数据，不存在历史核心运营数据和财务数据缺失的情况。2017年发行人信息系统中的一体化数据处理平台虽未留存生产日志信息，但并不影响发行人的业务开展及财务信息生成和处理，与财务报表相关的核心运营数据均在业务管理平台中有所记载留存，可支撑财务信息的生成，确保发行人内部控制管理措施的有效实施、财务数据的真实、准确、完整。

5、报告期内发行人对信息系统的内部控制有效，信息系统具备可靠性、不存在重大缺陷。

2、关于数据服务提供商

根据问询回复，发行人前五大数据服务提供商采购集中度较高且主要通过小外包进行采购。

请发行人说明：（1）结合发行人与服务提供商合作建立的过程、历次商务沟通记录和管理服务费率确定过程，说明管理服务费的定价依据及公允性；（2）报告期各期与数据服务采购相关留存的内外部记录及完整性，包括但不限于采购通知、采购过程沟通记录、采购量结算明细、日志记录、验收结算确认单据和银行流水等，采购人次和数量与相关业务数据、内外部记录的匹配关系及存在差异的原因；（3）报告期各期末是否存在正在执行过程中但尚未结算的数据服务成本及费用计提情况，相关内部控制制度及其执行情况，是否存在费用跨期的情形。

请保荐机构和申报会计师对上述事项进行核查并发表明确意见，进一步说明对与数据服务采购相关留存的内外部记录的核查程序、核查过程、核查比例和核查结论。

答复：

发行人说明事项

（1）结合发行人与服务提供商合作建立的过程、历次商务沟通记录和管理服务费率确定过程，说明管理服务费的定价依据及公允性；

发行人通过市场调研寻找目标数据服务提供商，经过电话等形式的初步洽谈形成合作意向后，双方进行正式商务谈判并形成会议纪要，后续根据会议纪要按照谈判结果拟定框架性协议供双方签署。在谈判和选定数据服务提供商的过程中，发行人主要考量以下因素：

因素	关注点
服务能力	短期内寻找及组织人员、关键节点把控、组织培训、及时支付报酬、配合程度、处理突发事件的能力
资源情况	公司的规模、资源情况（包括分布地域）、满足特殊需求的能力
按期交付能力	项目管理能力、沟通能力和效率

价格	人力资源成本和管理服务费率
----	---------------

在综合考量以上因素的基础上，发行人也参考市场同类服务的费率水平最终与数据服务提供商谈判确定管理服务费率。

发行人与报告期内主要数据服务提供商均建立了稳定的合作关系。对于境内主要数据服务提供商，发行人在商业谈判中普遍约定 11.5% 的服务费率。由于部分境内数据服务提供商管理成本上升的原因，经发行人与其进行商务沟通，2020 年起服务费率变更为 12%；由于在发行人与个别境内数据服务提供商的商务沟通过程中了解到其税务成本略低，因此约定采用 10% 的服务费率，后根据其税务成本变动情况，经发行人与其进行商务沟通，2020 年起服务费率变更为 10.5%。对于境外主要数据服务提供商，发行人在商业谈判中主要根据其所在地的税务成本和人力资源匹配成本，约定服务费率，没有相对统一的费率标准，差异略大于境内，具备商业合理性。

发行人与主要境内外数据服务提供商建立合作的时间如下：

区域	主要数据服务提供商名称	合作建立时间
境内	善世（广东）企业服务外包有限公司 （曾用名：广州善世人力资源服务有限公司）	2016年
	共青城宜邦人力资源服务有限公司	2019年
	芜湖善慧人力资源服务有限公司	2017年
	才星（广州）科技服务有限公司 （曾用名：广州金社宝网络技术有限公司）	2018年
	北京博睿驰信人力资源有限公司	2016年
	北京零点起航营销策划有限公司	2015年
	北京冠华英才国际经济技术有限公司	2015年
	信阳市联智网络科技有限公司	2016年
境外	恩泽尔商贸有限公司	2016年
	IVO EMEA Ltd.	2020年
	Voices.com Inc.	2017年
	Atrium Studio Medienproduktion GmbH	2018年
	SAKURA JAPAN LLC	2018年
	TWIN STUDIOS	2018年

THINKWILD STUDIOS, S.L.	2018年
-------------------------	-------

发行人与主要境内外数据服务提供商约定的管理服务费率符合发行人对市场同类人力资源服务公司和劳务公司的市场询价区间（9-15%）。根据广州善世、芜湖善慧、北京博睿、北京零点、信阳联智等主要境内数据服务提供商出具的说明，其向其他客户收取的服务费率区间为 9-13%，根据恩泽尔等主要境外数据服务提供商出具的说明，其向其他客户收取的服务费率区间为 5-8%，与该等数据服务提供商向发行人收取的服务费率基本一致，不存在明显差异。

综上所述，发行人与数据服务提供商约定的管理服务费率，充分考虑了数据服务提供商在服务能力、资源情况、按期交付能力等方面的特点，并参考了市场同类价格水平，由双方经过充分商务沟通后确定，具有较为充分的定价依据，定价水平公允。

（2）报告期各期与数据服务采购相关留存的内外部记录及完整性，包括但不限于采购通知、采购过程沟通记录、采购量结算明细、日志记录、验收结算确认单据和银行流水等，采购人次和数量与相关业务数据、内外部记录的匹配关系及存在差异的原因；

一、数据服务采购相关留存的内外部记录及其完整性情况

发行人数据服务采购相关留存的主要内外部记录及其完整性如下：

内外部记录名称	所属采购环节	记录内容	完整性情况
采购通知	下达采购通知★	项目简介、服务工期、符合项目预算的原料数据采集、标注服务量、采集或标注的人员要求和特别要素（例如被采集人的年龄、籍贯、性别、口音特点、采集场地，标注人员的语言背景、专业技能、学历要求）等。	发行人根据项目预算，对于 10 万元以上的数据服务采购下达采购通知。报告期各期采购通知分别留存 612 份、838 份、1,016 份和 910 份 ，占当期数据服务采购金额的比例分别为 77.32%、84.82%、81.39%和 86.96% 。对于 10 万元以下的数据服务采购，由于其金额较小占比较低，出于工作便利方面的考虑，发行人与数据服务提供商对接人员直接沟通确认，具体形式包括电话、通讯软件等。
采集信息记录	供应商提供服务	项目信息、被采集人基本信息或语料文本字词/句数。其中被采集人的身份证号、电话、籍贯等信息在 2019 年发行人加强终端人员管理后才成为必须留存的信息。	随着发行人项目管理和信息系统的日益完善，对采集信息记录的留存要求逐步提高。报告期各期采集信息记录分别留存 89 份、174 份、324 份和 315 份 ，占当期数据采集服务采购金额的比例分别为 37.53%、56.77%、97.71%和 98.97% 。

报销明细	验收结算及费用报销★	项目编号、财务编号、项目支出的类型，对应的项目任务环节和采购的任务需求，对应采购的数量、单价和报销金额	发行人所有数据服务采购均有对应报销明细。报告期各期报销明细分别留存 4,447 笔、6,178 笔、5,705 笔和 4,226 笔 ，占当期数据服务采购金额的比例均为 100%。
验收结算确认单	验收结算及费用报销★	项目编号、财务编号、采购劳务内容、采购劳务的数量及对应金额	发行人所有数据服务采购均有对应验收结算确认单。报告期各期验收结算确认单分别留存 3,048 份、3,968 份、3,737 份和 2,969 份 ，占当期数据服务采购金额的比例均为 100%。
银行流水	供应商付款	收款人信息、付款时间、付款金额	发行人所有数据服务采购均有对应银行流水。报告期各期银行流水分别发生 7,956 笔、6,464 笔、5,116 笔和 3,552 笔 ，占当期数据服务采购金额的比例均为 100%。

注 1：上述过程中的关键控制节点以★标示。

注 2：对于同一个项目，如存在向不同数据服务提供商采购的情形，则需要分别发出采购通知，按照多份采购通知计算。

注 3：根据项目类型，采集信息记录具体表现为项目采集信息表、语料确认邮件等不同形式，每个留存了采集信息记录的项目按照 1 份采集信息记录计算。

注 4：报销明细、验收结算确认单、银行流水均覆盖全部数据服务采购金额，但份数有所差异是由于各个记录的记录口径不同。发行人向数据服务提供商支付的管理服务费单独报销，因而报销明细份数多于验收结算确认单。发行人直接向终端人员采购数据服务时，按照项目产生报销明细和验收结算确认单，按照向单个终端人员的支付情况产生银行流水，因而银行流水笔数明显多于报销明细和验收结算确认单份数。

关于采购通知的形成过程和信息基础，具体说明如下：

发行人对每个项目均编制详细的项目预算，根据“以销定采”的业务逻辑，通过确定每个项目具体的采集损耗率和标产比，进而确定该项目需要的采集数据量和标注有效工时，并根据项目具体情况确定采集、标注的特别要素。项目预算经过内部审批后，相关信息会单独填报于发行人的业务管理平台。因此，每个项目采购通知中涉及的采集数据量和标注有效工时等信息均以发行人业务管理平台中记载留存的已经审批的项目预算为基础，经内部审核后下达给供应商，采购通知以会议纪要或邮件的形式保存留痕。

发行人制定了严格的采购预算管理制度，对于全部业务采购支出均会进行预算管理。在数据服务采购支出发生前，发行人会对每个项目编制详细的项目预算，一般项目为三级审批制，重大项目为五级审批制。具体为：所有项目的预算均须业务部门经理、分管业务副总经理及财务审核岗三级审批，若项目预算金额大于 20 万元人民币，则需要追加财务经理四级审批，若项目对应合同/订单金额大于 50 万元人民币或预算毛利率为负（由于业务拓展、客户维护等原因），则需要追加财务总监五级审批。对于项目预算中数据服务费部分，采集数据量会参照“以销定采”的

业务逻辑，根据销售合同或订单约定的交付数据量并考虑合理的采集损耗率后确定，标注量（即有效工时）根据采集数据量并参照合理的标产比后确定；标产比根据项目启动前的自测标产比或与客户约定的标产比确定；采集单价和标注单价参照公司采购价格体系确定；项目负责人编制预算时需要将数据服务费的采购量、采购单价均按工作环节逐项、单独填报于业务管理平台，经过业务部门、财务部等多层级审批通过后方可执行。

此外，发行人数据服务采购相关的其他内外部记录还包括采购过程沟通记录和日志记录。采购过程沟通记录是发行人与数据服务提供商在采购过程中关于价格、工期、数据质量、需求变更及其他突发问题等事项进行沟通所形成的记录，通常由发行人项目负责人与数据服务提供商对接人员直接沟通，具体形式包括电子邮件、补充确认单、通讯软件记录等。由于各项目的具体情况有所不同，因此采购过程沟通记录的内容、形式和数量并不一致，其沟通确认的情况最终主要体现在验收结算确认单之中。日志记录是发行人数据服务采购过程中信息系统记录的各项过程数据，随业务类型的不同而有所差异，且较为繁杂，通常不作为直接核对的证据性文件，而是供复核人员对报销数据合理性进行分析时作为参考。

除上述内外部记录之外，发行人于 2019 年在内部管理流程中加入终端人员对提供服务量进行签字确认的环节，从而进一步确认采购量数据。2019 年和 2020 年 1-9 月，留存了终端人员签字的项目收入占发行人总体收入（不含训练数据产品重复销售）的比例分别为 74% 和 98%。

二、采购量与相关业务数据的匹配关系及存在差异的原因

发行人数据服务采购过程呈现“以销定采”、“同类业务标产比可比”的业务逻辑，因此训练数据定制服务的销售量和训练数据产品的产量等销售产出业务数据作为重要业务数据与采购量之间存在匹配关系。

（一）采集服务的采购量与销售产出业务数据的匹配关系

采集损耗率是发行人采集数据量相对于交付给客户的数据量的损耗比例，例如对于语音识别采集项目，采集损耗率=（供应商结算总人数-客户验收发音人总人数）

/客户验收发音人总人数，对于语音合成项目，采集损耗率=（供应商结算字数-客户验收字数）/客户验收字数。报告期内，发行人采集损耗率总体控制在 5%以内，说明采集数据量与交付给客户的数据量之间具备匹配关系。

采产比是发行人采集数据量和销售量/产出量之间的比例关系。对于训练数据定制服务，销售量为实际销量数据；对于训练数据产品，产出量为当期研发的训练数据产品新增产出量。发行人有采集环节的主要业务类别（包括智能语音全部细分类别语音识别采集、语音合成，计算机视觉主要细分类别手写体、光学字符识别）在报告期内的一般情况采产比如下：

业务类别	2020年 1-9月	2019年	2018年	2017年	波动原因
语音识别采集	1.0	0.9	1.0	1.1	较为平稳，无明显波动
语音合成	0.8	0.6	1.0	1.1	2019年和2020年1-9月无采集环节的语音合成数据标注项目较多
手写体	2.2	2.6	2.5	2.6	较为平稳，无明显波动
光学字符识别（有采集）	1.1	1.0	1.0	1.0	较为平稳，无明显波动

注：采集数据量单位、销售量/产出量单位、采产比含义如下：

语音识别采集：人、小时、每小时产出所需人数

语音合成：万字词、小时、每小时产出所需万字词数

手写体：采集工时、千词组、千词组产出所需采集工时

光学字符识别（有采集）：百张、百张、百张产出所需百张数

发行人各业务类别在报告期内的采集数据量和销售量/产出量之间普遍存在上述比例关系，整体波动较小，匹配关系合理。此外，部分特殊项目的采产比与一般情况有所差异，是由于项目特殊要求、发音人资源稀缺、产品完善与扩建无采集环节等特殊原因所致。

（二）标注服务的采购量与销售产出业务数据的匹配关系

标产比是发行人标注有效工时和销售量/产出量之间的比例关系。对于训练数据定制服务，销售量为实际销量数据；对于训练数据产品，产出量为当期研发的训练数据产品新增产出量。发行人主要业务类别（包括智能语音全部细分类别语音识别采集、语音识别转写、语音合成，计算机视觉主要细分类别手写体、光学字符识别、行车相关视频和自然语言主要细分类别语义理解）在报告期内的一般情况标产

比情况如下：

业务类别	2020年 1-9月	2019年	2018年	2017年	波动原因
语音识别 采集	2.7	3.9	3.0	5.2	2017年和2019年自由对话类项目（标注复杂度高）较多，2020年1-9月部分项目标注复杂度低或无需标注
语音识别 转写	13.4	11.4	12.5	11.1	2018年和2020年1-9月长句音频项目（标注复杂度高）占比较高，2020年1-9月个别主要客户的项目包含测试集数据，准确率要求高，需要投入较多的质检工时，以及部分项目需要进行无效数据标注并注明无效原因
语音合成	25.9	45.5	45.9	67.3	发行人高效数据处理技术逐步成熟和普遍应用，提高了发行人数据处理工作效率，有效减少了标注过程中的人工工作量，使标产比整体呈下降趋势。2020年1-9月标产比较低还由于当期大部分项目无需进行音素边界标注，因此投入标注工时较少
手写体	0.2	-	-	1.3	2017年和2020年1-9月项目语种不同，且2020年1-9月部分项目标注环节仅需简单的数据整理，标注工时很少
光学字符 识别（有 采集）	2.2	-	0.5	0.5	2020年1-9月主要项目采集对象为实况图片及财务文档、作文图片，标注复杂度高
光学字符 识别（无 采集）	27.0	1.2	0.9	1.0	2020年1-9月均为标注难度很高的项目
行车相关 视频	0.06	0.07	0.09	-	较为平稳，无明显波动
语义理解	1.3	1.0	1.2	1.2	2019年开始发行人采取了机器辅助标注手段，进行预标注后再由人工校对，降低了人工投入

注：标注有效工时单位、销售量/产出量产出、标产比含义如下：

语音识别采集：小时、小时、每小时产出所需标注时长

语音识别转写：小时、小时、每小时产出所需标注时长

语音合成：小时、小时、每小时产出所需标注时长

手写体：小时、千词组、千词组产出所需标注时长

光学字符识别（有采集）：小时、百张、百张产出所需标注时长

光学字符识别（无采集）：小时、百张、百张产出所需标注时长

行车相关视频：小时、十个目标、十个目标产出所需标注时长

语义理解：小时、小时、每小时产出所需标注时长

发行人各业务类别在报告期内的标注有效工时和销售量/产出量之间存在上述比例关系和波动原因，匹配关系合理。此外，部分特殊项目的标产比与一般情况有所差异是由于客户要求较高、涉及较难语种/较难口音/网络用语、地名等内容导致转写难度较大等特殊原因所致。

上述采产比和标产比在报告期内的波动原因及特殊项目情况的具体分析，请参见第一轮审核问询函问题 10.1 “发行人说明事项/（6）采集数据量、标注数据量和标产比相关的确定方法和具体依据，是否存在随意修改或调整的情形，同类型项目的采集数据量、标注数据量和标产比是否存在差异并分析具体原因”的回复内容。

三、采购量与内外部记录的匹配关系及存在差异的原因

发行人采购量与前述主要内外部记录的匹配关系和差异原因如下：

内外部记录名称	与采购量匹配关系	差异原因
采购通知	与采购量存在口径差异	采购通知记载执行采购前根据项目预算确定的预计采购量，由于项目实际执行与预算可能不完全一致，因此与实际采购量可能存在差异。
采集信息记录	与采购量存在口径差异	采集信息记录记载执行采购过程中发生的采集服务量，由于采集过程中可能存在数据不符合要求等情况而不予结算，因此与实际采购量可能存在差异。
报销明细	与采购量一致	-
验收结算确认单	与采购量一致	-
银行流水	根据采购量计算采购金额	-

以 2019 年某 300 人多设备多场景语音识别采集项目为例，该项目采集服务采购量 320 人，标注服务采购量 1,209 小时，各项主要内外部记录与采购量之间的匹配关系如下：

1、采购通知记载预计采集服务采购量 315 人、预计标注服务采购量 1,212.5 小时，其中预计采集服务采购量和预计标注服务采购量与实际采购量不同，是项目预算情况与实际执行情况的差异所致；

2、采集信息记录记载采集人数 327 人，与实际采集服务采购量不同，是由于采集过程中部分被采集人的数据不符合要求而未实际结算所致；

3、报销明细记载采集服务采购量 320 人、标注服务采购量 1,209 小时，采集服务采购金额 89,820.61 元（不含税）、标注服务采购金额 20,070.00（不含税），与实际采购量和采购金额一致；

4、验收结算单记载采集服务采购量 320 人、标注服务采购量 1,209 小时，采集服务采购金额 $89,820.61 \times (1+6\%) = 95,209.85$ 元（含税）、标注服务采购金额

$20,070.00 \times (1+6\%) = 21,274.20$ 元（含税），合计采购金额 116,484.05 元（含税），与实际采购量和采购金额一致；

5、发行人银行流水及供应商银行回执显示支付采购金额合计 116,484.05 元，与实际采购金额一致。

(3) 报告期各期末是否存在正在执行过程中但尚未结算的数据服务成本及费用计提情况，相关内部控制制度及其执行情况，是否存在费用跨期的情形。

报告期内，数据服务提供商在服务提供完毕并经发行人验收后，与发行人结算数据服务费，每季度末发行人对数据服务提供商在服务提供过程中已产生但未结算的数据服务费进行暂估计提。由于发行人与数据服务提供商之间的结算周期较短，且发行人于每季度执行严格的数据服务采购相关的内部控制程序以保证数据服务费支出进度同项目执行情况的一致性，因此发行人报告期各期末正在执行过程中但尚未结算的数据服务费的暂估金额较小。发行人于报告期各期末暂估计提的数据服务费金额及占各期数据服务采购金额的比例情况如下：

单位：万元

项目		2020年1-9月	2019年	2018年	2017年
期末暂估的数据服务费	A	77.54	36.59	3.14	6.54
当期采购的数据服务费	B	5,646.32	6,099.95	7,352.66	4,057.31
比例	A/B	1.37%	0.60%	0.04%	0.16%

报告期内，发行人主要通过制定和执行严格的项目预算和项目决算相关管理制度及审批流程来保证数据服务费的入账完整性。

在项目执行之前，发行人需要对每个项目编制详细的项目预算，并经过业务部门直属负责人、财务部审批通过后方可执行。在项目进行过程中，数据服务费的采购严格按照项目预算执行，如果超出预算，则需要经过业务部门分管负责人、财务负责人批准后方可执行。

每季度末，发行人根据项目预算及项目执行情况综合判断预计总成本和项目支出进度，判断过程由项目负责人与财务部共同进行，经其签字确认后作为会计档案

保存。同时，财务人员对已结算入账的数据服务费进行复核，对已发生未结算的数据服务费根据项目预计总成本和项目实际进度情况进行暂估计提，并经业务部门负责人、财务部门负责人审批通过。

项目完工后，要求及时进行项目决算。项目决算应在项目全部验收后进行，最晚不得晚于验收后10个工作日，且季度内验收的项目必须在季度末完成决算、当年验收的项目应在当年完成决算。项目决算后不得再对项目费用进行报销。

发行人通过建立并确保有效执行上述与数据服务费核算相关的内控制度及审批流程，保证了数据服务费核算的完整性，不存在费用跨期的情形。

中介机构核查事项

(1) 核查程序

对于上述说明事项，保荐机构、申报会计师已经执行了与数据服务采购相关留存的内外部记录的核查程序，具体请参见本问题中介机构说明事项。另外，针对上述发行人说明事项，保荐机构、申报会计师进一步执行了如下核查程序：

保荐机构、申报会计师执行的主要程序如下：

1、通过与业务人员访谈，了解发行人与数据服务提供商的合作建立过程，获取历次的商务沟通记录，了解管理服务费率确定过程；

2、获取发行人对市场上同类人力资源服务公司和劳务公司的管理服务费率询价结果，与发行人的管理服务费率进行比较，评价管理服务费率的定价是否公允；

3、抽取发行人报告期内的项目，检查业务管理平台中记录的项目预算审批情况是否符合发行人项目预算审批相关的内部控制制度；

4、获取发行人报告期内的销售产出业务数据，抽取项目，检查采购量同销售产出业务数据的匹配关系；

5、了解和评价与数据服务费相关的内部控制及其执行情况，并执行控制测试；

6、从资产负债表日前后的数据服务费采购支出明细中选取样本执行截止性测

试，通过检查双方盖章确认的验收结算单，来评价接近资产负债表日前后入账的数据服务费采购交易是否记录于恰当期间；

7、获取并查阅报告期各期发行人暂估数据服务费用的明细及暂估依据，检查相关费用后续实际支出和核算情况，确认与原暂估金额是否存在重大差异。

(2) 核查结论

经核查，保荐机构、申报会计师认为：

1、发行人与数据服务提供商合作建立的过程符合发行人的实际情况，发行人的管理服务费的定价依据充分，定价水平公允；

2、报告期各期发行人与数据服务采购相关的主要内外部记录中除采集信息记录外均完整留存，采集信息记录在报告期内留存比例逐步提高，2019年和2020年1-9月已基本完整留存，采购量与相关业务数据相匹配，采购量与采购通知、采集信息记录的核对存在差异的原因合理，采购量与报销明细、验收结算确认单及银行流水具有匹配关系；

3、发行人报告期各期末存在少量正在执行过程中但尚未结算的数据服务费的暂估计提，相关的内部控制制度健全并得到有效执行，发行人不存在数据服务费跨期入账的情形。

中介机构说明事项

(1) 核查程序、核查过程及核查比例

对于报告期内主要数据服务提供商（各期采购金额占比5%以上）的数据服务采购交易，按项目维度逐个进行全样本穿行核查，检查的内外部证据包括：（1）发行人向供应商下达的采购通知（针对10万元以上的数据服务采购），采购通知的内容包括项目简介、服务工期、符合项目预算的原料数据采集、标注服务量、采集或标注的人员要求和特别要素等；（2）采集信息记录，内容包括项目信息、被采集人基本信息或语料文本字词/句数等；（3）报销明细，发行人对供应商提供的劳务服务验收合格后进入采购结算环节，项目人员会在系统中进行报销，相关信

息会录入报销明细，报销明细会载明项目财务编号、项目支出的类型，对应的项目任务环节和采购的任务需求，并记录对应采购的数量、单价和报销金额；（4）验收结算确认单，系统根据报销明细表生成验收结算单，并发送给供应商，由供应商确认后作为双方结算依据。验收结算单是发行人与供应商的结算依据，验收合格后双方签订验收结算单。验收结算单会载明项目财务编号、采购劳务内容、采购劳务的数量及对应金额，并由双方加盖公章；（5）发行人向供应商付款对应的银行流水信息。

根据数据服务采购流程，以采购通知为起点、银行流水为终点，通过逐项目核对采购通知、采集信息记录、报销明细、验收结算确认单及银行流水等内外部记录，确认上述内外部记录中记载的供应商名称、劳务服务采购量及采购金额等信息是否与该项目的实际采购情况相匹配。其中，对于采购通知和采集信息记录，核对其与劳务服务采购量的差异情况，确认是否存在异常差异；对于报销明细、验收结算确认单和银行流水，核对其与劳务服务采购量及采购金额是否一致。

上述核查采购交易金额的覆盖比例如下：

年份	核查数据服务费采购支出金额（万元）	数据服务费采购支出总额（万元）	覆盖比例
2020年1-9月	4,439.50	5,646.32	78.63%
2019年	5,110.19	6,099.95	83.77%
2018年	6,262.67	7,352.66	85.18%
2017年	3,159.40	4,057.31	77.87%
合计	18,971.76	23,156.24	81.93%

此外，检查数据服务采购相关的采集信息记录以及2019年、2020年1-9月增加终端人员劳务量签字确认环节后的签字文件，并核对采集信息记录的采购量是否与发行人记录的实际采购量一致。

上述各项内外部记录的具体的核查过程和核查比例如下：

1、获取项目的采购通知，检查采购通知中是否包含了项目编号、供应商名称、劳务内容，以确认是否与发行人的业务相关，检查采购通知中的采购量同发行人记

录的实际采购量是否存在重大差异，并访谈业务人员了解差异原因。采购通知的核查比例如下：

年份	核查采购通知单金额 (万元)	数据服务费采购支出总额 (万元)	覆盖比例
2020年1-9月	3,703.16	5,646.32	65.59%
2019年	3,822.56	6,099.95	62.67%
2018年	5,097.14	7,352.66	69.32%
2017年	2,209.04	4,057.31	54.45%
合计	14,831.90	23,156.24	64.05%

注：上述核查的采购通知为报告期内主要数据服务提供商（各期采购金额占比 5%以上）10 万元以上数据服务采购交易所涉及的采购通知（不足 10 万元的数据服务采购交易无采购通知），报告期内该类采购交易金额占主要数据服务提供商全部数据服务采购交易金额（即前述“核查数据服务费采购支出金额”）的比例分别为 69.92%、81.39%、74.80%、**83.41%**。

对于上述核查的报告期内主要数据服务提供商（各期采购金额占比 5%以上）10 万元以上数据服务采购交易所涉及的项目，进一步核查其中在报告期内立项且完成的项目的数据服务采购预算金额和数据服务采购验收结算金额的差异情况。该等项目的数据服务采购预算金额为 **11,460.40 万元**，数据服务采购验收结算金额为 **11,305.57 万元**，差异率不超过 **1.5%**，预算与实际执行不存在重大差异。

同时，对于报告期内主要数据服务提供商（各期采购金额占比 5%以上）10 万元以下数据服务采购交易所涉及的项目，同步核查其中在报告期内立项且完成的项目的数据服务采购预算金额和数据服务采购验收结算金额的差异情况。该等项目的数据服务采购预算金额为 **3,217.32 万元**，数据服务采购验收结算金额为 **3,128.14 万元**，差异率不超过 **3%**，预算与实际执行不存在重大差异。

2、获取项目的采集信息记录，检查采集信息记录中是否包含了项目编号、劳务内容、劳务提供时间，以确认是否与发行人的业务相关，检查采集信息记录的采购量同发行人记录的实际采购量是否存在重大差异，并访谈业务人员了解差异原因。采集信息记录的核查比例如下：

年份	核查采集信息记录金额 (万元)	主要采集环节的数据服务费 采购支出总额(万元)	覆盖比例
2020年1-9月	1,625.95	2,247.27	72.35%

2019年	1,174.53	1,987.54	59.09%
2018年	1,334.38	2,504.31	53.28%
2017年	574.02	1,597.47	35.93%
合计	4,708.88	8,336.59	56.48%

注 1：主要采集环节的数据服务费采购支出包括智能语音全部采集环节采购支出和计算机视觉、自然语言以“工作小时”为单位进行结算的采集环节采购支出，与招股说明书披露一致，其占全部采集环节的数据服务费采购支出的比例约为 90%。

注 2：数据服务采购包括采集环节和标注环节，上述采集信息记录核查仅涉及采集环节。

3、获取报销明细，检查报销明细中是否包含了项目编号、劳务内容，以确认是否与发行人的业务相关，检查报销明细中的采购数量、采购金额是否同发行人实际记录的实际采购量和采购金额一致。报销明细的核查比例如下：

年份	核查报销明细金额 (万元)	数据服务费采购支出总额 (万元)	覆盖比例
2020年1-9月	4,439.50	5,646.32	78.63%
2019年	5,110.19	6,099.95	83.77%
2018年	6,262.67	7,352.66	85.18%
2017年	3,159.40	4,057.31	77.87%
合计	18,971.76	23,156.24	81.93%

4、获取验收结算确认单，检查验收结算确认单中是否包含了项目编号、供应商名称、劳务内容，以确认是否与发行人的业务相关，检查验收结算确认单中记录的采购数量、采购金额是否同发行人实际记录的实际采购量和采购金额一致。验收结算确认单的核查比例如下：

年份	核验收收结算确认单金额 (万元)	数据服务费采购支出总额 (万元)	覆盖比例
2020年1-9月	4,439.50	5,646.32	78.63%
2019年	5,110.19	6,099.95	83.77%
2018年	6,262.67	7,352.66	85.18%
2017年	3,159.40	4,057.31	77.87%
合计	18,971.76	23,156.24	81.93%

5、获取发行人支付供应商的银行流水及供应商对外支付劳务费用的银行回执，检查银行流水记录的收款人信息是否与发行人的记录一致，银行流水对应的采购金

额是否同数据服务费采购金额一致，供应商对外支付劳务费用金额是否与扣除管理服务后数据服务费采购金额一致。银行流水的核查比例如下：

年份	核查银行流水金额 (万元)	数据服务费采购支出总额 (万元)	覆盖比例
2020年1-9月	4,439.50	5,646.32	78.63%
2019年	5,110.19	6,099.95	83.77%
2018年	6,262.67	7,352.66	85.18%
2017年	3,159.40	4,057.31	77.87%
合计	18,971.76	23,156.24	81.93%

6、对于提供数据服务的终端劳务人员，发行人2019年要求其对于提供的劳务工作量等信息进行签字确认，以完善数据服务采购管理并充分留痕。对于发行人一体化数据处理平台上线终端人管理系统前，终端劳务人员签字确认主要为线下纸质形式；上线终端人管理系统后，采用线上电子签名的形式进行确认。因此获取2019年及**2020年1-9月**的终端人员的线下签字单据和线上签字单据，检查上述签字文件是否包含了项目编号、劳务内容、劳务提供时间，检查上述签字文件是否有终端劳务人员的签字或电子签名，以确认是否与发行人的业务相关，是否经终端劳务人员签字确认。

已执行上述核查程序的2019年及**2020年1-9月**终端人员的核查比例如下：

核查了2019年及**2020年1-9月**终端人员确认所提供劳务工作量的签字文件，并根据项目编号查找与核查的终端人员签字文件相关的项目收入金额。2019年及**2020年1-9月**已核查的终端人员签字文件相关的项目收入占发行人当期总体收入（不含训练数据产品重复销售）的比例分别为74%及**98%**。

(2) 核查结论

经核查，保荐机构、申报会计师认为：报告期各期发行人与数据服务采购相关的主要内外部记录中除采集信息记录外均完整留存，采集信息记录在报告期内留存比例逐步提高，2019年和**2020年1-9月**已基本完整留存，采购量与采购通知、采集信息记录的核对存在差异的原因合理，采购量与报销明细、验收结算确认单及银

行流水具有匹配关系。

3、关于收入确认政策

根据问询回复：（1）发行人未明确披露训练数据定制服务取得的验收确认凭证；（2）多次交付、多次验收的情形下发行人按照完工进度确认收入；（3）新收入准则下发行人将每单元定制化数据认定为单项履约义务且属于在某一时点履行的履约义务。

请发行人补充披露：（1）各类业务的具体交付物及交付方式，各类型收入确认取得的具体验收确认凭据；（2）旧收入准则下确定完工进度的具体方法；（3）新收入准则下发行人评估客户取得商品控制权时点及其判断依据，相关交易价格以及分摊至各单项履约义务所采用的方法；（4）定制服务成本的核算流程和结转方法。

请发行人说明：（1）对华为技术有限公司的服务或销售收入是否存在初验验收和最终验收的相关约定，发行人对华为技术有限公司服务或销售收入的具体确认时点和取得的验收依据，是否存在提前确认收入的情形，是否存在其他类似情形，如有，请详细说明；（2）结合交付物内容及其自身的特征和实际业务执行过程，说明将每单元定制化训练数据识别为单项履约义务的具体依据以及充分性；（3）结合相关合同规定，说明交易价格分摊至各单项履约义务所采用的计算方法和依据；（4）发行人的收入确认政策是否符合企业会计准则的规定。

请保荐机构和申报会计师对上述事项进行核查并发表明确意见，并进一步说明发行人的相关信息披露和会计处理是否符合新收入准则的相关规定。

答复：

发行人补充披露事项

（1）各类业务的具体交付物及交付方式，各类型收入确认取得的具体验收确认凭据；

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主要采用的会计政策和会计估计/（十七）收入/4、各类业务的具体交付物及交付方

式，各类型收入确认取得的具体验收确认凭据”中补充披露下述内容：

一、各类业务的具体交付物及交付方式

1、训练数据定制服务和训练数据产品

发行人两大主要业务类型训练数据定制服务和训练数据产品覆盖了智能语音、计算机视觉及自然语言处理三大 AI 核心领域。前述服务和产品均以软件化的训练数据集为核心，成品训练数据集主要由数据文档、说明文档、技术文档三部分构成。不同业务领域下的具体交付物总结如下：

（1）智能语音业务的交付物一般包括原始采集形成的音频文件、与音频文件对应的带有时间戳的标注文件，训练数据集相关的设计文档、训练数据集说明、发音词典、数据集参数信息文件等；

（2）计算机视觉业务的交付物一般包括图像和视频等数据文件、标注文件和说明及技术文档；

（3）自然语言业务的交付物一般包括对话文本等数据文件、标注文本文件和说明及技术文档。

上述业务类型的实际交付方式主要为以 FTP 传输方式交付给客户。

2、训练数据相关的应用服务

发行人提供 AI 算法模型相关训练服务，交付物一般包括开发完成或者授权使用的应用程序和配套硬件（如有），交付方式主要为以 FTP 传输方式交付给客户或交付至客户指定地点并现场安装调试。

二、各类型收入确认取得的具体验收确认凭据

1、训练数据定制服务和训练数据产品

发行人训练数据定制服务的各类型收入确认方式，包括三种类型，即（1）一次交付，一次验收；（2）多次交付，一次验收；（3）多次交付，多次验收（分次验收）。训练数据产品已经开发完成，可多次销售并获取授权许可收入，收入确认

方式为产品交付验收后一次确认。

上述收入确认方式下，发行人取得的具体验收确认凭据均为客户的验收邮件。项目完成或阶段性工作完成后，发行人项目负责人将服务或产品交付客户，并向客户项目负责人邮件发送验收申请，验收申请附有项目验收报告，列明项目/产品名称及编号、服务/产品内容、本次验收的工作量等具体待验收信息。客户完成验收后，正式邮件回复确认服务或产品已交付，验收内容无误。

发行人训练数据定制服务和训练数据产品主要客户为长期合作的大型科技公司、人工智能企业及科研机构，客户选择邮件验收主要是考虑到办公习惯及工作效率方面的原因，且邮件验收与签字盖章的验收单同样具有验收作用，符合行业惯例。

2、训练数据相关的应用服务

应用软件的定制开发或者授权使用、以及配套硬件的销售，由于通常涉及简单的安装调试等现场实施服务工作，一般现场取得盖章纸质验收证明作为收入确认凭据；如无需现场实施服务，则采取邮件验收的方式。

（2）旧收入准则下确定完工进度的具体方法；

发行人在 2020 年 1 月 1 日之前，执行财政部于 2006 年颁布的《企业会计准则第 14 号——收入》及《企业会计准则第 15 号——建造合同》（统称“旧收入准则”）。旧收入准则下，发行人采用完工百分比法确认训练数据定制服务的收入，完工进度按照已经提供的劳务占应提供劳务总量的比例来确定。在资产负债表日，发行人统计已经按照合同或订单提供给客户并经客户验收和确认的服务量占服务总量的比例，计算和确定完工进度。

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主要采用的会计政策和会计估计/（十七）收入/1、原收入准则/（2）收入确认的具体方法/①训练数据定制服务收入”中补充披露下述内容：

发行人确定完工进度的具体方法为：

（1）对于一次交付、一次验收的训练数据定制服务，在验收之前，完工进度

为零，在验收之后，完工进度为 100%；

(2) 对于多次交付、一次验收的训练数据定制服务，在验收之前，完工进度为零，在验收之后，完工进度为 100%；

(3) 对于多次交付、多次验收的训练数据定制服务，在资产负债表日根据已经验收的服务量占合同或订单要求提供的服务总量的比例来计算完工进度。例如对于某中英文女声语音合成项目，合同要求提供的训练数据定制服务总量为 50 小时，在资产负债表日如果已验收的训练数据定制服务量为 45 小时，则完工进度为 45 小时/50 小时=90%。

(3) 新收入准则下发行人评估客户取得商品控制权时点及其判断依据，相关交易价格以及分摊至各单项履约义务所采用的方法；

一、发行人评估客户取得商品控制权时点及其判断依据

根据财政部于 2017 年颁布的《企业会计准则第 14 号——收入（修订）》（以下简称“新收入准则”）第十三条的规定：

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主要采用的会计政策和会计估计/（十七）收入/2、新收入准则/（1）收入确认的一般原则”中补充披露下述内容：

“对于在某一时点履行的履约义务，企业应当在客户取得相关商品控制权时点确认收入。在判断客户是否已取得商品控制权时，企业应当考虑下列迹象：（一）企业就该商品享有现时收款权利，即客户就该商品负有现时付款义务。（二）企业已将该商品的法定所有权转移给客户，即客户已拥有该商品的法定所有权。（三）企业已将该商品实物转移给客户，即客户已实物占有该商品。（四）企业已将该商品所有权上的主要风险和报酬转移给客户，即客户已取得该商品所有权上的主要风险和报酬。（五）客户已接受该商品。（六）其他表明客户已取得商品控制权的迹象。”

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主

要采用的会计政策和会计估计/（十七）收入/2、新收入准则/（2）收入确认的具体方法/①训练数据定制服务收入”中补充披露下述内容：

在新收入准则下，发行人将每单元定制化训练数据识别为某一时点履行的单项履约义务。发行人评估认为，在将定制化训练数据向客户交付并经客户验收确认的时点，客户已取得相关商品的控制权，并满足收入确认的条件。发行人的主要判断依据是基于控制权的定义以及结合以下的迹象：（1）根据合同条款约定或者法律规定，通常情况下，发行人对已经交付并经客户验收认可的该部分工作量享有现时收款权，客户就该部分工作量负有现时付款义务；（2）数据在交付验收后，交付物的所有权即已转移给客户；（3）商品的实物在数据交付验收后即已转移给客户；（4）客户的验收表明其已接受该商品；（5）客户对于所交付并验收的定制化训练数据，可以自行决定如何以及何时用于训练人工智能模型并从中受益，数据交付验收后可能发生的减值或毁损等损失也由客户承担，因此客户享有与交付和验收的定制化训练数据相关的主要风险和报酬。

二、相关交易价格以及分摊至各单项履约义务所采用的方法

根据新收入准则第二十条的规定，“合同中包含两项或多项履约义务的，企业应当在合同开始日，按照各单项履约义务所承诺商品的单独售价的相对比例，将交易价格分摊至各单项履约义务。”

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主要采用的会计政策和会计估计/（十七）收入/2、新收入准则/（2）收入确认的具体方法/①训练数据定制服务收入”中补充披露下述内容：

新收入准则下，发行人将每单元定制化训练数据识别为单项履约义务。训练数据定制服务的合同或订单中一般包含多项履约义务，因此需要按照上述准则规定将交易价格分摊至各单项履约义务。发行人与客户签订的训练数据定制服务的合同或订单中，通常会约定提供的同类型的单元定制化训练数据的数量和总价格，或约定提供的同类型的单元定制化训练数据的单价和数量。由于同类型单项履约义务具有相同的性质和特点，因此对于每一个同类型的单元定制化训练数据，单

独售价也相同。发行人采用平均分摊的方法将交易价格分摊至单项履约义务，即按照约定的同类型单元定制化训练数据总价格除以数据总量，计算结果作为各个单元定制化训练数据的交易价格。如果约定了单元定制化训练数据的单价，也可以直接将单价作为各个单元定制化训练数据的交易价格。

(4) 定制服务成本的核算流程和结转方法。

发行人已在招股说明书“第八节 财务会计信息与管理层分析/四、报告期内主要采用的会计政策和会计估计/（十七）收入/6、训练数据定制服务成本的核算流程和结转方法”中补充披露下述内容：

报告期内发行人的定制服务成本的核算流程和结转方法如下：

报告期内，发行人训练数据定制服务的各项支出，包括职工薪酬、数据服务费、折旧与摊销等，均按照所属项目在“制造费用”中进行归集与核算。在项目分批验收或一次全部验收并相应确认收入后，将该项目归集的项目成本中与确认的收入对应的部分，由制造费用结转至营业成本。

按项目归集成本：发行人严格按照项目为单位进行成本归集，在项目执行之前，每个项目需要编制项目预算，根据训练数据定制服务的具体内容，对项目所需人员及每人投入工时等直接人工内容，采集数量、采集单价、标注工时、标注单价等数据服务采购内容，折旧与摊销、租赁等项目所需场地、设备内容，差旅、交通、办公等其他项目相关支出内容制定详细预算。项目执行过程中，所有支出填写和报销均比照预算类型执行，对预算变更、未及预算、超预算等情况设定完整的审批流程，并制定了严格的项目决算要求，在项目申请决算前，所有支出必须填写和报销完毕，经专人审核后项目进行项目决算。决算后，项目不允许发生额外支出。通过以上内部控制手段，保证每笔支出均可以按项目准确、完整归集。

成本结转：具体表现为两种情形，一种是在项目分批验收并确认收入的情况下，在每次验收并确认收入的同时，按照已经验收的服务量占服务总量的比例乘以预计项目总成本并扣除以前期间累计已结转成本后的金额来结转当期营业成本；另一种是在项目一次全部验收并确认收入的情况下，直接进行项目决算，按照实际发生和

归集的项目总成本全部结转至营业成本。制造费用中未结转至营业成本的部分，在财务报表中列示和披露为“存货-在产品”。

报告期内，各相关费用和人工成本在各个项目上的归集和分配的方法如下：

（一）职工薪酬

职工薪酬按照工时积分制归集到各个项目。每月业务人员按其实际工作情况将工时积分填写到相应项目上，每人每月填写至所有项目（包括训练数据定制服务项目和训练数据产品开发项目等）的工时积分共计为 8 分/日*月工作日。各个项目应分摊的人工成本的计算方式为：各业务人员填写在该项目的积分*该人员所属级别的人工成本单价=该项目的人工支出。

（二）数据服务费

数据服务费指训练数据定制服务项目所需的原料数据采集与标注服务采购支出，业务人员按项目实际发生情况进行支出申请和报销。在进行支出申请和报销时，必须填报项目编号、涉及环节、采购量、采购单价、报销金额、供应商名称等关键信息，保证发生额直接归集到单个项目上，不得与其他项目存在交叉。因此数据服务费支出根据上述项目编号信息直接归集和分配至各个项目上。

（三）折旧与摊销

对于各项目使用的设备和软件的折旧和摊销费用，按各项目领用该设备、使用该软件时间的长短，由系统自动计算在各个项目上的分摊金额。计算方式为：项目折旧与摊销支出=项目领用天数/365 天*设备全年折旧额或无形资产全年摊销额。

（四）差旅费等其他费用

差旅费等其他费用按项目实际发生情况报销，在报销时，必须选择项目编号及费用类别，费用发生额直接归集到单个项目上，不得与其他项目存在交叉。因此差旅费等其他费用支出也是根据项目编号信息直接归集和分配至各个项目上。

发行人说明事项

(1) 对华为技术有限公司的服务或销售收入是否存在初验验收和最终验收的相关约定，发行人对华为技术有限公司服务或销售收入的具体确认时点和取得的验收依据，是否存在提前确认收入的情形，是否存在其他类似情形，如有，请详细说明；

发行人与华为技术有限公司签订的合同中存在初步验收和最终验收的约定。按照合同规定，初步验收的内容包括但不限于对数量、类型、数据隐私处理等验收，验收时间为收到交付物之后若干个工作日内，以乙方（即发行人）收到正式通知为准；最终验收的内容包括但不限于对数据数量和质量等验收，验收时间为通过初步验收之后若干个工作日内，以乙方（即发行人）收到正式通知为准。

在实际操作中，在项目完成或阶段性工作完成后，发行人将产品或服务交付客户，通知客户进行初步验收，并通过电话、邮件、即时通讯软件等渠道与客户保持密切联系，及时跟进产品或服务的验收状态。初步验收中如果发现数据存在问题，或者未达到客户的验收标准，客户会要求发行人对交付数据进行改进及修正，直至满足客户的验收标准。客户口头或书面通知可以进行最终验收时，可视为已通过初步验收。最终验收前发行人会填写验收报告，列明项目/产品名称及编号、服务/产品内容、本次验收的工作量等具体验收信息，以邮件的形式发送最终验收申请，客户完成最终验收后，正式邮件回复确认完成本次验收。

发行人向华为技术有限公司提供服务或销售产品的具体收入确认时点为最终验收完成的时点，取得的验收依据为华为技术有限公司的最终验收邮件，不存在提前确认收入的情形。发行人与其他客户签订的合同中不存在类似的初验验收和最终验收的约定。

(2) 结合交付物内容及其自身的特征和实际业务执行过程，说明将每单元定制化训练数据识别为单项履约义务的具体依据以及充分性；

一、交付物内容及其自身的特征和实际业务执行过程

发行人提供训练数据定制服务包括三种业务类型，即智能语音、计算机视觉和

自然语音。三种业务类型下训练数据定制服务均以软件化的训练数据集为交付物，成品训练数据集主要由数据文档、说明文档、技术文档三部分构成。不同业务类型下的具体交付物分别为：（1）智能语音业务的交付物一般包括音频文件、标注文本文件和说明及技术文档；（2）计算机视觉业务的交付物一般包括图像和视频等数据文件、标注文件和说明及技术文档；（3）自然语言业务的交付物一般包括对话文本等数据文件、标注文本文件和说明及技术文档。

上述三种业务类型下的交付物内容一般都包含了数据文档（包含音频文件、图像和视频文件、自然语言对话文本文件等数据文件及标注文件）、说明文档及技术文档。数据文档是核心交付物，也是客户用来进行人工智能模型训练的文件。说明及技术文档一般作为附加文件同时提供给客户，作为数据文档相关标准、规范、使用的说明，不直接应用于客户的人工智能模型训练，客户不能单独受益，不作为履约义务进行分拆。

发行人在实际业务执行过程中，以单元训练数据作为交付和验收的最小单位。单元训练数据是指按不同数据形态呈现的每计量单位的训练数据，如 1 小时语音、1 张图片等。例如，对于 10 小时中文方言语音识别采集项目，发行人向客户交付的训练数据集包括 10 小时中文方言音频和其对应的标注文本，其中，每 1 小时中文方言音频及其标注文本构成一个单元训练数据。发行人每次交付和验收的数据文档内包含多个单元训练数据，由于每个单元训练数据相对独立，它们之间也不会产生相互影响和重大整合的关系，例如每 1 小时的中文方言音频或每 1 张图片之间均为独立的数据文件，因此客户也不会对分次交付和验收的训练数据进行指定，发行人可以以单元训练数据为基础对定制服务进行分次交付、分次验收，客户也可以分次用于人工智能训练。

二、发行人将每单元定制化训练数据识别为单项履约义务的具体依据

根据新收入准则的相关规定：履约义务，是指合同中企业向客户转让可明确区分商品的承诺。企业向客户承诺的商品同时满足下列条件的，应当作为可明确区分商品：（一）客户能够从该商品本身或从该商品与其他易于获得资源一起使用中受

益。（二）企业向客户转让该商品的承诺与合同中其他承诺可单独区分。下列情形通常表明企业向客户转让该商品的承诺与合同中其他承诺不可单独区分：1、企业需提供重大的服务以将该商品与合同中承诺的其他商品整合成合同约定的组合产出转让给客户。2、该商品将对合同中承诺的其他商品予以重大修改或定制。3、该商品与合同中承诺的其他商品具有高度关联性。

根据前述交付物内容及其自身的特征和实际业务执行过程的分析，并结合上述新收入准则的规定，发行人将每单元定制化训练数据识别为单项履约义务的具体依据如下：

1、交付和验收的单元训练数据包含了核心的交付文件，即数据文档，因此客户可以将交付和验收后的每一单元训练数据用于训练人工智能模型，因此能够从该单元训练数据的使用中受益；

2、发行人转让每一单元训练数据的承诺在合同中可以明确区分，主要是因为（1）在转让每一单元训练数据之前，发行人无须提供重大服务以对单元训练数据之间进行整合。以智能语音业务为例，在语料文本的设计阶段之后，采集、加工和质检的环节均可以以单元训练数据为基础进行，并以单元训练数据为基础进行数据交付，不论是对于同一发音人采集的不同文本的数据，还是对于不同发音人采集的同一文本的数据，发行人无须再提供重大服务，对生产后的不同数据之间进行重大整合；（2）发行人转让的每一单元训练数据不会影响其他单元训练数据或其他单元训练数据做出重大修改。例如发行人在录制其他文本/发音人的内容时，不会对已录制完成的内容进行重大修改；（3）不同的单元训练数据之间不存在高度关联性，发行人可以对不同的单元训练数据分批提交，客户可以分批验收，而且分批提交已验收完成的数据可以用于客户的模型训练。

综合上述分析，发行人向客户提交的训练数据集包含多个独立的、可明确区分的单元训练数据，即按不同数据形态（音频、视频、图片、文本等）呈现的每计量单位（小时及句或字词、图片张数等）的训练数据（包括数据文件及标注文件），发行人根据新收入准则的规定将每单元定制化训练数据识别为单项履约义务的依据

充分。

(3) 结合相关合同规定，说明交易价格分摊至各单项履约义务所采用的计算方法和依据；

交易价格分摊至各单项履约义务所采用的方法请参见本问题“发行人补充披露事项/（3）新收入准则下发行人评估客户取得商品控制权时点及其判断依据，相关交易价格以及分摊至各单项履约义务所采用的方法/二、相关交易价格以及分摊至各单项履约义务所采用的方法”的回复内容。由于发行人销售合同下每项服务的定价与其单独售价基本相符，因此可直接采用合同约定的每项服务的总价格或单价计算各单项履约义务的价格。以下结合合同约定分别举例说明约定总价格和约定单价时交易价格分摊至各单项履约义务的具体方法和依据。

1、合同约定总价格

以某多语种语音识别采集训练数据定制服务合同为例，该合同涉及美式英语、韩语两种语言。合同约定，美式英语和韩语两个语种的数据需求量均为 70 小时语料的录制及标注，价格分别为 12 万元和 9.5 万元。因此，发行人根据合同约定将上述两个语种的定制服务分别核算，其中英语语音识别采集项目对应的金额为 12 万元，交付的数据量为 70 小时，将总价格 12 万元除以 70 小时计算得到的平均价格 1,714 元/小时作为单项履约义务的交易价格；韩语语音识别采集项目对应的金额为 9.5 万元，交付的数据量为 70 小时，将总价格 9.5 万元除以 70 小时计算得到的平均价格 1,357 元/小时作为单项履约义务的交易价格。

2、合同约定单价

以某语音识别转写训练数据定制服务合同为例，该合同中分别约定中文普通话音频数据标注单价为 550 元/小时，提供数据服务量为 200 小时；中文方言口音音频数据标注单价为 950 元/小时，提供数据服务量为 100 小时。因此对中文普通话和中文方言口音的语音识别转写项目分别按照约定的单价 550 元/小时和 950 元/小时作为单项履约义务的交易价格。

(4) 发行人的收入确认政策是否符合企业会计准则的规定。

在旧收入准则下，发行人以风险报酬转移作为收入确认时点的判断标准。其中：

(1) 训练数据定制服务属于提供劳务收入，且劳务交易的结果能够可靠估计，根据完工百分比法确认提供劳务收入。在资产负债表日，发行人统计已经按照合同或订单提供给客户并经客户验收和确认的服务量占服务总量的比例，计算和确定完工进度。并按照合同或订单约定的服务交易总额乘以完工进度扣除以前会计期间累计已确认提供的服务收入金额，确认当期的服务收入；(2) 训练数据产品收入属于让渡资产使用权，且不提供后续服务，在客户在对训练数据产品进行验收后，主要风险和报酬转移给客户时确认收入；(3) 训练数据相关的应用服务收入按照合同约定向客户交付开发完成或者授权使用的应用软件和配套硬件（如有），并经客户验收确认后，按照合同约定的交易金额确认收入。

在新收入准则下，发行人以控制权转移作为收入确认时点的判断标准。其中：

(1) 训练数据定制服务将每单元定制化训练数据识别为单项履约义务，且判断属于某一时点履行的履约义务；(2) 发行人不从事对训练数据产品有重大影响的后续活动，因此训练数据产品属于某一时点履行的知识产权许可授权；(3) 应用软件的定制开发或者授权许可、配套硬件的销售属于可明确区分的单项履约义务，并且属于在某一时刻履行的履约义务。对于上述均属于在某一时刻履行的履约义务，发行人均在向客户交付合同中约定的相关商品或服务，并经客户验收确认后，客户取得控制权时点时，按照合同约定的单项履约义务分摊的交易价格确认收入。

发行人的上述收入确认政策符合企业会计准则的规定。

中介机构核查事项

(1) 核查程序

保荐机构、申报会计师执行的主要程序如下：

1、了解并评价与产品服务交付验收及收入确认相关的关键财务报告内部控制的设计和运行有效性；

2、询问相关人员了解发行人的不同业务类型、具体交付物和交付方式。并针对不同业务类型，选取样本检查其具体交付物和交付方式；

3、询问相关人员了解发行人不同类型收入确认的具体验收确认凭据。并针对不同的收入确认类型，选取样本检查其验收确认凭据，以评价相关收入是否按照发行人的收入确认政策予以确认；

4、询问相关人员了解旧收入准则下完工进度的具体确认方法、新收入准则下发行人评估客户取得商品控制权时点及其判断依据、各单项履约义务的相关交易价格的具体分摊方法，并结合企业会计准则的规定和发行人实际业务情况，评价相关具体方法和相关依据是否充分合理；

5、询问相关人员以了解定制服务成本的核算流程和结转方法，并结合企业会计准则的规定和发行人实际业务情况，评价相关成本核算流程和结转方法是否恰当合理；

6、抽取检查相关合同了解发行人与华为技术有限公司所签订的合同中对于初验验收和最终验收的相关约定。询问相关人员了解业务实际执行过程中的验收情况，并进一步查看邮件往来、验收确认依据等具体信息，确认业务实际执行情况；

7、询问相关人员了解发行人与其他客户之间是否存在初验验收和最终验收的相关规定，并抽取主要客户合同进行检查，查看是否存在相关条款；

8、从资产负债表日前后的收入交易中选取样本，检查与收入确认相关的销售合同或订单、客户验收文件、发票等相关支持性依据，评价接近资产负债表日前后记录的收入是否记录于恰当期间；

9、抽取训练数据定制服务收入样本，查看合同规定并与相关业务人员访谈，了解交付物的具体内容和具体交付情况，结合企业会计准则的规定，评价发行人将每单元定制化训练数据识别为单项履约义务的具体依据是否充分；

10、抽取训练数据定制服务收入样本，检查交易价格分摊至各单项履约义务所采用的计算方法和依据，并与可观察到的单独售价或价格体系进行比较，结合企业

会计准则的规定，评价交易价格的分摊方法是否合理；

11、结合企业会计准则的规定和发行人实际业务情况，评价发行人的收入确认政策是否符合企业会计准则的规定。

(2) 核查结论

经核查，保荐机构、申报会计师认为：

1、发行人披露的各类业务的具体交付物及交付方式、各类型收入确认取得的具体验收确认凭据与发行人实际情况相符；

2、旧收入准则下发行人确认完工进度的具体方法合理，符合企业会计准则的相关规定；

3、新收入准则下发行人评估客户取得商品控制权时点的判断依据充分，相关交易价格以及分摊至各单项履约义务所采用的方法合理，符合企业会计准则的相关规定；

4、发行人定制服务成本的核算流程和结转方法合理，符合企业会计准则的相关规定；

5、发行人对华为技术有限公司提供训练数据定制服务的具体收入确认时点为最终验收完成的时点，符合企业会计准则的相关规定，不存在提前确认收入的情形，与其他客户签订的合同中不存在类似的初验验收和最终验收的约定；

6、新收入准则下发行人将每单元定制化训练数据识别为单项履约义务的具体依据充分，将合同交易价格分摊至各单项履约义务所采用的计算方法符合业务实际情况，符合企业会计准则的相关规定；

7、发行人的收入确认政策符合企业会计准则的规定。

中介机构说明事项

发行人的相关信息披露和会计处理是否符合新收入准则的相关规定

一、发行人执行新收入准则的相关信息披露

根据中国证监会《发行监管问答——关于申请首发企业执行新收入准则相关事项的问答》的要求，公司应披露新收入准则实施前后收入确认会计政策的主要差异以及实施新收入准则在业务模式、合同条款、收入确认等方面产生的影响。同时，企业应披露实施新收入准则对首次执行日前各年合并财务报表主要财务指标的影响，即假定自申报财务报表期初开始全面执行新收入准则，对首次执行日前各年（末）营业收入、归属于公司普通股股东的净利润、资产总额、归属于公司普通股股东的净资产的影响程度。

发行人已在招股说明书“第八节 财务会计信息与管理层分析/十三、资产负债表日后事项、或有事项及其他重要事项 /（四）执行新收入准则对公司的预计影响”中根据中国证监会《发行监管问答——关于申请首发企业执行新收入准则相关事项的问答》要求进行了披露，具体如下：

1、实施新收入准则对业务模式、合同条款的影响

公司实施新收入准则对公司各类型业务的业务模式、合同条款均无重大影响。

2、执行新收入准则对公司收入确认的主要影响

现行收入准则下，公司以风险报酬转移作为收入确认时点的判断标准。

新收入准则下，公司以控制权转移作为收入确认时点的判断标准。公司在履行了合同中的履约义务，将相关商品或服务的控制权转移给客户时确认收入。在满足一定条件时，公司属于在某一时段内履行履约义务，否则，属于在某一时点履行履约义务。合同中包含两项或多项履约义务的，公司在合同开始日，按照各单项履约义务所承诺商品或服务的单独售价的相对比例，将交易价格分摊至各单项履约义务，按照分摊至各单项履约义务的交易价格计量收入。交易价格是公司因向客户转让商品或服务而预期有权收取的对价金额，不包括代第三方收取的款项。公司依据新收入准则有关特定事项或交易的具体规定调整了相关会计政策。

总体而言，与现行收入准则相比，公司在新收入准则下的收入确认时点和确认金额均无差异。执行新收入准则对公司各类型业务的收入确认会计政策的具体影响如下：

业务类型	现行收入准则	新收入准则
训练数据定制服务	训练数据定制服务收入主要是指公司根据客户需求为客户提供训练数据定制服务而收取的收入。公司与客户通过合同或订单约定提供的服务内容、服务总量及交易总额。	
	公司采用完工百分比法确认训练数据定制服务的收入。完工进度按照已经提供的劳务占应提供劳务总量的比例来确定。在资产负债表日，公司统计已经按照合同或订单提供给客户并经客户验收和确认的服务量占服务总量的比例，计算和确定完工进度。	每单元定制化训练数据被识别为单项履约义务，属于某一时点履行的履约义务，公司在将定制化训练数据向客户交付并经客户验收确认时确认收入。
训练数据产品	训练数据产品收入主要是指公司授权客户永久使用公司已经开发完成并拥有所有权的训练数据产品而向客户收取的收入。公司与客户签订合同约定授权客户使用的训练数据产品及其交易金额。	
	公司按照合同约定向客户提交训练数据产品并经客户验收确认后，按照合同约定的交易金额确认训练数据产品收入。	公司不从事对训练数据产品有重大影响的后续活动，因此训练数据产品属于某一时点履行的知识产权许可授权，公司在向客户交付合同中约定的训练数据产品并经客户验收确认后，按照合同约定的交易金额确认训练数据产品收入。
训练数据相关的应用服务	训练数据相关的应用服务收入主要是指公司根据客户需求为其提供应用软件的定制开发服务或者是应用软件的授权使用，以及配套硬件的销售（如有）等，而向客户收取的收入。公司与客户通过签订合同，约定向客户提供的应用软件的定制开发或授权使用，以及配套硬件销售（如有）等服务内容及交易金额。	
	公司按照合同约定向客户交付开发完成或者授权使用的应用软件和配套硬件（如有），并经客户验收确认后，按照合同约定的交易金额确认收入。	应用软件的定制开发或者授权许可、配套硬件的销售属于可明确区分的单项履约义务，并且属于在某一时点履行的履约义务；应用软件的定制开发或者授权使用、以及配套硬件的销售，均是在向客户交付并经客户验收确认时确认收入，并按照合同约定的单项履约义务的交易价格确认收入金额。

3、新收入准则对首次执行日前各年度合并报表主要财务指标的影响

由于公司在新收入准则下的收入确认时点和确认金额和现行收入准则下均无差异，假定自 2017 年 1 月 1 日起开始全面执行新收入准则，对公司的合并财务报表中各年（末）营业收入、归属于公司普通股股东的净利润、资产总额、归属于公司普通股股东的净资产均无影响。

二、发行人的会计处理是否符合新收入准则的相关规定

发行人自 2020 年 1 月 1 日起执行新收入准则，以控制权转移作为收入确认时

点的判断依据，综合考虑合同条款和业务安排来评估新收入准则的影响。执行新收入准则后收入确认的会计政策具体如下：

对于训练数据定制服务收入，单元定制化训练数据被识别为单项履约义务，且属于某一时点履行的履约义务，因此在将单元数据向客户交付并经客户验收确认时确认收入。对于训练数据产品收入，因为训练数据产品属于某一时点履行的知识产权许可授权，因此在向客户交付训练数据产品并经客户验收确认后确认收入。对于训练数据相关的应用服务收入，应用程序的定制开发或者授权许可、配套硬件的销售属于可明确区分的单项履约义务，并且属于在某一时点履行的履约义务，因此均是在向客户交付并经客户验收确认时确认收入，并按照合同约定的单项履约义务的交易价格确认收入金额。

发行人执行新收入准则后，按照上述收入确认会计政策进行相关会计处理，符合新收入准则的相关规定。

4、关于非直接销售

根据问询回复，发行人存在通过贸易企业向最终客户销售训练数据产品的情形。

请结合说明事项补充披露非直接销售模式的相关内容、不同销售模式下销售收入及其占比。

请发行人说明：（1）结合相关合同条款和实际销售过程，说明发行人与贸易企业、最终客户的具体合作模式，收入确认时点及取得的具体收入确认依据，是否符合企业会计准则的规定，是否存在提前确认收入的情形；（2）不同模式下各类业务的销售收入及其占比、毛利率对比情况及差异原因，采用非直销模式是否具有商业合理性，是否符合行业惯例；（3）报告期各期非直销模式下主要客户的销售情况，包括客户名称、最终客户、销售内容及其销售收入和占比；（4）是否存在采用不同销售模式向最终客户进行销售的情形，说明原因及商业合理性。

请申报会计师对上述事项进行核查并发表明确意见。

答复：

发行人补充披露事项

发行人已结合说明事项在招股说明书“第六节/一、发行人主营业务、主要产品或服务的情况/（三）主要经营模式/6、销售模式”中进行了补充披露，具体如下：

（4）签约模式

报告期内，发行人均采用直销方式直接对接客户并销售，不存在采用经销模式通过经销商对外销售产品或服务的情况。

其中，仅在训练数据产品业务中，存在少量与最终客户非直接签约的情形。在该种情形下，最终客户出于其自身经营管理需要，指定其认可的第三方主体与发行人签约，发行人根据销售合同约定，直接向最终客户指定的签约主体交付训练数据产品，并由该指定签约主体对交付物进行验收、确认及按照合同约定支付款项。报告期内，仅4家客户涉及该种情形，且该情形收入占比很低。

上述客户指定第三方签约模式主要是客户自身经营管理需要所致，例如：部分最终客户的内部流程较为繁琐，出于预算计划安排、缩短项目周期、便捷采购等角度考虑，其指定认可的第三方签约主体向发行人执行采购。该模式为发行人应最终客户要求形成，具备商业合理性，符合行业惯例。

报告期内，不同签约模式的收入及占比情况如下：

金额单位：万元

签约形式	2020年1-9月		2019年		2018年		2017年	
	销售收入	收入占比	销售收入	收入占比	销售收入	收入占比	销售收入	收入占比
直接签约	13,431.23	94.07%	22,932.74	96.54%	19,265.77	100.00%	11,627.57	97.65%
客户指定第三方签约	846.13	5.93%	823.07	3.46%	-	-	279.52	2.35%
合计	14,277.35	100.00%	23,755.81	100.00%	19,265.77	100.00%	11,907.09	100.00%

发行人说明事项

(1) 结合相关合同条款和实际销售过程，说明发行人与贸易企业、最终客户的具体合作模式，收入确认时点及取得的具体收入确认依据，是否符合企业会计准则的规定，是否存在提前确认收入的情形；

报告期内，发行人均采用直销方式销售产品及服务，不存在采用经销模式通过经销商对外销售产品、服务的情况。发行人一般直接与最终客户签订合同，但在训练数据产品销售的少量情况中，最终客户出于自身安排或考虑，会指定相关的贸易企业等其认可的第三方主体与发行人签订训练数据产品销售合同并进行交易，因此存在少量的非直接签约（即客户指定第三方签约）的情况。2017年、2018年、2019年及2020年1-9月，发行人非直接签约收入占比很低，分别为2.35%、0%、3.46%及5.93%。

一、发行人与贸易企业、最终客户的具体合作模式

在客户指定第三方签约的情况下，训练数据产品销售合同的签约双方为最终客户指定的贸易企业等第三方主体和发行人，合同条款规定：发行人向该签约主体销售相关训练数据产品的使用许可，训练数据产品的最终使用方为该最终客户，签约主体无权再将该训练数据产品作为独立产品进行售卖。发行人向该签约主体交付训练数据产品，由其对交付的训练数据产品进行确认和验收以及向发行人支付合同约定款项。

在实际销售过程中，发行人直接对接、获取最终客户，与其直接沟通训练数据产品的具体需求；并结合客户需求，参考公司针对训练数据产品制定的价格区间向客户报价，与最终客户协商确定产品价格。

在最终的合同签署及执行环节，发行人依照最终客户的要求，与其指定的相关贸易企业等第三方签约主体签订销售合同。后续发行人根据销售合同约定，直接向最终客户指定的签约主体交付训练数据产品，并由该指定签约主体对交付物进行验收、确认及按照合同约定支付款项。前述签约主体由最终客户指定，最终客户认可其与发行人签约并执行对产品的验收程序、履行对发行人的支付义务。

二、收入确认时点及取得的具体收入确认依据

对于此类非直接签约方式销售训练数据产品的情况，发行人在最终客户指定的签约主体对交付产品进行验收后确认收入，具体收入确认依据为发行人获得的验收邮件（附有数据验收证明），列明训练数据产品名称及编号、产品内容及数据量、验收信息及完成情况说明等具体验收信息。

对发行人而言，不同签约模式仅为销售流程中最终客户关于签约主体方面做出的不同安排，并不影响其产品的销售模式以及具体生产、交付及验收方式，客户指定第三方签约的模式下发行人训练数据产品销售收入的确认方式、确认依据与直接签约模式一致，不存在差异。

三、收入确认符合企业会计准则的规定，不存在提前确认收入的情形

报告期内，发行人上述收入确认的时点符合企业会计准则的规定。2020年1月1日前，发行人执行财政部于2006年颁布的《企业会计准则第14号——收入》及《企业会计准则第15号——建造合同》（统称“原收入准则”）。在原收入准则下，发行人以风险报酬转移作为收入确认时点的判断标准。发行人销售商品收入在同时满足下列条件时予以确认，即：商品所有权上的主要风险和报酬已转移给购货方，收入的金额及相关成本能够可靠计量，相关的经济利益很可能流入发行人，发行人既没有保留通常与所有权相联系的继续管理权，也没有对已售出的商品实施有效控制。在非直接签约方式下，发行人直接与最终客户指定的签约主体签订合同，并按照合同约定向签约主体交付产品，由其进行验收确认。根据合同约定，发行人在交付产品并经签约主体验收确认后，即具有收款的权利。签约主体由最终客户指定，其对产品的验收受最终客户认可，发行人在向签约主体交付产品并取得验收确认后，与产品相关的主要风险和报酬已经转移，满足原收入准则规定的销售商品收入确认的相关条件。该收入确认的时点符合原收入准则的规定。

自2020年1月1日起，发行人执行财政部于2017年颁布的《企业会计准则第14号——收入（修订）》（统称“新收入准则”）。在新收入准则下，发行人以控制权转移作为收入确认时点的判断标准。发行人在履行了合同中的履约义务，即在客户取得

相关商品或服务的控制权时确认收入。在非直接销售方式下，签约主体由最终客户指定，其对产品的验收受最终客户认可，发行人在向签约主体交付产品并经其验收确认后已经完成了合同约定的履约义务，产品的控制权已经转移，因此在该时点确认收入符合新收入准则的规定。

综上所述，发行人在非直接销售方式下的收入确认时点符合企业会计准则的规定。

因此，发行人在实际账务处理中，根据最终客户指定的签约、验收及付款主体出具的验收邮件作为收入确认的依据，验收依据充分；签约主体对交付的产品进行验收受客户认可。发行人不存在提前确认收入的情况。报告期内，不存在签约主体验收后，最终客户不认可签约主体验收结果而向发行人要求修改产品或赔偿的情况。

(2) 不同模式下各类业务的销售收入及其占比、毛利率对比情况及差异原因，采用非直销模式是否具有商业合理性，是否符合行业惯例；

一、仅训练数据产品业务涉及少量客户指定第三方与发行人签约的情况，不同签约模式下毛利率不存在差异

报告期内，在训练数据定制服务、训练数据产品、训练数据相关的应用服务三类业务中，仅训练数据产品业务类别存在少量最终客户指定第三方与发行人签约的情况。

各业务类别在不同签约模式下的销售收入及占比、毛利率情况如下：

金额单位：万元

业务类别及签约形式	2020年1-9月			2019年		
	销售收入	收入占比	毛利率	销售收入	收入占比	毛利率
训练数据产品	7,013.88	49.13%	100.00%	9,176.47	38.63%	100.00%
客户直接签约	6,167.75	43.20%	100.00%	8,353.39	35.16%	100.00%
客户指定第三方签约	846.13	5.93%	100.00%	823.07	3.46%	100.00%
训练数据定制服务	6,962.82	48.77%	38.73%	14,232.91	59.91%	51.12%
客户直接签约模式	6,962.82	48.77%	38.73%	14,232.91	59.91%	51.12%

训练数据相关的应用服务	300.66	2.11%	49.72%	346.44	1.46%	68.30%
客户直接签约模式	300.66	2.11%	49.72%	346.44	1.46%	68.30%
合计	14,277.35	100.00%	69.06%	23,755.81	100.00%	70.25%
业务类别及销售模式	2018年			2017年		
	销售收入	收入占比	毛利率	销售收入	收入占比	毛利率
训练数据产品	6,601.67	34.27%	100.00%	5,489.31	46.10%	100.00%
客户直接签约	6,601.67	34.27%	100.00%	5,209.79	43.75%	100.00%
客户指定第三方签约	-	-	100.00%	279.52	2.35%	100.00%
训练数据定制服务	12,369.55	64.20%	45.64%	6,297.23	52.89%	47.87%
客户直接签约	12,369.55	64.20%	45.64%	6,297.23	52.89%	47.87%
训练数据相关的应用服务	294.55	1.53%	78.70%	120.55	1.01%	82.90%
客户直接签约	294.55	1.53%	78.70%	120.55	1.01%	82.90%
合计	19,265.77	100.00%	64.77%	11,907.09	100.00%	72.26%

在直接签约、客户指定第三方签约的两种情形下，发行人均与最终客户直接对接沟通产品需求，后续应客户要求与其指定的签约主体签署合同并向其交付。对发行人而言，不同签约模式仅为销售流程中最终客户关于签约主体方面做出的不同安排，并不影响发行人产品的具体生产、交付过程，也不影响最终毛利率。两种情形下，训练数据产品业务的毛利率均为100%，不存在差异。

二、客户指定第三方签约的模式具备商业合理性，符合行业惯例

客户指定第三方签约这一模式主要是客户自身经营管理需要所致，如部分最终客户的内部流程较为繁琐，出于预算计划安排、缩短项目周期、便捷采购等角度考虑，其指定认可的贸易企业等第三方签约主体向发行人执行采购。

该模式为发行人应最终客户要求形成，具备商业合理性，符合行业惯例。

(3) 报告期各期非直销模式下主要客户的销售情况，包括客户名称、最终客户、销售内容及其销售收入和占比；

报告期内，客户指定第三方签约模式涉及的客户及其销售情况如下：

金额单位：万元

直接签约主体	最终客户	营业收入				销售内容
		2020年 1-9月	2019年	2018年	2017年	
客户 1	最终客户 1	490.48	136.79	-	97.36	智能语音训练数据产品
客户 2	最终客户 2	210.52	666.66	-	-	智能语音、自然语言训练数据产品
Tegara Corporation	最终客户 4、最终客户 1	145.13	19.63	-	-	智能语音训练数据产品
Adecco Korea Co., Ltd.	最终客户 3	-	-	-	182.17	智能语音训练数据产品
合计		846.13	823.07	-	279.52	
占营业收入比重		5.93%	3.46%	-	2.35%	

报告期内，客户指定第三方签约模式涉及的客户数量、收入规模及占比均较低，报告期内各期占比均不超过6%。

(4) 是否存在采用不同销售模式向最终客户进行销售的情形，说明原因及商业合理性。

报告期内，对于同一最终客户主体，发行人不存在采用两种签约模式同时向其进行销售的情形。但对于极少数客户品牌（报告期内仅3个），发行人存在对该品牌下的个别主体采用非直接签约模式、对其他主体采用直接签约模式销售的情形。前述情形产生的原因主要是：相关客户品牌均为大型机构，在全球多地均有运营主体，其中部分主体决策相对独立，出于自身经营管理需要（考虑内部审批流程、预算分配管理、项目周期等因素），指定其认可的第三方与发行人签约、验收发行人提供的产品并向发行人支付具备商业合理性。

申报会计师核查事项

(1) 核查程序

申报会计师执行的主要程序如下：

1、访谈发行人业务人员、主要管理人员，了解发行人与贸易企业等第三方签约主体、最终客户的具体合作模式，检查发行人与第三方签约主体的销售合同以及涉及最终客户的相关合同条款，检查发行人与第三方签约主体以及最终客户的沟通

记录，了解该合作模式的商业合理性、确认其是否符合行业惯例；

2、检查第三方签约主体的背景信息，分析是否存在异常；

3、获取并检查发行人与第三方签约主体的销售合同，识别与风险、报酬转移或控制权转移相关的合同或订单条款，检查发行人的收入确认政策是否符合企业会计准则的相关要求；

4、在抽样的基础上，将发行人与第三方签约主体的各期记录的收入核对至相关的销售合同或订单、验收文件、发票等相关支持性依据，以评价相关收入是否按照发行人的收入确认政策予以确认；

5、获取报告期各期发行人向贸易企业等第三方签约主体销售的收入明细，检查其与同类业务下其他销售合同相比，收入和毛利率是否存在明显差异；

6、获取报告期各期通过第三方签约主体向发行人采购的最终客户明细，并与发行人各期的销售合同签约主体进行核对，检查发行人是否存在以不同签约方式向同一最终客户或客户品牌进行销售的情形，并了解其商业合理性。

(2) 核查结论

经核查，申报会计师认为：

1、报告期内，发行人与贸易企业等第三方签约主体、最终客户的合作模式符合实际情况、具备商业合理性，收入确认符合企业会计准则的规定，不存在提前确认收入的情形；

2、报告期内，发行人仅训练数据产品业务涉及少量客户指定第三方签约的情形，该情形收入占比较低；其他业务均直接与最终客户签约。与最终客户直接签约、与客户指定的第三方签约两种模式下的销售收入、毛利率情况不存在差异，与客户指定的第三方签约这一模式具有商业合理性，符合行业惯例；

3、报告期内，对于同一最终客户主体，发行人不存在采用两种签约模式同时向其进行销售的情形；但对于少数客户品牌，发行人存在对该品牌下的个别主体采用非直接签约模式、对其他主体采用直接签约模式销售的情形，该模式具有商业合

理性。

5、关于软件产品增值税即征即退

根据问询回复，发行人训练数据产品及其相关的应用服务收入享受软件产品增值税政策优惠。

请发行人说明：（1）软件产品具体的用途和实现功能，与训练数据产品及其相关的应用服务的具体关系；（2）结合实际销售过程、产品服务的交付内容和形式说明前述产品和服务符合软件产品增值税政策优惠的原因及依据，是否存在以软件产品的名义销售不符合相关优惠政策要求的产品，是否符合税收相关法律法规的规定；（3）模拟测算提供特许权使用服务情形下对发行人经营成果的具体影响。

请申报会计师对上述事项进行核查并发表明确意见。

答复：

发行人说明事项

（1）软件产品具体的用途和实现功能，与训练数据产品及其相关的应用服务的具体关系；

发行人训练数据产品相关的软件产品的功能和具体用途如下：

软件功能	具体用途
结构设计	支持对相同类型训练数据产品的结构统一设计与直观展示，例如针对语音识别训练数据产品会限定数据集结构、数据格式定义以及相关技术文档的格式规范等，从而保证语音、图形图像、文本、词典等训练数据产品的结构一致性。
数据管理与索引	支持对训练数据产品中的数据统一管理 with 索引，提供用户根据不同关键词、不同筛选条件对训练数据产品中符合特定条件的数据进行快速定位、实时查询和按需编辑的能力。例如，用户在使用语音数据产品进行模型训练时，常常需要定位一些特定语音以调整模型参数优化数据对模型的适配。
数据在物理磁盘上的存储	支持训练数据（包括音频文件、文本文件、图像图片、视频文件等）在物理磁盘上的高效存储，支持多种数据类型的压缩与加密等功能。
定制化展示和导出	支持不同类型训练数据产品的不同展示方式，以方便用户在使用训练数据产品时可以直观快捷地查询、编辑和处理，同时支持用户根据不同的训练需求对训练数据产品进行特定的数据导出、格式转换等功能。

发行人训练数据产品在完成设计、采集、加工、质检等生产环节后，需要经过软件化过程，将训练数据通过软件架构形成可供客户存储、查阅、调取、处理的完

整软件产品进行销售。因此软件产品是训练数据产品经过软件化后的销售形态。

发行人训练数据相关的应用服务相关的软件产品的功能和具体用途如下：

软件名称	软件功能	具体用途
电子书记员系统服务软件V1.0.0	语音内容识别	法院庭审场景下，应用语音识别人工智能技术实现在庭审过程中对审判长、原告、被告、当事人、证人等各方陈述的语音内容自动实时识别成文字，提升庭审的效率，提升庭审的流畅性，大幅减轻书记员的工作强度。
智能音频采集系统软件V1.1.4	语音内容识别	同时录制多个发言人的声音，并应用语音识别人工智能技术将录制的声音实时转写成文字，并匹配发言人角色，自动过滤特定词汇。
智能会议系统软件V1.0.1	语音内容识别	针对会议记录场景，将语音实时转换成文字并记录成文档，同时支持一键投屏并人工实时修改、查找、标记、替换等功能。

发行人训练数据相关的应用服务的服务内容中包括为客户定制开发特定应用领域的 AI 算法模型，并通常以软件形式或软硬件一体化形式进行销售，因此软件产品是训练数据相关的应用服务的销售内容或销售内容的一部分。

（2）结合实际销售过程、产品服务的交付内容和形式说明前述产品和服务符合软件产品增值税政策优惠的原因及依据，是否存在以软件产品的名义销售不符合相关优惠政策要求的产品，是否符合税收相关法律法规的规定；

根据《财政部 国家税务总局关于软件产品增值税政策的通知》财税〔2011〕100号第二条的规定：

“本通知所称软件产品，是指信息处理程序及相关文档和数据。软件产品包括计算机软件产品、信息系统和嵌入式软件产品。嵌入式软件产品是指嵌入在计算机硬件、机器设备中并随其一并销售，构成计算机硬件、机器设备组成部分的软件产品。”

发行人所销售并交付客户的软件产品，即为带有前述功能用于处理训练数据的信息处理程序及训练数据相关的文档和数据，以及在特定应用领域应用 AI 算法模型的信息处理程序及相关文档和数据，符合软件产品增值税政策第二条对于软件产品的界定。

根据《财政部 国家税务总局关于软件产品增值税政策的通知》财税〔2011〕100号第三条的规定：

“满足下列条件的软件产品，经主管税务机关审核批准，可以享受本通知规定的增值税政策：

1. 取得省级软件产业主管部门认可的软件检测机构出具的检测证明材料；
2. 取得软件产业主管部门颁发的《软件产品登记证书》或著作权行政管理部门颁发的《计算机软件著作权登记证书》。”

发行人所销售并交付客户的软件产品，均已取得中国软件评测中心出具的《软件产品登记测试报告》，并均已获得国家版权局颁发的《计算机软件著作权登记证书》，符合软件产品增值税政策第三条所列条件。

税务部门对于享受软件产品增值税优惠的软件产品销售，通常要求其销售合同、发票注明已完成优惠备案的软件产品著作权编号。发行人在训练数据产品或训练数据相关的应用服务的销售过程中，与客户进行具体商业谈判后，根据谈判确定的条件与客户签订销售合同，合同中包含了训练数据产品或训练数据相关的应用服务所涉及的软件著作权编号。签订合同后，发行人将训练数据产品或训练数据相关的应用服务涉及的软件产品以 FTP 传输、软硬件一体化发货等方式交付给客户，并向客户开具发票，发票中也包含该等软件著作权编号。因此，发行人在销售过程中的合同、发票均明确了已备案软件产品增值税优惠的软件著作权编号，符合税务部门对于享受软件产品增值税政策优惠的具体要求。

综上所述，发行人享受软件产品增值税政策优惠的训练数据产品和训练数据相关的应用服务符合相关优惠政策的要求，不存在以软件产品的名义销售不符合相关优惠政策要求的产品，发行人符合税收相关法律法规的规定。

(3) 模拟测算提供特许权使用服务情形下对发行人经营成果的具体影响。

报告期内，发行人享受软件产品收入的增值税优惠政策，即软件产品收入按 17%（2018 年 5 月 1 日起变更为 16%，2019 年 4 月 1 日起变更为 13%）税率征收增值税后，对其增值税实际税负超过 3% 的部分实行即征即退，于实际收到即征即退税款时计入当期损益。如果假设发行人的软件产品收入归类为提供特许权使用服务收入，并适用 6% 的增值税税率，则对发行人经营成果的具体影响分析如下：

金额单位：万元

项目	适用税率	计算公式	报告期合计	2020年1-9月	2019年度	2018年度	2017年度
合并报表中享受即征即退优惠政策的软件收入	17%	①	7,118.88	-	-	2,270.12	4,848.76
	16%	②	4,416.40	-	1,969.94	2,446.46	-
	13%	③	12,305.19	6,611.33	5,693.86	-	-
	小计	④=①+②+③	23,840.47	6,611.33	7,663.80	4,716.58	4,848.76
销项税(17%/16%/13%)		⑤=①*17%+②*16%+③*13%	3,516.50	859.47	1,055.39	777.35	824.29
含税合同总额		⑥=④+⑤	27,356.97	7,470.80	8,719.19	5,493.93	5,673.05
提供特许权使用服务情形下不含税收入	6%	⑦=⑥/(1+6%)	25,808.45	7,047.92	8,225.65	5,182.95	5,351.93
销项税(6%)		⑧=⑦*6%	1,548.52	422.88	493.54	310.98	321.12
实际收到即征即退税款		⑨	2,752.93	736.46	638.49	747.86	630.12
对利润总额的影响净额		⑩=⑦-(④+⑨)	-784.95	-299.87	-76.64	-281.49	-126.95
利润总额		⑪	26,139.43	5,796.79	9,135.95	7,212.51	3,994.18
占当期利润总额的比例		⑫=⑩/⑪	-3.00%	-5.17%	-0.84%	-3.90%	-3.18%

报告期内，在发行人的软件产品收入含税合同总额不变的情况下，假设为提供特许权使用服务，适用6%的增值税税率，则发行人确认的收入金额增加，同时因不再享受增值税即征即退政策而导致当期确认的其他收益金额减少。综合来看，报告期内利润总额分别减少126.95万元、281.49万元、76.64万元和**299.87万元**，占利润总额的比例分别为3.18%、3.90%、0.84%和**5.17%**。因此，特许权使用服务情形下对利润总额的影响金额及占比均较小，不会对发行人的经营成果产生重大影响。

申报会计师核查事项

(1) 核查程序

申报会计师执行的主要程序如下：

1、与发行人业务部门相关人员访谈，了解发行人软件产品具体的用途和实现功能，与训练数据产品及其相关的应用服务的具体关系；了解发行人软件销售业务类型，实际销售过程、产品服务的交付内容和形式，并执行了穿行测试和控制测试

程序；

2、获取发行人享受增值税优惠政策的软件产品清单，及相应的计算机软件著作权登记证书、增值税优惠备案回执、销售合同、发票等相关文件，检查是否符合增值税优惠政策的相关规定；

3、获取并检查提供特许权使用服务情形下对发行人经营成果具体影响的模拟测算过程。

(2) 核查结论

经核查，申报会计师认为：

1、发行人软件产品具体的用途和实现功能，与训练数据产品及其相关的应用服务之间存在匹配关系，符合发行人的实际情况；

2、发行人享受软件产品增值税政策优惠的训练数据产品和训练数据相关的应用服务符合相关优惠政策的要求，不存在以软件产品的名义销售不符合相关优惠政策要求的产品，发行人符合税收相关法律法规的规定；

3、假设发行人的软件产品收入归类为提供特许权使用服务收入，并适用 6% 的增值税税率的情形下，不会对发行人的经营成果产生重大影响。

6、关于收入季节性

根据问询回复，报告期各期发行人第四季度营业收入占比分别为 31.73%、37.11%和 42.86%，逐年上升，2019 年第四季度发行人营业收入为 10,180.58 万元，显著高于同期各季度的营业收入。

请发行人补充披露：训练数据定制服务和产品的收入季节性分布，2018 年度和 2019 年度第四季度收入及其占比较高的原因及合理性。

请发行人说明：（1）第四季度各月营业收入的分布情况，第四季度毛利率是否显著高于其他季度；（2）2018 和 2019 年度第四季度主要客户的销售内容、收入确认金额、合同签订日期、产品或服务的交付时间和交付方式、收入确认日期和具

体依据以及是否满足合同约定验收条件，是否存在提前确认收入的情形；（3）关于收入截止性制定的内部控制措施及执行情况，是否存在提前确认收入的情形。

请申报会计师对上述事项进行核查并发表明确意见。

答复：

发行人补充披露事项

训练数据定制服务和产品的收入季节性分布，2018年度和2019年度第四季度收入及其占比较高的原因及合理性。

发行人已在招股说明书“第八节 财务会计信息与管理层分析/九、经营成果分析/（一）营业收入分析/7、营业收入的季节性分布”中补充披露下述内容：

报告期内，发行人训练数据定制服务和训练数据产品的收入季节性分布如下：

单位：万元

训练数据定制服务								
项目	2020年1-9月		2019年度		2018年度		2017年度	
	金额	占比	金额	占比	金额	占比	金额	占比
第一季度	1,137.75	16.34%	2,039.75	14.33%	1,451.06	11.73%	609.45	9.68%
第二季度	2,995.03	43.01%	3,847.99	27.04%	2,142.81	17.32%	2,395.90	38.05%
第三季度	2,830.04	40.65%	2,186.35	15.36%	3,382.80	27.35%	1,079.90	17.15%
第四季度	-	-	6,158.82	43.27%	5,392.88	43.60%	2,211.98	35.13%
合计	6,962.82	100%	14,232.91	100%	12,369.55	100%	6,297.23	100%
训练数据产品								
项目	2020年1-9月		2019年度		2018年度		2017年度	
	金额	占比	金额	占比	金额	占比	金额	占比
第一季度	1,988.34	28.35%	1,082.90	11.80%	1,370.70	20.76%	598.54	10.90%
第二季度	3,772.14	53.78%	3,129.43	34.10%	1,493.57	22.62%	1,687.68	30.74%
第三季度	1,253.40	17.87%	1,273.88	13.88%	2,075.88	31.44%	1,679.49	30.60%
第四季度	-	-	3,690.27	40.21%	1,661.52	25.17%	1,523.61	27.76%
合计	7,013.88	100%	9,176.47	100%	6,601.67	100%	5,489.31	100%

发行人部分年度的第四季度收入及其占比较高，主要原因是部分客户的年度预

算制定及结算需求和项目研发周期安排，使得其在第四季度会相对更为集中地进行验收。2018年，发行人训练数据定制服务在第四季度的收入金额及其占比较高，是由于部分主要客户的智能语音项目验收所致；2019年，发行人训练数据定制服务在第四季度的收入金额及其占比较高，是由于个别主要客户的语音合成项目验收所致，训练数据产品在第四季度的收入金额及其占比较高，是由于部分客户购买语音识别训练数据产品所致。

发行人说明事项

(1) 第四季度各月营业收入的分布情况，第四季度毛利率是否显著高于其他季度；

2017-2019年，发行人第四季度各月营业收入情况如下：

单位：万元

月份	2019年度第四季度		2018年度第四季度		2017年度第四季度	
	金额	占比	金额	占比	金额	占比
10月	1,050.43	10.32%	1,016.57	14.22%	281.51	7.45%
11月	4,669.14	45.86%	3,176.47	44.43%	1,910.48	50.57%
12月	4,461.00	43.82%	2,956.18	41.35%	1,585.73	41.98%
合计	10,180.58	100.00%	7,149.23	100.00%	3,777.72	100.00%

2017-2019年，发行人各季度的毛利率如下：

月份	2019年度	2018年度	2017年度
第一季度	62.61%	69.15%	71.73%
第二季度	72.90%	63.77%	68.21%
第三季度	64.17%	69.48%	80.37%
第四季度	72.85%	59.95%	70.95%
合计	70.25%	64.77%	72.26%

2017-2019年，发行人第四季度毛利率根据具体期间的完成项目和产品销售情况而有所波动，不存在显著高于其他季度的情形。

(2) 2018 和 2019 年度第四季度主要客户的销售内容、收入确认金额、合同签订日期、产品或服务的交付时间和交付方式、收入确认日期和具体依据以及是否满足合同约定验收条件，是否存在提前确认收入的情形；

发行人 2018 年和 2019 年第四季度前五大客户的前十大销售合同收入占 2018 年和 2019 年第四季度的销售收入的比例分别为 64.68% 和 59.94%。上述主要客户第四季度前十大销售合同的销售内容、收入确认金额、合同签订日期、产品或服务的交付时间和交付方式、收入确认日期如下：

2019 年度第四季度：

单位：万元

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
2019 年第四季度客户 1	合同 1	2018 年 3 月 9 日	训练数据定制服务	语音合成	2,278.28	2019 年 10 月-2019 年 11 月	FTP 传输	2019 年 11 月 20 日
	合同 2	2018 年 8 月 23 日	训练数据定制服务	语音合成	526.34	2019 年 12 月	FTP 传输	2019 年 12 月 11 日
	合同 3	2019 年 7 月 10 日	训练数据产品	语音识别	309.78	2019 年 7 月-2019 年 8 月	FTP 传输	2019 年 10 月 2 日
	合同 4	2018 年 12 月 12 日	训练数据定制服务	语音识别采集	87.08	2019 年 8 月	FTP 传输	2019 年 10 月 11 日
	合同 5	2017 年 3 月 7 日	训练数据定制服务	语音合成	79.79	2019 年 12 月	FTP 传输	2019 年 12 月 18 日
	合同 6	2017 年 12 月 5 日	训练数据定制服务	语音合成	33.68	2019 年 12 月	FTP 传输	2019 年 12 月 25 日
	合同 7	2019 年 8 月 26 日	训练数据定制服务	语音识别采集	28.21	2019 年 12 月	FTP 传输	2019 年 12 月 27 日
	合同 8	2017 年 12 月 5 日	训练数据定制服务	语音合成	23.39	2019 年 12 月	FTP 传输	2019 年 12 月 18 日
				自然语言	2.88	2019 年 2 月	FTP 传输	2019 年 12 月 30 日
合同 9	2019 年 2 月 5 日	训练数据定制服务	语音识别采集	24.95	2019 年 10 月	FTP 传输	2019 年 11 月 26 日	

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
	合同 10	2016 年 11 月 5 日	训练数据定制服务	语音合成	21.32	2019 年 12 月*	FTP 传输	2019 年 12 月 26 日、 2019 年 12 月 30 日
				自然语言	2.04	2019 年 8 月	FTP 传输	2019 年 10 月 29 日
2019 年第四季度客户 2	合同 11	2019 年 6 月 7 日	训练数据产品	语音识别	475.72	2019 年 7 月- 2019 年 12 月	FTP 传输	2019 年 10 月 24 日、 2019 年 12 月 24 日
				自然语言	43.82	2019 年 10 月	FTP 传输	2019 年 10 月 24 日
	合同 12	2019 年 11 月 15 日	训练数据产品	语音识别	158.93	2019 年 12 月*	FTP 传输	2019 年 12 月 24 日
	合同 13	2019 年 5 月 20 日	训练数据定制服务	语音合成	156.05	2019 年 5 月- 2019 年 12 月	FTP 传输	2019 年 10 月 21 日、 2019 年 11 月 6 日、 2019 年 11 月 11 日、 2019 年 11 月 26 日、 2019 年 11 月 27 日、 2019 年 12 月 23 日、 2019 年 12 月 24 日、 2019 年 12 月 25 日、 2019 年 12 月 30 日
	合同 14	2018 年 2 月 15 日	训练数据定制服务	语音合成	89.86	2019 年 7 月- 2019 年 12 月	FTP 传输	2019 年 12 月 31 日
	合同 15	2019 年 4 月 26 日	训练数据定制服务	语音合成	79.85	2019 年 10 月	FTP 传输	2019 年 10 月 17 日
	合同 16	2017 年 3 月 17 日	训练数据定制服务	自然语言	72.81	2019 年 12 月	FTP 传输	2019 年 12 月 31 日
	合同 17	2019 年 10 月 28 日	训练数据产品	语音识别	25.12	2019 年 11 月- 2019 年 12 月	FTP 传输	2019 年 12 月 19 日
	合同 18	2017 年 11 月 24 日	训练数据定制服务	计算机视觉	21.52	2019 年 8 月- 2019 年 11 月	FTP 传输	2019 年 11 月 5 日、 2019 年 11 月 19 日、 2019 年 12 月 16 日、 2019 年 12 月 24 日
	合同 19	2019 年 3 月 25 日	训练数据定制服务	语音识别转写	14.99	2019 年 9 月- 2019 年 11 月	FTP 传输	2019 年 11 月 5 日、 2019 年 11 月 13 日
自然语言				2.21	2019 年 11 月	FTP 传输	2019 年 12 月 24 日	

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
	合同 20	2019 年 4 月 1 日	训练数据定制服务	语音合成	13.79	2019 年 12 月	FTP 传输	2019 年 12 月 30 日
2019 年第四季度客户 3	合同 21	2017 年 11 月 21 日	训练数据产品	计算机视觉	324.07	2019 年 4 月-2019 年 12 月	FTP 传输	2019 年 11 月 7 日、2019 年 11 月 14 日、2019 年 11 月 20 日、2019 年 11 月 21 日、2019 年 12 月 24 日、2019 年 12 月 26 日
			训练数据定制服务	语音合成	219.41	2019 年 12 月*	FTP 传输	2019 年 12 月 26 日、2019 年 12 月 27 日
				计算机视觉	94.40	2019 年 7 月-2019 年 12 月	FTP 传输	2019 年 11 月 29 日、2019 年 12 月 27 日
				自然语言	13.50	2018 年 12 月	FTP 传输	2019 年 12 月 27 日
	合同 22	2019 年 8 月 5 日	训练数据产品	计算机视觉	3.30	2019 年 8 月	FTP 传输	2019 年 10 月 11 日
2019 年第四季度客户 4	合同 23	2019 年 4 月 3 日	训练数据定制服务	语音合成	121.59	2019 年 10 月-2019 年 11 月	FTP 传输	2019 年 10 月 25 日、2019 年 11 月 27 日
	合同 24	2019 年 4 月 2 日	训练数据定制服务	语音识别采集	76.42	2019 年 10 月	FTP 传输	2019 年 10 月 29 日
	合同 25	2017 年 8 月 2 日	训练数据定制服务	语音识别采集	53.26	2019 年 6 月-2019 年 9 月	FTP 传输	2019 年 10 月 23 日、2019 年 11 月 6 日
	合同 26	2017 年 8 月 2 日	训练数据定制服务	语音合成	40.95	2019 年 8 月-2019 年 11 月	FTP 传输	2019 年 11 月 7 日、2019 年 11 月 11 日、2019 年 11 月 14 日
	合同 27	2017 年 9 月 9 日	训练数据定制服务	自然语言	39.40	2019 年 8 月	FTP 传输	2019 年 10 月 11 日
	合同 28	2019 年 4 月 4 日	训练数据定制服务	自然语言	29.10	2019 年 9 月-2019 年 12 月	FTP 传输	2019 年 10 月 31 日、2019 年 11 月 4 日、2019 年 11 月 25 日、2019 年 12 月 27 日、2019 年 12 月 30 日
语音识别采集				9.38	2019 年 9 月	FTP 传输	2019 年 11 月 12 日	

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期	
	合同 29	2019 年 4 月 15 日	训练数据定制服务	语音识别采集	34.59	2019 年 10 月-2019 年 11 月	FTP 传输	2019 年 10 月 29 日、2019 年 11 月 12 日	
	合同 30	2019 年 6 月 20 日	训练数据定制服务	语音合成	27.78	2019 年 12 月	FTP 传输	2019 年 12 月 17 日	
				自然语言	2.82	2019 年 10 月	FTP 传输	2019 年 11 月 27 日	
	合同 31	2019 年 7 月 4 日	训练数据定制服务	语音识别转写	25.44	2019 年 9 月	FTP 传输	2019 年 12 月 24 日	
				语音合成	1.90	2019 年 10 月-2019 年 12 月	FTP 传输	2019 年 11 月 12 日、2019 年 12 月 25 日	
				自然语言	0.11	2019 年 12 月	FTP 传输	2019 年 12 月 27 日	
	合同 32	2019 年 5 月 10 日	训练数据定制服务	语音识别采集	22.64	2019 年 9 月	FTP 传输	2019 年 10 月 23 日	
	2019 年第四季度客户 5	合同 33	2019 年 1 月 1 日	训练数据定制服务	语音识别转写	149.24	2019 年 8 月-2019 年 12 月	FTP 传输	2019 年 11 月 1 日、2019 年 11 月 6 日、2019 年 11 月 14 日、2019 年 12 月 6 日、2019 年 12 月 23 日
					语音合成	84.64	2019 年 5 月-2019 年 12 月	FTP 传输	2019 年 10 月 29 日、2019 年 11 月 5 日、2019 年 11 月 11 日、2019 年 12 月 18 日、2019 年 12 月 24 日、2019 年 12 月 25 日
					语音识别采集	47.17	2019 年 10 月	FTP 传输	2019 年 11 月 7 日
自然语言					10.22	2019 年 5 月-2019 年 10 月	FTP 传输	2019 年 11 月 19 日、2019 年 11 月 21 日、2019 年 12 月 26 日	
计算机视觉					0.87	2019 年 7 月	FTP 传输	2019 年 11 月 5 日	
训练数据产品					自然语言	16.98	2019 年 9 月	FTP 传输	2019 年 10 月 30 日
		语音识别	4.72	2019 年 9 月	FTP 传输	2019 年 10 月 12 日			

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
	合同 34	2019 年 1 月 1 日	训练数据定制服务	语音合成	32.48	2019 年 11 月-2019 年 12 月	FTP 传输	2019 年 11 月 18 日、2019 年 12 月 24 日、2019 年 12 月 25 日
			训练数据产品	语音合成	19.82	2019 年 11 月	FTP 传输	2019 年 11 月 28 日
				语音识别	13.27	2019 年 12 月	FTP 传输	2019 年 12 月 23 日
				计算机视觉	6.19	2019 年 11 月	FTP 传输	2019 年 12 月 24 日
	合同 35	2018 年 11 月 23 日	训练数据定制服务	计算机视觉	3.31	2019 年 10 月	FTP 传输	2019 年 10 月 24 日
	合同 36	2019 年 10 月 31 日	训练数据定制服务	语音识别转写	0.49	2019 年 11 月	FTP 传输	2019 年 11 月 20 日

注1：带有*号的交付时间表示当月有多次交付。

注2：对于同一合同下同一类销售内容的所有项目合并披露。

2018年度第四季度：

单位：万元

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
2018 年第四季度客户 1	合同 1	2018 年 6 月 29 日	训练数据产品	语音识别	629.84	2018 年 7 月*	FTP 传输	2018 年 10 月 23 日、2018 年 10 月 24 日
			训练数据定制服务	自然语言	3.26	2018 年 10 月	FTP 传输	2018 年 10 月 23 日
	合同 2	2018 年 1 月 3 日	训练数据定制服务	语音识别采集	356.60	2018 年 5 月	FTP 传输	2018 年 11 月 14 日
				语音合成	178.22	2018 年 11 月-2018 年 12 月	FTP 传输	2018 年 12 月 13 日、2018 年 12 月 25 日
				语音识别转写	63.55	2018 年 6 月-2018 年 8 月	FTP 传输	2018 年 10 月 10 日、2018 年 10 月 11 日
				自然语言	24.01	2018 年 7 月-2018 年 9 月	FTP 传输	2018 年 11 月 30 日、2018 年 12 月 13 日
	合同 3	2017 年 8 月 2 日	训练数据定制服务	语音识别转写	300.30	2018 年 6 月-2018 年 8 月	FTP 传输	2018 年 12 月 12 日、2018 年 12 月 23 日

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
				自然语言	27.34	2018年5月-2018年9月	FTP传输	2018年10月15日、2018年11月12日、2018年11月26日
	合同4	2017年8月2日	训练数据定制服务	语音合成	119.04	2018年12月	FTP传输	2018年12月25日
				语音识别转写	92.73	2018年7月-2018年10月	FTP传输	2018年10月17日、2018年10月31日、2018年11月30日
	合同5	2018年8月7日	训练数据定制服务	计算机视觉	135.90	2018年10月	FTP传输	2018年11月5日
	合同6	2018年5月14日	训练数据定制服务	语音识别转写	122.32	2018年5月-2018年10月	FTP传输	2018年11月27日、2018年12月12日
				自然语言	6.89	2018年7月-2018年10月	FTP传输	2018年10月12日、2018年10月18日、2018年11月30日
				语音识别采集	2.08	2018年5月-2018年9月	FTP传输	2018年10月15日、2018年11月12日、2018年11月20日
				语音合成	0.33	2018年8月	FTP传输	2018年10月29日
				计算机视觉	0.23	2018年9月	FTP传输	2018年10月24日
	合同7	2018年3月12日	训练数据定制服务	语音合成	96.38	2018年10月	FTP传输	2018年11月15日
				自然语言	1.98	2018年5月	FTP传输	2018年11月15日
	合同8	2018年11月20日	训练数据定制服务	计算机视觉	95.38	2018年12月*	FTP传输	2018年12月21日、2018年12月25日
	合同9	2018年7月23日	训练数据定制服务	计算机视觉	94.34	2018年10月	FTP传输	2018年11月5日
	合同10	2017年6月5日	训练数据定制服务	语音合成	93.21	2018年6月-2018年11月	FTP传输	2018年11月21日
2018年第四季度客户2	合同11	2018年4月18日	训练数据定制服务	计算机视觉	232.89	2018年8月-2018年12月	FTP传输	2018年10月9日、2018年10月24日、2018年11月28日、

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
								2018年12月13日
	合同 12	2018年7月3日	训练数据产品	语音识别	180.12	2018年9月-2018年10月	FTP 传输	2018年11月19日
	合同 13	2017年6月30日	训练数据定制服务	语音合成	102.40	2018年10月-2018年11月	FTP 传输	2018年11月22日
	合同 14	2018年4月9日	训练数据定制服务	计算机视觉	94.34	2018年12月	FTP 传输	2018年12月26日
	合同 15	2017年5月19日	训练数据定制服务	语音合成	83.15	2018年2月-2018年11月	FTP 传输	2018年12月27日、2018年12月28日
				自然语言	1.16	2018年8月	FTP 传输	2018年11月12日
	合同 16	2018年4月9日	训练数据定制服务	自然语言	67.45	2018年12月	FTP 传输	2018年12月26日
	合同 17	2018年4月9日	训练数据定制服务	计算机视觉	29.25	2018年12月	FTP 传输	2018年12月26日
	合同 18	2017年11月24日	训练数据定制服务	计算机视觉	8.65	2018年9月-2018年11月	FTP 传输	2018年10月23日、2018年11月21日、2018年12月3日
	合同 19	2018年10月25日	训练数据定制服务	自然语言	5.66	2018年12月	FTP 传输	2018年12月12日
	合同 20	2016年6月27日	训练数据定制服务	自然语言	3.51	2018年9月-2018年12月	FTP 传输	2018年11月28日、2018年12月13日、2018年12月24日、2018年12月25日
2018年第四季度客户 3	合同 21	2018年3月9日	训练数据定制服务	语音合成	653.74	2018年12月*	FTP 传输	2018年12月26日
	合同 22	2015年4月8日	训练数据定制服务	语音识别转写	1.85	2018年10月	FTP 传输	2018年11月1日
2018年第四季度客户 4	合同 23	2018年8月7日	训练数据产品	语音识别	124.14	2018年8月	FTP 传输	2018年11月7日
				自然语言	28.71	2018年8月	FTP 传输	2018年10月9日
	合同 24	2018年4月11日	训练数据定制服务	语音识别采集	73.30	2018年7月	FTP 传输	2018年11月13日

客户名称	合同	合同签订日期	销售内容		收入金额	产品和服务的交付时间	交付方式	收入确认日期
	合同 25	2018 年 8 月 8 日	训练数据产品	语音识别	72.41	2018 年 8 月	FTP 传输	2018 年 11 月 6 日
	合同 26	2014 年 1 月 1 日	训练数据定制服务	语音合成	46.23	2018 年 12 月	FTP 传输	2018 年 12 月 27 日
	合同 27	2018 年 7 月 15 日	训练数据定制服务	自然语言	22.64	2018 年 8 月-2018 年 12 月	FTP 传输	2018 年 11 月 5 日、2018 年 11 月 15 日、2018 年 12 月 26 日
	合同 28	2017 年 7 月 10 日	训练数据定制服务	自然语言	5.65	2018 年 4 月-2018 年 11 月	FTP 传输	2018 年 11 月 5 日、2018 年 11 月 6 日、2018 年 12 月 26 日
	合同 29	2018 年 10 月 17 日	训练数据产品	语音识别	5.43	2018 年 12 月	FTP 传输	2018 年 12 月 25 日
	合同 30	2018 年 9 月 13 日	训练数据定制服务	语音识别采集	1.42	2018 年 9 月	FTP 传输	2018 年 10 月 23 日
2018 年第四季度客户 5	合同 31	2018 年 9 月 6 日	训练数据产品	语音识别	167.68	2018 年 9 月	FTP 传输	2018 年 12 月 5 日
			训练数据定制服务	语音识别采集	42.58	2018 年 11 月	FTP 传输	2018 年 11 月 16 日
	合同 32	2018 年 10 月 26 日	训练数据产品	语音识别	125.76	2018 年 11 月	FTP 传输	2018 年 12 月 5 日

注1：带有*号的交付时间表示当月有多次交付。

注2：对于同一合同下同一类销售内容的所有项目合并披露。

上述客户的合同中对训练数据定制服务和训练数据产品的交付内容、交付时间与验收标准等进行明确规定，发行人按照合同约定的交付时间交付服务或产品，客户按照合同约定的交付内容和交付标准对发行人交付的服务或产品进行验收，并在上述服务或产品满足合同约定的交付内容和交付标准时出具验收邮件，发行人将获得的验收邮件作为收入确认的具体依据，在验收邮件记载的验收时点确认收入，不存在提前确认收入的情形。

(3) 关于收入截止性制定的内部控制措施及执行情况，是否存在提前确认收入的情形。

与收入截止性相关的内控措施具体如下：

报告期内发行人的业务人员在客户对产品或服务进行验收确认后，在业务管理平台发起验收申请，在系统中填写项目验收具体信息，包括验收数量、验收金额、验收时间等，并将交付通知、验收确认的邮件作为交付及验收依据上传至业务管理平台。业务部门负责人、财务部相关人员审核业务管理平台记载的验收信息的准确性和完整性，并检查验收信息与上传的交付及验收依据是否一致。2019年11月之前，业务管理平台尚未实现与财务系统的直接对接，因此由财务人员每月导出业务管理平台中当月的项目验收记录，核对验收时间后，将其作为财务记账的依据，并在财务系统中进行收入确认的会计处理，相关会计凭证由财务经理进行复核。2019年11月之后，业务管理平台与财务系统直接对接，实现了系统自动化传输和记账处理，每日业务管理平台提取当日验收记录自动传输至财务系统，并自动生成记账凭证。每月月初，财务人员从业务管理平台中导出验收时间在上月的验收信息明细，并与上月财务账面的收入金额进行核对，检查是否一致。

报告期内发行人已经针对收入截止性风险制定了上述内部控制措施，并确保其有效执行。报告期内发行人不存在提前确认收入的情形。

申报会计师核查事项

(1) 核查程序

申报会计师执行的主要程序如下：

1、通过与业务人员访谈，了解发行人2018年和2019年第四季度收入占比较高的原因，与发行人的业务模式进行比较，分析其合理性；

2、获取发行人报告期内第四季度各月收入情况，以及报告期内各季度的收入和毛利率情况，分析是否存在第四季度毛利率显著高于其他季度的情形，并访谈相关业务人员，了解季度间波动的原因和合理性；

3、在抽样的基础上，检查2018年和2019年第四季度记录的收入对应的销售合同或订单，识别与风险和报酬转移或控制权转移相关的合同或订单条款，评价发行人的

收入确认政策是否符合企业会计准则的相关要求；

4、在抽样的基础上，将 2018 年和 2019 年第四季度记录的收入核对至相关的销售合同或订单、客户验收文件、发票等相关支持性依据，以评价相关收入是否按照发行人的收入确认政策予以确认；

5、了解和评价与收入截止性相关的内部控制及其执行情况，并执行控制测试。

6、从资产负债表日前后的收入交易中选取样本，检查与收入确认相关的销售合同或订单、客户验收文件、发票等相关支持性依据，评价接近资产负债表日前后记录的收入是否记录于恰当期间，是否存在收入确认的截止性差异。

(2) 核查结论

经核查，申报会计师认为：

1、发行人训练数据定制服务和训练数据产品的收入季节性分布符合发行人的实际情况，2018 年度和 2019 年度第四季度收入及其占比较高的原因合理；

2、发行人 2018 年度和 2019 年度第四季度各月的收入分布情况、第四季度的毛利率情况，符合发行人的实际情况且具有合理性，发行人第四季度毛利率不存在显著高于其他季度的情形；

3、发行人 2018 年度和 2019 年度第四季度主要客户的销售内容、收入确认金额、合同签订日期、产品或服务的交付时间和交付方式、收入确认日期和具体依据均符合发行人的实际情况，收入确认满足合同约定的验收条件，发行人不存在提前确认收入的情况；

4、发行针对收入截止性制定的内部控制措施恰当并得到有效执行，发行人不存在提前确认收入的情形。

7、关于整体变更时个人所得税事项

根据问询回复：（1）中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数的自然人合伙人未就发行人整体变更缴纳完毕个人所得税；（2）发行人律师未明确回复是否符合税收法律法规的规定。

请发行人说明：整体变更时前述股东个人所得税的申报和缴纳情况，是否符合税收相关法律法规的规定。

请发行人律师和申报会计师就整体变更时涉及的个人所得税的申报和缴纳事项进行核查并发表明确意见。

答复：

发行人说明事项

根据《财政部、国家税务总局关于将国家自主创新示范区有关税收试点政策推广到全国范围实施的通知》（财税[2015]116号）第三条，“2.个人股东获得转增的股本，应按照“利息、股息、红利所得”项目，适用20%税率征收个人所得税。”

根据《国家税务总局关于<关于个人独资企业和合伙企业投资者征收个人所得税的规定>执行口径的通知》（国税函[2001]84号）第二条，“个人独资企业和合伙企业对外投资分回的利息或者股息、红利，不并入企业的收入，而应单独作为投资者个人取得的利息、股息、红利所得，按“利息、股息、红利所得”应税项目计算缴纳个人所得税。以合伙企业名义对外投资分回利息或者股息、红利的，应按《财政部 国家税务总局关于印发《关于个人独资企业和合伙企业投资者征收个人所得税的规定》的通知》（财税〔2000〕91号）所附规定的第五条¹¹精神确定各个投资者的利息、股息、红利所得，分别按“利息、股息、红利所得”应税项目计算缴纳个人所得税。”

发行人2017年11月整体变更为股份有限公司，整体变更前海天瑞声科技有限公司的注册资本为654万元，整体变更后发行人的注册资本为3,000万元，发行人存在整体变更时以未分配利润、盈余公积、资本公积转增股本的情形。发行人前述股东（中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数）的自然人合伙人应就整体变更时转增股本申报缴纳个人所得税。

根据中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数申报纳税的

¹¹ 《财政部 国家税务总局关于印发《关于个人独资企业和合伙企业投资者征收个人所得税的规定》的通知》（财税〔2000〕91号）所附规定第五条内容为：“个人独资企业的投资者以全部生产经营所得为应纳税所得额；合伙企业的投资者按照合伙企业的全部生产经营所得和合伙协议约定的分配比例确定应纳税所得额，合伙协议没有约定分配比例的，以全部生产经营所得和合伙人数量平均计算每个投资者的应纳税所得额。前款所称生产经营所得，包括企业分配给投资者个人的所得和企业当年留存的所得（利润）。”

银行回单、《税收完税证明》等，中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数的自然人合伙人已经申报并缴纳了发行人整体变更时转增股本相应的个人所得税。

综上，发行人整体变更时的合伙企业股东中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数的自然人合伙人已经就整体变更时转增股本申报并缴纳个人所得税，符合税收相关法律法规的规定。

中介机构核查事项

(1) 核查程序

发行人律师、申报会计师主要履行了如下核查程序：

- 1、查阅发行人整体变更时的工商登记档案资料、《验资报告》；
- 2、查询有关整体变更个人所得税的法律法规；
- 3、查阅贺琳、唐涤飞经北京市海淀区地方税务局第四税务所备案的《〈个人所得税分期缴纳备案表（转增股本）〉及其填报说明》，并核查发行人是否符合中小高新技术企业的条件；
- 4、向中瑞安、中瑞立、清德投资、上海丰琬、上海兴富、杭州银杏数了解其申报缴纳个人所得税的相关情况，查阅中瑞安、中瑞立、清德投资、上海丰琬、上海兴富、杭州银杏数申报缴纳的个人所得税的银行回单、《税收完税证明》等资料。

(2) 核查结论

经发行人律师和申报会计师核查，发行人整体变更时涉及申报缴纳个人所得税的股东包括自然人股东贺琳、唐涤飞以及合伙企业股东中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数的自然人合伙人。其中，自然人股东贺琳、唐涤飞已办理个人所得税分期缴纳备案，符合税收相关法律法规的规定；中瑞安、中瑞立、清德投资、上海丰琬、上海兴富和杭州银杏数的自然人合伙人已经申报并缴纳了发行人整体变更时转增股本相应的个人所得税，符合税收相关法律法规的规定。

8、关于函证程序

根据问询回复，保荐机构和申报会计师存在通过邮寄和电子邮件两种方式收发函证的情形。

请保荐机构和申报会计师说明：（1）通过电子邮件方式收发函证的原因，发函和回函方式不一致的函证情况，涉及的客户及其收入和往来情况；（2）针对上述情形履行的具体函证控制程序，是否符合相关准则和问题解答的要求。

答复：

中介机构说明事项

（1）通过电子邮件方式收发函证的原因，发函和回函方式不一致的函证情况，涉及的客户及其收入和往来情况；

一、电子邮件收发函证

（一）通过电子邮件收发函证的原因

通过电子邮件发送函证的原因：受境内外新冠病毒疫情影响，对于少量海外客户或受疫情影响无法纸质寄达的国内客户，中介机构通过电子邮件方式发送函证¹²。

通过电子邮件回收函证的原因：①受境内外新冠病毒疫情影响，少量海外客户或受疫情影响不便寄送纸质回函的国内客户以电子邮件形式回复函证；②考虑到被询证者处理询证函的习惯做法及回函的可能性，中介机构在寄送的纸质函证中载明了电子联系方式，或在寄送纸质函证的同时追加电子邮件函证，客户选择以电子邮件方式直接向中介机构回函。

除**2020年1-9月**外，报告期内涉及电子邮件收发函的情形占比较低。**2020年前三季度**电子邮件函证及回函比例有所提高主要是境内外新冠病毒疫情影响所致。

（二）通过电子邮件收/发函证涉及的客户及其收入和往来情况

中介机构涉及电子邮件收发函的共有下述两种情形：

¹² 因保荐机构参与了发行人两次科创板申报工作、申报会计师仅参与了发行人本次科创板申报工作，导致双方针对2017及2018年函证的发函时间不同，所以发函回函形式存在略微差异。保荐机构2017年、2018年函证均在2019年发送；而申报会计师2017年-2019年函证则均在2020年发送，受函证寄发时新冠病毒疫情影响，2017年、2018年函证也存在电子邮件形式发放的情形。

电子邮件收发函证的情形	发函方式	回函方式
情形 1	电子邮件	电子邮件回函或未回函
情形 2	纸质寄送	电子邮件

注：对于同时发送了纸质函证、电子邮件函证且被询证对象仅以电子邮件形式回复函证的情况，计入情形1核算。

情形1涉及的客户及其收入和往来情况具体请见下方列示，情形2涉及的客户及其收入、往来情况请参见下文“二、发函回函方式不一致的函证情况”。

1、情形1

(1) 保荐机构函证涉及的客户情况

金额单位：万元

序号	客户名称	2020年1-9月		2020年9月30日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	732.91	5.13%	127.70	3.35%
2	Oracle America, Inc.	279.91	1.96%	90.78	2.38%
3	某大学	251.57	1.76%	-	-
4	NTT DOCOMO, INC.	242.51	1.70%	0.01	-
5	客户 3	217.42	1.52%	-	-
6	客户 2	210.52	1.47%	-	-
合计		1,934.84	13.55%	218.49	5.73%
序号	客户名称	2019年		2019年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	Oracle America, Inc.	312.40	1.32%	-	-
合计		312.40	1.32%	-	-

注1：上表中列示的2020年9月30日应收账款余额中包括合同资产；

注2：保荐机构2017及2018年函证不涉及上述情形。

(2) 申报会计师函证涉及的客户情况

金额单位：万元

序号	客户名称	2020年1-9月		2020年9月30日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	732.91	5.13%	127.70	3.35%
2	Oracle America, Inc.	279.91	1.96%	90.78	2.38%
3	某大学	251.57	1.76%	-	-
4	NTT DOCOMO, INC.	242.51	1.70%	0.01	-

5	客户 3	217.42	1.52%	-	-
6	客户 2	210.52	1.47%	-	-
合计		1,934.84	13.55%	218.49	5.73%
序号	客户名称	2019 年		2019 年 12 月 31 日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	Oracle America, Inc.	312.40	1.32%	-	-
合计		312.40	1.32%	-	-
序号	客户名称	2018 年		2018 年 12 月 31 日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	Qualcomm Technologies, Inc.	163.95	0.85%	-	-
合计		163.95	0.85%	-	-
序号	客户名称	2017 年		2017 年 12 月 31 日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	西北工业大学	133.89	1.12%	78.33	2.89%
合计		133.89	1.12%	78.33	2.89%

注1：上表中列示的2020年9月30日应收账款余额中包括合同资产。

注2：因保荐机构参与了发行人两次科创板申报工作、申报会计师仅参与了发行人本次科创板申报工作，导致双方针对2017及2018年函证的发函时间不同，发函回函形式存在略微差异。保荐机构2017年、2018年函证均在2019年发送；而申报会计师2017年-2019年函证则均在2020年发送，受函证寄发时新冠病毒疫情影响，2017年、2018年函证也存在电子邮件形式发放的情形。

二、发函回函方式不一致的函证情况

报告期内发函和回函方式不一致的函证情况为：

发函方式	回函方式
纸质邮寄	电子邮件

前述情形产生的原因主要是：①受境内外新冠病毒疫情影响，少量客户虽已形成纸质回函文件，但无法及时进行邮寄，其通过电子邮件将回函直接发送至中介机构的邮箱；②少量被询证方处理询证函的惯例做法为电子邮件形式回函。

报告期内，该类情形涉及的客户对应的收入、往来金额及占比均较小，相关客户及其收入、往来情况具体如下：

1、保荐机构函证涉及的客户情况

金额单位：万元

序号	客户名称	2020年1-9月		2020年9月30日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	北京小米智能科技有限公司	233.22	1.63%	214.10	5.62%
2	北京小米移动软件有限公司	107.95	0.76%	67.34	1.77%
3	Samsung R&D Institute India-Bangalore Pvt. Ltd.	97.13	0.68%	4.59	0.12%
4	科大讯飞股份有限公司	54.72	0.38%	-	-
5	北京吉云互动科技有限公司	49.22	0.34%	41.86	1.10%
合计		542.24	3.79%	327.89	8.61%
序号	客户名称	2019年		2019年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	920.39	3.87%	-	-
2	Samsung R&D Institute India-Bangalore Pvt. Ltd.	275.90	1.16%	-	-
3	北京联想软件有限公司	185.84	0.78%	80.93	1.74%
合计		1,382.13	5.81%	80.93	1.74%
序号	客户名称	2018年		2018年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	NAVER Corporation	327.93	1.70%	46.09	1.49%
2	Info-communications Media Development Authority	327.05	1.70%	108.90	3.52%
3	某国际消费电子产品厂商	278.97	1.45%	30.88	1.00%
4	Capio, Inc.	268.32	1.39%	0.96	0.03%
5	Qualcomm Technologies, Inc.	163.95	0.85%	-	-
6	Advanced Media, Inc.	28.91	0.15%	0.18	0.01%
合计		1,395.13	7.24%	187.02	6.04%
序号	客户名称	2017年		2017年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	580.16	4.87%	-	-
2	NAVER Corporation	353.33	2.97%	-	-
3	Adecco Korea Co., Ltd.	182.17	1.53%	-	-
4	Capio, Inc.	117.14	0.98%	-	-
5	Advanced Media, Inc.	87.02	0.73%	26.64	0.98%

合计	1,319.82	11.08%	26.64	0.98%
----	----------	--------	-------	-------

注1：上表中列示的2020年9月30日应收账款余额中包括合同资产；

注2：因保荐机构参与了发行人两次科创板申报工作、申报会计师仅参与了发行人本次科创板申报工作，导致双方针对2017及2018年函证的发函时间不同，所以发函回函情况存在略微差异。

2、申报会计师函证涉及的客户情况

金额单位：万元

序号	客户名称	2020年1-9月		2020年9月30日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	北京小米智能科技有限公司	233.22	1.63%	214.10	5.62%
2	北京小米移动软件有限公司	107.95	0.76%	67.34	1.77%
3	Samsung R&D Institute India-Bangalore Pvt. Ltd.	97.13	0.68%	4.59	0.12%
4	科大讯飞股份有限公司	54.72	0.38%	-	-
5	北京吉云互动科技有限公司	49.22	0.34%	41.86	1.10%
合计		542.24	3.79%	327.89	8.61%
序号	客户名称	2019年		2019年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	920.39	3.87%	-	-
2	Samsung R&D Institute India-Bangalore Pvt. Ltd.	275.90	1.16%	-	-
3	北京联想软件有限公司	185.84	0.78%	80.93	1.74%
合计		1,382.13	5.82%	80.93	1.74%
序号	客户名称	2018年		2018年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	278.97	1.45%	30.88	1.00%
2	北京联想软件有限公司	205.13	1.06%	-	-
3	Samsung R&D Institute India-Bangalore Pvt. Ltd.	153.99	0.80%	-	-
合计		638.09	3.31%	30.88	1.00%
序号	客户名称	2017年		2017年12月31日	
		营业收入	营业收入占比	应收账款余额	应收账款余额占比
1	某国际消费电子产品厂商	580.16	4.87%	-	-
合计		580.16	4.87%	-	-

注1：上表中列示的2020年9月30日应收账款余额中包括合同资产；

注2：因保荐机构参与了发行人两次科创板申报工作、申报会计师仅参与了发行人本次科创板申报工作，导致双方针对2017及2018年函证的发函时间不同，所以发函回函情况存在略微差异。

(2) 针对上述情形履行的具体函证控制程序，是否符合相关准则和问题解答的要求。

根据《中国注册会计师审计准则第1312号——函证》及问题解答的相关要求：当实施函证程序时，注册会计师应当对询证函保持控制，包括：（1）确定需要确认或填列的信息；（2）选择适当的被询证者；（3）设计询证函，包括正确填列被询证者的姓名和地址，以及被询证者直接向注册会计师回函的地址等信息；（4）发出询证函并予以跟进，必要时再次向询证者寄发询证函。对以电子邮件或传真形式收到的回函，由于回函者的身份及其授权情况很难确定，对回函的更改也难以发觉，因此可靠性存在风险。注册会计师可与被询证者联系以核实回函的来源及内容。

针对上述通过电子邮件收发函证、发函和回函方式不一致（即纸质邮寄函证、电子邮件形式回收函证的形式）的情形，中介机构执行的具体函证控制程序如下：

中介机构综合考虑海天瑞声的经营环境、内部控制的有效性、账户或交易的性质、被询证者处理询证函的习惯做法及回函的可能性等，各自对海天瑞声报告期内的营业收入、应收账款的函证程序进行了恰当的设计，确定了函证的内容、范围、时间和方式。针对以电子邮件方式收发函证的情形，中介机构采取下列措施实施了控制：

一、函证程序设计控制情况

1、中介机构在电子邮件发放的询证函中列示的信息与纸质邮寄发放的询证函一致，均通常包括：截至资产负债表日对该客户的应收账款（或预收账款）余额、当年向该客户实现的销售额、当年向该客户开具的发票明细、当年收到该客户付款的明细，且已与发行人财务账套数据及有关资料核对一致，以核实公司与客户之间的交易金额、期末往来余额信息；

2、选择适当的与发行人有日常业务往来且对函证信息知情的联系人作为被询证者；

3、纸质询证函由各中介机构独立设计制作并打印，经发行人管理层授权并加盖公章；并在询证函中指明被询证对象应直接向中介机构回函；中介机构在所有询证函中均将中介机构办公地址作为函证回函收件地址进行列示。中介机构将前述纸质函证扫描件作为电子邮件附件，使用标准的邮件样式向被询证对象发送电子邮件函证。电子邮件函证由保荐机构及申报会计师分别使用中介机构业务邮箱独立发送，在邮件正文中阐明需要对方回复的具体内容，并要求直接回函至保荐机构及申报会计师的发件邮

箱；

4、若在合理的时间没有收到询证函回函，中介机构将在重新核实原被询证者邮箱地址的准确性后再次发出邮件并予以跟进。对于回函存在差异的客户，中介机构根据回函差异情况编制回函差异调节表，对回函差异原因予以核实。对于未回函客户，执行相应的替代程序。

二、发函程序控制情况

对于通过电子邮件发放的函证，保荐机构、申报会计师分别独立执行的控制措施：

1、对电子邮件函证收件人身份进行核查：将收件人邮箱的后缀与被询证者官网显示的企业邮箱后缀进行比对；获取发行人与客户在日常沟通时的往来邮件，确认该邮箱地址的使用者确实是被询证者与发行人日常业务往来的联系人且知晓双方的交易情况。

保荐机构、申报会计师对电子邮件函证收件人身份的核查情况汇总如下：

(1) 保荐机构

序号	客户名称	电子邮件函证收件人的邮箱地址后缀	电子邮件函证收件人的职务	函证期间
1	客户 3		采购人员	2020 年 1-9 月
2	客户 2		销售人员	2020 年 1-9 月
3	NTT DOCOMO, INC.	@nttdocomo.com	开发人员	2020 年 1-9 月
4	某国际消费电子产品厂商		商务人员	2020 年 1-9 月
5	某大学		某大学人类语言中心主任	2020 年 1-9 月
6	Oracle America, Inc.	@oracle.com	高级总监	2020 年 1-9 月、2019 年

经核查，上述收件人的邮箱地址后缀与函证对象的企业/机构邮箱或官网地址一致，收件人均均为函证对象业务职能部门员工，知晓双方业务往来或交易情况，不存在异常。

(2) 申报会计师

序号	客户名称	电子邮件函证收件人的邮箱地址后缀	电子邮件函证收件人的职务	函证期间
1	客户 3		采购人员	2020 年 1-9 月
2	客户 2		销售人员	2020 年 1-9 月

3	NTT DOCOMO, INC.	@nttdocomo.com	开发人员	2020年1-9月
4	某国际消费电子产品厂商		商务人员	2020年1-9月
5	某大学		某大学人类语言中心主任	2020年1-9月
6	Oracle America, Inc.	@oracle.com	高级总监	2020年1-9月、2017-2019年
7	Qualcomm Technologies, Inc.	@qualcomm.com	研发总监	2017-2019年
8	西北工业大学	@nwpu.edu.cn	计算机学院教授/博导	2017-2019年

经核查，上述收件人的邮箱地址后缀与函证对象的企业/机构邮箱或官网地址一致，收件人均均为函证对象业务职能部门员工，知晓双方业务往来或交易情况，不存在异常。

2、将核对后的盖章纸质函证扫描成PDF文件作为邮件附件。

3、由保荐机构、申报会计师分别独立直接向被函证的客户发送函证，并要求对方直接回复至发件的中介机构。

三、回函程序控制情况

对于通过电子邮件收到的回函，保荐机构、申报会计师分别独立核对：

1、发出回函的发件人是否与保荐机构、申报会计师发函时的收件人一致。

2、检查电子邮件地址后缀、邮件中所附的发件人名片或落款等，将发件人邮箱地址后缀与被询证者官网显示的企业邮箱后缀进行比对，查询邮件属性中显示的邮箱服务器IP地址，并获取日常往来文件记录进行验证，判断回函发件人是否为该公司员工。

保荐机构、申报会计师对电子邮件回函发件人身份的核查情况汇总如下：

(1) 保荐机构

序号	客户名称	回函的邮箱后缀地址	邮箱服务器IP地址	IP地址分布地区	电子邮件回函发件人的职务	函证期间
1	客户3		211.***.***.54	日本	采购人员	2020年1-9月
2	科大讯飞股份有限公司	@iflytek.com	36.*.***.15	合肥	业务人员	2020年1-9月
3	北京吉云互动科技有限公司	@bytedance.com	209.**.***.50	美国谷歌云地址	业务人员	2020年1-9月
4	某大学		128.***.**.120	美国	某大学人类语言中心主任	2020年1-9月
5	NTT DOCOMO, INC.	@nttdocomo.com	202.***.***.74	日本	开发人员	2020年1-9月

6	北京小米智能科技有限公司	@xiaomi.com	207.***.***.123	香港	人工智能部商务人员	2020年1-9月
7	北京小米移动软件有限公司	@xiaomi.com	207.***.***.123	香港	人工智能部商务人员	2020年1-9月
8	某国际消费电子产品厂商		17.***.10、 17.***.66	美国	商务人员	2020年1-9月、2019年、 2017-2018年
9	客户2		211.***.25	日本	贸易及市场部销售人员	2020年1-9月
10	Samsung R&D Institute India-Bangalore Pvt. Ltd.	@samsung.com	203.***.***.33、 203.***.***.34	韩国	IT采购高级经理	2020年1-9月、2019年
11	Oracle America, Inc.	@oracle.com	141.***.***.78、 156.***.***.85	美国	高级总监	2020年1-9月、2019年
12	北京联想软件有限公司	@lenovo.com	104.***.***.11	美国	语音测试助理	2019年
13	Advanced Media, Inc.	@advanced-media.co.jp	122.***.***.225	日本	跨境商务部-科技发展经理	2017-2018年
14	Capio, Inc.	@capio.ai	209.***.***.181	美国谷歌云地址	运营经理	2017-2018年
15	Info-communications Media Development Authority	@imda.gov.sg	160.***.***.115	新加坡	财务人员	2017-2018年
16	NAVER Corporation	@navercorp.com	125.***.***.213	韩国	财务人员	2017-2018年
17	Qualcomm Technologies, Inc.	@qualcomm.com	65.***.***.72	美国	财务人员	2017-2018年
18	Adecco Korea Co., Ltd.	@adecco.com	165.***.***.124	新加坡	服务采购部门总监	2017-2018年

经核查，上述电子邮件回函发件人的邮箱地址后缀均与被询证主体的企业邮箱或官网地址一致，相关发件人为被询证主体的业务/财务职能部门员工，知悉发行人与其业务往来、交易情况；发函、回函均为电子邮件形式的客户，其回函发件人与发函收件人一致，不存在异常。除北京吉云互动科技有限公司、Capio, Inc.邮箱服务器部署于谷歌云因而显示为谷歌云IP地址外，其余客户邮箱服务器IP地址对应地区与其公司所在地、集团公司主要经营地一致，且与发行人IP地址不存在重合，不存在异常。

(2) 申报会计师

序号	客户名称	回函的邮箱后缀地址	邮箱服务器IP地址	IP地址分布地区	电子邮件回函发件人的职务	函证年份
1	客户3		211.***.***.54	日本	采购人员	2020年1-9月
2	科大讯飞股份有限公司	@iflytek.com	36.***.15	合肥	业务人员	2020年1-9月
3	北京吉云互动科技有限公司	@bytedance.com	209.***.***.50	美国谷歌云地址	业务人员	2020年1-9月
4	某大学		128.***.***.0	美国	某大学人类语言中心主任	2020年1-9月
5	NTT DOCOMO, INC.	@nttdocomo.com	202.***.***.74	日本	开发人员	2020年1-9月

6	北京小米智能科技有限公司	@xiaomi.com	207.***.***.123	香港	人工智能部商务人员	2020年1-9月
7	北京小米移动软件有限公司	@xiaomi.com	207.***.***.123	香港	人工智能部商务人员	2020年1-9月
8	某国际消费电子产品厂商		17.***.***.10	美国	商务人员	2020年1-9月、2017-2019年
9	客户2		211.*.**.25	日本	贸易及市场部销售人员	2020年1-9月
10	Samsung R&D Institute India-Bangalore Pvt. Ltd.	@samsung.com	182.***.**.39、203.***.***.34	韩国	IT采购高级经理	2020年1-9月、2017-2019年
11	Oracle America, Inc.	@oracle.com	141.***.***.79、156.***.***.75	美国	高级总监	2020年1-9月、2017-2019年
12	北京联想软件有限公司	@lenovo.com	104.***.***.11	美国	语音测试助理	2017-2019年
13	Qualcomm Technologies, Inc.	@qualcomm.com	199.***.***.39	美国	研发总监	2017-2019年
14	西北工业大学	@nwpu.edu.cn	222.**.**.171	陕西	计算机学院教授/博导	2017-2019年

注：某大学发件人为函证对象业务部门员工，其业务部门员工将相关函证转交其财务部门员工处理，并由财务部门员工回函且抄送业务部门员工。

经核查，上述电子邮件回函发件人的邮箱地址后缀均与被询证主体的企业邮箱或官网地址一致，相关发件人为被询证主体的业务/财务职能部门员工，知悉发行人与其业务往来、交易情况；发函、回函均为电子邮件形式的客户，其回函发件人与发函收件人一致，不存在异常。除北京吉云互动科技有限公司邮箱服务器部署于谷歌云因而显示为谷歌云IP地址外，其余客户邮箱服务器IP地址对应地区与其公司所在地或集团公司主要经营地一致，且与发行人IP地址不存在重合，不存在异常。

3、对于以扫描的PDF文件作为邮件附件的回函，检查回函的签名印鉴是否与日常往来文件记录一致，客户回复的函证附件是否与保荐机构、申报会计师各自独立发送的原始发函文件一致。

中介机构对选择被询证者、设计询证函以及发出和收回询证函的全过程保持了独立、有效控制，并对函证的可靠性进行了核实。经核查，电子邮件回函的发件人均与被函证单位员工且知晓双方交易情况，发件人邮箱后缀与被询证客户的单位名称匹配一致，且与被询证单位和发行人日常业务往来邮件或文件的邮箱信息匹配一致。电子邮件回函的收件人均中介机构邮箱，即：被询证方直接向中介机构回复电子邮件函证。因此，采用邮件回函的发件人的身份可以得到验证，回函的结果真实可靠。

四、小结

综上，中介机构在函证程序设计、发函程序控制、回函程序控制方面均设置了有效的控制程序，符合《中国注册会计师审计准则第1312号——函证》及问题解答的要求。

9、保荐机构在充分核查基础上的总体意见

对本回复材料中的发行人回复（包括补充披露和说明的事项），本保荐机构均已进行核查，确认并保证其真实、完整、准确。

（本页无正文，为《关于北京海天瑞声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第二轮审核问询函的回复》之签章页）

发行人董事长（签名）：贺琳

贺琳



发行人董事长声明

本人已认真阅读北京海天瑞声科技股份有限公司本次审核问询函回复的全部内容，确认审核问询函回复中不存在虚假记载、误导性陈述或重大遗漏，并对其真实性、准确性、完整性、及时性承担相应法律责任。

发行人董事长（签名）：贺琳

贺琳



（本页无正文，为华泰联合证券有限责任公司《关于北京海天瑞声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第二轮审核问询函的回复》之签章页）

保荐代表人：

张鹏

葛青

张 鹏

葛 青

华泰联合证券有限责任公司

2020年12月18日

保荐机构总经理声明

本人已认真阅读北京海天瑞声科技股份有限公司本次审核问询函回复的全部内容，了解审核问询函回复涉及问题的核查程序、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，本回复中不存在虚假记载、误导性陈述或重大遗漏，并对其真实性、准确性、完整性、及时性承担相应法律责任。

保荐机构总经理（签名）：_____



马 骁

华泰联合证券有限责任公司

2020年12月18日

