



FY22 供应商业务近况汇报

--标贝科技 数据事业部



“用极致态度, 做极致数据”

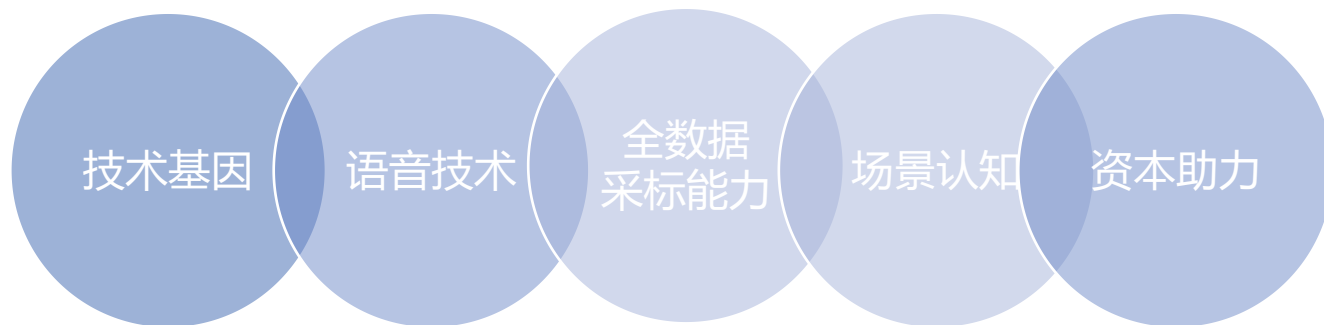


标贝（北京）科技有限公司

—国内领先的人工智能全数据服务商—



官网地址：www.data-baker.com



- 国内领先的AI全数据（音像义）服务商；
- 愿景：“语音连接场景，数据服务技术”；
- 国家高新技术企业，中关村高新技术企业；
- 致力于为客户提供专业、高效的数据服务，具备全场景、多类型采标能力；
- 标贝科技开创了TOBI国际标注体系在中国商业领域运用，并通过自主研发的TTS评测系统，为客户提供快速、有效、专业的语音数据服务。

发展轨迹

AI语音技术加持+深厚数据基础+资本助力

初创

2016-10
获得中关村高新企业认证

2016-02
成立于北京，核心团队来自腾讯、微软、百度等知名业内企业。

起步

2017-11
完成Pre-A轮千万级融资
获国家高新技术企业认定

2017-05
设立长春、天津标注基地

发展

2018-12
A轮5000万融资(深创投领投)

2018-04
亮相ICASSP2018

2018-03
加入中国语音产业联盟
上线TTS评测系统

沉淀

2019-12
标贝悦读智能配音服务平台上线

2019-10
承接HUAWEI出海项目(千万量级)

2019-09
黄金级赞助
INTERSPEECH2019, 语音论文入选

2019-08
入选腾讯5G生态计划&微软AI加速器

2019-01
深圳分公司成立, 杭州办事处成立

上升

2020-08
HUAWEI严选上架(数据服务)
3D平台建设/优化

2020-07
完成数千万A+轮融资

2020-05
多语种、多模态能力储备
深入场景的自有库建设
语音、图像、文本多能力发展

2020-05
TATH专项组模式

飞跃

2021-07
某国家级项目, 单项目金额500W+

2021-06
27701认证等
采集/标注平台能力晋级

2021-03
海外业务持续拓展, 多语种合成/识别项目承接

2021-01
数据事业部独立运营
山东、吉林标注基地规模化扩张
一总部、两分公司、三基地(山东、吉林、天津)基本成型, 具备大型项目承载能力

2016

2017

2018

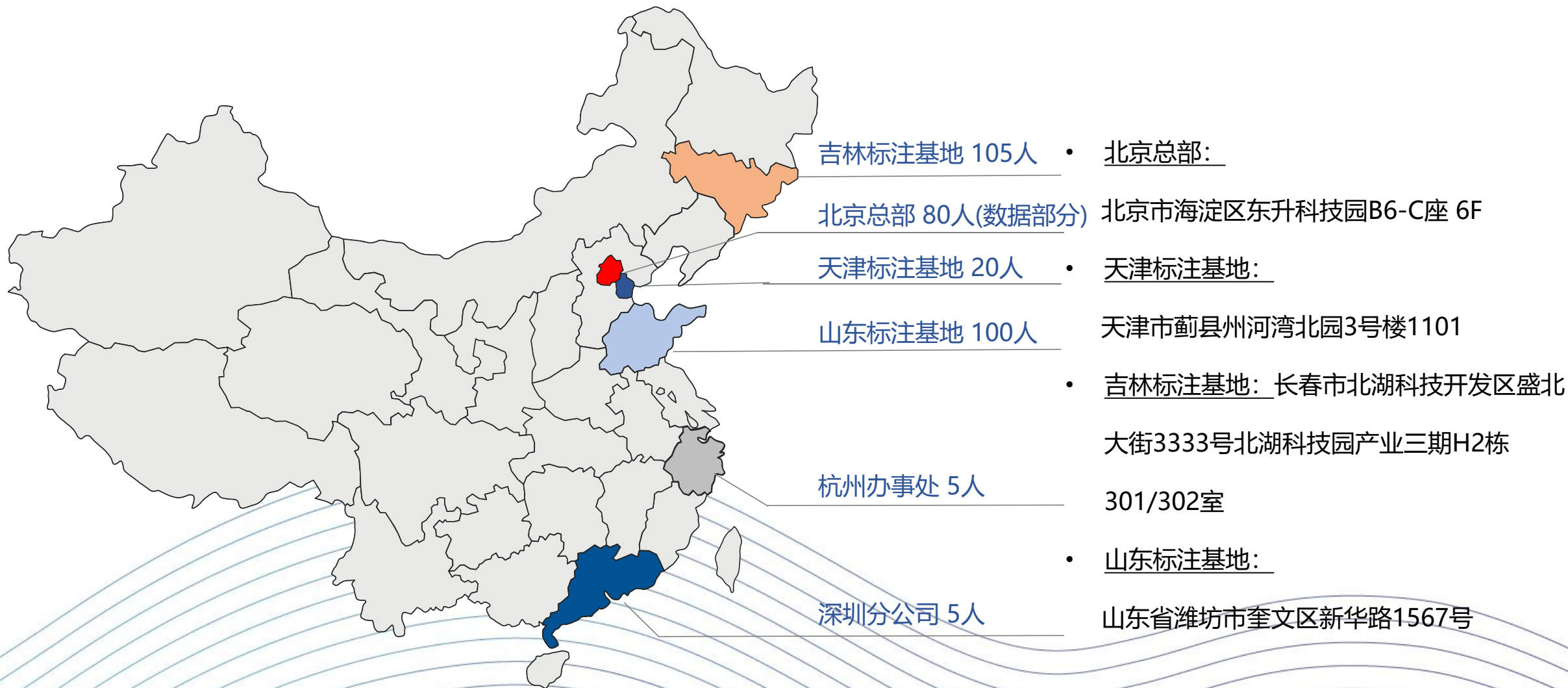
2019

2020

2021

数据事业部组织架构

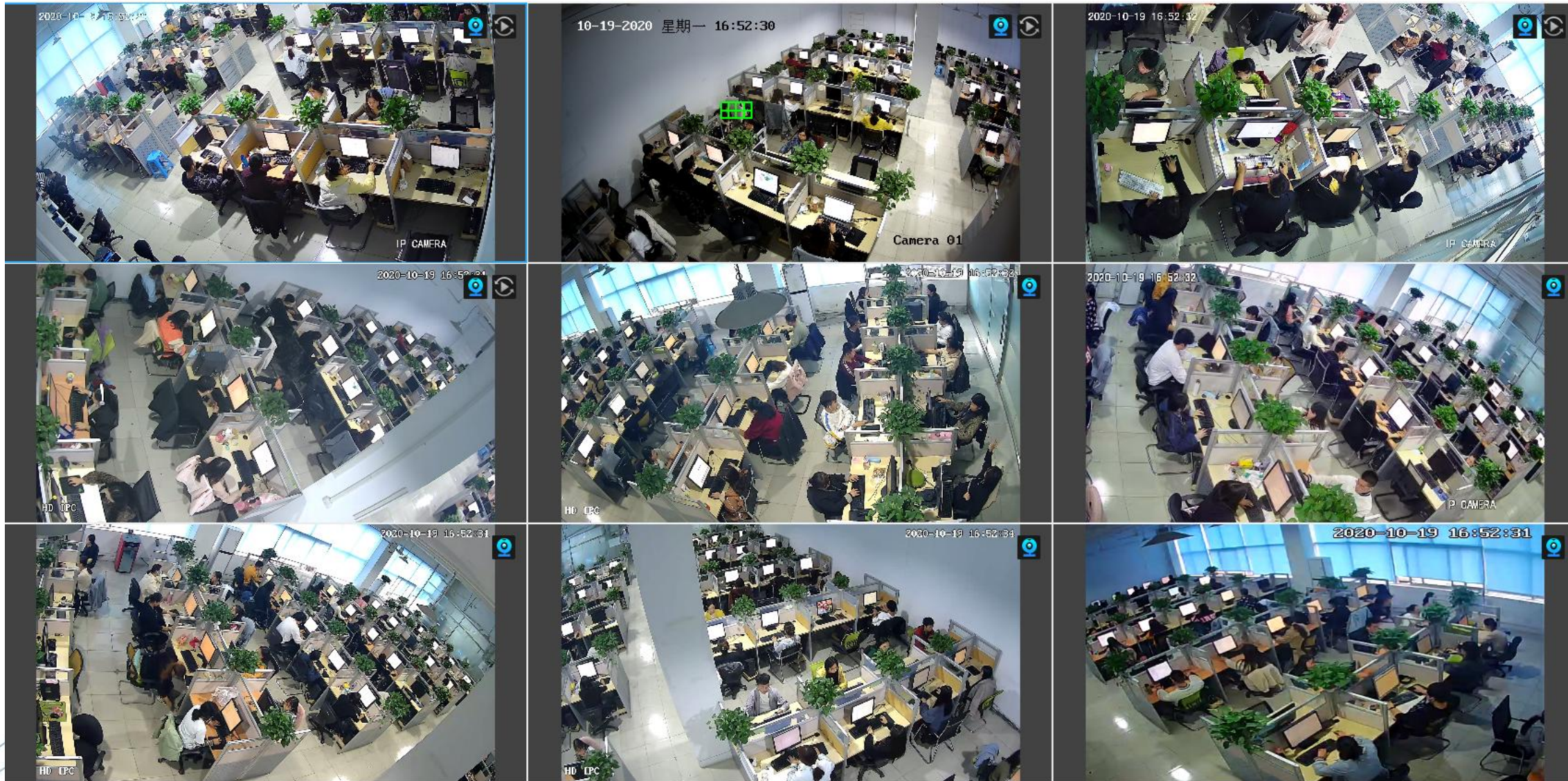
—国内领先的人工智能全数据服务商—



数据事业部组织架构



标注基地-俯视图



行业资质

质量/安全等体系认证



27001
信息安全管理体系认证



ISO20000
信息技术服务管理体系认证

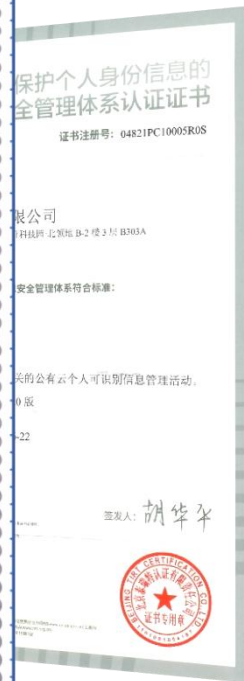
环境管理
体系认证证书



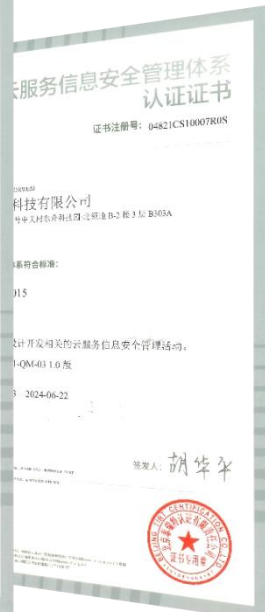
27701
隐私信息管理体系认证



ISO9001
质量管理体系认证



公有云中保护个人身份信息的
信息安全管理体系认证证书



云服务信息安全管理体系的
认证证书

行业资质

--CMMI3&高新技术认证--



CMMI 3
Maturity Level 3



国家高新技术企业
2017年获得国家高新技术企业认定



中关村高新技术企业
2016年获得中关村高新技术企业认定

私有平台部署-信息安全等级认证2.0

信息系统安全等级保护 备案证明

依据《信息安全等级保护管理办法》的有关
规定，标贝（北京）科技有限公司 单位
的：

第 2 级 标贝数据标注网站 系统

予以备案。

证书编号：11010850001-20001

中华人民共和国公安部监制



荣誉奖项

—国内领先的人工智能全数据服务商—

行业联盟

中国语音产业联盟会员
中国人工智能产业发展联
中国服务机器人产业联盟

生态合作

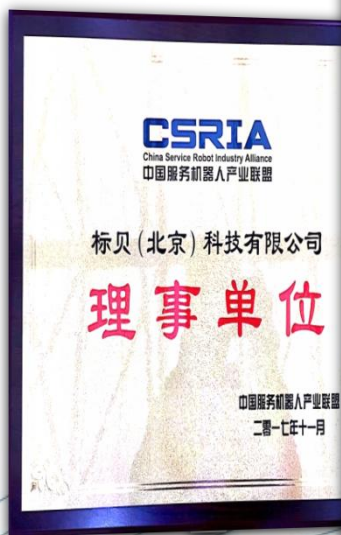
百度AI生态合作伙伴
腾讯生态战略合作伙伴

荣获奖项

2017中国信息化和
软件服务风云企业奖
2018创业黑马TOP100

软著数量

软著数量 9个
软件证书数量 3个
专利数量 30+



全方位数据采集标注能力

合作企业
100+

合作项目
1000+

服务领域
20+



语音
采集·标注



文本
采集·标注



图片
采集·标注



视频
采集·标注

标贝科技—业务能力

数据产品自有库-自有库规模

■ 语音

TTS: 80个 1500+

ASR: 29个 100000+

■ 图像

50W+

■ 语义

2000W+

自有库分类	TTS自有库List							ASR自有库List					文本List	图像List		
	中文女声音库	中文男声音库	中英童声	方言/外语	客服男女声	前段训练集&发音词典	歌曲音库	成人中文&英语	成人方言	儿童中文&英文	8K识别库	多语种识别库	垂域文本	人体&人脸	视频类	
Quantity	20	13	13	18	6	10	500	4	12	4	1	8	3	4	5	
TTL	80							500	29					3	9	

TTS音库覆盖场景: 小说、游戏、客服、夜读、广告、新闻、故事、唐诗宋词、电台解说

TTS音库覆盖风格: 亲切、甜美、妩媚、严肃、暖男、沧桑、正太、萝莉、中性

标贝科技业务能力

AI数据服务形式

1

定制化数据采标服务

每个客户在模型识别领域所进行的科研、测试和产品开发都有不一样的数据需要，可根据不同客户的不同需求为客户量身定制，提供多种类型的采集、标注等服务。

2

语音合成数据

包括音库制作和纯文本标注。音库制作包含多语种的合成音库录制，基于语音文件的音字标注、韵律标注、音节音素标注等内容；纯文本标注包含基于文本进行分词、词性、多音字、韵律、数字符号标注等。

3

语音识别数据

可满足各语种、各方言的数据采集及标注；指定设备（如手机/高保真麦克/录音板）及指定环境下（如商场、餐厅、医院、超市、车站等）的采集及标注。

4

图像识别数据

文字图片、人脸图片、车辆图片、人体行为轨迹视频、车辆视频等类型的采集和标注；以及基于客户指定的其他采标要求。

5

文本数据

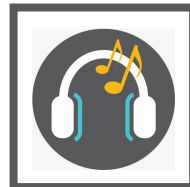
表单采集，实体标注，标签分类，词性标注，情感分析，意图判断等。

标贝科技业务能力



语音数据采集

语音数据采集，可支持多语言语种，多场景，多方式的定制化数据采集（中文男声，中文女生，中文方言，童声等等）



歌曲数据采集

歌曲数据采集，我司有专业的录音棚以及专业的歌曲数据采集团队，可提供歌曲数据采集



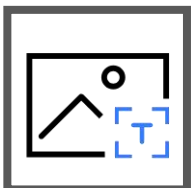
动物物体图像采集

动物以及物体的数据采集



智能驾驶数据采集

智能驾驶数据采集，通常用于ADAS的训练模型使用，包含路况采集，DMS采集等。



OCR数据采集

OCR数据采集，可支持发票，海报，水电费凭证等数据采集



街景图像数据采集

街景图像数据采集，可支持国内外指定城市的街景采集



人脸数据采集

人脸数据采集，通常用于人脸识别模型训练，可支持多人种、各年龄段的人脸数据采集



安防监控数据采集

安防监控数据采集，通常用于公共安全防范类型的识别模型训练，可支持火灾数据，危险行为数据，交通监控类数据的采集。

标贝科技业务能力

数据标注服务（语音） 1/3



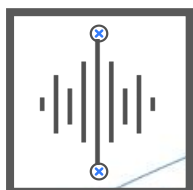
语音转写

语音转写是根据音频播放的内容转写为对应的文本，通常用于语音识别模型训练，可支持普通话、方言、英文和小语种的语音转写



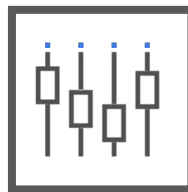
语音清洗

语音清洗通过技术清洗空音频，通过人工监听，筛选出符合要求的音频，用于语音识别模型训练



语音切分

语音切分是对长音频进行监听，标注音频中说话人的起始点，用于语音识别模型训练



音素标注

音素标注是对音频进行监听，转写文本同时对文字的音标进行标注，常用于语音合成技术



情绪判定

语音情绪是对音频进行监听，根据说话人的情绪状态进行标注，常用于语音识别、合成技术



声纹标注

声纹标注是对音频数据中说话人进行辨别或确认，用于加密解锁传输数据等技术

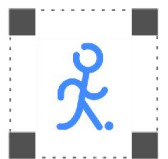
标贝科技业务能力

数据标注服务（图像） 2/3



图片分类

图片分类，根据目标属性，可实现千万量级的图片清洗分类，对图像划分不同分类，依据您的需求可对您提供的图片集做清洗和属性归类



图片框选

图片框选可助力图像识别模型训练，用2D、3D、多边形框等标注出图像或视频数据中的指定目标，常见于对人脸、人体、障碍物、红绿灯的框选

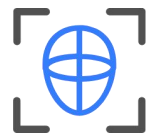


图像语义分割是标记图像中存在的内容及位置，对复杂不规则的图片进行区域划分并标注对应属性，根据属性进行像素级分割，支持单属性，多级属性。



线段标注

线段标注支持贝赛尔曲线和普通线段标注，扶持对车道进行贝赛尔曲线标注，使用线段将图像目标的边缘、轮廓用线段标注



关键点标注

关键点标注，对人脸五官、人体骨骼关键点等其它物品关键点进行标注，用于图像识别模型训练



3D点云标注

3D点云标注是将点云数据图像中的多类指定对象使用3D框进行标注



2D3D融合

2D3D融合标注是对2D平面图与3D点云图映射的数据组进行标注，支持自动贴边、跨帧复制、测距、2D3D图片标注分离等多种功能



目标跟踪

目标跟踪，连续帧标注常用于自动驾驶及视频图像识别模型的训练，通过对视频进行抽帧，并对每一帧图片中的目标物体进行连续标注

标贝科技业务能力

数据标注服务（文本） 3/3



OCR转写

OCR转写是对图片中的文字内容进行标框和转写，支持中文、英文及小语种的图片转写，助力图片和文本识别模型



情感判断

情感判断是对文本表达情绪倾向进行判断，分类积极和消极的文本



文本清洗

文本清洗是按照客户的规则对文本进行筛选，挑出符合要求的数据，可实现千万级文本的清洗



词性标注

词性标注，基于特定文本数据进行词性理解标注，如动词、名词、形容词。



文本富集

文本富集是围绕主题进行文本撰写，使得对于同一主题，文本表达方式多样且贴合实际



意图判断

意图判断，通常全用分类的方法将句子分为相应的意图种类，单意图或多意图等，进而判断语义意图



信息抽取

信息抽取是把文本里包含的信息进行分类和打标签等结构化处理。



NLP标注

NLP标注，NLP标注是对文本语法的标注，包括槽位提取、文本关系等

标贝科技业务能力

图像业务范围-可承接

生物识别：
包括人脸采集、指纹采集、虹膜采集
等及对应标注服务



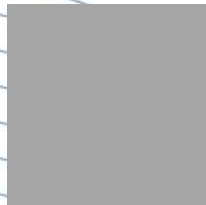
OCR文档/场景数据采集：
包括多场景图片轮廓标注、多语言文
档图片标注、类别标注、属性标注等

智能交通&自动驾驶类：
包括车载视频采集、卡口图片收集、街
口视频收集、自动驾驶相关标注、轨迹
跟踪标注、街景像素级标注



智慧医疗：
包括医疗影像病灶区域、人体骨骼关键
点、人体区域等标注，并根据医疗影像
进行医学参考分类

人机交互类采集：
包括离散情感图像视频的采集、动作识
别数据的采集标注



安防监控：
红外相机、摄像头、深度相机等。
安防场景中，监控视频的采集及对视
频中特定的物体和事件的标注

标贝科技业务能力

行业解决方案-自动驾驶



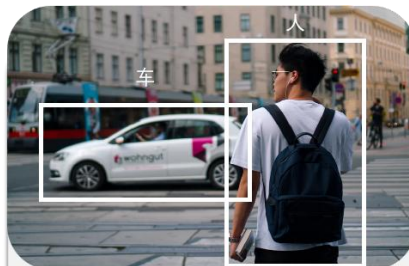
车道线标注

对道路地面标线进行的综合标注，标注包括区域标注、分类标注及语义标注，应用于训练自动驾驶根据车道规则进行行驶。



车辆多边形标注

对车辆进行区域标注以及分类标注，主要应用于对车辆的识别，例如面包车、卡车、大客车、小轿车等，训练自动驾驶在道路行驶时选择性跟车或者变道操作。



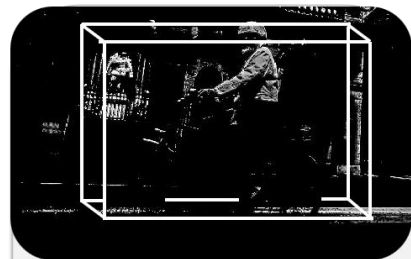
2D车辆/行人标框标注

这种标注在自动驾驶中是最基础也是应用最广的一种标注方式，主要应用于对车辆与行人的基础识别。



指示牌/信号灯标注

这种标注是一种对道路悬挂指示牌/信号灯进行的综合标注，标注包括区域标注、粉类标注及语义标注，应用于训练自动驾驶根据交通规则进行行驶。



3D雷达标注

根据镜头反求原理，将视频场景模拟成3D图像，通过3D图像标注出标注物的位置及大小。主要应用在自动驾驶虚拟现实（VR）训练场景的搭建。



3D车辆标注

将2D图片中的车辆进行3D标注，主要应用于训练自动驾驶对会车或超车的体积判断。



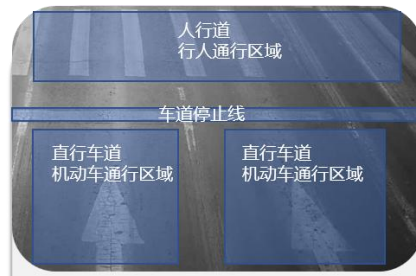
行进方向标注

这是一种对标注物前进方向的预判性标注，需进行标框标注及方向预判标注，应用于训练自动驾驶判断行人或车辆前进方向，规避行人或车辆。



3D车辆标注

将2D图片中的车辆进行3D标注，主要应用于训练自动驾驶对会车或超车的体积判断。

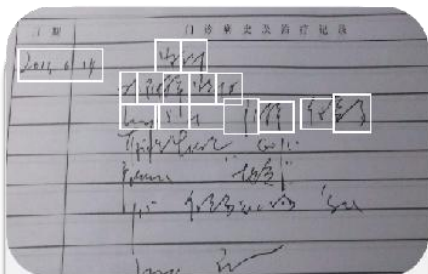


区域分割标注

这是一种对道路区域进行的综合标注，标注包括了区域标注、分类标注以语义标注，应用于训练自动驾驶根据交通规则进行行驶。

标贝科技业务能力

行业解决方案-智慧医疗&智能安防



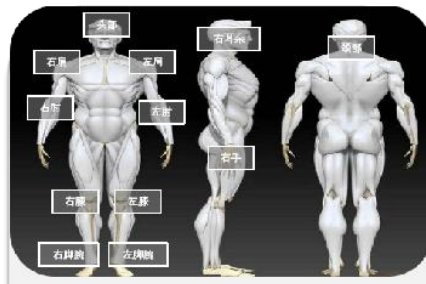
病例文本标注

病历文本标注是对病历信息进行文本标框标注，通过对病历内容的文本转录实现电子病历系统建立。



骨骼点标注

将人体运动的关节点进行描点标注，多应用于建立健康档案，人工智能通过对骨骼点标注的学习，可以快速锁定病灶关节。



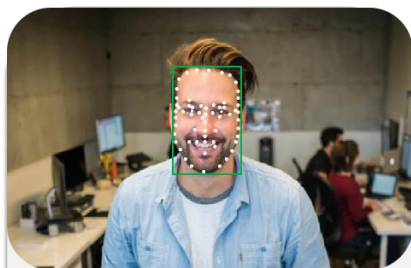
人体标框标注

根据人体不同部位进行标框标注，多应用于远程医疗外伤诊断。



医疗影像标注

对医疗影像进行区域标注及分类标注，多应用于辅助临床诊断。人工智能通过学习大量医疗影像标注数据，可以辅助医生进行临床诊断以及提出治疗方案。



人脸标注

广泛应用并且在不断发展中的数据标注，在智能安防中，如今描点标注已经从简单的29点发展到超过108点，主要应用于人脸识别与身份识别。



表情分析

表情分类是一种分类标注，在机器学习时，需要配合人脸标注进行。在智能安防中，表情分析是智能安防系统从被动防御向主动预警发展的关键技术。



行人标注

对行人进行标注，主要应用于进出人数统计，一般在商场、超市、学校等人员容易密集的场所需要通过进出统计判断容纳人员是否已经饱和，也可以有效的防范因为人员密集而造成危险。



行为标注

对特定行为进行区域标注和分类标注，主要应用于对危险行为的监控，如打架、晕倒、车祸等，视频监控系统识别出危险行为后，可以及时报警。



物品标注

将物品进行标框标注及分类标注，在智能安防中，物品标注需要和行为标注结合。

项目经验

语音合成-过往项目经验

所属行业	项目规格	项目量级	项目难点&亮点	经验总结
某平台类KA	垂类场景音色录制	6W句	多场景海量声优筛选, 目标音色获取 大数据量合成音库文本设计, 情感评分 数据质量保证, 交付进度、工期限制	海量垂类声优资源储备 发音人大语料量录制经验 项目进度控制, 交付进度
某国家级合成项目	小语种合成音库	200H	多语种声优严格筛选, 非大语种声优获取 海外声优沟通, 标准传递 海外语种熟悉程度, 质量控制 (声优质量) 合成项目并行管理能力, 项目进度保证	小语种资源丰富度 多合成并行项目经验 海外声优项目控制
某消费电子厂商	多情感音色录制	2W句	符合特定场景的声优获取 多风格、多情感语料的录制 半自由、全自由式录制规格	半自由、全自由合成项目录制经验
某BAT客户	歌曲类采集标注	3500首	近45H歌曲量采标项目, 对并行人力要求较高 客户指定标注方式 交付进度, 项目工期保证	大量歌曲自有库 (带标注) 储备 歌曲大标注量并行, 团队培养/储备

项目经验

语音识别-过往项目经验

所属行业	项目规格	项目量级	项目难点&亮点	经验总结
某消费电子	英国英语采标	5000+人 3000h+	海外执行（英国全境），多场地（各地域口音分布）采标并行，跨域管理难度大； 多场景（安静/车载）、多设备（耳机、标麦、手机等）录音采集； 项目专属场地（数据安全&管理） 数据质量&多并行团队管理（培训机制） 交付进度（交付进度、需求变更管理）保证， 数据安全（GDPR）	大型海外采标经验 数据安全/合规的经验 海外大型项目执行经验，应急预案 完善的项目管理流程&质量保证 项目整体交付能力，交付控制能力 海外当地文化融入&本地资源
某品牌车厂	欧洲6国语种采标	1000+	多语种车载采标并行，对资源丰富度、项目并行管理能力提出较大挑战； 综合海外项目管理能力； 海外语种熟悉程度，质量控制； 项目进度（交付进度、交付节点）保证； 大规模采集人员组织能力，短时间内人员管理能力； 法律风险规避经验	欧洲语种成熟海外采标经验 海外丰富资源&稳定海外采标团队 海外项目整体难度的认知&把握 成本&效率的控制与平衡 项目执行中影响质量&交付的关键节点把握 采集人完善招募渠道，完善标注培训体系，多国家法律法规遵守、合规经验， 数据安全

项目经验

NLP类-过往项目经验

所属行业	所属类型	项目规格	项目量级	需求解读	项目表现
消费电子	某消费电子厂商	NLG文本撰写&标注	百万	多轮文本撰写&槽位标注&意图标注	语音助手有效优化
银行/金融	某互联网金融科技公司	NLP标注	百万	智能设备控制领域的NLU标注&多垂类意图 (intent) 和槽位 (slot) 标注	助力技术模型训练&优化
消费电子	某消费电子厂商	NLP话术撰写	百万	根据场景、人设、上下文、query定义等做多轮对话撰写	语音助手有效优化
消费电子	某科技公司	文本审核	近百万	文本内容鉴黄、鉴反等并打对应标签	助力企业规避经营风险

项目经验

运动健康类-过往项目经验

所属行业	所属类型	项目规格	需求解读	项目表现
消费电子	某消费电子厂商	心率跳绳采集	跳绳过程中心率的采集	智能穿戴设备优化
消费电子	某消费电子厂商	耳温、动态心率采集	手表、心率带等设备的采集	智能穿戴设备优化
消费电子	某消费电子厂商	血氧采集	手表、血氧仪、指夹血氧信号采集器等	智能穿戴设备优化
消费电子	某科技公司	游泳，滑雪等心率采集	各种运动类型中的心率采集	各种运动智能设备优化

NLP类-过往项目案例

项目概述:

知识库设计, 根据角色要求用指定工具完成文本采集, 按agent与customer

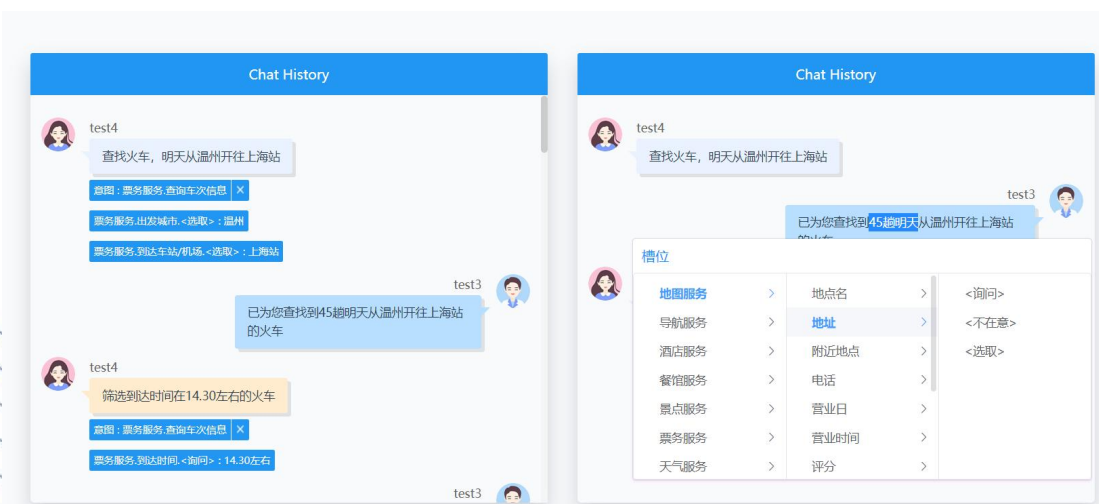
角色定义完成对话。Agent需要模拟语音助手的身份, 回复用户的请求/指令

对话原则: 每轮对话表达一个意图, 情景项目完整, 符合日常习惯

标注规则: 对每轮对话中的customer的文本内容进行意图、槽位、情感的

标签属性标注

阶段	评价项目		评价标准
对话采集	任务完成度		100%
	错别字比例		<1%
	日常语用习惯	指代使用频率	待定
		省略使用频率	待定
	平均对话轮数		4.5-5.5轮 (待定) / 任务
对话标注	标注完成度	意图标注完成度	100%
		槽位标注完成度	>99%
	标注正确率	意图标注正确率	100%
		槽位标注正确率	>99%



过往承接项目

图像标注&采集-过往项目经验

所属行业	所属类型	项目规格	项目量级	需求解读	项目表现
智能驾驶	AI产品厂商	全景语义分割	百万张	智能驾驶场景下的全景语义分割	助力街景识别表现
新零售	某互联网科技公司	人脸拉框聚类	百万框	视频抽帧后，标注人脸框，同时将同一个人的人脸框进行聚类标注	助力技术模型训练&优化
消费电子	某消费电子厂商	夸张表情采集	千人	17个角度联动采集，按照要求做出夸张表情，带动特定肌肉群拉伸变形	特殊表情识别效果有效优化
安防厂商	某AI巨头公司	安防场景采集	万人	针对算法薄弱人种，采集印度人安防场景数据	海外落地执行，助力算法填补空缺

项目能力

小语种项目能力

多语种采标能力

- 支持全球范围70多语种采集、标注
- 全球多国家海外执行能力
- 多国家、多语种、多场地并行&交付能力

过往项目经验



- 过往海外小语种大中型项目执行经验
- 过往丰富海外采集、标注&并行经验&大业务量发音人组织经验
- 丰富海外实操经验，数据采标过程风险规避

团队&项目能力

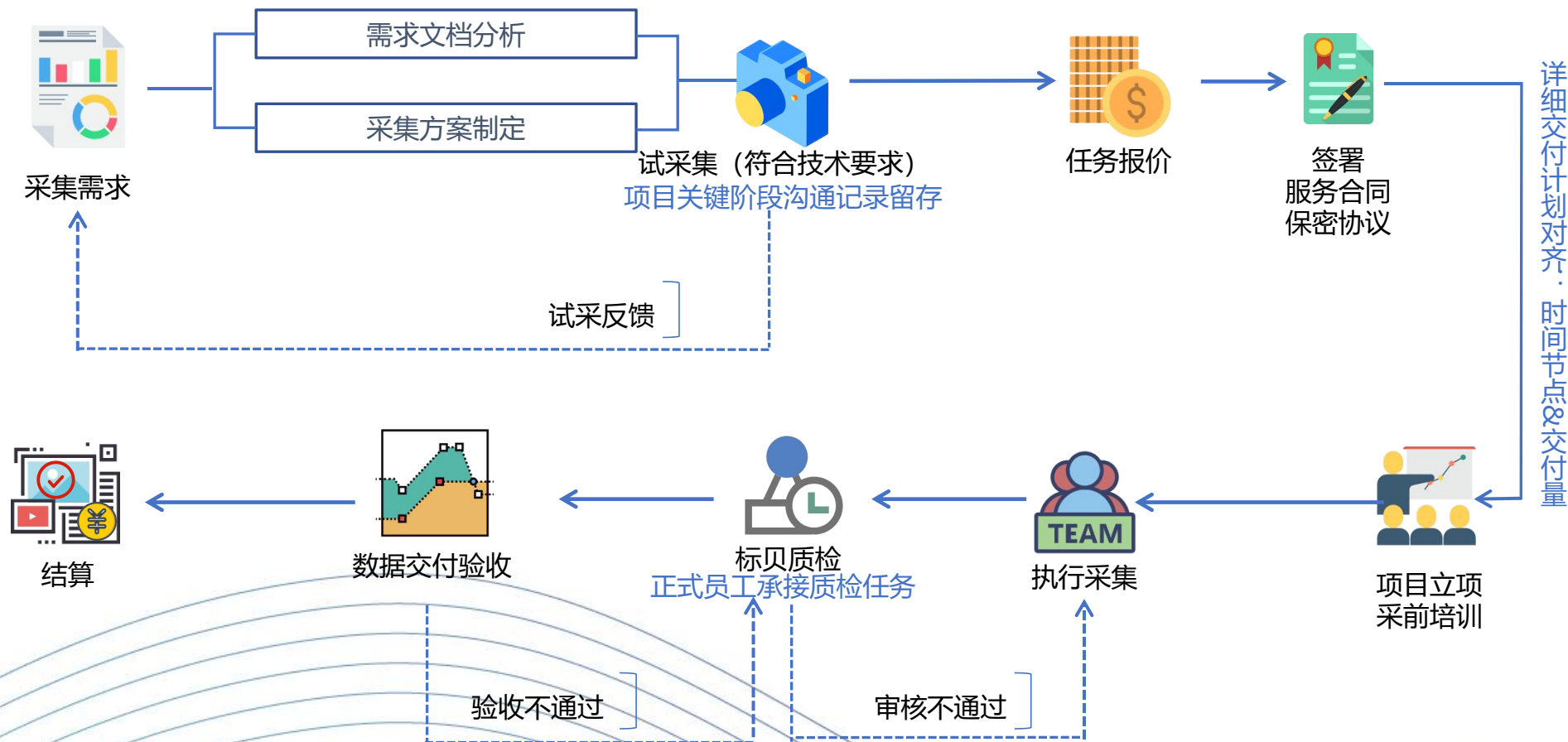


- 团队丰富海外项目执行能力，项目经理丰富小语种采标经验
- 团队可承接较大规模小语种采标项目，支持海外多场地并行
- 各采集国家发音人组织能力，丰富当地国采标资源
- PM+PA (native) 业务模式，保证业务解读与沟通，进而保证质量与交付进度

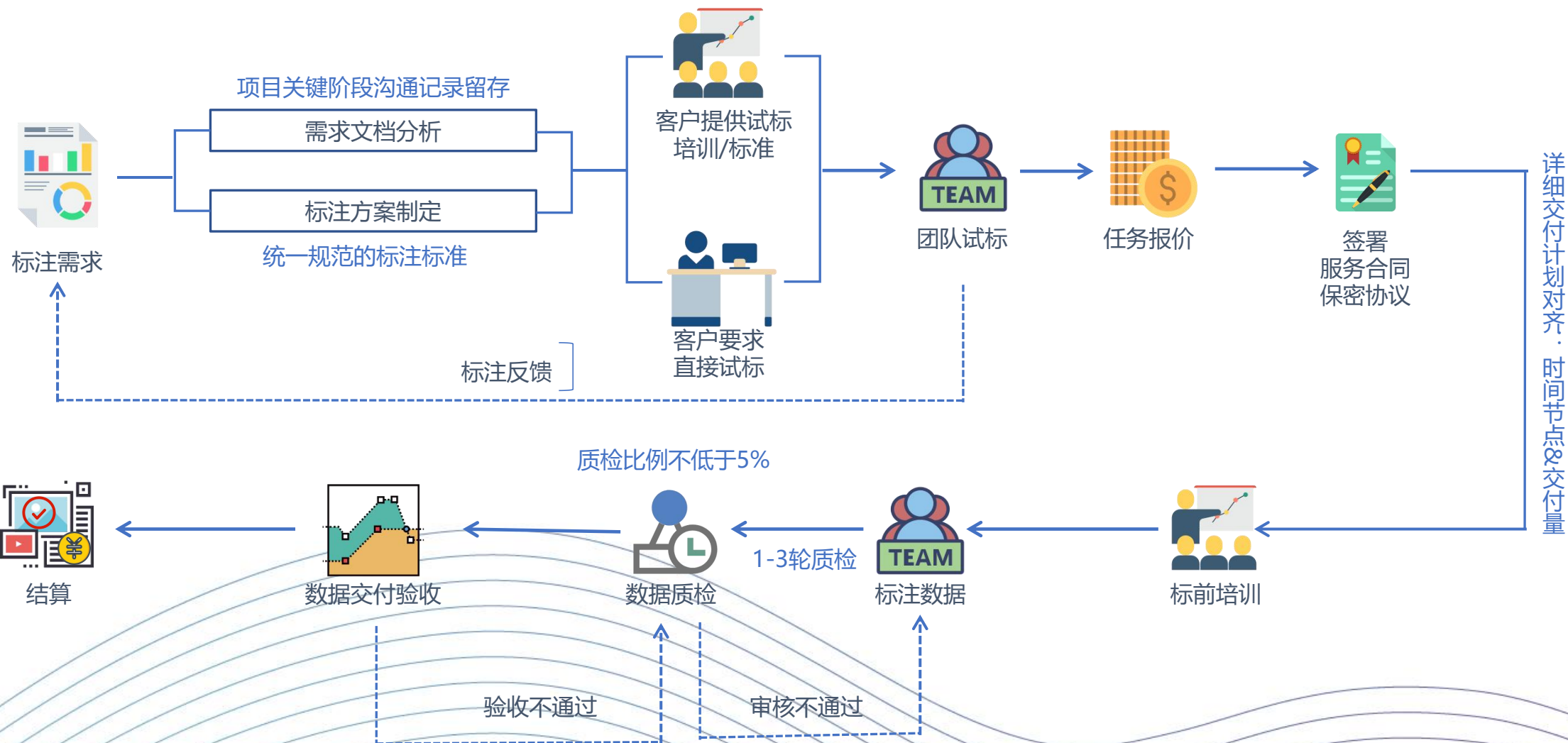
数据安全

- GDPR理解，数据安全海外多国家实操
- 海外执行场地风险规避，各国家风俗习惯融入、适配
- 数据合规、流程可溯

采集服务合作流程

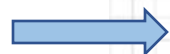


标注服务合作流程



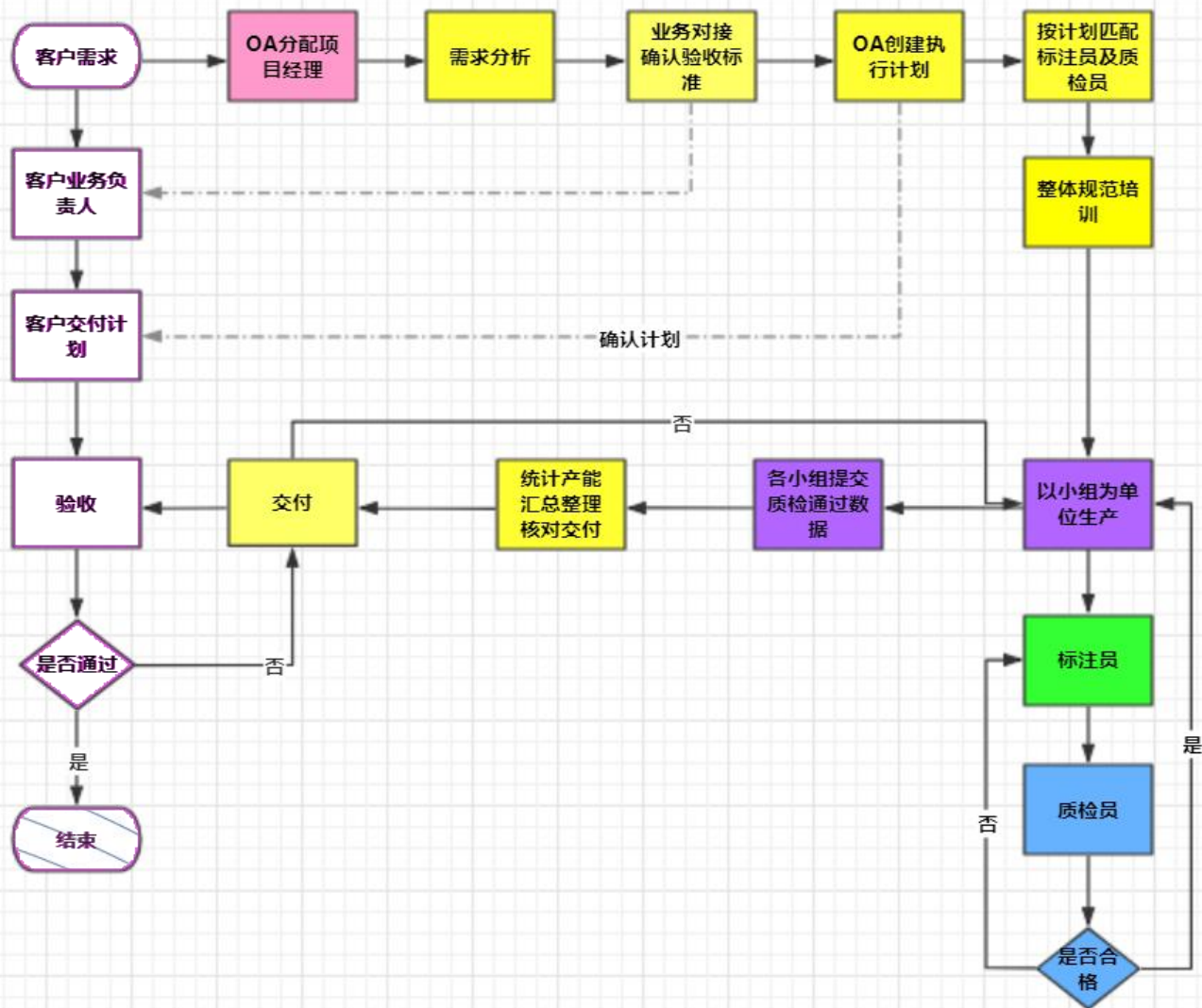
交付管理-识别类

项目执行全过程生命周期示意图



- ✓ 项目项目经理专项负责
- ✓ 实际执行按标准业务流进行
- ✓ 交付流程:

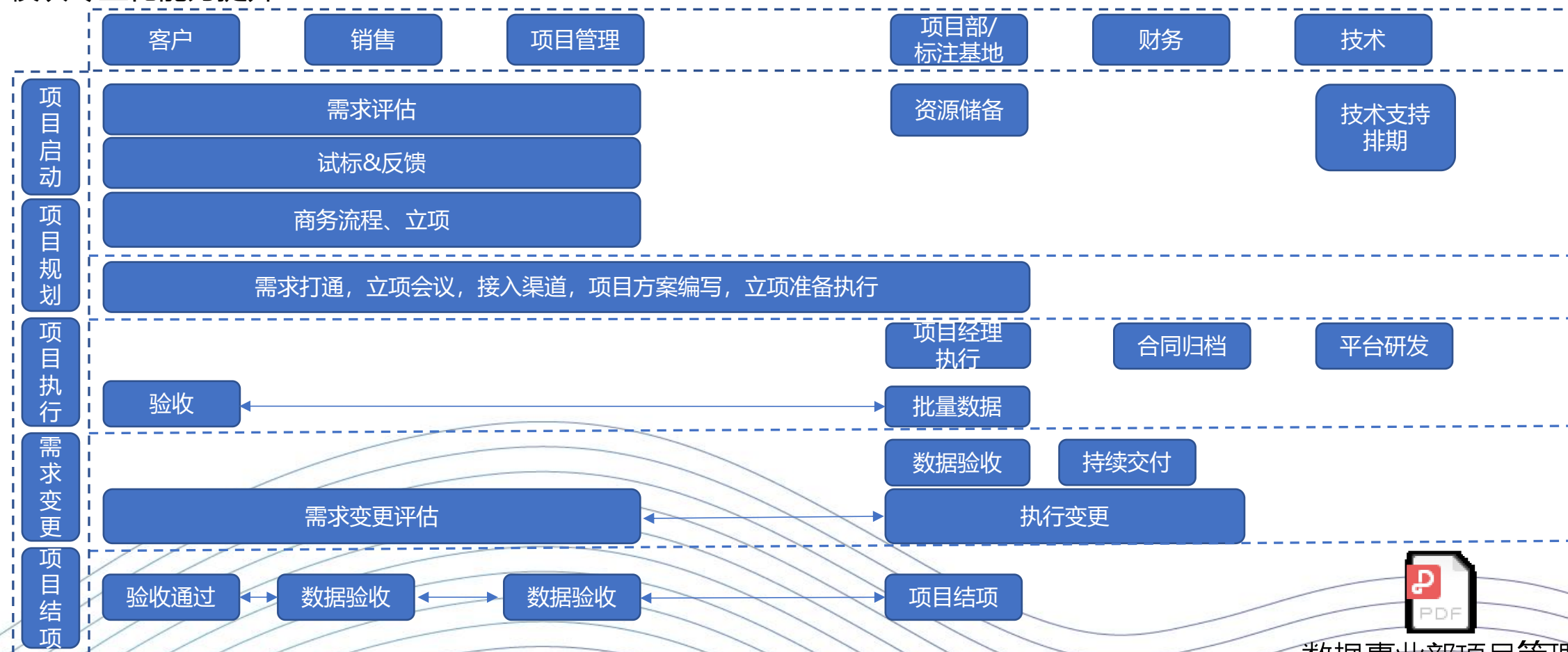
需求分析-确认交付计划-落地生产-结果交付
-售后数据维护



项目管理全流程

数据项目标准化流程-一个标准

流程清晰、职责清晰的模块化项目流程（各司其职）
模块专业化能力提升

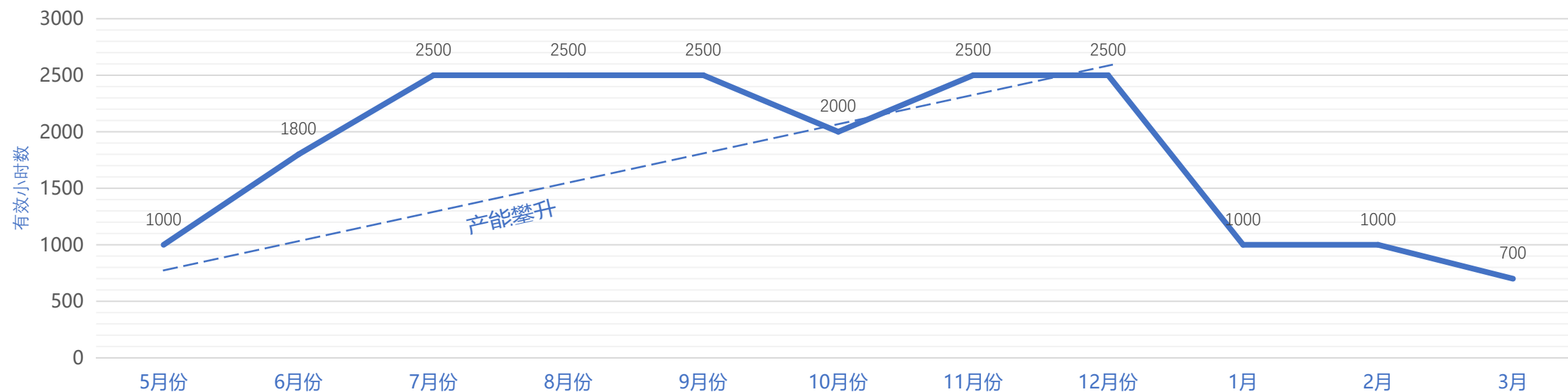


项目管理全流程

项目管理流程-交付计划(举例)

四阶段：启动（项目整体流程设计）--试标/磨合阶段—正式标注阶段—收尾阶段

ASR语音标注项目计划产量（2W小时）



试标&标准明晰&对齐 培训体系完善 数据持续交付和进度监控

团队组建

数据安全保证

标注&效率持续提升

中期复盘、计划调整总结

按质、按计划

提交数据

数据整理

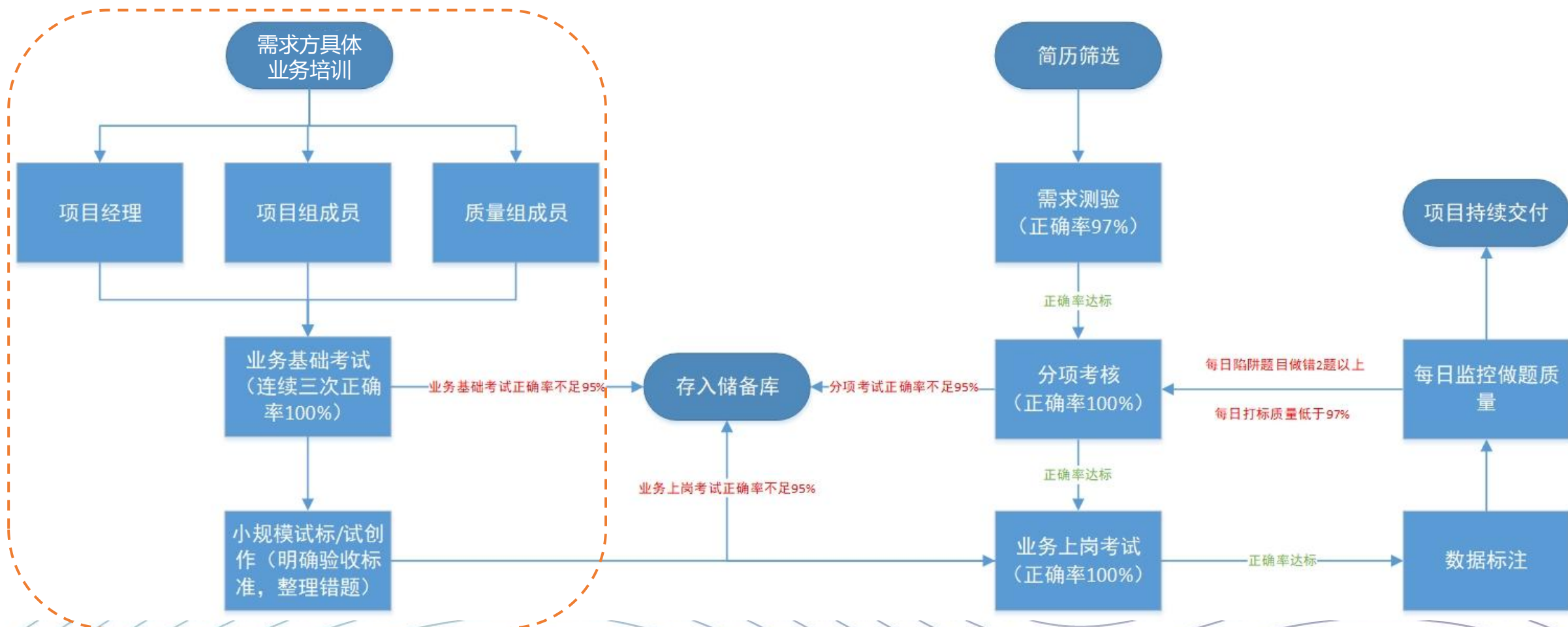
和可交付物的移交

项目结项

项目管理全流程

项目管理流程-培训细节 (驻场)

完善的培训、管理体系助力项目如期、按质交付



标贝科技业务能力

标贝数据工场
创新打造数据可视化众包平台

首次将数据格式算法融入其中，通过操作平台严格控制前端采集格式，避免数据在采集和上传的过程中因压缩出现质量损失问题，大幅度提升了数据的高质量和准确性。

样本收集承载量

1000万+

单日样本采集

10万+

采标效率提升

50%+

准确率

95%+



数据采集

可根据定制化需求，对各类规定文本、指定图片，各种环境下的语音、视频进行采集；采集过程中可实现对其内容的筛选、文本化等相关任务。可承载千万级别以上样本的收集，单日完成10万+样本采集。

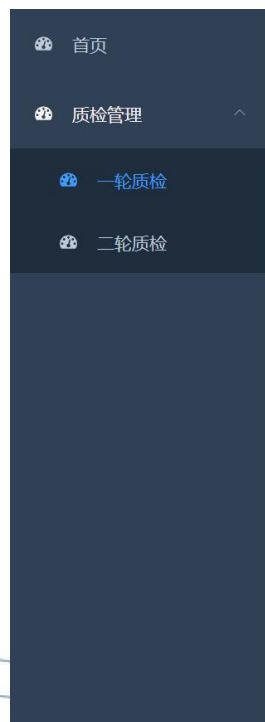
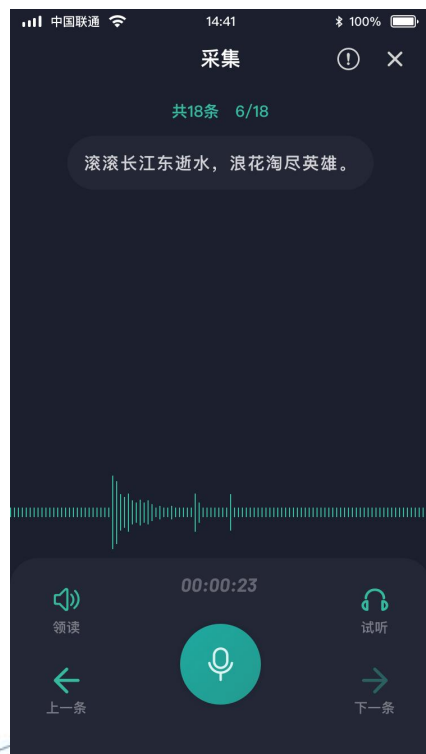


数据标注

通过对图像、文本、音频、视频等信息进行搜集、评估、归类，最终完成标注；标注过程中可实现对内容进行关键词等内容的提取、清洗、脱敏、校验等相关任务。企业客户通过使用标贝数据工场采标效率可提升50%。

标贝科技业务能力

标贝易采
创新打造数据可视化众包平台



文本 → 新建项目 → 录制 → 质检 → 验收

XML文本、采集模板、任务分配

登录APP、加入团队、录制任务、提交

一检、二检

标贝科技业务能力

语音标注平台

[音频列表](#) / [音频详情](#)
标注员: CAPTAIN 验收员: A0001 数据编号: 912378 音频名称: 音频1 状态: 质检驳回

[A](#) [保存](#) [通过](#) [?](#)



总时长: 0秒 开始: 0秒 总截取: 0秒
结束: 0秒 未截取: 0秒 截取: 0秒

自动进入下一条

0.8X 1X 1.2X 1.5X 2X

自动吸附 1.5X

文本信息

如果此段状态为“已完成、未完成、质检合格、质检不合格、质检不合格-已修改”则按钮为“合格、不合格”。如果此段状态为“验收不合格、验收不合格-已修改、验收不合格-质检不通过、验收不通过-质检不通过”则按钮为“复查通过、复查不通过”点击后改变语音段序号颜色。

性别: 男 地区:

画段标注

音频是否有效* 是 否

如果此段状态为“已完成、未完成、质检合格、质检不合格、质检不合格-已修改”则按钮为“合格、不合格”。如果此段状态为“验收不合格、验收不合格-已修改、验收不合格-质检不通过、验收不通过-质检不通过”

音频口音* 官话 四川话 粤语 闽南话

性别* 男 女

儿童* 是 否

订单编号*

音频是否有效* 这里是文案 这里是文案

全局标注

音频是否有效* 是 否

段落编号 全部 不合格 合格 已修改 未完成

1	2	3 <small>检</small>	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48												

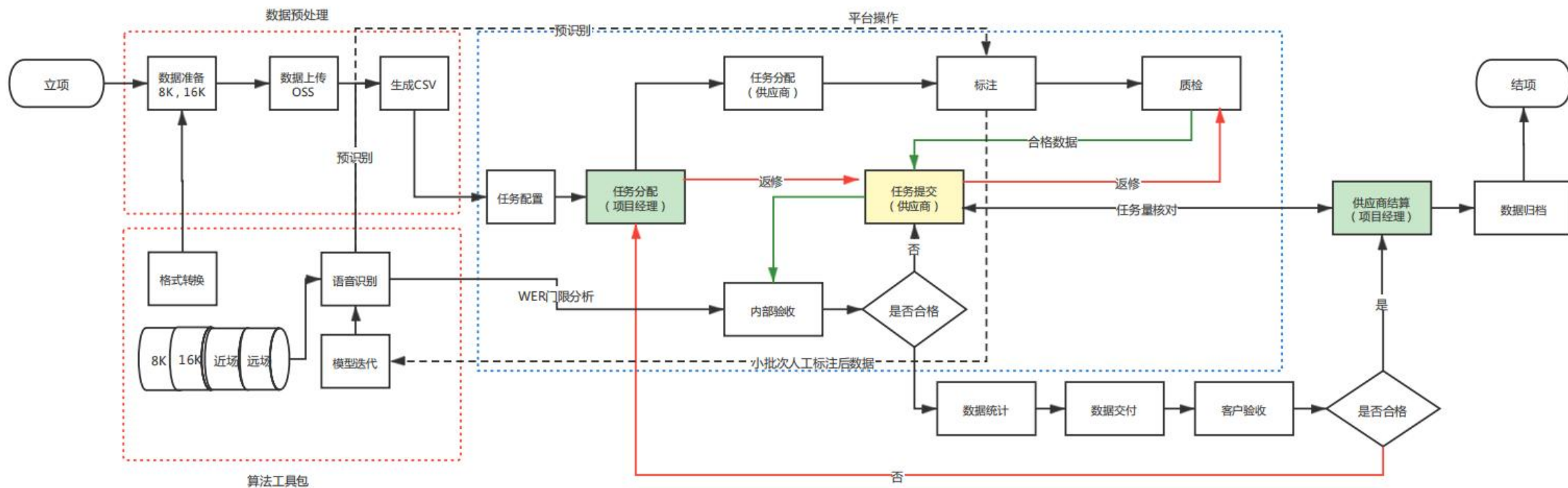
质检意见

验收意见

[不合格](#) [合格](#)

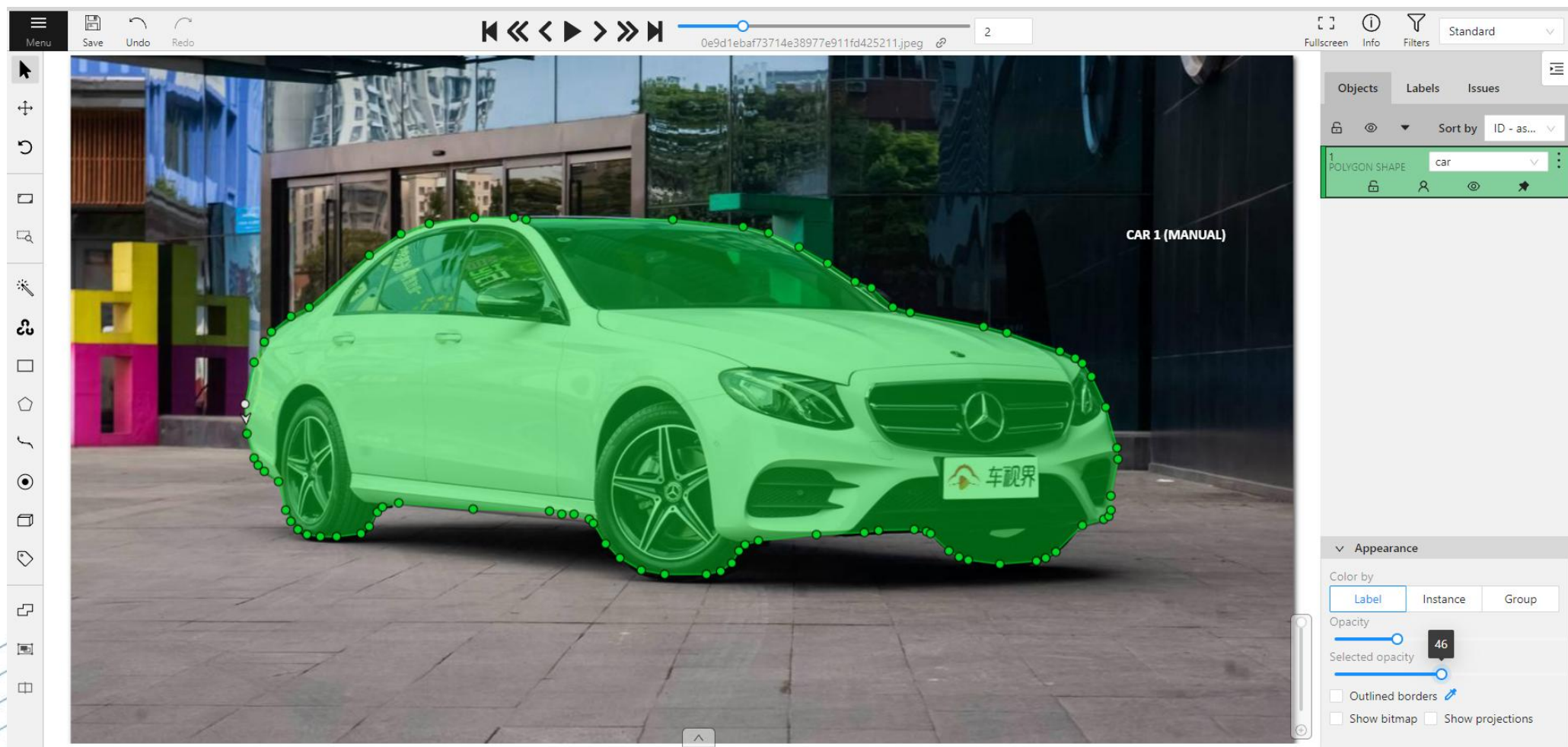
标贝科技业务能力

语音标注业务逻辑



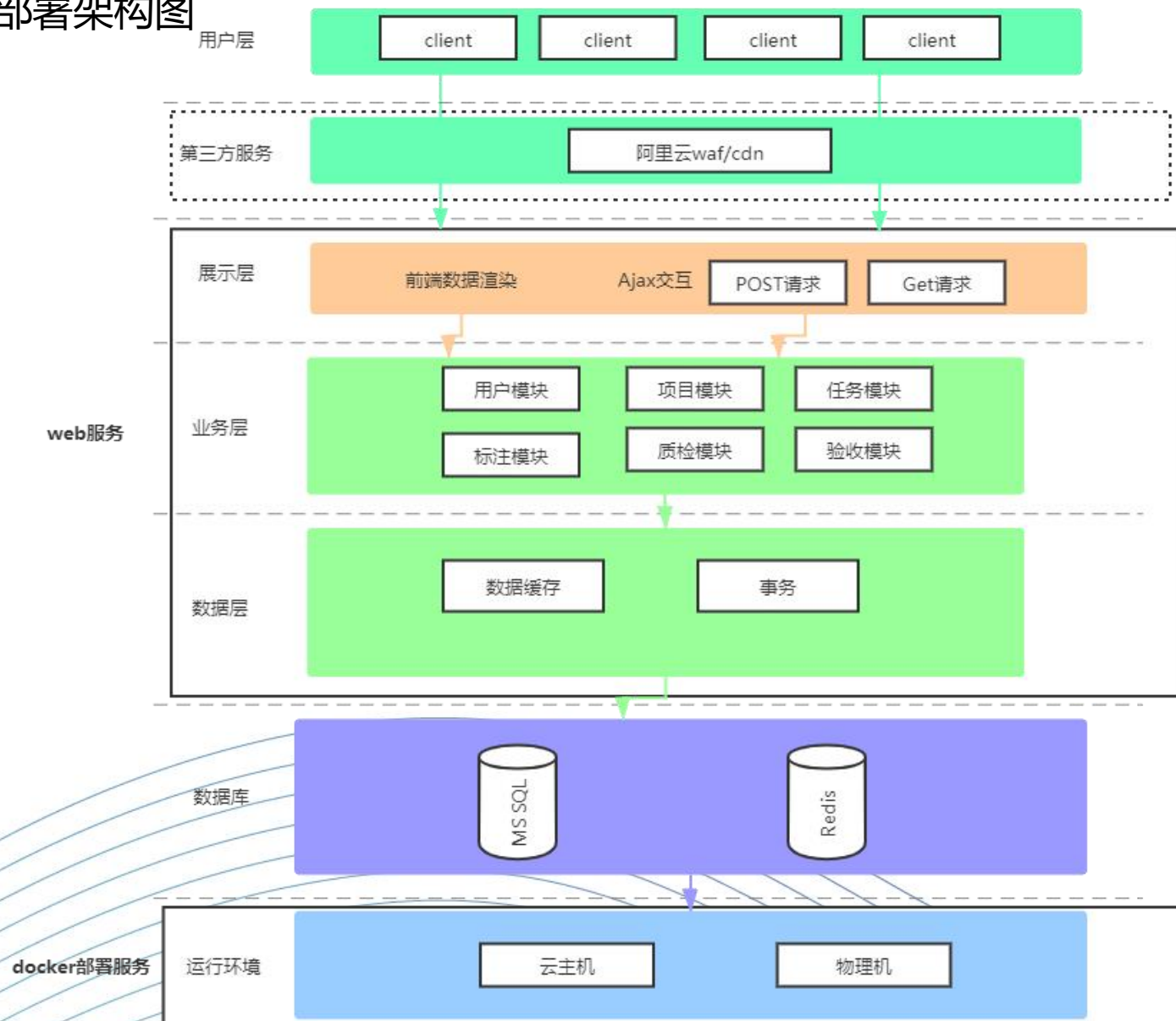
标贝科技业务能力

图像标注平台



标贝科技业务能力

私有平台部署架构图



- 数据传输采用HTTPS 协议，支持对网络报文进行加密（RSA加密进行报文加密 公钥加密 私钥解密 确保数据在网络传输的绝对安全）
- 实现时间：单周级别
- 数据格式：数据到达客户端后导出后仍然是明文，直接可以播放使用。

标贝科技—业务能力

标贝录音环境-合成业务



合作录音明星

- 柳岩
- 桃宝
- 包贝尔
- 韩乔生
- 金沙
- 王璐丹
- 乔杉
- 黄轩
- 杨洋
- 魏瑾
- 罗云熙
- 冯提莫



应用场景



消费电子

唤醒/命令词、多语种、多场景识别/合成、语义分析等



泛娱乐/电子商务

多语种、多方言语音识别、黄反监测，声纹识别



银行/金融

文字翻译、语义分析、图像标注等



医疗行业

人体标框、3D画框、骨骼点标记、病历转写



智能客服

语音识别/语义检测



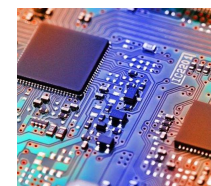
智能安防

人脸标注、视频分割、语音采集、行人标注等



车载/自动驾驶

自动驾驶，矩形框、语义分割、3D点云



芯片/微电子

通用/垂类识别语料等

全数据服务优势

合作优势

业务场景深入理解与多能力经验

- √多垂类业务场景&具体承接业务的深入理解
- √SPEECH+NLP+CV全能力，成熟的数据采标团队(+海外能力)
- √过往项目实操经验，丰富中大型项目管理经验

Advantages

合作优势

扎实的技术能力、平台能力、项目基础

- √标贝全数据服务能力，技术平台能力，稳健经营能力
- √标贝大量垂类资源储备
- √进行中项目管理能力、执行能力、交付能力、数据安全等方面

良好合作基础、强烈的合作意愿，良好的现金流

- √过往对标贝的认知，标贝强烈的合作意愿
- √KA客户业务专项组，优秀的服务能力
- √公司营收持续向好，充沛现金流

Partners

—国内领先的人工智能全数据服务商—

ByteDance
字节跳动

HUAWEI

Tencent 腾讯

MI

OPPO

VIVO

HONOR

Alibaba.com
阿里巴巴

Microsoft

Baidu 百度

JD.COM 京东

美团

bilibili

科大讯飞
iFLYTEK

快手

网易
NETEASE

Haier

美的 Midea

长虹

Hisense

TCL

Partners

—国内领先的人工智能全数据服务商—



Lenovo



贝壳
找房大平台



平安科技
PING AN TECHNOLOGY



Sogou 搜狗



思必驰
AI SPEECH



云知声
Unisound



iQIYI
爱奇艺



海康威视
HIKVISION



度小满金融
Du Xiaoman Finance



WeBank



马上消费金融



喜马拉雅FM



360
www.360.cn



出门问问



HUDUN



中国科学技术大学
University of Science and Technology of China



清华大学
Tsinghua University





标贝科技-人工智能数据服务商

--短期的营收能力、长期的市场格局!