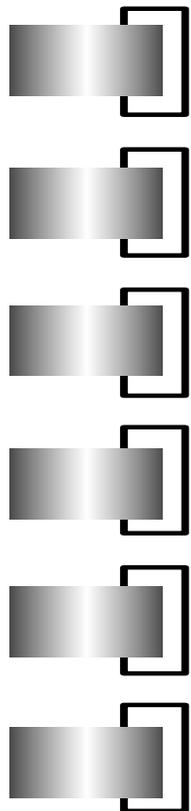




点石平台产品介绍

百度 大数据部

目录



- 百度公司介绍
- 点石平台产品介绍
- 点石平台案例分享

百度：不只是一家搜索公司

百度一下

百度：让人们最平等便携地获取信息找到所求

最常用的中文搜索引擎

最具影响力的互联网公司



92.1% 约有92.1%的网民使用百度PC端

667 Millions 百度手机端月活跃用户数约有6.67亿

90.3% 约有90.3%的无线网络用户使用百度

72 Billions 每天在百度地图上约有720亿个位置服务请求

6 Billions 每天响应约60亿次的搜索

百度：用科技让复杂的世界更简单

十八年的健康成长，百度已不只是“搜索”



848亿元
2017年全年营收



约900亿美元
目前市值



40,000+人
员工总数，其中
工程师超过三分之一



每天响应数十亿次搜索请求
移动搜索6.65亿活跃用户



月活跃用户3.48亿
每天为用户导航超过2亿公里
覆盖海外209个国家



超过15亿注册用户
2162万个主题吧

立足技术创新，引领产业发展

● 注重技术积累，不断加大投入

5

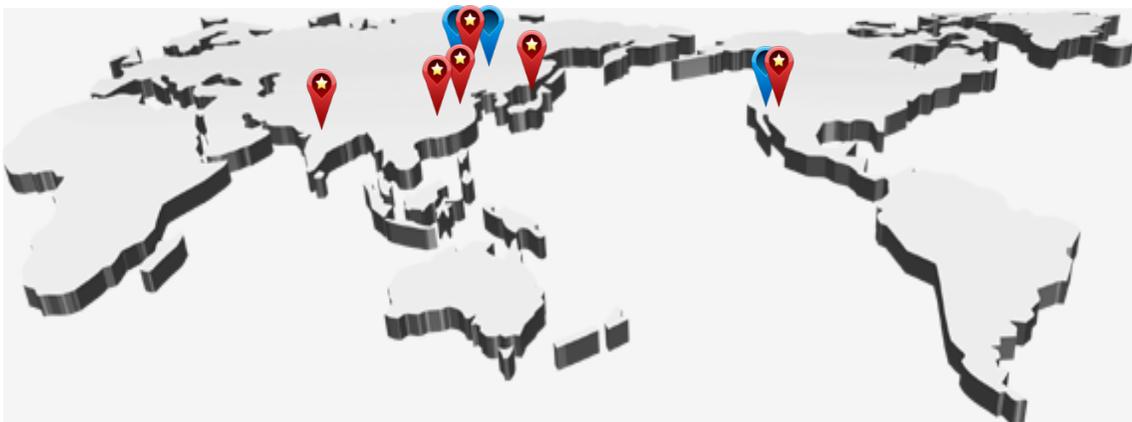
大实验室

6

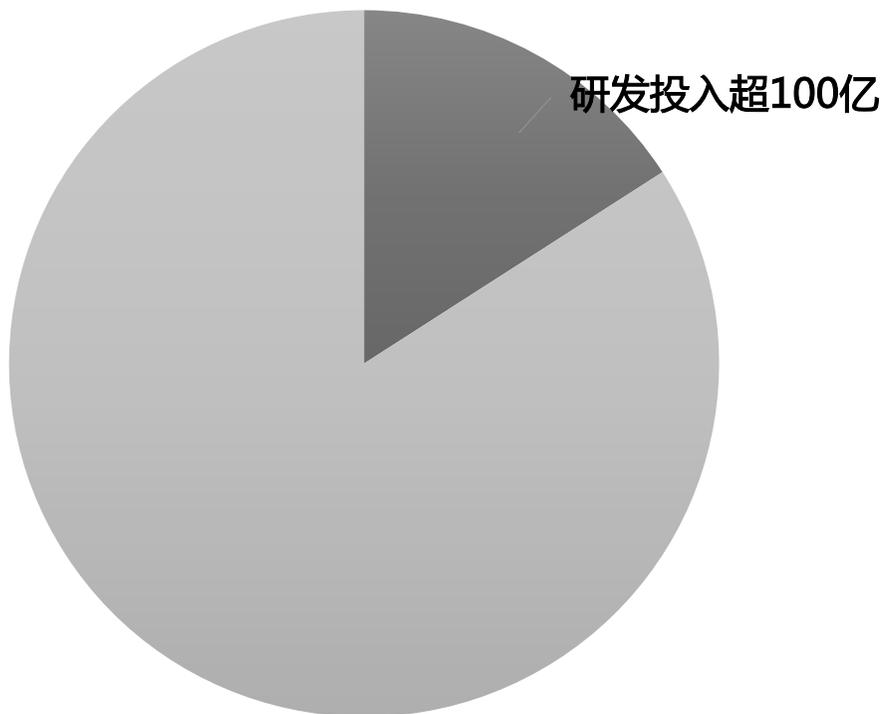
大研发中心

32

项国家重点科技专项



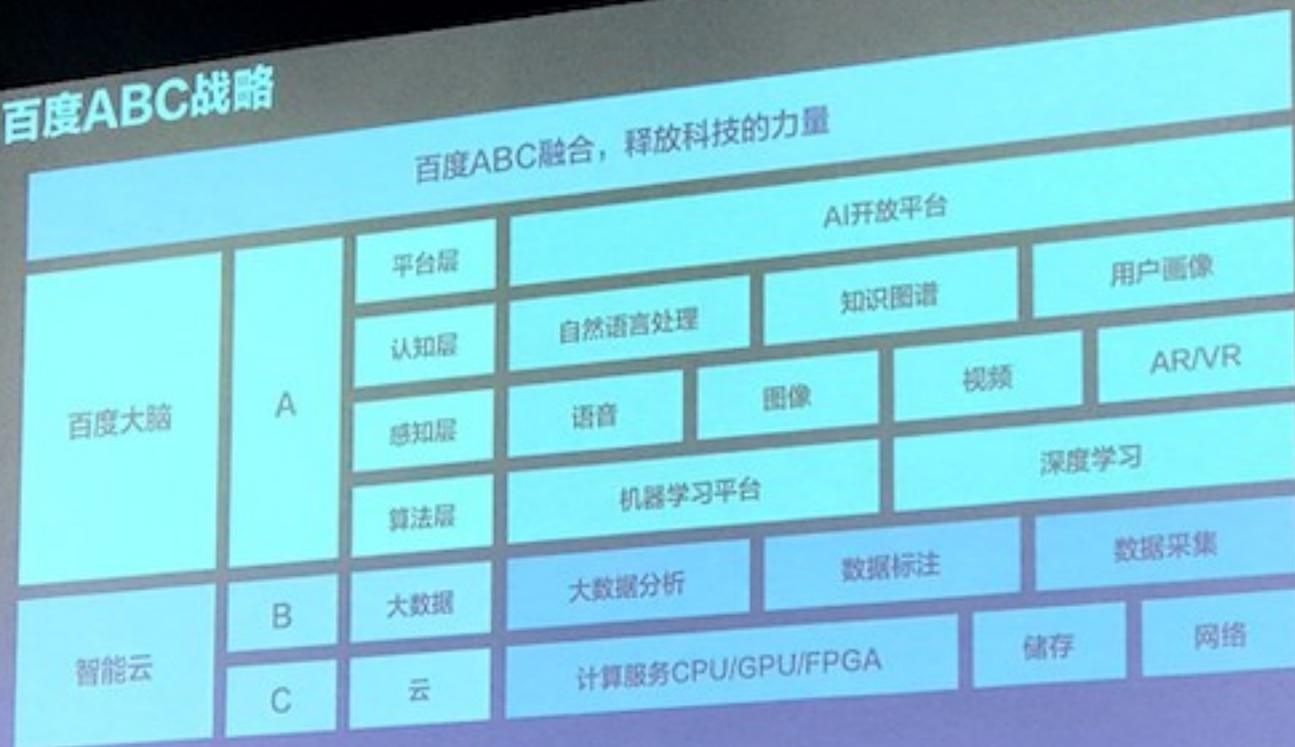
百度基础研究领域的投入



- 2015、2016、2017连续三年研发费用超过100亿人民币
- 2016年12月27日，欧盟委员会发布“2016全球企业研发投入排行榜”，百度作为唯一中国互联网企业上榜
- 2017年，研发费用占公司营收的15.84%，与谷歌等企业持平，远超中国企业500强1.48%的平均研发占比

I 百度ABC战略

百度ABC融合，释放科技的力量



百度在人工智能领域发展迅猛，储备雄厚

技术储备

2000人研发团队

人工智能研发团队

2000项专利

申请人工智能专利数量

50%出自百度

20位引领中国AI变革的技术专家

ABC优势

AI

世界最大 深度神经网络 (DNN)

全球第10 超级计算机 MINWA (600TFlops)

世界最大 深度机器学习开放平台

Big Data

世界第一 通用推荐引擎

中国第一 数据预测开放平台 1,000+ PB

数据存储, 100+ PB 日数据处理量

Cloud Computing

10,000+ 单集群服务器规模, 100+亿文件

IDC资源管理平台, 100,000+ 服务器管理量

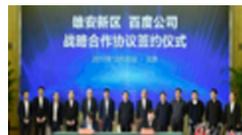
PUE <= 1.22 , 10,000+ 日交付服务器数量

发展里程碑



2018.3 量子计算研究所

开展量子计算软件和信息技术应用业务研究



2017.12 雄安战略合作

雄安新区与百度公司签署战略合作协议



2017.3 国家实验室

深度学习技术及应用国家工程实验室百度揭牌



2016.11 首次试运营

百度无人车乌镇首次试运营



2015.12 点赞无人车

世界互联网大会上国家领导人点赞百度无人车



2015.3 中国大脑计划

李彦宏政协提案建议设立“中国大脑”计划

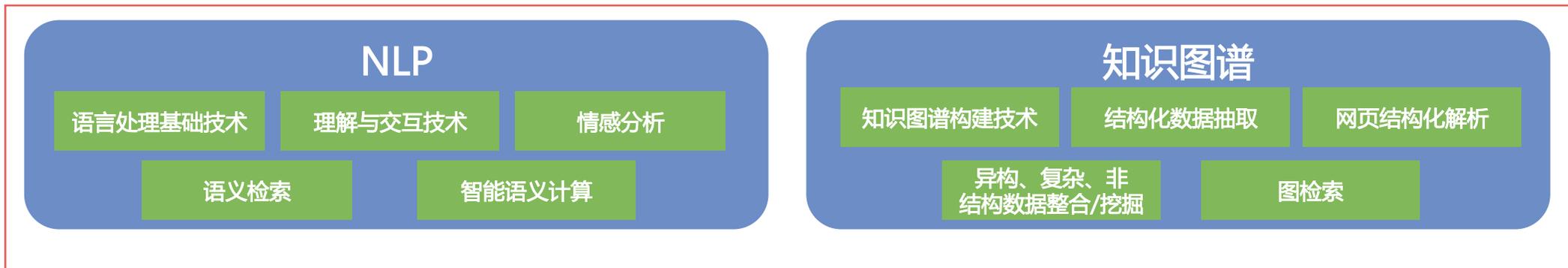
打造开放生态布局，助力中国AI产业发展



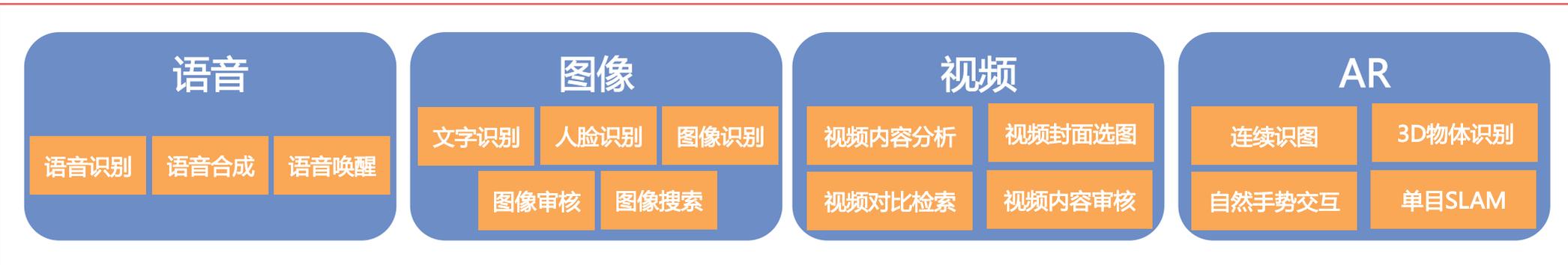
7月5日，百度在北京国家会议中心举办了首届“百度AI开发者大会”，吸引了超过6000名开发者及合作伙伴参加，是中国工业界第一次大规模的AI 开发者大会，首次向行业公布了百度最完整的AI生态开放战略

百度AI技术能力积累

认知智能



感知智能



数据智能



数据智能从1.0到2.0

行业解决方案	政务		零售		保险		工业 IoT		
	舆情	城市综治 明察	商家经营大脑	场内分析 智客	数字化营销 觅客	保险风险定价	保险反欺诈	异常检测	
	智慧警务 明镜	智慧环保 度环境		场外分析 汇客	可视化大屏 炫客	精准营销	智能保顾	边缘计算模型	
对外开放产品	鲁班				孔明				
	数据处理		数据分析		营销分析		统计分析	知识理解	其它
	数据传输 Minos	数据工厂 Pingo	搜索分析 Baidu-ES	数据仓库 Palo	客情	司南	网站统计	智能知识库 智周	智能推荐
	数据治理 Dayu		报表可视化 Habo	数据分析开发 Jarvis	商情	司南户外	移动统计	知识生产工具 智珠	
基础能力	数据工程			数据科学					
	分析型数据库	批量和流式处理	数据采集与传输	用户理解	舆情分析	风控建模			
	可视化分析	搜索分析	异构计算	知识图谱	推荐预测	深度关系模型			
	数据治理	自动机器学习	可信数据计算	时空大数据	领域知识理解	精算建模			

点石平台产品介绍

百度一下

百度点石：大数据众智开放平台

百度点石是百度大脑面向企业客户的大数据服务平台，面向个人与企业开发者提供从数据收集、存储、建模、部署上线等全流程的开发与交易服务，为大数据生态赋能。



<http://dianshi.baidu.com/gemstone/home>



交易市场
海量数据集集成，满足定制化需求



大数据竞赛
挑战机器学习难题，赢取丰厚奖励



数据安全流通
数据存证、数据流通、数据血缘



Jarvis
数据强大易用的交互式数据开发工具

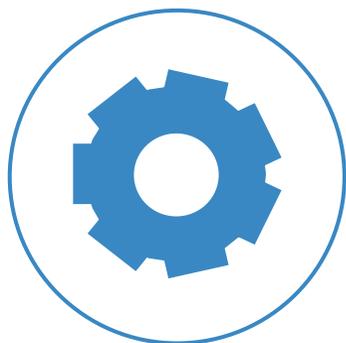
点石优势

实现各种垂直应用和各种用户数据聚合，在不同应用的数据库匹配过程中产生边际效益，增强各个用户的用户黏性，实现“1+1>2”。只有构建起强大的数据库，具有强大的存储能力、计算能力、分析能力，才能够把大数据真正用起来。



数据优质

百度自有数据
合作方授权数据
开放数据集



服务多维

百度AI基础能力开放
开发者可二次开发
垂直行业解决方案组建



环境安全

区块链数据确权
数据安全隔离域
多方安全计算

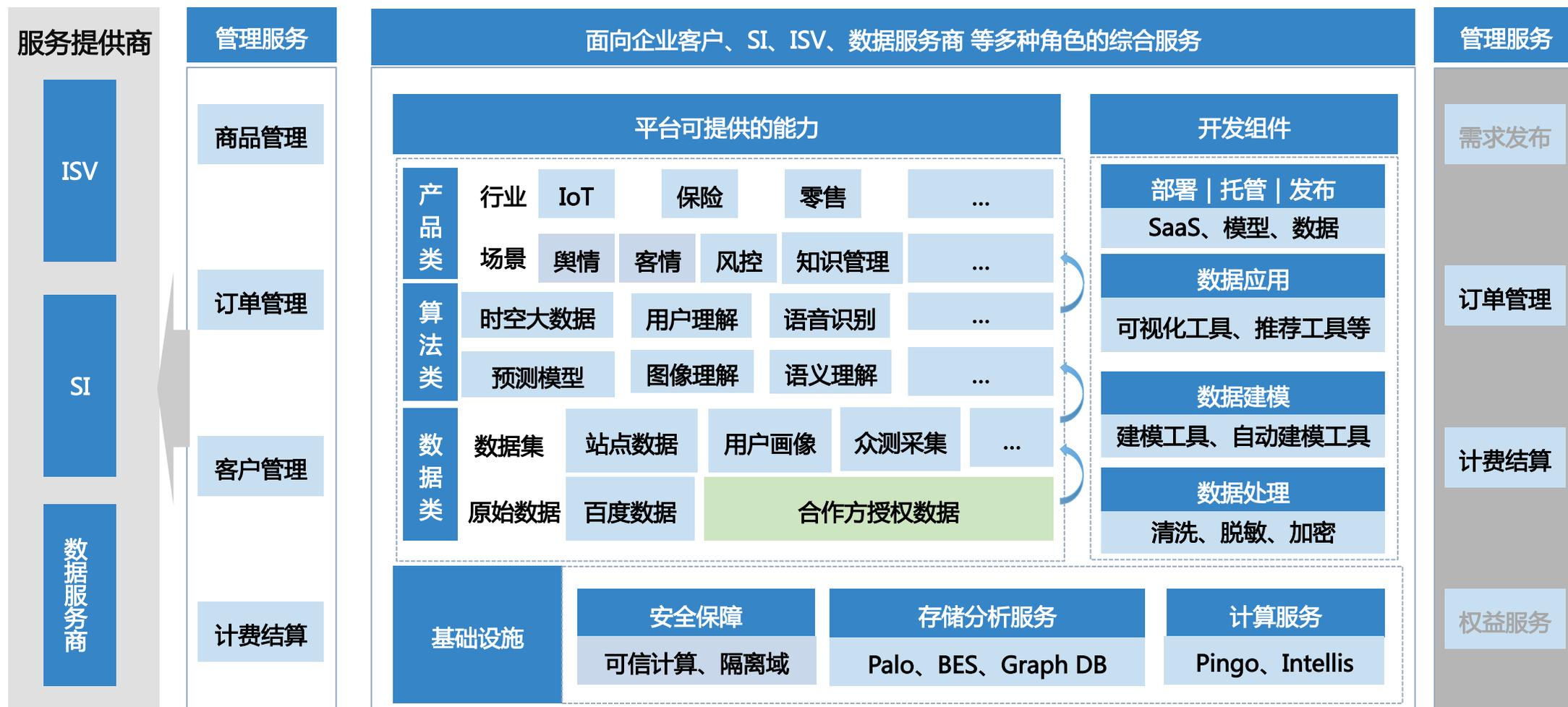


便携运维

一站式机器应用开发
百度云服务托管运维

整体架构图：一站式数据开发、交易平台

致力于打造以数据为驱动的一站式大数据分析和人工智能应用在线开发、服务托管和交易流通的生态。



产品功能概览：全流程数据开发与托管服务

应用层

账号管理

订单管理

交易管理

权限管理

应用管理

托管层

镜像管理

容器管理

服务运维

应用部署

开发层

应用开发

机器学习建模

服务编排

持续集成

接入层

数据管理

资源管理

资源层

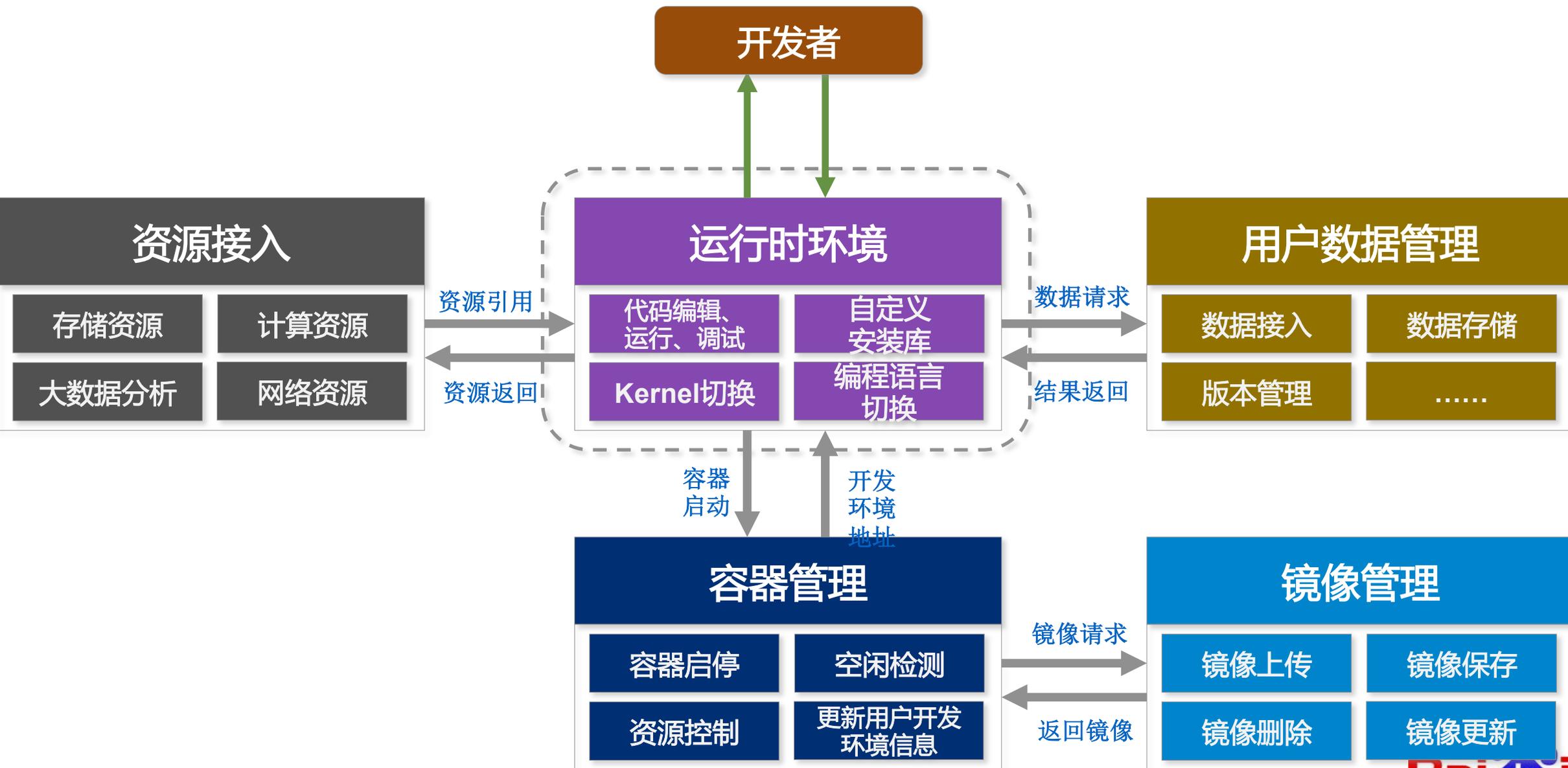
云平台资源

组件库

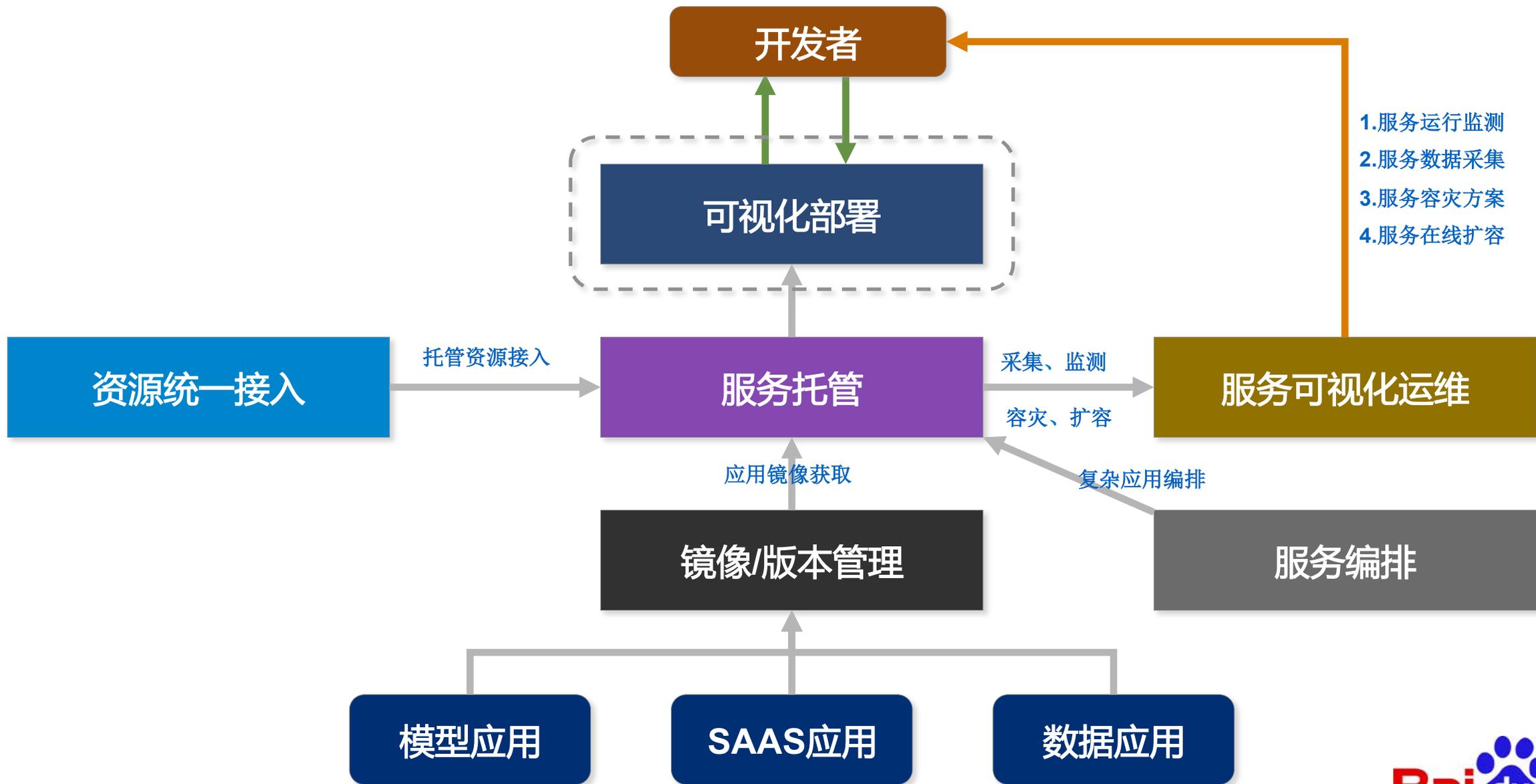
算法库

安全方案

主要功能建设方案：应用在线开发技术建设方案

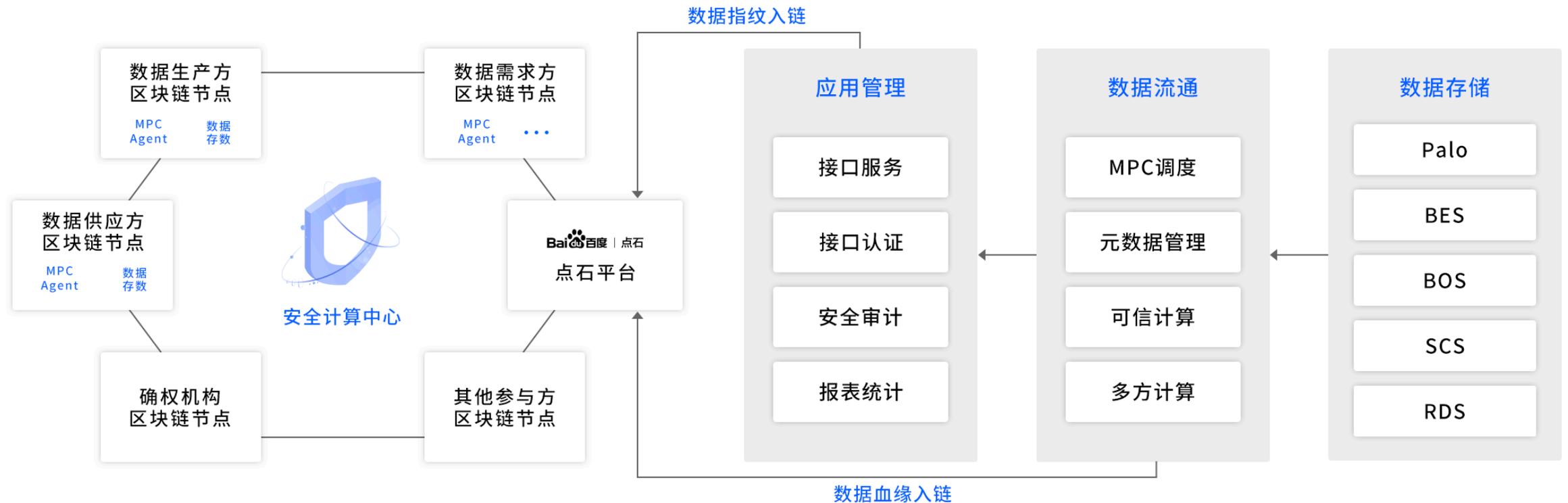


主要功能建设方案：服务托管技术建设方案



如何保障数据安全流通？

数据安全流通通过结合区块链、可信计算、智能合约等技术，为用户提供数据存储、托管、确权、认证加工、血缘图谱等服务，保证数据安全流通，为数据生产方提供数据隐私和权益的保障，为数据需求方提供高价值数据的使用。



如何高效进行数据开发？

Jarvis是为开发者打造的强大、易用的交互式在线数据开发工具。支持Python等多语言的交互式开发环境，秒级启动；集成百度PaddlePaddle以及TensorFlow等优秀开源深度学习框架。



交易市场商品类型

- API pipeline
- API组合

高级API/web service

- 舆情大屏

可视化大屏

- 第三方独有数据集
- 公开数据集

数据块/集

- 如计算资源CCE
- 存储资源BOS

云平台资源

- 情感分析基础模型
- 图像迁移学习基础模型

base模型

- 基础机器建模
- 自动化机器建模

机器学习应用/模型预估服务

- 舆情应用
- CRM

SAAS

- 保险投顾
- 基金人群定位

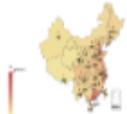
行业/垂类解决方案

交易市场页面展示

来源: **百度** 开发者

标签: **全部** 新闻 传播图 情感 热搜词 头条 门户网站 高管 话题 热门 股东

排序方式: 时间 ↓ 价格 ↓ 收藏量 ↓ 购买量 ↓ | 价格: **全部** 付费 免费 共 28 条数据

 <p>新闻地域分布</p> <p>0 元</p> <p>浏览量: 2563 收藏量: 0 购买量: 0</p>	 <p>网页新闻传播图</p> <p>0 元</p> <p>浏览量: 1033 收藏量: 0 购买量: 0</p>	 <p>头版头条</p> <p>0 元</p> <p>浏览量: 822 收藏量: 0 购买量: 0</p>	 <p>行政处罚</p> <p>0 元</p> <p>浏览量: 700 收藏量: 0 购买量: 0</p>	 <p>情感分布</p> <p>0 元</p> <p>浏览量: 338 收藏量: 0 购买量: 0</p>
 <p>实时热点</p> <p>实时热点话题</p> <p>0 元</p> <p>浏览量: 281 收藏量: 0 购买量: 0</p>	 <p>企业工商信息</p> <p>0 元</p> <p>浏览量: 197 收藏量: 0 购买量: 0</p>	 <p>词云</p> <p>0 元</p> <p>浏览量: 184 收藏量: 0 购买量: 0</p>	 <p>分类热点新闻</p> <p>0 元</p> <p>浏览量: 146 收藏量: 0 购买量: 0</p>	 <p>分地域热搜词</p> <p>0 元</p> <p>浏览量: 130 收藏量: 0 购买量: 0</p>

谁是百度数据生态的合作伙伴？

独立软件供应商 (ISV) 系统集成商 (SI)

将自身产品服务与百度大数据产品服务相结合，为客户提供基于百度大数据、AI能力的通用行业解决方案或垂直行业解决方案，与点石平台合力打造良性的大数据生态。

解决方案提供商

数据采集商 数据标注商

可根据特定领域、特定场景的客户需求，提供定制化的数据获取与加工方案的设计与执行服务，数据类型涵盖文本、图像、音频、视频、网页等。

数据服务商

线上、线下培训合作伙伴 课程制作合作伙伴 企业内训合作伙伴

结合自身在教育培训方面的优势、对百度大数据产品及技术感兴趣的企业及用户提供在线培训、线下教学、企业内训等服务。

课程培训机构

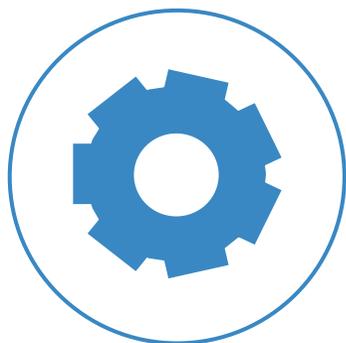
为什么要成为百度大数据生态合作伙伴？

实现各种垂直应用和各种用户数据聚合，在不同应用的数据库匹配过程中产生边际效益，增强各个用户的用户黏性，实现“1+1>2”。只有构建起强大的数据库，具有强大的存储能力、计算能力、分析能力，才能够把大数据真正用起来。



数据优质

百度自有数据
合作方授权数据
开放数据集



服务多维

百度AI基础能力开放
开发者可二次开发
垂直行业解决方案组建



环境安全

区块链数据确权
数据安全隔离域
多方安全计算



便携运维

一站式机器应用开发
百度云服务托管运维

如何成为百度点石平台的合作伙伴



点石商家商品开发步骤6步走

01

注册与学习

提供文档教程、使用说明。

02

购买云资源

资源供开发工具或应用部署。

03

创建项目选择开发工具

机器学习jarvis、大屏展示炫客、联合建模天链。

06

发布商品

部署完成的应用上架到点石商城

05

部署应用

将开发完成的应用部署到点石线上环境。

04

进入开发工具进行开发

交互式开发环境轻松开发应用。

大数据竞赛：解决大数据业务难题



大数据竞赛优势

01

业内领先的大数据平台

百度基于多年技术积累沉淀，已推出大数据基础套件，包括：数据存储、清洗等底层基础组件，交互级可视化开发环境等。

03

卓越的大数据行业专家

横跨多行业业务场景，深挖企业核心需求，拥有丰富经验的大数据解决方案专家，将业务分析落地到解决方案。

02

丰富的大数据处理经验

从对内支撑上百条海量数据产品线，到对外建立十余行业多场景大数据解决方案标杆，十载摸索历练沉淀出扎实的大数据处理经验。

04

完备的实战型培训

多年大数据实战经验沉淀，孵化出完备的培训体系，夯实基础联合实战项目，线上课程配合线下实训，引领开发者进阶大数据专家。

合作案例

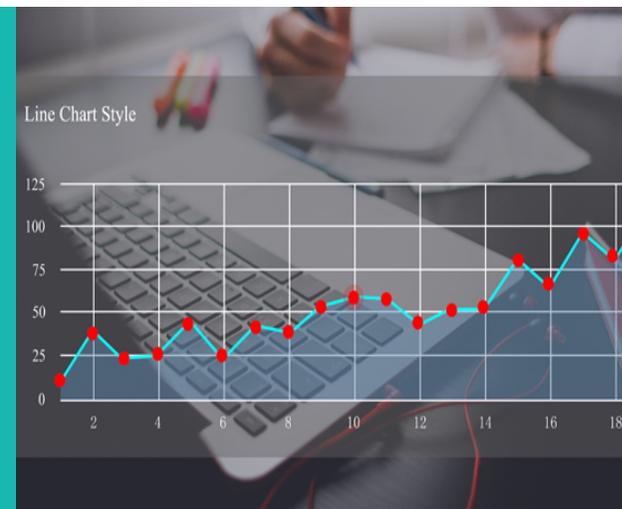


百度地图LBS海量服务支撑

百度地图的LBS服务对于POI的更新速度要求很高，以往通过实地人工采集方式成本大效率低，因而期望通过机器学习、深度学习等方法分析街采或用户上传的商家招牌图片以获得商家名称信息。通过竞赛甄选出的top10方案效果优异，获得百度地图LBS团队高度赞赏。

百度舆情行业风险评估支撑

网络上用户对商家的评论数据，几乎全部带有正中负3种情感极性，在一定程度上可反映商家产品和服务的质量，是商家自我监督的预警灯。评论情感极性判别的人工成本高、效率低，为解决该问题，百度舆情团队希望通过机器学习、深度学习等方法自动化获得用户评论数据的情感极性。通过竞赛甄选出的top10方案效果优异，得到百度舆情团队高度评价，并强调受到重大启发。



点石平台案例分享

百度一下

案例一：商家招牌的分类与检测

大数据竞赛是由百度和西安交通大学联合主办，旨在通过全国高校范围内，挖掘、发现大数据人工智能尖端人才，并为其提供更好的发展环境和实践机会。2018年为第四届，大赛结合图像处理技术，通过街采的或用户在该地点上传的包含该商家招牌的照片，机器检测出图片中的招牌，并且对招牌进行分类，实现降低人工成本。

主办方

西安交通大学 百度

目标

品牌影响

扩大百度在专业人才中的品牌影响力

人才挖掘

聚集顶尖大数据人才，提供学习，实践，奖励和就业机会

业务难题

源自真实的百度地图业务场景，解决实际业务难题

百度-西交大 大数据竞赛2018

赛题：商家招牌的检测与分类

竞赛流程

报名网址：dianshi.baidu.com
报名时间：2018年4月20日-6月7日

报名方式

- 2018年4月20日00:00至2018年6月7日24:00，用百度账号登入点石官网dianshi.baidu.com
- 1-5人组队报名，每人只能参加一支队伍，且每队至少一名在校学生
- 保证参赛队员报名信息准确有效，否则取消参赛资格与奖励
- 竞赛官方交流方式：1、QQ群：482699741；2、百度点石官网社区板块

奖项设置

一等奖	二等奖	三等奖	优秀奖
¥50000	¥30000	¥10000	¥1000

其他奖励

- 百度云智学院免费在线课程培训（百度搜索“百度云智学院”）
- 进入决赛队伍额外获得最高价值5千元百度云资源
- 参赛过程中选手如使用PaddlePaddle深度学习框架并进入决赛，则有机会获得百度相关部门实习机会
- 进入决赛的选手有机会获得百度面试绿色通道（免笔试）

案例一：商家招牌的分类与检测

解决方案

赛题设计

百度地图商家招牌自动分类更新：百度地图的LBS服务对于POI的更新速度要求很高，以往通过实地人工采集方式成本大效率低，因而期望通过机器学习、深度学习等方法分析街采或用户上传的商家招牌图片以获得商家名称信息。本次大赛分为初赛、复赛和决赛三个阶段，其中：初赛为招牌图像分类竞赛，由参赛队伍下载数据在本地进行算法设计和调试；复赛为招牌图像检测+分类竞赛，提供线上开发环境，参赛队伍可依自身需求选择使用；决赛要求参赛者进行现场演示和答辩。

赛事运营



赛事推广

人群精准定位 多渠道高ROI投放 线上线下结合 定向资源邀请 连续系列报道



案例一：商家招牌的分类与检测

成果

参赛队

1139

支

覆盖范围

26

省份

参赛人数

1822

人

参赛高校

263

所

案例二：充电桩智能运维项目

充电桩故障分类与检测

充电桩对于新能源汽车的广泛使用至关重要，其中充电桩的故障检测及稳定性保障需要充电桩服务商投入很多人力成本。本次比赛通过数据挖掘的方式预测分析充电桩的工作状态及影响因素，助力降低充电装人工维护检修成本、帮助充电装改造提供参考建议。



准确率

99%



95%



数据分析时间

2天



1秒



节省成本

1000元/天*2天*2次

*30,000个桩

(北京地区)

=4000万



百度大数据 用科技让复杂的世界更简单

合作咨询：dianshi-help@baidu.com

