



大数据业务简介

定位于领先的数字化转型技术与服务提供商

网易数帆：网易集团TOB企业服务品牌

定位于领先的数字化转型技术与服务提供商

来自哪里

背靠网易杭研

网易技术科研和创新应用的前端
网易集团业务公共技术支撑团队
网易技术对企业客户开放的接口

网易关系

TOB业务排头兵

隶属数字产业事业部，垂直企业
数字化转型技术与服务领域。
与网易智企、网易灵犀办公、网
易伏羲共同担当网易TOB业务

使命与愿景

品牌使命

帮助所有客户成功实现数字化转型

品牌愿景

成为客户首选的数字化转型
技术与服务提供商



依托集团公司优势，大数据产业链横纵向布局，提供高质量服务

客户覆盖区域图

中国地图



审图号：GS(2019)1826号

自然资源部 监制

京津冀经济圈

长三角经济圈

珠三角经济圈

依托网易集团优势

三大经济圈重点布局，服务半径辐射全国；

三大总部：杭州、广州和北京总部；

上海、深圳等地配备办事处；

业务专家+技术专家的双重保障

产品团队：瑞典大学、浙江大学等国内外知名院校；

数据团队：多名Apache committer技术专家；

技术团队：网易首席技术专家带队，硕博比例超过80；

业务咨询团队：各行业咨询实践专家团队；

良好的生态伙伴



一体化，高质量

网易数帆大数据的发展历史

2006

分布式数据库、分布式文件系统、分布式搜索引擎，三驾马车支撑了网易互联网2.0时代的产品

2014

大数据平台猛犸、网易大数据（有数）上线，加速了大数据规模化应用

2018

网易严选、考拉、音乐、新闻等业务相继开始数据中台构建，网易数帆发布“全链路数据中台”解决方案

2022

网易数帆发布数据治理和数据开发一体化“数据治理 2.0”解决方案

2009

开始基于Hadoop做数据分析以及运维

2017

网易数帆大数据正式对外商业化

2020

网易数帆提出“数据生产力”理念，倡导“人人用数据、时时用数据”

众多权威资质认证与外部荣誉认可

■ 国家级组织资质

国家高新技术企业
国家规划布局内重点软件企业
国家企业技术中心
国家级博士后科研工作站

■ 省级组织资质

浙江省重点企业研究院
浙江省企业技术中心
浙江省数字工厂标杆企业
浙江省民营企业数字化转型促进中心
浙江省云计算和大数据省级企业研究院
大数据系统软件浙江省工程实验室
浙江省网易大数据重点企业研究院

■ 管理体系相关资质

ISO 27001 信息安全管理体系认证
ISO 20000 信息技术服务管理体系认证
ISO 9001 质量管理体系认证
CSA STAR Certification 2013 服务管理体系认证
CMMI (三级) 认证

■ 大数据技术认证资质

国家工业信息安全发展研究中心评测鉴定所国产化测试
信通院商务智能工具基础能力国产化评测
华为鲲鹏技术兼容性认证
达梦技术兼容性认证
浪潮技术兼容性认证
麒麟技术兼容性认证
长江计算技术兼容性认证
人大金仓技术兼容性认证
SelectDB技术兼容性认证

■ 大数据团队资质

产研团队成员来大多来自瑞典大学、浙江大学等国内外知名院校，硕博比例超过80%
多名Apache Committer技术专家
业务咨询团队扎根行业多年，为百余家头部企业提供咨询实践服务

■ 权威三方机构及媒体荣誉认可

教育部科技进步一等奖
电子学会科技进步特等奖
《Gartner 2022中国中国分析平台市场指南》数据分析代表厂商
《Gartner 2021中国 ICT 技术成熟度曲线》数据中台领域标杆厂商
《Gartner 2020中国 ICT 技术成熟度曲线》数据中台领域标杆厂商
《Gartner 2019中国 ICT 技术成熟度曲线》Cloud ABI领域标杆厂商
OSCAR尖峰开源技术创新奖（二次开发）
浙江省大数据应用服务创新奖
2022CCF大数据与计算智能大赛优秀案例奖
信通院大数据'星河'案例-行业大数据应用优秀案例
2022爱分析国企数字化最佳实践案例
DTCC 第十一届中国数据库技术大会创新产品奖
InfoQ2021年度最具价值技术团队奖
华泰证券2021年度金融科技最具创新成果奖
2021数博会“十佳大数据案例”
2021长三角金融科技创新与应用大赛“中国金融科技·最佳供应商”
数字化观察网金数奖2021年度最具影响力数字化服务商
中国长三角数字经济大会“数字经济优秀案例企业”
思否 SegmentFault 中国技术品牌影响力企业
海比研究院综合中台优秀企业奖
2019CCF Top10大数据应用最佳实践案例
入选信通院数据治理产业图谱1.0
入选艾瑞咨询2022中国数据中台行业研究报告
入选艾瑞咨询《2022年中国制造业数字化转型研究报告》
入选InfoQ中国技术力量年度榜单
入选e-works-中国制造业SaaS服务市场生态地图
入选爱分析数据智能应用实践报告
入选爱分析数据智能厂商全景报告
入选互联网周刊2022数据中台TOP50
入选中国大数据网2022中国大数据分析行业研究报告
入选字母点评2022中国数据分析与商业智能平台采购指南
入选2021“科创中国”开源创新榜

■ 国家或行业标准制定

《数据治理工具图谱》，全国信息技术标准化技术委员会大数据标准工作组
《数据中台能力模型》，中国通信标准化协会大数据技术标准推进委员会
《DataOps成熟度框架》，中国通信标准化协会大数据技术标准推进委员会
《大数据标准化白皮书》，全国信息技术标准化技术委员会大数据标准工作组
《数据资产管理实践白皮书（6.0版本）》，中国通信标准化协会大数据技术标准推进委员会

数据生产力模型

聚焦：金融 | 制造 | 医药 | 流通 | 国企



模型定义： 充分利用大数据相关技术, 协助行业客户沉淀数据资产, 探索在业务场景中的数据应用价值, 并通过数据运营发挥长期效用。

模型内核：

DataOps：

DataOps是一种将软件工程CI/CD的方法融入数据开发的流程, 基于自动化的数据测试、任务发布等技术, 构建数据发布流水线, 使得数据开发效率更高、交付更加频繁, 交付质量更有保障。

DataFusion：

面向分析的数据治理, 在传统数据治理的基础上, 将数据治理的过程融入数据开发的过程中, 采用开发与治理一体化的方式, 实现数据的长效治理。同时, 结合数据中台的构建思想, 打破烟囱式数据架构, 构建统一的指标管理体系和企业公共数据模型层, 通过服务化的方式对外提供服务。

DataProduct：

数据产品化的核心是打通数据到决策的转化链路: 通过数据可视化直观呈现数据, 以交互式分析辅助决策, 以嵌入式分析随时随地提供场景化决策支持。通过数据门户方式组织数据, 通过决策引擎, 实现数据到决策的转化, 通过机器学习平台提升决策的精准性, 基于连接中心实现决策到业务系统的执行。

模型四要素：

数据技术：

用以数据采集, 存取, 处理, 统计分析, 挖掘, 预测和可视化等一系列操作相应的技术。

数据资产：

基于数据开发与治理一体化方法论, 通过构建企业数据中台, 重塑数据架构, 沉淀核心数据资产。

数据应用

基于行业用户应用场景, 通过数据应用建设方法, 让数据在业务场景中发挥价值, 实现数据价值变现。

数据运营：

基于运营平台, 结合企业自身数据应用要求, 关注与人才培养, 数据文化建设以及持续性运营工作。

大数据产品体系

1 我们是一个基础软件提供商

2 我们支持跨云的战略

3 未来的大数据软件市场一定是分层的软件市场

全面、稳定的大数据产品矩阵

消费者运营平台

内置客户管理、互动管理、营销自动化、数据分析等能力的企业全链路营销中台

用户洞察 会员管理 全域触达 营销自动化

标签画像

多对象标签管理与应用平台，通过规范化标签体系建设，帮助用户构建标签管理和共享机制

标签体系 智能画像 自定义标签 群组管理

流量分析

集数据采集、埋点管理与数据分析为一体的一站式流量分析平台

留存分析 归因分析 路径分析 用户分析

有数BI

大数据敏捷数据分析及可视化平台，满足数据收集、分析、展示等不同阶段需求，用数据连接组织各角色，实现整体效率提升

数据大屏 自助取数 数据门户 智能分析 移动端

机器学习平台

一站式、高性能的分析建模平台，功能覆盖算法开发、模型训练评估、在线推理预测建模全流程

多模态数据 高性能计算 低延时推理 一站式建模 零代码建模

数据开发及管理平台EasyData

一站式大数据管理及开发平台，覆盖大数据开发，任务调度，数据质量，数据治理及数据服务

指标管理 实时计算 离线开发 任务运维 数据服务 数据资产

大数据基础平台NDH

基于最新开源技术打造，内置多种存储计算引擎，并在Hadoop、Spark、Impala等多个核心组件做了功能及性能增强，实现智能运维和任务治理，支持企业级安全管控

HDFS Arctic Kudu Spark Impala Ranger等

他们都在用网易数帆大数据

零售



政企



物流



文创



制造



教育



医疗



金融



50%

合作客户中上市公司占比



11个

主打行业

金融	制造	央国企
流通	医药	物流
科技	能源	教育
文创	电商	...

他们都在用网易数帆大数据



大数据产品介绍内容

- **1.大数据基础设施：NDH**
- **2.数据开发治理平台：EasyData**
- **3.数据可视化平台：有数BI**

1.大数据基础平台(NDH)

网易大数据基础平台 (NetEase Data Hub) 基于最新开源技术打造，内置多种存储计算引擎，并在Hadoop、Spark、Impala等核心组件做了功能及性能增强，实现智能运维和任务治理，支持企业级安全管控



1.NDH六大特性

自主可控制

- 每年300+开源贡献
- 漏洞持续修复
- 国产平台支持
- Impala、Spark等重点组件加强

安全稳定

- 集成Ranger、Kerberos、LDAP等安全管理组件
- 行列权限控制、审计
- 数据加密、脱敏
- 增强回收站

开发便利

- SQL on Spark服务
- 细粒度权限控制
- 多租户
- Z-order支持

智能运维

- 任务异常诊断
- 性能分析优化
- 存储空间分析

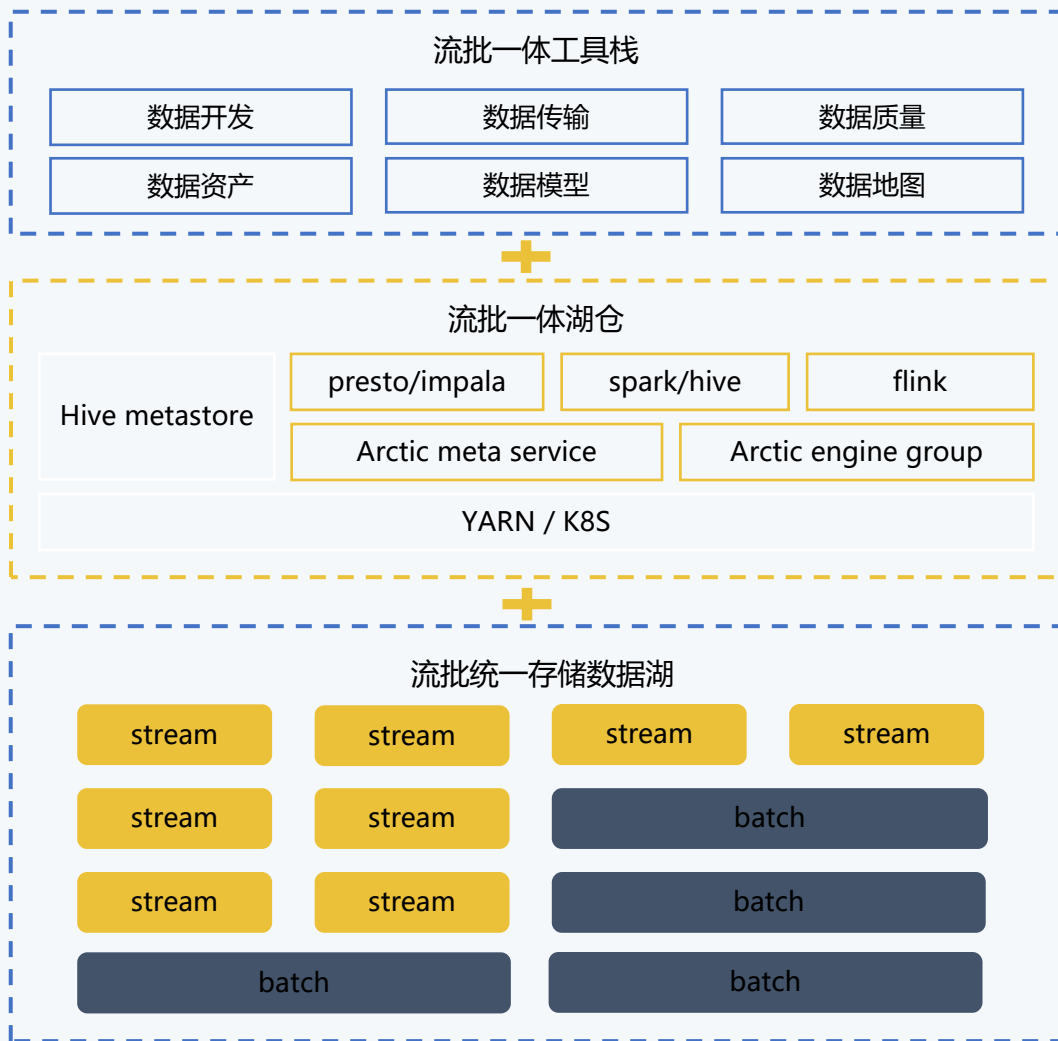
平滑迁移

- 提供迁移工具，支持数据、元数据、任务、权限迁移
- 提供迁移服务
- 不影响业务运转

流式湖仓

- 流式湖仓Arctic
- 统一数据湖存储，流批表一体
- 开源表格式兼容
- 数据自优化服务

1.流式湖仓Arctic架构及特性



Arctic key features

1-数据库特性

- 主键支持 (排序/聚合键)
- 约束保障和结构自优化
- 支持并发读写和数据一致性

2-开放式架构

- 实时读写: flink
- Ad-hoc SQL: spark
- spark/impala/presto 支持

3-服务标准化

- 标准控制台 + 元数据中心
- Hive / iceberg API 兼容

4-流批一体

- CDC 摄取 + 订阅
- 离线 join + 实时 join
- 分钟级数据延迟

1. Apache Kyuubi 成为 Apache 基金会顶级项目



Apache Kyuubi 是一个企业级分布式和多租户大数据服务网关，用于在 Lakehouse 上提供无服务器 SQL。项目于2018年3月由网易开源，于2021年捐赠给 Apache Software Foundation，**2022年12月通过ASF基金会投票决议，成为国际顶级项目。**

[RESULT][VOTE] Graduate Kyuubi from Apache Incubator as a Top Level Project

Posted to general@incubator.apache.org



Kent Yao - 2022年11月21日星期一 GMT+8 18:12:43

Hi IPMC,

The VOTE[1] for graduating Kyuubi from Apache Incubator as a Top-Level Project has passed with 13 binding +1 votes, 28 non-binding +1 votes, and no +0 or -1 votes. The results are (*=binding):

- PMC 13位，分别来自网易，阿里巴巴，eBay，华为等11家国内外企业及组织
- Committer 7位，分别来自Bilibili，网易，爱奇艺百度等7家不同机构
- Contributor 105位，分别来自腾讯、百度，Cisco，亚信、中国移动，爱奇艺，携程，小米等数十家国内外机构
- 落地阿里云、腾讯云、中国移动云，KtNexR, Zillow, Womply、广发证券等百余家企业

2.EasyData产品介绍

一站式大数据开发治理平台，包含大数据开发，任务调度，数据质量，数据治理及数据服务平台

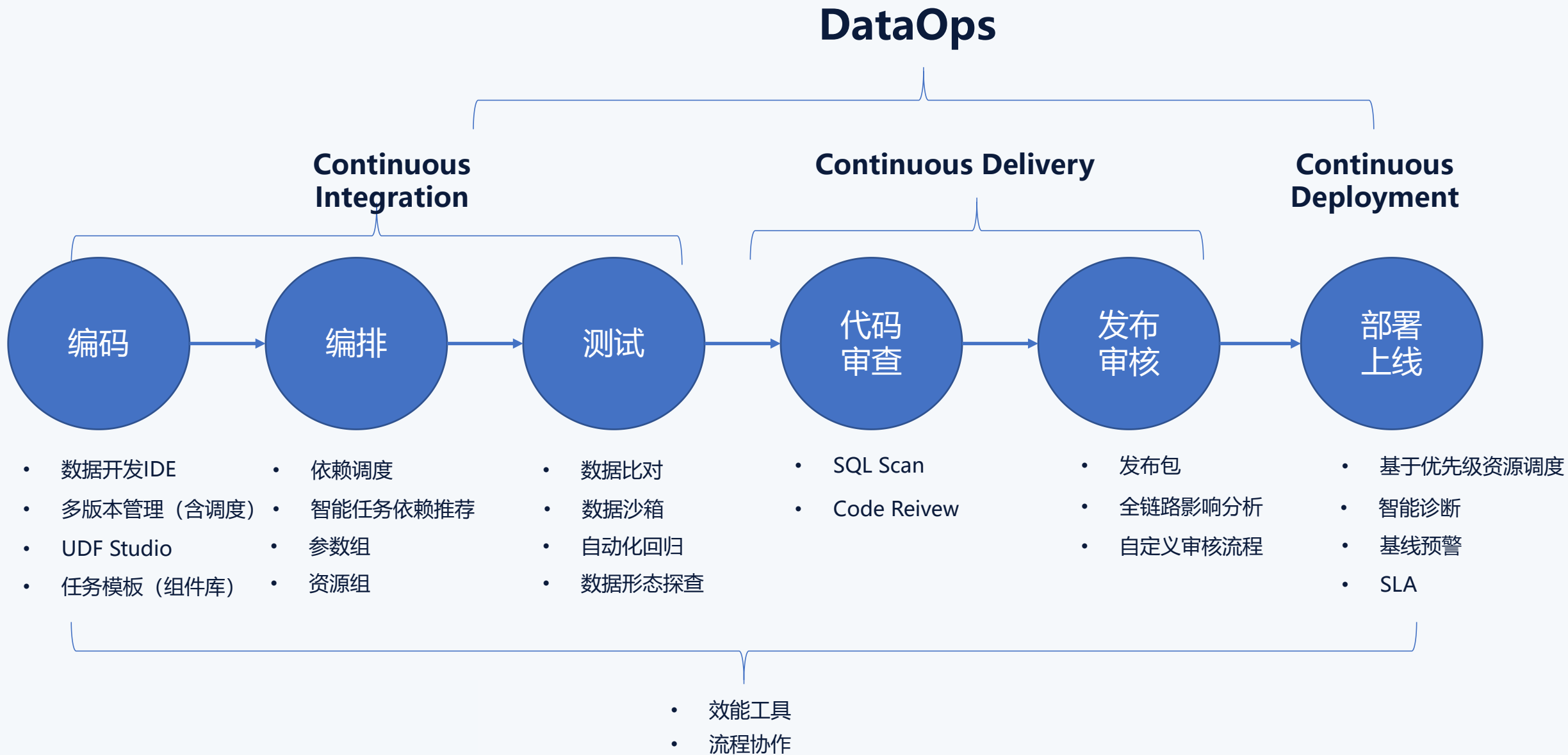


EasyData产品特色

- 基于DataOps的数据开发底座
- 数据开发治理一体化
- 基于DataFabric逻辑数据湖
- 数据治理360

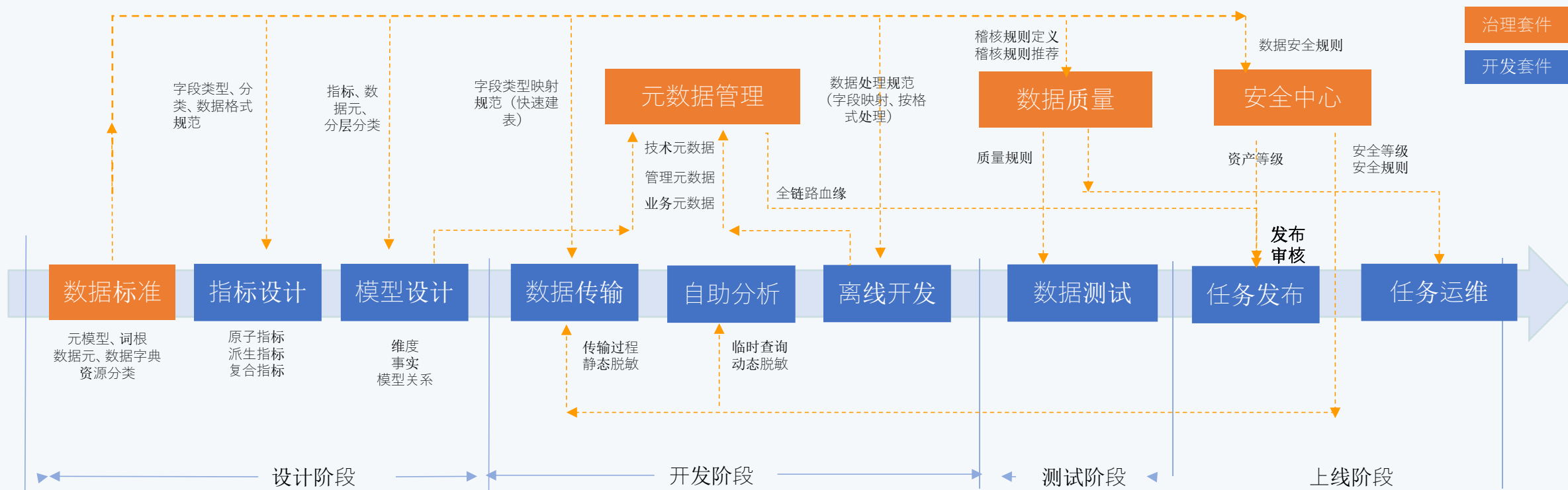
(以上特色案例在浙江电信都有深度应用)

2. 基于DataOps的数据开发底座



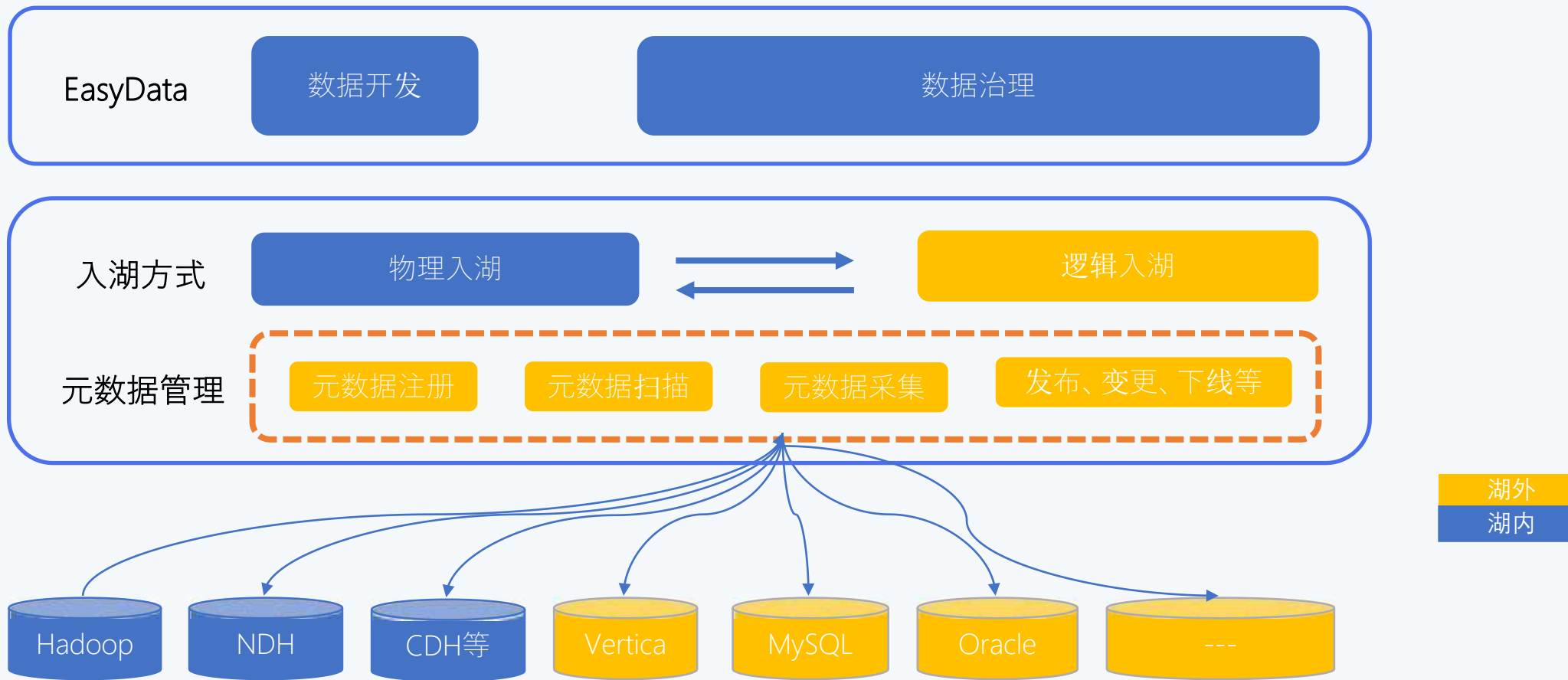
2.数据开发治理一体化

- “先设计，后开发，先标准，再建模”，数据标准成为研发治理一体化的核心
- 将数据治理的流程自然融入数据开发的全生命周期过程中，在数据开发的过程中就完成数据治理



2.构建“物理分散，逻辑统一”的数据中台体系

数据中台构建模式从“Collect”把所有数据收集起来，到“Connect”所有数据有机地连接起来



(浙江电信案例是Veritca，日任务数10000+)

2.数据治理360

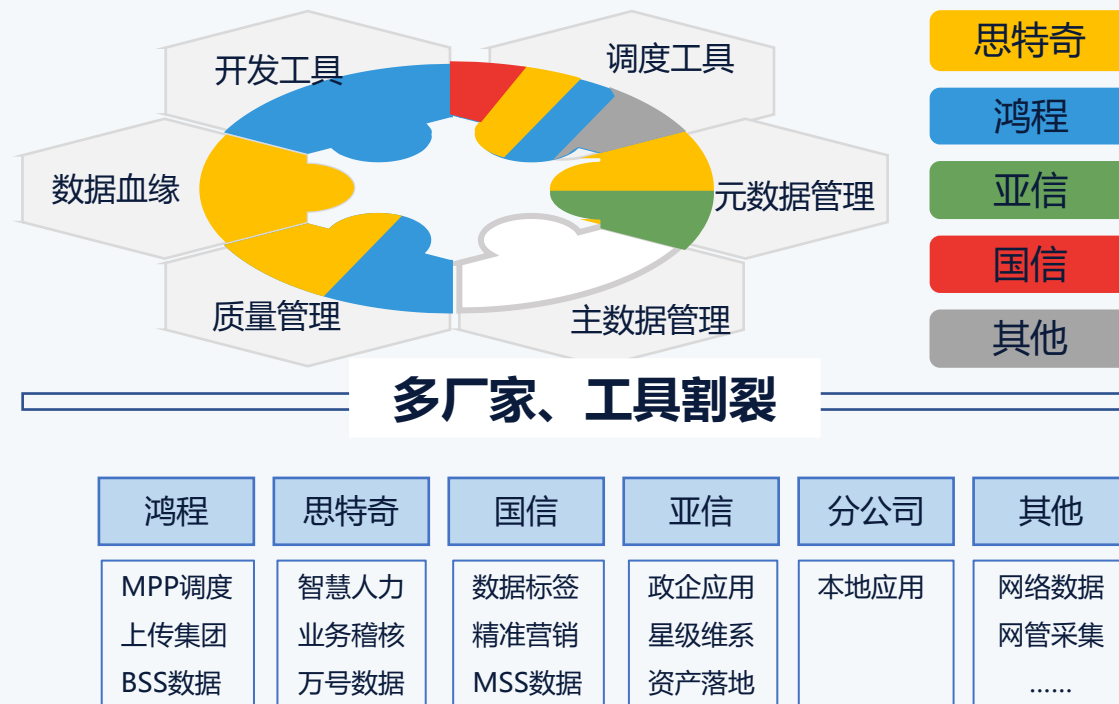
- 从成本、质量、安全、价值、规范五大方面360度评估大数据健康水平
- 核算每个任务、查询、表的计算、存储资源消耗，折算到钱，且分摊到每个数据报表，数据服务API应用层面
- “剥洋葱”式数据下线，从下游不再使用的数据应用开始，逐层向上游任务和下线归档
- 任务和查询成本预估，对于高消耗任务和查询，进行审批管控



2.浙江电信案例背景

在引入网易大数据平台前面临的问题：

- 数据标准、数据质量与数据开发严重脱节，规范只能停留在字典层面，无法融入数据生产的流程中，不能有效的落地执行和监督。
- 不同厂商，不同工具之间严重割裂，数据质量的稽核规则无法和数据标准中数据元的值域约束打通，数据标准中数据元无法跟数据建模工具无法联动，元数据管理中的数据安全等级和安全中心的数据脱敏无法联动。



2.浙江电信方案-数据开发与治理一体化

由数据中台统一为仓库、经分、网络集群提供数据采集、建模、开发、调度、治理等一体化能力。在生产过程中对于程序上下线，建表等操作实现在线化，流程化操作一方面减少人工提升效率，一方面完善数据管控的过程。



标准规范执行

在上线时按照标准规范进行控制，包括命名规范，信息完整性，合理性等，不符合标准规范则不予上线

检查项	检查内容	检查结果
检查指标中文名称	是否填写指标中文名称	OK
检查指标单位	是否填写指标单位	OK
检查指标精度	是否填写指标精度	OK
检查指标业务口径	是否填写指标业务口径	error
检查统计周期	是否填写指标统计周期	OK
检查指标相关部门	是否填写指标相关部门	error
检查指标应用类型	是否填写指标应用类型	OK

将数据开发与数据治理有机结合起来，既是对开发过程的管控，也是保障数据质量的有效方法。

数据中台已有用户数270+人，已迁移上线作业数15000+个，数据质量稽核数100+个，周自助分析累计30000+次，

已导入元数据模型数1500+套，具备作业调度和元数据统一管控能力。

2.浙江电信案例成果一览

质量

- 质量问题下降了63%
- 稽核规则覆盖率达到92%
- 质量稽核规则、数据模型、数据标准的统一，解决了稽核规则完备性的根本问题

价值

- 数据使用用户规模从30+扩展到500+
- 平台用户业务人员的比例从30%提升到98%

安全

- 梳理数据安全等级，治理非法权限策略1200+
- 数据标准与数据安全策略（权限、脱敏）保持一致

成本

- 下线无用数据20%，年约节省成本超过100W
- 构建了精细化的管理机制

标准

- 完成企业数据标准定标500+，发布电信行业数据标准
- 申报行业标准规范

效率

- 需求延期率下降到1%
- 平均需求交付时间从一周提升到3天

3.有数BI

- 企业级敏捷BI辅助用户更智能的完成数据分析
- 扩展数据大屏、数据填报、自助取数、数据门户、复杂报表、智能决策等子产品，帮助用户打造自己的数据产品



3.有数BI-丰富的数据源类型

支持丰富的数据源

简单配置即可实现数据快速接入



支持30多种主流数据源



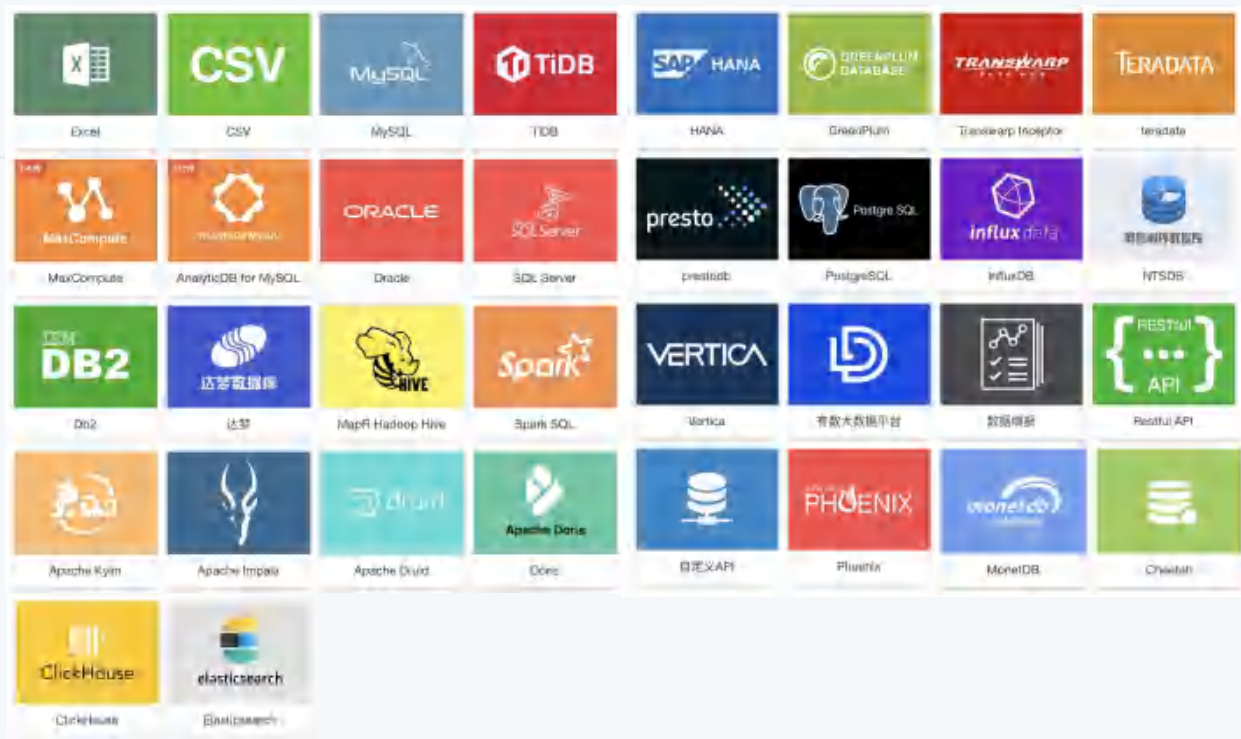
支持对接多种国产数据库：达梦，TIDB



支持以Restful API及自定义API方式连接数据



支持以Kerberos方式连接Hadoop



3.有数BI-数据准备助力用户构建轻量级数仓



拖拽式建模

通过鼠标拖拽直接进行数据之间的关联，快速实现模型建立



学习门槛低

无需代码基础，快速上手



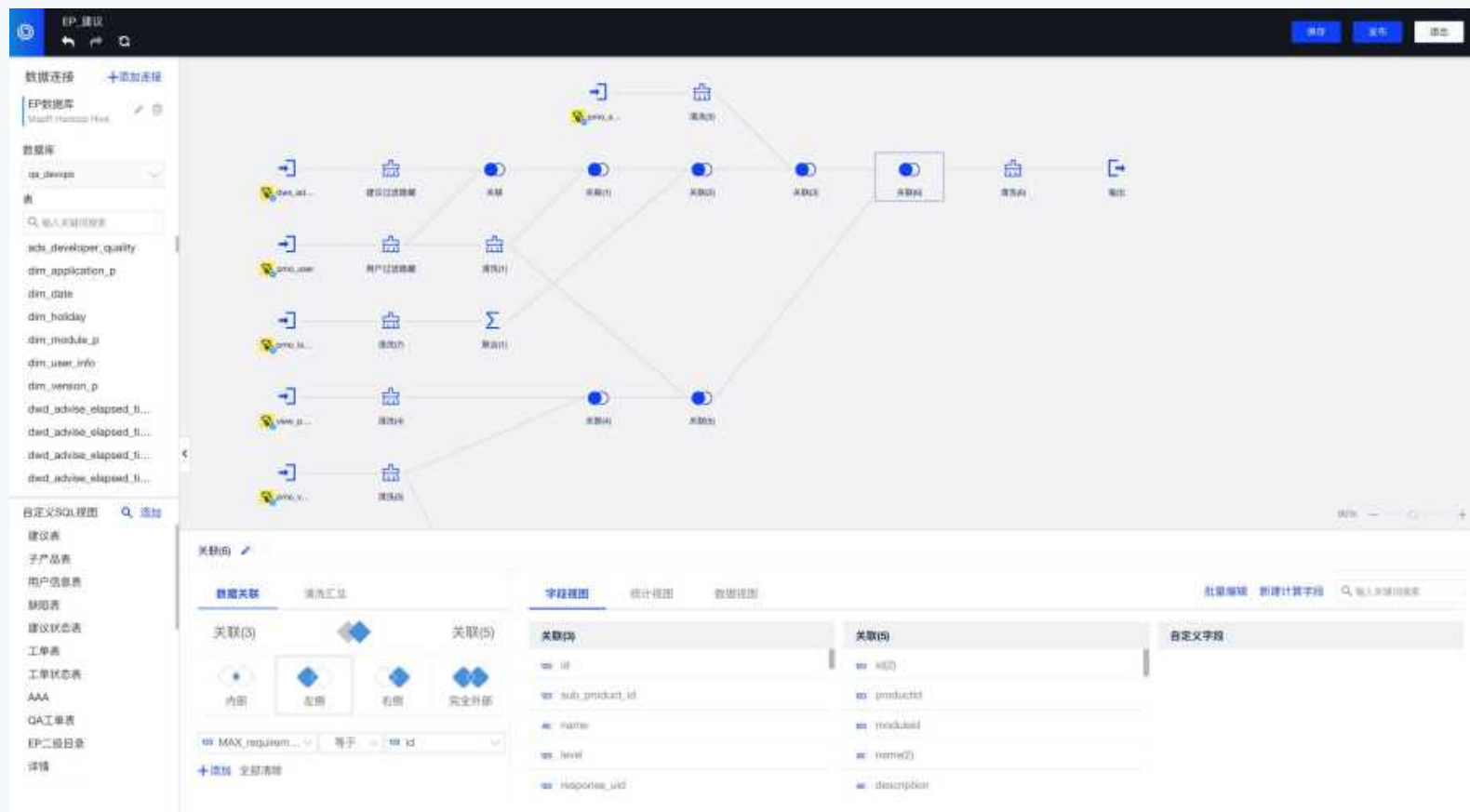
多源数据整合分析

不同来源数据统一建模，满足多源异构的数据分析



支持多种表关联方式

更灵活的建模方式，满足多样化的建模需求



3.有数BI-PPT模式的自助可视化分析报告

快速制作

美观优雅

组件丰富

智能排版



自助式报告制作: 丰富的可视化组件, 只需鼠标拖拽即可完成理想的报告



自由布局: 自由进行图表页面的布局, 支持图表的堆叠, 分组, 拖拽



主题颜色: 支持预置与整体调整报告主题

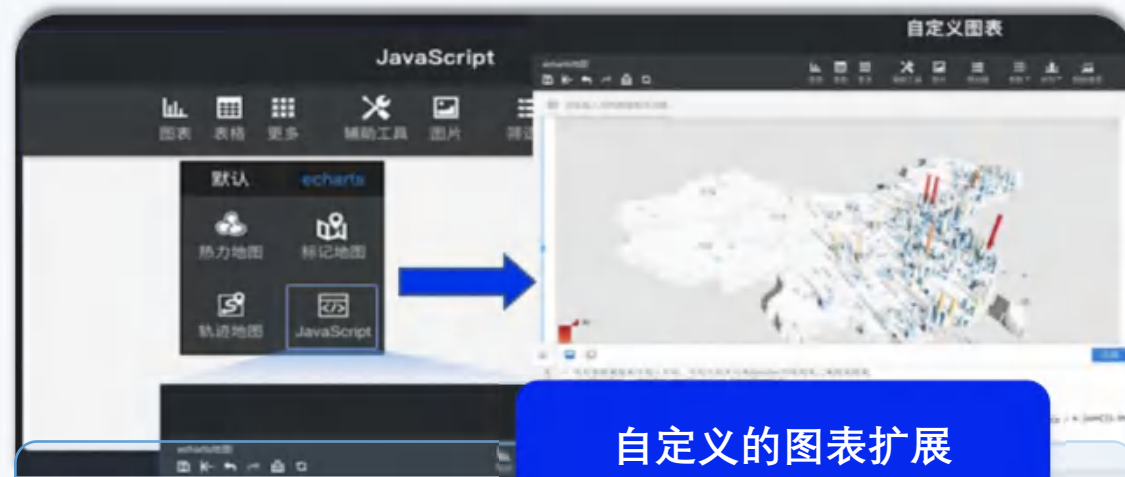


一键美化: 一键实现优雅布局、报表优化, 快速修改报告布局

3.有数BI-丰富的图表组件支持



- 比较类：柱状图、并列柱状图
- 构成类：堆叠柱状图、堆叠百分比柱状图、饼图、环图、玫瑰图
- 关系类：散点图、气泡图、透视表、突出显示表
- 趋势类：平滑折线图、标记折线图、柱线图、区域图
- 分布类：词云、气球图、树图、甘特图、帕累托图、多度量图
- 地图类：标记地图、填充地图、自定义填充地图、自定义背景地图
- 其他控件：表格（条件格式）、指标卡、仪表图、漏斗图、雷达图、桑基图、旭日图




有数BI将数据集形成标准的接口，实现了灵活的图表拓展功能：

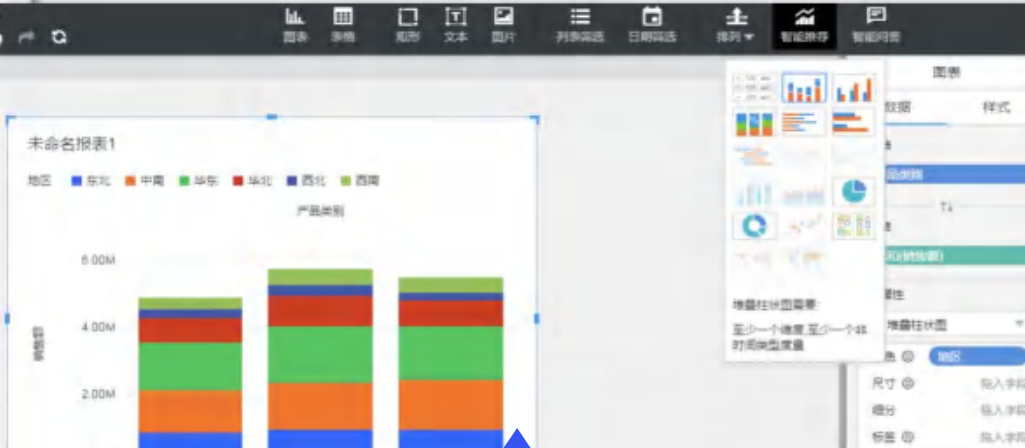
- JavaScript对接：系统支持直接通过JS对有数BI的数据集进行调用，可以支持多样的数据图表应用拓展；
- 图形开放：对接全量echarts图形，全面对外开放实现数据赋能；

3.有数BI-智能可视化分析

辅助下的高效洞察能力
增强分析




智能问答：自动生成并建立最佳数据洞察，帮助用户更好的进行数据分析



智能推荐：根据用户所选数据推荐更合适的分析图形



智能分析：帮助用户挖掘蕴藏在数表格中的价值



增强分析：支持时序预测；支持线性、指数、对数、多项式等趋势线分析

3.有数BI-移动端随时随地洞见数据



有数BI移动端采用跨平台技术，完美适配包括企业微信、钉钉等多套平台。同时通过独创的小程序式设计，用户在访问报告时可以获得沉浸式的数据阅览体验，随时随地享受数据的盛宴。

跨平台支持

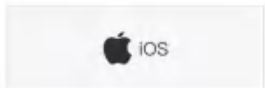
支持 iOS、Android 平台，对接钉钉、企业微信

沉浸式设计

沉浸式应用设计，降低用户对于数据分析的认知阻力，更加专注于数据的探索

移动数据分析

下钻、跳转以及参数等高级功能让移动端也有数据分析能力



3.有数BI-满足中国式报表制作与看数需求

2016年财政教育经费投入情况调查表

表头多层嵌套

学校类别	财政教育经费投入(万元)										其他投入					补充材料		
	合计	新办项目				基建 拨款	村办				社会捐助		信息化建设					
		总计	比上年 增长(%)	人员 经费	日常 公用 经费		项目 经费	其中		其中		本年 投入 总额 (万元)	信息化投入					
								标准化 建设	信息化 建设	合计	投入 总额 (万元)		本年 投入 总额 (万元)	本年 建成数				
合计																		
其中:小学																		
初中																		
高中																		
职业院校																		
民办																		

合计行在表格上方

中国式复杂报表设计器，敏捷企业级 Web 报表工具，类 Excel 交互设计，满足企业对中国式复杂报表的所有需求，拖拽式报表制作，轻松开发复杂报表，支持集成有数BI报告、数据门户，统一分析企业数据。

交互式的操作体验

类 Excel 交互，零代码设计，拖拽点选、快速响应、即时生效、灵活调整，简单易上手

数据源的统一管理

与报告数据源打通，一次性建模，多处共用，有效解决数据分散及数据壁垒问题

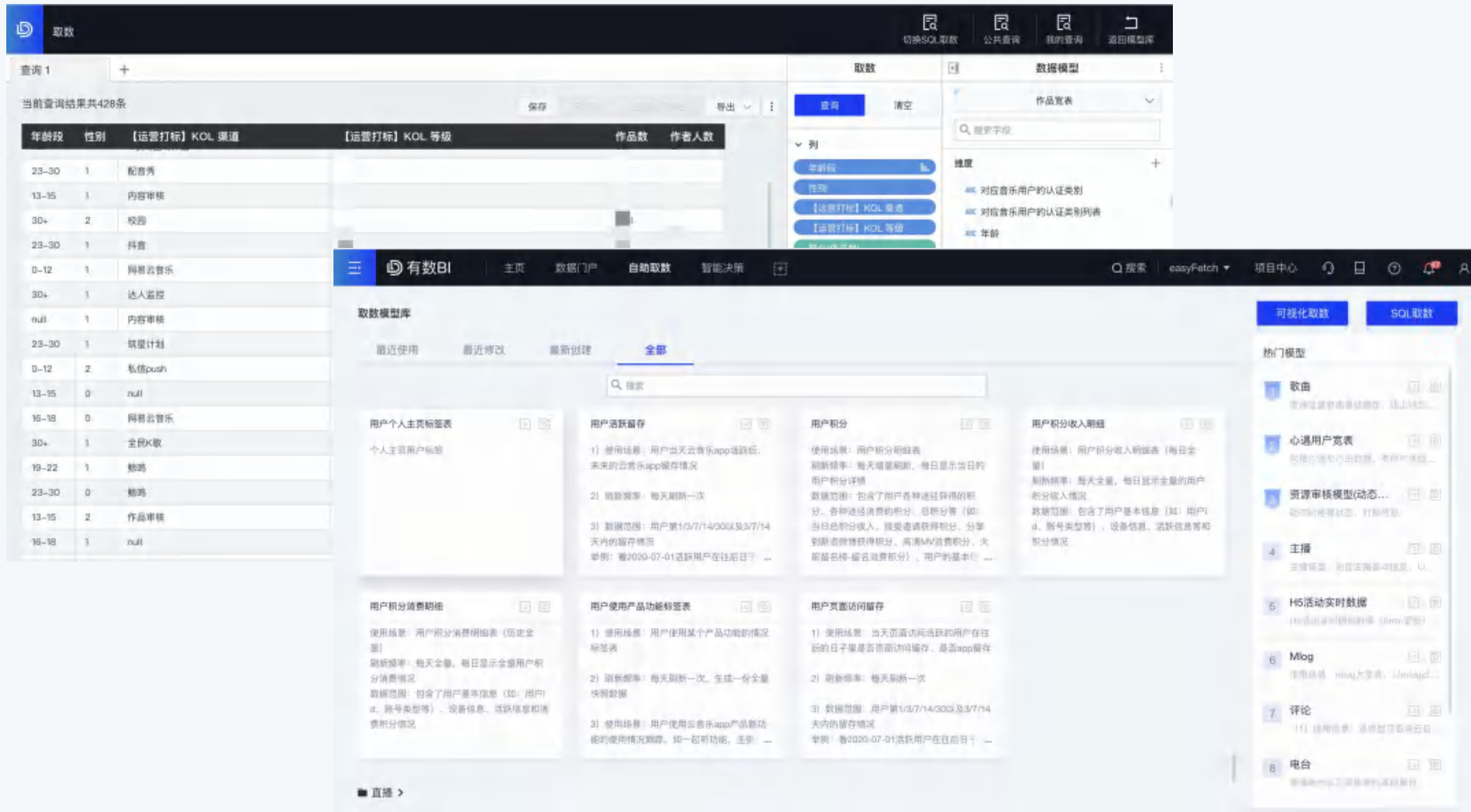
丰富的报表功能

多源分片、行列对称、冻结表头、不规则分组、动态格间计算，轻松应对中国式复杂报表

客户订购产品信息

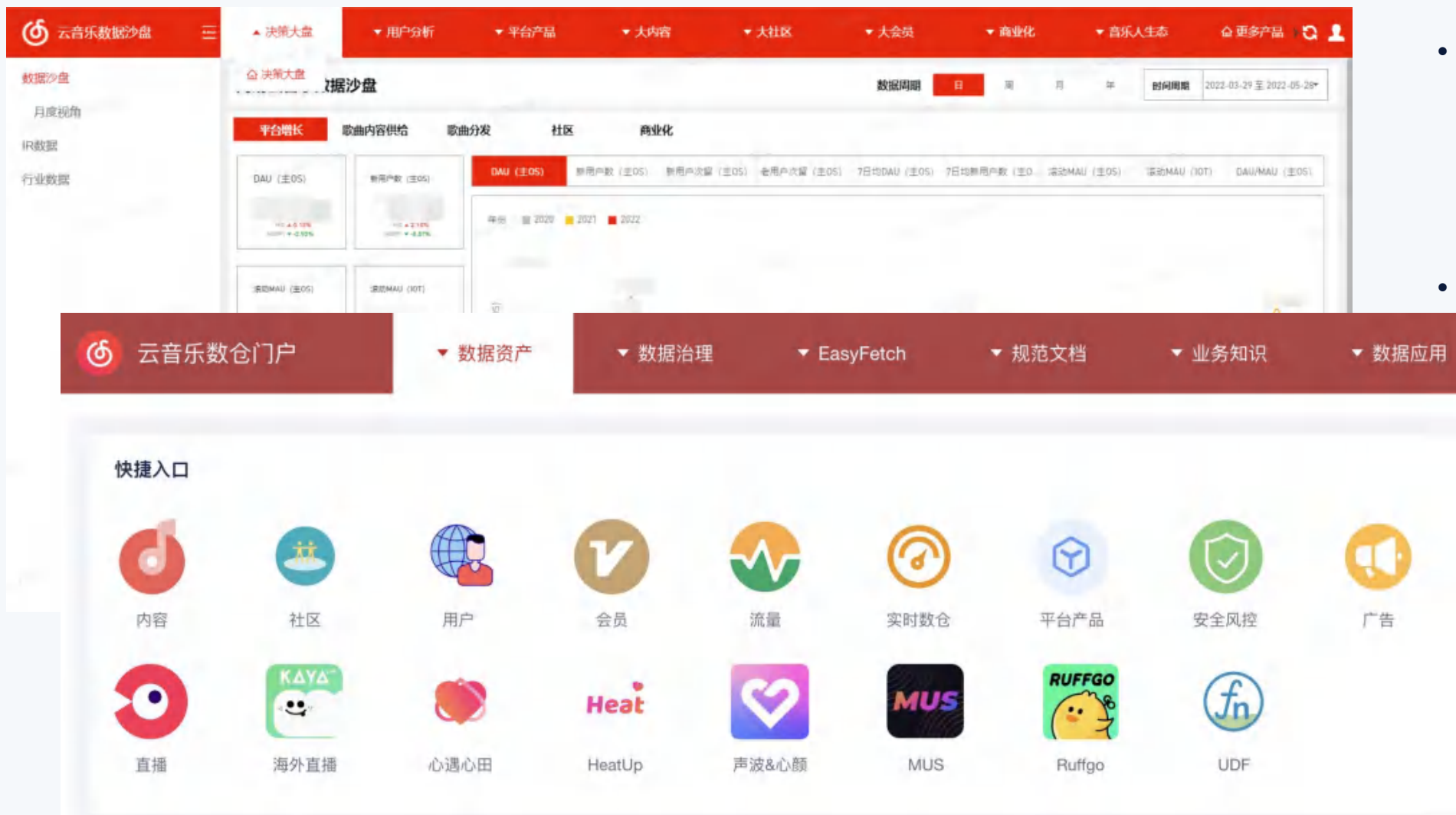
货主地区	货主城市	产品名称	客户名称	单价	订购数量	库存量	两订购量	联系电话
广西				27.4	12670	9990	2316	
广西				27.4	8335	4995	1158	
华南				27.4	12670	9990	2316	
华东				27.4	44345	24995	8198	
	常州			27.4	8335	4995	1158	
		蜜钱汁	顶上系统	18	224	200	15	(056) 6562722
		苹果汁	顶上系统	19	333	30	122	(056) 6562722
	南京			27.4	19005	14985	3474	
	青岛			27.4	12670	9990	2316	
	温州			27.4	6335	4995	1158	
		蜜钱汁	学仁贸易	18	224	200	15	(055) 5355939
		苹果汁	学仁贸易	19	333	30	122	(055) 5355939
华北				27.4	12670	9990	2316	
	秦皇岛			27.4	12670	9990	2316	
		蜜钱汁	利合材料	18	224	200	15	(0117) 1955333

3.零代码自助式-自助取数案例



- 用户完全自助取数，支持拖拉拽和SQL模式
- 支持数据探查、固化成报表，多结果集关联、webhook定时推送，数据找人
- 网易云音乐周uv200+，覆盖70%运营，每周10000次+取数

3.零代码自助式-数据门户案例



- 零代码建站工具，核心用户：分析师、数据产品经理、数仓研发
- 云音乐数据沙盘案例： Dau 150+， 峰值250+， Mau 500+

3.有数BI企业级产品特色



3.有数BI-数据大屏

(搭建平台)

(多端控制)



- 拖拉拽即可搭建可视化大屏
- 支持3D模型可视化
- 支持行政地图、街道地图等丰富的地理信息可视化
- 通过参数/事件功能，支持灵活的多组件、多屏幕、多终端交互能力
- 支持私有部署、离线导出独立运行等灵活的部署使用方式



网易数帆大数据产品矩阵

有数BI

数据门户

可视化报表

自助取数

移动端

数据填报

智能决策

驾驶舱

数据准备

数据大屏

复杂报表

大数据开发治理平台EasyData

数据服务

面向数据分析的数据治理体系

数据标准

数据元

数据字典

数据分类

标准发布

元数据管理

元数据采集

元数据注册

元数据扫描

元数据发布

数据地图

数据目录

数据检索

资产门户

数据血缘

指标系统

原子/派生

版本管理

指标字典

指标关联

数据质量

稽核监控

质量报告

强弱规则

质量工单

数据治理360

价值分析

成本分析

量化ROI

数据下线

模型设计中心

维度建模

量化评估

规范设计

发布审核

数据安全

数据加密脱敏

安全等级

敏感识别

权限申请

基于DataOps
数据开发

数据集成

数据开发

数据测试

任务运维

大数据基础平台NDH

Spark

Hive

Impala

Flink

Yarn/K8s

Arctic (实时数据湖)

HDFS/S3

自动化运维管理