



《智能信息系统建模》

Lecture3 信息架构

04.数据中台是数据管理、分析、服务的平台抓手

王东 张谦

上海交通大学软件学院RFID与物联网研究所

2024年秋季



第三部分：信息架构 提纲

- 概述
- 基于行业数据模板的数据模型设计
- 数据规划方法
- **数据中台是数据管理、分析、服务的平台抓手**





中台的缘起

——2015年底，阿里提出并启动“中台”战略



阿里在2015年面临的问题

- 三套电商体系的架构完全独立，各个应用独立开发和运维
 - 2008年淘宝的技术团队同时支持着淘宝和天猫两大电商平台
 - 1999年成立的B2B电商平台1688一直拥有自己的技术支持团队
- 为什么不共享，要独立建设和维护？
 - 不管是B2B、C2C或者B2C电商平台都需要提供会员服务、商品的信息、交易支付，这些是电商都具有的通用业务功能



各业务都是由以**交易**为核心关联的领域组成

- 交易的对象是**商品**
- 商品通过**店铺**售卖给**会员**
- **交易**的凭证是**订单**
- 在线交易需要**支付**
- 成单后需要**货品出库**和**物流派送**等
- 售前需要**营销**促销活动吸引流量并加强转化
- 售后用户会对**店铺**、**商品**进行**评价**等



阿里的业务中台组成

各业务都是由以**交易**为核心关联的领域组成

- 交易的对象是**商品**
- 商品通过**店铺**售卖给**会员**
- **交易**的凭证是**订单**
- 在线交易需要**支付**
- 成单后需要**货品出库和物流派送**等
- 售前需要**营销**促销活动吸引流量并加强转化
- 售后用户会对店铺、商品进行**评价**等

会员中心

商品中心

交易中心

评价中心

店铺中心

支付中心

营销中心

库存中心



阿里的业务中台组成：（1）会员中心

- 会员中心服务于用户的消费全生命周期，为用户提供特定的权益和服务，企业可以通过会员中心与用户进行互动，培养用户忠诚度。主要能力包括：
 - 会员运营管理：包括会员注册、个人信息维护、会员注销、会员卡办理等相关能力
 - 会员体系管理：包括会员体系的创建、积分规则、成长值规则、等级、权益等相关能力
 - 客户服务管理：包括客户的新增、导入、查询等相关能力
 - 积分交易管理：包括积分获取、核销、清零、冻结、兑换等相关能力



阿里的业务中台组成：（2）商品中心

- 商品中心提供管理商品核心数据的能力，围绕商品构建商品关联数据，诸如商品版本信息、商品品牌、商品属性、商品类目等。其主要能力有：
 - 品牌、类目、属性管理：包括对商品品牌维护、查询，前后端类目的维护，属性及属性组管理等相关能力
 - 产品数据管理：包括对产品模板的创建、编辑、查询、禁用等相关能力
 - 商品数据管理：包括商品创建、修改、查询等相关能力
 - 商品发布管理：包括商品发布、上下架（即时+定时）等相关能力



阿里的业务中台组成：（3）交易中心

- ④ 交易中心负责企业业务交易订单的整体生命周期管理，包括加入购物车→订单生成→合并分拆→流转→支付→发货→退换货→完成。所有电商业务的核心系统都是围绕交易订单进行构建的。其主要能力包括：
- 购物车管理：包括购物车商品添加、编辑、查询、校验等相关能力
 - 正向交易管理：包括交易订单生成、发起支付交易订单、商品发货管理、上门自提及核销等相关能力
 - 逆向交易管理：包括换货、退货、退款等相关能力
 - 订单数据管理：包括交易订单、支付记录、发货记录、换货记录、退款记录等数据管理能力
 - 交易流程编排：支持交易流程节点的配置化，便于根据业务场景的不同设置与之匹配的流程



阿里的业务中台组成：（4）评价中心

- 评价中心提供对评价主体对象、评价规则/等级、评价内容、评价操作的管理能力，从而满足不同角色的评价用户对评价内容的发布、追加、平台审核、平台申诉等需求。主要能力包括：
 - 评价内容管理：包括管理评价的主体对象、评价规则配置、评价等级、评价标签配置等相关能力
 - 评价操作能力：包括评价的发布、修改、追加、回复、申诉等相关能力
 - 评价监管能力：包括评价发布审核、申诉审核、评价屏蔽等监管相关能力



阿里的业务中台组成：（5）店铺中心

- ④ 店铺中心提供企业店铺主体管理、店铺管理、类型管理、经营对象管理等能力以支持企业为商户提供线上门店，同时支持商户管理、店铺会员、店铺会员等级管理、店铺装修等。其主要能力包括：
 - 商户管理：包括商户单个、批量开通，商户审核，商户基本信息维护等相关能力
 - 店铺管理：包括店铺开通、店铺基本信息维护、店铺审核、店铺会员等相关能力



阿里的业务中台组成：（6）支付中心

支付中心给下游商户输出标准的支付服务，提供代付代收、财务对账等服务。通过对接多个主流渠道，稳定输出微信、支付宝、银联等支付能力。其主要能力包括：

- 支付能力：包括创建支付订单、接收渠道通知、查询渠道订单等基本支付能力
- 支付路由：包括支付渠道管理、支付方式管理、支付商户和应用开通管理等相关能力
- 资金账户：包括资金账户管理、充值维护、提现等相关能力



阿里的业务中台组成：（7）营销中心

- 营销中心提供商家的活动计划、申报、审批、执行、核销的全链路管理，也提供基本的促销能力，如优惠券活动、满减买赠等。其主要能力包括：
 - 活动模板管理：包括提供营销活动的策略模板、规则配置、条件、动作模板等相关能力
 - 活动管理：包括提供具体活动的基本信息配置、人群圈选、商品管理、触发条件等相关能力
 - 优惠券管理：包括优惠券的发放、领取、查询、使用核销等相关能力
 - 赠品管理：对于满赠、买赠活动，提供赠品维护、查询、启用、禁用等相关能力



阿里的业务中台组成：（8）库存中心

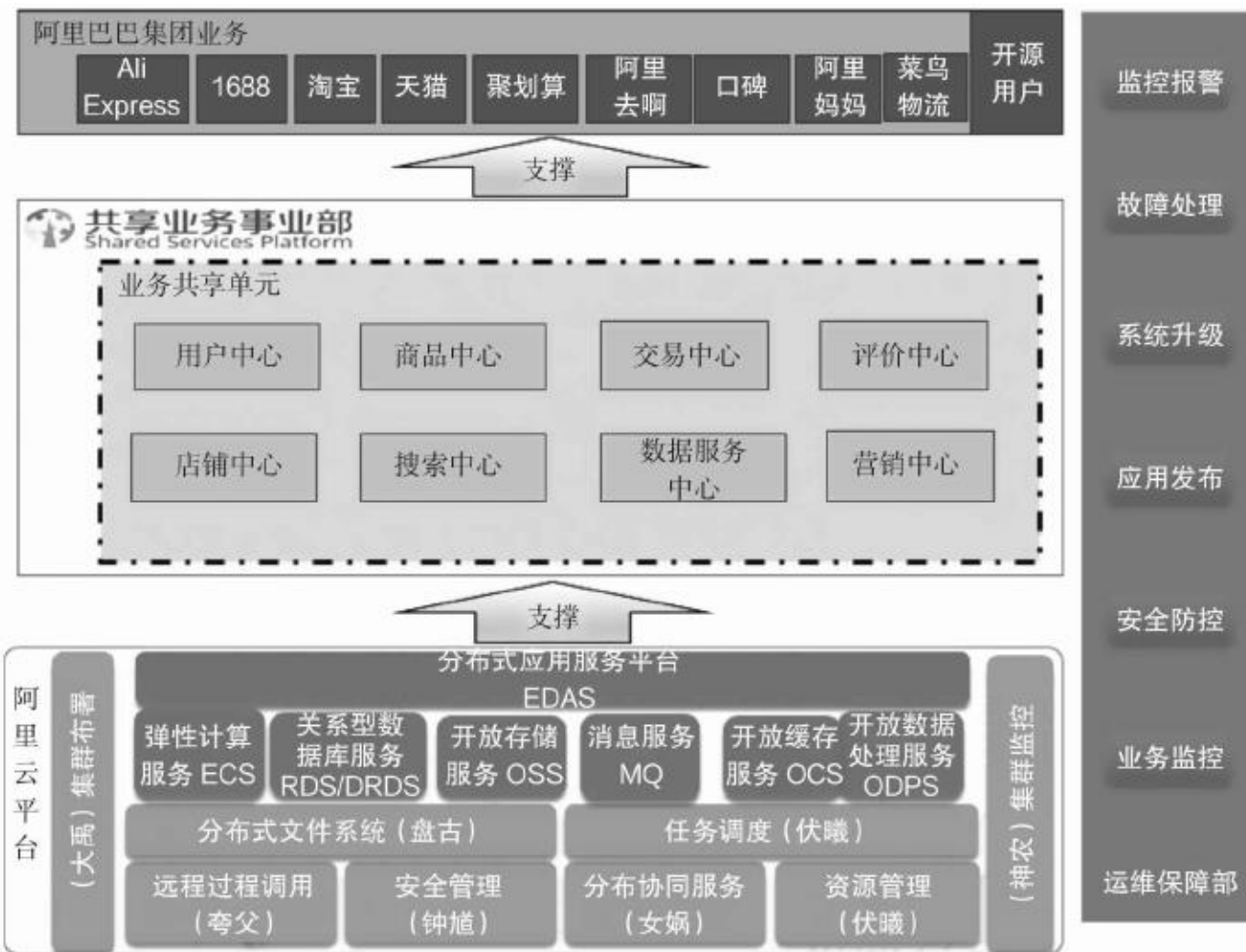
- **库存中心提供仓库、库存、货品、单据（入库单/出库单/盘点单/盘点盈亏单）、审核（调拨/盘点）、包裹、货品运费、物流运输、接入第三方物流公司的服务能力。其主要能力包括：**
 - **仓库管理：**包括服务区、仓库、仓位及其关联管理等相关能力
 - **货品管理：**包括货品进货入库、销售出库、调拨入库、调拨出库、调拨审核等相关能力
 - **货品盘点：**包括盘点单生成、审核、查询等相关能力
 - **履约管理：**包括库存检查、发货单创建及查询、包裹物流查询、运费管理、物流状态跟踪等相关能力



阿里共享业务事业部的发展史

- 阿里将前端业务中公共、通用的业务沉淀到了共享业务事业部，包含了用户中心、商品中心、交易中心、评价等十几个中心，共享业务事业部成了阿里的核心业务平台
- 目前阿里前端超过25个业务单元（如淘宝、天猫、聚划算、去啊等业务）均不是独立构建在阿里云平台之上，在后端阿里云技术平台和前端业务间有了一个“共享业务事业部”，为阿里各种前端业务提供相应服务中心领域内最为专业、稳定的业务服务

“厚平台、薄应用” 架构形态



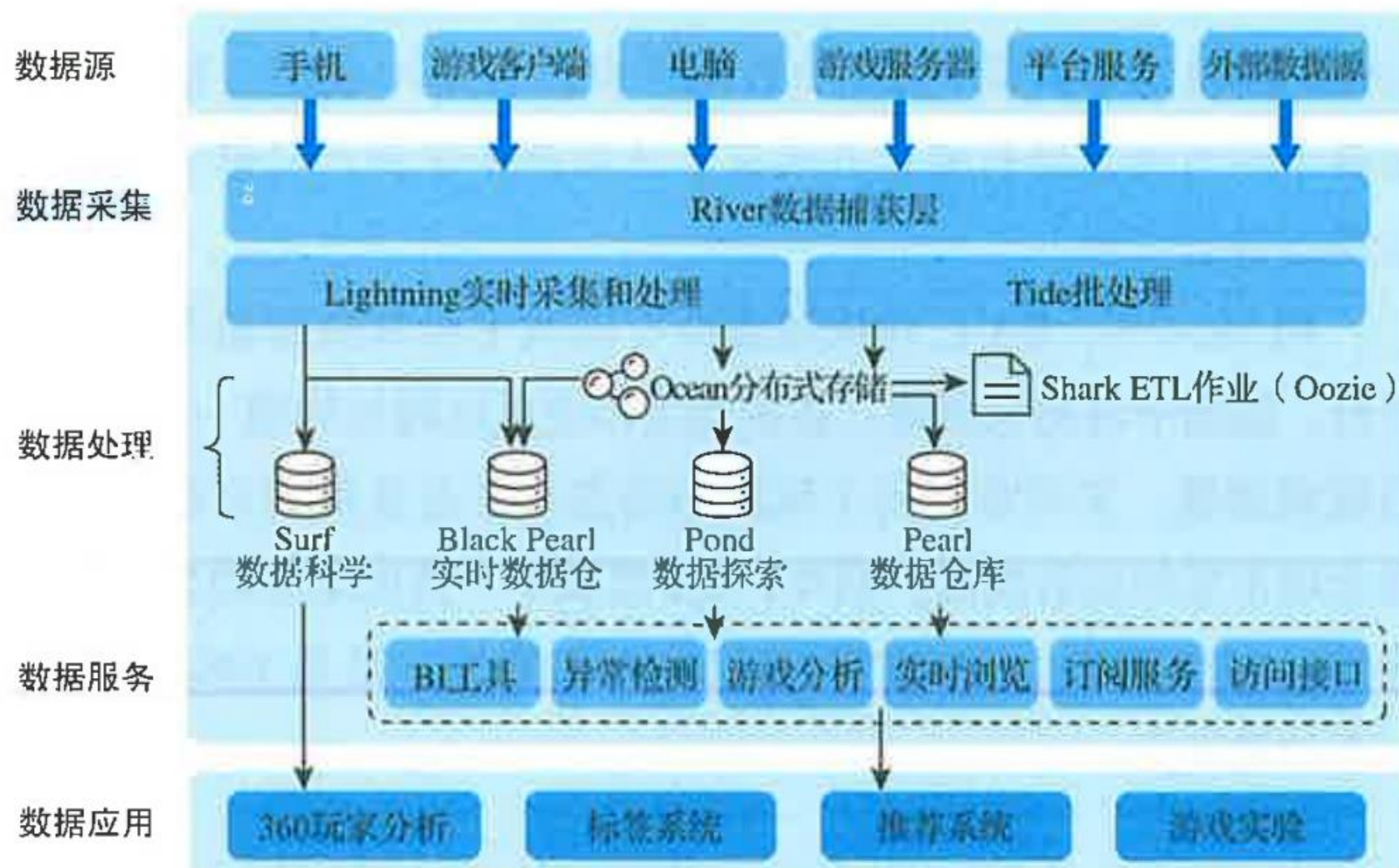


“数据中台”建设项目全景实录

——游戏巨头EA 的实践



“数据中台”是业界进行数据管理、分析、服务的平台抓手





EA 每天产生的数据量非常庞大

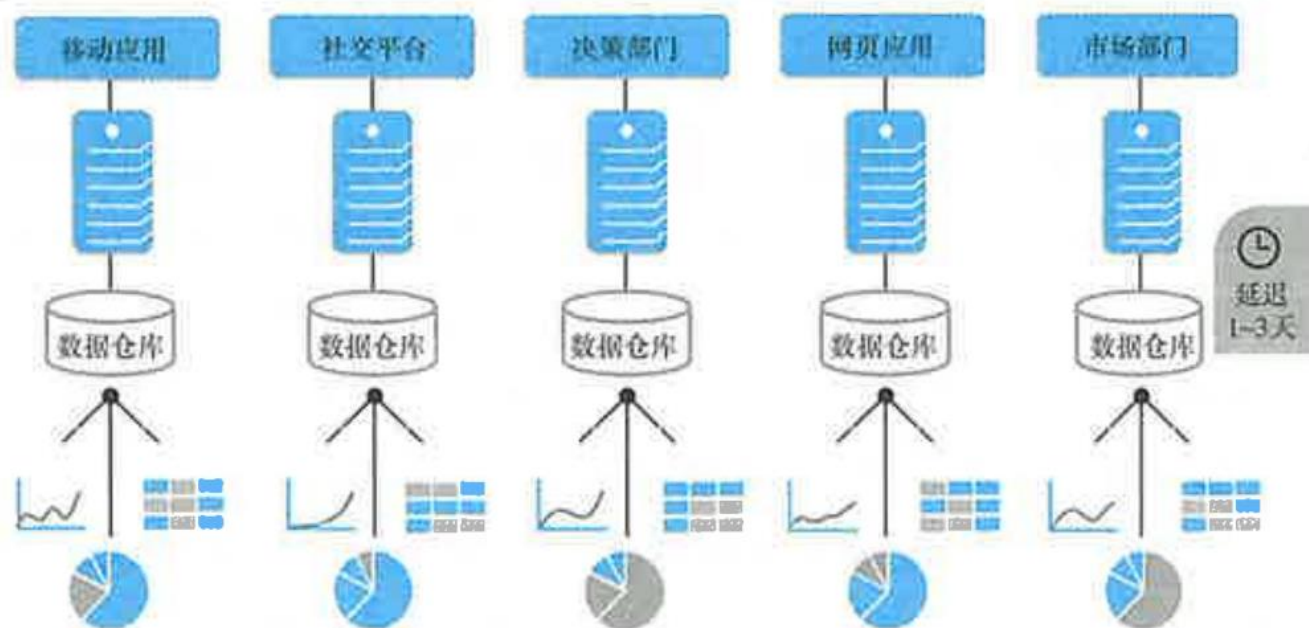
- 在2019财年,《FIFA 足球》在PC和游戏机平台上一共拥有4500万玩家,这些玩家平均每90分钟完成50万场比赛,射门300多万次,进球数达99万个
- 《战地》系列游戏在2017年拥有2100万玩家,平均每天产生1TB 的数据
- 《模拟人生》系列游戏在全球有超过2000万玩家,在20年间一共售出2亿份
- 2019年发布的《Apex 英雄》游戏,在72小时内的玩家数就达到了1000万,而在线玩家峰值突破200万



EA早期的数据分析架构

EA早期的数据分析架构是一种烟囱式的架构，每个游戏工作室(移动、社交、游戏机平台、PC 平台)、每个业务部门几乎都有一套自己的数据分析平台，且架构大同小异。形成这种状况是有其历史原因的

- 首先，EA旗下工作室大部分是收购而来的，它们原来就有一套自己的数据标准和数据分析平台
- 其次，EA一直没有统一的数据分析平台，各业务部门只能自行建设





2013年以来 EA 建设 “数据中台” 的全过程

- ④ **组织架构调整**
- ④ **建设过程**
- ④ **体系架构**
- ④ **数据治理**
- ④ **数据应用产品**
- ④ **“数据中台” 功能总结**



EA建设“数据中台”的第一步

成立EA数字化平台部门 (EA Digital Platform, EADP)

- EADP统一负责数据标准及数据规范的制定和数据中台的建设，打破各游戏工作室和各业务部门各自运营自己数据的孤岛模式
- EA 的大数据部门隶属于 EADP，简称EADP Data，该部门主要由以下几个团队组成

(1) 产品团队(Product Team)

- 该团队负责与EA各游戏工作室和各业务部门进行协调工作，制定和修订 EA 的数据标准和数据规范，同时收集各游戏工作室、各业务部门的大数据开发需求，并与大数据研发团队制定相应的数据应用和数据服务的开发及交付计划
- 产品团队每个季度都会与各游戏工作室、各业务部门进行为期两天的季度产品规划会议，统一梳理各条业务线提交的大数据产品需求，确定各种需求的优先级并落实整个季度可以交付的需求列表。通过季度产品规划会议，EADP 将整个 EA 大数据应用和数据服务的开发和交付纳入一个统一管理的框架

(2) 项目管理团队(Project Management Team)

- 该团队负责与产品组、各游戏工作室和各业务部门进行协调工作，确保数据应用和数据服务的开发和交付都能按时按质完成



EA建设“数据中台”的第一步

成立EA数字化平台部门 (EA Digital Platform, EADP)

(3) 大数据研发团队 (Data Platform Team)

- 该团队负责 EA 大数据平台的集群规划及建设、大数据组件的集成和开发、大数据应用和服务的开发和交付
- 这个团队划分为几个小组
 - 平台组负责大数据集群的规划和建设以及整个大数据集群的 DevOps 工作
 - 架构组负责开发大数据基础组件及服务，比如数据采集系统、ETL 作业调度系统、自助式大数据分析平台等
 - 应用组负责大数据应用及数据服务的开发，比如 360° 玩家画像、标签系统、统一的数据访问接口等
 - 人工智能组负责 EA 人工智能系统的开发，包括反欺诈系统、智能定价系统、各种推荐算法、统一推荐平台等



EA建设“数据中台”的第一步 成立EA数字化平台部门 (EA Digital Platform, EADP)

- EA大数据部门并没有专门的数据分析师团队。EA 所有负责数据分析的成员都分布在各游戏工作室和各业务部门，直接隶属于各个业务线，他们对各个业务线的业务非常熟悉，还掌握了一定的数据分析技能以及使用各种数据分析工具的能力。这些分布在各业务线的数据分析师与大数据团队密切合作，他们既是大数据分析需求的提出者，也是大数据应用和数据服务的使用者，通过这种合作，帮助EA大数据平台逐步成熟、逐步完善，同时在此过程中逐步提升自己的数据分析能力
- 这种组织架构的好处是，EA 大数据团队不需要过多地关注业务问题，因而可以集中精力开发大数据平台和大数据应用及服务，而各个业务线的数据分析师在统一的EA大数据平台进行数据能力的挖掘和开发工作，也方便进行数据能力的共享和复用



2013年以来 EA 建设 “数据中台” 的全过程

- ④ 组织架构调整
- ④ **建设过程**
- ④ 体系架构
- ④ 数据治理
- ④ 数据应用产品
- ④ “数据中台” 功能总结



EA“数据中台”的建设经历了4个重要阶段

1.快速迭代

- ③ 在建设初期，EA大数据部门首先花了近一个季度的时间，在AWS上搭建了30个节点的Hadoop集群作为数据中台的基础
- ③ 逐步将EA的游戏数据从各个游戏平台汇聚到大数据平台，这个平台名为“Ocean”，意思是所有数据都会汇聚到这个数据海洋里。通过AWS公有云和EA内网之间的高速专用网络，Ocean与EA的内部系统连接在了一起。最初，这个平台只提供一些基础服务，比如数据的浏览、查看和下载功能。在这个阶段，业务部门的数据分析能力还非常有限，它们只能通过操作界面将少量简单的数据下载到部门的服务器或笔记本上，然后导入Excel中做一些分析和报表
- ③ 复杂的数据分析可以通过两种方式完成：一种是大数据部门事先开发好MapReduce作业，进行统计分析，然后把结果转换为可下载的数据；另一种是业务部门的分析人员自己通过客户端连接到Hadoop集群(Hive)运行SQL语句进行自助式查询。第一种方式分析的数据是有限的，难以满足业务部门日益增长的数据分析需求；而第二种方式只有少数分析人员能够使用，而且自助式作业的运行时间也不能过长，否则客户端会失去与服务器的连接

虽然这个阶段的成果是有限的，但各业务部门快速见到了一些成效，对数据中台的建设也慢慢有了信心



EA“数据中台”的建设经历了4个重要阶段

2.工具开发

- ④ 数据中台的一个很重要的衡量标准是，业务部门要能够自助进行数据探索，要能够利用数据中台的工具快速发现业务的洞见，然后快速应对市场变化
- ④ 这个阶段的主要建设工作是开发一个自助式数据分析平台。EA 大数据部门建设了一个与生产集群并列的Hadoop 集群，取名叫“Pond”，因为它的规模要比生产集群Ocean小得多。Pond 集群每小时都从生产集群中抽取最近一小时的数据，保持集群中始终有最新的“热”数据，同时Pond还可以访问AWS S3所保存的所有历史数据，这样 Pond就具有了全量数据的访问能力，它本身的主要资源都用来做数据分析
- ④ EA 大数据部门还开发了一个交互式界面，让业务部门能够在Pond上自助运行Hive查询语句，并提交数据分析作业。数据分析作业完成后会以电子邮件的形式通知作业创建者。这个平台上线后，每天有两三百名EA 各部门的数据分析人员在上面运行数据分析作业，平均每天运行的查询作业有四五百个
- ④ 慢慢地，一些部门日常的数据分析工作也都转移到了Pond 上。这时，针对多租户的使用场景，大数据团队又研发了一些高级功能，比如作业优先级的管理、各部门资源使用总量限制以及数据分析的审计功能
- ④ 对于业务部门而言，这个自助式数据分析平台极大解放了它们的生产力，充分调动了它们使用数据中台的积极性。而通过 Pond的审计功能，管理层清楚看到了各部门数据驱动业务的发展程度以及各种数据资产的使用率



EA“数据中台”的建设经历了4个重要阶段

3.能力复用

- ④ 自助式的数据分析平台上线以后，EA 大数据部门肩上的压力瞬间减轻不少，**下一阶段的主要精力就放在一些复杂数据能力的开发上**
- ④ **首先要解决的是EA的主要营收贡献者《FIFA 足球》团队的两大痛点**
 - 第一，游戏欺诈问题。一些玩家在游戏中利用大量的僵尸账号来收集游戏币，然后在游戏外的黑市上售卖而获利。这个问题每个月给EA 造成的损失达数百万美元
 - 第二，市场推广难。EA 市场推广团队每天需要分析上千万玩家的数据，进行上百次操作，根据玩家的静态和动态数据，每次锁定几百个特定玩家，进行有针对性的市场推广活动，但自助式分析平台因为计算资源的限制，每天只能完成几次这样的操作



EA“数据中台”的建设经历了4个重要阶段



针对游戏欺诈的痛点

—EA 大数据团队开发了一个**基于图数据库的快速回溯检索工具**。《FIFA 足球》的数据分析人员将从黑市上买来的游戏币编号输入系统后，该检索工具就会快速将经手过这些游戏币的账号全部检索出来，并根据账号活跃度筛选出真正的僵尸账号，然后游戏运营团队快速将其查封，从而阻止游戏欺诈行为



针对市场推广的痛点

—EA 大数据团队开发了一个**基于 CouchBase和 Elasticsearch的实时标签系统**，将玩家相关的数据都存储在CouchBase中，然后将玩家的静态和动态数据都定义成玩家的标签，并在CouchBase中进行实时更新和计算，将标签计算的结果都存入Elasticsearch中，供《FIFA足球》的市场推广部门筛选玩家。这个系统上线后，市场推广人员只需简单拉取几个标签，就可以在几秒内完成一次玩家的筛选工作

在反欺诈系统和标签系统的功能逐步完善并得到足够多的反馈之后，EA 大数据团队将这些数据工具进一步抽象，升级成通用的系统进行共享和复用，将它们推广给其他游戏工作室。这里体现了数据中台的一个重要方法论：各业务部门不应该重复造轮子，数据应用和数据能力需要共享和复用



EA“数据中台”的建设经历了4个重要阶段

4.形成闭环

- 在数据分析平台初具规模、数据应用日益丰富以后，EA 大数据团队将注意力转向了EA 管理层制定的玩家优先战略的核心问题
 - 如何利用大数据、AI、机器学习技术提升玩家的游戏体验
- 从EA “数据中台”的建设初期，EA 大数据团队就开始了数据科学家的招募工作。由于团队人员逐步到位，数据中台中也积累了足够多的历史数据，EA 大数据团队在这一阶段的主要精力是建设EA的游戏推荐系统
 - 通过对大量历史数据的分析，利用人工智能和机器学习算法，为每个玩家提供个性化的游戏难度推荐，或者进行游戏玩家动态组队
 - 在推荐系统上线后，EA 玩家游戏参与度(玩家游戏时长、玩家留存率、玩家推荐率)提高了 10%
- 经过这个阶段，EA 的数据中台已经基本成型，数据驱动各个环节都已经打通，并形成了从数据采集到产品推荐和用户反馈的数据闭环

到2014年，EA 已建成覆盖所有游戏平台 and 所有游戏的数据中台



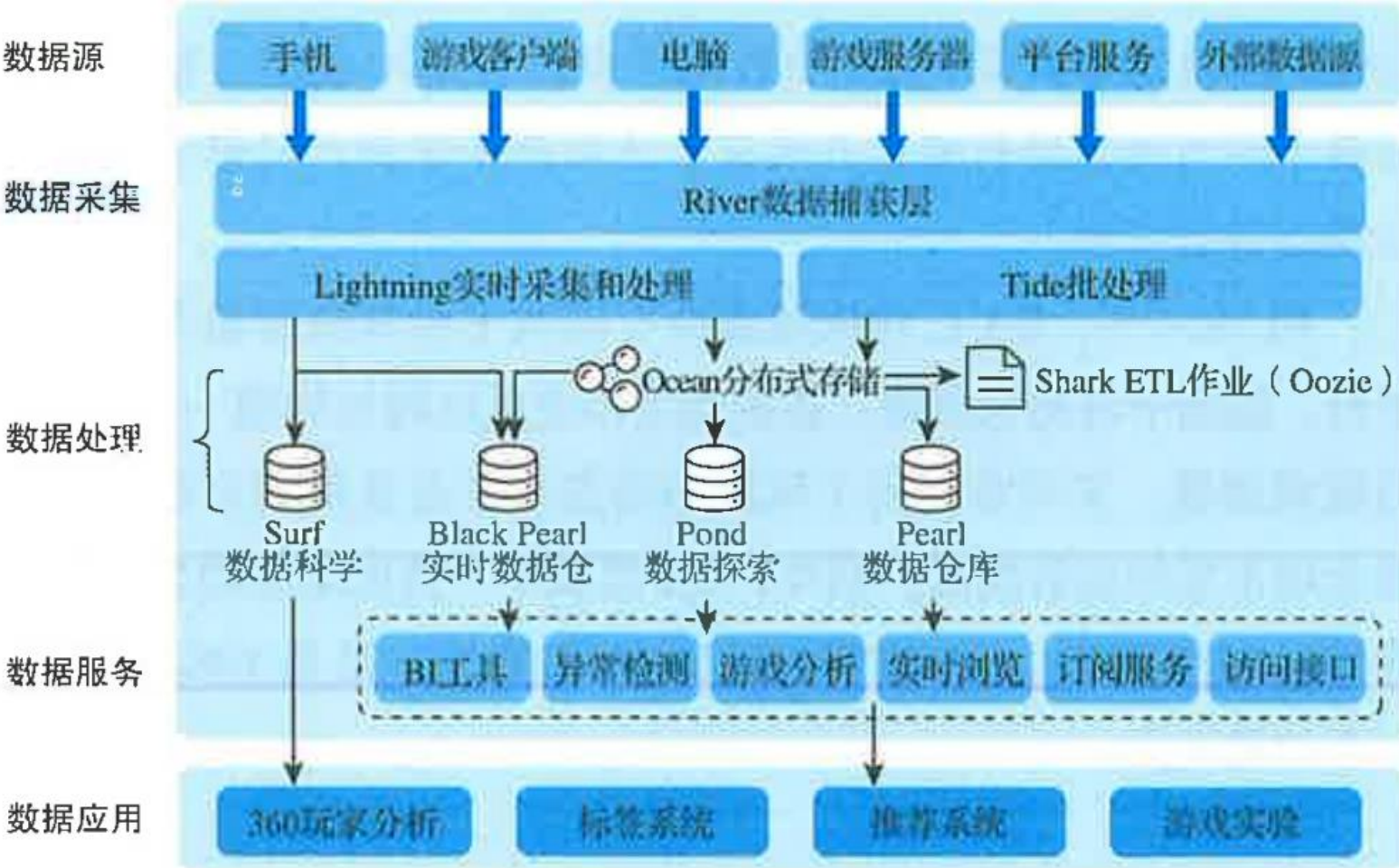
2013年以来 EA 建设 “数据中台” 的全过程

- 组织架构调整
- 建设过程
- 体系架构**
- 数据治理
- 数据应用产品
- “数据中台” 功能总结



EA “数据中台”的架构

架构从上至下分为4个层次





2013年以来 EA 建设 “数据中台” 的全过程

- 组织架构调整
- 建设过程
- 体系架构
- 数据治理**
- 数据应用产品
- “数据中台” 功能总结



1. 数据标准和规范

- 在综合考虑了数据分析的所有需求后，EA 大数据部门制定了第一版**Game Telemetry标准**和**游戏分析指标体系Taxonomy**
 - Game Telemetry 发送的每一个数据称为事件 (Event)，它以JSON 的格式发送和接收，主要内容如下
 - 通用属性**：通用属性是所有平台、所有游戏都需要上报的数据。通过通用属性的定义，可以轻松的分析EA跨平台的所有游戏的日活跃玩家数、玩家平均游戏时长、玩家游戏内平均消费金额等共性指标。通用属性包括玩家ID、平台型号、游戏名称、事件编号、事件发生时间、SDK 版本等
 - 特殊属性**：特殊属性是指每个游戏独有的属性，比如移动游戏中iOS 设备的UDID, 射击类游戏武器编号、《FIFA 足球》中玩家所选球员号码等。游戏工作室还可以将A/B 测试所需要的属性放在特殊属性里
- 随着EA 业务需求的变化，数据标准和规范也会逐步升级。每次升级，Game Telemetry都会有版本号，指标体系Taxonomy 也有与之相对应的版本号
- 每次Game Telemetry版本升级，游戏客户端及服务器端的Game Telemetry SDK软件会由EA 大数据部门统一开发并发布给各游戏工作室。各游戏工作室在制定游戏开发和发布计划的时候，必须与EA 大数据部门确定好 Game Telemetry SDK的版本及集成计划相符
 - 通过数据标准和规范的制定，以及Game Telemetry SDK的发布和实施流程的确定，EA 迈出了进行有效的数据资产管理的第一步



2. 元数据管理

- EA 每个版本的Game Telemetry 和游戏分析指标体系Taxonomy 都会在EA 统一的文档系统中发布，供各游戏工作室和业务部门查阅。EA 大数据部门在EA 的自助式查询平台开发了**数据目录管理**，公司内部的数据分析人员可以在这个平台上浏览和查询自己被授权访问的数据库及数据表的元数据
- 每个数据表都标注了来自哪个数据源、有哪些字段、有什么样本数据，如果是维度表(比如武器编码和名称、车辆编码和名称)，则标注了对应 Game Telemetry的版本
- 数据分析人员还可以查看主数据关键字段在不同时段的数值分布，这是由数据质量监控系统预先计算出来的。数据分析人员在自助式查询平台还可以对数据库和数据表进行进一步的标注，使元数据更加完善
- 元数据的一个重要功能是**数据血缘分析**
 - EA 的大数据平台采用开源软件Oozie 进行ETL 作业的调度，每个作业的配置文件及代码本身其实都描述了该作业的上下游数据，通过整合所有作业的元数据信息，可以构建全局的数据和ETL 作业的血缘关系。但是Oozie本身并不提供这个功能，为此EA 大数据团队基于Oozie开发了数据血缘分析功能，通过对 Oozie 保存的元数据的采集和处理，实时捕获系统中数据的依赖关系



3.数据质量管理

- ④ **EA的数据质量管理是通过EA 大数据团队自主研发的数据异常检测平台来实现的**
- ④ EA 大数据团队负责运维整个 EA 跨多个游戏平台的上百个游戏的数据，需要监控的游戏指标有上万个。为了适应如此大规模的数据质量管理需求，EA 大数据团队打造了一个自动化的数据异常检测平台Oasis
- ④ 通过接口调用注册一个新监控指标，后台监控引擎会对该监控指标进行一段时间的“适应”运行，自动构建合理的监控数据模型，并自动选择合适的算法进行后续的持续监控工作
- ④ 总体来说，**Oasis 监控系统主要检测以下几类数据的异常**
 - 数据表某些关键字段的数值分布，比如每个小时的Distinct Count。这对元数据计算及监控是同步完成的
 - 一些关键游戏指标的时序异常，比如某个游戏每分钟客户端请求数，该指标在某个时点出现显著下降
 - 一些组合性指标的异常检测，比如需要监控某一数据表中的若干个字段，异常检测算法会选择一个机器学习算法，计算出将这些字段的值组合在一起后的最大异常分值，该异常分值出现波动时会相应地触发监控报警



2013年以来 EA 建设 “数据中台” 的全过程

- 组织架构调整
- 建设过程
- 体系架构
- 数据治理
- 数据应用产品**
- “数据中台” 功能总结



依托“数据中台”，EA打造了一批数据应用产品供共享使用

- 为了更好地贯彻玩家优先、数字化和One EA 的战略目标，依托“数据中台”，EA着重打造了一批数据应用产品，供EA 各团队共享使用，其中比较重要的是推荐系统、动态游戏体验和标签系统

1. 推荐系统

- EA大数据团队构建了一个统一的用户推荐系统以支持所有类型的游戏及EA 的主页，这个推荐系统支持三种应用：推荐购买游戏、游戏地图推荐和游戏难度推荐
 - 推荐购买游戏**就像推荐电影或图书一样，向玩家推荐下一个可能令其心仪的游戏。虽然同一类型游戏(如射击、体育、角色扮演类)互相推荐的效果并不好，但推荐系统可以向玩家推荐不同类型的内容，比如游戏、额外下载的游戏包、游戏实况视频和游戏教学视频等
 - 游戏地图推荐**是在游戏内向玩家推荐游戏的模式和地图。很多在线游戏，特别是第一人称视角射击游戏和体育竞赛游戏，提供不同的游戏模式及地图以支持多样化的游戏体验。向新玩家推荐合适的游戏模式和地图是很重要的，这种推荐可以提供比较好的游戏体验(而非挫败感很强的体验)给这些新玩家，从而提高游戏玩家的留存率
 - 游戏难度推荐**是向玩家推荐游戏内关于难度的配置选项。玩家的经历、游戏技巧、学习速度及游戏风格千差万别，因此他们对于同一种难度设置的反应是不一样的。就算是同一个玩家，他对游戏难度的喜好也会随着时间的推移而发生改变。游戏难度选项的推荐会根据游戏玩家之前的游戏经历来给出建议或者调整，从而最大限度提高玩家的游戏停留时间

EA 的这个推荐系统在上线初期，使EA 游戏购买网页的CTR 提高了80%,平均游戏 停留时间提高了10%



依托“数据中台”，EA打造了一批数据应用产品供共享使用

2.打造动态游戏体验

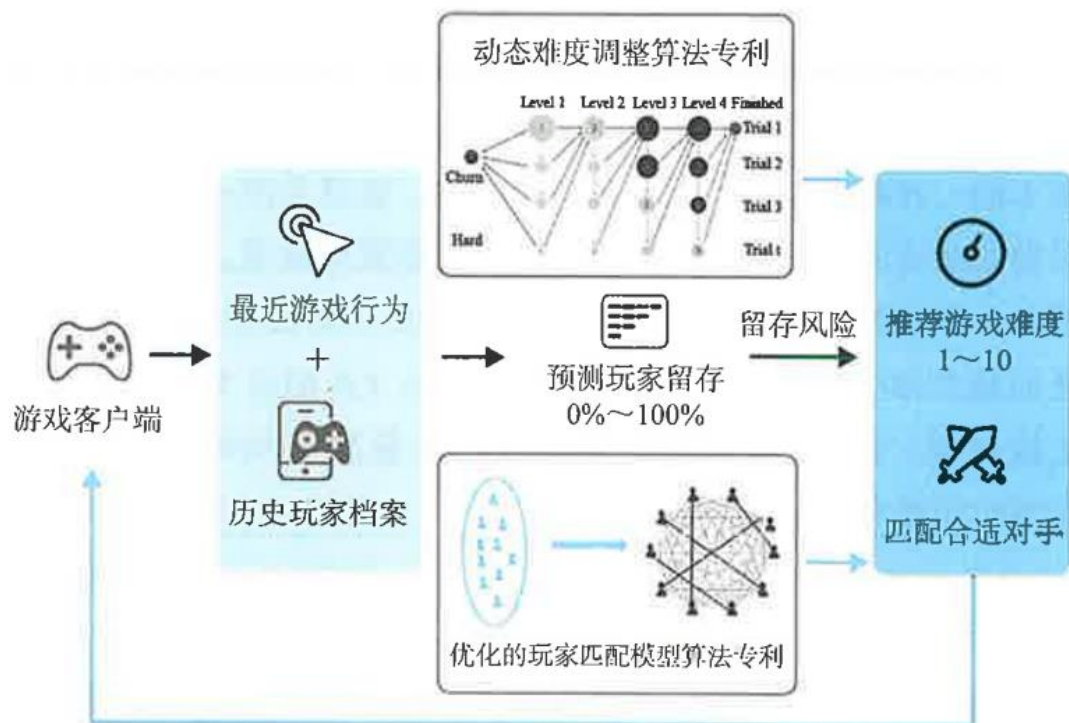
● 利用AI 技术提供动态游戏体验

- EA 的竞争对手动视暴雪(Activision Blizzard)在2017年申请了一项专利，**在游戏中提供动态体验来增加游戏内购买的次数**，但动视宣称，该专利还没有在游戏中实际应用
- EA 不仅在2016年和2017年相继申请了“**动态难度调整算法**”和“**优化的玩家匹配模型算法**”这两项专利，而且还把这些专利技术运用到了游戏中

● EA动态游戏体验的实现过程

传统动态游戏体验算法一般只考虑局部游戏数据(玩家玩前几关的数据)，追求的是一种短期效应，也就是通过算法来判断应该给玩家提供什么样的难度，使玩家在玩下一关时有个好体验

EA 的算法模型不仅考虑了玩家最近24小时的游戏行为，还结合了玩家的历史数据来建模。在决策效果方面，EA 的算法着眼的是全局优化，也就是提高玩家在整个游戏过程中的留存率，而不只是下一关的体验





依托“数据中台”，EA打造了一批数据应用产品供共享使用

- ④ 对于**动态难度调整**，简单地说，EA 的算法就是对玩家未来游戏全路径的一次探索，对每一条可能的路径计算出玩家的留存概率，则当前的难度选择就是对全局留存率最优的选择
- ④ 对于**动态玩家匹配**，EA 的算法是通过对所有玩家历史数据的分析建立一张图，图中每一个节点代表一个玩家，每一条边就代表边两端的玩家可以进行匹配。这张图还计算出了针对不同的匹配，每个玩家相应的游戏留存率，于是动态玩家匹配就是在这张图中寻找对两个玩家的留存率来说都最优的选择

这一动态游戏体验框架在EA 的多个游戏中得到了应用，游戏玩家停留时间最多得到了9%的提升，并且没有影响到游戏内的消费行为



依托“数据中台”，EA打造了一批数据应用产品供共享使用

3.标签系统及游戏运营

- 游戏运营是指对玩家的数据进行分析，找出特定的玩家群体，对他们进行营销，以提升他们的游戏体验，提高他们的游戏留存率。在这个过程中，给用户个体打上一些标签，以便有针对性地为其提供服务。标签系统是实现这一目标的常用工具
- 在 EA“数据中台”建设初期，《FIFA 足球》的游戏运营部门做一个营销活动要花费很多时间，通常要一两天
 - 运营人员首先要在自助式的大数据分析平台重运行复杂的 SQL 语句，从上千万玩家中筛选出几万个玩家作为目标群体。因为这样的计算往往涉及玩家的历史消费数据及历史行为数据，所以它的计算量很大，但在一个多租户的平台上，为了保证资源的公平共享，分配给这个计算作业的资源往往是有限的，整个计算过程要花费几个小时
 - 在计算完成后，运营人员将计算结果数据下载下来，然后根据筛选出来的玩家，通过邮箱进行营销推送
 - 这样的流程显然是不能满足运营部门的需求的，因为他们需要在同一时间段进行多个营销活动，并进行比较，以确定在这个时间段内最合适的营销方式



2013年以来 EA 建设 “数据中台” 的全过程

- ④ 组织架构调整
- ④ 建设过程
- ④ 体系架构
- ④ 数据治理
- ④ 数据应用产品
- ④ “数据中台” 功能总结



EA 的“数据中台”在数字化运营中实现的提升

- ④ 第一，实现了全部游戏数据和业务线的覆盖，提高了系统对整体业务的还原度，所有游戏工作室和所有业务部门的数据都在一个平台上管理。这样就提高了效率，避免了每个工作室重复造轮子，也允许以前很难实现的跨平台 (Cross-Platform)、跨产品(Cross- Product) 推荐 (Promotion) 成为可能，真正实现了开源节流
- ④ 第二，提供了不同层次的自助式工具，方便各业务线的数据分析人员进行数据能力的开发和共享，公司整体数据能力得到极大提高， 基于数据的思考方式被真正落实到业务人员
 - 通过在自助式分析平台上运行Hive程序进行自助查询
 - 通过统一的数据服务接口抽取数据
 - 通过数据仓库支持的BI 工具制作可视化报表
 - 通过标签系统快速锁定玩家群体.....



EA 的“数据中台”在数字化运营中实现的提升

- ③ 第三，整个数据中台有一个量化指标体系来衡量平台上数据和人员的活跃情况，比如自助式分析平台上每个部门计算资源的使用量、每个Hive 表被查询的次数、各部门调用数据服务接口的次数、访问数据仓库的次数等。这个量化指标体系可以让EA 大数据部门进行适当的调整
 - 一是保证核心业务部门(比如EA 营收大户《FIFA足球》背后的团队)能获得相对多的平台资源， 以保证他们的业务能创造更多的价值
 - 二是确保整个数据中台的投入花在值得投入的地方并产生可量化的价值
- ④ 第四，整个EA 数据中台的建设中，各个子系统基本都采取了统一的顶层设计、从特殊到一般的构建过程。也就是说，先根据某个业务部门的需求进行设计和开发，设计过程中会充分考虑以后的扩展要求，当该子系统运行稳定以后，再进行升级改造， 提供给其他业务部门使用。这种“顶层设计，单点突破，快速落地，全面开花”的策略使得数据中台的快速见效成为可能



不要做“过早优化”(Premature Optimization)

- 国内类似的概念“技术中台”就源于中间件、PaaS 平台
- 但是这种中间件、平台、中台的功能一般并非由一个顶层设计得出，而是一步步建立起来的
- 硅谷企业有一个重要理念是不要做“过早优化”(Premature Optimization)，也就是说，**不要在不需要的时候进行优化，一定要先完成功能再优化**。因此不需要中台的时候没有必要刻意建一个大而全的中台
- 当然，在建设数据中台的不同阶段可以使用不同的技术，只要保证中台建设能够平滑过渡即可



上海交通大学RFID与物联网研究所

上海东川路800号交通大学软件大楼5313室

王东

18621365573

wangdong@sjtu.edu.cn

张谦

13816047437

qwert3472@sjtu.edu.cn