

企业数字化转型新引擎 解码数据中台最佳实践

——爱分析·中国数据智能应用趋势报告

2020年10月





企业数字化转型新引擎 解码数据中台最佳实践

——爱分析·中国数据智能应用趋势报告

2020 年 10 月

报告编委

报告指导人

黄 勇 爱分析 首席分析师

报告执笔人

王小娜 爱分析 分析师

徐碧云 爱分析 助理分析师

外部专家（按姓氏拼音排序）

陈吉平 袋鼠云 创始人&董事长

崔晓波 TalkingData CEO

方 毅 每日互动 创始人&CEO

贺鸿富 爱数 总裁

李振宏 餐道 创始人&CEO

林永峰 山西证券 金融科技部总经理&山证金科副总经理

刘 晨 御数坊 创始人&CEO

刘译璟 百分点 CTO

孙元浩 星环科技 创始人&CEO

王 华 互道 创始人&CEO

汪祥斌 DataEye CEO

王新义 逸迅科技 研发总监

杨 磊 滴普科技 销售服务体系总裁

摘要

数据中台是支撑数据智能落地的基础设施

随着企业业务发展与技术成熟，企业面临海量增长的数据，结合大数据、人工智能物联网等技术，企业对数据的应用从监测、洞察逐步向决策迈进，进入了数据智能应用的时代。

传统企业面临多种数据应用的困境：内部数据孤岛现象严重，且数据碎片化，无法统一融合赋能业务；企业内部通常是 IT 部门负责处理和管理数据，业务部门无法直接使用数据，影响业务运营与规划；数据开发工作成本高，不同需求导致企业存在数据服务能力重复建设的问题等。

而数据中台能够实现数据汇聚、统一标准与口径，形成数据资产，进而为业务提供高效的数据服务；同时数据中台面向业务场景而建，将数据抽象封装成服务，能够实现数据的资产化、服务化，具有跨部门的普适性业务价值能力，赋能业务部门人员进行数据分析和数据应用；连接数据前台和后台，实现企业的数据开发能力的复用。因此，数据中台成为企业实现数据智能应用的核心。

以业务场景应用为指引规划数据中台建设

企业建设和运营数据中台需要从业务战略目标出发，以业务场景应用为指引贯穿数据中台建设与运营全流程，包括顶层战略规划、应用场景规划、数据治理体系搭建、技术平台建设、业务价值实现、持续运营等。

第一，企业数字化转型过程中，数据中台建设的核心目的是服务于企业的整体战略目标与业务目标，企业建设数据中台建设要从顶层战略规划出发，根据业务目标，从战略层面规划数据中台建设路径、调整组织架构、调配相关资源。

第二，数据中台的价值最终需要通过在业务场景的数据应用来体现，不同行业和企业处在不同阶段所需要的数据中台是不同的，因此，应用场景需要规划先行，企业需要明确应用场景实现的优先级，率先实现部分场景应用，打造业务场景的闭环，实现业务收益。

特别鸣谢 (按拼音排序)

DataEye

TalkingData

AI SHU 爱数®
— For a smarter future —

Percent 百分点

YEEXUN

餐道

袋鼠云
DTSTACK

滴普科技
DEEPEXI

互道
驱动商业未来

个推
GETUI

明略科技™
MININGLAMP TECHNOLOGY

TRANSWARP
星 环 科 技

御数坊
DGWorkshop

目录

1. 数据智能助力企业数字化转型	8
2. 数据中台支撑数据智能落地	32
3. 数据智能应用趋势展望	62
结语	73
关于爱分析	74
研究咨询服务	74

CHAPTER

01

数据智能

助力企业数字化转型

1. 数据智能助力企业数字化转型

1.1 数据驱动是数字化企业的核心特征

数字化转型，是指企业利用数字技术，通过商业模式、业务流程、组织架构等重塑，为客户创造新的价值。近年来，数字化转型已经成为各行业的共识，企业在数字化转型上已经投入了大量资源，部分行业领先企业已经成功地通过数字化实现良好的客户体验，并实现敏捷运营。

与信息化企业不同，数字化企业的核心特征是数据驱动。在信息化阶段，企业主要是进行业务系统等 IT 基础设施的建设，实现经营流程的标准化和线上化，提升业务效率。但在这个阶段，数据主要是流程的副产物，并没有真正发挥价值。

在数字化阶段，企业的核心特征在于数据驱动业务，即能够在信息化的基础上产生和汇集更多高价值数据，在数字世界监测和洞察业务真实状况，并构建数据分析和决策模型以驱动业务运行。在数字化阶段，各业务部门都有数据分析和数据决策的需求，数据成为驱动业务的生产资料。

近年来，互联网、移动互联网、物联网等技术的应用产生了海量数据，使得企业进入数据爆发阶段，而人工智能等技术的发展也降低了数据应用的门槛，驱动企业数字化转型加速。

从应用深度来看，企业对于数据的应用是一个逐步深化的历程，整体发展经历了收集、监测、洞察、决策四个阶段，数据的应用价值不断提升，逐步从业务数据化走向业务智能化。

图 1：企业数据应用的发展历程



收集、监测和洞察阶段是业务数据化过程，即通过数据描述跟踪业务的发展，主要的技术应用分别是大数据平台及数据服务、BI 与可视化和数据行业应用。决策和重塑阶段是业务智能化过程，数据、算法可以赋能、改变业务流程，提升业务效率，企业的数据应用走向数据智能与人机协同。当前，数据应用已经处于业务决策阶段，正在向业务重塑阶段迈进。

1.2 企业迈入数据智能时代

当前，企业的数据应用已经进入到业务决策阶段，机器能够基于数据和算法自主进行决策或给出决策建议，带来更高效的业务反馈和更大的业务价值。

例如，在互联网领域，滴滴、美团的调度系统可以根据前端订单等业务数据，自动完成决策，给出最佳任务调度方案和路线，并将任务指令下发给司机和骑手，整个决策过程极为高效精准，但几乎没有业务人员参与。

随着数字化转型的推进，这种数据驱动业务决策的场景也越来越多地渗透到传统行业。例如，某线下商超利用大数据和智能算法进行销量预测，进而实现智能补货，大幅降低门店缺货率，并提升库存周转率。AI 模型替代采购专家经验，基于数据和算法给出销量预测结果，使得预测准确率大幅度提升，并提升了采购效率。

1.2.1 数据智能的定义

爱分析认为，数据智能是指基于中台体系，融合大数据、人工智能、云计算、物联网等多种技术，利用数据实现智能决策。

图 2：数据智能的定义



图：爱分析绘制



为了实现数据驱动决策，企业需要具备统一管理数据、快速配置开发业务的能力，从企业视角形成数据资产，进行数据能力的输出，充分释放数据价值，数据中台逐步在企业内部形成。

数据中台使数据具有数据服务化的能力，同时，还需要对数字资产进行运营优化，使数据资产赋能企业能力快速迭代升级。这需要与业务密切相关的技术平台的支持。

技术中台包括数据采集与预处理、数据存储、数据分析、数据可视化与 BI 等环节的所有工具和平台，包括大数据基础平台、NLP 与知识图谱、数据科学平台、BI 与可视化系统等。

但数据中台输出的是同质化的数据服务，仅有数据中台和技术中台无法赋能不同业务场景，需要与业务系统结合，才能真正地将数据用起来。因此，数据智能以数据中台和技术中台为底座，结合业务中台，支撑企业上层智能决策和应用。

中台汇聚企业的业务数据，通过自动化、智能化的数据采集与汇聚，将实时与离线数据打通关联，深度挖掘数据价值，并开放数据服务至各业务场景中，具备汇聚整合、数据提纯加工、数据服务可视化、数据价值变现等核心能力。

在数据智能阶段，企业可以基于数据中台和业务中台构建业务智能化闭环。第一步是业务数据化，即采集和沉淀业务运行过程中的数据；其次是数据资产化，即通过数据中台的统一汇集和治理，将业务数据转化为可用的数据资产；资产场景化，是基于数据中台将数据资产加工和输出为可应用于业务场景的数据服务，通过业务中台实现数据资产在场景中应用，最终实现业务场景的智能化。而业务场景实现智能化之后，也能不断的提供新的数据，优化整个系统，形成业务智能化的闭环。

图 3：基于数据中台和业务中台构建业务智能化闭环



1.2.2 数据智能的技术应用体系

为了实现数据驱动决策，需要汇聚并处理大量结构化、半结构化和非结构化数据，基于复杂网络进行推理和决策，单纯依靠简单的数理统计已经不足以满足需求。因此，数据智能是大数据、人工智能、云计算、物联网等多种技术融合的技术应用体系，贯穿数据采集与预处理、数据存储与计算、数据分析与挖掘、数据展现与交互的全生命周期。

图 4：数据智能的技术应用体系



图：爱分析绘制



（1）数据采集与预处理

随着互联网、移动互联网、物联网等技术的发展，企业面临海量数据，且数据多存在于彼此独立的系统与平台中，不同系统的数据采集方式是不同的。

数据采集具有一定的数据生产属性，线上数据通常来自数据库、PC 网页、H5、App 应用、应用及服务、智能可穿戴设备等资产层，采集方式主要有客户端 SDK 埋点、服务端埋点和网络爬虫；线下数据主要是通过硬件设备采集，例如传感器、摄像头、Wi-Fi 信号采集等。

数据采集之后，需要经过中间系统的流转写入目标存储中，进行数据汇聚。多种数据源类型可基于离线数据同步和实时数据同步两种策略进行数据交换。

但企业数据大多是脏数据，无法直接用于计算或模型训练，需要在建模前对数据进行预处理，即 ELT（Extract-Load-Transformation，抽取-加载-转换）模式，对数据进行清洗和处理。

人工智能技术也开始应用于数据处理的过程中，AI 可以反哺大数据的处理能力，通过机器学习技术识别哪些数据有问题，如果确认数据特征，机器可自动对数据进行处理，提升大数据处理效率。

（2）数据存储与计算

数据汇聚之后，企业根据数据规模、数据应用方式等选择合适的存储组合满足企业的数据战略与数据应用需求。

企业最早的数据存储计算技术采用数据库，主要分为操作型数据库和分析型数据库。随着数据库大规模应用，数据量飞速增长，为了研究数据之间的关系，挖掘数据隐藏的价值，面向分析型数据处理的数据仓库成为广泛应用的数据存储技术，数据仓库将数据按照所需要的格式提取，再进行数据转换、清洗和装载，实现多个异构的数据源有效集成。

数据仓库主要处理历史结构化数据，但随着技术发展，企业面临越来越多的半结构化数据和非结构化数据，数据湖应运而生。数据湖能处理所有类型的数据，且包含数据更多的相关信息，拥有足够强的计算能力。

数据计算能力与数据存储紧密相关，随着数据规模不断增加，除了存储能力需要升级，数据计算框架也需要根据用户的需求场景进行细分，例如批量计算和流式计算。

(3) 数据分析与挖掘

数据分析与挖掘是数据智能最核心的环节，基于对数据的分析和挖掘，企业可以从数据中提炼、发掘、获取有提示性和可操作性的信息，从而为智能决策提供支持。

数据分析指根据分析目的，用适当的统计分析方法及工具，对数据进行处理与分析，提取有价值的信息。主要采用对比分析、分组分析、交叉分析、回归分析等常用分析方法，多面向结构化数据。

数据挖掘指从大量的数据中，通过统计学、人工智能、机器学习等方法，挖掘出未知的、且有价值的信息和知识的过程。主要采用关联规则、分类与标签、聚类、决策树、序列模式、神经网络等方法。数据挖掘主要面向非结构化数据的分析，主要应用的支撑技术为机器学习、深度学习、NLP、知识图谱、计算机视觉等。

机器学习平台等支持算法可视化建模与交互式开发环境代码编写的平台也成为重要支撑，使得数据分析与挖掘也逐步呈现自动化、低门槛的发展趋势。

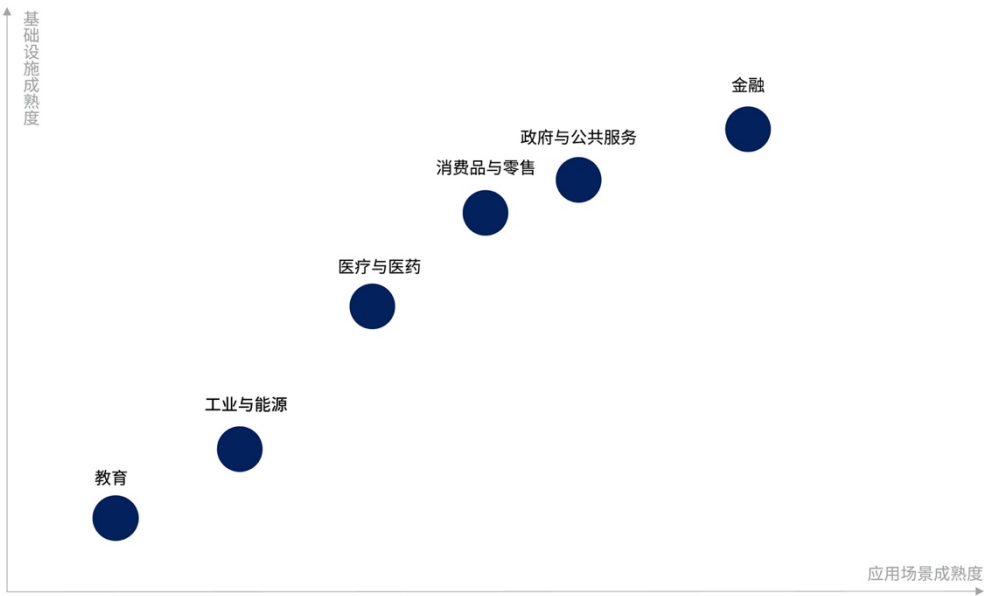
(4) 数据展示与交互

在数据服务体系中，对数据的计算经过封闭生成 API 服务，让数据快速应用到业务场景中。用户需要以直接的方式获取数据分析结果，同时可根据业务需求与系统进行交互，获取任意维度的数据分析挖掘服务，以支持数据化运营和分析决策等场景。常见的数据展示和交互应用有数据大屏、数据报表等。数据可视化技术是此阶段的主要支撑技术，典型的技术是 BI 工具等。

1.3 数据智能的应用场景

数据智能已经在各行业的业务场景中落地，但各行业的应用成熟度差异较大。

图 5：各行业数据智能应用成熟度



图：爱分析绘制

ifenxi

数据智能在不同行业的应用场景范围、渗透率均不同，这与行业本身的数字化基础设施成熟度、数据智能与行业应用场景的结合度等因素相关。

数字化基础设施成熟度主要包括各行业的信息化、数据资产化、云化等方面的成熟度；数据智能与行业应用场景的结合度指各行业数据智能的应用阶段、业务环节改造程度和数据智能应用场景覆盖度等。

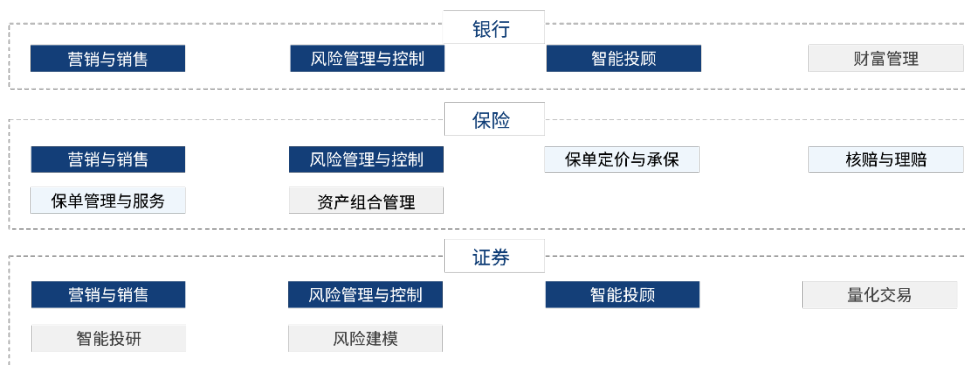
在行业应用场景中，金融、政府与公共服务属于数据智能应用的高渗透行业，金融行业已经能够实现基于数据的智能决策，并实现显著的效益，逐步向数据驱动阶段过渡，较为成熟的应用场景有智能营销、智能风控等；政府与公共服务领域的数字智能场景主要有智慧政务、智慧安防等；消费品与零售正在加速应用数据支撑其业务，整个消费品与零售行业的商业模式也发生了剧变，主要体现在智能营销与全渠道中台应用场景；工业与能源等传统制造领域数据来源与形式复杂，数字化程度较低，数据智能渗透率低，但企业已经意识到数据的价值，整体行业处于数据智能应用的起步阶段，目前主要在智能制造方面发力。

随着 5G、物联网、人工智能等技术的发展与成熟，非结构化数据的融合和应用受到各行业的重视，部分原本数字化程度很低的领域出现了基于数据智能技术应用的新业态，例如自动驾驶、在线教育等。

(1) 金融

金融领域数据智能应用成熟度高，其基础设施和应用场景的成熟度都领先于其他行业。

图 6：数据智能在金融行业的应用场景



图：爱分析绘制

ifenxi

在基础设施方面，金融行业业务对数据强依赖，因此，各金融企业基本都已经完善的数字化基础设施和数据应用，技术团队强大，数字化投入金融大，已经广泛应用大数据、云计算等技术。

在传统金融应用场景方面，营销、风控、反欺诈等场景是非常成熟的数据智能应用，已经覆盖信贷完整业务场景。目前，已经可以实现机器自动审批与放贷。金融领域的数据智能应用场景正在逐步向客户服务、产品设计等业务场景延伸，例如，银行可以根据用户数据，在用户不同人生阶段设计不同类型的产品、不同风险等级的产品；也可以基于企业的不同阶段开发设计产品。

在证券行业，机器学习及大数据将会成为新的投资策略持续产生的源泉。例如，智能投顾可基于客户理财需求，基于投资组合理论，通过算法和产品搭建数据模型，完成理财顾问服务。

在保险行业，大数据在风险评估与定义、交叉销售、客户流失管理、理赔欺诈检测与理赔预防与缓解等环节都能产生影响。例如，通过可穿戴设备实时监控人体健康，从而弥补对细分群体生老病死的概率评估的不足，保证此类保险业务的推进。

(2) 政府与公共服务

政府与公共服务领域数据基础设施相对成熟，其数据应用成熟度目前处于数据决策阶段。

图 7：数据智能在政府与公共服务行业的应用场景



图:爱分析绘制



政务领域具有最完备的数据源，信息化投入大。智慧城市的推进，带来更加丰富的数据。同时，政府云化进程快，数据资产成熟。各地政府正在加速出台数字政府规划计划，建设大数据中心，加速数字化基础设施的完善。

在应用场景方面，早期政府主要应用数据大屏，实现业务监测。近年来，随着技术和数据资产的成熟度提升，政府部门能够实现数据洞察与数据决策，例如，在智慧政务领域，可基于大数据和人脸识别、语音识别等技术提升业务办理的效率；在智慧安防领域，基于视频监控数据和人脸识别、机器学习等技术，实现快速破案。此外，精准扶贫、环境治理等场景也是数据智能渗透率较高的场景。

(3) 消费品与零售

受到电商的影响，消费品与零售是数据智能应用渗透率提升最快的领域之一。

图 8：数据智能在消费品与零售行业的应用场景

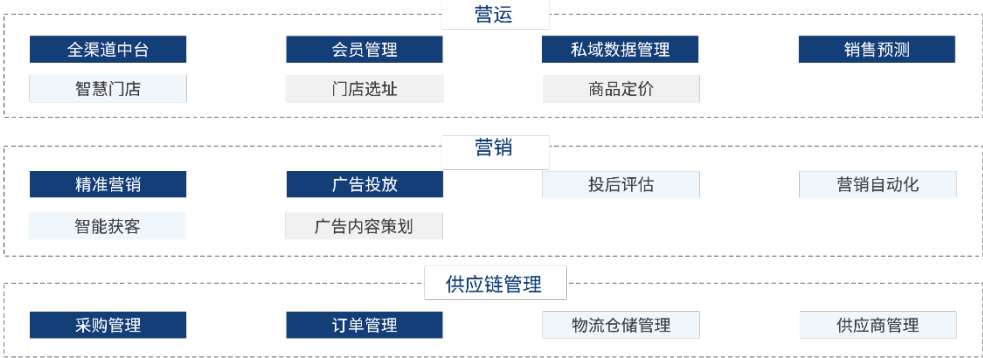


图:爱分析绘制



在基础设施方面，消费品和零售行业的数据丰富，但一般都来源于多端多渠道，包含门店、APP、小程序、服务号等渠道，对渠道整合有强需求。零售行业目前的云化和数据资产化正在快速推进。

在应用场景方面，已经可以对用户进行全生命周期跟踪，企业可以基于海量数据形成消费者画像，线上化智能营销的应用在逐步趋于成熟。例如，基于用户画像进行千人千面的推广和精准营销。企业根据需求收集精准营销需要的数据，打通所有相关数据后，通过算法，实现对用户偏好的挖掘，从而实现客户的定制化营销及服务。

未来，智慧门店等线下渠道的数据智能渗透率会逐步提升。

一体化数据分析平台，赋能德克士创新连锁品牌运营新模式

数据显示，目前中国餐饮市场规模达到4万亿，且以每年20%的速度增长。但餐饮业面临着堂食流量减少、运营成本增加、渠道拓展困难等问题，同时，餐饮业逐步向生鲜、团餐等多元化模式发展，数字化平台和外卖系统更加成熟，疫情更是加剧了这种趋势的发展。因此，餐饮行业逐步发展为数据驱动的业务形态，餐饮企业亟需能够敏捷应对客户需求的数字化系统。

德克士成立于1988年，是知名的快餐品牌，其在全国的门店数量已经突破了3000家。面对餐饮行业的变革与挑战，德克士也在推进数字化转型，而实现数据驱动的业务经营分析是基础。在业务经营数据分析方面，德克士面临着以下挑战。

第一，德克士连锁店覆盖多个城市，德克士连锁店和总部以及各业务系统的数据分布在多个城市的门店，有加盟店和非加盟店，涉及多个部门及系统供应商，如包括收银系统、财务系统、会员系统、外卖平台的运营，以及不同的小程序供应商等，存在大量数据孤岛情况，且其数据结构与规则均不相同，要保障在数据安全的前提下实时地向总部传输数据，对企业是一个巨大的挑战。

第二，原有系统采用多供应商，系统供应商的系统老旧，不易做大规模修改与延展，在原有的系统和数据结构下做数据分析会对企业的系统造成很大的压力；同时，旧系统难以为新形态的数据交互提供快后且准确的交互模式，企业需要不断对接不同的应用服务，重复造轮子。

第三，当前的数据系统无法支持实时运营数据及客户反馈可视化，德克士需要系统具备可视化能力，以更有效率地指导品牌运营。

为了解决上述问题，打造准确、敏捷、易用的数据分析平台，德克士与餐道合作启动了BI系统的建设。

餐道是一家智能餐饮系统服务商，产品服务涵盖全渠道管理SaaS平台+业务数据双中台，通过“系统+数据+服务”赋能餐饮零售企业线上业务，为企业提供一体化数字解决方案。截至2020年8月，餐道服务于国内外连锁品牌超过2000个，覆盖了全国400多个城市的55000多家门店，包括汉堡王、哈根达斯、DQ等连锁餐饮集团，以及京东便利店、壳牌便利店等新零售企业。

餐道的全渠道中台连接全渠道，解决客户业务中“数据孤岛”的问题、统一输出可视化信息助力客户高效分析业务，同时，提供共享API服务与服务对接，为餐饮新零售提供业务系统重构基础，连接前后端应用，满足数据实时互通、快速拓展新渠道业务需求。

图 9：餐道全渠道中台体系



在不断服务客户的过程中，餐道也逐步完善和迭代其产品模块。餐道有标准化数仓平台，基于其中台系统，有超过 30 套大型分析工具，包括产品分析、销量分析等，共计 100+功能项。在数据源方面，餐道可以接入不同渠道的数据，不限制接入平台数量。因此，餐道具备 平台、数据、对接等多方面的能力，赋能餐饮企业不断优化运营管理。

在数据治理方面，餐道为德克士打造数据中台与应用中台的“双中台”系统，打通数据，将德克士线上渠道和自有渠道的分散数据汇集在平台中，包括外卖数据、餐单数据、制作数据、配送数据、客户评论等，同时提供标准数据结构、定义数据内容，为不同渠道的数据制定规则，保证各平台数据的格式与定义一致，打破“业务孤岛”与“数据孤岛”，避免重复造轮子。餐道平台提供标准信息交互接口 API，持续抓取各端采集数据并转换为标准格式。

同时，基于统一的数据中台与应用中台，德克士可以灵活地拓展业务，并选择最优服务供应商。

在可视化方面，基于统一的数据结构，餐道为德克士建立 BI 系统。

餐道提供两种 OCRM 报表：标准报表和自定义报表。标准报表从运营、销售、财务、用户这四个板块分别展现了不同维度的明细数据，可通过数据分析出门店日常运营情况，找出运营问题所在，是否是配送不达标或者是营运 SOP 不规范；自定义报表可以根据商家特定需求，如与平台对账或者自定义所需字段和计算逻辑，快速得出汇总数据。德克士可以根据业务需求将数据分派到不同业务部门里德报表分析。

例如，餐道会把各个平台的舆论、客户评价等数据采集到 BI 系统里面，德克士可以很方便地在系统对各个平台客户对德客士的评价与反馈进行综合分析。

目前，BI 系统上线后已经覆盖德克士 2190 多家门店，其中 80%都是加盟店。

在具体项目实施过程中，德克士积累了宝贵的餐饮连锁店数字化转型经验。

首先，餐饮行业的未来是以数据为驱动的，餐饮品牌要注重数据价值，所以企业需要将数据放在战略和管理层面的首要位置。

其次,企业需要由上到下地推动数字化转型,更新组织架构,协调各方参与者,包括 IT 部门、业务部门和企业不同的供应商,让不同部门迅速地融合并加快项目实施进度。

未来,德克士计划进一步上线订单管理系统,接入德克士会员系统、配送系统、客服中心以及外部其他接口,并上线接单管理、聚合配送、营运与监控管理、营销管理等 SaaS 服务。

数字化零售平台，打造汽车品牌高端用户体验

某全球化汽车公司致力于通过提供高性能的汽车产品与极致用户体验，为用户创造愉悦的生活方式，打造全球范围内的“用户品牌”。其旗下品牌用户中心致力于为车主及其朋友打造一体化的自由空间，提供包括轻餐、精品、书籍租赁、空间租赁等一系列服务，发展多种业态服务，力求为用户带来线上线下一致性的体验。

该品牌拥有完备的开发团队和庞大数量的业务系统，在提升多业态服务质量的过程中，面临的主要挑战来自于多业态服务系统造成的前端用户体验不佳和后端运营效率低。

在用户体验方面，第一，用户支付繁琐，现金、刷卡、移动支付、企业币等多支付方式在一台设备无法集成；第二，多渠道数据割裂，不同渠道的订单、库存、促销、会员权益等数据和规则均不一致，客户多渠道消费的体验不佳。

在业务运营方面，由于存在多业态系统，且数据割裂，一方面，前端支付数据无法在后端业务系统中整合，餐饮和零售的收银和财务结算需要在两个系统中完成，操作烦琐，运营效率低；另一方面，实时数据无法及时反馈至前端业务人员，前端无法快速灵活地响应。

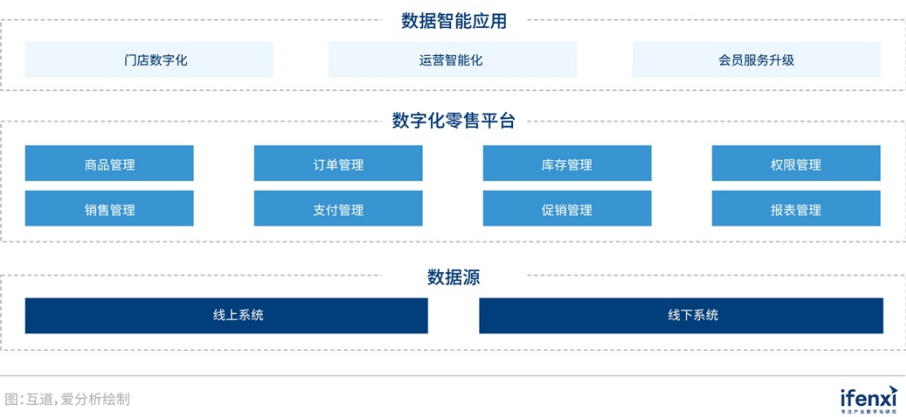
为了实现线上线下融合并满足消费者日益升级的消费需求，该品牌开启全渠道业务布局。多业态服务发展中所关注的商品管理、订单管理、库存管理、支付管理、报表管理等都有了数字化升级需求，需要解决客户收银体验一致性、线上线下促销玩法一致、门店运营的数据实时性等问题。

该品牌需要将分散在各个系统的数据打通，但庞杂的系统对接周期长且十分困难，完全舍弃或替换原有系统则意味着巨大的投入成本。因此，品牌希望有兼容性更强的系统，在兼容原有框架的同时具有高度扩展性，不仅能够完美对接企业原有系统，将割裂的数据打通，统一管理，更重要的是能够将数据实时反馈到前端的各个业务触点，支持业务的高效运营；同时需要前端对接数字化工具，赋能品牌提供更高端细致的服务。

该品牌数字化运营部门与互道信息密切合作，成立项目组，共同解决其品牌用户中心的后端运营和用户体验问题。互道信息技术(上海)有限公司是一家行业领先的数字化零售解决方案供应商。互道融合创新技术，为企业提供高性能、高可靠的技术引擎和实时、可复用的全渠道业务引擎，支持零售业务场景快速创新与拓展，助力全球众多中大型零售品牌商实现数字化转型。

针对品牌用户中心的诸多痛点，项目组以强大的中台产品——互道数字化零售平台DataForce，为品牌实现多业态零售的数字化升级。

图 10：汽车品牌数字化零售平台解决方案



(1) 数字化零售平台技术赋能——打通数据壁垒实现业务联动

在不替换企业原有系统的前提下，互道帮助品牌对接了企业原本割裂的众多系统，将商品数据、库存数据、订单数据、会员数据等打通，利用数字化平台进行统一精细的管理，并将实时数据及时反馈至前端业务，为业务运营提供决策依据，让品牌用户中心的前端业务运营获得更高灵活性和更快响应速度，提升运营效率。数字化零售平台可复用的能力帮助企业减少了重复对接的成本，提高了系统对接效率及数据一致性，并支持第三方伙伴跨语言、跨平台的扩展能力，能够从原来的单点作战转换为并行作战，为企业未来随市场而不断变化的业务需求也能快速满足，快速落地创新业务场景。

(2) 数字化零售平台业务赋能——开启全渠道一体化业务运营

基于数字化零售平台实现数据融通，互道为品牌搭建了全渠道业务管理中心，包括订单、库存、促销、会员等，优先为全渠道业务处理零售业务核心数据，提升业务运营效率的同时，消费者服务体验得到大幅提升。例如，支持各种复杂促销规则的全渠道促销引擎，帮助品牌实现了线上线下营销优惠统一，品牌导购端可以直接查询用户优惠，快速核销，保障用户体验。全渠道订单路由为品牌实现了 O2O 全渠道订单的智能接单拆单派单，提升了订单履约能力，而消费者可以在 APP 内查看到自己的全渠道订单状态，客户体验得到提升。同时基于订单通、库存通、商品通等，支持了多业态品牌商品的录入，让品牌全渠道一体化业务运营可以更高效开展。

(3) 门店数据智能终端 ShopForce——数字化门店伴随式服务

此外，针对品牌希望在业务前端以数字化工具赋能，为用户群体提供更高端细致服务的需求，互道为品牌搭载了门店数据智能终端 ShopForce。通过 ShopForce，门店店员得以摆脱收银台束缚，为顾客提供随时随地的服务，提供移动化的点单支付体验。ShopForce 支持全面的支付方式，并可组合支付，支持导购端积分抵现，只要 1 次识别用户，即可享受全流程收银服务，保证了用户体验的一致性，让品牌客户在用户中心空间内，享受到作为高端用户的极致服务体验。

在系统端，数字化零售平台的功能在持续迭代，基于互道搭建的中台系统可以快速对接新的应用，不用再对接后台多套业务系统。例如，一期先实现了零售与餐饮的标准功能，二期逐步增加了餐饮 BOM 管理功能，赋能企业做好餐饮的精细化管理；三期增加虚拟商品、车商城、订单流程优化等功能，同时支持云仓，实现线下购买线上实现配送流程的功能。

在用户端，项目上线后，用户体验大幅提升。原来用户具有 POS 机、微信、支付宝、积分购买等多个支付场景，通过对接企业支付体系及优化购物流程，前端扫描用户会员二维码识别用户后，一线人员可以用 PAD 用满足大部分支付场景，用户在 15 秒内即可完成支付。

用户体验提升的背后是整个系统的优化，包括操作人员使用系统的便捷性、后台数据处理与查询、物流紧急补货调货等，才能提升线下整体的体验流程。例如，会员在线下购买之后，能实时地同步到 APP，系统可以自动识别用户积分、购买历史、订单数据等，赋能提升用户体验。

在业务端，基于数字化零售平台，店员和运营人员都可以获取多角度的数据分析。一线服务人员可以获取会员、门店、销售等数据，赋能其更好地服务用户；后端可对商品、订单、库存、促销数据进行统一管理和实时分析，运营人员可获取深度分析报表，优化业务运营。同时，总部可对各门店端的一线人员进行远程培训，且系统提升整体运营效率，帮助企业精简运营人员，降低人员成本。

图 11：数字化零售平台实现的价值

原业务系统	数字化零售平台
餐饮和零售需要到不同收银台	贴身服务，一站式购买
完成一笔订单支付需要2分钟	15秒完成一笔订单支付
不同支付方式需要不同设备	IPAD覆盖95%支付场景
员工培训需要半天	15分钟员工即可学会
收银报表需要单独后台	各类报表实时呈现
新的前端应用需要对接多套系统	利用中台标准API对接新应用

图：互道，爱分析绘制

ifenxi

在整个项目过程中，实时数据同步是核心难点，而该汽车品牌的解决方案也可以为其他企业提供一定的借鉴意义。

该品牌原有采用 T+1 非实时 BI 系统，要保障实时数据运营，需要将实时数据和后端数据打通，需要根据系统的情况，采用不同类型的对接方式。不同于传统 SaaS 产品仅开放 API 的方式，互道采用“千系统千面”的系统对接接口，根据具体情况制定具体的对接方案，保障接口响应时间符合用户需求。例如，互道可以采用订阅分发的方式，保障数据同步的同时减少系统本身的压力。

(4) 工业与能源

工业与能源行业的基础设施薄弱，数据智能应用处于起步阶段。

图 12：数据智能在工业与能源行业的应用场景



在基础设施层面，工业与能源领域的数据来源多，且存在大量非结构化数据，不易采集，随着物联网、5G 等技术的成熟与普及，工业与能源数据才开始具备完整的数据采集基础。

在应用场景方面，工业与能源行业具有复杂的生产运营体系，需要协调大量的设备与资源，过往主要依据人工经验进行观察与决策，业务的数字化改造程度低。目前初步实现了业务自动化，但在数据智能方面相对滞后。

但是，工业与能源领域是数据智能最有潜力的应用领域之一，工业的智能化升级是国家的重要战略。工业互联网汇集设备、产线和企业数据，可以结合机器学习、物联网技术，深入挖掘数据背后的应用价值，将会极大提升企业的运营效率，降低各类成本。

例如，目前在部分企业可以实现产品质量控制智能化，基于生产过程中的大数据，分析生产流程，一旦流程偏离标准工艺，系统即产生警报信号，从而快速发现错误。同时，可对工业产品的生产过程建立虚拟模拟模型，仿真并优化生产流程。

(5) 医疗与医药

医疗与医药行业的具备一定的数字化基础，数据智能应用处于决策阶段。

图 13：数据智能在医疗与医药行业的应用场景



图：爱分析绘制

ifenxi

在基础设施方面，2019 年全国卫生总费用约 63800 亿，其中医疗信息化投入占比约为 1%。医疗机构已经实现了基本的信息化建设，医疗大数据主要有诊疗数据、研发数据、患者数据、支付数据等，其中诊疗数据占比 90%。目前数据正在快速地电子化，但数据格式繁杂、目前还没有实现多源头数据的整合与分析，下一步需要完善数据的互联互通。

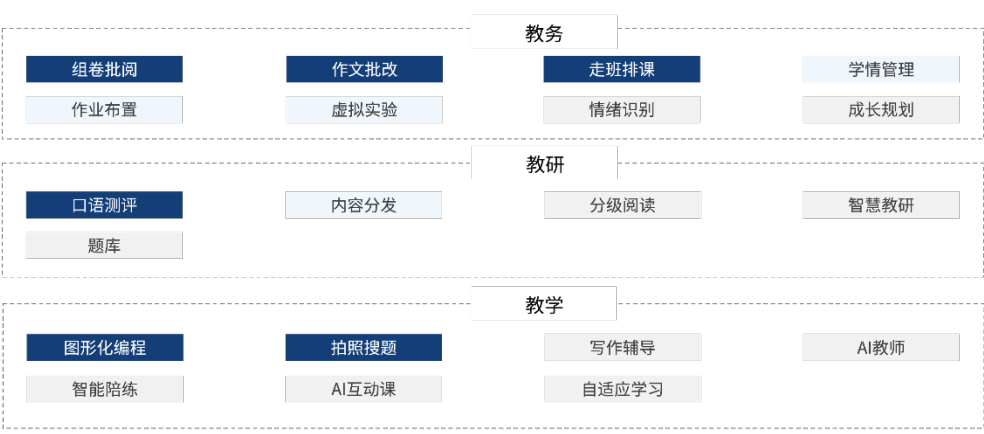
在应用场景方面，临床操作与研发是数据智能最有潜力的应用方向，例如基于用户健康数据以实现健康管理；整合临床数据和基因数据实现癌症风险评估、预防和治疗；基于大数据预测治疗效果降低花费。

除了临床决策以外，基于大数据还可实现计算机远程监护以进行慢病管理，以及利用大数据预测流行病趋势。例如，谷歌通过流感相关关键词的搜索趋势，推出了“流感指数（Google Flu Trends）”产品，曾成功地在甲型 H1N1 流感爆发几周前成功预测其在全美范围内的传播。

(6) 教育

教育行业的数字化转型正在加速进行，数据智能应用处于起步阶段。

图 14：数据智能在教育行业的应用场景



图：爱分析绘制

ifenxi

教育行业已近经历了广泛的在线化历程，随着 K12、语培等领域的一大批在线教育公司崛起，教育行业初步积累起大量的数据。目前，教育行业的数据积累与核心教学环节相关的数据积累还相对较少，对于非结构化数据的处理分析技术还不成熟，整体落地成熟度还有较大提升空间。

在应用场景方面，目前主要的应用场景主要是智慧校园管理，但数据智能技术已经开始渗透到师资教研、教学服务等业务流程当中，例如，个性化教学产品可以根据老师的教学内容与学生学习的行为数据自动输出学生的教学计划与学习内容，同时可自动批改并利用智能学习引擎自动分析学生的学习效果，形成教学闭环。另外，在师资教研环节，已经有头部的在线教育机构通过语音识别技术分析老师的上课过程，挑出最优秀老师的最佳实践进行分析，沉淀到内容当中，从而提升教学质量和教学的标准化。

(7) 通用职能部门场景：营销

在营销领域，线上营销场景已经实现了完全在线化、数据化，线下营销场景也随着智能终端及智能手机的普及逐步向数字化转型。各品牌也加速数据中台等基础设施的建设。在应用场景方面，数据智能已经开始渗透品牌商营销链路各环节。

图 15：数据智能在营销领域的应用场景



图：爱分析绘制



在用户洞察环节，品牌商通过建立客户数据平台（CDP）整合线上、线下自有用户数据，统一用户画像，补充第三方数据管理平台（DMP）的用户标签，为营销活动开展提供支撑；在营销策划环节，品牌商可借助 AI 辅助图片类、文案类创意内容生成；在用户触达环节，在全渠道下可基于用户数据进行广告投放，同时，借助对效果广告的用户转化链路分析，指导品牌和效果广告的人群精准定向；在用户转化环节，利用 CRM 系统管理用户在线上、线下的留资数据，实现营销自动化；在用户运营环节，针对不同人群执行差异化的营销策略，实现用户购买升级，单客 LTV 提升。

“大数据+AI 视觉”助力游戏行业破解营销买量和创意痛点

随着中国互联网用户增长红利的见顶，内容创意营销对各行业的广告投放越来越重要。以游戏行业为例，中国的游戏行业市场并不大，数据显示，每年在买量端参与竞争的厂商只有约 500 多家。但是对于游戏的各垂直品类，各家游戏公司的投放预算都比较大，广告投放抢量程度比其他行业都更为激烈。

早期广告买量投放建立壁垒都是依赖于信息不对称和规模化的优势，近些年广告买量竞争越来越激烈的原因就在于信息趋向于透明，媒体渠道的规模化优势在锐减。成熟渠道遭遇瓶颈，需不断开拓买量渠道，获客成本变高，游戏厂商的压力日趋增大。

此外，与电商或者教育等行业追求有效订单数不同的是，游戏是虚拟的、线上的产品，希望获得的是即时反馈，对投放效果的追求远甚于其他行业。围绕着效果营销，游戏推广环节对能带来良好营销效果的优质内容创意的需求越来越大。但是，行业也面临着创意短缺、产能不足、素材制作成本高等问题亟待解决。

图 16：游戏营销行业痛点

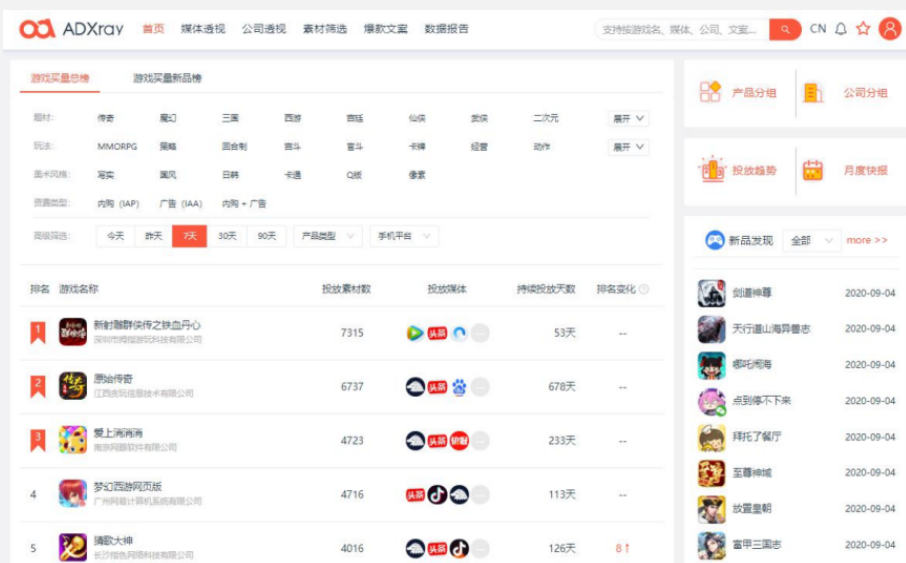


由此，当前游戏行业的买量现状呈现出游戏买量更注重品效合一、买量市场头部效应显著、游戏买量进入短视频时代等特征，成本高、创意难、渠道难成为游戏行业买量痛点，精品化、智能化、精细化成为未来的趋势。

DataEye 是一家专注于移动广告情报分析的大数据公司，为垂直领域提供移动广告情报分析的数据工具服务及定制化数据服务，推动行业移动营销的创新及发展。DataEye 的聚焦方向是效果的内容营销端，提供与数据相关的行业情报分析工具和买量创意的全案营销服务。

DataEye-ADXray 是一款专注于游戏行业广告投放的情报分析工具，利用海量数据存储、AI 视觉智能等核心技术，通过抓取并持续追踪手游广告素材以及营销创意，针对性解决游戏行业渠道难、创意难等买量痛点。

图 17: DataEye-ADXray 工具



DataEye-Tidea 添羿则是一项专注于移动买量创意的全案营销服务，借助 DataEye-ADXray 的海量数据支持，利用多年的技术及移动广告数据沉淀，深入手游买量市场，基于精准的买量素材分析，针对不同类型游戏定制爆款创意，有效解决手游买量创意短缺、素材产能不足的难题。

添羿基于“大数据+AI 视觉”的解决方案，能够提供情报分析洞察以及内容创意服务。添羿首先对外部投放内容进行数据化分析，为客户提供大量的数据分析研究，对广告投放数据和趋势做出解读，以数据驱动确定客户的广告内容创意创作方向。接下来的创作环节，剧本、脚本等都由 DataEye 提供。

广告、短视频等素材都是添羿很重要的挖掘方向，对短视频的分析包括两个方面：一是识别广告投放主体以及投放时间等原始数据，用 DataEye 自己的算法预估投放的概率；另一方面 DataEye 会针对文案内容做分析，把市面上所有的视频在自己的系统里还原成脚本，应用 AI 技术对其进行拆解，通过文本语义挖掘的方式找到内容创意的亮点，打上与游戏营销卖点密切相关的更多高频词标签，最终提升内容创意生产制作效率。

1) 基于 ADXray 海量数据，拥有庞大的数据能力

营销创意监测产品 ADXray 拥有强大的后端视觉和视频标注系统以及便于训练各种模型的训练系统，核心技术主要在于海量视频数据的存储以及对视频、视觉数据的理解。ADXray 90% 以上的数据标注都是由机器完成。

目前 ADXray 已和国内所有主流研发商、发行商达成合作，行业覆盖率近 100%，每日分析游戏创意数据超 10 万条，ADXray 创意库累计收集分析超过 8000+ 万条短视频创意，覆盖 2017 年以来所有的游戏优秀投放案例。

2) 基于 ADXrayPro 受众人群众分析&AI 智能处理，具备数据分析能力

ADXrayPro 基于大数据统计分析技术，对在主流短视频平台投放的视频广告进行剖析，剖析维度包含视频广告的热门评论、评论人群属性、覆盖人群偏好的广告类型以及受众人群偏好的视频类型。

相比于 ADXray 的基本数据分析，添羿对数据做了进一步的深度挖掘，除了单一的视频持续投放时间维度，还对视频进行了深入的理解，扩展了更多的分析维度。通过这些技术添羿能够第一时间呈现创意市场的迭代和趋势，帮助客户分析定位游戏的创意方向，以此提升爆款率。

3) 采用买量策略脑图&竞对分析，具有丰富的创意策划能力

在进行创意之前，添羿团队将根据 ADXray 以及 ADXrayPro 的统计数据对该游戏类型进行全方位剖析，在游戏简介、玩家需求、游戏爽点以及爆点等多个维度对游戏进行理解来定位创意方向以及买量策略方向。

此外，在内容创意方面，DataEye 有自己的创意生产线，目前已经有两个创作基地，此外还有来自外部的视频制作供应商合作。客户可以直接从 DataEye 采购内容。DataEye 最大的客户例如网易、腾讯、恺英网络，每个月定期从 DataEye 采购大量用于广告投放的创意内容，两大素材制作基地正是后端支撑。

目前，DataEye 与客户合作升级《梦幻西游》、《梦幻西游网页版》、《梦幻西游三维版》买量营销打法，为客户提供了受众分析、竞品分析、市场基本面分析、素材策略等多维度策略分析及广告素材制作服务。

《梦幻西游》系列自 2020 年大范围买量营销推广以来，招回大量“老玩家”的同时吸纳大量“新玩家”，《梦幻西游网页版》除 Android 端、移动端 H5、PC 端微版等，仅 iOS 端月流水便已过亿。

鉴于不同垂直领域的营销内容创意都需要解决信息透明问题的共同痛点，未来 DataEye 会将在游戏行业的商业模式直接复制到诸如电商、教育、金融、网服等行业，向广告投放的其他领域有延伸拓展。

某欧洲老牌药妆零售商借力每日互动智能营销解决方案入局国内市场

数据显示，中国美妆市场正在迅速发展壮大，目前已经成长为全球第三大美妆市场。在天猫平台，美妆已经超过母婴、食品，成为增长最快的细分品类。同时，电商平台与数字创新在全球的领先发展，让不少海外品牌看到了机遇。

某欧洲药妆品牌作为一家历史悠久的老牌药妆零售商，最初以药店起家，其产品已经通过代购和出境游等方式被中国消费者所熟知。该品牌看到了中国美妆市场的巨大潜力，于 2018 年 9 月首次进入中国市场，在天猫国际开设了官方旗舰店。但面对中国美妆市场的激烈竞争与快速变化，该品牌面临着以下挑战：

- 1) 在消费升级的趋势下，中国消费者的高端美妆需求高涨，该品牌大众化的定位使其不仅面临国际高端美妆的冲击，更要应对中国本土美妆品牌的日益崛起。
- 2) 作为一家欧洲美妆品牌，它需要先了解亚洲市场消费者的消费偏好，再制定定制化营销策略，快速占领美妆市场。

为了解决上述问题，该品牌选择了每日互动作为合作伙伴。每日互动成立于 2010 年，是专业的数据智能服务商，其推出的新一代数据中台以行业数盘的形态为互联网运营、品牌营销、城市规划等各个行业的客户提供服务。该药妆品牌在广告投放过程中就使用了每日互动行业数盘中聚焦于品牌营销领域的“个灯数盘”。

“个灯数盘”涵盖多维度用户画像分析、潜客拓展、客流分析等功能，可全方位满足品牌人群洞察需求，在每日互动“精准投放定向服务”的配合下，可帮助品牌全面提升营销价值。

具体来说，个灯数盘从人口属性、兴趣爱好、媒体偏好等维度，对指定人群进行全面画像分析，洞察目标人群特征。基于前沿 AI 机器学习，实现目标人群 lookalike 扩量，帮助品牌找到更多具有相似特征和偏好的潜在消费者。而在程序化投放过程中，在个灯数盘的基础上，每日互动基于移动端数据筛选流量，依托智能算法，可以一键按需智能排序选量，并持续优化 TA 浓度，实现精准高效的媒体曝光。

以该欧洲美妆品牌的一款产品广告投放为例，品牌方借助个灯数盘，精准找到目标消费者，挖掘潜在消费者，并联合站内站外资源进行大规模精准投放，提升了品牌认知度，提高了产品销量。

图 18：基于个灯数盘的精准投放



第一步：洞察目标消费者特征，分组测试确定营销重点人群

在洞察环节，品牌方借助个灯数盘的算法技术，对天猫站内产品相关用户数据进行充分挖掘，找出消费者的行为特征并生成了用户画像标签，再基于属性、行为特征等维度，对种子用户进行聚类细分，将用户分成了精致女孩、积极女性和知识白领三类种子人群。之后，品牌方从中各筛选出小批量用户进行投放测试。通过分析广告的曝光和点击情况，该品牌最终将“知识白领”确定为后续大规模站外营销活动的重点开展人群。

由于站内的激烈竞争使得流量获取成本不断攀升，为避免站内红海的正面争夺，品牌方听从个灯数盘的建议将眼光转向站外，借助每日互动在站外的丰富优质流量进行大规模投放，以优质的成本有效引流。

第二步：扩量算法挖掘潜在消费者，智能选量实现站外大规模精准投放

在投放环节，品牌方先依托每日互动的数据库能力和机器学习，找出高潜力用户“知识白领”的显著特征，并通过 lookalike 扩量算法，找到了 3000W+ 具有相似特征和偏好的潜在消费者。之后，该品牌基于个灯数盘能力，择优选择知识白领人群偏好明显的短视频、新闻类移动媒体作为投放途径，并与国内主流 DSP 服务商完成对接。

图 19：扩量算法挖掘潜在消费者



在投放前，品牌方采取每日互动独创的“子弹夹模式”对媒体流量和 TA 之间进行契合度排序，确保广告都能匹配到优质的流量进行精准曝光。每日互动会在流量匹配过程中，基于全网覆盖的移动端数据，对流量的真实性进行判断，过滤掉虚假流量。

投放时，该品牌在每日互动指导下，通过紧密联动站外站内，将站外的优质流量精准导入站内，提高了目标消费者从曝光到购买的转化率。

第三步：通过品牌 Databank，对投放结果进行评估验证

在评估验证环节，品牌方在营销活动结束后，将数据导入品牌数据银行（Databank），对广告效果进行评估验证。结果显示，此次投放人群的品牌关联度、购买/复购用户占比明显高于普通电商活跃人群，实现了品牌宣传与线上销售的“品效合一”。数据显示，本次营销活动不仅提升了品牌认知度，线上销量还比预期高出了 127%。

此次合作中，每日互动充分发挥自身数据和技术优势，帮助品牌方全面洞察目标消费者，激发品牌新的营销思路和创意，不仅为品牌抢占消费者心智加码，更为深化中国美妆市场的数字化营销挖掘出更多可能性。

CHAPTER

02

数据中台
支撑数据智能落地

2. 数据中台支撑数据智能落地

2.1 数据中台是数据应用的基础设施

企业在实现数据应用的过程中，面临着大量技术和业务组织层面的挑战。数据中台作为一种新兴的架构，是企业实现数据驱动业务的必经之路，是支撑数据应用的新基础设施，对企业的数字化转型具有重要的意义。

2.1.1 企业数据应用挑战与数据中台价值

对于数据中台的概念和实践，目前行业内尚未形成统一的认知和通用标准，但随着企业的数据意识与认知增强，越来越多的企业对于数据中台的价值形成了共识。

数据中台不仅仅是技术或产品体系，更是一种战略选择和组织形式，是依据企业特有的业务模式和组织架构，通过有形的产品和实施方法论支撑，构建一套持续不断把数据变成资产并服务于业务的机制。

在企业推进数字化转型和数据应用的过程中，通常会面临以下几方面的挑战：

第一，企业内部有大量数据孤岛，传统企业的数据大量存在各个孤立的数据业务中，且数据碎片化，无法统一融合赋能业务。数据中台通过对跨平台海量数据进行采集、计算、存储、加工，同时统一数据标准和口径。数据中台把数据统一之后，会形成标准数据，再进行存储，形成数据资产，进而为业务提供高效的数据服务。

第二，传统企业内部通常是 IT 部门负责处理和管理数据，业务部门无法直接使用数据。数据中台是聚合和治理跨域数据，将数据抽象封装成服务，提供给前台以业务价值的逻辑概念，能够实现数据的资产化、服务化，赋能业务部门人员进行数据分析和数据应用。

第三，一旦推进数据智能应用，企业可能也会面临大量的数据开发工作，技术门槛高、投入成本高，但面对前端的业务部门源源不断的需求，企业存在数据服务能力重复建设的问题。数据中台连接数据前台和后台，为企业提供更灵活、高效、低成本的数据分析挖掘服务，实现企业的数据开发能力的复用，避免企业为满足具体某部门某种数据分析需求而投入大量高成本、重复性的数据开发成本。

同时，数据中台实现将通用的数据服务在业务部门间共享，提升跨部门的普适性业务价值能力和企业运营效率，能帮助企业更好地响应业务需求的快速变化。

因此，数据中台可以解决客户在数据整合、数据资产管理、数据应用和数据运营等方面的问题，包括一切与企业数据资产使用相关的平台、工具、数据、组织、流程与规范等。以数据中台建设为抓手推进数据智能应用是非常有效的方式。

2.1.2 数据中台的架构

数据中台连接前后与后台，基于底层数据存储计算基础设施建设，通过数据开发、数据治理体系与数据资产管理形成可对外进行服务的数据资产，再通过数据服务体系将数据资产转变为数据服务能力，支撑企业上层数据应用。数据运营体系、数据安全管理体系则保障数据中台的持续运营。

图 20：数据中台的架构



图：爱分析绘制



1) 数据开发

数据汇聚到数据中台需要经过数据加工以及过程管控的工具进行数据处理，才可以快速将数据加工成对业务有用的数据。数据开发平台提供实时开发和离线开发工具，以及智能调度、智能运维、监控告警等一系列工具，以提升数据开发人员和分析人员的效率。

经过数据开发模块，可以形成不同业务可用的数据体系，数据体系具有一致性和可复用性的特点。

2) 数据资产管理

为了使企业业务人员更容易理解和应用数据资产，数据资产管理模块对数据资产进行管理和质量控制。数据资产管理包括数据标准管理、数据模型管理、元数据管理、主数据管理、数据质量管理、数据安全管理体系、数据价值管理和数据共享管理等职能。数据资产管理包含了传统数据治理体系。

数据资产管理的核心不在于技术与产品，更在于企业的战略、组织与制度，需要企业具有明确的战略规划、有效的组织架构、良好的数据意识、合理的制度与业务流程，才能逐步形成统一的数据标准与规范，实现良好的数据资产管理体系。

3) 数据服务

数据服务体系是连接数据资产和数据应用的枢纽，将数据资产转换为业务可用的数据服务。数据通过计算逻辑的封装生成 API 服务，数据应用可以对接数据服务 API 获取数据服务。企业的服务需求大部分是定制化的，因此，大部分数据服务并不是数据服务体系本身自带的，而是通过数据服务体系提供的快速的服务生成、管控能力来提供的。

基于数据服务体系，数据服务可通过对数据进行封装快速支撑上层应用，降低了数据接口的重复建设，保障数据服务输出的一致性、及时性、高效性。同时，通过数据拓展和数据封装形式的拓展，数据能力能不断进行扩展。

4) 数据运营体系

数据中台建成之后，需要持续运营、不断迭代，才能依据企业不断变化的业务和数据产生持续的价值。数据运营体系保障数据中台得以持续运转和迭代。

5) 数据安全管理体系

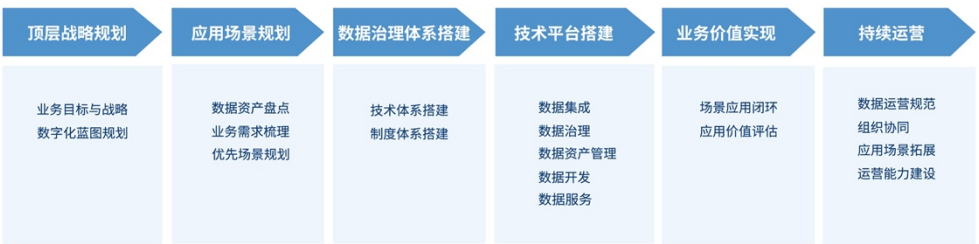
数据安全管理体系是数据中台体系的基础，包含大量复杂的技术手段，涉及数据产生、数据存储、数据使用、数据共享等全生命周期。企业数据中台汇聚了企业全量数据，应当建设多方面、多层次保障的数据安全管理体系。

和数据中台建设类似，数据安全管理体系不仅是依赖于产品与技术，而是包括了安全战略、安全组织、安全过程管理等多方面，需要企业兼顾多方面。

2.2 数据中台建设与运营指南

数据中台不仅是一个复杂的技术架构，是一套数据资产体系和数据服务能力，更是一种业务和组织架构和业务运营的流程与机制。爱分析经过多家企业的数据中台的案例调研，总结了建设数据中台过程中部分共性的最佳实践方法论。

图 21：数据中台建设与运营关键环节



图：爱分析绘制

ifenxi

2.2.1 顶层战略规划

数据中台是企业数字化转型的基础设施，是企业各部门共同的数据平台和数据服务体系。因此，数据中台建设的核心目的是服务于企业的整体战略目标与业务目标。例如，零售企业为了增加营业收入，则需要对客户进行深入洞察，进一步需要建设数据中台汇聚并利用全渠道数据。

此外，数据中台的建设不仅仅涉及技术架构，还会涉及企业的业务模式也组织架构。因此，企业应当从顶层战略为起点，根据业务目标规划数据中台的建设蓝图与路径。

同时，数据中台不仅面向技术人员，更面向业务人员。传统企业的各部门间部门墙问题明显，要实现部门间的沟通协作、共建数据中台的挑战巨大，因此，建设数据中台需要站在企业战略层面进行推进，在组织架构和资源方面给予统一的调配与支持。

因此，企业建设数据中台建设要从顶层战略规划出发，根据业务目标，从战略层面规划数据中台建设路径、调整组织架构、调配相关资源。

影儿时尚集团打造 CDP 客户数据平台，助力三年百亿目标

随着消费需求的不断细分、科技创新的深入演进，市场与业态的快速变化冲击着传统服装业的经营模式，业务运营方式呈现出从粗放到精细、从线下到全渠道融合、从注重规模到注重效果的特点，而数字化转型和智能化重塑则是传统服装行业提升竞争力核心所在。

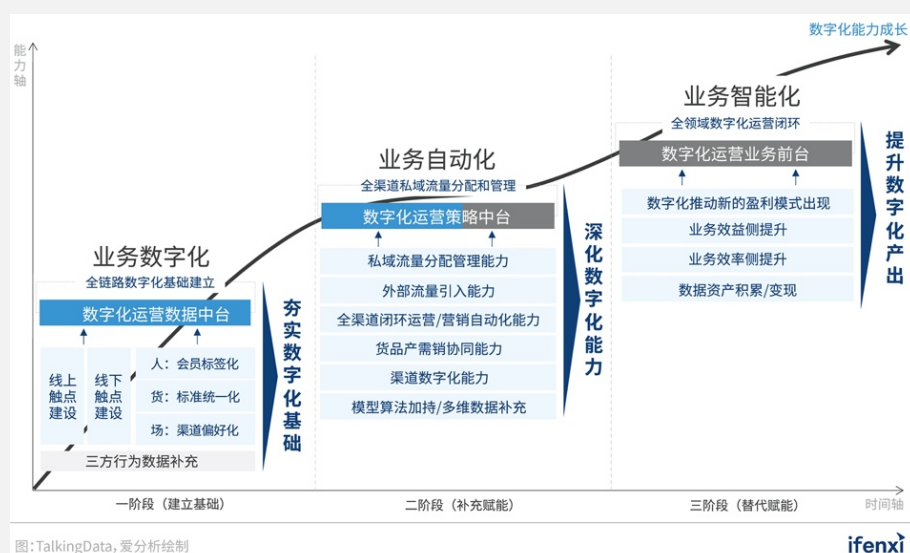
影儿时尚集团是一家大型服装企业，成立于 1996 年，旗下拥有六大独立品牌，在全国拥有 15 大区域销售分公司、1500 多家直营门店和数百万会员。2018 年下半年，影儿集团制定了未来三年收入增长至百亿的目标，而增长命题有三个重要元素：商品、市场、消费者。随着用户为王时代的到来，影儿集团希望把焦点放在消费者身上，让消费者管理成为增长的主要驱动力。

鉴于消费者运营是非常适合数据化转型落地并快速产生效果的场景，影儿集团选择了 TalkingData 作为数字化运营合作伙伴，启动数字化会员项目，构建起 CDP 客户数据平台，通过数据优化和完善全业务流程。

TalkingData 是国内领先的数据智能服务提供商，围绕 SmartDP 数据智能平台构建“连接、安全、共享”的数据智能应用生态，帮助商业企业和现代社会实现以数据为驱动力的智能化转型。

影儿集团携手 TalkingData，明确了从集团到品牌、从区域到终端、从基础平台搭建到系统自动化和系统智能化的建设路线，经过业务数字化、业务自动化和业务智能化三个阶段逐步推进，最终实现集团整体数字化能力的成长。

图 22：影儿集团数字化转型项目三阶段



第一阶段：数据化基建，用“数据驱动型业务”验证 RGM 增长模式

服装企业数字化转型的基础，就是对“人、货、场”进行数字化的重构。在平台建设层面，影儿集团与 TalkingData 密切配合，确保数据采集、数据处理、平台能力与业务应用相匹配，

实现了线上、线下触点数据合规采集以及数据的标准化、标签化脱敏处理，并打通从数据到业务应用的输出通路，确保业务侧能够准确、高效地应用数据。

在项目实施过程中，为了保证数据更业务化、平台能力更智能化，集团以“获新”为目标，先以“半人工”的方式尝试了一次数字化运营全流程。这次营销活动基于 TalkingData 建设的一方 CDP 客户数据平台，通过微生态裂变的方式打通了数字化运营闭环，从事前策略、事中调优、事后总结三方面实现了核心效率的提升。

该次“获新”活动取得了推文曝光量超 10 万、公众号新增粉丝超 4 万、新增注册会员超 4000、活动转化销售额超 800 万的亮眼成绩，然而成本支出仅为 3 万元。这次尝试在验证项目的阶段性成果的同时，也为后续数字化运营闭环的全面推广奠定了基础。

第二阶段：数据释能渠道&品牌，驱动业务变现

数据化基础能力构建完成后，影儿集团逐步将平台能力释放到各品牌、各大区域分公司以及线上销售渠道，随之而来的是数据处理需求的激增。CDP 平台提供的数据权限功能、标签可视化配置功能很好应对了该问题，不仅能够同时满足多业务部门的数据需求，还实现了业务人员自主完成数据标签加工的灵活性，既提升工作效率也降低信息部门的沟通成本。

除了运用平台完成老客运营之外，TalkingData 基于自身的数据产品帮助影儿集团进行用户画像、指导门店用户洞察以及选址策略，辅助影儿集团搭建起潜客数字化运营的基础。

有了数据平台的支持，2018 年的双 11 活动中，影儿集团线上商城抛弃过往基于历史经验的“人海战术”式营销，通过 CDP 平台进行客群洞察及人群拆分，针对不同人群设置不同的主推权益、触达渠道、测试文案，再分群分组的进行用户触达，以营销活动全链路数字化促进消费转化，最终实现销售额提升超过 90%、客单价提升超过 20%、新客销售额提升 2 倍的亮眼成绩。

第三阶段：复购模型赋能终端导购，大幅提升运营效率

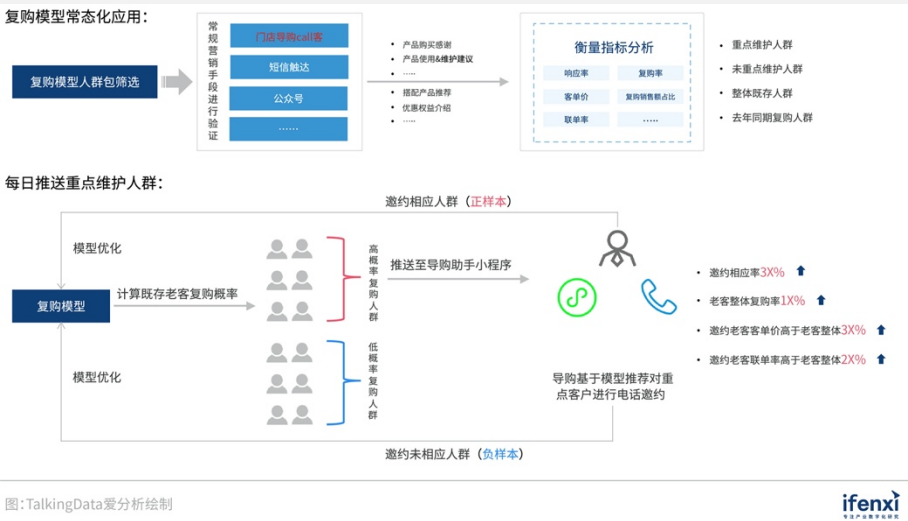
随着项目逐步深入，从集团到品牌、渠道，平台能力逐级下沉，将支持的业务领域推进至终端导购层面。为此，影儿集团将导购助手平台与 CDP 平台对接，使导购助手数据能够回流至 CDP 平台，实现平台赋能导购的运营闭环。

为了满足实际业务场景下的需求，平台开发了大量定制化数据分析报表，用于追踪各项运营活动的效果，再叠加用户复购模型，为导购排序会员维护的优先级，指导导购判断高概率复购人群，既提升了导购侧的用户运营效率和终端交易转化，又能同时追踪导购的业绩完成情况，实现导购管理的智能化。

通过第三阶段的项目建设，导购加复购模型效果的常态化运营已经深入到每个导购的日常工作中，真正实现了用数据来指导业务运营决策，用智能来提升效率和收入，也大幅提升了客户满意度。

通过老客复购模型，影儿集团改变过往完全基于经验判断的运营方式，大大提高了老客运营效率，取得了电话邀约率提升超过 30%、老客整体复购率提升超过 10%、定向邀约老客客单价高于普通老客 30%以上、定向邀约老客联单率高于普通老客 20%以上的效果，真正实现了从集团到终端的 RGM 数字化转型。

图 23：影儿集团数字化运营深入终端导购



未来，影儿集团将在流量、运营、数据、平台能力方面继续拓展生态合作，整合多方资源为业务持续赋能。

2.2.2 应用场景规划

数据中台的价值最终需要通过在业务场景的数据应用来体现，因此，必须应用场景规划先行。不同行业、不同企业处在不同阶段所需要的数据中台是不同的，企业基于明确的战略和业务目标建设数据中台，进而也有可预期的业务场景落地。

数据中台的应用场景规划需要考虑以下关键要素：

首先，企业需要评估业务需求数据与企业数据现状。从具体的业务需求场景梳理相关的业务线、相关岗位和业务流程，梳理其中的业务需求；同时，对企业数据资产进行评估，理清企业已有哪些数据、需要补充哪些数据等。

其次，企业需要明确数据中台的应用场景实现顺序，即哪些优先级和紧急度比较高的场景可以应用数据中台解决业务问题。企业需要建立数据智能应用场景实现顺序的评估模型，可基于企业战略与业务目标、可实现的业务价值、数据应用的实现成本、数据应用的可行性等方面进行评估。

最后，企业在建设数据中台之前应当尽可能多地进行相关调研，并参考同行实践案例，总结相关经验，同时，可借助有成熟经验的数据中台建设厂商帮助企业解决相关问题。

基于数据中台，某航空公司重塑航空燃油分析体系

民航机组的“直接运行成本”（DOC）包主要由固定成本、燃油成本、时间成本三个部分构成，其中，燃油成本约占总成本的 30-40%。2019 年，民航局公布咨询通告《航空承运人燃油政策优化与实施指南》（CCAR-121-R5），对不可预期燃油政策进行优化调整，航空公司可以申请 5%及以下的不可预期燃油政策。从 10%的不可预期燃油降低到 5%及以下，可以避免多带油造成油耗油的成本浪费，极大地降低公司运营成本。

随着油价的持续攀升和航空业竞争的不断加剧，某航空公司迫切需要通在保障飞行安全的前提下，基于大数据进行运行分析和数据挖掘，从而降低燃油消耗、节省运行成本，以达到符合 R5 最低燃油标准的精细化运行管理的目的，提升航司综合竞争力。

QAR 是航司最重要的数据资产之一，完整记录了飞行过程的重要参数，数据质量高，可满足为公司提供高效运营决策支持的要求，从而提升公司成本控制水平、管理能力和竞争力。

因此，依据 QAR 再加上传统运行的 CARS、A-CDM、CDM 等关键运行核心生产数据建立统一数据共享服务平台，实现运行成本的多维度管理，从而达到节能减排和降低公司运行成本，已成为公司应对挑战和把握机遇的一个重要手段。

该航空公司传统的数据分析体系存在多个痛点：

第一，该航空公司原有数仓是基于 Oracle 构建，随着业务发展，数据量增多，数据种类不断丰富，数据融合的难度加大，运行分析的维度和精细化场景不断复杂深化，原有数仓无法有效支撑性能和灵活性，运行分析和燃油数据难以发挥应有价值。同时指标口径不一、数据资产缺乏管理、数据服务化水平不足等问题严重制约着该航空公司的数字化转型。

第二，飞机大多数传感器的解码和接入标准都掌握在国外飞机或发动机制造商手中。因此，民航公司多与其开展 QAR 数据分析合作，不仅成本高，而且只能提供部分场景的数据分析且不支持场景拓展，无法满足航空公司日益增长的需求。

第三，公司过往的指标分析，主要是通过数据分析团队进行手工跑数，但是业务人员无法进行数据分析，造成数据分析人员工作量过大，而业务人员无法灵活地分析数据。

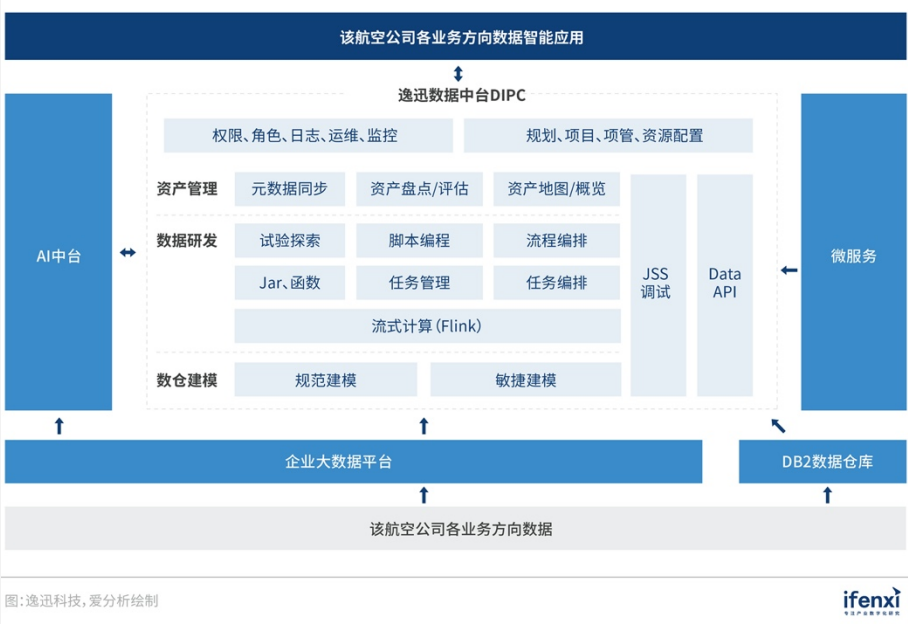
基于此，该航空公司联合逸迅科技共同研发针对该公司的数据中台，并打造指标管理平台。

上海逸迅信息科技有限公司是提供数据治理和数据中台的全栈方案供应商，依托自主研发的数据治理、数据模型和机器学习模型开发工具等产品，提供采集、存储、分析和挖掘大数据的高效数据平台和服务，为各行业客户提供定制化解决方案，服务领域覆盖智慧城市、智慧航空、智慧物流、智能制造等行业，帮助用户快速实现大数据应用价值。

该航空公司分两期项目推进方案，一期基于该航空公司的业务系统打造数据中台，二期基于数据中台进行 R5 燃油分析，并构建了指标管理平台，完善指标管理体系。

逸迅科技自主研发的数据中台产品 DIPC，是数仓构建、指标体系构建及数据研发的一站式集成环境（开发、部署），涵盖数仓可视化建模、数据研发（实时和离线处理）、数据服务等数据处理和服务全流程，实现智能的数据开发和数仓建模等流程。

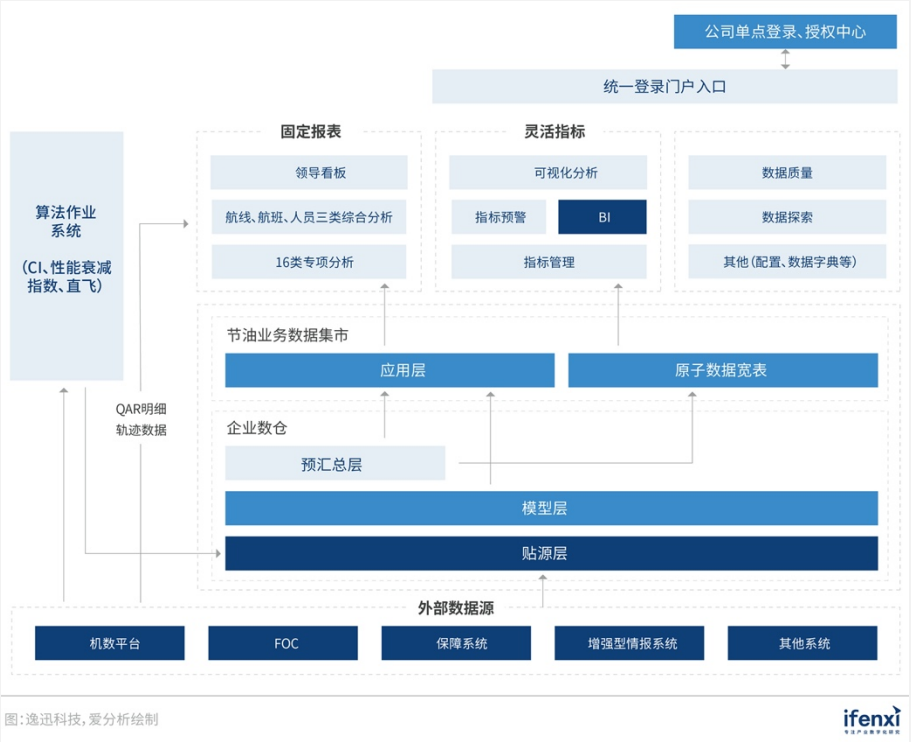
图 24：数据中台架构



在一期项目中，基于逸迅数据中台产品 DIPC，该航空公司通过规范化的建模体系和统一集成的数据开发能力，连接数据后台与业务前台，有效解决前端业务场景对数据的灵活需求与后端数据开发较慢的能力不匹配问题，提升开发和运行效率；建立统一指标/标签体系，构建该航空公司数据资产目录，实现该航空公司数据资产化，强化数据资产管控能力；引入微服务架构建设数据服务体系，实现数据资产化管理和服务化开放，高效赋能前端业务场景需求，帮助该航空公司提升数据资产运营的能力。

数据中台成为了该航空公司一站式数据研发、数据服务和数据资产管控平台，提升企业内部的数据研发效率，强化数据研发流程和数据资产管控能力，为数字化转型提供关键支撑。

图 25：运行分析系统



在二期项目中，基于该航空公司业务和高品质的数据成果，公司通过 Hadoop 大数据平台重构该航空公司数仓体系，并基于 XData 和 XGov 对数据进行迁移、汇聚和数据治理。基于 Hadoop 的大数据平台计算性能提高将近百倍，同时灵活的扩展性可支撑该航空公司未来数年内不断增长的数据体量和业务需求，从根本上解决了原有数仓的性能问题。

基于高品质的数据成果，该航空公司完成 R5 不可预期燃油分析、最优进近轨迹分析等功能，同时优化 R5 燃油分析算法，丰富和完善燃油分析体系，辅助业务部门进行更加精确的节油提效，通过指标管理与自主分析的模块构建，赋能业务部门更加灵活自主的分析能力。

该航空公司还搭建丰富的指标体系，逐步增加和飞行器相关的传感数据分析，例如飞行员模拟机训练、维修物料的流程监控、维修数据统计分析，为该航空公司构建了数百个运行分析指标和数十张分析报表，数据查询的性能响应时间由小时级别提升为分钟级别甚至秒级别，同时赋能业务人员更加准确地分析，作为燃油加注、驾驶行为等飞机运行的支撑。除了丰富指标体系以外，该航空公司还构建了指标管理平台，使业务人员能够自主进行指标的自由组合，指标的灵活分析，释放数据分析人员的工作量，实现精细化分析，同时提升业务人员 KPI 考核全面性和精细度，进而提高业务人员的主观能动性。

该航空公司利用燃油数据整合成果，推广节油分析功能到集团，为其带来长远价值和效益。

该航空公司在项目过程中的经验积累对于同类企业十分具有借鉴意义。企业建设数据中台之前首先需要有明确的应用场景规划，并选取部分应用场景同步推进数据中台建设和上层应用开发，打造业务场景闭环，率先实现部分业务价值。基于此，不断发掘新的业务场景和系统需求，获取企业各部门的持续投入和支持，完善数据中台体系和应用场景的建设。

2.2.3 数据治理体系搭建

数据治理是数据资产建设的基础，目前行业实践中关注较多的是数据汇聚之后的数据治理，聚焦于数据模型与标准开发，但实际上很多数据在汇聚到数据中台之前的业务系统中就已经存在质量问题。如果数据在源端出现问题，即使在数据中台中进行治理，也无法保障最终数据质量。

数据治理的核心在于数据全生命周期管理，企业需要在数据采集之前在源端业务系统中即关注数据治理，覆盖数据中台与数据中台之外的数据资产。因此，要获得高质量的数据资产，企业需要关注源端系统的数据治理。

数据治理应当结合业务场景进行，明确区分数据治理是面向源系统还是数据中台。通过数据资产盘点、数据质量管理、数据标准管理、数据探索分析、数据安全管控，以及源端数据质量进行数据治理，要保证核心业务之间没有断点，提升数据质量，实现数据安全可用。同时，企业需要同时通过技术与制度搭建企业的数据治理体系，基于制度保障员工按规范的数据治理方法工作，基于技术提升数据治理的工作效率。

源端数据治理是根本，某电信运营商的数据治理实践

面对快速增长的数据，某电信运营商省公司从战略和业务层面都亟需提升数据管理能力，建立从上到下的数据治理体系。

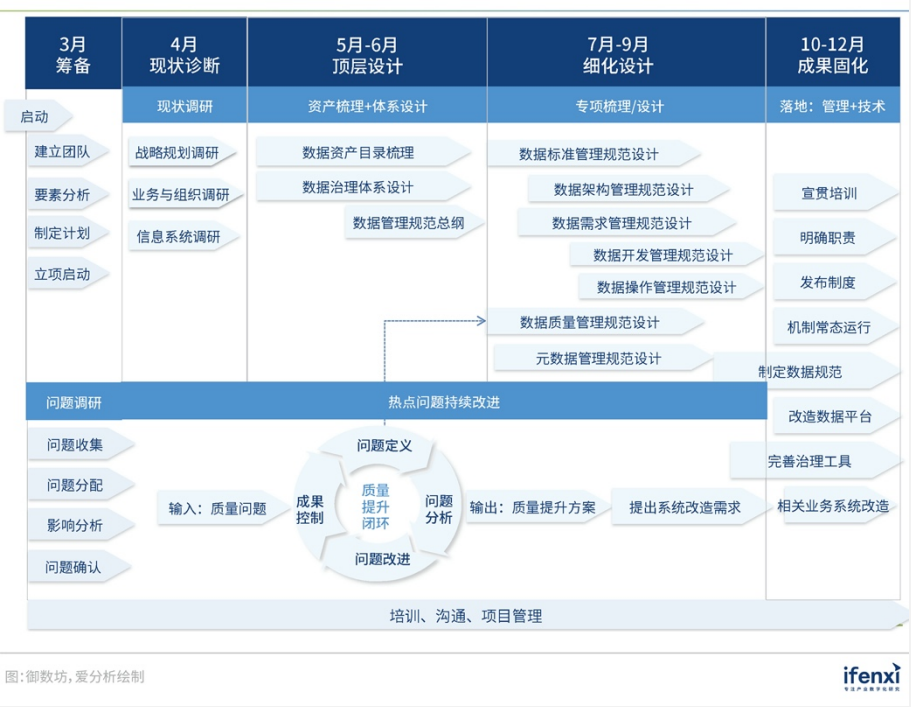
从战略层面，集团总部对省级大数据平台的数据治理能力提出了更高的要求：数据治理需要做到可视化、高效化、可靠化。同时，该分公司大数据发展规划提出构建全新的数字化创新战略体系。从战略规划来看，高价值数据整合及共享的要求促使公司必须提升企业数据管理能力。

在业务层面，数据质量问题也影响着业务开展。一方面，该分公司已发展自有渠道 3 百余家、社会渠道 3 万余家，但是各系统间分类体系不统一、重点业务数据不一致、渠道酬金结算不准确不及时等问题，影响了渠道的健康发展，造成了业务部门对信息系统数据质量的投诉与质疑；另一方面，市场部每天有上百个指标，指标按月对比波动，管理层无法知晓指标波动大背后的原因。

因此，该分公司启动了数据治理项目，与御数坊共同合作，组成数据治理项目组，结合其自身战略规划和具体业务问题，制定公司数据治理的方案规划。

御数坊是专业的数据治理培训、咨询及软件服务商，以“咨询服务+软件产品”的方案为企业提供全生命周期的数据治理解决方案，帮助客户建立数据治理能力、建立数据资产目录、对多源异构数据进行标准化，提高数据中台的数据质量，创造数据资产价值。

图 26：企业级数据治理项目建设方案及工作计划



图：御数坊, 爱分析绘制

ifenxi

项目组采用顶层设计与热点业务问题结合的建设方案，分 5 个关键阶段推动项目。顶层设计涉及资产梳理和体系设计，业务问题解决方案主要结合大数据平台解决具体业务问题。

整体项目主要分为数据管理现状分析、数据质量根因分析、专项数据治理、数据治理体系建设等阶段，短期解决公司的业务实际问题，长期提升企业的数据管理能力。

（1）数据管理现状分析

在现状分析阶段，分别进行了顶层设计和具体业务问题的现状分析调研。项目组基于业务发展重要性及数据质量提升紧急性，优先选择家庭宽带和渠道两个模块，针对酬金结算不准和订单数据不准等场景作为切入点，对相关的 5 个系统累计进行 13 余次深入调研。

其中，流程梳理主要分为数据流和业务流梳理。在数据流梳理过程中，重点分析接口数据、基础数据现状，进行自下而上数据流分析。例如，分析渠道收益阶段的指标在数仓是如何加工的、由哪些数据加工、基础数据的分布情况、各部门在哪些系统修改创建数据等；在业务流梳理过程中，项目组关注数据承载业务及流程前后衔接、上下对应关系，识别各类数据管理问题及风险。

基于 DAMA 数据管理体系，项目组确定工号类、渠道类和组织类数据是主数据，并提出 19 项初步数据治理优化建议。同时，针对短期问题解决要求，设计包括组织保障、数据规范、制度流程等方面的数据质量问题管理机制。

（2）数据质量根因分析

项目组选取发卡数据不一致、家宽酬金数据不准确、酬金 BI 接口问题、4A 基础数据准确性、问题管理机制不完善等 5 类与核心业务发展、与内外部客户满意度密切相关的问题，将其纳入根因分析范围，确定质量问题根本原因，并制定相应改进计划。

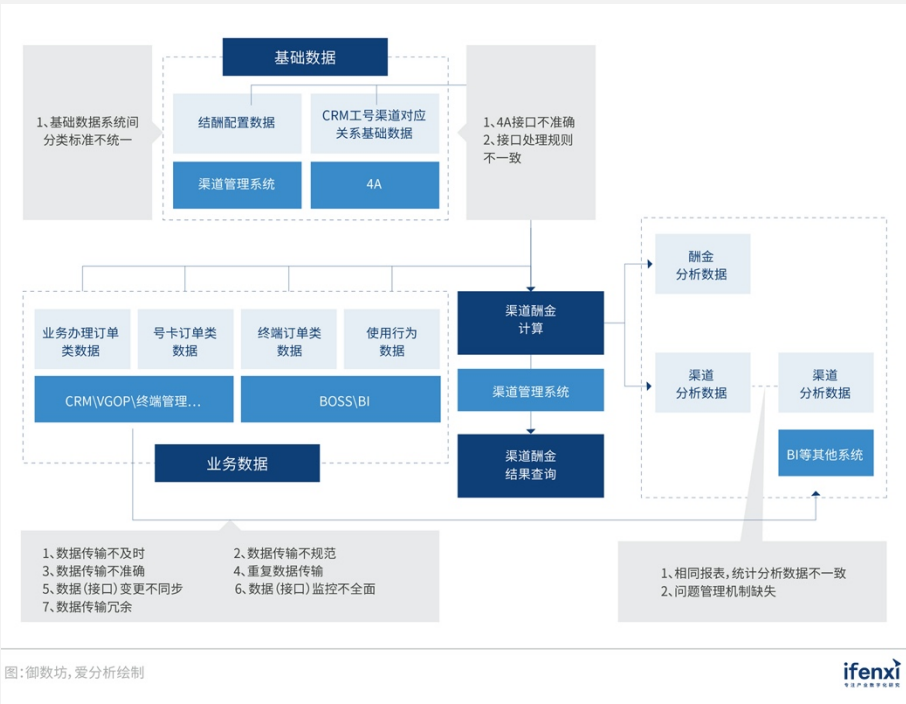
（3）专项数据治理

数据治理并非仅仅清洗数据并存入数仓，而是要结合业务场景进行深度治理，在源端改变业务人员的职责、业务流程和业务系统自身的数据规范和接口，同时改变后台数据仓库的规则、标准和模型。项目组选取出现频繁、与项目紧密相关的基础数据，设计“基础数据数据治理方法论”模型，设计数据标准和数据质量规则，进行基础数据专项治理，提升数据质量。

在数据标准管理方面，项目组制定术语标准和与基础数据相关的参考数据与主数据标准；在数据质量管理方面，根据基础数据数据价值链分析，制定系统间数据校验规则，以加强数据质量管理。

专项数据治理工作由该分公司信息系统部大数据支撑中心牵头，会同市场部相互协作，历时 2 个月，保证了基础数据优化及改进工作。同时企业级数据治理项目组协同各系统厂商、市场部业务人员详细分析了基础数据在重点系统间的信息价值链，绘制了渠道酬金基础数据信息价值链全视图，有效分析了数据系统间使用冲突，并根据信息价值链分析原理，给出了科学的数据源权威系统建议。

图 27： 基础数据在渠道酬金结算数据流中关键问题分析



专项数据治理完成了 3 大类基础数据数据治理，并由市场部牵头面向 4A 系统、渠道管理系统集中下发改造需求。

(4) 数据治理体系建设

数据治理的需求是全局的，来自于企业管理层和业务部门，需要从管理层到业务层均采取行动，并优化其组织架构和业务体系。数据治理体系分为数据治理组织体系与数据治理能力体系。

项目组首先结合具体业务问题，优化公司组织架构，并动态地调整和增加组织成员。数据治理体系建设由公司高层领导挂帅，业务与信息部门协同的数据治理组织体系，包括数据治理委员会，总体协调组，各业务部门、信息系统部、各分公司和各项目组。其中数据治理委员会是最高议事机构，负责全面协调、指导和推动公司的数据治理工作。

图 28： 企业级数据治理能力体系框架



长期来看，该分公司建立 1 个总纲、7 个管理分册、1 套数据规范的“1+7+1”数据治理体系，规定数据治理目的、范围、原则、组织职责及工作内容、工作评估方法。基于此，该通信运营分公司从顶层设计和业务运营方面实现了数据治理效果提升。

在顶层设计层面，公司建立了数据治理组织架构，提升数据管理能力，编制数据治理体系制度，为数据治理运营管控提供基础保障。建立了长效机制，指导后续工作常态开展。

在业务层面，公司实现了三方面的效果。

第一，改进底层数据质量，通过 DAIC 数据质量问题管理机制，开放业务部门反馈数据质量问题渠道，有效分析并解决了 17 类市场部关注数据质量问题，完成了发卡数据不一致、家宽酬金数据不准确等 5 类热点问题的根因分析及改进方案制定。

第二，消除了业务流程中的断点，改善业务运营。

第三，提升业务绩效。项目组实现了数据治理体系创造收益的量化分析。例如，可基于公司原本的每月错误业务归属订单评估数据治理为业务带来的价值，对该项目进行业务 ROI 测算。数据统计的指标长期稳定，波动减小，和营业厅的数据核对差异很小。

该数据治理项目成功的核心在于和市场部门的业务场景紧密结合，率先实现了部分业务收益，获取管理层和市场部分的进一步投入与支持，进而能够持续数据治理体系的建设。

2.2.4 技术平台建设

根据前述的数据中台架构，数据中台的技术平台搭建重点包括数据开发、数据资产管理、数据服务、数据安全管理体系等。

目前，数据中台的技术平台搭建涉及到的技术已经较为成熟，且大量开源技术降低了开发门槛。此外，也有大量兼具技术产品能力与实施落地经验的厂商，可以帮助企业进行技术平台搭建。企业可以根据自身需求和规划，选择自研或与厂商合作开发。

基于成熟的开源组件，一方面，企业系统可以更加稳定，且平台更加工程化，可支撑企业大规模应用，中台可根据业务需求灵活拓展；另一方面，企业可以大幅降低建设数据中台的成本，提升开发效率，让企业内部开发人员更加专注解决业务问题。

2.2.5 业务价值实现

数据中台建设属于底层基础设施，往往难以直接产生效益，加之涉及较多横向模块，建设周期长、投入大，如何衡量投入产出是一个问题。

围绕企业的应用场景规划，企业需要明确应用场景实现的优先级，率先实现部分场景应用，打造业务场景的闭环，实现业务收益。

一方面，基于数据中台首先实现业务能够量化 ROI，在短期内即可实现部分收益，企业管理层和业务部门看到实际的价值，有利于持续投入并配合推进数据中台的后续建设与优化。

另一方面，部分应用场景的落地是对数据中台的有效性验证，在过程中更容易发现数据中台建设和运营方面的问题，可以在后续的拓展中持续优化，避免前期建设底层基础设施的成本浪费。

2.2.6 持续运营

数据中台完成建设之后，需要保证数据资产在持续稳定地赋能业务，以实现数据资产价值的最大化。实现部分业务场景的闭环和业务价值之后，企业逐步将单个场景的数据智能应用延伸至多个场景，因此，企业需要进行数据中台运营与优化，将数据资产和应用流程固化，拓展更多的场景，最大程度地发挥数据中台的价值。

数据中台的持续运营是基于数据智能的相关技术体系，构建一套完整的、体系化的数据处理及服务流程，源源不断地把数据变为资产并服务于业务的一种可持续让企业数据用起来机制。

数据中台的运营需要从多方面进行：

(1) 数据运营规范

建立数据中台之后，首先需要保障企业体系化、标准化的数据采集、存储和打通，搭建企业的数据资产；其次，在数据质量方面，需要建立体系化的数据汇聚、加工与应用流程，通过运营手段完善数据管理制度与规范，保障数据资产的高效输出与循环落地机制。

(2) 组织协同

数据中台是企业数字化转型的核心支撑，业务模式不同于传统模式，应当根据数据流、业务流对组织架构进行调整；同时，企业开始在管理层设立数据管理委员会等岗位负责数据机制的建设与管理，为数据资产的建设提供支持。另外，企业成立专门的数据资产运营部门，保障数据资产的应用，同时维护平台中数据的真实、稳定、准确和高效。

(3) 数据应用场景拓展

数据中台可对外提供统一、标准化的数据服务能力，企业可以在此基础上探索各类业务应用，并利用数据中台提供数据服务支撑，包括原有业务的优化与业务创新。

(4) 数据运营能力建设

企业应当建立完整的数据运营体系，包括数据资产质量评估、数据资产安全管理、数据成本运营、数据持续运营等。同时，数据运营能力还包括企业的组织调整与制度规范，保障企业全员具备数据意识，调动部门积极性持续优化中台机制，保障数据资产的持续运营。

2.3 各行业数据中台实践案例

数据智能在各行业逐步渗透，不同行业的数据基础设施、业务需求均不同，企业遇到的数据应用问题与相应的解决方案也有很大的差异。近几年来，数据中台已经开始在各个行业落地，一些头部企业经过不断探索，已经在各个垂直行业形成了成熟的解决方案，探索出了一些成功的实践经验。

本章节挑选了部分重点行业的典型数据中台建设及应用案例，为同行业企业提供相应的借鉴。

2.3.1 金融

基于数据中台，数据智能驱动中型证券公司数字化转型

某中型证券公司是国内第一批拿到从事证券投资业务牌照的证券公司。随着业务发展，面临着数据开发与应用方面的挑战。原有的大数据平台能够解决大部分数据应用问题，但由于业务变化快，系统上线的速度也越来越快，基于原有的大数据平台的服务支持存在诸多问题，不仅导致企业资源浪费，还对信息部门的开发工作造成不便。

第一，具有多种并存的数据库，且业务系统大多是烟囱式的，所以数据存储分散，交叉使用非常不便；若上线新系统可能造成重复建设，资源浪费。

第二，数据处理缺乏统一的规范，技术人员的处理方式不同，技术能力难以在团队内复用，也难以拓展至其他项目。

第三，单点开发问题严重，企业文档记录不完善，数据知识共享不便，导致重复造轮子。

第四，任务调度难。随着开发任务量增加，如何安排新任务执行才能保证系统不崩溃、不锁表，是开发人员的一大挑战。

为了能够建立可科学管理、可便捷开发的数据服务，该证券公司基于原有的大数据平台升级开发数据服务中心，打造自身的数据中台，为前台业务提供快速强力的能力支持，同时实现数据资源复用，业务能力沉淀。

该证券公司采用自研+外购的方式建设数据中台，在原有的大数据中台的基础上进行升级。数据中台的建设原则是可复用、标准化、安全可靠、能力开放、应用简单及智能化。中台整体架构分为大数据基础平台、大数据治理平台、数据服务中心三部分。其建设核心在于两方面，一是数据资产，二是数据智能。

图 29：数据中台架构



大数据基础平台是数据的存储计算中心，数据来源于柜台系统、业务系统、子公司的数据库等多渠道，基础平台建立了统一的数据接入、存储和计算方案。基础平台具有数据工作的统一入口，可以避免开发人员单点开发的问题，以实现工作成果积累和知识共享。

在大数据基础平台中，采用了星环科技的分布式技术，主要技术包括流处理、批处理及 AI 集群。离线计算模块采用 Transwarp Inceptor，是可实现完整 SQL 支持的分析型分布式批处理数据仓库，用于构建数据湖、大规模离线批处理系统、企业级数据仓库。在线计算采用 Transwarp Slipstream SQL 编程规则引擎，是在事件驱动计算引擎上支持批处理等复杂编程模型的流处理引擎。算法平台基于数据挖掘平台 Transwarp Sophon Discover 打造，能够进行灵活的数据挖掘分析探索，包含丰富的分布式算法库，提供 R、Python、Spark 的集成开发环境。Transwarp Sophon Discover 还整合 Notebook 工具实现可视化的模型训练和数据分析，支持团队协作，内置多种行业模板和基础架构辅助用户构建智能应用，如精准营销、流失预警、文本挖掘等。三者共同构成统一计算模块，为该券商数据中台提供强大的算力。

此外，统一存储层使用了星环科技的操作型数据库 Transwarp Hyperbase，属于 NoSQL 操作型数据库，可构建海量数据的在线存储和查询系统，支持高并发实时数据入库场景、索引和高并发的数据查询业务、非结构化数据的存储分析。

同时，星环的标准产品进行了容器化，促进了大数据基础平台的环境双核部署和安全认证。基于星环的产品，极大节省了探索平台底层的性能的时间，星环已经将本地环境配置好，使得整个系统能够高效稳定地运行。

数据治理平台是该证券公司的自研平台，数据经过数据治理即转换为数据资产，形成了客户、用户、产品、企业、资讯等数据集市。不同于其他企业自上而下的数据治理原则，其采用向下而上的原则驱动集团数据治理，在进行数据治理的同时兼顾业务特殊性，这种方式更适合以技术趋动数据治理的中小券商。

基于数据资产，数据中台将对外的数据服务层包装成数据服务中心，以微服务的方式构建，通过多种渠道支持前端系统，同时对前端系统开放了 API 服务、产品化服务以及数据库的服务。数据服务中心提供了丰富的数据接口和数据字典，不管是业务人员还是开发人员，都能够比较顺利接入到底层的数据。

数据智能不仅体现在前端的业务应用支撑，也体现在对数据中台的支撑。开发人员接到业务需求之后，平台会进行灵活的数据字典配置，开发人员只需要配置资源数据字典，任务即自动运行，实现开发运营化。这不仅降低专业人员的开发门槛，还使得平台的底层代码池健壮，且系统非常稳定。再基于任务智能调动系统调整任务，最后形成应用，对接业务系统或赋能前端客户。

基于强劲数据中台底座，赋能多元化智能应用

在应用层面，基于数据中台，为客户赋能、业务赋能、管理赋能提供支持。

在客户赋能方面，利用大数据和智能算法，打造专业智能客户投资工具，例如，在交易服务中结合客户投资的整个过程，打造交易前、中、后的投资服务；在零售业务线中基于用户画像对线上零售客户实现精准营销、精准精细化运营服务。

在业务赋能方面，通过报表和各种分析体系等数据可视化平台，帮业务人员更高效地拓展业务，提升业务效率。例如，提供了业务和领导人员的数据获取能力，目前已经形成了领导驾驶舱、业务报表、绩效考核、监管报表和用于自身的 IT 平台运营监控。

在管理赋能方面，基于数据中台进将风险数据监控和信用风险各个系统进行有机融合，搭建风险模型，支持企业风控、合规和审计等。

项目进行过程中，与星环合作，直接应用其模块化产品，降低了数据中台底层技术平台的成本投入，使其能够在自己专长的领域里有针对性投入资源，包括业务服务、数据分析和数据挖掘技术等，同时，提升了开发效率。

目前，该数据中台基本上已经具备完整的功能。未来，将会扩充外部数据，例如舆情、资讯等数据，并基于知识图谱丰富公司业务中台中的风控以及商机发现部分。同时，会更加侧重于非结构化数据处理，例如，投行业务中的合同管理、底稿管理以及业务流程，需要大量图像识别。

中小证券数据中台建设经验借鉴

在项目过程中，该中型证券公司的经验积累可为其他中小券商数据中台建设借鉴。

第一，数据中台建设须围绕客户以及业务开展，一方面更好地解决业务需求，另一方面获取业务部门的支持与配合。

第二，需要选择正确的技术方向。首先选择主流、开放的技术，在 POC 阶段，要让技术人员实际尝试使用；其次，选择的技术要有先进性，能够提供很多便捷的开发方法；最后，需要具备可持续发展性和自主承接的能力。

第三，注重中台建设的节奏，本着“有用才做”的原则，需要有一些业务场景率先落地，在这个过程中逐步将中台建设起来。

2.3.2 政府与公共服务

某市应急管理局构建统一大数据平台，实现智慧应急

应急管理涉及众多场景，传统管理模式下管理职责较为分散，难以制定整体的应急规划。因此 2018 年 3 月，全国人大通过组织机构改革办法提出了大应急体系，把安监、消防、水利等部门都融合到了应急管理部门，从国家层面作出整体规划，不再局限于某一领域，而是力求全方位的安全。

部门整合之后，应急管理仍然面临多个场景数据标准不统一且业务系统碎片化的问题。某市应急管理局面临着以下挑战：

- 1) 应急管理对象涉及地震灾害、地质灾害、森林火灾、草原火灾、火灾事故等，数据资源繁杂，标准不一致，缺乏统一治理和管理，数据孤岛严重。
- 2) 虽然全市积累了大量安全生产、自然灾害和综合防范等数据资源，但由于数据多头管理，权责不明确，数据挖掘与分析不足，数据价值和效能未能最大化释放。
- 3) 数据应用水平低，无法高效支撑对全市运行状态的全面感知、态势预测、事件预警和决策辅助等业务应用。

为了解决上述问题，该市应急管理局提出构建一个统一的大数据平台，并支撑上层数据应用。该市应急管理局与百分点合作，构建智慧应急统一大数据平台。

百分点作为一家数据智能技术企业，在应急管理领域，按照“数用分离、智能驱动”的思路，依托大数据全栈技术和产品，构建符合应急管理需求的应急大数据治理体系，实现数据接入、处理、存储、应用等全生命周期的管理。

图 30：应急管理局统一大数据平台



市应急管理局首先通过构建智慧应急统一大数据平台，消除了数据标准不一致，实现了全市应急大数据全方位获取、全网络汇聚、全维度整合的整体信息资源规划。其次，在大数据平台之上通过数据融合和挖掘分析，提高了各种数据资源的利用率。同时，通过建模分析预测风险，实现全域感知，将数据资源转换为应急决策依据，提升了应急管理决策水平，有效支撑了全市及各区智慧应急应用体系建设。

1) 建设智慧应急统一大数据平台

为了建设应急数据治理体系，市应急管理局率先构建了城市大数据中心应急专项数据资源体系，各类业务应用得到了统一的数据支撑；通过大数据平台的建设成果获取数据，做到了“数用分离”，避免数据重复建设，提升数据使用效率；同时，将数据治理与应急业务结合起来，真正实现了业务系统数据之间的交互、共享、互补、协同。

市应急管理局构建的智慧应急大数据平台，统一接入局内 10 个业务系统数据、21 家“安委办”单位应急数据、市域物联网感知数据和互联网数据的信息资源，实现全市应急基础信息汇聚、治理、主题库构建及信息展示。针对安全生产、自然灾害、城市生命线等多个领域的 591 类、1.6 亿条应急数据进行了勘探、稽核、质量分析和提升，实现了应急数据的标准化和全生命周期化管理。

在项目落地的过程中，遇到的难点主要有核心数据来源以及数据质量问题。应急数据有不同的数据层级，获取数据的来源主要是三方面：一是应急业务内部的数据，可能是跨层级的；二是可以汇聚外部各个委办局的数据；三是可以向上级申请各部门上报的数据。此外，针对数据质量并不高导致无法使用的问题，百分点与各个业务单位持续沟通，提升数据质量的标准要求。

2) 数据融合、数据挖掘分析及可视化展现

在完成大数据平台建设后，市应急管理局对全市应急全要素数据进行融合和关联分析，能够全面实现灾前风险分级、灾害事故识别和预警、综合研判、预案关联匹配、灾害趋势自动分析、救援方案推荐、智能辅助决策、灾后回溯评估等业务。针对全市安全生产、三防灾害、城市生命线等特定业务场景，实施风险监测预警和主动防范。在可视化方面，市应急管理局基于 GIS 一张图展示的全市应急信息资源，可实现应急数据可视化决策。

3) 构建智慧应急业务模型

市应急管理局还基于大数据平台，将人工智能算法与应急业务深度结合，面向安全生产、防灾减灾、应急救援等业务域构建应急业务模型，实现危化品生产、运输、存储、使用的全生命周期管理，自然灾害及次生灾害风险隐患的动态监测、预测预警等，全面提升应急管理事前、事发、事中、事后四个过程域内风险预测预警能力和联动处置能力。

应急管理局通过对全市应急全要素数据资源的统一融合治理，极大提高了应急数据资源的综合利用率，实现了数据资源和业务应用态势的可视化综合展示，为城市“智慧应急”提供数据基础，释放了应急数据价值和效能。

图 31：应急管理大数据高分可视化系统



智慧应急统一大数据平台上线后，应急管理局达成了以下目标：

- 1) 通过数据资源池的建立，规范和完善了 28 大类应急数据应用标准，实现应急数据标准化和体系化管理，提高了数据质量，促进了各系统之间应急信息资源的开放共享和有效应用。
- 2) 在不同来源数据融合治理形成的信息资源库基础之上，构建了全域实时风险地图，汇聚了全域各种应急数据，全面提升了风险管控和监测预警能力。
- 3) 通过算法模型，叠加多源数据分析预测隐患，实现灾害链的关联分析和全域感知，推动实现“事前预警、事中可控、事后联动、可追溯可评估”的应急全流程管理，从而将被动的应急管理变成主动的应急“智”理，有力提升了科学预防、安全监控、快速反应、智能研判、有效决策的水平。

2.3.3 消费品与零售

数据中台赋能传统体育零售商，应对数字化转型挑战

滔搏体育（以下简称滔搏）是中国最大的体育零售商，作为耐克、阿迪达斯最大的分销商，滔搏共代理 11 家国际运动品牌。同时作为百丽国际旗下面向体育业务的子公司，滔搏从上世纪 90 年代涉足体育运动产品的经营业务。滔搏将自己定义成以消费者为核心的运动零售及服务平台，从一家传统的经销代理向运动服务提供商、体育活动组织方以及运动社交分享平台三大方向转型。打通原来分散割裂的门店、渠道、消费者、会员等数据，也成为了实现转型不能越过的先决条件，这也对这家拥有近 30 年历史的传统零售代理商的数字化建设提出了新的挑战。

作为一家业务模式相对传统、历史积淀厚重的零售商，在滔搏业务扩张发展过程中，公司内部信息系统大量建设且缺乏统一管理，各类数据标准化程度低、对接难，组织间信息壁垒和数据孤岛逐步形成，加之数据统计口径不一致，长期存在业财差异无法解决。

同时，滔搏原有的数仓存在性能瓶颈，数据指标计算时间长，无法满足新业务场景的需求，且原有的数仓架构扩展性差、维护成本高。

在拆分上市后，滔搏将数字化转型、线上线下消费者互动的深入整合、基于客户需求的供应链升级作为其业务战略的重点方向。为满足战略规划的实现，建立一套高效、全面、业务导向、敏捷响应的数据中台是滔搏迫在眉睫的现实需求。

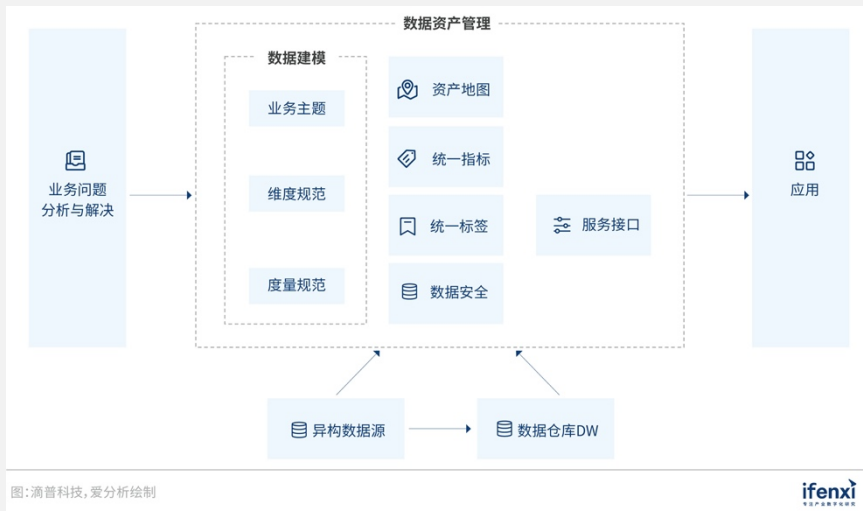
滔搏数据中台搭建模式

在数据中台搭建中，滔搏选择了滴普科技作为合作伙伴。滴普科技是一家全场景数据智能服务商，有深厚的技术和平台建设经验积累，为包括零售、快消、汽车、3C、工业等在内的众多不同行业的头部客户提供了标杆性的数字化转型服务。

滴普科技为滔搏设计并建构了一整套以互联网中台技术架构为蓝本的数据中台，打通了滔搏包括门店、财务、CRM、OA、HR、会员、运输管理等在内的 23 个业务系统，通过数据资产化、资产服务化、服务业务化，实现数据赋能零售业务目标。

在中台的落地层面，在综合考虑业务/财务维度的数据标准和数据模型基础上进行平台建构，同时统一滔搏各业务线、前后端的数据采集、数据处理、数据治理、模型管理、数据服务，从源头到应用端到端提升标准化水平。

图 32：滴普的数据中台建设架构



得益于互联网架构，滔搏数据中台的架构系统自主可控、高可用且易扩展，提供例如高性能分析能力、面向业务的服务管理等支撑快速迭代的能力。在确保数据准确易用的同时，满足了对业务敏捷响应的数字化平台建设初衷。

同时，新的数据中台拥有健全的安全保护机制，提供多租户管理、角色权限管理、认证鉴权、数据加密脱敏、安全审计等功能，保证滔搏旗下众多部门仅可以保仅可访问允许的数据，避免数据泄露。

数据中台赋能滔搏数据服务体系

通过建构数据中台，滔搏日常运营分析需要的多个数据源系统得以整合，数据体系设计得以规范，数据差异风险有效降低，赋能业务的综合分析能力迅速提升。滔搏的数据资产管控体系也得以同步建立，通过对现有数据资产的盘点进行数据资产运营，提升数据的有效性、可追溯性，并逐步改善数据准确性问题。

通过对数据中台的升级，滔搏原有的数据仓库处理与响应效率也大幅提升，响应时间从原来的 2 小时缩短到 1 小时左右，增强了对前端业务快速变化的响应能力。

通过数据中台进行全量数据封装透出，支撑各个数据业务，保障数据获取的及时性和稳定高效，并可支持数据能力不断扩展，形成可靠的数据服务体系，为查询服务、分析服务、检索服务、精准营销、个性化推荐、大数据风控等应用场景提供良好的数据支撑。

2.3.4 工业与能源

国网某省级电力公司借助数据中台建设能源互联网

我国电网规模已超过美国，跃居世界首位。但是随着人们用电量需求的日益增加以及现代化进程的不断推进，人们对电力系统的要求越来越高——需要提高运行水平和运行效率、保障可靠性和安全性、减少系统故障率和故障时间。

当前，数字革命与能源革命相融并进，电网形态逐渐向能源互联网转型。国网从总部层面开始统一规划，提出数字化转型和能源互联网的战略目标，要求自上而下搭建数据中台。

国网某省级电力公司，有庞大的业务系统和组织架构体系，无论是数据量还是业务系统数量都是一个庞大的数字，在数据治理的过程遇到以下问题：

1.从数字化现状来看，前期的信息化建设在支撑各类专业业务开展的过程中积累了大量高价值的数据资产，但是系统的相互独立也造成了业务壁垒的存在。

2.数据应用方面也存在瓶颈，主要表现在四个方面：各业务系统数据交互难、企业数据资产综合利用效率低、各系统数据冗余和不一致以及跨系统应用难以获得稳定的数据支撑。

为了解决上述问题，国网某省公司与明略科技合作，开展了数据中台系统的建设，以达到基于中台实现数据资产管理、通过数仓实现各业务域标准业务模型、建成高效业务数据服务体系的目的。

明略科技是一家数据中台和企业智能决策提供商，致力于通过大数据分析挖掘和认知智能技术，推动知识和管理复杂度高的大中型企业进行数字化转型。明略科技的数据中台以多维数据的感知和汇聚能力、基于知识图谱的数据融合打通能力、基于智能数据引擎的数据自服务能力，赋能客户的中台建设，实现客户数据的资产化、智能化和服务化。

国网某省公司前期针对数据治理层面的问题构建了数据中台；中后期的项目建设在数据中台基础上进行了改善，做了一些基于数据中台的应用，利用知识图谱技术同时解决了数据质量、数据关联问题和业务壁垒问题等。

1) 数据中台建设

图 33：明略集团数据资产管理的总体架构



数据中台的建设任务是针对业务需求建立数据模型，形成标准化、高质量的企业数据资产，并对外统一开放服务。

在数据资产管理方面，国网省公司从数据汇聚、数据加工、数据管理、数据服务整体上规划，以数字仓库建设作为治理重点，基于数据仓库的业务模型建设和管理，在整体上提升数据中台和数据资产管理的数据服务能力。

在数据仓库建设方面，国网省公司结合 SG-CIM 模型与维度建模方法，按主题域和业务活动进行业务模型建设和管理，建立了以资源驱动为特征的近源层、以业务驱动为特征的明细层、以需求驱动为特征的加工层以及以应用驱动为特征的应用层。在这四个层级之上，国网省公司完成了客户、设备、项目等 9 大主题域、590 多个数据模型、300 多个业务模型的搭建，并通过这些模型来支撑各类上层应用的高效应用。至此，数据中台能力逐渐显现。

图 34：按主题域和业务活动进行业务模型建设和管理



2) 构建知识图谱

基于数据中台的典型应用是知识图谱构建，例如设备全景图谱和设备拓扑图谱都是基于数据仓库构建。在知识图谱的支撑下，再赋能数据应用，实现精确计算停电范围等智能应用。

以计划性停电工作为例，国网省公司将电网相关设备的拓扑图以知识图谱的形式进行了存储、表示，定义了实体类型、属性、关系等，根据需求和现有数据，共设计了 15 类实体、16 种关系、68 个属性，通过从不同来源、不同结构的数据中进行知识抽取，形成知识存入知识图谱，服务于诸如停电范围自动计算电力行业实际工作。

在该知识图谱中，国网省公司利用图计算在“图”上通过深度检索、广度检索、最短路径检测以及停电模式相关的数据挖掘，实现了对停电范围的自动检测。工作人员只需输入需要计划性维护、排查的台区，根据设备依赖关系即可自动分析停电范围与受影响的台区，精确计算出停电范围。与传统的人工手工对照 CAD 一次接线图以及范围影响分析的方法相比，大大节省了所花费的时间，同时也提高了分析的准确性。

在精确计算出停电范围后，国网省公司基于语音技术进一步搭建的语音短信下发系统可提前通知台区内的小区居民停电检修时间。

通过整条链路的打通，国网省公司大幅缩短了从停电范围计算到短信下发的时间，实际解决了供电公司配网、调度过程中的重复劳动、效率低下问题，打通了各业务部门的业务系统。

数据中台和知识图谱平台上线后，截止目前，明略科技协助国网省公司一共打通了 136 个业务系统，接入了 6300 多张报表，2800 多亿条数据。

未来，国网省公司会将更多业务系统打通，同时接入更多实时数据，建设实时数据仓库，进行更多数据创新应用探索。

CHAPTER

03

数据智能应用趋势展望

3. 数据智能应用趋势展望

3.1 知识图谱推动非结构化数据价值挖掘

企业以往的数据智能应用大多是基于大数据技术，对结构化数据进行分析和挖掘，在金融、营销等数据基础设施比较完善的行业迅速落地。

然而，企业所拥有的数据中，70%为非结构化数据，包括图片、视频、文档等。未来，5G、物联网等技术的加速落地也会催生更多的非结构化数据。对非结构化数据的价值挖掘是数据智能应用需要解决的问题。

非结构化数据的处理技术门槛更高，需要应用机器学习、知识图谱、NLP 等技术。其中，知识图谱技术可以实现从海量非结构化数据中抽取结构化知识，有效推动非结构化数据的应用。尤其是知识密集型行业中，知识图谱可以将复杂的知识领域及知识体系通过数据挖掘、信息处理、知识计量和图形绘制显示出来，是企业知识运营的重要支撑。

未来，企业将针对非结构化数据搭建和优化底层数据架构，同时，融合知识图谱、机器学习等技术，挖掘非结构化数据的价值赋能企业业务。

筑森设计基于非结构化数据中台赋能知识运营体系

筑森设计是一家位于行业前列的建筑设计公司，总部位于常州，是新城集团的子公司。

对于建筑设计公司来说，知识的运营和传承至关重要。新员工入职后通常需要经过专门培训才能上岗，企业需要知识运营系统帮助新员工自主学习，了解此前的优秀建筑设计是如何完成的，以提高企业内部生产力。

目前筑森设计的知识库类应用是独立的知识门户类系统，比较封闭，且与生产系统脱节。该知识库内的内容来源大多依赖于人工整理、手动上传等方式，造成了知识库运营的压力较大，且传播、推广的知识没有针对性，新员工无法有效学到有实际意义的知识。也有很多企业对图纸进行加工，加上一些文字描述或者文案以提供指导性意见和技术要点，实现知识运营。但这类知识运营不仅工作量大，而且对于负责知识运营的人员的水平和要求非常高。同时，仅仅对图纸进行标注无法完全还原设计全过程。

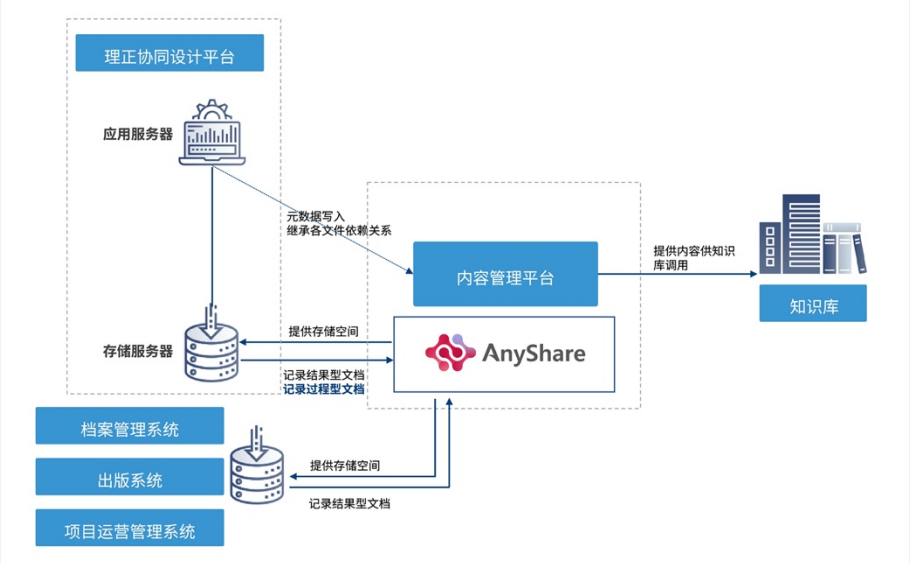
筑森设计的核心业务系统为北京理正协同设计系统，是图元级的二维协同设计系统，能与工作流系统协同，实现图纸的质量控制和追踪。该系统所管理的对象是项目图纸及相关资料，其中的设计文稿、标杆项目的设计过程，都是宝贵的数字资产。如何基于协同设计平台中的数字资产实现知识运营，有针对性、指导性的进行专项知识输出是筑森面临的一大难题。在此背景下，筑森联合爱数建设其知识运营体系，以期还原设计的生产过程，实现高效的知识运营。

上海爱数信息技术股份有限公司是一家大数据基础设施提供商，致力于为政府、公共事业及企业数字化转型赋能。爱数基于 AnyShare Family 7 帮助筑森搭建其知识运营平台。

知识运营系统需要还原整个生产过程，具备思路还原的能力，核心在于两方面。第一是要还原设计关键点，例如需要记录设计师易犯错误的环节、常犯错误的版本更新与迭代等；第二是还原设计过程，需要记录协同设计平台每个元数据的属性和相互的依存关系。

爱数打造知识运营系统，将其与北京理正协同设计平台进行对接。筑森的设计人员每天在协同设计平台上产出大量的设计图纸，这些设计图由大量的 dwg 小文件组成，知识运营系统需要“抓取”这些 dwg 文件的属性及依赖关系，并同步写入 AnyShare Family 7 元数据引擎中，记录它们之间的关系从而记录整个协同设计的过程。例如，平台可显示设计师在项目周期中的某个点犯了什么错误，在后期的工作中要注意哪些问题。爱数将筑森设计在项目设计过程记录并“转化”为宝贵知识，开放给知识库应用，将这些来源于生产的“经验”转化为知识，借助知识库应用对这些知识进行有针对性、有借鉴意义的知识运营。

图 35：筑森设计项目建设图



图：爱数，爱分析绘制

ifenxi

AnyShare Family 7 具备两个独特的优势。一是由爱数非结构化数据中台所提供的 AnyShare Family 7 集群 NAS 功能可以面向业务系统提供存储空间，替代理正协同设计系统将数据存放于服务器本地硬盘上，在保障数据安全的同时考虑未来数据量爆炸增长带来的挑战，构建属于筑森的企业内容总线。基于分布式存储，爱数 AnyShare Family 7 可较容易地实现扩容、升级、数据迁移，甚至用户都不需要数据迁移，未来只需要搭建应用系统。

二是 AnyShare Family 7 具有元数据引擎，元数据引擎具有对接功能，主要通过 API 形式与筑森设计协同设计系统对接。系统抓取海量 dwg 文件之间的“依存关系”并记录于 AnyShare 的元数据引擎中；同时，元数据引擎能够对元数据进行加工，增加元数据的注释性内容。根据实际知识运营工作中遇到的“新增需求”，可以新增更多维度的元数据，实现有针对性、有指导性的知识运营体系，将“设计过程”记录并存储，实现非结构化数据的结构化管理。

在传统模式中，新员工可能经过数月的培训才能正式接手工作，基于爱数非结构化数据中台解决方案，首先，筑森构建了围绕知识数据的内容管理平台，打通各个子业务系统，实现内容统一管理，知识统一运营，实现从项目管理、设计协同、图纸出版到最后项目归档的内容知识流转复用；

其次，为筑森设计打造端到端知识运营体系架构，还原优秀项目设计过程、传递成功经验、赋能一线设计人员，使得原线下、脱离项目实际情况的的知识运营体系转变为线上、连接业务、复制真实项目成功经验并进行传播的知识运营体系，同时，筑森可以根据相应的用户开放不同的查看权限，员工可以自主学习，相比于传统知识运营，更具有指导性的意见，有效提高人员生产力，可以极大减少从培训、学习到上岗的时间周期。其他的日常设计师也可以对自己不足的知识查询补漏。

最后，由于设计协同过程是基于项目管理进行的，筑森可跟着项目进度实现知识运营，项目完结即自然形成项目设计知识资料，可供新项目参考复用，打造了灵活高效的业务流程和创新的商业模式。

图 36：基于爱数非结构化数据中台的知识运营体系



未来，爱数将继续为筑森提升知识运营体系的优化。项目一期主要为理正协同设计平台做底层开发及企业内部体系梳理；二期规划中，筑森所有的应用系统数据都会迁移到 AnyShare Family 7 中，并应用知识图谱技术，实现以文搜图、以图搜文的功能；三期规划中，爱数会根据筑森的实际需求构建企业内部的数据湖，实现针对于非结构化数据的数据洞察，例如，过去筑森为客户进行土地竞拍设计时，都是基于日常经验和预算进行加价，但基于爱数数据湖的解决方案，系统可以通过海量非结构化数据分析，为筑森提供土地竞拍设计服务的时候提供指导性的意见。

3.2 数据应用实时性需求不断提升

随着数据应用的深入，企业对数据驱动业务决策的实时性要求在不断提升，以实现更敏捷高效的业务运营，例如更及时的客户响应。而实时计算与流处理等技术的成熟，也帮助企业能在越来越多的场景提升实时性。

企业需要从技术和业务流程两方面进行实时性数据应用的落地。在技术层面，可以建立基于流处理和实时计算的数据仓库，保障实时监控环境；在业务流程方面，需要明确业务层面的实时性需求，并设计相应的实时指标，以便业务人员监控业务变化，并及时做出实时反馈与处理。

某头部券商实时数据中台，支撑上千实时业务指标监控

某头部券商国内拥有证券市场业务全牌照的一流券商，致力于“全面提升核心竞争力，打造具有国际竞争力的中国最佳投资银行”。但随着业务、技术的发展，该券商在内部流程和外部业务方面都面临着挑战，也提出了相应的数字化升级需求。

该券商此前已经建设了成熟的大数据平台。随着数据应用走向深入，众多金融机构开始探索实时数据大数据平台，其中，券商对此需求更为迫切。不同于银行的存贷业务，证券的实时数据都是客户交易数据，且证券单客户交易频率远高于银行，因此，券商需要快速处理事件和响应客户需求，对系统实时分析的要求高。例如，其中关键业务场景之一是异常交易监控，监控是否有扰乱市场秩序的大单，如疯狂下单、拉抬股价、打压股价，大单委托、超额委托、超额撤单等扰乱市场秩序、欺骗股民的行为。

在此背景下，该券商的业务体系需要向实时化迭代，提高响应速度和运营精细化程度，这对数据中台和业务系统提出了要求：支持高并发、高准确性、实时反馈等。过去该券商的数据处理与反馈都是以批数据计算为主，未来需要以流计算模式进行，以实现场景触发式服务，只要客户触发一项功能就会触发后面的整套流程。

这些业务需求也对 IT 团队的开发运维提出了新的要求。一方面，基于 Flink 流计算框架开发的门槛高，目前该券商开发模块上传至生产环境，包括流程控制、数据源适配等，大多都要靠写代码实现，开发周期长，且企业内部缺乏足够的研发人员；另一方面，基于 Flink 的系统运维监控难。不同于传统较为成熟的数仓和 Hadoop 的监控体系，流数据处理的延时、反压不适用于传统的监控逻辑，需要监控整套工作流的各环节作业，该券商以往通过研发人员写脚本实现相应的功能，但不够平台化。

因此，该券商需要借助精通 Flink、技术能力强、能够提供实时计算技术支撑的外部厂商，构建实时数仓，为该券商搭建实时数据中台打好基础。同时，降低开发门槛，让更多的初级或中级开发人员甚至业务团队都能使用大数据平台，提升开发效率。

该券商选择与袋鼠云合作。袋鼠云是行业领先的数据中台服务商，致力于数字化基础设施建设。公司研发了云原生一站式数据中台 PaaS “数栈 DTinsight”，基于数据中台的理念和产品，帮助客户建设自主可控的企业数据中台和数据共享能力中心，挖掘数据价值，实现企业数智化转型。

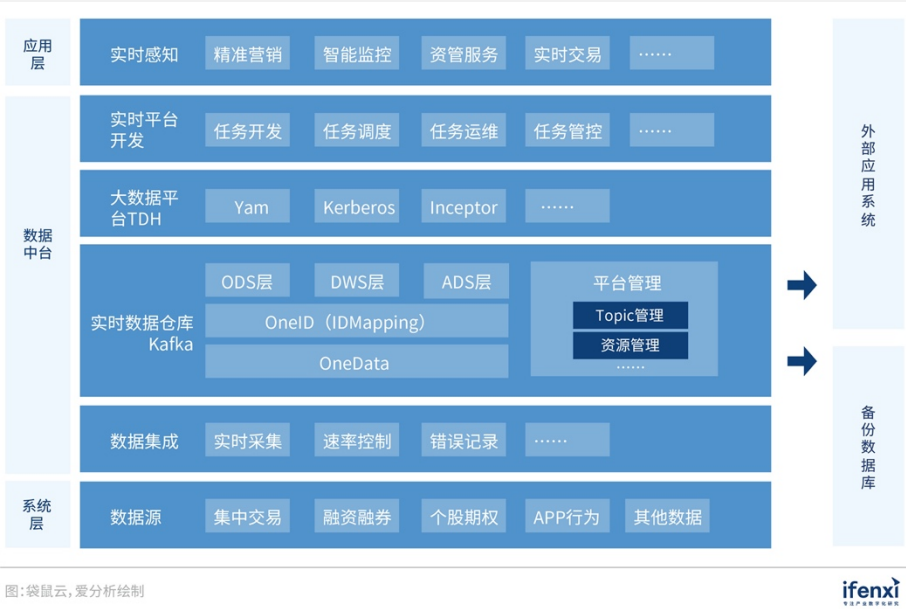
“数栈 DTinsight”具备天生云原生、自研核心引擎、金融级安全特性，覆盖数据源采集、离线计算、实时计算、数据挖掘、数据治理、数据资产管理、数据共享服务等整个数据中台生命周期，满足企业建设数据中台的多种需求；同时，兼容主流大数据平台，或基于自有计算引擎，可以快速完成大数据平台从 0 到 1 的搭建。

该券商实时数据中台项目分两期进行，一是实时数仓建设，二是大数据开发平台建设。

在实时数仓建设阶段，袋鼠云帮助该券商在前期制定标准、选择最终的技术栈，统一数仓建设的标准。基于数栈 DTinsight 帮助客户构建实时数仓，实时汇聚包含证券交易、融资融券、

个股期权等多个领域数据流，基于 Flink 技术提供实时数据源，推进企业 UCM（统一客户模型）体系构建，快速响应各业务系统数据应用需求，实现全域用户状态实时感知，为上层的业务方提供基于客户状态的实时数据处理，帮该券商重构底层实时数仓，打牢实时数仓的根基。

图 37：某券商实时数据中台架构



在大数据开发平台建设阶段，该券商将基于袋鼠云的标准化产品数栈 DTinsight 开发模块，提高开发效率、数据资产管理效率，加强其开发管理能力。

基于实时数据中台，该券商在实时数据计算与监控、系统稳定性和开发效率方面都有了显著提升。

在实时数据的计算和监控效率方面，过往，该券商的数仓 70% 以上的指标计算都不能在三秒钟以内完成，基于数栈 DTinsight 平台不断优化，该券商可以解决大规模数据计算问题，在计算指标数量大幅增加的前提下，也可以实现在 3 秒内完成 90% 的指标计算。例如，该券商可通过流计算搭建实时的反洗钱和风控模型，针对柜台、APP 的任何操作和用户行为系统可以马上给出结果并进行审批。若监控到用户交易异常，中央监控平台会即时出现弹窗，业务人员可及时处理；同时，也会有 MOT、短信、APP 弹窗提醒，提醒员工或客户及时处理。

基于实时数仓，该券商的实时业务指标数量大幅增加。2018 年该券商的业务只有几十个实时指标，从 2019 年至今，已有 1000 多个实时指标。例如，客户中签新股后是否有缴款就是新增的实时指标，投顾无需像过去一样不停地打电话询问客户是否已缴款，同时提升了企业运营效率和客户满意度。

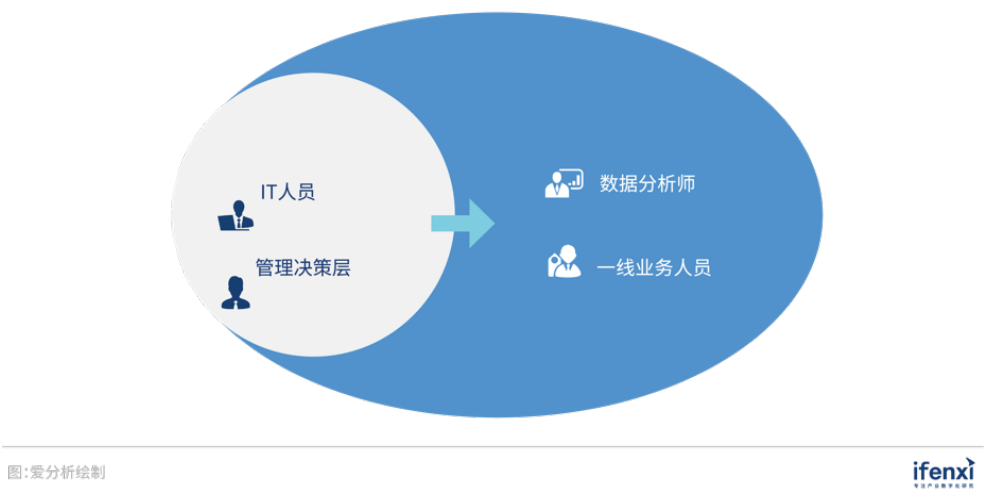
在稳定性方面，在应用数栈 DTinsight 平台之前，该券商自行搭建数仓和 Flink 集群，但系统不稳定，不敢将关键应用迁移至实时数仓。在袋鼠云的技术支持下，该券商系统更加稳定，可以将类似反洗钱等重要应用都迁至实时数仓，支持企业的核心决策，系统报错率大幅降低。同时，数栈 DTinsight 平台的工程化更好，可以大规模使用。

在开发效率方面，数栈解决了研发标准化和研发便利化的问题。数栈 DTinsight 平台是基于开源技术框架的成熟商业化产品，该券商无需把时间花在复杂的流处理代码开发上，只需要关注业务逻辑及数据建模，通过 SQL 化开发，即可快速响应业务部门实时数据需求，实现实时开发的统一规范。基于数栈 DTinsight 平台，实时数据开发效率提升 2 倍以上，实时任务排错及解决时间减少 80%，实时业务消费数据时效从秒级提升到毫秒级。

目前该券商所有的客户数据都已经实时化。未来，在业务层面，该券商要实现产品数据、公司经营数据实时化，以消息流的模式驱动公司业务，甚至实现实时化决策；在技术层面，该券商正在努力实现批流一体化，以一套代码实现批处理与流处理。

3.3 数据应用走向平民化

图 38：数据应用走向平民化



随着数字化转型的推进，企业数据应用场景将不断拓展，且数据服务需要快速响应日益增长的业务需求，要求数据应用走向平民化。

一方面，数据应用不仅限于服务决策层，还要服务大量的一线业务人员和数据分析师。另一方面，更多业务人员需要参与到数据开发工作中，实现自助式的数据服务，以提升效率。在这样的需求之下，敏捷 BI、数据科学与机器学习平台等可以降低数据开发和使用门槛的技术应用的价值凸显。

3.4 领先企业走向产业赋能

图 39：数据智能应用领先企业产业赋能



越来越多的行业头部企业正在成为数据智能应用的行业领先企业。这些企业一般处于产业链中的核心地位，对产业链的掌控能力强，不仅具备较强的技术能力，还拥有更多场景的数据沉淀。这些领先企业不仅在自身业务中应用数据智能，也逐步将能力向外输出，有机会服务于上下游企业。

对于企业而言，赋能产业链上下游，构建数字化生态，一方面能持续通过数据智能的能力强化企业在行业内部的竞争壁垒；另一方面，也能逐步实现自身商业模式的升级。例如，传统零售企业可以通过数据智能的应用与对外赋能，逐步转变为一家人工智能公司。

结语

随着新基建的政策落地和加速推进，各行各业在都加快数字化转型的步伐，数据智能已经成为企业驱动业务增长、寻求转型的主要支撑。企业数据爆发增长、海量集聚，建设数据中台以支持数据资产管理和数据智能应用，是大部分企业当下的必经之路。

但是，目前数据中台的建设和数据智能的应用仍在探索初期，在技术与产业方面仍然面临着诸多挑战。在技术方面，5G、人工智能、物联网等技术融合度较低，与数据智能结合应用的成熟度还不足；在产业方面，目前仅有部分行业的部分场景实现数据驱动业务，仍有大量行业需要进一步加强业务的数字化改造，探索更加广泛的数据智能的应用场景。

未来，数据智能的应用不再仅限于部分企业内部。随着技术成熟和产业发展，数据智能的应用场景会逐步从企业自身应用拓展至全产业链的业务协同，这不仅需要产业上下游共同协作、统一标准、开放数据与场景，还需要行业中具有技术与产品领先性的优秀数据智能厂商，共同赋能，将成熟的数据智能产品与解决方案应用于多行业多场景，推动产业智能。

关于爱分析

爱分析是一家中国领先的产业数字化研究与咨询机构，成立于中国数字化兴起之时，致力于成为决策者最值得信赖的数字化智囊。

凭借对新兴技术和应用的系统研究，对行业和场景的深刻洞见，爱分析为产业数字化大潮中的企业用户、厂商和投资机构，提供专业、客观、可靠的第三方研究与咨询服务，助力决策者洞察数字化趋势，拥抱数字化机会，引领中国产业数字化升级。

研究咨询服务

技术研究

新兴技术研究，厂商能力调研，助力数字化最优决策

商业研究

基于研究、数据和案例调研积累，辅助业务可靠落地

客户洞察

企业用户需求及实践调研，辅助制定业务与市场策略

品牌&营销

权威背书，树立行业地位；教育市场，精准触达客户

行业研究

行业数字化趋势与实践研判，辅助业务与战略决策

投资研究

成熟方法论，一手数据，助力研判机会、稳健投资

法律声明

此报告为爱分析制作，报告中文字、图片、表格著作权为爱分析所有，部分文字、图片、表格采集于公开信息，著作权为原著者所有。未经爱分析事先书面明文批准，任何组织和个人不得更改或以任何方式传送、复印或派发此报告的材料、内容及其复印本予任何其它人。

此报告所载资料的来源及观点的出处皆被爱分析认为可靠，但爱分析不能担保其准确性或完整性，报告中的信息或所表达观点不构成投资建议，报告内容仅供参考。爱分析不对因使用此报告的材料而引致的损失而负上任何责任，除非法律法规有明确规定。客户并不能仅依靠此报告而取代行使独立判断。

北京爱分析科技有限公司 2020 版权所有。保留一切权利。

如欲了解更多爱分析精彩洞见，请关注我们的微信公众号



©北京爱分析科技有限公司 2020 版权所有

咨询/合作

微 信：ifenxi888

网 址：www.ifenxi.com

地 址：北京市朝阳区酒仙桥路兆维华灯大厦 A1 区 1 门 2 层 2017



ifenxi

专 注 产 业 数 字 化 研 究