 DOI: https://doi.org/10.12349/iser.v6i7.6691

Research and Application of Enterprise-level Data Middle Platform

Longquan Tan

CRRC Times Electric Co., Ltd.,Zhuzhou, Hunan, 412000, China

Abstract

Over the past 20 years, IT managers of domestic and foreign enterprises have built information systems such as ERP, PLM, CRM, MES, and MRO through the construction of information projects, achieving the goal of business informatization. However, the rapid independent iteration of each system has led to the frequent occurrence of the familiar "siloed" system architecture and the phenomenon of information islands. With the rapid development of digital technology, new business models are accelerating the transformation of the world around us and people's lives. We hope to leverage our own experience in digital transformation, be committed to developing and applying advanced ICT technologies, actively promote the independent research and development of data middle platforms, and achieve the collection, storage, computing, and processing of massive data, as well as unified standards and standards. Provide users with flexible, efficient, reusable and low-cost data services. Lowering the threshold and investment for digitalization helps enterprises quickly integrate into the digital world, enabling them to develop more rapidly, healthily and steadily.

Keywords

data center; innovative application; data governance; data model

企业级数据中台研究与应用

谭龙泉

中车时代电气股份有限公司，中国 · 湖南 株洲 412000

摘　要

在过去的20年，国内外企业IT管理者通过信息化项目建设，构建了ERP\PLM\CRM\MES\MRO等信息化系统，实现了业务信息化目标，但是各个系统独立的快速迭代，使得我们熟悉的”烟囱式”系统架构、信息化孤岛现象频频出现。随着数字化技术的快速发展，新型的商业模式正在加速改变周遭的世界和人们的生活，希望借助于企业自身的数字化转型经验，致力于开发和应用先进的ICT技术，积极推进数据中台自主研发，实现对海量数据进行采集、存储、计算、加工，统一标准和口径，为用户提供灵活、高效、可复用、低成本的数据服务。降低数字化门槛和投资，帮助企业快速融入数字化世界，使得企业可以更快速、健康、稳步发展。

关键词

数据中台；创新应用；数据治理；数据模型

1 概述

全球加速步入数字经济新时代，数字化、网络化、智能化、智慧化将成为未来经济发展的方向[1]，习近平总书记强调，要抓住产业数字化、数字产业化赋予的机遇，不断激发中国经济高质量发展的新动能。国资委印发了《关于加快推进国有企业数字化转型工作的通知》，就推动国有企业数字化转型做出了全面部署。离散制造业产品呈现品类众多、结构复杂，小批化等特点，企业经营管理当前面临了业务体系庞大、精细化管理难度大、业务流程运转“不通畅”、数据存在孤岛、产品质量数据难追溯等问题。公司在十四五规划中明确了建立“精细化管理”与“智能化管理”的企业应用，通过数字技术来实现数字化赋能目标。

2 数据分析痛点问题

公司推进数字化转型建设以来，以指标和业务牵引，数据模型开发已快速交付业务需求为首要目标，开发过程和数据模型管理存在以下几个问题：

“数据不好找”：模型有很多，业务人员查不到，对外的数据资产目录没有形成规范，从最早的“市场、研发、售后、采购、生产”到按一级流程，缺少集团化的支撑能力。

“数据不能持续可用”：缺少治理环节的参与，导致模型上线后一段时间后无人问津，变成呆滞资产。

“数据不易懂”：快速交付模式下，各专项独立开发，

【作者简介】谭龙泉（1986–），男，中国湖南株洲人，本科，工程师，从事数据技术组长，数据分析领域信息化解决方案交付和实施研究。

模型技术和业务元数据丢失、同样的维度代码不同，给数据消费端造成困扰。

“数据架构不完善”：未建设公司统一的数据语言，数据分析停留在简单报表，且时效性不高，难以实现全面的业务监控。

3 数据中台建设目标

通过数据技术，实现对海量数据进行采集、存储、计算、加工，统一标准和口径，形成数据资产，为用户提供灵活、高效、可复用、低成本的数据服务[2]。

3.1 数据采集方面

需要支持实时和批量数据采集，数据类型包括主流的关系型数据库含：oracle、sqlserver、mysql、hana、postgresql、doris、hive、Mongodb、实现图片、机台日志文件采集和解析、支持第三方接口数据源（restful/soap）、支持通过接入其他数据库驱动的方式自定义数据库。

3.2 存算平台方面

提供分布式存储系统、流式数据处理框架、批量数据处理框架等功能组件。支持结构化和非结构化数据抽取和存储[3]、计算、支持集群节点横向按需扩展、支持集团化多租户管理模式，各租户之间数据相互独立，公司各产业按租户申请使用。

3.3 数据资产方面

提供数据全生命周期管理能力，含源系统定义、元数据管理、数据标准、数据质量、数据资产目录、数据开发、数据脱敏等能力。

3.4 数据探索方面

提供自助分析工具，满足用户不同的自主化分析需求，包括数据集选择、报表字段定义、图表类型选择等功能。

4 数据中台功能设计

坚持走独立自主的技术路线，利用大数据、物联网、AI 等新一代 ICT 技术，以标准化用户入口、流程、规则、数据的管理过程为基础，并结合 XC 的要求，确保流程通畅、数据可管、可用，界面风格一致，通过 3 年的建设，已发布数据中台“DCO”解决方案。

平台定位为企业级数据技术平台，基于 Hadoop 生态体系和技术中台微服务架构体系自主研发的平台级解决方案，专注于数据采集技术、数据存储技术、数据开发技术、数据治理技术；提供数据全生命周期管理一站式服务，帮助整合企业内外部业务数据、半结构化与非结构化数据，低成本快速实现数据价值最大化。平台功能包含了大数据存储平台、数据开发系统、数据治理系统三大模块功能，主要功能如下：(1)大数据存储平台：提供各种异构网络、异构数据源的数据采集、集中存储和计算能力。(2)数据开发系统：提供数据集成、离线开发、实时开发、算法等一站式开发功能，支撑快速实现数据处理。(3)数据治理系统：提供包含元数据管理、数据标准、数据质量、数据安全、数据资产功能，实现数据资产服务化。对比同行具有“快速交付、简单灵活、自助可控”特点，如图 1 所示。

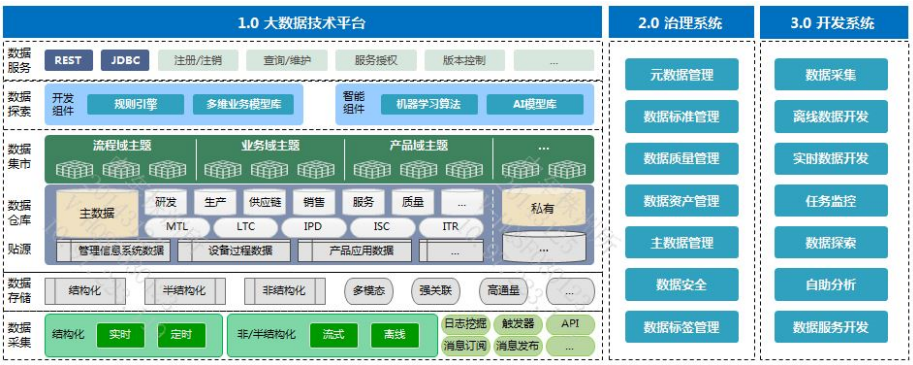


图 1：数据中台功能规划

5 数据中台在企业应用

平台历经 3 年研发现已投产使用，通过建立数据领域模型，完全拉通了企业研、产、供、销、服、管理等核心领域的业务数据链，形成企业级数据湖和数据管理应用。在助力企业创新发展方面：

数据中台方案助力企业经营管理能力提升，基于数据中台构建企业级 APP，通过接入 34 个核心信息系统的业务数据，实现了经营指标按公司、事业部、产品线、项目等多个维度的经营管理指标自上而下的分解贯通，提升产业协同发展的经营作战效率。

数据中台方案助力产品质量全生命周期管理，基于数据中台构建产品质量全生命周期追溯管理工业 APP，完全支撑了企业自产产品按批次管理到按序列号管理的业务精细化变革，打通了产品从物料寻源、采购、设计、工艺、外协、仓储、运输、装车的产品全生命周期管理，逐步将质控重心从生产过程环节向车载运行监控转移。

数据中台方案助力智能化生产能力提升，结合先进的 ICT 技术，依托 IOT 平台，接入智能化产线设备控制系统数据，采集自动化产线机台数据，实现生产制造过程透明化、产品良率分析和故障预警，提升了产线智能化控制水平及生

产效率。

数据中台方案的跨产业应用，数据中台方案从主营产业推广至新型发展产业，在开展不同产业园区管理、智能制造生产过程监控管理方面，提供了可差异化选配的数据应用解决方案。

6 存在的问题及对策

一是面临了大量个性化需求难满足：数字化平台应用单位反馈已有报表格式和图表形式不能满足个性化分析需求。

为满足用户不同的自主化分析需求，提供自助分析工具，包括数据集选择、报表字段定义、图表类型选择等功能，能对数据进行筛选、切割、排序、汇总等，自助灵活地达成期望的数据结果，建议基于开源技术组件，建设自助数据分析系统，完善数据源管理、数据集管理、数据可视化、自助数据模型开发能力，实现员工自定义数据分析。

二是技术难题亟待攻克识：大量 IOT 数据源采集存在技术问题，公共数据仓库贴源层数据对象存在重复抽取现象，众多业务源系统，面对海量数据的抽取时，给源系统带来一些性能压力。

组织开展技术平台共性组件开发与能力建设，以及技术规格、应用规范的标准化设计，形成信息化自主技术中台的技术路线、建设要求和推广计划；着力推动数据中台的数据可视化组件开发和数据权限管控能力升级，同步开展 IOT 库的 API 接口分析和数据采集通用程序开发，不断提升技术平台支撑力。完善大数据平台功能，按系统统一建设公司贴源层数据对象，并归集到数据资产目录统一管理。

三是数据治理能力待建立：数据治理工作尚未形成组织级长效机制与能力。

结合公司业务与信息系统应用现状，对标行业先进，从组织保障、文化建设和制度管控等三个维度构建公司数据治理体系。建立统一的数据治理管理制度、数据标准、数据质量规则，强调管理用数据说话、用数据支撑业务决策，落实数据管理的第一职责，建立考核机制，提升数据质量；同时，组织建设数据管理支撑系统，包括元数据管理、数据资产管理、数据质量管理、数据服务管理等系统平台，为公司产业数字化系统平台提供数据服务。

四是人力资源短缺，公司分支机构普遍存在 BP&IT 人员资源短缺问题，未设置匹配支撑业务流程优化和信息化支撑的组织。业务流程无人优化，导致业务运行低效低质量，同时，因 IT 人员缺乏，无法完全支撑业务流程的落地和业务效率的提升。

站在公司视角，做好信息化人才三年规划，全盘考虑信息化人力资源整合和集约化管理，盘活现有内部资源。

7 关于进一步完善平台的构想建议

立足公司“十五五”发展战略，为公司数字化转型提供一个技术领先、自主可控、经济适用的数据中台解决方案，实现搭建“一个数据技术平台”，研发“一个数据治理系统、一个数据开发系统”，为数字化转型提供技术平台支撑。

7.1 大数据平台服务支持能力扩展

数据底座持续扩展离线处理、实时流处理、实时检索、交互分析、MPP 技术数据库领域组件能力、建立平台运维和开发规范 , 降低数据存储与分析成本 [4]；构建基于大数据技术平台数据开发服务能力，实现数据采集、实时数据开发、离线数据开发、数据服务发布、数据自助分析等功能，提升中台开发效率和成本。

7.2 数据治理能力持续建设

构建基于大数据技术平台数据治理服务能力，实现元数据管理、数据标准管理、数据质量管理、数据资产管理等功能，提升数字化平台的数据管理能力和数据质量。规划并构建全所数据资产服务目录，面向各企业单位的数据资源开放共享、数据交换、数据分析等场景，为数据提供方与数据需求方之间共享数据提供平台，提供统一的注册、发现、查询、定位、获取数据资源的入口，发挥数据资产价值。

7.3 数据脱敏建设

构建一套数据安全技术防护体系 , 根据数据的分类分级规则，建立对应的数据安全管控策略，实现对数据全生命周期进行管控，包含数据从采集、传输、存储、处理、分享直到销毁的全部过程。

7.4 创新应用建设

顶层规划、分布推进：聚焦经营管理数字化建设目标，结合企业经营实际，统筹规划建设方案，分梯次、分阶段、分步骤完成建设。

统一标准 数据贯通：以一体化经营指标体系为牵引，以业务流程为主线，集成市场、研发、制造、采购、售后、财务等各项经营情况数据，按照“一套标准、一套数据”要求，基于统一的数据入湖标准，实现数据贯通。

8 结语

本研究立足企业数字化转型过程中数据迫切需求，基于 ICT 技术发展方向提出数据中台解决方案，即为企业提供数据“采、存、管、用”全生命周期平台管理能力。介绍了数字化转型过程中数据分析业务痛点、数据中台建设推进过程中存在的问题及对策思考。帮助整合企业内外部业务数据、半结构化与非结构化数据，低成本快速实现数据价值最大化。

参考文献

[1] IDC.《中国数据空间市场分析》[R],2025:45-47.

[2] 张照, 关晓峰, 崔伟. 基于数据中台的企业数字化转型分析[J].数字通信世界, 2024, (01): 170-172.

[3] 陈晓宇, 杨川, 胡陈啸. 深入浅出 Prometheus：原理、应用、源码与拓展详解 [M]. 北京: 电子工业出版社, 2019: 60-79.

[4] 华为公司数据管理部.华为数据之道,2020,10:98-99.