

航空发动机大数据平台应用研究

Application Research of Aero Engine Big Data Platform

李雪莉 于明达 何为 / 中国航发动所

利用数据治理、集成、大数据存储和分析等技术，实现航空发动机研发过程业务数据深度融合，打通数据壁垒，构建统一的大数据平台和航空发动机全域数据资产目录，利用运营可视化思想和数据分析技术，对大数据平台的信息进行展示、分析、挖掘，助力航空发动机研发。

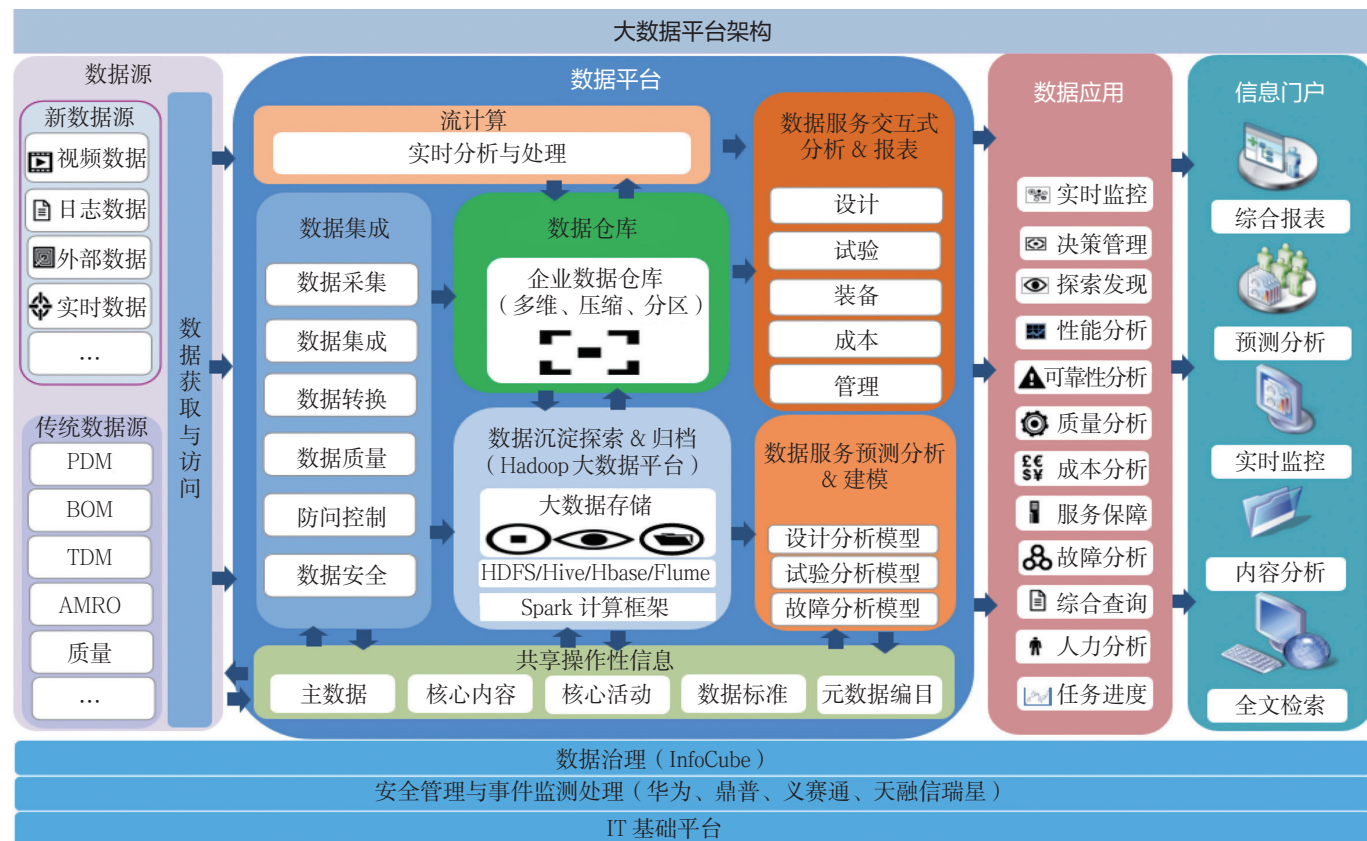
数据是航空发动机研制中的重要战略资源，对完善、改进设计的重要性日益凸显。中国航发动所积极推进研发系统建设，夯实了信息化基础，积累了大量的设计、加工、装配和试验数据，迫切需要综合整理这些数据，让海量的数据形成智慧资产并

创造价值，形成助推科研的关键要素。根据实际需要，应利用数据治理、数据集成、大数据存储和分析等技术，实现业务数据的深度融合，构建统一的大数据平台和航空发动机全域数据资产目录，并在此基础上，利用运营可视化思想和数据分析技术，对大数据平台的信息进行展示、

分析、挖掘，可提升航空发动机整体研发能力。

航空发动机大数据平台设计

通过构建航空发动机大数据平台，建立覆盖设计、装配、工艺、试验、维修等全生命周期的大数据系统，从业务系统及外部数据源中获取数



航空发动机大数据平台总体架构

据中心所需的数据。大数据平台位于数据源和工程数据中心平台之间，屏蔽了数据中心及数据处理与数据源间的直接关系，通过系统地管理发动机的设计生产、使用各阶段数据，打破按专业管理数据的壁垒，使各维度的数据之间建立关系/纽带，解决发动机研制过程中的复杂问题，积累设计、工艺、生产经验，促进航空发动机自主研发体系建立。

航空发动机大数据平台的架构总体上应满足高效率、稳定、可配置、可维护、可监控、可扩展等特性，需要一套合理的、健全的、成熟的、统一的监控调度策略，以保证整个系统安全、稳定、简单地运行。大数据平台的建设应包含数据整合、转换、存储和应用等基础能力。

数据源整合。应用数据仓库技术(ETL)，大数据平台应能够整合各应用系统源数据，自动进行数据实时和批量采集；同时针对缺失的数据提供手工补录功能，能够分析缺失数据的源头并针对数据源提出合理的改造方案。

数据转换。对从不同数据源采集到的数据，根据数据模型的要求，进行数据的转换、清洗、拆分、汇总等处理，保证来自不同系统、不同格式的数据的一致性和完整性，为应用平台提供高质量的数据服务。

数据存储。应实现发动机各研制阶段(全生命周期)、各设计专业、制造/装配/工艺、各类试验数据融合统一的平台的能力，针对各种数据按一定规则进行存储汇聚，并充分考虑动力所历史数据的存储及在线查询。

数据应用。按特定规则实现对复杂航空发动机数据进行分析、挖

掘的目的，以实现发动机设计规律和内在联系显性化的目标，从而实现对发动机设计、制造装配工艺和保障维护能力的提升。应实现对外、内场使用的海量发动机进行趋势预测，应能实现对外、内场制造及使用海量发动机数据进行实时监控和故障诊断行。

大数据平台技术特点

支持“湖仓一体”的多维度、多层级、多主题先进多维数据架构

航空发动机多维度数据包含大量的结构化和非结构化数据，并具有多样性、专业性、关联性、流程性、时序性和解析性等特点。在构建发动机大数据平台时，采用多层级数据湖和数据仓库相结合的存储架构，充分发挥数据仓库和数据湖的优势，纵向按管理、设计、加工、装配、试验分类存储，横向按发动机研制过程建立设计模型，通过模型指导数据的采存管用，从各层级研制需求出发，形成类型研制主题，对数据开展综合查询分析、发动机研制过程管控、发动机振动与故障诊断、滑油光谱超标分析诊断、发动机性能衰减评估等应用分类存储，整体架构利用“湖仓一体”采用松耦合方式，保障系统的灵活性和可扩展性，打通航空发动机全生命周期数据链，形成统一的航空发动机大数据平台和灵活定制的数据应用场景。

灵活高效的分布式计算框架和自动化数据融合与处理流程

利用Spark分布式计算框架的内存计算、Cache缓存机制和有向无环图(DAG)特性，通过迭代式计算和数据共享，基于大数据平台，实现发动机多维度数据的批处理、计

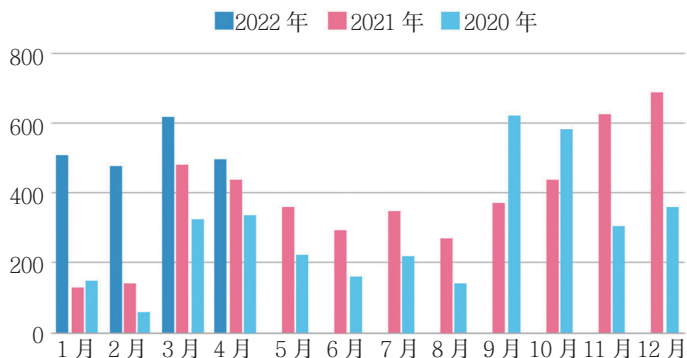
算分析、交互式计算等。基于Spark分布式计算框架，高效、灵活地构建发动机试验数据分析模型，对发动机振动与故障诊断、滑油光谱超标分析诊断和性能衰减评估等进行快速计算分析，反哺发动机的研发设计。

形成面向全生命周期的单一数据源

针对发动机研发过程中产生周期长、阶段特性显著、数据量大、关系复杂等数据特性，构建了从需求到使用的面向研制多个阶段的数据库，通过统一标志、多层次、面向快速更改的发动机数据模型，形成发动机各阶段的数据阶段视图，用于索引和连接发动机研发过程中的不同种类数据，对相关型号研制进程中设计、加工、装配、试验等数据进行建模，对研制进程中关键领域进行数据分析与集中展现，为科研及管理决策人员提供稳定、可靠、客观的数据支持，支持在发动机研制各阶段为各类应用系统提供统一、完整的数据源。

航空发动机大数据平台建设与应用

大数据平台的设计应当充分考虑数据质量的问题，准确的、可信的分析结果建立在准确的数据基础之上，应该有良好的机制确保数据的准确性，能够尽早发现和定位数据质量问题，尽可能提高数据的质量。在架构和技术层面，大数据平台和外围业务系统应保持松耦合的关系，确保数据运行不会对关键业务系统的性能和稳定性有影响，最重要的就是体现在如何从源系统抽取数据，既要保证对源系统的影响最小，同



试验信息统计

时也可适应源系统的数据源的变化，这就要求数据抽取这一层的设计具备足够的灵活性、稳定性和源系统的无关性。随着数据量的增加，应该提供有效的数据全生命周期管理策略和高性价比的水平扩展和垂直扩展能力，确保效率和成本的可控。根据上述分析和航空发动机大数据平台的设计需求，主要从数据整合与存储、动态查询分析和大数据分析预测等方面进行大数据平台的建设。

数据整合与存储

基于发动机维度相关数据源的特点，建立数据采集整合策略和数据补录机制，依据数据标准和数据质量规则，在数据采集整合过程中对数据进行处理，建立数据存储机制，形成跨专业、跨地域及跨研发阶段的数据关联整合，分层存储解决复杂的系统问题。

面向研制过程主题域的动态多维度查询分析

通过航空发动机大数据平台，建立各系统之间的数据关联，构建科研管理、试验、人力、财务、质量、服务保障、企业文化等主题数据应用和主题查询分析看板。通过跨不同维度的数据查询分析，找出发动机研发过程的规律，建立统一的数

据管理机制，实现对数据有效的积累，从时间维度上可以总结发动机规律，从数据类型维度上避免了原来系统间数据统一问题。

大数据分析预测

航空发动机大数据分析技术是利用包括统计分析、支持向量机、数据可视化技术、人工神经网络在内的多种数据挖掘与分析手段对收集的航空装备状态数据进行分析与挖掘，提炼数据中隐含的反映装备运行状态的规律，从而实现了对复杂数据规律的显性化。通过知识发现，利用现代的数学方法和原理，借助先进的信息技术手段，把隐藏在复杂数据信息之间的已知、未知或想知的知识、原理、模型等提炼出来。同时，数据分析仍在不断改进升级，具备不断扩充分析判据、分析方法的功能。

大数据平台应用效果

航空发动机大数据平台目前整合了包括科研管理、CDM、PDM、TDM、人力、财务等在内的十几个系统数据，数据量达到TB级，打通了系统之间壁垒，建立了系统之间的数据关联，建立了近百个各种主题看板分析查询应用，提升了发动机研发

管理能力和研发效率；建立了噪声、发动机性能衰减、滑油光谱超标分析等专业主体的数据分析挖掘应用，将相应领域的数据分析效率提高了几倍到几十倍，加速了发动机研发过程。

结束语

航空发动机大数据平台的建设与航空发动机设计研发涉及的各项业务形成良性的互动与支撑模式，满足用户基于不同业务层面的诉求，实现了单位内部不同业务的数据互通，提升各类主题数据的增值服务，促进了管理机制的转型与升级。同时，建立了标准化的数据治理机制，通过不同主题领域内数据的融合与治理，建立面向多源系统的异构数据管理机制，促进了核心业务环节数据标准化，具备优化数据模型和知识重构的能力，实现数据驱动关键业务的创新服务模式，从被动定期服务发展成为主动实时服务，为航空发动机大数据分析场景的全面应用和发动机数字孪生的建设奠定数据基础。

航空动力

(李雪莉，中国航发动力所，工程师，从事软件开发设计和数据应用管理)