Integrated Management and Application of Product Support Data

Weibo Pan, Jialong Chen*, Rui Zhang, Ting Wang

Wuhan Second Institute of Ship Design and Research, Wuhan, 430064, China

1538466240@qq.com

Keywords: Product Support, Data Integrated Management, Data-driven, Intelligent Support

Abstract: To meet the demand for data-driven transformation of product support, the integrated management and intelligent application of product support data were studied in this paper. In response to existing challenges such as “data silos”, different standards, difficulties in integration, and insufficient value exploration, an overall architecture comprising “One Data Base, Two Systems, and Three Layers of Application” was proposed to establish a data resource system that is logically unified, standardized, safe and controllable. Key technologies such as multi-source heterogeneous data fusion, “Data Lakehouse” storage, and the integration of large-scale models with intelligent technologies were discussed in this paper. Taking examples of typical scenarios such as product lifecycle traceability, intelligent failure localization, and precise resource scheduling, the transformation path for data into support capability was demonstrated. The systematic data integration and intelligent application could drive the evolution of support modes from being fragmented and reactive to being integrated and proactive, thereby providing robust support for creating combat effectiveness.

1. Introduction

Along with increasing modernization and intelligence of products, the structural and technical complexity of products continues to rise, causing complexity to the support of modern products [1]. Product is the material foundation for combat effectiveness, and the level of modernization in its management and support capabilities is directly associated with combat readiness and sustained combat capabilities. Data is the core factor of production for new-type support capability. The effective integration, unified management, and intelligent application of product management support data to break down the barriers of “siloed” systems and explore the underlying patterns of value within the data, is the only path to transition from an “experience-driven” to a “data-driven” support mode and achieve the goal of “precision support”. At present, products are rapidly evolving towards system of systems coordination and intelligent autonomy, leading to an exponential increase in product complexity, and enabling the data required for lifecycle management and support to be massive, multi-source, heterogeneous, and highly dynamic [2].

2. Current Situation and Challenges

New characteristics of warfare and weapon products set higher requirements for training systems of products, demanding reforms in training methods and enhanced training on high-tech support products [3]. Currently, product support data management is faced with the challenge of transformation from a traditional, fragmented model to an integrated, intelligent one. Covering the entire product lifecycle, data, due to its diverse sources and fragmented business processes, is faced with several challenges in such key business domains as product maintenance information management, management of equipment and consumable parts, technical information and training management, product operation and reliability management, and failure information and experience feedback management.

First, system barriers cause “data silos”, and restrain the effectiveness of business collaboration. Business data is stored in independent systems across different departments and organizational levels, resulting in significant differences in technical architecture and data models, as well as difficulties in interoperability. This physical and logical isolation creates difficulties in smooth data sharing and

  DOI: 10.25236/iwmecs.2025.063

process integration for cross-domain business collaborations (such as dynamic adjustment of equipment supply and maintenance plans according to failure predictions), which severely affect response efficiency and resource coordination capabilities.

Second, the lack of standards leads to “data dialects” and diminishes the value of data integration. A unified standard was not established for product support data elements, coding systems, and semantic specifications on a whole. Different departments define data based on their own practices, causing inconsistencies in the connotation of information. For instance, the same part is identified differently in factory records and ledgers, and the failure descriptions are unmatched in maintenance records and technical information. Such “data dialects” not only generate redundancy but also hinder cross-stage and cross-business data association and deep integration, resulting in the loss of data value at the source.

Third, weak association produces limitations on in-depth analysis and difficulties in knowledge consolidation and reuse. There is a lack of an effective mechanism of semantic associations and integration in multi-source and heterogeneous data, including product operation, maintenance history, failure modes, equipment replacements, and staff training. The absence of deep connections makes it difficult to systematically trace the root causes of failures, identify defects across a product family, and evaluate the actual impact of training on reliability. Constrained by unstructured text, a large volume of experience knowledge could not be effectively extracted, correlated, and reused, which hinders the continuous optimization of support capabilities.

Fourth, the superficial application brings obstacles to intelligent upgrades, and insufficient decision-support capabilities. The current data applications are primarily limited to query statistics and report generation within a single system, without deep exploration and the construction of intelligent models based on domain-wide integrated data. The potential of data in predictive maintenance, precision equipment replenishment, optimization of support plans, and evaluation of system effectiveness is far from being fully exploited. Decision-making remains heavily dependent on personal experience instead of data-driven scientific insights, making it difficult to achieve a fundamental transformation of support mode from “reactive response” to “proactive forecasting and precise regulation”.

3. General architecture design

In response to the current situation and challenges, a general architecture comprising “One Data Base, Two Systems, Three Layers of Application” was proposed in this paper, aiming to establish a data integration management system with unified logic, physical distribution, standardized specification, and shared services, thereby providing robust support for the digital and intelligent transformation of product management support. In terms of the core concept, this architecture was related to aggregating, governing, and managing dispersed and heterogeneous data resources in a unified platform, on such basis achieving diverse intelligent applications, and ultimately serving the entire process of support command, operational management, and decision-making, so as to achieve a leap in value from “data resources” to “data assets” and finally to “support capabilities”, as shown in the Figure 1.

(1) A Base

As the core of the product support data system, “A Data Base” involved the architecture of “Data Lakehouse” to achieve virtualized data integration and centralized logical management. A comparison of data warehouse, data lake, and data lakehouse is shown in Table 1 [4]. The data base logically consisted of three layers. The layer of Data Access connected to business systems at all levels via a variety of adapters, enabling real-time or quasi-real-time data access, and ensuring a continuous and stable flow of data. The layer of Storage and Computing employed a hybrid storage architecture, where distributed objects stored massive amounts of raw data and an analytical database managed high-value structured data. It integrated multiple computing engines and enabled elastic scaling of resources and intelligent scheduling of tasks via a containerized platform, supporting diverse computing demands from real-time monitoring to in-depth analysis. The layer of Data Governance served as a central system spanning the entire data lifecycle. Its core functions include implementing unified data standards and models for data consistency; performing automatic auditing and remediation of data quality via rule engines and machine learning; enabling full-process data traceability via lineage analysis; and creating an intelligent data asset catalog for data cataloging, classification, search, and evaluation, thereby realizing data transformation from resources into operational assets.

Table 1 Comparison of Data Warehouse, Data Lake, and Data Lakehouse

	Data Type	Collection Process	Access Methods	Reliability and Safety	Applicable Scenarios
Data Warehouse	Structured, processed data	Schema on Write	Primarily SQL, with limited API support	High data quality and high security	BI, SQL applications, reports, etc.
Data Lake	Structured, semi-structured, and unstructured raw data	Schema on Read	Open API	Poor data quality and security, easily leading to data swamps	Data Science, Machine Learning
Data Lakehouse	Structured, semi-structured, and unstructured raw data	Schema on Read, Schema on Write	Open API	High data quality and high security	Diverse scenarios

(2) Two Systems

The “Two Systems” comprised the system of standards and specifications and the system of safety operation and maintenance, serving as the foundation for the stable, reliable, and efficient operation of the data base.

The system of standards and specifications is the fundamental solution to breaking down “data silos” and eliminating “data dialects”. This system shall encompass at least four aspects. First, basic data standards, including unified Unique Identifiers (such as UIDs) for products and their components (assemblies), unified equipment codes, unified failure mode codes, and unified codes for personnel and organizational units, which provide basic values for data associations; second, the technical interface specifications define the protocols, formats, and security requirements for data exchange and service calls between business systems and the data base, offering interconnectivity; third, data models and semantic specifications, the most critical step, required the construction of a domain ontology covering the core business concepts of product support, and clear definition of entities such as “product”, “maintenance activity”, “failure”, “equipment”, and “training course”, as well as their attributes, relationships, and business rules, so as to provide data with consistent and explicit semantics, and enable automatic understanding and intelligent association of data across different sources; fourth, full-process management regulations clearly defined the responsible parties, operational procedures, and quality requirements for stages of the data from generation, input, transmission, processing, application, archiving, and destruction, so as to specify regulations for data governance.

As a robust shield for the data lifecycle, the system of safety operation and maintenance emphasized both techniques and management, implemented multi-layered security measures covering network, host computer, application, and data regarding technical protection, including access control, data encryption, permission management, and dynamic data masking. A mechanism was established for data classification and grading, as well as full-lifecycle security management. With respect to operational management, platform monitoring, log auditing, fault alerting, and emergency response mechanisms were improved for stable operation; a system of regular data backups and disaster recovery drills was implemented to achieve business continuity and data reliability.

(3) Three Layers of Application

The "Three Layers of Application" is a layer of realizing the value based on the foundation, following the progressive logic from basic services to common business and then to intelligent innovation, with the purpose of gradually promoting data capabilities to be service-oriented, scenario-based, and intelligent.

Located at the bottom of the architecture, the Support Service Layer served as the "standard pipeline" and "capability foundation" for releasing the value of data. It encapsulated the general-purpose data capabilities of the base and provided standardized data query, exchange, computation, and visualization services to upper-layer applications and external systems in the forms of API services, message queues, data subscriptions, and so on.

The Common Application Layer focused on common business requirements in the domain of product support, developed core analytical applications based on integrated data across all domains to address the issue of cross-domain analysis that is difficult to achieve with a single system. Typical applications included product lifecycle digital archives, health status assessment and prediction systems, a unified visualization map of support resources, intelligent failure knowledge bases and case retrieval systems, which directly promoted daily support management.

At the forefront of deep data mining, the Domain Intelligence Layer employed artificial intelligence technologies to build advanced models and applications, driving the decision-making from being "experience-driven" to being "data- and model-driven". It included intelligent failure diagnosis and predictive maintenance, precision supply chain optimization and intelligent replenishment, intelligent planning and simulation of maintenance solutions, and simulation and evaluation of product system performance, enabling proactive predications and precise decision-making.

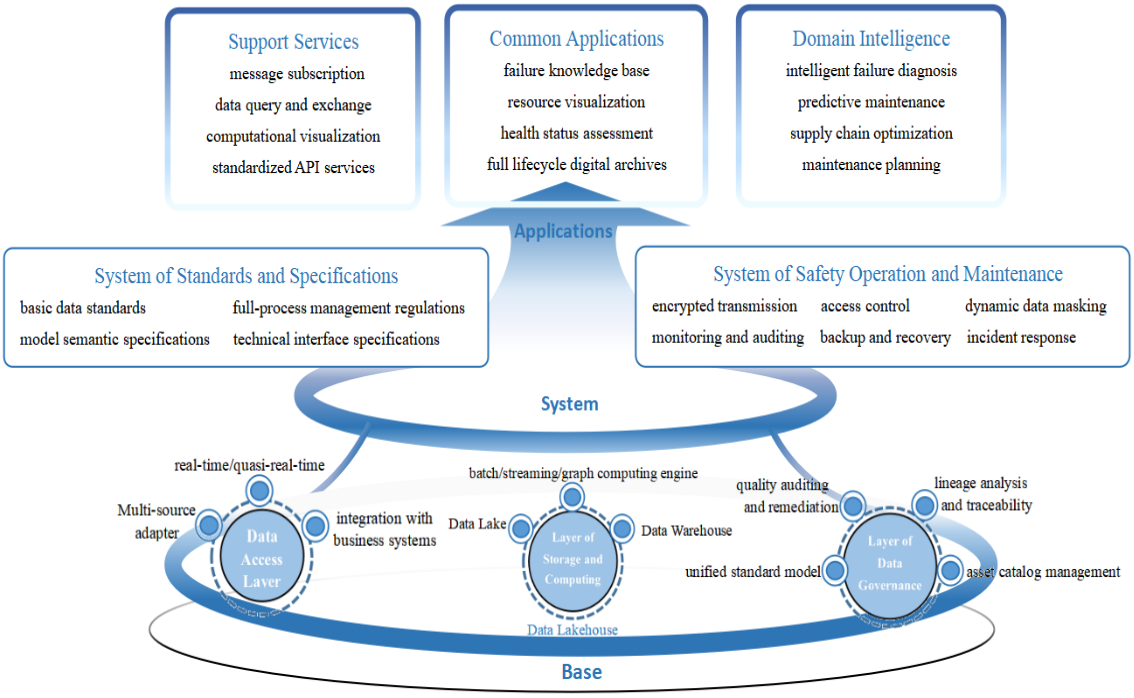

Fig. 1 General architecture of data integration management

4. Key technical pathways

According to product characteristics and support requirements, focusing on the objectives of building intelligent support capability and intelligence, as well as the key technical pathways for integrated management of support data [5], a technical framework was required to cover the entire data lifecycle.

The technology for the integrated access and fusion of multi-source, heterogeneous data was fundamental to breaking down “data silos”. With a metadata-driven, flexible access framework, the standardized access to various data sources was enabled via adapters, including relational databases, time-series databases, and file systems. As for unstructured data such as drawings, technical manuals, and maintenance reports, the automated extraction and structured processing of key information was achieved by combining Optical Character Recognition (OCR) with Natural Language Processing (NLP). At the level of data fusion, ontology was built for the domain of product support to enable the intelligent linking and integration of cross-source data via semantic mapping and entity alignment, thereby creating a unified data view covering the entire product lifecycle.

Data governance and technologies for building high-quality data assets are essential to ensure reliable and usable data. A governance mechanism was established to span the entire data lifecycle of access, processing and application, and enable data traceability and impact analysis by data lineage analysis. The quality control method combining rules engine with machine learning was adopted for automatic detection and recovery in data completeness, consistency, and accuracy. On such basis, a dynamic, intelligent data asset catalog was constructed to enable the automatic cataloging, classification and grading, and valuation of data resources, thereby facilitating the transformation of data from raw resources into standardized assets.

The flexible storage and collaborative computing technologies based on the architecture of “lakehouse” provided the infrastructure for massive amounts of multimodal data. It integrated the flexible storage of data lake with the efficient analytical capabilities of data warehouse, adopted distributed object storage to store multi-source raw data and made use of a distributed analytical database to manage high-value, integrated data. By virtue of a unified metadata service layer, seamless data flow and consistent management were realized between data lakes and data warehouses. The computing framework required integrating various engines, including batch processing, stream processing, and graph computing, to meet diverse computing needs such as real-time monitoring, offline analysis, and complex correlation relying on a containerized resource scheduling platform.

The technology of deeply integrating large models with domain intelligence is reshaping the technical paradigm of data applications. In data governance, large models can deeply understand the semantics of product technical documentation, automatically extract entity relationships, and generate structured summaries, thereby significantly enhancing the efficiency of knowledge construction. In terms of analysis services, intelligent parsing and automated responses to business requirements were achieved based on the technology of shifting from natural language to query language, thereby lowering the barrier to data analysis. In the field of intelligent applications, the effective combination of large models with simulation systems and knowledge graphs enabled the development of decision-support agents with complex reasoning capability, providing intelligent solutions for scenarios such as product health management, maintenance plan optimization, and support resource scheduling.

Data security and privacy computing technologies formed security defense for data sharing and utilization, where data classification and categorization strategies were implemented to achieve fine-grained data access control in combination with dynamic data masking and access control technologies. In cross-domain data collaboration scenarios, privacy computing technologies such as Secure Multi-party Computation and Federated Learning were employed for achieving joint modeling and analysis under the condition of “data available but not visible”. Taking advantage of the immutability of blockchain, key data operations were recorded and traced to ensure the trustworthy and controllable data lifecycle.

5. Profound application scenarios

Relying on key technological pathways, the core challenges faced by data integration and management were precisely addressed and resolved, thereby transforming disparate data resources into quantifiable and assessable new support capabilities, and achieving a leap in value from data elements to combat effectiveness.

(1) Connection and Traceability of Product Lifecycle Status Based on a Unified Data View

A unified data view centered on a unique product identifier was outlined to semantically connect design data, production records, operational monitoring, maintenance history, and equipment replacement records to form a coherent “digital mainline of product lifecycle”, which supported quality problem close loop, failure tracing, and performance evaluation, and enabled precise data-driven attribution.

(2) Root Cause Analysis and Identification of Systemic Weaknesses Based on Intelligent Fusion Analysis

The in-depth correlation and graph analysis were conducted on integrated operational data, maintenance data, equipment consumption data, and failure reports. Failure propagation networks were automatically constructed to identify high-frequency failure patterns, and uncover potential systemic factors (such as common issues with specific equipment batches), thereby forming the capability to identify the systematic health status of the product and risks.

(3) Precise Support Resource Scheduling Based on Predictive Models and Optimization Algorithms

The real-time product status, failure patterns, task schedules, and inventory distribution data were integrated, and machine learning was employed to predict demands for spare parts and maintenance workloads. The application of optimization algorithms could dynamically generate optimal inventory pre-deployment plans, spare parts delivery routes, and maintenance resource deployment plans, thereby evolving from “stocking and waiting” to “predictive pre-deployment”.

(4) Enhanced Effectiveness of Cross-domain Collaboration Support Based on Privacy Secure Computing

In joint support, taking advantage of technologies such as Federated Learning, participating parties could jointly train failure prediction models or collaboratively optimize support plans without sharing sensitive raw data, thereby constructing a trusted collaborative support capability where “data stays while value moves”.

(5) Support Command and Decision-making Assistance Based on Natural Language Interaction and Intelligent Generation

An intelligent support command assistant was provided. The commander could use natural language to directly inquire about complex operational situations, and the system automatically interpreted the intent, accessed integrated backend data and models related to maintenance, equipment, operations, and failures, and rapidly generated comprehensive reports containing risk analysis and recommended solutions, thereby enhancing the agility and scientificalness of support command.

6. Conclusions

The integrated management and profound application of product management support data is the core initiative for enhancing product support capabilities in the era of digitalization and intelligence. In this paper, the relevant challenges were systematically analyzed, an integrated management architecture centered on a logically unified data base was constructed, and key technologies and application scenarios were outlined. Its implementation would undoubtedly drive a fundamental transformation in product support modes from a crude, reactive manner to a refined, proactive, and targeted way, laying a solid data foundation for the continuous improvement of combat effectiveness. In the future, with the continuously evolving technologies, an intelligent support ecosystem based on all-domain data fusion will inevitably become the direction of development.

References

[1] Wang Yabin, Wang Shuai, Wang Jinguo, et al. A Review of Key Technologies for Equipment Support Based on Digital Twin [J]. Science Technology and Engineering, 2022, 22(28): 12233–12241.

[2] Guo Lu, Mao Yong, Yao Huiju. Research on Big Data Systems for Integrated Equipment Support Design [J]. Air & Space Defense, 2019(1):70-76.

[3] Chen Feng, Zhang Zijun, Hong Dongpao. Exploration of Comprehensive Support Models and

Development for Weapons and Equipment [J]. Aerospace Industry Management, 2020, (08): 50–55.

[4] Chen Qing, Song Shiwei. A Study on the Lakehouse Architecture from the Perspective of Data Governance [J]. Digital Library Forum, 2023(4): 19–28.

[5] Gao Zilong, Zhao Lei, Chen Benjun, Liu Zhao, Gao Wangsheng. Research on the Construction of an Intelligent Support System for Weapons and Equipment [J]. Aerospace Industry Management, 2022, (07): 41–45.