Automated Data Transformation in the Execution of ETL Operations in the Aviation Industry

Maksim Pivovar [1]

[1] *Transport and Telecommunication Institute, Lomonosov str. 1, Riga, LV-1019, Latvia*

Abstract

The aviation maintenance industry faces significant challenges in integrating maintenance information systems (MIS) due to the diverse nature of data formats and the increasing volume of information. This paper introduces the aviation maintenance transformation platform (AMTP) as a novel framework designed to standardize and streamline ETL (extraction, transformation, loading) processes within the aviation sector. By developing a universal data ontology and incorporating advanced data quality assessment algorithms, the platform aims to enhance the reliability and efficiency of data integration. The platform's architecture includes predefined transformation rules and a robust analytics system, enabling the automation and optimization of data handling procedures. The research highlights the critical need for a unified approach to data management, addressing key research questions related to system interoperability and data transformation automation. The AMTP offers a scalable framework adaptable to various MIS environments, thereby facilitating more effective decision-making and operational efficiency when implementing ETL projects. The paper describes the concept and the high-level architecture of the AMTP.

Keywords

ETL, extraction transformation loading, maintenance information system, data integration, aviation maintenance

1. Introduction to the Field of Research

The field of aviation maintenance, particularly the integration of maintenance information systems (MIS), is pivotal in ensuring the operational readiness and safety of aircraft [1-3]. The essence of MIS in aviation revolves around the systematic management of maintenance data which encompasses records of upkeep, repairs, inspections, and compliance with aviation regulations. This data is critical not only for maintaining the airworthiness of aircraft but also for planning, forecasting, and implementing preventative maintenance that mitigates potential failures [4].

Baltic DB&IS Conference Forum and Doctoral Consortium 2024
m.pivovar@sntechnics.com (M. Pivovar)
0000-0002-1191-8197 (M. Pivovar)

Historically, the evolution of MIS in aviation has been closely tied to advances in information technology and regulatory changes [5]. The early stages of MIS utilized basic record-keeping methodologies, primarily paper-based, which were cumbersome and prone to errors. As the aviation industry expanded, the complexity and volume of data grew, necessitating more sophisticated systems. The introduction of computerized systems in the late 20th century marked a significant advancement, providing more reliable and accessible data storage and retrieval systems [6-8].

In the contemporary landscape, the drive toward digital transformation in aviation maintenance is fueled by the need for higher operational efficiency and improved safety measures [1, 9-12]. The integration of advanced technologies such as big data analytics, cloud computing, and artificial intelligence (AI) into MIS is transforming how data is processed and used. These technologies enable more dynamic and predictive maintenance strategies, which are essential in an industry where the cost of unplanned maintenance can be exorbitantly high.

The implementation of MIS faces several challenges, primarily due to the heterogeneous nature of data sources and formats in the aviation sector and the need to manage a growing amount of information [1, 13-16]. Different stakeholders, such as aircraft manufacturers, maintenance providers, and airline operators, often use disparate systems that are not natively compatible with one another [17, 18]. This diversity creates significant challenges in data integration, requiring robust and flexible ETL (Extraction, Transformation, Loading) processes to ensure data is accurately consolidated and usable [19].

This paper explores the development of an aviation maintenance transformation platform (AMTP) aimed at addressing these challenges. By proposing a universal ETL environment and a flexible data ontology, this platform seeks to standardize the approach to data integration in aviation maintenance, enhancing the efficiency and reliability of MIS across the industry. The research delves into optimizing these integration processes, offering a scalable framework that can adapt to various data formats and systems, thus pushing the boundaries of what is currently achievable in aviation maintenance management.

2. Research Questions and Objectives

The main thrust of this research revolves around addressing the complexities of integrating diverse MIS within the aviation maintenance sector, by proposing a standardized and scalable AMTP. This study is driven by several critical research questions:

- RQ1. How can the integration of disparate data formats and systems in aviation maintenance be streamlined through a universal ETL environment? This question explores the potential for creating a flexible yet standardized approach to handling varied data sources, ensuring that they are compatible with a unified system for maintaining and analyzing aircraft data.
- RQ2. What are the challenges associated with automating the ETL process in the aviation maintenance industry, and how can these be overcome? This question

delves into the specific obstacles faced when automating data transformations, including technical limitations, data diversity, and system interoperability issues.

- RQ3. In what ways can a pre-configured analytics system within the platform enhance the efficiency and effectiveness of data transformation processes? This explores the impact of advanced analytics and pre-configured settings on optimizing the data transformation process, aiming to improve project resource management and operational decision-making.

To answer these questions, the research aims to achieve several key objectives:

- Design and implement a comprehensive data ontology that accommodates all potential source data formats, thereby standardizing data storage and processing across different systems and platforms within the aviation maintenance industry.
- Develop algorithms that can effectively analyze the quality of source data, adjusting for various requirements and system goals. These algorithms should support the automation of quality control in data transformations, enhancing the reliability and accuracy of the integrated data.
- Implement an analytics system that is preconfigured to calculate and optimize data transformation results based on specific maintenance and operational needs.

3. Research Methodology

The paper explores the conceptual framework of the AMTP, addressing the need for a standardized ETL process within the aviation maintenance industry on the basis of the next methodological approach:

- The paper leverages existing literature to establish the need for the AMTP, citing key challenges and gaps in current ETL processes as identified in previous studies. This foundational review supports the rationale behind the development of a more efficient and integrated system.
- The paper outlines a detailed design of the proposed platform, including the creation of a universal data ontology and predefined algorithms for data quality analysis. This indicates a developmental approach where theoretical constructs have been used to propose solutions tailored to the unique needs of aviation maintenance data integration.

These components suggest a methodological approach that, while primarily theoretical and design-focused, sets a strong groundwork for understanding and addressing the complexities associated with ETL processes in aviation maintenance.

To enhance the practical application and empirical grounding of the AMTP, future research activities will be directed toward the following areas:

- The simulation of ETL processes using synthetic and real-world data would allow for the evaluation of the platform's performance under controlled conditions, providing insights into its functionality and efficiency in various scenarios.
- It is crucial to implement the platform in real-world settings through collaborative projects with aviation maintenance organizations and aircraft operators.

- Comprehensive data collection and analysis should be conducted to assess the effectiveness of the platform. This should involve both quantitative metrics, such as integration speed and error rates, and qualitative feedback from end-users. Analytical methods, including statistical testing, will be essential to validate the results and ensure the reliability of the findings.
- Based on the feedback and data collected from initial testing and case studies, the platform should undergo iterative refinement. This process will involve adjusting the data ontology and transformation algorithms to better meet the users' needs and to cope with the complexities of diverse data formats.

By pursuing these research directions, the project can transition from a theoretical framework to a practical tool that significantly enhances ETL processes and data integration within the aviation maintenance industry, providing a robust evidence base to support its claims of improved efficiency and adaptability.

4. State of the Art of the Domain

The scientific literature on the state of the art in aviation MIS and data transformation highlights the critical role of digitization in revolutionizing aircraft maintenance and engineering.

Traditional approaches in aircraft maintenance heavily rely on manual processes and reactive strategies, leading to operational disruptions and financial losses. The shift towards digitization, predictive maintenance, and the use of digital twins are imperative for enhancing operational efficiency, reliability, and cost savings in the aviation industry [9, 20, 21]. Digitization of the aviation industry is fundamentally driven by the collection, processing, and analysis of data. However, the utility of this data depends entirely on its quality [22]. Research studies focus on developing autonomous systems for maintenance planning, fault prediction, and decision-making processes. These systems leverage data analytics, machine learning, and real-time data to enable proactive maintenance scheduling, reduce unscheduled ground time, and optimize maintenance strategies [23].

The scientific literature on aviation ETL processes highlights the critical role of data quality and the challenges associated with implementing effective ETL systems in the aviation industry. The case study [24] demonstrates the complexities involved in transitioning from paper-based to digital maintenance processes. Research [25] underscores the importance of assessing data quality characteristics at each stage of the ETL process. The case study [26] highlights the optimization of ETL processes at a major global airline addressed to improve operational efficiency. The survey [27] emphasizes the need for a systematic approach and the consideration of data quality factors.

Current MISs often struggle with interoperability, lack flexibility in adapting to varied data formats, fall short of maintaining consistent data quality, and still rely heavily on manual processes for data transformation.

The study addresses a critical gap in existing research by developing the AMTP, which targets the integration challenges of disparate MIS in the aviation industry filling significant voids in the current landscape of aviation maintenance technology.

5. Aviation Maintenance Transformation Platform Concept

5.1. Key Aspects of the Platform Concept

The aviation maintenance sector, particularly in the context of maintaining airworthiness and servicing aircraft and components, is characterized by a limited and finite number of key entities and mandatory attributes, which are established and regulated by industry standards and aviation authorities [18, 28]. Despite the vast variability in data formats and transformation methods, these core elements remain stable, making it crucial to align data management and ETL strategies with these foundational aspects [29].

Therefore, developing data management and ETL strategies that align with these stable core entities and attributes can offer a more reliable and sustainable approach to handling the complexities of aviation technical data transformation and integration.

The concept of the platform involves:

- Creating a universal data ontology for maintenance, repair, overhaul (MRO), and airworthiness for aircraft and components, possessing the necessary flexibility to store all possible source data formats.
- Developing predefined algorithms for analyzing the quality of source data by calculating certain data quality indicators (proposed by the author in [25]), managed through fine-tuning following the requirements of the target system and project goals.
- Establishing flexible settings for data transformation rules and quality indicators to meet transformation needs without altering the fundamental core of the data ontology.
- Developing the pre-configured analytics system designed for calculating data transformation results and solving optimization tasks (proposed by the author in [30]) to refine system settings and make management decisions.

The basic data ontology and associated data analysis systems will address the critical task of reusing accumulated knowledge and developments without significant changes when transitioning from project to project through platform configuration.

Data ontologies have proven to be valuable in the aviation technology industry, particularly in the areas of big data analysis, air transport network management, forecasting, and machine learning [31]. However, the field of aviation research is still grappling with issues related to the diversity, availability, tractability, applicability, and sources of data [32].

5.2. The Platform's Architecture

Figure 1 illustrates the structure of the platform concept, depicted using the UML components diagram format. The platform is divided at the top level into four subsystems (Modules).

The Extraction Module contains components responsible for defining the source system data and requirements, as well as for extracting the initial data. The subsystem must define a customized Source System Ontology Structure (component Source System Adaptor) based on the basic Predefined Ontology Structure. To achieve this, the customized ontology must describe the initial data structure and source system requirements and constraints regarding the order of data storage and processing (component Source System). The component Source System Extractor is responsible for extracting the initial data from the source system and uploading them into the Source System Ontology Structure.

The Target System Adaptor component of the Loading Module is responsible for defining the customized Target System Ontology Structure, based on the Target System Structure and Requirements (component Target System and the basic Predefined Ontology Structure. This customized ontology must describe the target data structure, target system requirements, and constraints regarding the order of data storage and processing.

Both customized ontologies must be created on top of the basic Predefined Ontology Structure (component Aircraft Maintenance Data Ontology of the Transformation Module). The basic ontology, as mentioned earlier in the paper, is built, managed, and revised on aviation standards, regulatory requirements from manufacturers and authorities, as well as best manufacturing practices.

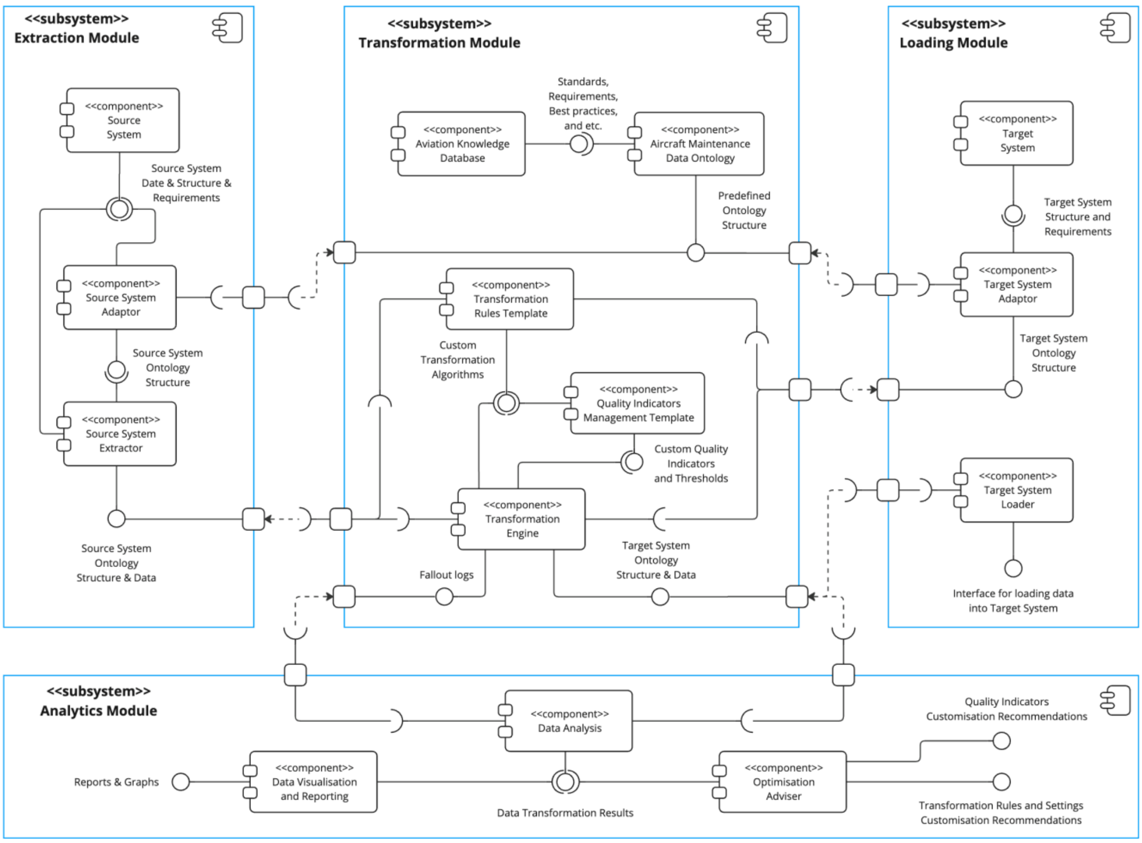

Figure 1: Comparison of a standard ETL process and a suggested ETL process.

The completeness and extensibility of the basic data ontology is a great complexity that needs to be addressed. Despite the finite nature of the basic data ontology, a significant challenge is the depth of detail in the aircraft and MRO basic data representations. Numerous parameters that are not mandatory from a regulatory standpoint may be required based on the project goals. To address this issue, it is necessary to develop an ontology data architecture that allows for the expansion of the ontology based on the basic structure, to accommodate the unique requirements of aircraft operators and maintenance organizations for both Source and Target systems.

Although Source and Target System Adapters components are nearly unique due to their strict linkage with the parameters and characteristics of the Source and Target Systems, they have the potential for subsequent reuse when implementing ETL projects that involve similar source and target systems.

The task of the Transformation Engine component of the Transformation Module is to transform the initial data loaded into the Source System Ontology into the Target System Ontology based on the defined transformation algorithms (component Transformation Rules Template) and the configured data quality indicators (component Quality Indicators Management Template). As a result, the initial data that has been successfully transformed is loaded into the target system through the Target System Loader component of the Loading Module. All data that fails validation due to not meeting the structural or quality requirements of the target system is recorded in special fallout log files.

According to the presented concept and original methodology, described by author in the study [33], it's assumed that the loading of initial data into the Source System Ontology and loading of transformed data from the Target System Ontology into the target system must be error-free. All the possible errors and data dropouts should be strictly registered in the fallout logs. This ensures the accurate calculation of data transformation quality indicators [25]. The Data Analysis component of the Analytics Module will implement all the necessary basic mechanisms for calculating data quality indicators, which will be managed through transformation algorithms and the settings of the Source and Target System Ontologies.

The project's limited resources present a complex optimization challenge for the platform. The Optimization Adviser component of the Analytics Module is designed to solve these optimization tasks. The result of the component's work includes recommendations for setting up Transformation Rules and Settings Customization, as well as Quality Indicators Customization.

The Data Visualization and Reporting component of the Analytics Module is designed to visualize transformation results and analytical data generated by the platform.

6. Discussion

The aviation industry is marked by its stringent safety requirements and the critical need for precise maintenance and operational integrity. The implementation of MIS is essential for ensuring airworthiness and efficiency in maintenance operations. However, the integration of diverse MIS across various stakeholders presents significant challenges, primarily due to the non-uniformity of data formats and the complex requirements for data processing. The proposed AMPT seeks to address these challenges by introducing a comprehensive, standardized solution for the ETL of data across different systems.

The core of the proposed platform is the development of a universal data ontology that can accommodate all potential source data formats found within the industry. This ontology is designed to serve as a robust framework for the uniform representation of MRO data, regardless of the originating MIS. By standardizing data representation, the platform facilitates more seamless integration and interaction between disparate systems, enhancing data consistency and reliability.

A critical aspect of the platform is its focus on maintaining high data quality. The platform includes predefined algorithms designed to analyze the quality of data transformation. These algorithms assess data on multiple dimensions, such as accuracy, completeness, and timeliness, ensuring that only data that meets stringent quality criteria are processed and integrated. This approach minimizes the risk of errors and enhances the reliability of maintenance decisions based on the integrated data.

The platform features a sophisticated analytics system pre-configured to evaluate the efficiency of the data transformation processes and to optimize these processes based on the specific needs of the project. This system allows users to calculate the effectiveness of data transformations, providing valuable insights that can be used to refine and adjust the ETL processes. By doing so, the platform not only improves its own operational efficiency but also enhances the overall performance of the MIS it supports.

Recognizing the dynamic nature of aviation maintenance needs, the platform is designed with a high degree of flexibility and adaptability. It allows for modular adjustments to its core algorithms and ontology structure, enabling it to easily adapt to new requirements or changes in industry standards without extensive overhauls. This adaptability is crucial for maintaining the long-term viability of the platform.

7. Conclusion

This paper presents the AMTP, a novel approach aimed at resolving the prevalent challenges in the ETL processes within the aviation maintenance industry. By introducing a universal data ontology, modular structure and predefined algorithms for data quality analysis, this platform seeks to increase the efficiency, flexibility, and adaptability of data management practices in aviation maintenance.

The successful deployment in diverse operational environments confirms the AMTP's capacity to meet the evolving needs of the aviation maintenance sector.

These results and conclusions reinforce the original hypothesis that a standardized, automated approach to data integration can significantly improve the efficiency and reliability of maintenance operations in the aviation industry.

The AMTP is set to provide a foundation for further innovations in aviation data management. The development of the data ontology architecture and the integration of advanced technologies such as AI and machine learning will further enhance the platform's capabilities, making it a pivotal tool in the advancement of automated data transformation processes in aviation maintenance.

Acknowledgments

For the preparation of this article, I would like to express my gratitude to my academic supervisors: Boriss Misnevs and Igor Kabashkin. The feedback and recommendations received from them have been invaluable in improving the quality and content of this work.

References

[1] X. Lian, "Application Prospects of New Information Technology in Aviation Maintenance Industry," in Proceedings - 2021 7th International Symposium on Mechatronics and Industrial Informatics, ISMII 2021, 2021. doi: 10.1109/ISMII52409.2021.00068.

[2] M. Zieja, H. Smoliński, and P. Gołda, "Information systems as a tool for supporting the management of aircraft flight safety," Archives of Transport, vol. 36, no. 4. 2015. doi: 10.5604/08669546.1185211.

[3] "Aviation Data Migration," URL: https://bytzsoft.com/services/aviation-data-migration/.

[4] Y. M. Chinyuchin, E. D. Gerasimova, and A. I. Zhelannikov, "Principles of Information Management Implementation in the Processes of Aircraft Technical Maintenance," Russian Aeronautics, vol. 64, no. 1, 2021, doi: 10.3103/S1068799821010025.

[5] U. Durak, J. Becker, S. Hartmann, and N. S. Voros, Advances in aeronautical informatics: Technologies towards flight 4.0. 2018. doi: 10.1007/978-3-319-75058-3.

[6] Z. Hong-bao, "A preliminary discussion on the evolution and development of MIS," East China Economic Management, 2001, URL: https://api.semanticscholar.org/CorpusID:184364559

[7] N. Bisset, "MIS trends and demands: software," 2012, URL: https://api.semanticscholar.org/CorpusID:113005219

[8] T. Bo, "The Design and Realization of MIS for Airforce Aeronautical Servicing Material," Journal of Changsha Aeronautical Vocational and Technical College, 2003, URL: https://api.semanticscholar.org/CorpusID:111954944

[9] I. Alomar and I. Y. Jackiva, "Digitalization in aircraft maintenance processes," Aviation, vol. 27, no. 2, 2023, doi: 10.3846/aviation.2023.18923.

[10] E. Kivanç, Ö. Vayvay, and Z. T. Kalender, "Digital Transformation Approaches for Aircraft Maintenance Operations," 2021. doi: 10.4018/978-1-7998-2402-2.ch011.

[11] S. Chang, Z. Wang, Y. Wang, J. Tang, and X. Jiang, "Enabling Technologies and Platforms to Aid Digitalization of Commercial Aviation Support, Maintenance and Health Management," in Proceedings of 2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, QR2MSE 2019, 2019. doi: 10.1109/QR2MSE46217.2019.9021222.

[12] R. Arnaldo Valdes and V. F. Gómez Comendador, "Aviation 4.0: More safety through automation and digitization," in WIT Transactions on the Built Environment, 2018. doi: 10.2495/SAFE170211.

[13] O. Candell, R. Karim, and A. Parida, "Development of information system for e-maintenance solutions within the aerospace industry," International Journal of Performability Engineering, vol. 7, no. 6, 2011.
[14] E. al. Nityash Solanki, "Challenges of Big Data Technology in Aviation Management," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 9, 2023, doi: 10.17762/ijritcc.v11i9.9956.
[15] N. Berkani and L. Bellatreche, "A variety-sensitive ETL processes," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2017. doi: 10.1007/978-3-319-64471-4_17.
[16] G. Burmester, H. Ma, D. Steinmetz, and S. Hartmannn, "Big Data and Data Analytics in Aviation," in Advances in Aeronautical Informatics: Technologies Towards Flight 4.0, 2018. doi: 10.1007/978-3-319-75058-3_5.
[17] R. M. Keller, "Ontologies for aviation data management," in AIAA/IEEE Digital Avionics Systems Conference - Proceedings, 2016. doi: 10.1109/DASC.2016.7777971.
[18] S. Bruni, M. Freiman, C. Weiss, D. Ward, S. Lynch, and K. Kay, "A 7-dimensional framework for technical data in high-intensity vital environments and its application to aircraft maintenance," in Proceedings - 2020 IEEE International Conference on Cognitive and Computational Aspects of Situation Management, CogSIMA 2020, 2020. doi: 10.1109/CogSIMA49017.2020.9216152.
[19] A. Simitsis, S. Skiadopoulos, and P. Vassiliadis, "The History, Present, and Future of ETL Technology," in CEUR Workshop Proceedings, 2023.
[20] Sander de Bree, "Leveraging data for digital aircraft airworthiness management & MRO", URL: https://www.linkedin.com/pulse/leveraging-data-digital-aircraft-airworthiness-mro-sander-de-bree/.
[21] Sander de Bree, "The Digital Transformation in Aircraft Maintenance & Engineering: Driving Efficiency and Cost Savings,", URL: https://www.linkedin.com/pulse/digital-transformation-aircraft-maintenance-driving-cost-de-bree-mt5fc.
[22] Y. Liu, "Aircraft maintenance information system design and verification," 2019, URL: http://theses.insa-lyon.fr/publication/2019LYSEI133/these.pdf
[23] M. J. Scott, W. J. C. Verhagen, M. T. Bieber, and P. Marzocca, "A Systematic Literature Review of Predictive Maintenance for Defence Fixed-Wing Aircraft Sustainment and Operations," Sensors, vol. 22, no. 18. 2022. doi: 10.3390/s22187070.
[24] Rich O"Mara, "Case Study: Thomas Cook eTechLog (ETL / ELB) - Aircraft IT", URL: https://www.aircraftit.com/articles/thomas-cook-etechlog-etl-elb/.
[25] M. Pivovar, L. Rozhkova, B. Misnevs, and I. Pticina, "Data Quality Indicators of ETL in the Process of Implementation of Information Systems for Aircraft Maintenance and Operation," in Lecture Notes in Networks and Systems, Springer Science and Business Media Deutschland GmbH, 2021, pp. 130–139. doi: 10.1007/978-3-030-68476-1_12.
[26] Gajanan Gaidhane, "Optimizing ETL Processes at a Major Global Airline", URL: https://www.datavail.com/blog/optimizing-etl-processes-at-a-major-global-airline/.
[27] A. Kabiri and D. Chiadmi, "Survey on ETL processes," J Theor Appl Inf Technol, vol. 54, no. 2, 2013.
[28] SkyBrary, "Safety Regulation", URL: https://skybrary.aero/articles/safety-regulation.

[29] T. Hoppe, H. Eisenmann, A. Viehl, and O. Bringmann, “Shifting from data handling to knowledge engineering in aerospace industry,” in 2017 IEEE International Symposium on Systems Engineering, ISSE 2017 - Proceedings, 2017. doi: 10.1109/SysEng.2017.8088312.
[30] M. Pivovar, “ETL for Aviation Maintenance and Operations as a Multi Objective Optimization Task,” in Lecture Notes in Networks and Systems, 2022. doi: 10.1007/978-3-030-96196-1_14.
[31] S. H. Chung, H. L. Ma, M. Hansen, and T. M. Choi, “Data science and analytics in aviation,” Transportation Research Part E: Logistics and Transportation Review, vol. 134. 2020. doi: 10.1016/j.tre.2020.101837.
[32] M. Z. Li and M. S. Ryerson, “Reviewing the DATAS of aviation research data: Diversity, availability, tractability, applicability, and sources,” J Air Transp Manag, vol. 75, 2019, doi: 10.1016/j.jairtraman.2018.12.004.
[33] M. Pivovar, B. Misnevs, and I. Pticina, “Methodology for Calculating ETL Indicators in the Process of Implementation of Aircraft Maintenance Information Systems,” in Lecture Notes in Networks and Systems, vol. 117, Springer, 2020, pp. 44–53. doi: 10.1007/978-3-030-44610-9_5.