Aerospace Qualification Services Knowledge Graph: A Leap towards Enhanced Data Management

Raed Awill[1,2], Wajahat Ali Khan[1], Maqbool Hussain[1], Sadiq Zada[1,2] and Ben Anderson[2]

[1]*University of Derby, Derby, UK*
[2]*AddQual LTD, Derby, UK*

Abstract
In the field of aerospace product qualification services, AddQual's operation relies on the effective coordination of data from diverse sources. This is presently achieved through a relational model, which has displayed numerous limitations, including the lack of semantic enrichment, data heterogeneity, and horizontal data evolution. To overcome these challenges, we meticulously designed, developed, and implemented an Aerospace Qualification Services Knowledge Graph (AQS-KG), moving from the existing relational model of AddQual. During this transition, we devised OwlFly, an in-house tool specifically designed to facilitate the mapping between the two models. This solution adeptly manages complex aerospace data, precisely aligning with AddQual's strategic objectives, which include a plan to increase from 1000 scanned parts with AQS-KG in 2022 to 20000 parts by 2026, a substantial reduction in carbon emissions, and introducing automation in the repair procedures. This research paves the way for forthcoming innovations, driving AddQual's transition to adapt a digital twin framework.

Keywords
Knowledge Graph, Data Heterogeneity, Relational Model, Aerospace Industry, Semantic Mapping

1. Introduction

AddQual receives shipments of various jet engine parts from different clients for feature measurement and reporting. The clients had numerous engine sets of blades that had suffered significant wear. Consequently, they require further information to assess dimensional criteria and understand defects. They seek insights on causes and solutions for comprehensive part analysis. To respond to their needs, the aim is to identify repairable parts based on the gathered data. AddQual serves diverse clients globally, resulting in varying data representations. Heterogeneity arises as clients use different terms; e.g., 'Bucket' or 'NGV' for a nozzle vane blade and 'airfoil' for an aerofoil in America. The current approach of detailed inspection plans within the relational model leads to conflicts in data storage. Its lack of semantic enrichment and trouble handling data heterogeneity, especially with diverse component features, necessitate frequent schema updates and horizontal database expansion. The designed AQS-KG outperforms the limitations of the relational model [1] by streamlining data integration and accommodating diverse formats. Its flexible, self-evolving design promotes horizontal development and interoperability.

SemREC'23: THE 22ND INTERNATIONAL SEMANTIC WEB CONFERENCE, City of Athens, Greece
100470771@unimail.derby.ac.uk (R. Awill); w.khan@derby.ac.uk (W. A. Khan); m.hussain@derby.ac.uk (M. Hussain); sadiq.zada@addqual.com (S. Zada); ben.anderson@addqual.com (B. Anderson)

2. Methods

Our application of AQS-KG followed an iterative, four-phase process, designed for continuous refinement of the graph. The process incorporated Data Acquisition (DA), Knowledge modeling and Representation (KMR), Data Persistence (DP), and Data Visualization (DV), with an emphasis on graph validation across the KMR phase. Throughout the DA phase, we prepared the company's data by examining the existing model and collaborating with individuals across various departments. This stage involved defining the project's scope and identifying data sources. In the KMR phase, we created a robust, semantically organized AQS-KG. Concurrently, we developed OwlFly in Java and Apache Jena, which creates a KG model and uploads instances from any data source. It establishes a connection with databases, e.g., PostgreSQL, and extracts triples to form the AQS-KG model. The data extracted is used to define an intermediate data model (IDM), map it to the triples, and load individuals into RDF, enabling precise and efficient integration of data into the graph structure. To enhance the capabilities of AddQual, we utilized a graph database during the DP and DV phases. This not only secured the data storage but also granted us the capacity to visualize it in graphs. Such a strategic approach equipped AddQual to generate insightful reports about their data, such as pinpointing parts susceptible to significant defects.

3. Results and Conclusion

Since the inception of AQS-KG at AddQual, we've noticed a cut in the time needed to execute complex queries by an impressive 35%, enhancing the data querying process. This efficiency increase is crucial as we plan to increase the number of processed parts from 1,000 in 2022 to 20,000 by 2026. It becomes even more significant when considering the rich semantics of the 34 entities and the 31 relationships we have established within the graph, in contrast to the 12 relations and 5000 tuples present in the previous relational model. Notably, we identified 100 features with differing names but identical meanings. To address this data heterogeneity, rather than using disparate fields in the relational model, we employed the graph's semantic labeling to unify similar entities and individuals. When considering future integration with machine learning and artificial intelligence techniques, we anticipate an additional 20% increase in the accuracy of our predictive capabilities. This could potentially lead to a reduction in downtime and unnecessary expenditure. Furthermore, the current yield for complex component repair parts stands at approximately 50%. We anticipate that integrating the knowledge graph with our robotics systems could potentially increase this yield to 70-75% and concurrently achieve a 20% reduction in associated costs. Additionally, we foresee that integration of the graph model will improve AddQual's digital twin repair approach, leading to the maximization of operating performance and reducing fuel consumption by 0.5%.

References

[1] S. Batra, C. Tyagi, Comparative analysis of relational and graph databases, International Journal of Soft Computing and Engineering (IJSCE) 2 (2012) 509–512.