FEDERAL UNIVERSITY OF TECHNOLOGY - PARANÁ
POSTGRADUATE PROGRAM IN MECHANICAL AND MATERIALS ENGINEERING

ARTHUR EDUARDO DE FREITAS LAROCCA

DE DADOS PARA VALOR: UM MODELO PARA O GERENCIAMENTO DE DADOS EXPERIMENTAIS COLETADOS EM TESTES DE DURABILIDADE DURANTE O PROCESSO DE DESENVOLVIMENTO DO PRODUTO

DISSERTATION

CURITIBA
2021

ARTHUR EDUARDO DE FREITAS LAROCCA

DE DADOS PARA VALOR: UM MODELO PARA O GERENCIAMENTO DE DADOS EXPERIMENTAIS COLETADOS EM TESTES DE DURABILIDADE DURANTE O PROCESSO DE DESENVOLVIMENTO DO PRODUTO

From data to value: a model to manage experimental data collected from durability tests within the product development process

Dissertation presented to the Postgraduate Program in Mechanical and Materials Engineering of the Federal University of Technology – Paraná, as fulfillment of the requirements for Master degree in Engineering.

Concentration area: Manufacturing Engineering.

Supervisor: Prof. Milton Borsato, Ph.D.

CURITIBA

2021

Ministério da Educação
Universidade Tecnológica Federal do Paraná
Campus Curitiba

ARTHUR EDUARDO DE FREITAS LAROCCA

DE DADOS PARA VALOR: UM MODELO PARA O GERENCIAMENTO DE DADOS EXPERIMENTAIS COLETADOS EM TESTES DE DURABILIDADE DURANTE O PROCESSO DE DESENVOLVIMENTO DO PRODUTO

Trabalho de pesquisa de mestrado apresentado como requisito para obtenção do título de Mestre Em Engenharia da Universidade Tecnológica Federal do Paraná (UTFPR). Área de concentração: Engenharia De Manufatura.

Data de aprovação: 25 de Junho de 2021

Prof Milton Borsato, Doutorado - Universidade Tecnológica Federal do Paraná

Prof David Carlos Romero Diaz, Doutorado - Tecnológico de Monterrey (Itesm)

Prof John Stark, Doutorado - Plm Institute - (Product Lifecycle Management)

Documento gerado pelo Sistema Acadêmico da UTFPR a partir dos dados da Ata de Defesa em 25/06/2021.

To Manuela

ACKNOWLEDGEMENTS

Throughout this research I have received a great deal of support and assistance.

First and foremost, I would like to thank my supervisor, Professor Milton Borsato, whose expertise was invaluable for the development of this work. Professor Borsato consistently allowed this research to be my own work, but precisely steered me in the right direction and encouraged me during all this process.

I would like to acknowledge Volvo do Brasil and all my colleagues from Volvo Group who collaborated and supported this research. I would particularly like to single out my managers Santiago Moran and Rafael Barreto, who supported me and this research from its beginning. Thank you also to the durability community members within Volvo Group who invested their time and expertise in this work.

I must express my very profound gratitude to my friend Pablo Kubo, who not only supported this work but also inspired and encouraged me to pursue this journey. Thank you, Pablo.

I also acknowledge the contribution from both the Federal University of Technology - Paraná (UTFPR) and the Federal University of Paraná (UFPR) in my personal and professional development.

Also, I would like to thank my parents, Paulo and Roselis, for always being there for me.

Finally, a big thank you to my wife Luna, whose support (both academic and emotional) was essential for this work.

RESUMO

Junto à pressão da concorrência, regulamentações e a necessidade geral de melhorias, as companhias começaram a reconhecer os seus dados como um ativo de negócios. Embora as empresas tenham mais dados do que nunca à disposição, convertê-los em conclusões significativas e valor é um desafio. Dentro do contexto do processo de desenvolvimento de produtos (PDP) e de engenharia colaborativa, o gerenciamento de dados experimentais foi identificado como uma oportunidade de pesquisa. Apesar de vários estudos abordarem a gestão de dados e colaboração em ambientes de tecnologias auxiliadas por computador (CAx), o gerenciamento dos dados medidos foi notado como um tópico escasso na literatura. Além disso, dentre todas as verificações e validações realizadas no desenvolvimento de produtos, análises adicionais em uma empresa multinacional de manufatura indicaram uma maior prevalência de testes de durabilidade realizados em protótipos. Sob esta ótica, o objetivo deste estudo é propor um modelo para gerenciar arquivos de medição coletados em testes de durabilidade que fomente a reutilização dos mesmos através de diferentes projetos e ao longo da organização de desenvolvimento de produto. Para tal, esta pesquisa qualitativa aplicada trata este tema adotando a abordagem metodológica de Design Science Research Methodology (DSRM). Baseado em referências da literatura e considerando os obstáculos que impedem a reutilização dos dados, um modelo de dados lógico foi desenvolvido e documentado como um diagrama de classe na notação UML (Unified Modeling Language). Tal artefato encompassa um conjunto de propostas que foram desenvolvidas para ensejar um ciclo de vida completo para os dados mais recorrentes em testes de durabilidade, tais quais: força, deformação, aceleração e deslocamento. A principal estratégia do modelo proposto é a utilização extendida de metadados nos arquivos de medição. Isto permetiria que as informações relevantes ao experimento, as quais normalmente são encontradas de forma não estruturadas em relatórios de teste, sejam anexadas e estruturadas nos próprios sinais medidos. Diversos conceitos foram desenvolvidos baseados nesta estratégia, como o de diferentes versões para os dados experimentais e a nova abordagem para a documentação da posição da instrumentação baseada em coordenadas cartesianas. Uma prova de conceito fora demonstrada usando dados e informações reais para as partes interessadas da indústria e do meio acadêmico. Pesquisas de satisfação sugeriram que o artefato proposto é completo e aborda um tema relevante. Essas pesquisas também sugeriram que o modelo proposto é melhor que o modelo (ou forma de trabalho) atual nos quesitos de consitência interna, nível de detalhe e robustez. A relevância deste trabalho reside na sua natureza aplicada e potenciais contribuições para um processo de desenvolvimento de produtos mais colaborativo e eficiente no que diz respeito à verificação e validação de novos produtos.

Palavras-chave: Gestão de Dados. Dados Experimentais. Arquivos de Medição. Durabilidade. Processo de Desenvolvimento de Produtos. Engenharia Colaborativa.

ABSTRACT

Coupled with the pressure from competition, regulations, and the overall need for improvements, organizations started to recognize their data as a business asset. Although companies have more data than ever at their disposal, actually deriving meaningful insights and value from them is easier said than done. Within the context of the product development process (PDP) and collaborative engineering, the management of experimental data has been identified as a research gap. While many studies have tackled data and collaboration based on computer-aided technologies (CAx) environments, the management of the measured data has been observed as a scarce topic in the literature. Moreover, among all verifications and validations performed during development, further analyses in a multinational manufacturing company indicated a higher prevalence of durability tests performed in prototypes. In this concern, the main objective of this work is to propose a model to manage measurement files collected in durability tests which would enable and foster their reuse across projects and throughout the product development organization. For this purpose, this qualitative applied research tackles this challenge based on the methodological framework of the Design Science Research Methodology (DSRM). Built upon prior literature and considering the constraints that hinder data reuse, a logical data model was developed and documented as a class diagram in the Unified Modeling Language (UML). This artifact comprises a set of approaches that have been designed to enable the complete lifecycle of the most recurrent data in durability testing: load, strain, acceleration, and displacement. The main strategy of the proposed model is the enhancement of metadata into the measurement files, so important information, that is usually unstructured within test reports, is stamped and structured into the test signals. Several concepts were developed based on this strategy, such as the concept of data versions and the novel approach for the documentation of the instrumentation positioning as cartesian coordinates. A proof-of-concept has been demonstrated using real measurement files and other related information to key academic and industrial stakeholders, who also took part in the development and evaluation of the solution. Feedback and satisfaction surveys suggested that the proposed model is complete and faithful to real-world phenomena. They also suggested that the proposed artifact is better than the current model (or way of working) regarding its internal consistency, level of detail, and robustness. The relevance of this work lies in its applied nature and potential contributions towards a more collaborative and efficient product development process regarding the verification and validation of new products.

Keywords: Data Management. Experimental Data. Measurement File. Durability. Product Development Process. Collaborative Engineering. Knowledge Management.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

BP	Bibliographic Portfolio
CAD	Computer-Aided Design
CAE	Computer-Aided Engineering
CAM	Computer-Aided Manufacturing
CAx	Computer-Aided Technologies
CE	Concurrent Engineering
CMM	coordinate measuring machines
CSCD	Computer Supported Collaborative Design
DAQ	Data Acquisition
DM	Data Management
DMU	digital mockup
DoE	Design of Experiment
DS	Design Science
DSR	Design Science Research
DSRM	Design Science Research Methodology
FMEA	Failure Mode and Effect Analysis
GPS	Global Positioning System
IS	Information Systems
IT	Information Technology
LDM	Logical data model
LHS	left-hand side
NPD	New Product Development
PDA	Product Data Analytics
PDF	Portable Document Format
PDM	Product Data Management
PDP	Product Development Process
PLM	Product Lifecycle Management
QFD	Quality Function Deployment
RDM	Research Data Management
RHS	right-hand side
SDM	Simulation Data Management
UML	Unified Modeling Language

SUMMARY

1 INTRODUCTION

In a continually changing market, in which innovation is considered as the driving force of rapid changes and competitive advantage, companies have to improve their product development process (PDP) to stay profitable (MINAVAND; LORKOJOURI, 2013; SU, 2014; TAN; VONDEREMBSE, 2006). While Ernst (2002) summarized the most important findings and success factors of new product development through a broad literature review, more recent researches have explored the extensive context of Concurrent Engineering (CE) (EBRAHIMI, 2011; HAMBALI et al., 2009; HAUG, 2012, p. 3; SAPUAN; OSMAN; NUKMAN, 2006). In the automotive industry, for instance, this approach is shifting the traditional and sequential design-build-test cycle to a combined and synchronized task approach, mainly driven by front-load analysis, simulation, and testing (MILBURN, 2004).

Within the context of CE, Borsato and Peruzzini (2015) recognized the concept of collaborative engineering and explored the application of Computer Supported Collaborative Design (CSCD). The authors highlighted the importance of an integrated approach to connect different software tools in product design, simulation, and manufacturing to a successful implementation of CSCD.

Coming alongside, new technologies are gathering more data than ever before. Coupled with the pressure from competition, regulations, and the overall need for improvements, organizations started to consider data as a business asset. Yet, many are still looking for better ways to extract value from their data (FISHER, 2009, p. 16). LaValle et al. (2011) connect performance and the competitive value of analytics in a business context. They found that top-performing organizations (companies identified for substantially outperforming their industry peers) use analytics tools five times more than lower performers. Their research also pointed out that rather than data quality and technology, the adoption barriers are mostly managerial and cultural.

Many authors have tackled the importance of data management (DM) within the New Product Development (NPD) (ANDRADE-VALBUENA; MERIGO, 2018; NALLUSAMY et al., 2015; ZHAN et al., 2018; YU; YANG, 2016) and recent researches focused on the so-called “big data” from a more holistic perspective (ARDITO et al., 2019; KALANTARI et al., 2017; MISHRA et al., 2018; RIALTI et al., 2019; CUI; KARA; CHAN, 2020). However, within this context, a literature review identified the management of experimental data as a research opportunity. While many studies have tackled data and collaboration based on computer-aided technologies (CAx) environments, very few studies examined the management of the measured data collected during the verification and validation stages of a product. Even though studies regarding experimental data management are available, its association with the PDP is scarce in the literature. In this concern, this work aims to contribute to the literature by providing more insight into how experimental data shall be integrated with the framework of CE and leveraged as a catalyst towards a more efficient PDP.

1.1 RATIONALE

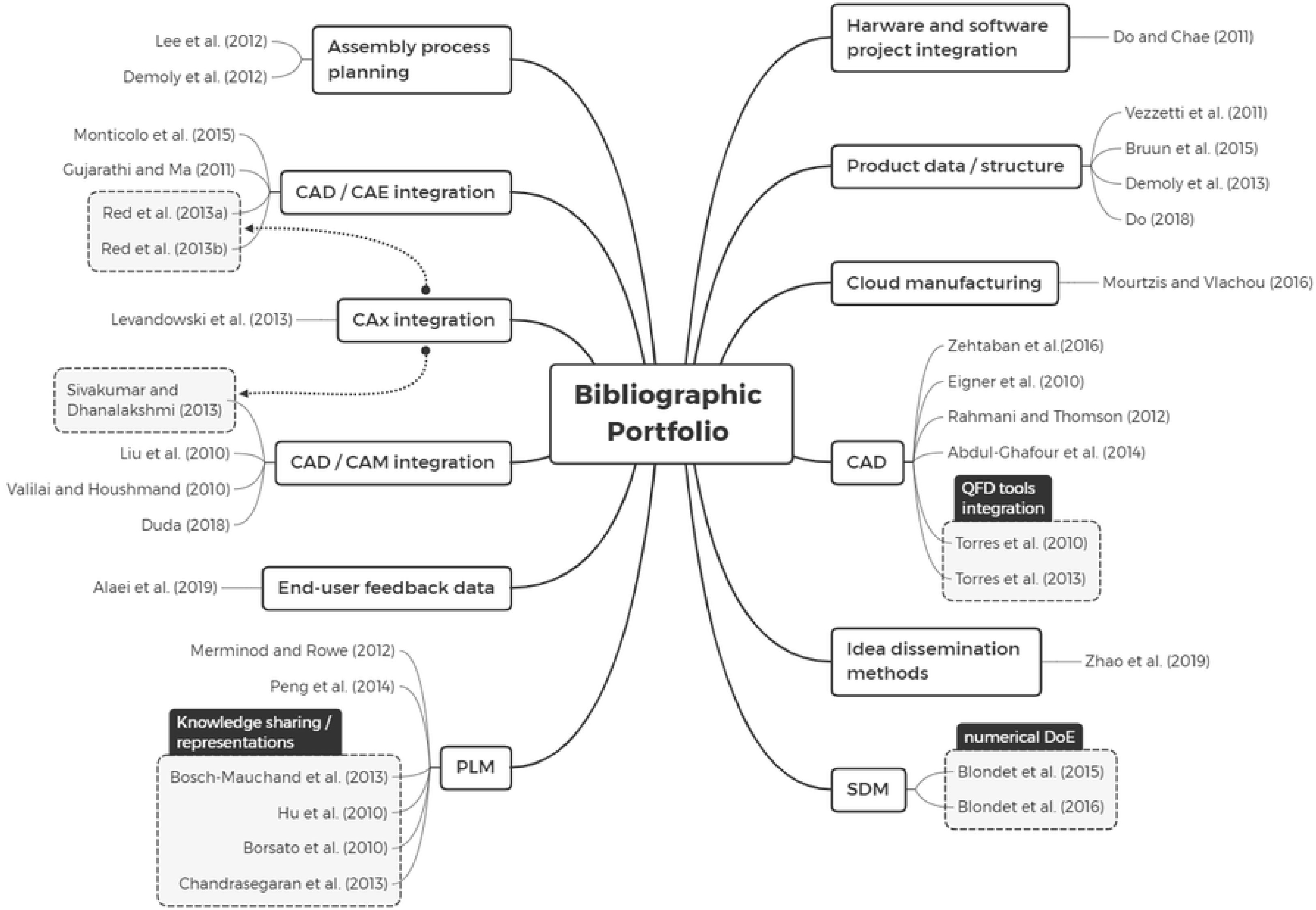
The relevance of the experimental data is justified by the high cost of product verification and validation – which, together with the prototype building, can allocate a significant amount of resources in the development of new products in the automotive industry (PASCOAL; SILVA, 2010; ROZENFELD; AMARAL, 2006, p. 378). Moreover, despite the increase of simulation, testing costs have risen in this industry due to the increase in product complexity and variety (BARTELS; ZIMMERMANN, 2009). Finally, while many of the studies explored methods to reduce test samples or complete prototypes as a way of minimizing PDP costs (CHELST et al., 2001; BOBER, 2005; PASCOAL; SILVA, 2010), Steyer, Voight and Hering (2005), and Braden and Harvey (2014) stress the need for data-driven approaches led by knowledge databases and data mining solutions.

Within the scope of this dissertation, a research article entitled “Data Management within New Product Development and Collaborative Engineering: a Bibliometric and Systemic Analysis” has been published in the VINE Journal of Information and Knowledge Management Systems by Larocca et al. (2021). In this paper, bibliometric and systemic analyses have been carried out using the methodological procedure ProKnow-C, which provides a structured framework for the literature review. A bibliographic portfolio (BP) was consolidated with 33 papers that represent the state of art in the subject. One of the main implications of this work was to provide a fresh and relevant source of authors, journals, and studies for researchers and practitioners interested in the domain of data management applied to NPD and collaborative engineering. Moreover, it was expected that the identified research gaps and opportunities may contribute to future studies tackling the efficiency in the PDP, which is the case for this dissertation.

The systemic analysis revealed that despite the variety of studies, from specific case studies to more holistic perspectives, most of the authors highlight the importance of data management within concurrent and collaborative engineering practices. Moreover, a strong association has been noted between the research subject and the concept of Product Lifecycle Management (PLM) – which is defined as the “business activity of managing, in the most effective way, a company’s products all the way across their lifecycles; from the very first idea for a product all the way through until it is retired and disposed of” (STARK, 2015, p. 1).

Most recent researches within the bibliographic portfolio (BP) indicate new trends and paradigm shifts in this area of research, tackling subjects such as the Internet of Things (IoT), cloud computing, big data analytics, and digital twin. However, most of the selected papers proposed frameworks regarding the management of Computer-Aided Design (CAD) data. At the same time, many researchers explored the contributions of data management to knowledge capture and re-use in the PDP. Concerning the applied methodologies, ontology-based approaches were the most recurrent methods proposed by the authors within the BP. Furthermore, researches in Simulation Data Management highlighted the importance of enhancing metadata of data models and ontologies (LAROCCA et al., 2021). The diagram in Figure 1 illustrates the researches and their respective topics, within the BP.

Figure 1 – Diagram illustrating the covered topics within the bibliographic portfolio



Source: Larocca et al. (2021)

Research gaps identified within the portfolio included the lack of data automation and the absence of a common framework for the overall integration of systems and tools of the PDP. However, from a general outlook of the BP, the management of experimental data was noted as a knowledge boundary. While many studies have tackled data and collaboration based on computer-aided technologies (CAx) environments, no study examined the management of the measured data collected during the verification and validation stages of a product. Even though studies regarding experimental data management are available, its association with the PDP is scarce in the literature.

Moreover, among all verifications and validations performed during development, further analyses in a multinational manufacturing company (further detailed in section 4.3) indicated a higher prevalence of durability tests. In this regard, a consultation in a collaborating company revealed some important perspectives:

- The measurement files are perceived as the foundation of many of the tests performed in a durability concern. If they were properly archived and identified, they would have the potential to eliminate (some of) the need for a new physical test for future demands/projects.
- The use of past measurements, when occurred, was only possible given previous experiences from the team members involved in the task. Simply put, this knowledge is embedded in the people that took part in the assignment, but not in the organization itself.

In these concerns, this research seeks to answer the following question: How durability measurement files collected within the PDP should be structured so they could be used up to their potential and consolidated as a business asset?

This research is positioned within the Smart Manufacturing Program from the Postgraduate Program in Mechanical and Materials Engineering of the Federal University of Technology - Paraná (UTFPR). This program is detailed as several work packages that are required to fulfill specifics demands and their respective challenges, based on NGMTI (2005). This work is connected to three (out of ten) demands and to six challenges (two for each demand), which are described below:

Demand #1: Smart visualization and conceptualization

One-click access to all the possible needed analyses, e.g. requirements analysis, performance evaluation, manufacturability, material behavior. It encompasses the following challenges:

- Integrate visualization with modeling and simulation, analysis, and knowledge-based tools to drive the manufacturing and planning of the process (not just visualize product concepts).
- Enable a smart evaluation of many product options, simulate different processes, and support the interactive evaluation of multiple perspectives.

Demand #4: Guiding Factors for Knowledge-Based Planning and Design

A standardized knowledge capture environment, capable of directing product planning and design, which includes the following challenges:

- Develop methods and frameworks to capture, maintain and manage knowledge with quality.
- Access to data that represents corporate knowledge.

Demand #8: Complete Virtual Product Documentation

Virtual documentation that provides a complete technical data package that captures product data as planned, designed, built and used. It includes the following challenges:

- Identify a complete and flexible pattern to capture and communicate the product model definition;
- Store and maintain data in the long term;

1.2 OBJECTIVE

By providing more insight into how experimental data shall be integrated with the framework of CE and used as a catalyst for innovation, the long term goal of this work is a series of research contributions that would pave the way towards a more collaborative and efficient product development process regarding the verification and validation of new products.

1.2.1 General Objective

The general objective of this research is to propose a model to manage measurement files collected in durability tests within the PDP which would enable their reuse across projects and throughout the product development organization.

1.2.2 Specific Objectives

To achieve the proposed objective, this research will address the following specific objectives:

1) Identify methods and approaches for the management of the experimental data within a research context.
2) Identify the main constraints regarding the use (and mainly reuse) of durability measurement files within the PDP.
3) Structure a logical model to handle measurement files collected in durability tests within the PDP.
4) Identify potential technologies that could operationalize the proposed model.
5) Simulate the applicability of the model based on real data.
6) Assess the advantages/disadvantages of the proposed model based on the demonstrated results.

1.3 DISSERTATION OUTLINE

The remainder of this dissertation is structured as follows. Section 2 provides a conceptual background on the concept of data and knowledge management, the context for durability testing in the industry, and a review of research data management. While section 3 explores the theoretical framework of the proposed research, section 4 details the proposed research methodology for reaching each one of the delineated specific objectives. Section 5 presents the proposed model while Sections 6 and 7 the obtained results. Finally, section 8 presents the conclusions, final remarks, and suggestions for future works.

2 LITERATURE REVIEW

The subsections below complement the literature review from Larocca et al. (2021) by focusing on data and knowledge management, the concept of durability and the context for durability testing in the industry, and key research in the domain of research data management.

2.1 DATA AND KNOWLEDGE MANAGEMENT

Data management is defined by Brackett and Earley (2009, p. 4) as being the business function of planning for, controlling, and delivering data and information assets. It is an important multi-disciplinary function to organizations regardless of their size and purpose. Defined as the discipline that fosters organizational learning and the management of intellectual capital, Brackett and Earley (2009, p. 3) connect knowledge management to data management, as they are both dependent on high-quality information. Through an empirical approach, Grillenberger and Romeike (2017) systematically determined the central aspects of data management and presented a model of its key concepts. This model was divided into its practices, design principles, mechanics, and core technologies. Given its function of representing current developments and research progress, the aspect of the core technologies was the only one perceived as not being constant over time (GRILLENBERGER; ROMEIKE, 2017). Finally, through an overview of highly cited reviews, recent research by Shah, Naeem and Bhatti (2020) explored the emerging trends of data practices and data management.

Knowledge, information, and data are conceptually different though, too often, the words are used interchangeably (NONAKA; TAKEUCHI, 1995, p. 58). Marchand (1998) defines information as data provided with context, thus situational relevance. While knowledge derives from interpreting the incoming and circulating information flow, leading to descriptive understandings and prescriptive beliefs.

According to Plessis (2007), besides having no formalized way to access it, organizations are generally not conscious of the amount of tacit knowledge available to them. Knowledge management (KM) tools can assist in codifying tacit knowledge to make it explicit and readily available for future applications. Overall, KM plays a major role in facilitating collaboration, which supports the sharing of tacit knowledge.

Plessis (2007) argues that the second major role of knowledge management is related to explicit knowledge. Although it does not play such an important role as tacit knowledge (given that explicit knowledge is usually accessible to competitors), explicit knowledge is also an important component of innovation. In developed science processes, explicit knowledge is strongly present in the research and development process with a rich exchange with tacit knowledge. However, this process demands the capability to convert tacit and explicit product and process knowledge into explicit models. While the knowledge from upstream research and development discoveries is generally tacit in nature, Cardinal, Alessandri and Turner (2001) and Scarbrough (2003) suggest that knowledge downstream in the value chain is most explicit and codifiable. For this reason, they argue that organizations need to build resources and capabilities that will allow them to capture and codify knowledge and product development routines.

Within the context of the new product development (NPD), Pitt and MacVaugh (2008) present a holistic interpretation of the scope of knowledge management. According to the authors, no matter the NPD model evaluated, much of the required organizational knowledge is distributed, rather than centralized. Tsoukas (1996), Alavi and Tiwana (2002), and Kristian (2002) add that much of this knowledge is also tacit, located in the minds of relatively few specialists. Pitt and MacVaugh (2008) conclude that knowledge and information reside at different organization levels and locations that are accessible with varying degrees of complexity. Explicit, codified information is suitable for computer-based capture, storage, and dissemination, where IT solutions may play a valuable role.

Following the studies from Nonaka and Takeuchi (1995) and Marchand (1998), Pitt and MacVaugh (2008) pointed out the four forms of knowledge conversion that constitute collective learning – shown in Table 1.

Table 1 – Process of knowledge conversion

Transition from	Transition to	
	Information (explicit knowledge)	Knowledge (tacit knowledge)
Information (explicit knowledge)	(Re)combination and diffusion by acquiring, analyzing, and organizing documents, files, messages, etc. into databases and other forms of accessible repository and publishable report intended for extended access.	Internalization by individuals who read documents and e-mail, attend presentations by others, access databases, and then absorb and reflect on the contents of all of these.
knowledge (tacit knowledge)	Externalization by articulating the personal knowledge of teams and individuals and creating documents, databases, presentations, etc. derived from this knowledge.	Socialization among individuals and teams who share knowledge and understating by articulating, demonstrating, exchanging, and negotiating ideas among themselves in a variety of settings (networking, ad hoc conversations, etc.) without directly codifying what has been shared.

Source: Pitt and MacVaugh (2008)

As information and knowledge span technical and non-technical (social, cultural, procedural) areas, Table 2 complements Table 1 by distinguishing four important levels of knowledge exchange and transfer that affect NPD within the focal organization.

Table 2 – Information/knowledge transfer among organization levels

Transition from	Transition to			
	External sources	Organization level	Team level (Q)	Individual (Y)
External sources		Download	Download	Download
Organization level	Upload	Circulate	Download	Download
Team level (P)	Upload	Upload	Exchange/transfer	Download
Individual (X)	Upload	Upload	Upload	Exchange/transfer

Source: Pitt and MacVaugh (2008)

It is possible to connect the rationale of this work to some levels of knowledge exchange from Table 2. Regarding the collected durability measurement files, it has been noted that the exchange occurs exclusively on the individual level, by transferring the files through shared or external drives. Hence, this work will focus on the information (data with context) exchange from individual to the team and organization levels. Although it might sometimes occur, the transfer to external sources (e.g. suppliers) is out of the scope of this study for confidentiality reasons.

2.2 DURABILITY

A fundamental element of this work is the very meaning of durability. According to Bennebach and Cawte (2007), durability connects to the resistance of a component or structure under different damage mechanisms such as fatigue, corrosion, wear, creep, etc. Overall, it expresses the capacity of a component to withstand its operating environment for a target duration. Nevertheless, of all these failure mechanisms, fatigue is responsible for most in-service failures and, for this reason, the term durability will be used herein to describe mainly fatigue performance. For this research, given its applied nature within a company of the transportation segment, durability may be defined as the ability of a vehicle, system, or component to preserve its designed function for its intended service life.

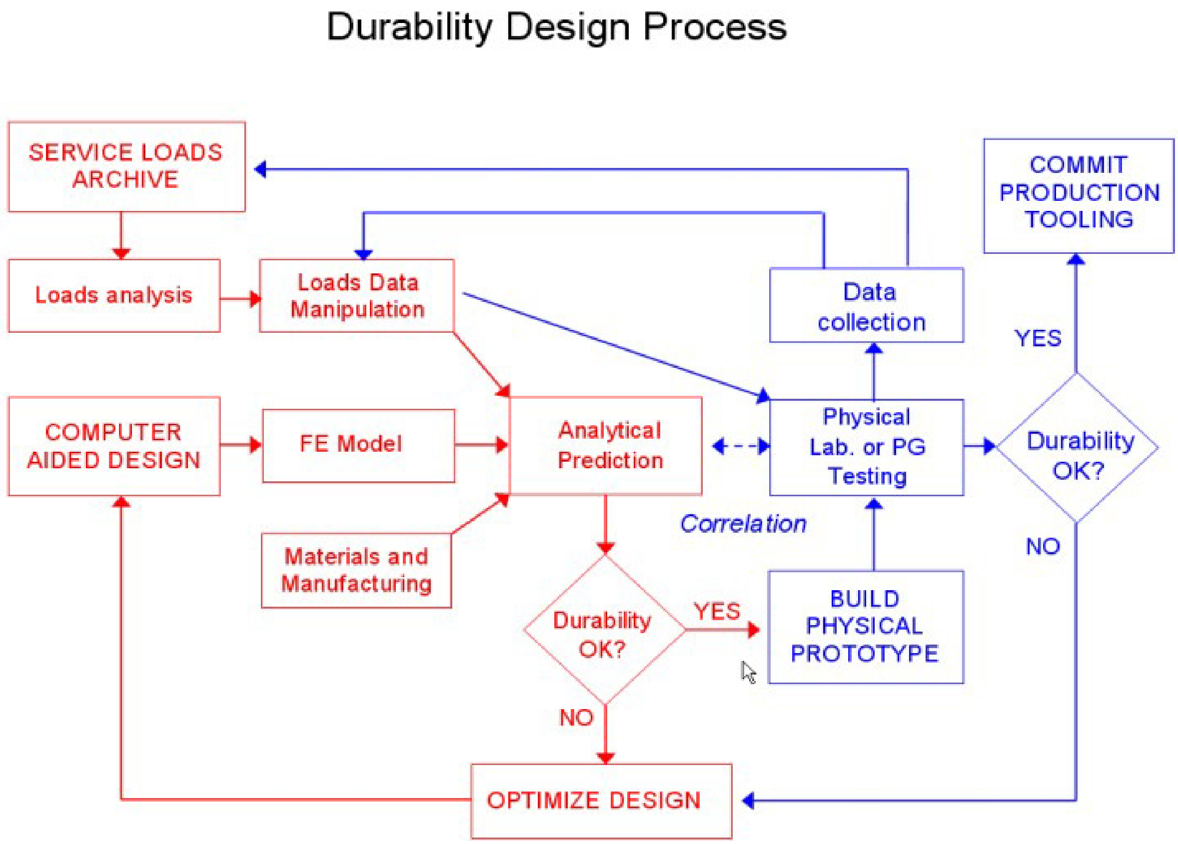
Johannesson and Speckert (2013, p. 5) highlight that in the process of designing a robust and reliable product that meets customers' demands, it is vital not only to assess the life of a component but also to investigate and take into account the sources of variability. According to them, there are mainly two aspects influencing the components' life: the load the component is exposed to, and the structural strength of the component. Statistical methods present useful tools to assess and quantify the variability of those aspects. While the variability in the structural strength depends on both the material scatter and the geometrical variations, the customer load distribution may be influenced by, for example, the application of the product, the operator behavior, and the market.

Johannesson and Speckert (2013, p. 6) state that developments in information technology and its integration into vehicles have presented new possibilities for in-service measurements. Additionally, the design process has also shifted to the computer. Both these activities, alongside demands for lightweight design, require a refined view on loads and lead to a renewed interest in load analysis to:

- Assess and quantify the customer service loads.
- Define design loads for vehicles, subsystems, and components.
- Define verification loads and test procedures for the verification of components, sub-systems, and vehicles.

While analytical techniques are available for load generation, stress state calculation, and fatigue analysis, Bennebach and Cawte (2007) argue that these methods rely on complex models and need realistic measured test data as inputs. The diagram in Figure 2 illustrates how analytical and physical approaches work in synergy to successfully analyze product performance. In this way-of-working, simulation techniques must be employed at the earliest possible stage, so that the design can be developed, optimized, and above all understood through analysis. Key elements in making reliable durability calculations, apart from the use of adequate models, is the accuracy of inputs loads – measured, virtually derived, or based on a previous design. Simulation allows sensitivity analyses to be conducted until acceptable durability is obtained. These analysis loops give an insight into the predominant influencing factors and the distribution of fatigue lives. At the same time, physical tests are conducted in the field or laboratory. These tests provide the necessary information such as loads or materials data and allow correlation with simulation which is critical to ensure high confidence in the results. Due to the inherent scatter of the fatigue phenomenon and all possible sources of inaccuracy & variability, the physical test represents the final validation of the design and is often time-consuming. At this stage, several test acceleration techniques may be applied to further optimize the process (BENNEBACH; CAWTE, 2007).

Figure 2 – Example of the durability design process within product development



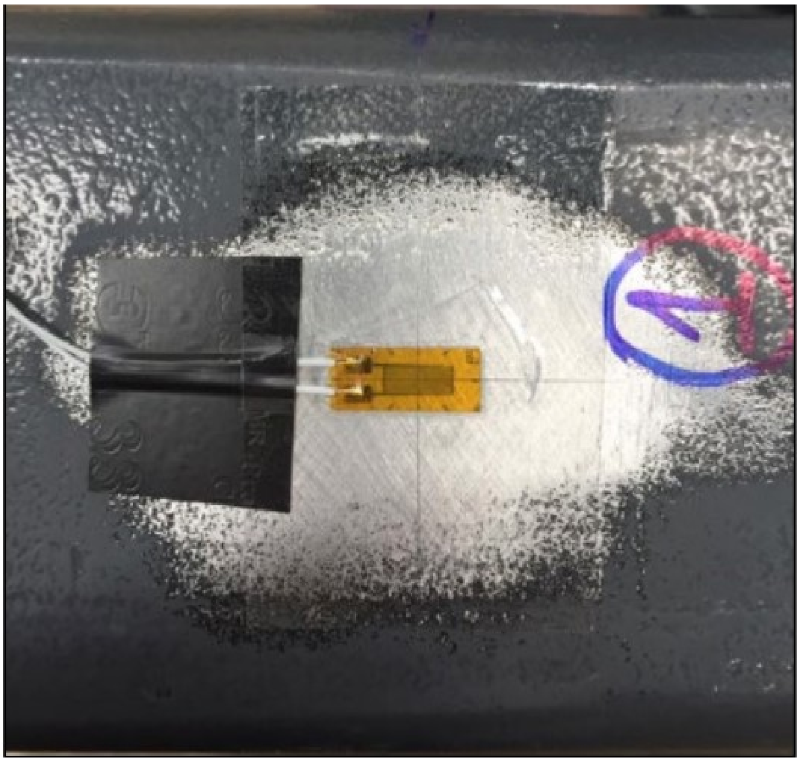
Source: Bennebach and Cawte (2007)

However, according to Bennebach and Cawte (2007), knowing the loads that are being transferred to a structure in real conditions is not a trivial task. Attention shall be paid when evaluating or measuring these loads to ensure that they are representative of the in-service usage. Physical loads can be obtained from in-service measurements using devices such as force transducers (Figure 3), strain gauges (Figure 4), accelerometers (Figure 5), and displacement sensors (Figure 6) from instrumented prototypes.

Figure 3 – Wheel force transducer installed in a commercial vehicle

Source: The author (2021)

Figure 4 – Strain gauge installed in the spring leaf of a commercial vehicle

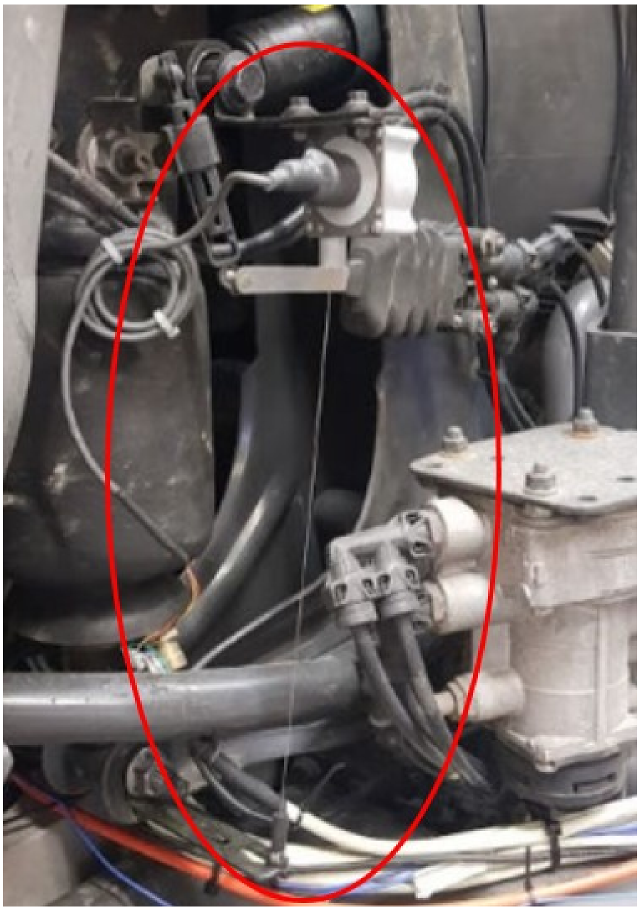


Source: The author (2021)

Figure 5 – Triaxial accelerometer installed in the frame of a commercial vehicle

Source: The author (2021)

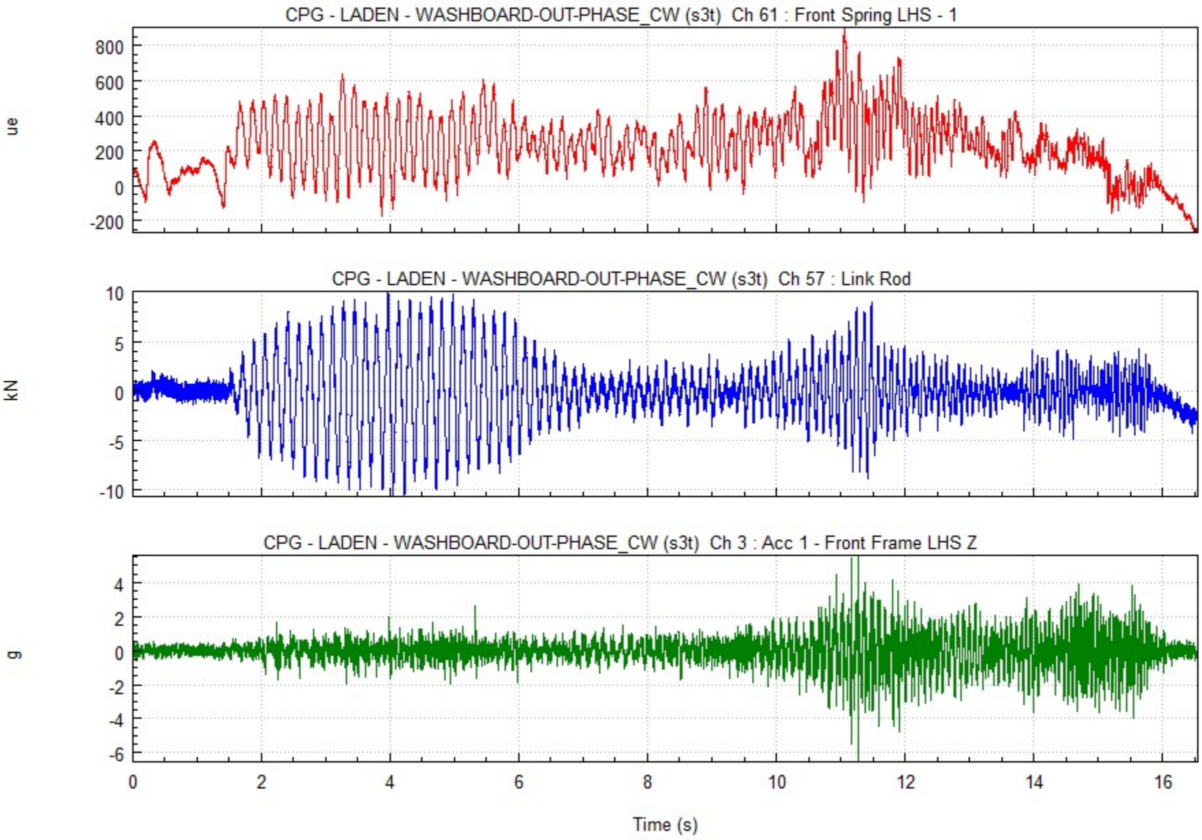
Figure 6 – Displacement transducer installed in a suspension system



Source: The author (2021)

While it is the best source of data possible, the measured data require long-term logs and many statistically representative samples. Typically, the referred data are measured time signals. An example of such signals is shown in Figure 7, showing the strain results on a truck's spring leaf, the force on a steering link, and the vertical acceleration on the front frame of the vehicle.

Figure 7 – Example of measured load signals (strain, force, and acceleration) - time history data



Source: The author (2021)

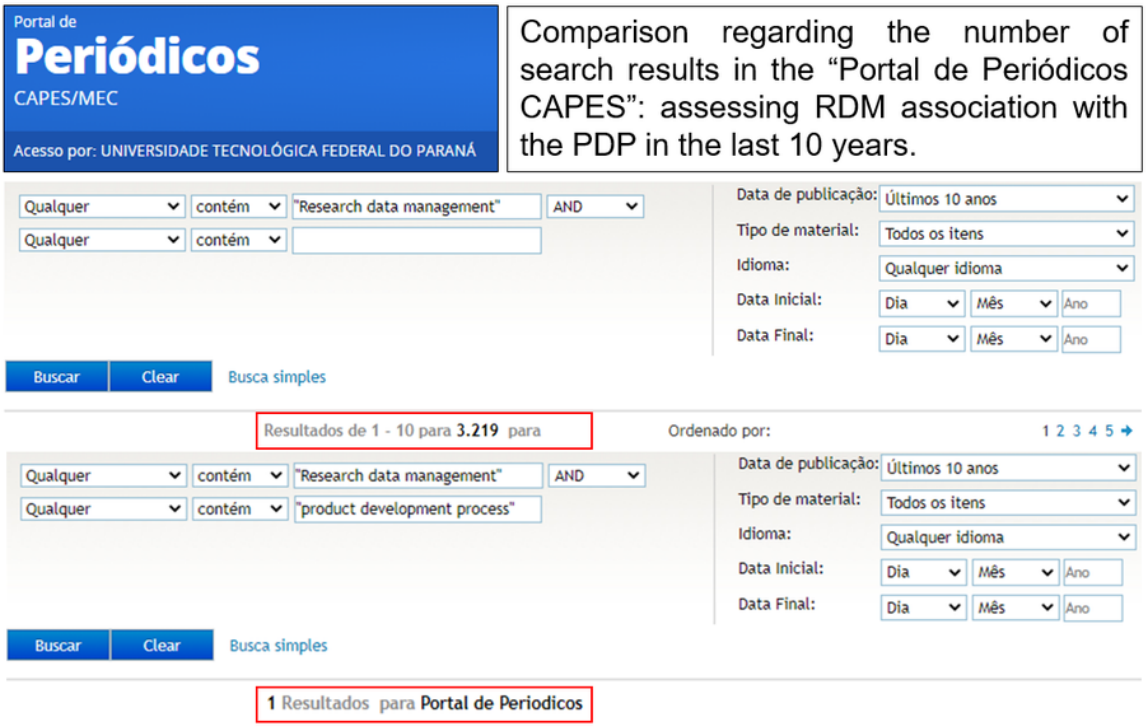
In a mathematical setting, the load is thus described by a function x(t), where t ranges over the time interval of interest. Typically, the values of x are known only at certain discrete points in time (samples). Due to the measurement processes, the samples are more often than not equally spaced with a sampling time and a sampling rate. In the following, such a set of data is called a time signal, a time series, or a time history (JOHANNESSON; SPECKERT, 2013, p. 32).

2.3 RESEARCH DATA MANAGEMENT

Additional researches revealed that the management of measured data has been widely discussed in the literature as Research Data Management[1] (RDM) but focusing on an academic context of projects – as illustrated by the search results comparison in Figure 8.

The association between experimental data and the PDP is not tackled in the bibliographic portfolio presented by Larocca et al. (2021) and is scarce in the broad literature, emphasizing the relevance of the present research project.

Figure 8 – RDM association with PDP: comparison of search results in the literature



Source: The author (2021)

Research data management (RDM), a term that encompasses actions connected to the storage, organization, documentation, and sharing of data, is vital to efforts to maximize the value of the scientific investment and to address the concerns of the research process integrity (COLLINS; TABAK, 2014). Yet, when assessed directly, researchers often acknowledge their lack of skills and experience to manage

[1] "Research Data Management is the care and maintenance of the data that is produced during the course of a research cycle. It is an integral part of the research process and helps to ensure that your data is properly organized, described, preserved, and shared." (THE UNIVERSITY OF SHEFFIELD, 2020, p. 1)

and share their data effectively (TENOPIR et al., 2016; FEDERER et al., 2015; BARONE; WILLIAMS; MICKLOS, 2017).

This disconnection reinforces the need for tools that bridge the gap between the research community, data service providers, and other data stakeholder groups (BORGHI et al., 2018).

Effective management of data provides rewards throughout and beyond the life of a research project. While data need to be discoverable, accessible, and intelligible to allow their long-term reuse, such values are equally critical during the research-active phase. The increasingly collaborative nature of research is a pressing argument for RDM services, as researchers need to exchange data across diverse platforms and demand effective systems to store, access and share data securely across multi-institutional (JONES; PRYOR; WHYTE, 2013).

A survey from Pinfield, Cox and Smith (2014) shed a light on the main drivers for RDM developments at an institutional level:

1) Storage: storage facilities for a wide variety of data at a scale that anticipates future demands.
2) Security: confidential or sensitive data should be held securely with relevant authentication and authorization mechanisms in place.
3) Preservation: medium and long-term archiving of data with associated selection protocols and preservation activities.
4) Compliance: comply with the requirements and policies of relevant agencies, as well as legal obligations, such as data protection.
5) Quality: maintain and improve the quality of research to prove the robustness of findings and allow results verification and reproducibility.
6) Sharing: share data amongst targeted users and provide mechanisms and systems to enable open access to data where appropriate.
7) Jurisdiction: professional narrative around the necessity to be involved in RDM and how it impacts other stakeholders in the institution.

Tuyl and Whitmire (2018) approached non-academic corporations and institutions to assess how data is managed in those organizations and how they compare to the academic RDM context. Their research revealed that data management practices in non-academic corporations were reminiscent of the types of practices and challenges experienced by academic researchers. Tuyl and Whitmire (2018) observed that even in well-resourced companies, data workflows and management regimes were highly disorganized. While there are many parallels between academic research data management and data management in non-academic settings, the authors highlighted the importance of engaging end-users, defining processes, workflows, and documentation for a successful process of data management.

On the other hand, research from Akers and Doty (2013) revealed that the data management needs of researchers vary substantially across disciplines. Besides datasets that differ in size and content, researchers from different domains are also connected to diverse research cultures and communities of practices with different attitudes toward data sharing and archiving (AKERS; DOTY, 2013). Overall, Akers and Doty (2013) conclude that different disciplines differ widely in their context (research funding, technical infrastructures, collaboration networks, methodologies, types of research outputs, and others). Effective data curation, therefore, requires services that are tailored to different populations of academic researchers (CRAGIN et al., 2010).

Borghi et al. (2018) prepared a set of materials called “Support Your Data”, which aims to support research communities in a broader effort to improve their data management-related practices. Among this set of materials, the authors provide one-pager guides for each of the following RDM-related activities: planning, organizing (Figure 9), saving, preparing, analyzing, sharing (Figure 10).

Figure 9 – One-pager guide for organizing data

Organizing Data

Organizing data involves ensuring that you can find your data and other research materials (including documentation, code, and physical samples) when you need to and ensuring that data and materials that go together are connected in a meaningful way.

What does it mean to organize data?

Organizing data means arranging your data and other research materials so they can be found – by yourself and by others – as needed. Here are four factors to consider when organizing data. Remember, you can't use data you can't find.

Names	Data should be labelled using a consistent and descriptive file naming system. Your system should allow you to immediately and uniquely identify the contents of your files.
Structures	Data should be organized within a consistent and easy to navigate file structure. Maintaining such a structure can help reduce the risk of data loss and unnecessary replication.
Connections	Connections give context. Data and other materials should be organized in a manner that emphasizes the links between them. This may refer to different versions of the same file or different files related to the same aim or project.
Documentation	You should document how you organize your data and other research materials and refer back to and update your documentation often. When thinking through how to organize your files, make sure you also consider how you include all of the related description and documentation (e.g. notes, data dictionaries, metadata).

Source: Borghi et al. (2018)

Figure 10 – One-pager guide for sharing data

Sharing Data

Sharing data is more involved simply uploading files somewhere for other researchers to find. The methods you use to share your data will depend on a number of factors including the size and content of your data, mandates from the entities that fund and publish your research, and assumptions and requirements related to future use. If you make your data and other materials available, you should make sure that other researchers can find and use them.

What does it mean to share data?

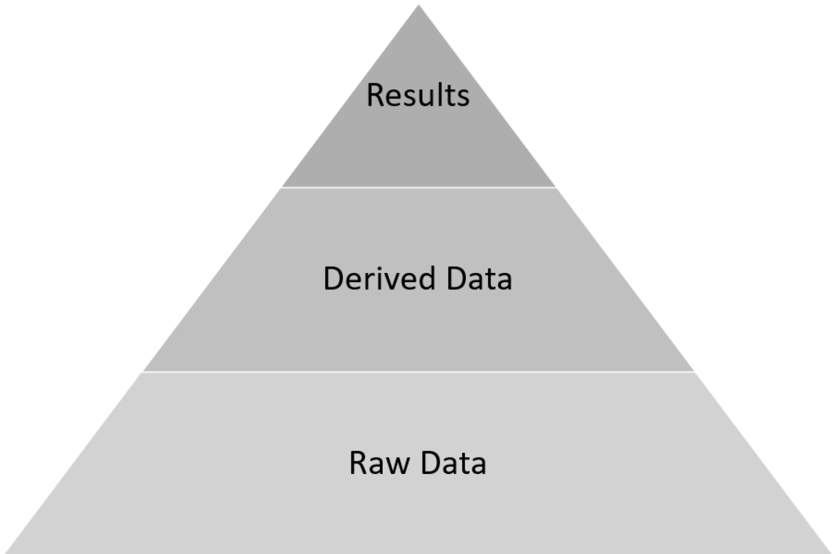
Sharing data means making your data available so that they can be accessed and used - by yourself or by others - in the future. Here are three factors to consider when sharing your data.

Format	Data should be shared in a usable format. This may mean sharing raw data instead of prepared data (or vice versa) or ensuring that data are saved in common or open file formats.
Completeness	Remember that notes, documentation, and other information about your data are part of your data. To ensure that your shared data is useful, make sure these elements are included.
Location	When choosing a method for sharing your data, consider how other researchers will find and use it. The storage options you use to save your data as you work on it will probably be different than the options you use to share it, especially over the longer term.

Source: Borghi et al. (2018)

Paton (2008) brings the concept of a data pyramid (Figure 11) and explores the decision process to identify what to store, how to store it, and why. Raw data, produced directly from an instrument, are subject to one or more steps of post-processing to yield derived data. Some portion of this derived data is then selected for consumption as the ‘result’ of the experiment in a subsequent analysis activity.

Figure 11 – Different data types: raw data, derived data, and results



Source: adapted from Paton (2008)

Paton (2008) provides different motivations to store each one of these data:

- Raw data: it might be difficult or impossible to reproduce the experiment again and that it may be desirable to rerun different analyses over the data at a later date (or project).
- Derived data: It explains why specific results were produced. Though, if the raw data are properly stored, it may be sufficient to describe how the derivation was carried out.
- Results: the most frequently consumed part of the pyramid.

In terms of how data should be stored, according to Paton (2008), this depends mainly on the form of access that is required. Raw data are generally produced in proprietary file formats and read in specific analysis programs; as a result, such data are typically stored in the original format, or in some standard representation that can be used by multiple analysis programs.

Paton (2008) concludes his work by presenting a review of the many pitfalls of scientific data management. One example is the over-ambition in this domain, where is common to observe the design of complex systems for capturing comprehensive descriptions of tests that are time-consuming to populate and consequently never used.

Back to the key aspects of the development of RDM services and data catalogs, Jones, Pryor and Whyte (2013) provide over their research a summary of key actions, such as:

- Assign an RDM team with proper responsibilities and authority to undertake actions.
- Define the current status of the RDM, including service and support gaps, to identify requirements.
- Design services that meet internal and external customers, and suit the organizational culture.
- Pilot services and involve end-users to ensure they are fit before wider roll-out.
- Define the metadata you need to record research datasets.
- Establish a system for capturing and displaying a record of research data holdings.
- Integrate systems to benefit from options to mine data and embed metadata creation into existing workflows.

Built upon the discussed literature regarding research data management (RDM) and through a consensus-building approach, Table 3 below summarizes the cornerstones that will ground the solution design of the data management model within this research:

Table 3 – Pillars for the design of the data management model

Cornerstones/ Key aspects	**References**
Data management practices and challenges in non-academic corporations are similar to the ones experienced by academic researchers. Hence, RDM's best practices and guidelines could be extended to the industry and more specifically to this research scope.	Tuyl and Whitmire (2018)
Data management needs vary substantially across disciplines. Top-down initiatives from companies, aiming for a unique solution, might fail due to their over-ambition and disregard of specifications and requirements from each discipline. Therefore, even in the test domain, one size does not fit all. This justifies the need for a bottom-up approach from each feature - such as the management of durability measurement files, which is the scope of this research.	Akers and Doty (2013) Paton (2008)
Given the fact that the success of data management systems relies on the acceptance and usage of the end-user, their engagement and involvement in the development process are vital.	Tuyl and Whitmire (2018) Jones, Pryor and Whyte (2013)
Raw data has to be properly stored: it might be difficult, costly, or even impossible to reproduce physical experiments. Different projects, with different demands, might need to rerun different analyses for their specifics needs. Raw data are the base of the data pyramid. The results shall continue to be stored in the format of test reports, detailing how the raw data has been post-processed to yield the derived data.	Paton (2008)
Although raw data are generally produced in proprietary file formats and read in specific analysis programs, it is important to set a standard that can be accessed by end-user and requesters.	Paton (2008) Borghi et al. (2018)
Metadata provides context to raw data. It plays a vital role in data retrieving and reuse.	Borghi et al. (2018) Jones, Pryor and Whyte (2013)

Source: The author (2021)

Complementary to this literature review, the next section will address the theoretical framework in which this research is positioned.

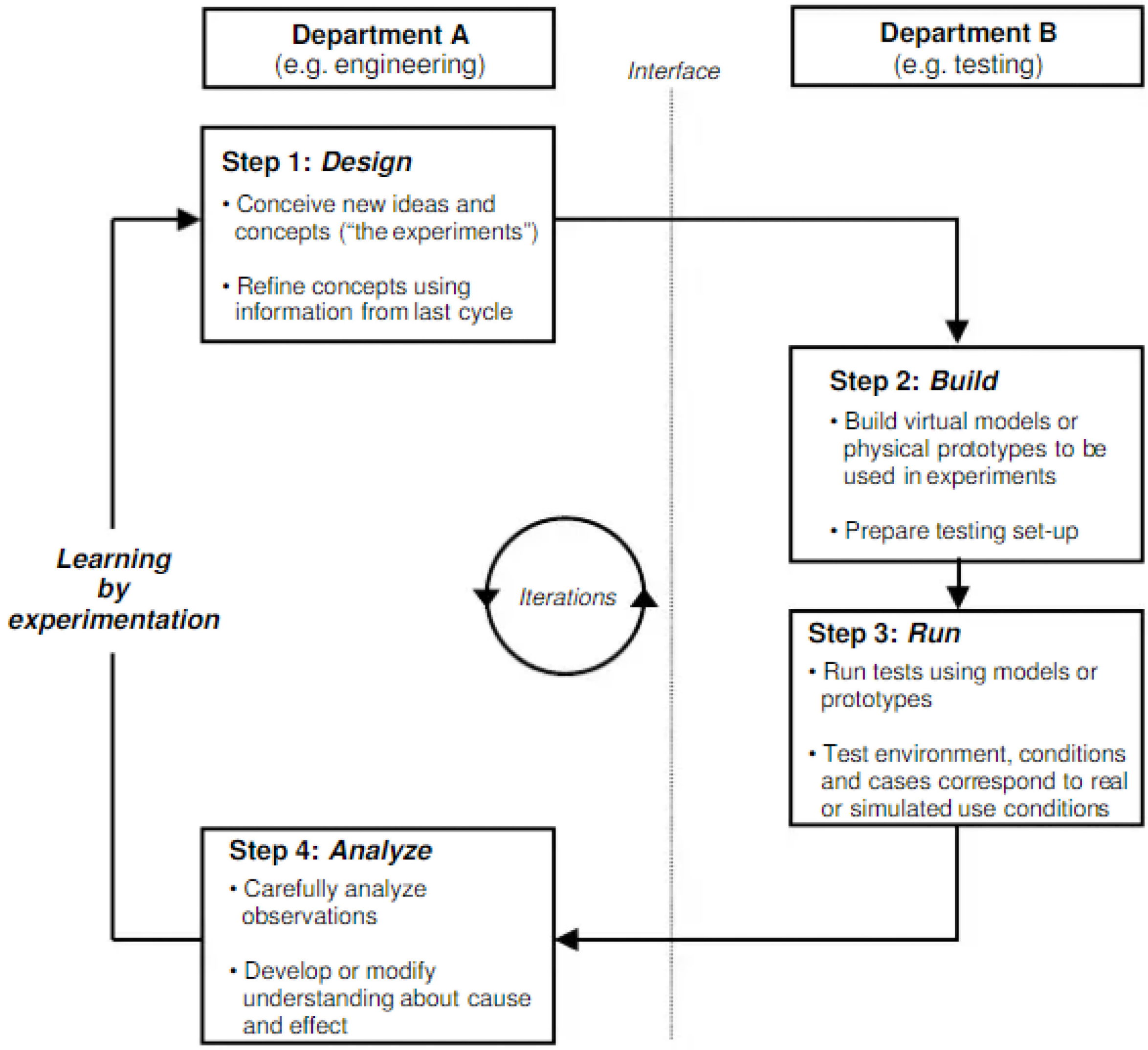
3 THEORETICAL FRAMEWORK

This section aims to provide the researcher’s particular perspective, or lens, through which the present research problem will be explored. It also targets to position the present dissertation regarding its contributions when compared to related work in the field. By providing a conceptual grounding of a study, this framework is built on a combination of tacit and formal theories. It introduces and describes the theory and related work that explains why the research problem under study exists (CAMP, 2001; UNIVERSITY OF SOUTHERN CALIFORNIA, 2020).

3.1 PROTOTYPING AND TESTING WITHIN THE PDP

Prototyping, testing, and experimentation usually comprise iterating attempts to find the direction in which a solution might lie (ALLEN, 1966; MARPLES, 1961; HIPPEL; TYRE, 1995). According to Thomke (2008), this process is generally triggered by the selection or creation of one or more potential solution concepts, which are then tested under an array of requirements and constraints. The iterative nature of testing is highlighted by the fact that new information and learning are produced when the outcomes of the experiment are not (or cannot be) foreseen in advance. These outcomes are then applied to review and improve the solutions under development. Thomke (2008) demonstrates this process through the four-step iterative cycles illustrated in Figure 12 below.

Figure 12 – Experimentation as four-step iterative cycles



Source: Thomke (2008)

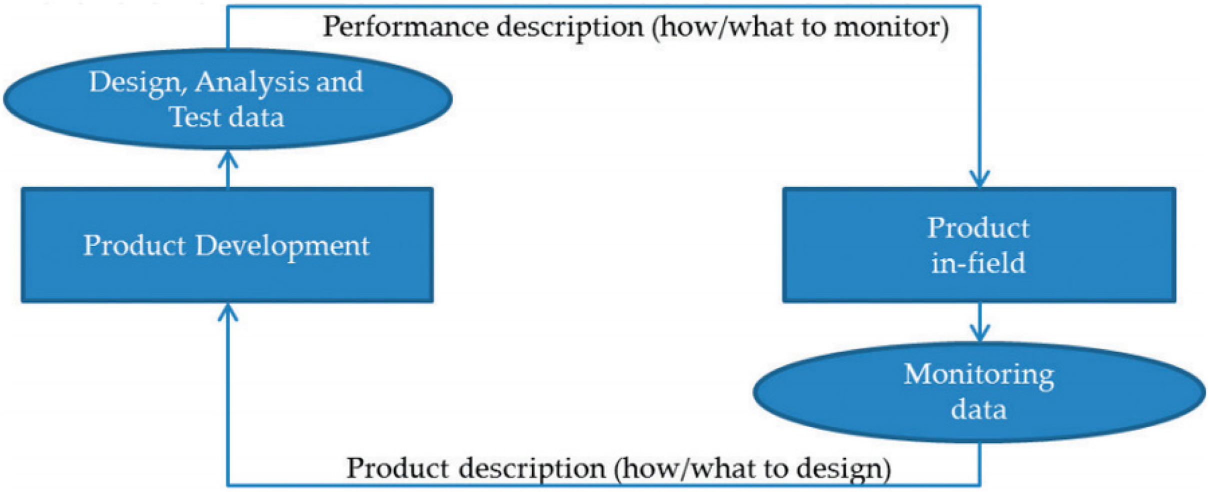
Moreover, Thomke (2008) recognizes the managerial challenges promoted by those iterations, as the steps are performed by individuals and teams that are usually divided across different functional departments with different objectives, incentives, and resources.

3.2 PLM AND SYSTEMS INTEGRATION

From an interdisciplinary collaboration perspective, Vornholt, Geist and Li (2010) presented a categorization of approaches for concurrent virtual engineering. Collaboration techniques and technical methods to handle heterogeneous data were categorized in order to classify and compare the advantages and disadvantages of existing data management solutions. More recently, research from Tahera and Earl

(2018) connected process and product models in product development regarding testing and PLM. According to them, PLM has a designed approach in which the processes of PD effectively trigger the verification and validation demands of a design proposal. Conversely, Tahera and Earl (2018) argue that testing and design are equal partners in the PDP and suggest that a testing and use data view of PLM can actually drive the design, as illustrated in Figure 13.

Figure 13 – Iterative updating of product and performance through the use of test and field data



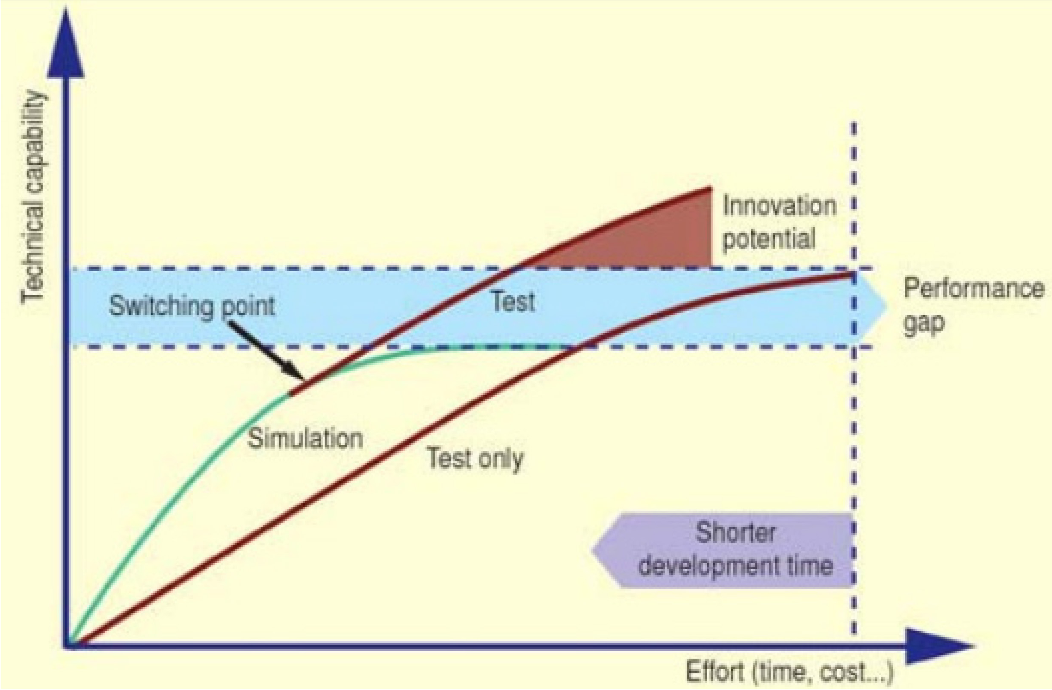
Source: Tahera and Earl (2018)

The study from Tahera and Earl (2018) aims to provide an overview of the importance of testing and the mismatch between several models of product development, which tend to relegate testing as a late step in the design process or primarily connected to quality issues. Although Tahera and Earl's (2018) research shed a light on the relevance of test data within the PDP, according to the authors "considerable further research is required both in theoretical methods and in industry cases to optimize the costly and time-consuming processes of testing, simulation and field data collection as well as integrating them with PLM systems" (TAHERA; EARL, 2018, p. 18).

Concerning this integration, Auweraer and Leuridan (2005a) advocate for the combined use of test and simulation, which would allow solving engineering problems faster and more precisely compared with exclusive use of one or the other. According to them, at each phase of the PDP, test data and test-obtained models contribute to increase the accuracy and even speed up the design process. Figure 14 shows how

Auweraer and Leuridan (2005a) connect the combination of test and simulation with product requirements, delivery, and innovation.

Figure 14 – Combining test and simulation to deliver innovation



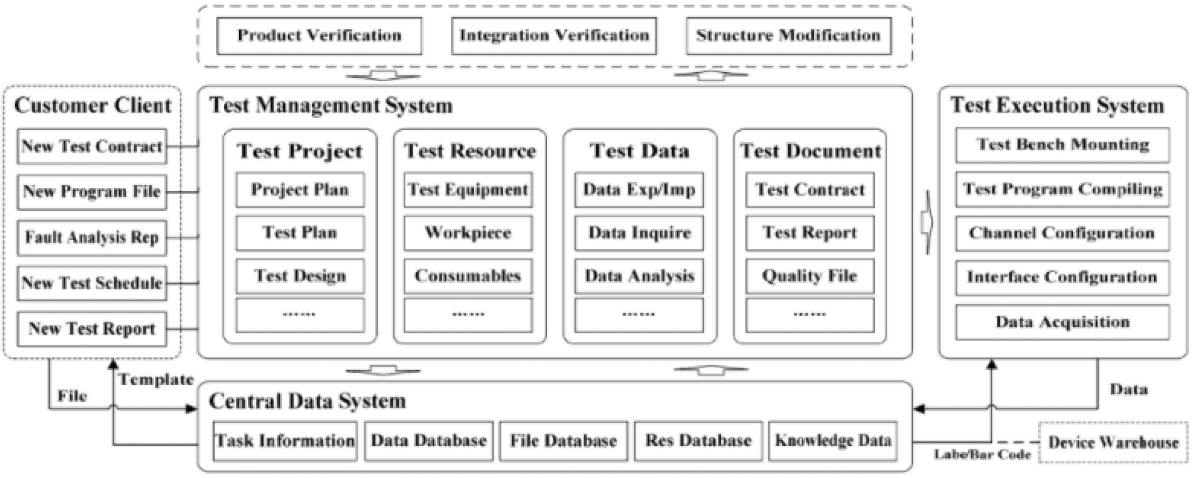
Source: Van der Auweraer and Leuridan (2005a)

As required technical capability is available, the traditional (test) method can take over where simulation reaches its limits. Such a case is common for system verification, where physical tests will be used to validate and calibrate simulation models. Auweraer and Leuridan (2005a) stress this as a goal, given that the required capability will increase in future development programs as a result of constant product innovation.

3.3 TEST DATA

Focusing on a testing perspective, Ji et al. (2017) present the design of a test system (schematic drawing in Figure 15) that includes test project, test resource, test data, and test document management.

Figure 15 – Proposed test management system by Ji et al. (2017)



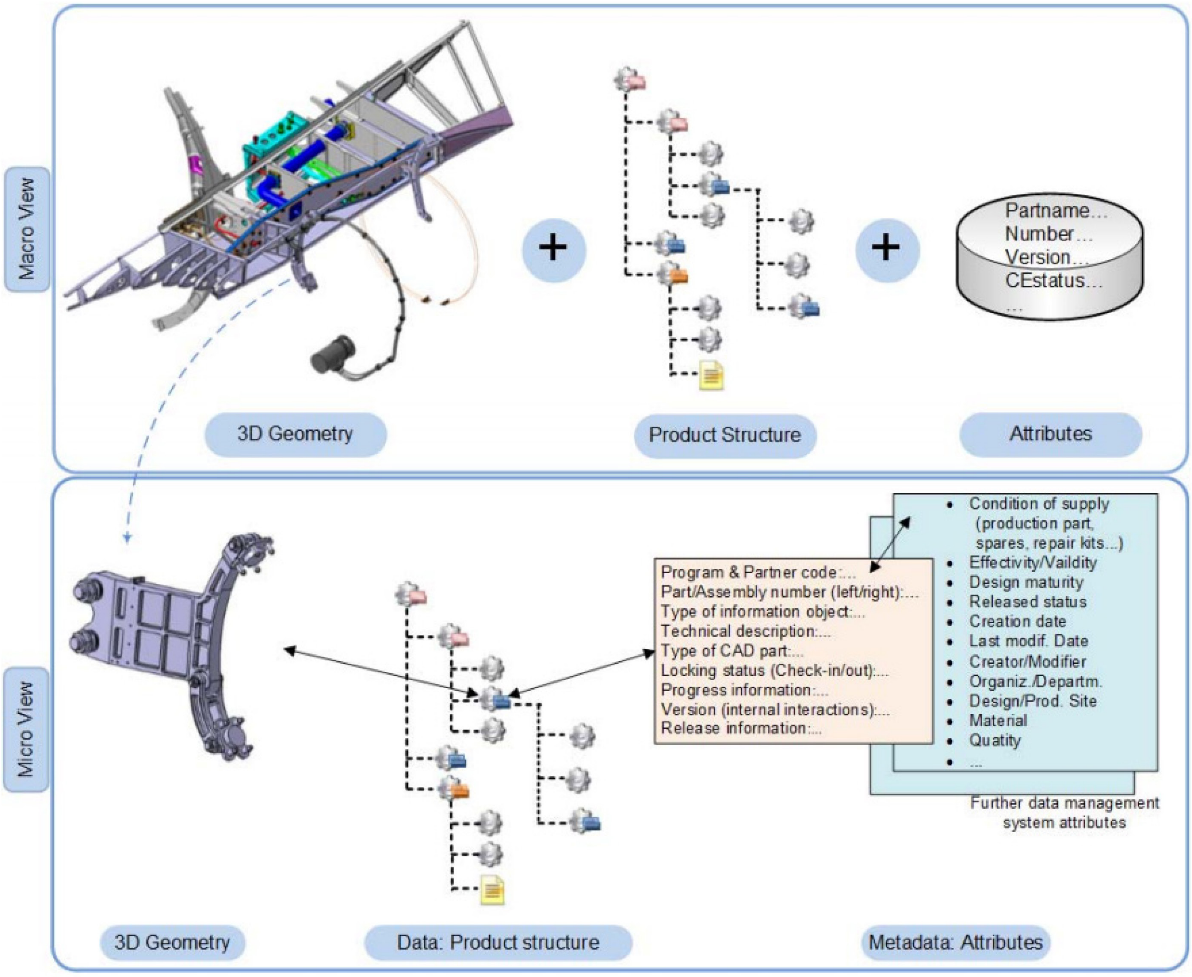
Source: Ji et al. (2017)

Concerning the measurement files, Ji et al. (2017) describe a management unit in which test data could be published, examined, approved, archived, and compared. The system should support unstructured data and enable their format change and exportation to the client. Moreover, according to the authors, the user should be able to “look over the data in the form of a curve, picture, and table in the test management system.” (JI et al., 2017, p. 42). Although the authors recognized relevant concepts and requirements regarding the management of measurement files, their work does not provide details of how the data should be structured or how the proposed system should be modeled.

Auweraer and Leuridan (2005b) present in a second work an interesting summary regarding the validation of the measured data. According to them, the key concern with acquiring experimental data is that these data must be valid. Detecting test data problems has to be done on several levels, which are detailed in their work. Auweraer and Leuridan (2005b) recommend that nonconformities have to be recognized, and decisions must be made on whether to redo part of the test, correct data afterward, or just flag them as being invalid or with problems. They recognize that time can be so limited that redoing the test might not be feasible, but at least one should be aware when using any ‘flagged’ data. Auweraer and Leuridan (2005b) conclude by stating that test engineers need to have easy and online data validation tools and that, upon completing a test, their report must include an assessment of the test data plausibility.

From a practical standpoint, previous research by McSorley (2014) has recognized the difficulties in ensuring that information gathered during development testing is efficiently shared with designers and made available for reuse in future projects (reinforcing what has been discussed in the rationale section of this work). Within this context, Toche et al. (2017) took part in a project entitled ‘Collaborative Development for Product Lifecycle Management’, based on a partnership between five universities and five major aerospace companies based in Canada. With similar motivation, their work is the most related to this research project that has been found in the literature. Toche et al. (2017) propose a framework that leverages digital mockup (DMU) configurations and PDM data (illustrated in Figure 16) to support the management of prototyping and testing information in a PLM perspective.

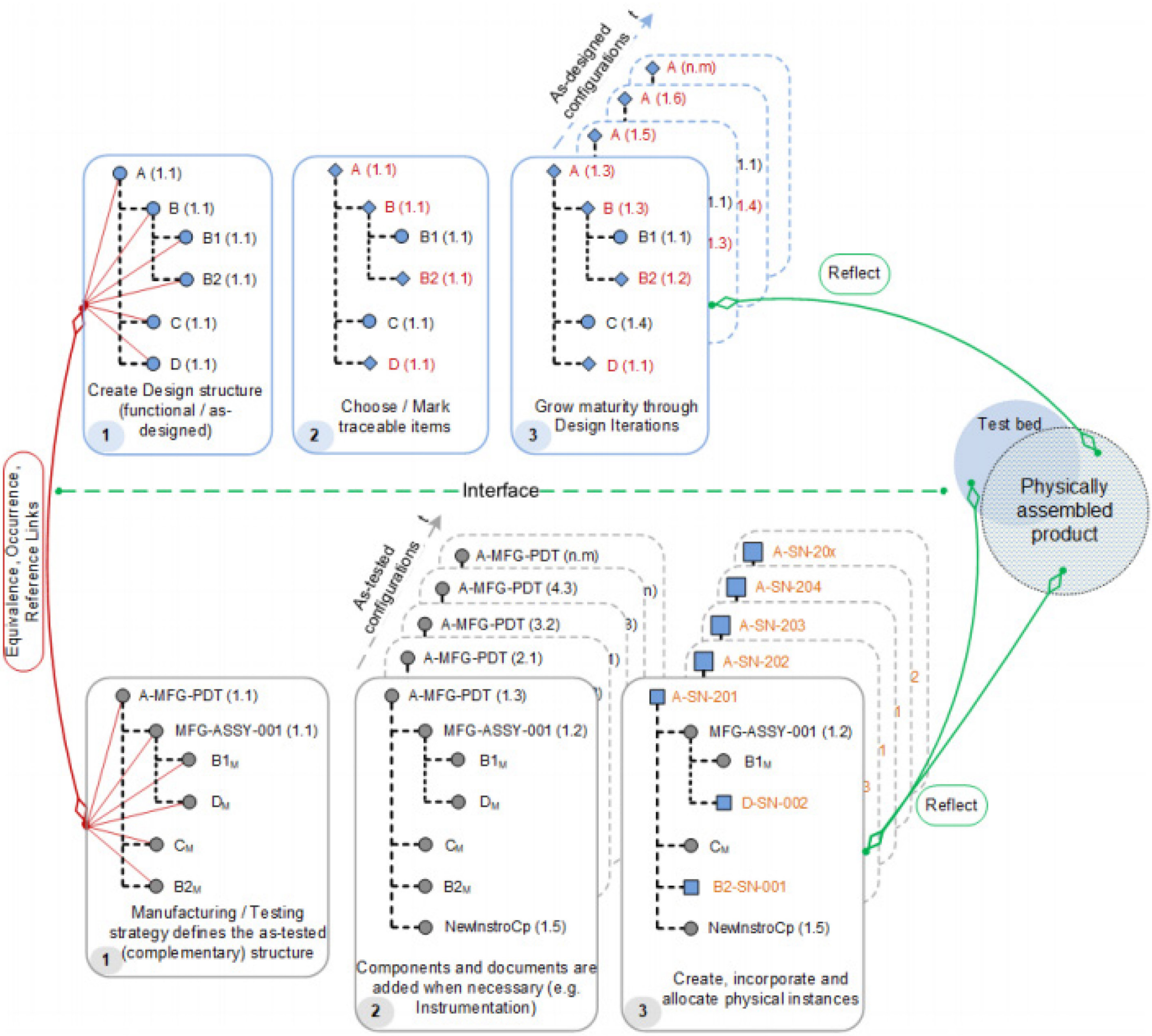
Figure 16 – Example of the relationship between geometry, product structure data, and metadata



Source: Toche et al. (2017)

Their approach is based on the development of explicit links between as-designed and as-tested complementary product structures, as illustrated in the diagrammatic analysis in Figure 17.

Figure 17 – Diagrammatic analysis of proposed information structures by Toche et al. (2017)



Source: Toche et al. (2017)

Toche et al. (2017) suggest that this framework would support collaboration between design and test engineers, as well as the management of the links between physical and virtual models. According to them, this approach would enable data from prototyping and testing activities to be mapped and merged with design activities to maintain product configuration and organizational needs in a cross-functional setting.

3.4 RESEARCH POSITIONING

Table 4 summarizes the main contributions from the related work in the field.

Table 4 – Contributions from related work in the field

Research	Main contribution
Tahera and Earl (2018)	Highlights the mismatch between several models of product development regarding the use of experimental data.
Auweraer and Leuridan (2005a)	Integration between measured and simulated data.
Ji et al. (2017)	Design of a holistic test management system.
Auweraer and Leuridan (2005b)	Importance and steps for validating measured data.
Toche et al. (2017)	Framework to connect as-designed and as-tested objects.

Source: The author (2021)

All the studies address the relevance of the measured data within the PDP and highlight the need for future research in the subject. However, only the works from Ji et al. (2017) and Toche et al. (2017) actually propose frameworks for test management. Although Ji et al. (2017) bring interesting concepts for the management of measurement files (such as the concept of an approval process for them), their work does not provide details or a specific model for their management – focusing on a more holistic perspective of the test domain. On the other hand, Toche et al. (2017) propose an interesting framework to connect test and product data during the development process. Their key objective is to ensure matching specifications between designed and tested objects. The present research will build upon the work from Toche et al. (2017) but focusing on catalyzing the reuse of measurement files and not the broad concept of test data (which include test reports, requisitions, measurement files, and other data types). It will also focus on the management of the files collected during durability tests of the PDP, which will allow an investigation of the main constraints that hinder the reuse of those files (which might include aspects such as sensors positioning, acquisition sample rate, test location, and object specification).

This section provided the researcher's perspective and positioned the contributions of the present dissertation compared to related work in the same domain. The next section will present the research design and methods, including the characterization of the research, the methodological framework and procedures.

4 RESEARCH DESIGN AND METHODS

In this section, the research is characterized, the methodological framework is described, and the performed procedures are detailed. The methodology was delineated in incremental steps that individually address the specific objectives, and that together, tackle the general objective of this research.

4.1 RESEARCH TYPE

Given the concept from Mills and Gay (2019, p. 7), this work can be classified as a qualitative research study, since it explores a specific situation to better understand a phenomenon within its original context and the perspectives of the stakeholders. This categorization is reinforced by Bui's (2013, p. 14) work, in which qualitative researches are characterized by collecting nonnumerical data (such as interviews, observations, and others) to answer the research question(s). According to Gil (2010, p. 134), a characteristic of this classification is its basis on inductive and descriptive analyses, without the need for statistical tools/methods.

Based on the lack of previous research regarding the management of experimental data in the PDP, this research can be classified as exploratory. This research type is characterized by Brotherton (2008, p. 12) as an attempt to generate some initial insights and understanding of the problem. It is intended to surface the key issues and questions as it would help to make the situation clearer and, possibly, set the research agenda. However, given the defined objectives of this work, this research can also be classified as prescriptive. This type of research has an important position in the domain of the information system (IS) discipline, as prescriptions are essential to apply theory in practice and to fulfill organizational improvements (CHANDRA; SEIDEL; GREGOR, 2015).

Finally, because of its practical nature and the fact that some of the stakeholders of this research (from a multinational manufacturing company) are from outside the academic discipline, this work is classified as an applied research. According to Brotherton (2008, p. 14), this type of research is generally concerned with tackling real-world problems. In this perspective, it is much more focused and goal-oriented than pure research and, therefore, more utilitarian. Collis and Hussey

(2013, p. 6) reinforce this concept by stating that the driving goal of applied research is to have practical payoffs or uses for the research output. In this concern, a collaboration from a multinational manufacturing company from the transportation segment has been set.

4.2 METHODOLOGICAL FRAMEWORK

The scientific perspective of design arises from the concepts described in the work from Simon (1969, p. 3), in which the purpose of design is defined as the act of changing existing situations into preferred ones. Simon's perspective of design science (DS) encompasses three central aspects: an imperative or prescriptive logic, a search for alternatives, and the evaluation of design (PRIES-HEJE; BASKERVILLE; VENABLE, 2008). March and Smith (1995) described the outputs of Design Science Research (DSR) as artifacts and framed them on the following categories:

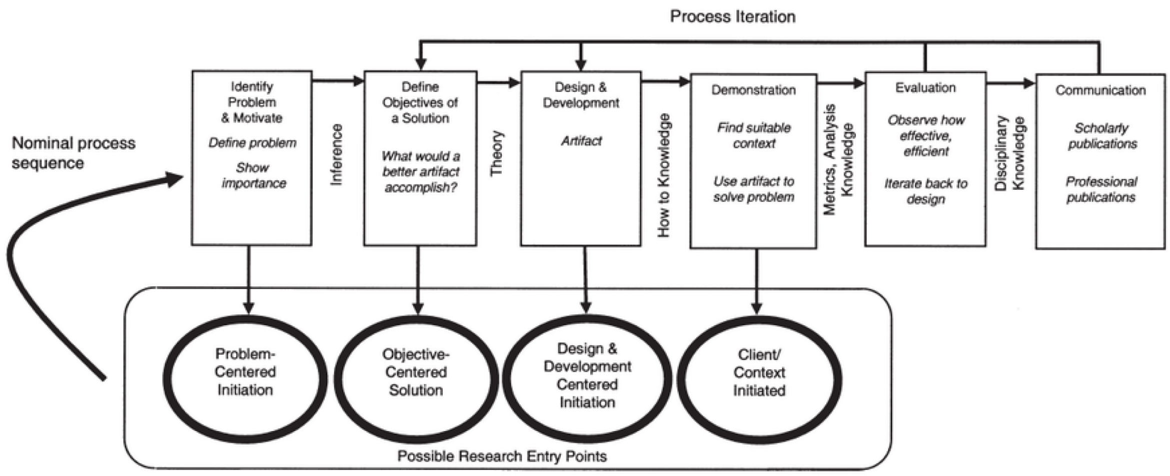
- Constructs: a conceptualization to describe problems within the domain and to specify their solutions.
- Models: a representation of how things are. A model expresses the relationship among constructs, its concern is the utility and not truth (so it is not a synonym for theory).
- Methods: a set of steps to perform a task, such as an algorithm or guideline. Methods can be linked to specific models in that steps take parts of the model as input.
- Instantiations: the realization of an artifact in its environment, which operationalizes constructs, models, and methods.

Vaishnavi and Kuechler (2004) argued that designed artifacts must be evaluated as to their use and performance as possible explanations for changes in the behavior of systems, people, and organizations. According to Geerts (2011), relevance and novelty are important characteristics of design science artifacts. First, an artifact must solve an important problem (being relevant). Second, to distinguish DSR from routine design, Hevner et al. (2004) state that design science research should tackle an unsolved problem in a unique way or a solved problem more effectively or efficiently.

In this concern, the methodological framework proposed for this research project is the Design Science Research Methodology (DSRM). This framework was proposed by Peffers et al. (2007) targeting to provide a standard process for the conduct of design science research, build upon prior literature about DS in information systems (IS), and provide researchers with a mental model for the characteristics of research outputs. According to Geerts (2011), while consistent with the principles and guidelines of design science research consolidated in previous works such as Nunamaker, Chen and Purdin (1990), Walls, Widmeyer and Sawy (1992), and Hevner et al. (2004), the DSRM enhances the production, presentation, and evaluation of DS research.

Figure 18 presents the result of Peffers et al.'s (2007) synthesis: a DSRM process model consisting of six activities in a nominal sequence.

Figure 18 – DSRM process model



Source: Peffers et al. (2007)

Although this process is structured in sequential order, Peffers et al. (2007) highlight that there is no expectation that researchers would always proceed in sequential order from activity 1 through activity 6. In fact, they may start at almost any step and move outward – this is illustrated by the possible research entry points in Figure 18. Each of the activities is detailed by Peffers et al. (2007) as below:

1) Problem identification and motivation: Describe the research problem and justify the relevance of a solution. The solution can benefit from the

breakdown of the problem conceptually. The relevance of a solution attains two things: it encourages the researcher and the audience to seek the solution and to accept the results and it supports the reasoning associated with the researcher's understanding of the problem.

2) Define the objectives for a solution: Infer the goals of a solution rationally from the problem definition and knowledge of what is possible. The objectives can be quantitative or qualitative. In this regard, the knowledge of the state of problems and current solutions are included in the resources required.
3) Design and development: Develop the design research artifact. This can either be constructs, models, methods, or instantiations or "new properties of technical, social, and/or informational resources" (JÄRVINEN, 2007, p. 49). Theoretically, an artifact can be any designed item in which a research contribution is embedded in the design. Besides creating the artifact, this activity includes defining its wanted functionality and architecture.
4) Demonstration: Simulation, experimentation, case study, or any other activity that demonstrates the usage of the developed artifact to solve one or more instances of the defined problem.
5) Evaluation: Observe and quantify how well the created artifact fulfills the solution to the problem. This assessment could take many forms (satisfaction surveys, customer feedback, or simulations). Such an evaluation could include any suitable empirical or logical evidence. There is a decision point at the end of this activity: the researcher can decide whether to iterate back to the design and development stage to attempt to improve the artifact or to proceed to the next activity (communication) and leave additional improvement to future works.
6) Communication: Communicate the problem and its relevance, the created artifact, its value and originality, the consistency of its design, and its effectiveness to researchers and/or practitioners.

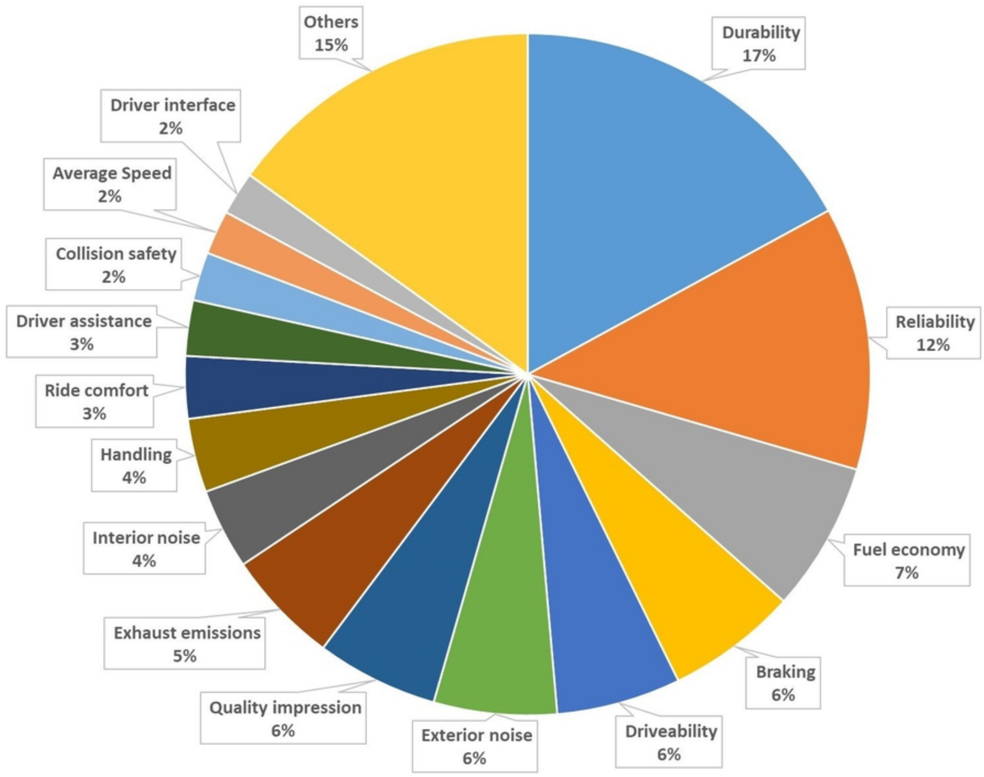
4.3 DEFINING OBJECTIVES OF A SOLUTION

From this standpoint, the first activity described in the DSRM's framework (Problem identification and motivation) is deemed to be completed by the research published in the VINE Journal of Information and Knowledge Management Systems by Larocca et al. (2021).

Regarding the second activity (defining the objectives for a solution), a consultation has been carried out to assess the current challenges and opportunities of collaborative approaches for experimental data within one multinational manufacturing company from the transportation segment. Besides manufacturing, this company invests in the product development of different products, such as trucks, buses, construction equipment, and engines.

Firstly, an assessment has been carried out within the global database of test reports from the evaluated company. The objective was to identify the most prevalent test type within their product development processes. As shown in Figure 19, the results indicated that, over the last five years (2015-2020), durability tests presented a higher prevalence when compared to other verification and validation activities.

Figure 19 – Feature tests in a multinational manufacturing company

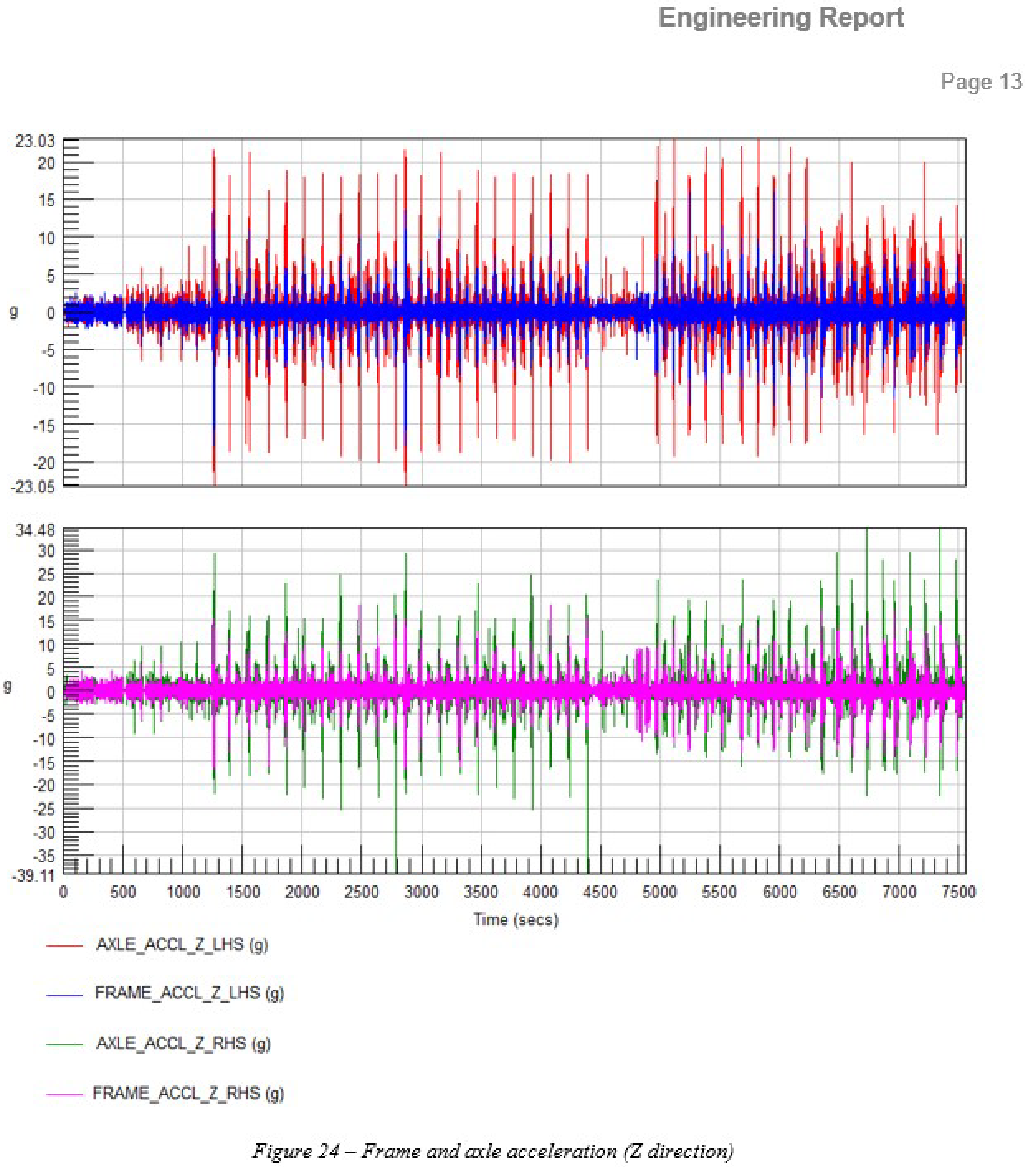


Source: The author (2021)

Secondly, an inquiry has been carried out with the top requesters of durability tests within the company's Brazilian subsidiary (within the same timespan: 2015-2020). The objective was to understand their perspective on the challenges and opportunities regarding the management of the experimental data used by them.

The survey revealed that test reports still an important way of formally reporting the results, covering commercial and legal aspects of the business. Nevertheless, the information within these reports loses value rather quickly as it tends to be specific to the request that has been made at that point in time. Moreover, the format of the reports brings additional challenges, as their quality and content are solely dependent on the issuer. These reports are usually stored in a Portable Document Format (PDF), which do not provide the readers the level of flexibility and control desired, such as selecting the signals/variables in a chart, calculate basic statistics, change graph scales, and others – as illustrated in an extract from a report presented in Figure 20.

Figure 20 – Extract from test report: static charts are a limitation for the reader



Source: The author (2021)

. Nonetheless, according to the requesters, the management of the measurement files is even more critical and important than the test reports. They argue that while old reports are used as a source of lessons-learned, to inspire a new test demand/method, it is unlike that the information in it could be capitalized to an extent where a new test could actually be avoided. The measurement files, on the other hand, are seen as the foundation of many of the tests performed in a durability context. If properly archived and identified, they would have the potential to enhance the design of future experiments and to eliminate the need for a new physical test in some cases.

For instance, we could imagine the following scenario: within a new development project, one engineer needs to validate a component in a bench test. With no additional information, this demand would probably unfold in two test requisitions:

1) Data acquisition of the input loads (strain/acceleration/displacement) on a test track or durability reference.
2) Bench test of the component using the measured signals as the target.

Now, if the engineer had access to a database of measured files, he would know that signals on similar positions and vehicle specifications have already been measured on a past project. In this case, there would be no need to measure these data again and one physical test could be avoided – saving project cost and time. Unfortunately, this is not how it usually happens.

Durability measurements often generate very large files of unstructured time-series data. In the company's Brazilian subsidiary, these measurement files are stored in one of the following ways:

- Attached to the test report within the engineering reports' database.
- Stored in a local shared driver of the Verification & Validation team.
- Stored in the local disk or external hard drive of the test engineer.

These files are generated in a variety of extensions, depending on the data acquisition (DAQ) system used to perform the measurement or the software used for data post-processing. This brings an additional problem, as very few people in the assessed organization (mainly test/analysis engineers) have the necessary software/license to read those file formats – pushing the design engineers even further from their requested data.

Furthermore, no metadata is incorporated in those files. Test reports are helpful to a certain extent, but they miss details that, in most cases, hinder the reuse of the measured data. It has been observed that, when occurred, the use of past measurements was only possible given previous experiences from the team members involved in the task. In other words, this knowledge is embedded in the people that took part in the assignment, but not in the organization itself.

Those issues are connected to the company's PLM and go beyond the management of CAD models, which are perhaps the focus for traditional vendor solutions. These challenges are exemplified within some of the topics and issues described in the book from Stark (2020, p.427), such as "data silos", "legacy data problems", "conflicting copies of the same data", and others.

4.4 METHODOLOGICAL PROCEDURES

Based on the framework of DSRM, this research presents from this standpoint a design and development centered approach, given the "existence of an artifact that has not yet been formally thought through as a solution for the explicit problem domain in which it will be used" (PEFFERS et al., 2007, p. 56). More precisely, the existent artifacts are the constructs based on the consultations and observations performed in a multinational manufacturing company (described in the previous section) and the state of the art summarized in section 1.1 (based on the bibliometric and systemic analyses presented by Larocca et al. (2021)). Those constructs are deemed to have fulfilled activities 1 and 2 of the nominal sequence from Figure 18: "problem identification and motivation", and "define the objectives for a solution", respectively.

The resulting artifact of this research can be framed as a logical data model (LDM) for the management of the measurement files collected in durability tests within the PDP. This is aligned with the definition of a model as a "solution component to an information requirements determination task and a problem definition component to an information system design task" (MARCH; SMITH, 1995, p. 256). This emphasizes the scope of the present research, in which the development of an algorithm (method) and the realization of the artifact in its environment (instantiation) are not expected. Instead, those steps were considered as a possible continuation of this work and are suggested for futures researches. An LDM delivers the specifications for data that define the concepts, relationships, and interpretation of values of data. A logical data model does not define the physical structures in which the data may be stored in databases or diffused between service units. In other words, an LDM is a business abstraction of the data specifications (CUMMINS, 2010, p. 132). According to Sherman (2015, p. 175), there are three levels of data models, which grow increasingly complex: conceptual, logical, and physical. Their relation is illustrated in Figure 21.

Figure 21 – The three levels of data models

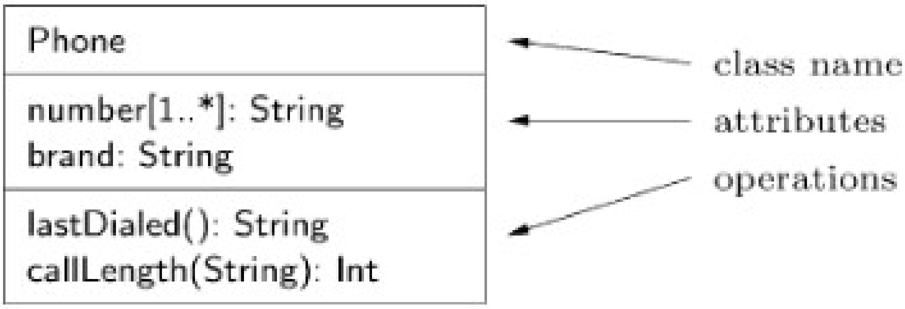
Source: adapted from Sherman (2015)

The logical model is used as a bridge from the application designer's perspective to the database design and the developer's specifications. This model is used to validate if the resulting applications fulfill business and data requirements (SHERMAN, 2015, p. 175).

The representation of this logical model was based on the Unified Modeling Language (UML), which is "probably the most widely known and used notation for object-oriented analysis and design" (ALI; SHUKUR; IDRIS, 2007). Within the UML notation, this research opted for the use of UML class diagrams as they "allow for modeling, in a declarative way, the static structure of an application domain, in terms of concepts and relations between them" (BERARDI; CALVANESE; GIACOMO, 2005).

According to Purchase et al. (2001), classes (represented as rectangles) are a description of concepts and may present attributes and operations in them. A class is divided into three parts (see example in Figure 22). The first part presents the class name, which has to be unique in the whole diagram. The second part contains the attributes of the class, each denoted by a name, while the third part contains the operations of the class (BERARDI; CALVANESE; GIACOMO, 2005).

Figure 22 – An example of UML class



Source: Berardi, Calvanese and Giacomo (2005)

The following subsections describe the methodological procedures of this research based on the framework of the Design Science Research Methodology.

4.4.1 Design and Development

The design and development process started by building upon what researchers pointed in key prior literature regarding research data management (RDM). A narrative literature review has been carried out, targeting a consensus-building approach in views about RDM and extending these to an industry perspective and a PDP environment. For this, full-text access to scientific journals was provided through the Brazilian national electronic library consortium for science and technology (Portal de Periódicos CAPES). This step was documented in the literature review of this dissertation, section 22

Secondly, this work followed some aspects from the participatory development methodology, which ". . . . (involves) stakeholders, particularly end-users, as much as possible in system development to ensure that requirements are met" (GAMMACK; HOBBS; PIGOTT, 2007, p. 345). It started with a requirements-gathering process, in which a diverse set of potential end-users and experts participated, resulting in the identification of the main constraints that hinder the reuse of durability measurement files within the PDP. This was addressed in the format of an ideation workshop, as the one described by Richter et al. (2018).

Subsequently, grounded on prior literature and the defined requirements, the model to handle measurement files collected in durability tests within the PDP has been structured. It focused on enabling the complete lifecycle of the most recurrent data in durability testing: force, strain, acceleration, and displacement. The following resources have been used for this step:

- Collaborating company's PLM infrastructure to leverage the interconnections already in place between their different systems.
- A set of measurement files collected in durability tests and relevant information about them (such as project, test report, vehicle specification, date, location, etc.).
- Software nCode GlyphWorks[2], which was used to manipulate both data and metadata from the measurement files.
- Software Enterprise Architect for the visual modeling and design based on the Unified Modeling Language (UML).

Besides the proposed UML model, its architecture and the data approaches were also detailed in the text. This was done so readers that are not familiar with the Unified Modeling Language could understand the defined concepts. Another reason is that, although only one model is proposed, the different and independent concepts therein could potentially be implemented separately or in different phases - depending on the business context. Moreover, as the details and concepts were delineated, this research also provided an outlook of the already available or potential technologies that could operationalize the proposed model. This included the mention of commercial hardware and software products that could support future works in the instantiation of the proposed artifact.

---

[2] "nCode GlyphWorks is a post-processing system that contains a comprehensive set of standard and specialized tools for analyzing measured data to increase product durability and performance" (HBM PRENSCIA INC., 2020, p. 1).

4.4.2 Demonstration

In this stage, a proof-of-concept has been demonstrated using real durability measurement files and other related information (such as vehicle specification, test report, and any other relevant data required by the artifact). The artifact demonstration included:

- The data and metadata structure.
- The searching process in the proposed database structure for durability measurement files.
- The connections between the proposed model with other PLM systems from the collaborating company.
- An outlook of how the measurement files, once consolidated within the proposed model, could be streamlined into virtual simulations/models.

Thus, this step demonstrated the operationalization of the proposed artifact and clarified how the measurement files collected in durability tests within the PDP would be available for reuse across projects and throughout the product development organization. For demonstration purposes, a database and user interface have been modeled using Microsoft Power BI.

This demonstration was done to key stakeholders from the collaborating company: a manufacturing company from the transportation segment. To support the evaluation of the extent of this work, this demonstration involved cases and employees from two different business units: trucks and buses. This reinforces the fact that the proposed solution was not modeled to address a single unit or department issue. Moreover, the partner company performs a variety of data acquisitions (measurements), from singles components, to full systems and complete vehicles. Therefore, besides tackling the challenges faced by the manufacturing company, it may also be applicable for smaller companies, such as the components' suppliers, which take part in the development process as well.

Some specific concepts from the proposed model might not be relevant for all types of industry and academic research. For instance, a company that develops home appliances or a university lab might not be interested in connecting their measured data to GPS data. However, they might still be interested in some key concepts proposed for data management in this work: data versions, data approval, metadata enhancement, and/or the virtual documentation of their instrumented test objects/specimens.

So, although this demonstration does not cover all industry and academic environments, we stress the modularity of the proposed artifact and the wide application (from a measurement perspective) of the partner company.

4.4.3 Evaluation

Given the fact that the implementation of the proposed artifact is not in the scope of this research, it was not possible to measure how much cost and time would be saved by the reuse of durability measurement files within the PDP (which would have been fostered by the research model).

From the standpoint of the actual output of this research, the evaluation phase consisted of the assessment of empirical evidence or logical proof (PEFFERS et al., 2007) regarding the functionality of the proposed artifact. Table 5 presents some of the design evaluation methods proposed by Hevner et al. (2004).

Table 5 – Design Evaluation Methods

Observational	• Case Study: Study artifact in-depth in a business environment • Field Study: Monitor use of artifact in multiple projects
Analytical	• Static Analysis: Examine the structure of the artifact for static qualities (e.g., complexity) • Architecture Analysis: Study fit of artifact into technical IS architecture • Optimization: Demonstrate inherent optimal properties of the artifact or provide optimality bounds on artifact behavior • Dynamic Analysis: Study artifact in use for dynamic qualities (e.g., performance)

Experimental	• Controlled Experiment: Study artifact in a controlled environment for qualities (e.g., usability) • Simulation: Execute artifact with artificial data
Testing	• Functional (Black Box) Testing: Execute artifact interfaces to discover failures and identify defects • Structural (White Box) Testing: Perform coverage testing of some metric (e.g., execution paths) in the artifact implementation
Descriptive	• Informed Argument: Use information from the knowledge base (e.g., relevant research) to build a convincing argument for the artifact's utility • Scenarios: Construct detailed scenarios around the artifact to demonstrate its utility

Source: Hevner et al. (2004)

This study opted for the experimental approach proposed by Hevner et al. (2004) and described in Table 5. Initial "proof-of-concept" level validation was presented to experts and potential lead users from the collaborating company, characterizing both a controlled experiment and simulation (as the artifact has been executed with both real and artificial data).

Feedback and satisfaction surveys were the selected tools used to compare the proposed artifact to the current model and to measure the stakeholders' perception regarding the model's:

- Completeness: the state or condition of having all the necessary or appropriate parts.
- Fidelity with real-world phenomena: how well the model's goal describes a desire not yet realized (VAISHNAVI; KUECHLER, 2015, p. 24).
- Internal consistency: based on internal relations. Something consistent cannot violate the rules that have been established before (BOLLEN, 1984).
- Level of detail: the abstraction level, the overall state of your information model (VAISHNAVI; KUECHLER, 2015, p. 24).

- Robustness: the ability to cope with errors during execution. The model has to keep an acceptable behavior in execution conditions (FERNANDEZ-REYES; HERMOSILLO- VALADEZ; MONTES-Y-GÓMEZ, 2018).

The aspects above were defined based on the DSR evaluation criteria proposed by March and Smith (1995) for models, as shown in Table 6 below.

Table 6 – Evaluation criteria for DSR artifacts

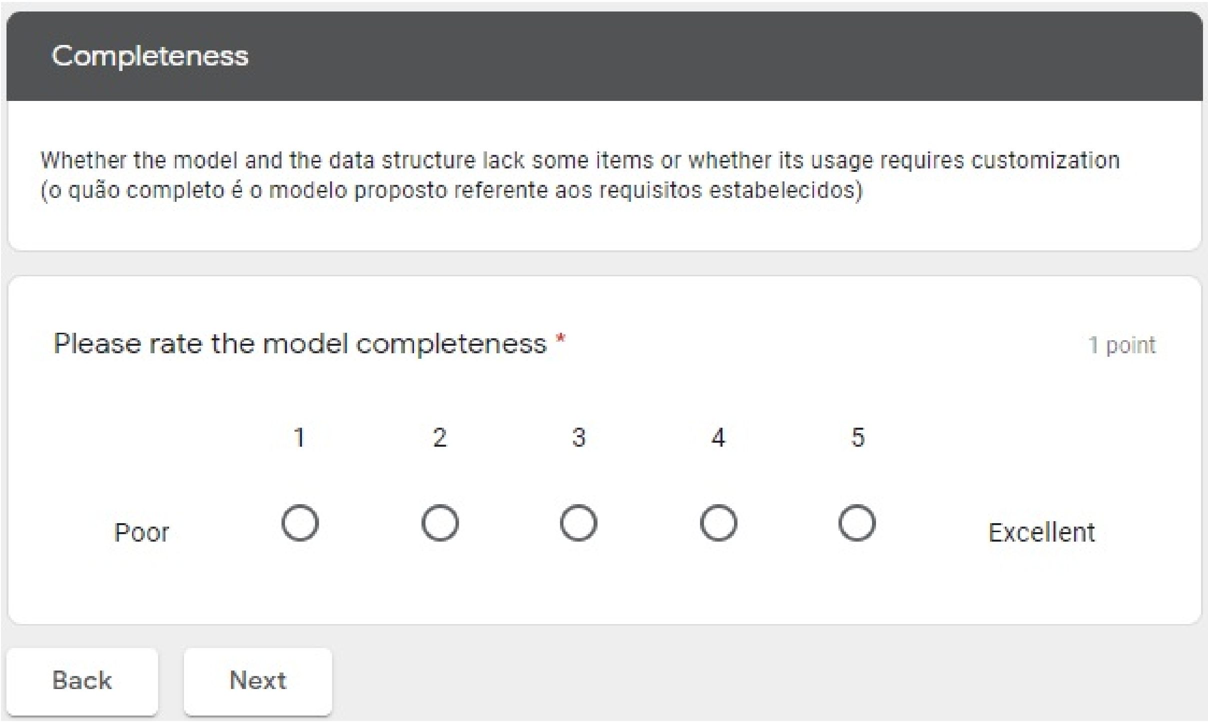
Criterion	Construct	Model	Method	Instantiation
Completeness	x	x		
Ease of use	x		x	
Effectiveness				x
Efficiency			x	x
Elegance	x			
Fidelity with real-world phenomena		x		
Generality				x
Impact on the environment and the artifact's users				x
Internal consistency		x		
Level of detail		x		
Operationality			x	
Robustness		x		
Simplicity	x			
Understandability	x			

Source: March and Smith (1995)

These criteria were selected to balance the interests of practitioners, who are interested in the applicability and usefulness of the model, and the researchers, who focus on the validity of the artifact and the process rigor (SONNENBERG; BROCKE, 2011). The target is to obtain an average perception that the proposed model is potentially better than the current model or way of working. Until this is achieved, this methodology would be iterated back to activity 3 (design and development) to improve the effectiveness of the artifact. After that, this work proceeded to the communication, the last phase of the DSRM.

These aspects were evaluated based on a survey (through Google Forms) sent to stakeholders that took part in the demonstration phase. As shown in Figure 23, Figure 24, and Figure 25, they were asked to rate the model's "completeness", "fidelity with real-world phenomena", and "internal consistency" based on a scale from 1 (poor) to 5 (excellent).

Figure 23 – Evaluation of model's completeness



Completeness

Whether the model and the data structure lack some items or whether its usage requires customization
(o quão completo é o modelo proposto referente aos requisitos estabelecidos)

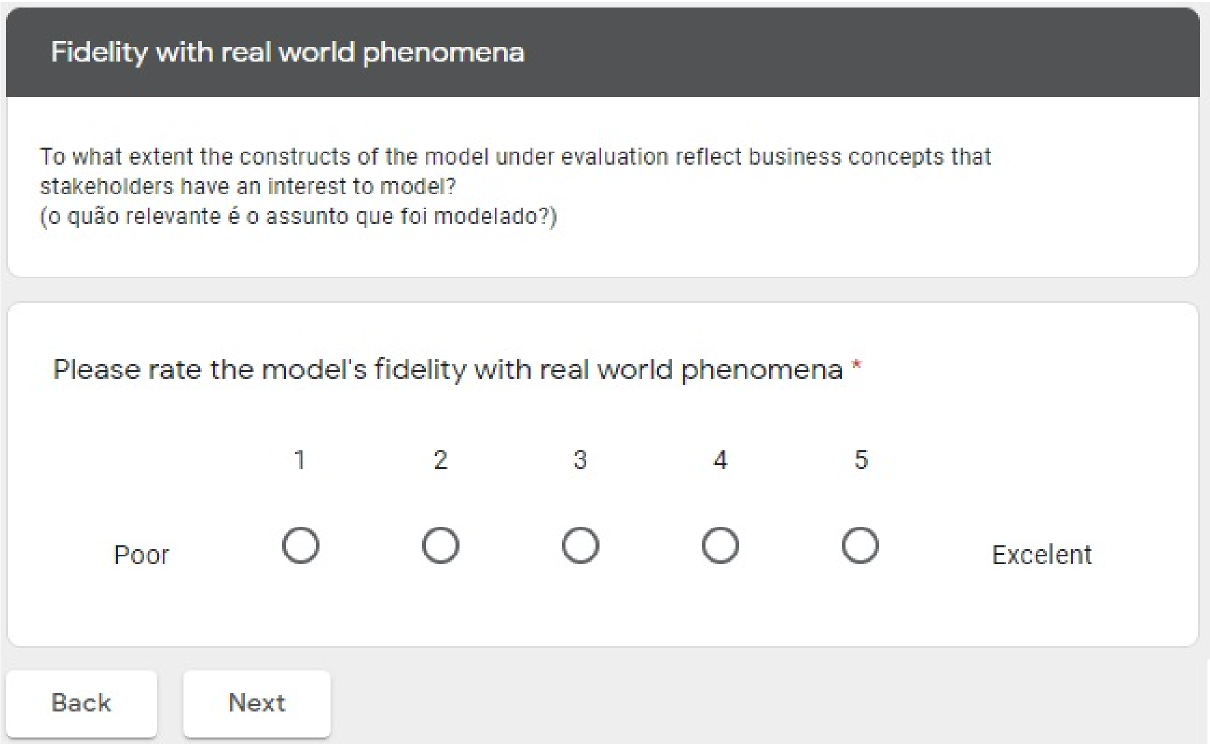
Please rate the model completeness * 1 point

	1	2	3	4	5	
Poor	○	○	○	○	○	Excellent

Back Next

Source: The author (2021)

Figure 24 – Evaluation of model's fidelity with real-world phenomena



Fidelity with real world phenomena

To what extent the constructs of the model under evaluation reflect business concepts that stakeholders have an interest to model?
(o quão relevante é o assunto que foi modelado?)

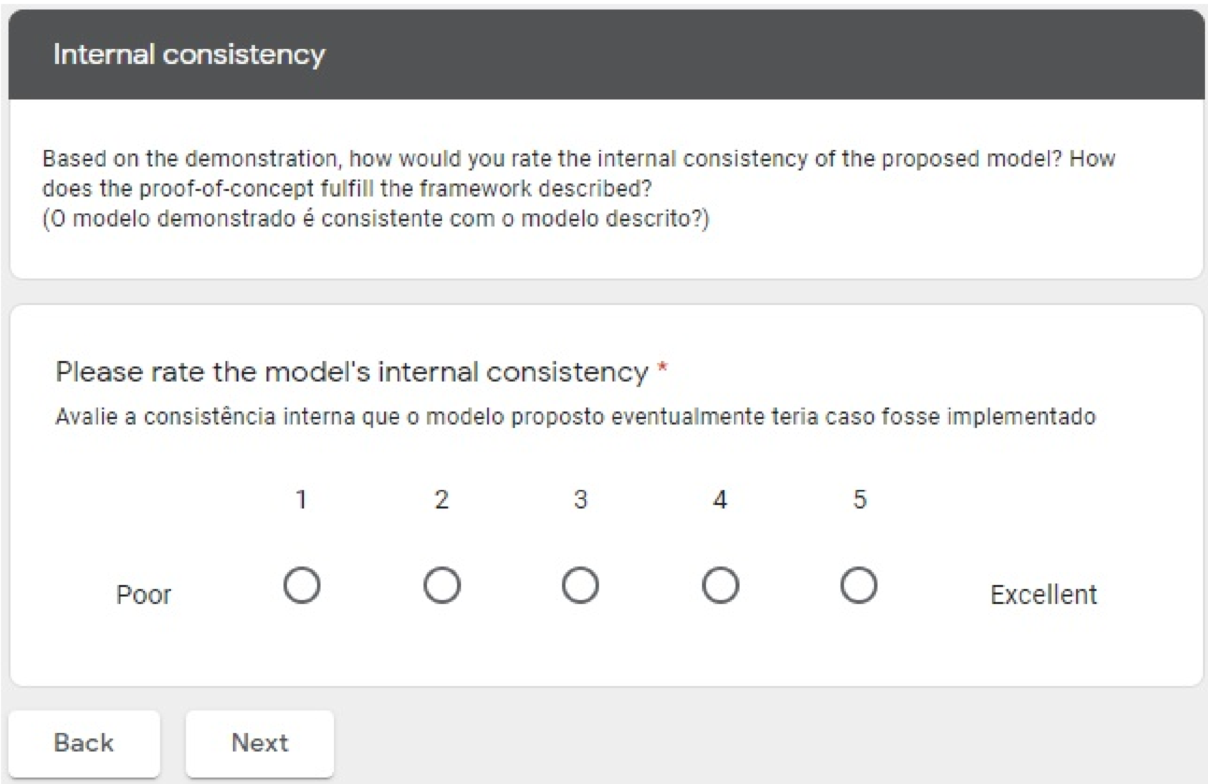
Please rate the model's fidelity with real world phenomena *

	1	2	3	4	5	
Poor	○	○	○	○	○	Excelent

Back Next

Source: The author (2021)

Figure 25 – Evaluation of model's internal consistency



Internal consistency

Based on the demonstration, how would you rate the internal consistency of the proposed model? How does the proof-of-concept fulfill the framework described?
(O modelo demonstrado é consistente com o modelo descrito?)

Please rate the model's internal consistency *

Avalie a consistência interna que o modelo proposto eventualmente teria caso fosse implementado

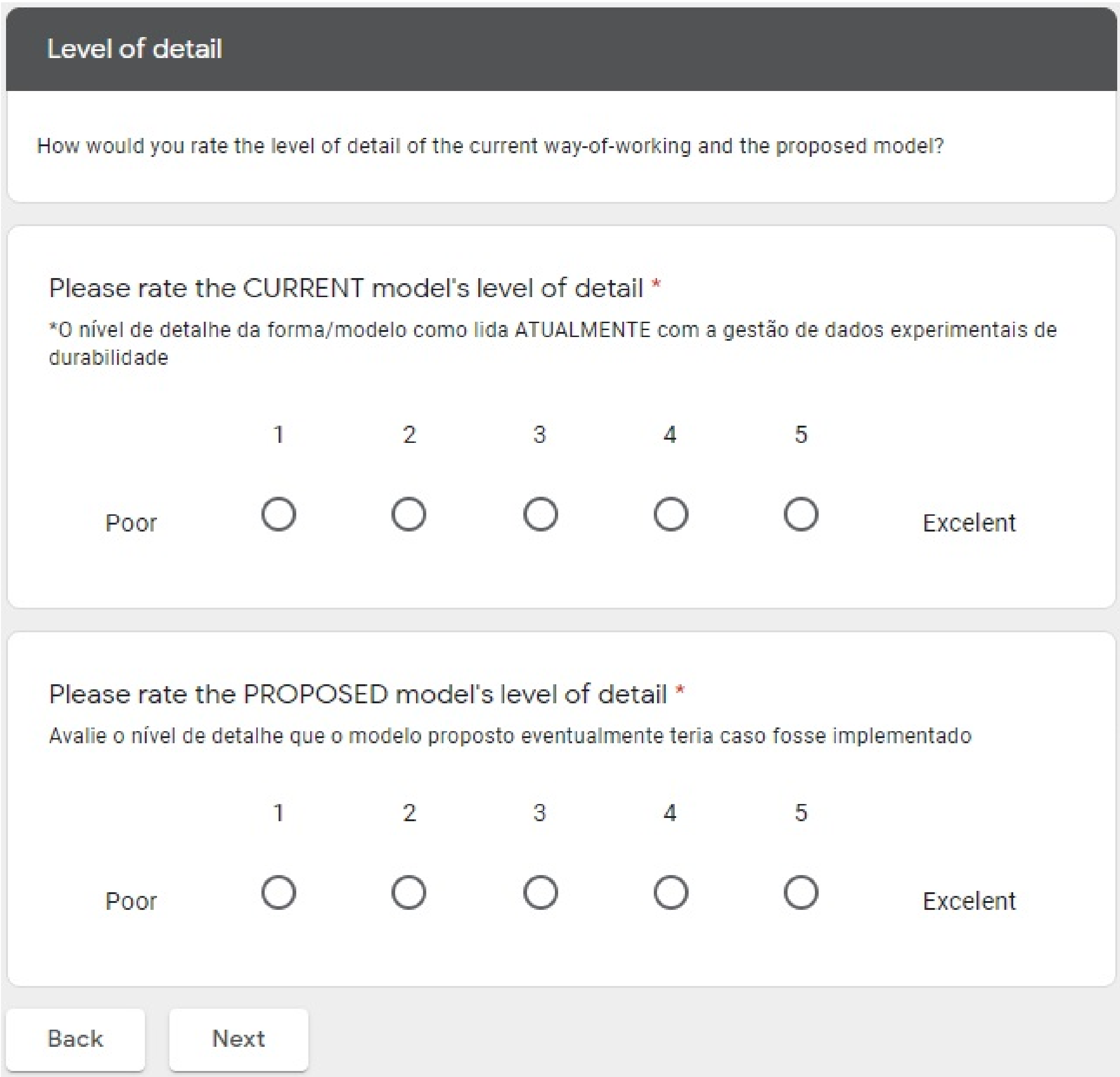
	1	2	3	4	5	
Poor	○	○	○	○	○	Excellent

Back Next

Source: The author (2021)

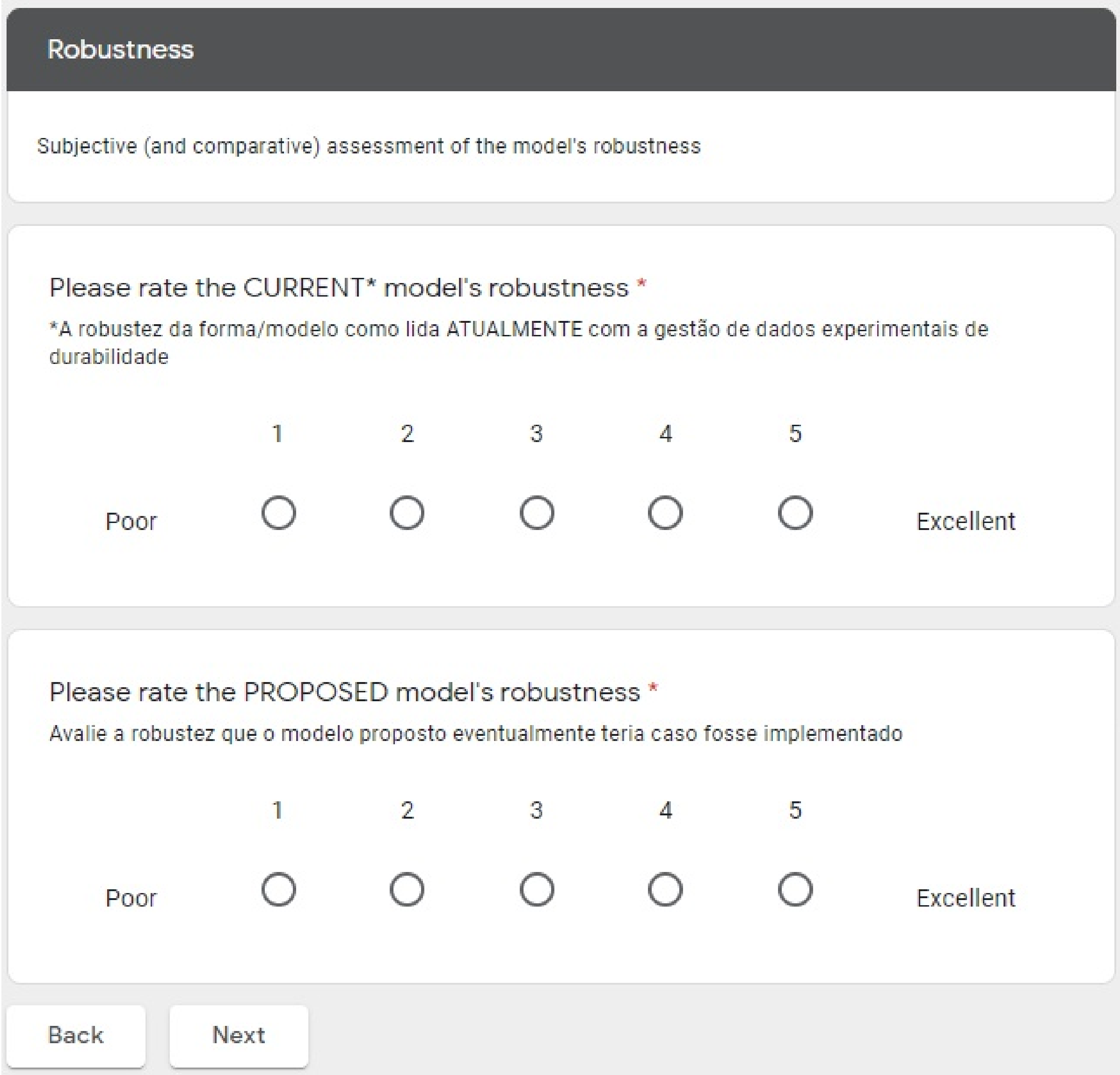
Furthermore, as shown in Figure 26 and Figure 27, regarding the model's "level of detail" and "robustness", they were asked to compare the proposed model with their current model/way of working for the management of durability measurement files.

Figure 26 – Comparative evaluation of the model's level of detail



Source: The author (2021)

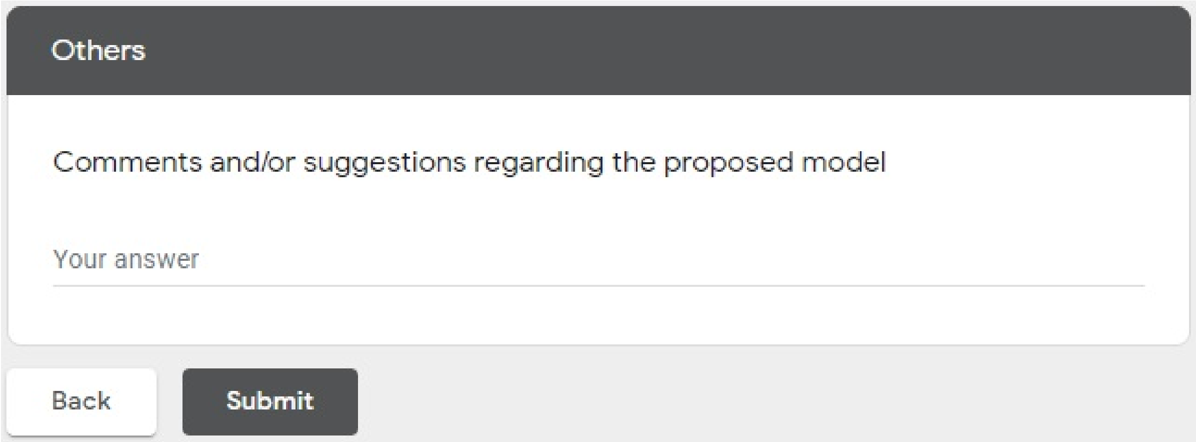
Figure 27 – Comparative evaluation of the model's robustness



Source: The author (2021)

Finally, a text field was available to the evaluators to (optionally) contribute with specific comments and suggestions, as shown in Figure 28 below.

Figure 28 – Free text field for comments and/or suggestions from the evaluators



Source: The author (2021)

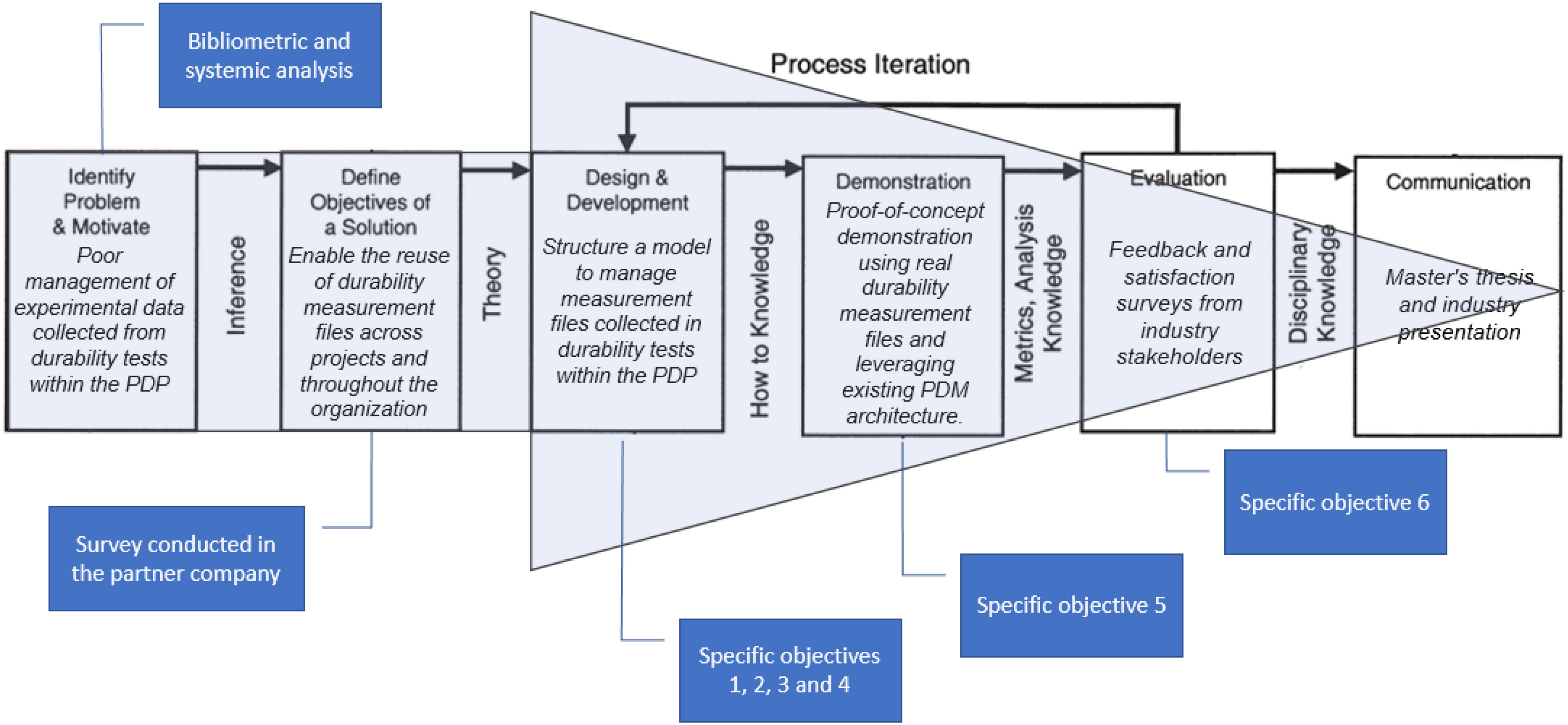
4.4.4 Communication

This last phase includes the communication of the problem and the created artifact. It stresses the relevance of the problem and the value, originality, and potential effectiveness of the artifact to researchers and practitioners. The communication will be done in three different formats:

1) Research results to stakeholders within the collaborating company (presentation format).
2) Master's dissertation (written format – the present report).
3) Master's dissertation defense (presentation format).

Figure 29 summarizes the Design Science Research Methodology process that has been followed and how it relates to the specific objectives delineated in Section 1.2.2. The next section will present the results of the design and development of the proposed artifact.

Figure 29 – DSRM process of the proposed research project



Source: adapted from Peffers et al. (2007)

5 RESULTS: DESIGN & DEVELOPMENT

5.1 MODEL REQUIREMENTS

For three months, several meetings and workshops were carried out with local end-users and global durability experts from the collaborating company. Besides new ideas to the model development (such as defining a standard file format for the measurement files and highlight the need of enhancing metadata to the files), the main output from those exchanges is the understanding of the main constraints that hinder the reuse of durability measurement files within the PDP. This understanding was then framed as the following list of information or filtering requirements that must be included in the model:

- Object (vehicle) specification, including variants and bill of materials (BOM).
- Measurement location: GPS data is mandatory and need to be included within the measurement file.
- The position of the measuring sensors positioning.
- The loading of the tested object.
- The trailer, if any, that have been attached to the truck for the measurement.
- The sample rate of the acquired data.
- The test report number of the measurement.
- The project in which the test was carried out.
- If data from a customer, the application specification (such as timber, mining, etc.).

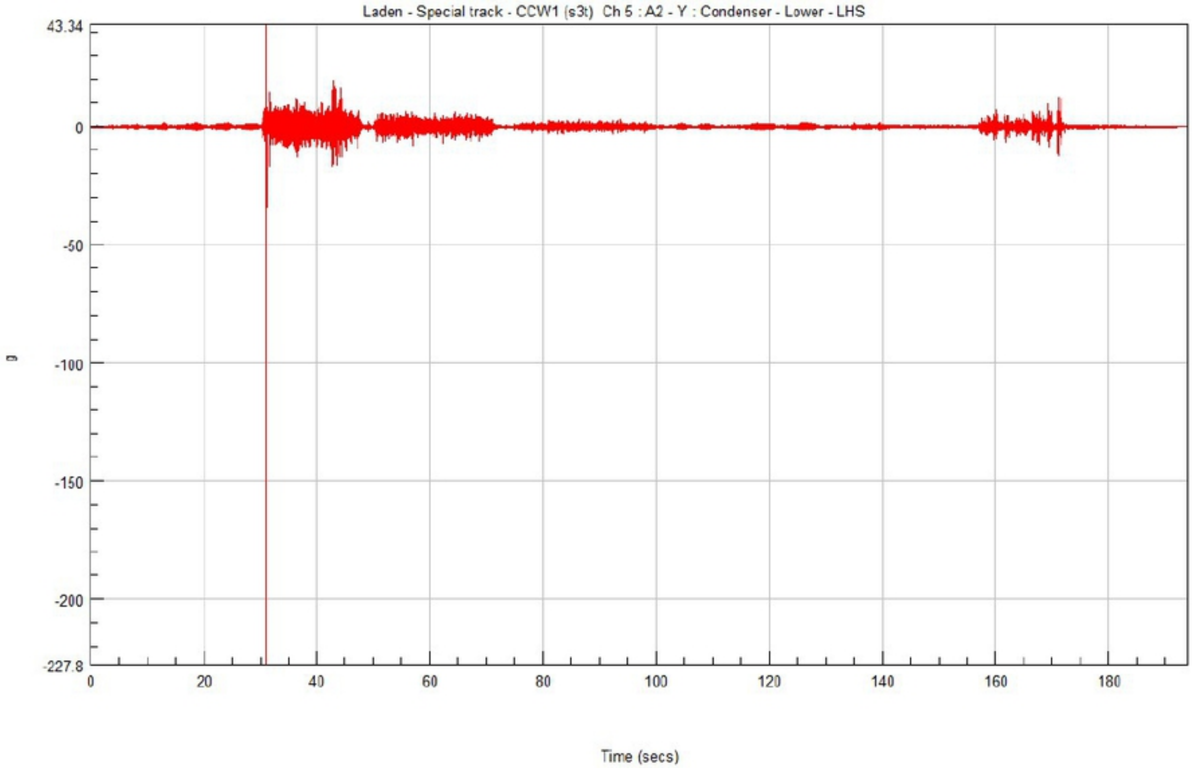
5.2 LOGICAL DATA MODEL

As detailed in the methodology section, the development of this logical data model was built upon relevant literature (section 2.3) and based on the requirements and inputs from the consulted employees of the collaborating company. Although the complete logical model is formalized in a UML notation, the following sections describe in detail some of the most relevant aspects and strategies proposed in the model. Although they can be seen as the pillars for the model, these strategies could also be implemented independently or at different phases. As it is clear that one size does not fit all (AKERS; DOTY, 2013; PATON, 2008), the main goal of this research is not to present a final and inflexible model, but rather to present a model that brings different approaches, concepts and tools that could be used according to the applicable business context.

5.2.1 Signals Version: Laying Out the Lifecycle of Measured Data

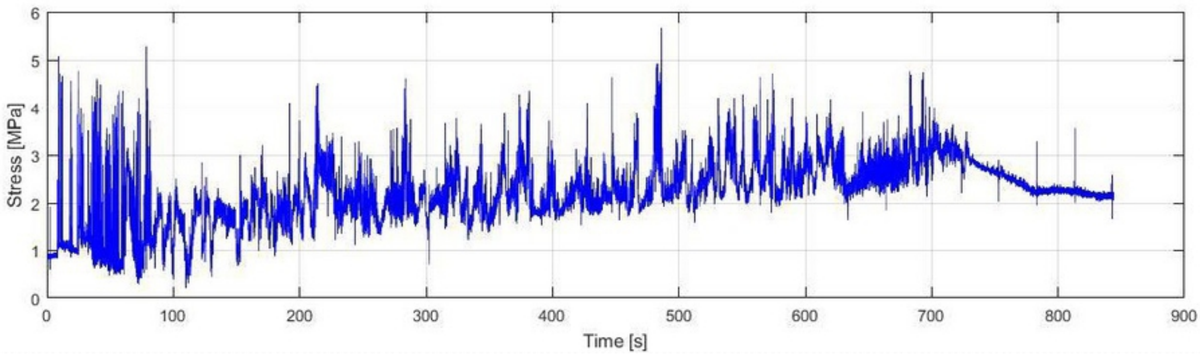
The concept of a data pyramid from Paton (2008) has been brought to discussion during the workshops and exchanges with durability experts. Although analysts and experts usually work on the migration of derived (post-treated) data to results (Figure 11), it is a common understanding that the availability of the raw data is important (BORGHI et al., 2018). The reason for this is that the translation of raw data to derived is not trivial and some important information could be lost (or added) to the data within this process. Data post-treatment is usually done by the test engineer who collected the data. This process includes the removal of measuring errors, such as signal noise, spikes (Figure 30), and drift (Figure 31).

Figure 30 – Example of a spike in a measured acceleration signal (time-series)



Source: The author (2021)

Figure 31 – Example of signal drift in a measured strain (converted to stress) time-series

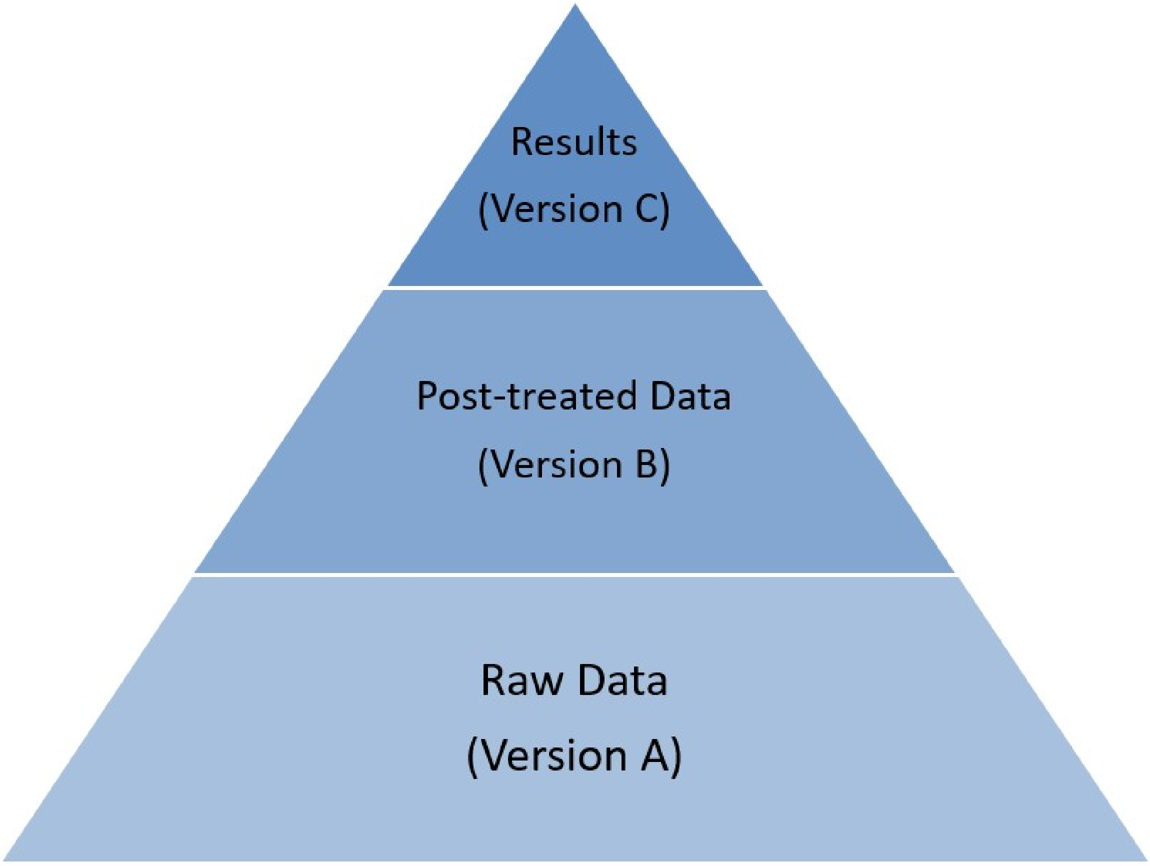


Source: The author (2021)

Although the examples above illustrate clear cases of measurement problems (signal spike and drift), in practice these issues are not always that evident. A small error or wrong assumption in this phase could yield very different or unexpected results. Obviously, the analysts and experts will not be able to recheck this process for every single channel, but they should be able to have easy access to the raw data in case of specific doubts in the results.

To tackle this issue, we bring the concept of structuring data versions (or revisions) as a path to enable the lifecycle of measured data. In Figure 32 we propose an adaptation of the data pyramid presented by Paton (2008), where we group the different stages of the measured data:

Figure 32 – Proposed versions for measured data



Source: adapted from Paton (2008)

By connecting the different stages of the data, the goal is to enable the lifecycle tracking of the measured data. A typical example would be the following: one test engineer would upload the raw data of a measurement (version A) and its respective post-treated data (version B). This version B would then be used for some kind of analysis/calculation by an expert, yielding the results (version C). Although starting from version B, the analyst would still have easy access to version A, in the case of any doubts or questions. Another important point is that different revisions shall be possible to be uploaded within each version (e.g. B01, B02, etc.) – representing the maturity of the signal and allowing the tasks to be done in several sittings if needed.

However, for this to work, a set of standards shall be defined to delimit and contextualize the different data versions:

Raw data: Minimum data processing is expected to be done with the raw data. However, it does not mean that the data would be exactly the one acquired during the data acquisition (DAQ). Raw data preparation may include:

- Splitting the time-series data into different segments that are desired.
- Changes in channel naming and order.
- Correction of acceleration directions (X, Y, and Z) to match the company's standard.
- Units conversions.
- Any other change that does not impact the data signal.

Post-treated data: The data processing that is expected to be done by the same person that performed the measurement. The goal is to provide a high-quality signal, free of measurement errors. The post-treatment may include:

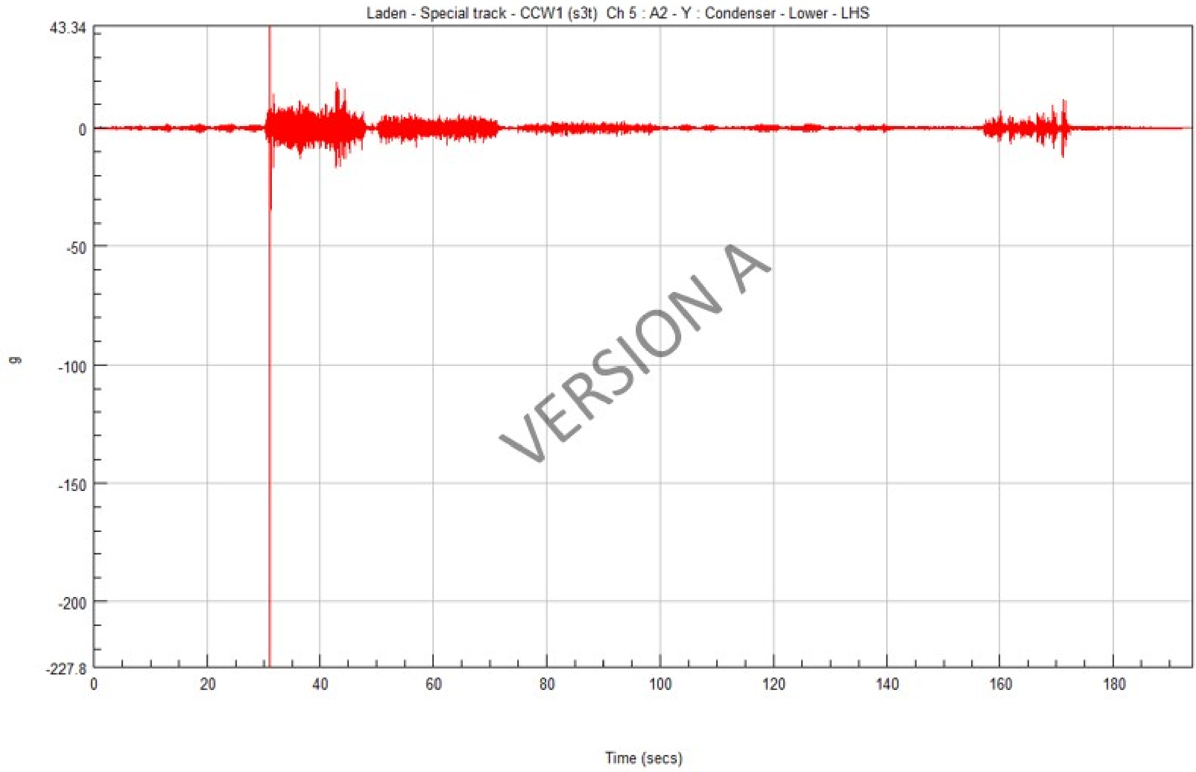
- Addition of static offsets (to compensate for the static load, for instance).
- Removal of signal drift, spike, or noise.
- Data filtering.

Results data: Derived from the calculation of the post-treated data. They are generally performed by analysts or data experts and may include:

- Stress results, calculated from strain gauge data.
- Damage / pseudo-damage results;
- Frequency spectrums;
- Other types of durability analysis.

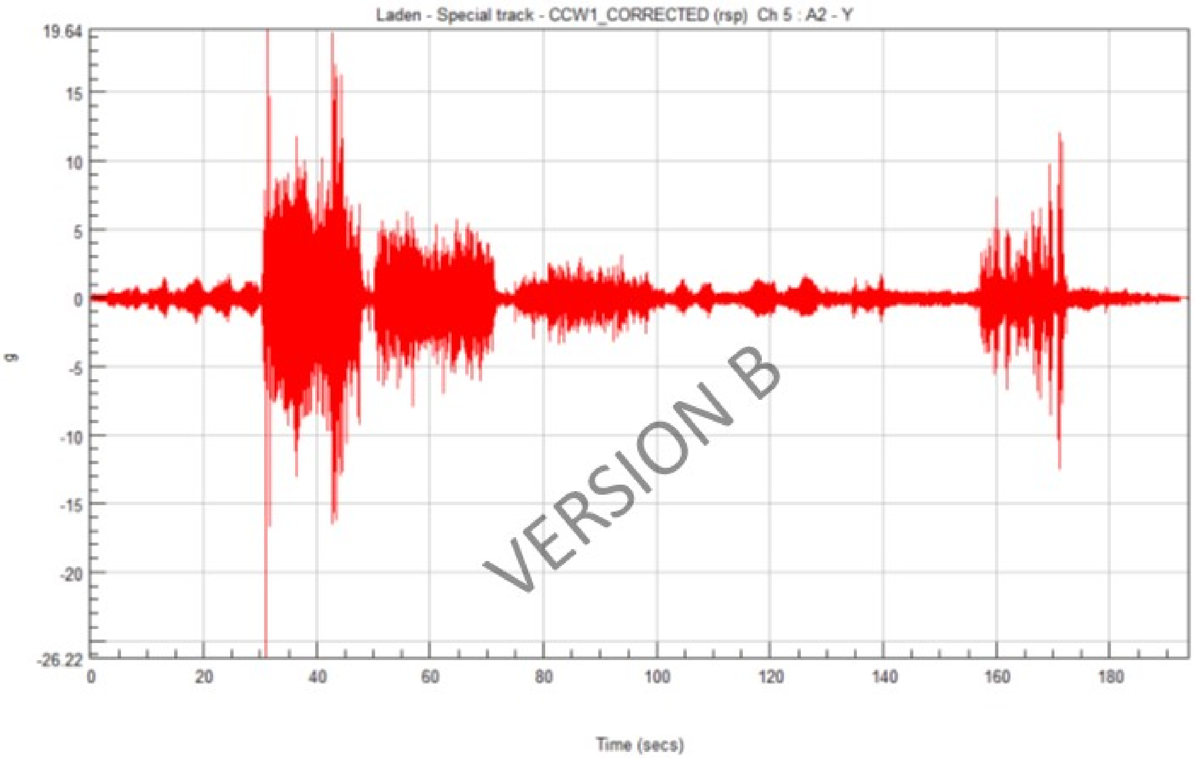
Figure 33, Figure 34, Figure 35 illustrate the proposed versions for the measured data. Version A, in Figure 33, is the raw acceleration data from a measurement performed in a proving ground. Version B, in Figure 34, is the post-treated data of the same signal – as it can be noted, the spike has been removed from the raw data signal. While version C, in Figure 35, presents the results of a shock response spectrum analysis that has been carried out on version B.

Figure 33 – Raw acceleration data (spike around the 30s of the time series)



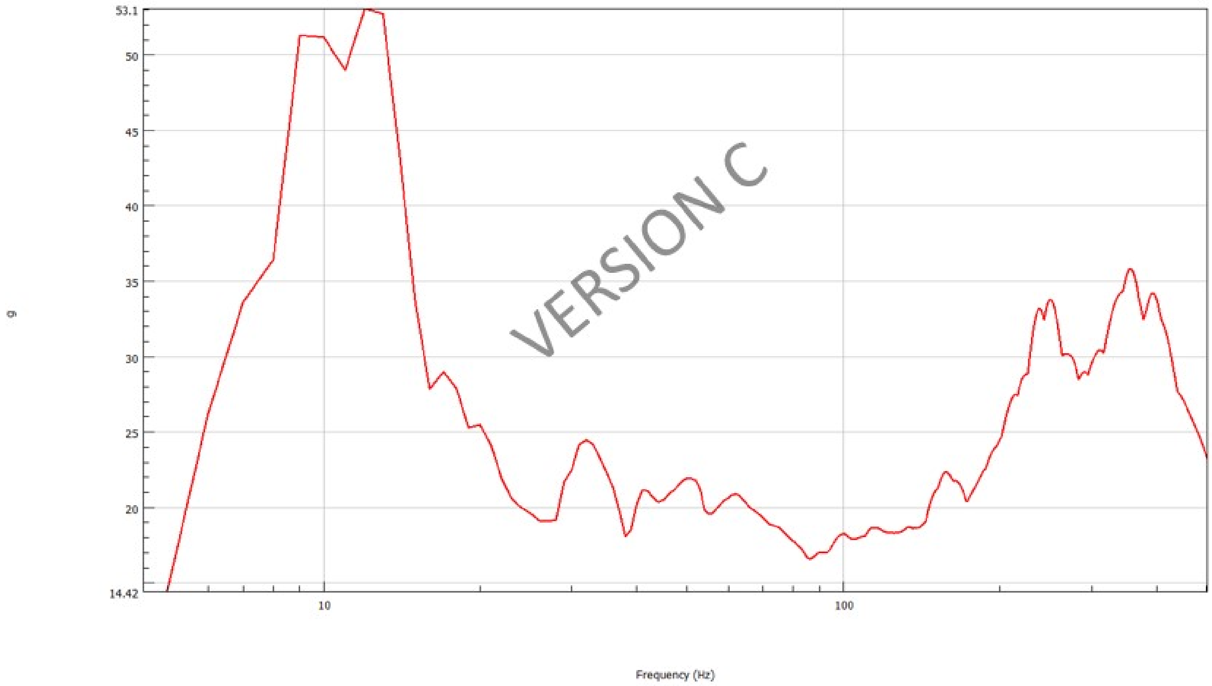
Source: The author (2021)

Figure 34 – Post-treated acceleration data (spike removed)



Source: The author (2021)

Figure 35 – Results data: shock response spectrum analysis based on the acceleration signal from



Source: The author (2021)

This proposed concept also tackles Auweraer and Leuridan's (2005b) concern regarding the validation of measured data and their argument that test data problems shall be verified on several levels.

5.2.2 Approval Process for Measured Data and Roles Description

Given the concerns regarding signal quality and reliability, and based on the work from Ji et al. (2017), we propose to include an approval process for the uploading of the test data into the model. This is not supposed to be different from the drawing checks that are performed in CAD data and, even closer, approval of test reports. The goal here is not to add a bureaucratic and hierarchical process, but rather to foster a structured peer-review check for all the data entering the system. Regardless of the model or management system, reliable data is fundamental. By that, we assume an increase in the discussions regarding the data/analysis and, therefore, expect that signals with errors will be less likely to be found.

From a complete model perspective, we propose the identification and split of different user roles as follows:

Test engineers have rights to:

- Upload a new measurement file for approval.
- Update data version or revision for approval.
- Search, view, and download data within the database.

Durability analysts/experts have rights to:

- Update data version or revision for approval.
- Approve/reject data.
- Search, view, and download data within the database.

Approvers have rights to:

- Approve/reject data.
- Search, view, and download data within the database.

Data viewers have rights to:

- Search, view, and download data within the database.

5.2.3 Enhancing Metadata

One key approach of the proposed model is the metadata enhancement within the measured data. As highlighted by Borghi et al. (2018) and Jones, Pryor and Whyte (2013), metadata provides context to raw data, therefore playing a vital role in data retrieving and reuse. It is important to highlight that, in this research, we split the metadata information into (1) test metadata and (2) channel metadata. In both cases, the objective remains the same: stamp and structure relevant information (that is normally documented in the test reports) to the measurement files. Figure 36 and Figure 37 present some examples of this relevant information that, currently, is only documented within the test reports.

Figure 36 – Extraction of test report: information regarding the specification of the test object

Page 3

2.4 Test truck

Picture 2 – test truck loaded

Table 2 – test truck data from "VDA"

Vehicle	VM Rigid, 32 ton; 6*4
Chassis / ID	E163593
Engine	MWM7B330; 7 litres, 325 hp
Transmission	AT2612F, I-Shift, 2600 Nm; 12 speed
Front axle load	8.0 tonnes
Tyre	12; R22.5; Michelin X Works
Steering gear ratio	20.0 (ZF8098)

Source: The author (2021)

Figure 37 – Extraction of test report: photo of accelerometer positioning in the test object

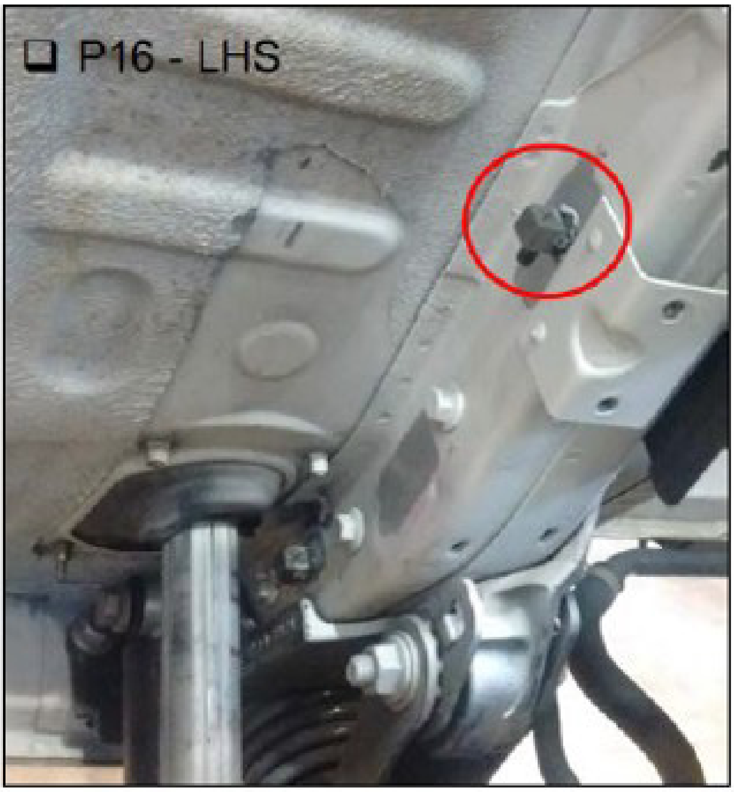


Figure 29 – Transducer position on BiW – run 2

Source: The author (2021)

5.2.3.1 Enhancing test metadata

Test metadata are common conditions that apply to all measured channels, such as the measuring date, ambient temperature, test object, and many others. The proposed test metadata to be included in the model tackle the concerns and requirements established in section 5.1. Each parameter to be added as test metadata is detailed below:

- Measurement ID: a unique identifier of the measurement file. This identifier must be automatically generated by the physical model and attached to the test metadata.
- Vehicle ID and phase: a unique identifier of the test object (vehicle, in the case of this research) connected to the company's Test Object Management tool. As a prototype or vehicle can undergo different retrofits, the version information is relevant as well. The goal is to connect the collected measured files to a specific list of variants and bill of materials (BOM) from the test object.
- Front axle load (FAL): the measured (or estimated) load in the front axle of the test vehicle.
- Rear-axle load (RAL): the measured (or estimated) load in the rear axle of the test vehicle.
- Gross Combination Weight (GCW): the measured (or estimated) GCW of the test vehicle (front axle load + rear axle load + trailer load).
- Trailer: Unique identifier of the trailer (if any) used for the measurement. The trailer ID shall be selected through a drop-down list. The system administrator will be responsible for the inclusion/exclusion of the trailers from the listing. The goal of this is to document in which trailer the DAQ has been performed, as different trailers might have a different impact on the measurements (as models vary the number of axles and the center of gravity).
- Report ID: Unique identifier of the test report that describes the performed measurement.

- Project: The code of the project that requested the performed measurement.
- Test location: Selection of the test location (e.g. potholes track in a specific proving ground). For specific cases, reverse geocoding (read Yin et al., 2020) based on the GPS data might be an (advanced) alternative.
- Application (if field): Selection of customer application if “Field data” was selected on “Test Location” (e.g.: timber, long haul, mining, etc.).
- Comments: Free text field for comments regarding the data.

These metadata can be included before, during, or even after the DAQ. This will depend on the acquisition equipment and the test engineer’s preference. The only exception shall be the measurement unique identifier, which must be automatically generated by the system. The important is to ensure that this information is included in the raw data uploaded in the database.

5.2.3.2Enhancing channel metadata

Channel metadata is the information related to a specific measured channel, in other words, that does not apply to all channels that together comprise the measurement file. Basic channel metadata are generally the following:

- Channel name.
- Channel number.
- Sample rate.
- Unit of measurement.

In addition to this standard metadata, the developed logical data model also proposes the inclusion of a metadata field for the channel version and revision – as the concept described in section 5.2.1. In this way, within one measurement file, different channels might be on different versions and revisions. That would be the case, for example, if the data could be post-processed (migrated from version A to B) for some channels only.

From this point, more precise metadata shall be added according to the signal type. So basically, the signals shall be split as follows: strain, load, acceleration, displacement, and GPS data.

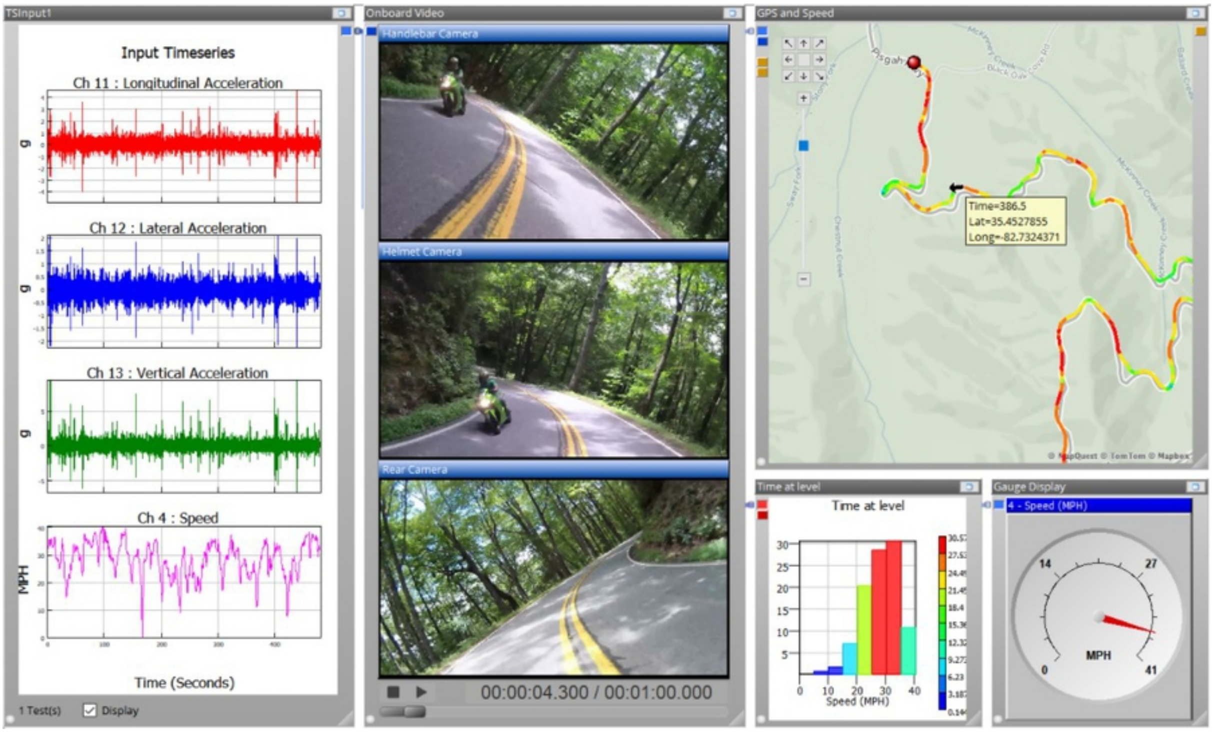
5.2.3.2.1 GPS data

Although GPS data (latitude, longitude, and speed time-series) must be present in all measurement files, no specific metadata is defined for these channels. On the other hand, an engine must be capable of detecting these channels (either by name, channel number, or unit) and plotting their data on a map. The signals database must be able to connect to an Application Programming Interface (API) so these plots can be available for the data viewers. Several engines and APIs are available for this physical implementation, such as Microsoft MapPoint, Bing Maps, Google maps, the open-source Open-StreetMap, and many others.

From an end-user perspective, there are two main functionalities of the model that rely on this process:

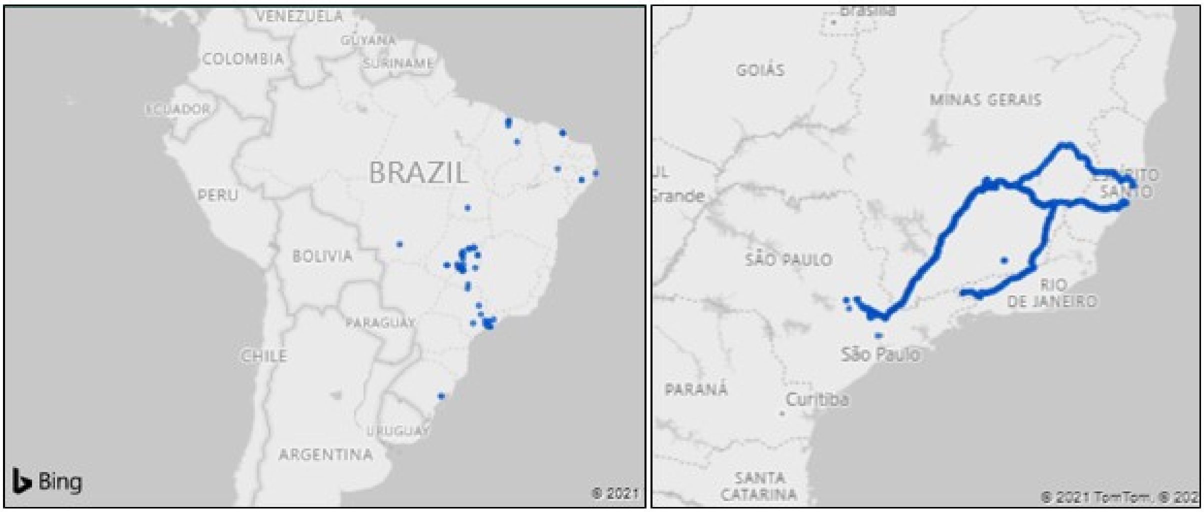
1) To be able to easily visualize time-synced GPS and channel data, allowing a better understanding of what happens in the real world – and where it happened (example in Figure 38).
2) To be able to search for experimental data on points or regions of interest – for instance, looking for data collected on a specific road or specific country (example in Figure 39).

Figure 38 – Time-synced GPS and channel data example



Source: Hottinger Brüel & Kjær (2017)

Figure 39 – Example of a dynamic map where measurement files could be retrieved based on their location



Source: The author (2021)

5.2.3.2.2 Durability data

Currently, the discretization of the sensors positioning in the measurement is only available in an unstructured format within their respective test report or channel name. Some examples are highlighted in Figure 40, Figure 41, and Figure 42 below:

Figure 40 – Example of channels list - sensor position is described in the channel naming but lacks precision

Table 6 – Channels list.

Channel Number	Channel Title	Type	Unit
1	Acc 1 - ADB150 Tank Upper Front LHS - X	Acceleration	g
2	Acc 1 - ADB150 Tank Upper Front LHS - Y	Acceleration	g
3	Acc 1 - ADB150 Tank Upper Front LHS - Z	Acceleration	g
4	Acc 2 - ADB150 Tank Upper Rear LHS - X	Acceleration	g
5	Acc 2 - ADB150 Tank Upper Rear LHS - Y	Acceleration	g
6	Acc 2 - ADB150 Tank Upper Rear LHS - Z	Acceleration	g
7	Acc 3 - ADB150 Tank Upper Front RHS - X	Acceleration	g
8	Acc 3 - ADB150 Tank Upper Front RHS - Y	Acceleration	g
9	Acc 3 - ADB150 Tank Upper Front RHS - Z	Acceleration	g
10	Acc 4 - ADB150 Tank Upper Rear RHS - X	Acceleration	g
11	Acc 4 - ADB150 Tank Upper Rear RHS - Y	Acceleration	g
12	Acc 4 - ADB150 Tank Upper Rear RHS - Z	Acceleration	g
13	Acc 9 - Front Frame RHS - X	Acceleration	g
14	Acc 9 - Front Frame RHS - Y	Acceleration	g
15	Acc 9 - Front Frame RHS - Z	Acceleration	g
16	Acc 10 - Lower Corner Radiator RHS - X	Acceleration	g
17	Acc 10 - Lower Corner Radiator RHS - Y	Acceleration	g
18	Acc 10 - Lower Corner Radiator RHS - Z	Acceleration	g

Source: The author (2021)

Figure 41 – Example of accelerometer positioning presented within a test report

Source: The author (2021)

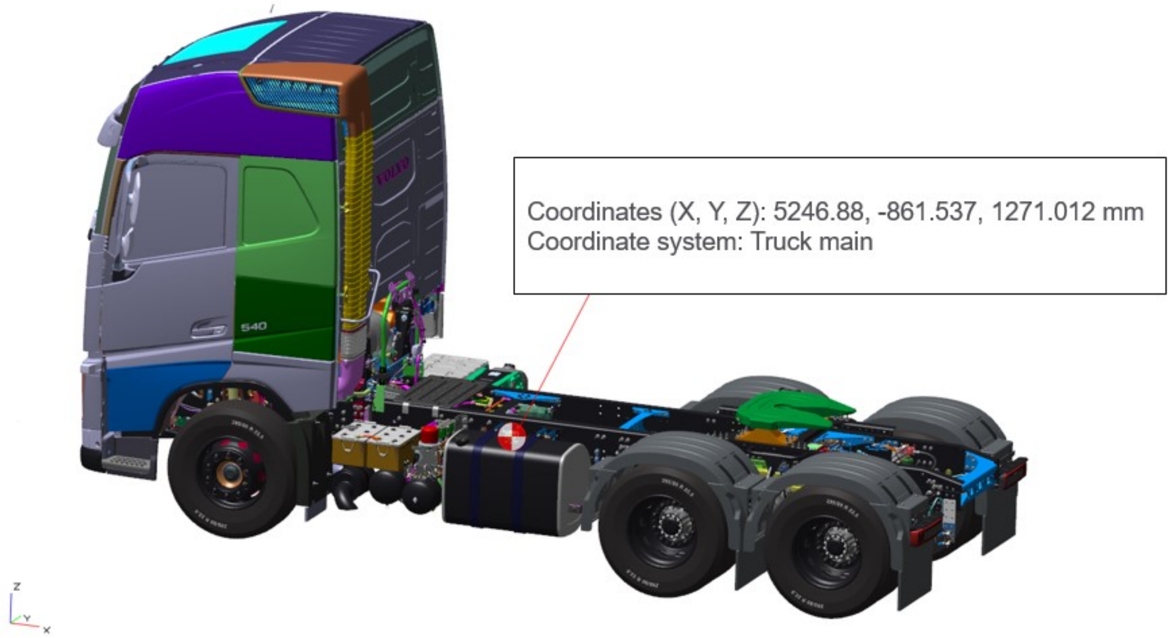
Figure 42 – Example of accelerometer positioning presented within a test report

Source: The author (2021)

This leads to a difficult scenario, where data readers are not able to easily interpret where that data is really coming from. This uncertainty was observed as something that hinders data reusing. For most of the users, it would be easier to request a new test than it would be to dive into all the details and reports of an old measurement. Therefore, similar to some parameters described for test metadata, the objective here is to remove this burden (documentation of the sensors positioning) from the test reports and add another layer to the measurements files themselves – in the format of channel metadata.

The proposed concept is quite simple: discretize the sensors positioning in a test object as X, Y and Z coordinates from a known Cartesian coordinate system and add this information as channel metadata. Obviously, the coordinate system shall be recognizably by the companies' CAD and CAE environments. Note that, in the automotive industry, the layout for a complete vehicle usually serves as a basis for coordinate dimension – one example of these coordinates is shown in Figure 43.

Figure 43 – Example of a truck's DMU and a selected position on the fuel tank - coordinates and coordinate system are described



Source: The author (2021)

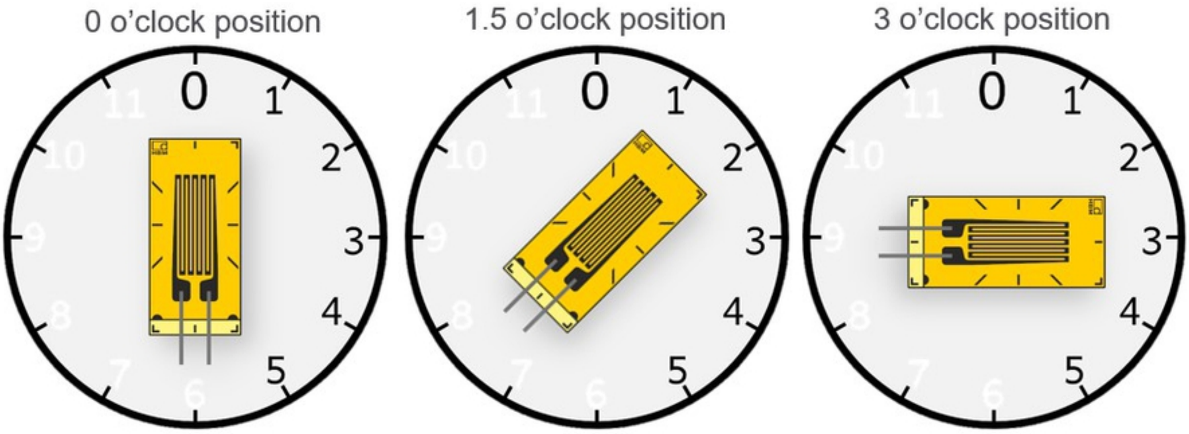
Although the concept is straightforward, some specifications depending on the data type are detailed in the sub-sections below. Moreover, a final subsection provides an outlook of the current challenges and potential technologies that could be used for this type of implementation.

Strain data

Strain data are generally measured by strain gauges, which are available in a variety of models, sizes, and specifications. As per Hoffmann's (1989, p. 208) work, the alignment of these sensors has a major influence on their results. Therefore, the X, Y, and Z positioning from where these sensors were installed are not sufficient. In this concern, we proposed to add an additional metadata field for strain gauges, in which it would be possible to indicate its angular position within the component.

For this, we propose to indicate the strain-gauge alignment as a clock position. The discretization is proposed to be in steps of 15° (equivalent to 30 minutes in the clock) and from 0h to 5h30. There is no need to go beyond 5h30, as a gauge installed at 6:00 o'clock position will present the same results as one installed at 0 o'clock, a gauge at 6:30 o'clock will be equivalent to one at 0:30, and so on. Figure 44 below illustrates the proposed angular positioning standard for strain gauges.

Figure 44 – Examples of angular positioning proposed for strain gauges



Source: The author (2021)

Load data

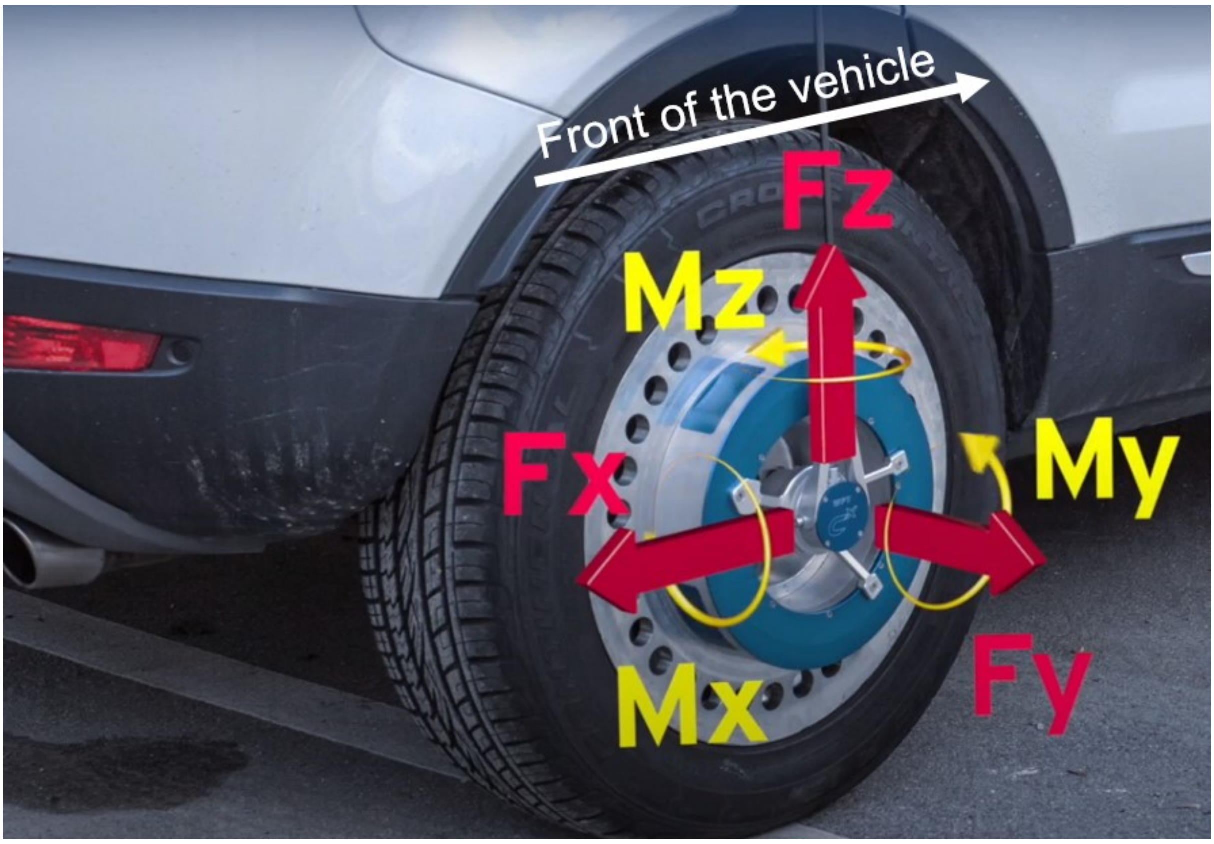
Although load data (forces and moments) are generally measured through the use of strain gauges, the context needed for this type of data is different. In this case, the interest is not on the strain-gauge or transducer alignment/angle, but rather on the nature and direction of the measured load (axial force, shear force, bending moments, etc.).

So, in addition to the specific coordinates and coordinate system of the sensor, we propose the following parameters that could be selected as additional metadata for each load channel:

- Fx (Force component on the x-axis).
- Fy (Force component on the y-axis).
- Fz (Force component on the z-axis).
- Mx (Bending moment on the x-axis).
- My (Bending moment on the y-axis).
- Mz (Bending moment on the z-axis).
- Axial.
- Torsion.
- Shear (optional, as it is rarely measured).

For instance, in the case of a wheel force transducer (Figure 12), each one of its channels must be identified with either Fx, Fy, Fz, Mx, My, and Mz – which are the standard outputs from this sensor. It is important to highlight the importance of the direction of this type of sensor. In our model, we propose the axes conventions illustrated in Figure 45 (adapted from Datron Technology LTD (2016)) below as the standard to be followed.

Figure 45 – Proposed axes convention for load data



Source: Adapted from Datron Technology Ltd (2016)

Another example would be the force measurement on a steering or suspension system. In this case, the usual procedure is to use the system links/rods as load cells. This is done by instrumenting these components with strain gauges and correlating the strain reading with a known applied load in a test bench – as shown in Figure 46.

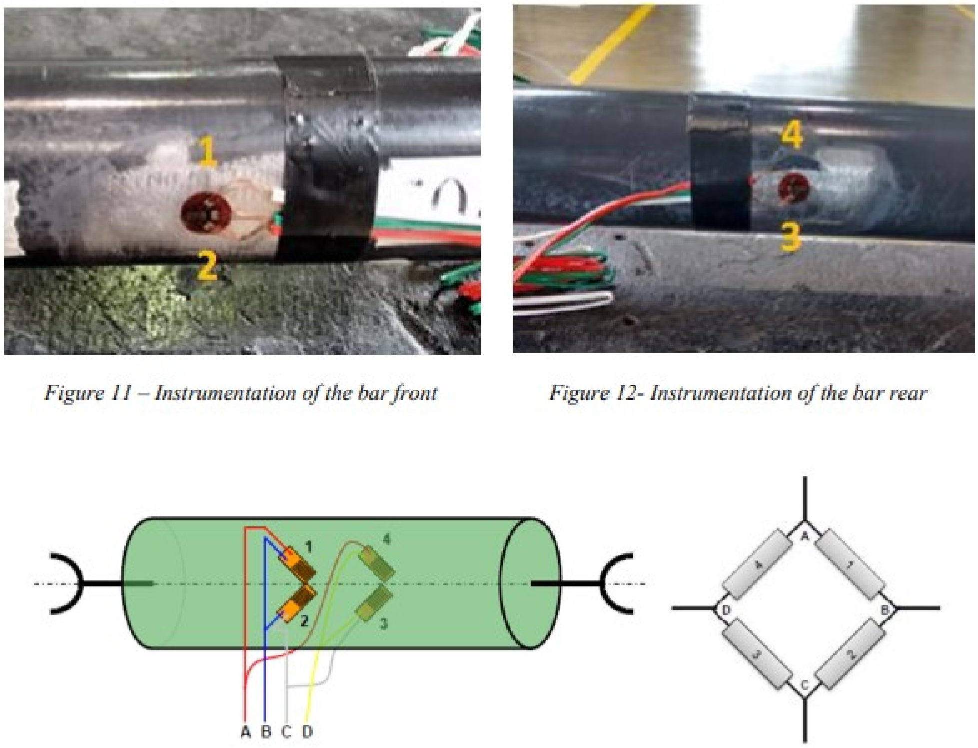
Figure 46 – Steering link calibration for axial forces in a test bench

Source: The author (2021)

These types of loads shall be identified as “Axial” in the metadata, as these signals refer to loads that are being transferred along their respective components’ axis.

The “Torsion” would apply, for instance, for stabilizer bars (or any other bar) instrumented with full-bridge strain gauges mounted for the torsion measurement (HOFFMANN, 1989, p. 218) – as illustrated in the test report extraction in Figure 47 below.

Figure 47 – Extraction of test report: full-bridge strain gauge wired for torsion loads

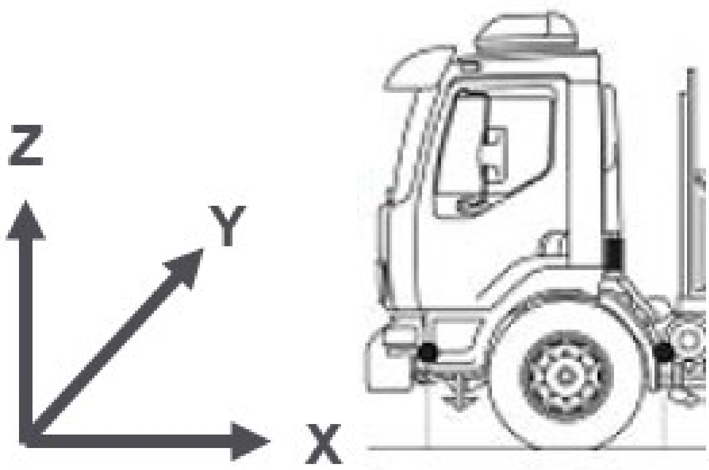


Source: The author (2021)

Acceleration data

Besides X, Y, and Z coordinates from a known cartesian coordinate system, acceleration data require one additional field of channel metadata: the acceleration direction. Our model proposed the following axes convention illustrated in Figure 48 below.

Figure 48 – Proposed axes convention for acceleration measurements



Source: The author (2021)

Displacement data

Different from the other data types, the displacement data cannot be specified from a single Cartesian point because of its relative (from point A to point B) nature. In this case, we've identified that displacement data are generally measured concerning the suspension travel (connected to the vehicle's shock absorbers). Therefore, we propose that suspension metadata shall identify in which suspension (primary or secondary), position, and the side that has been measured. Some examples:

- Primary suspension (axle-frame), front suspension, first axle, left-hand side.
- Secondary suspension (frame-cab), rear suspension, right-hand side.
- Primary suspension (axle-frame), rear suspension, second axle, left-hand side.

5.2.3.2.3 Sensors positioning overview

For each durability data channel, a 3-D cad model shall be generated that represents the sensor type (strain gauge, accelerometer, etc.) and be positioned according to its defined coordinate position and coordinate system. As an example, Figure 49 illustrates the cad model generated from an acceleration signal that had the following channel metadata:

- X coordinate (mm): 5262.95
- Y coordinate (mm): -1024.97
- Z coordinate (mm): 1283.71
- Coordinate system: Main
- Axis direction: Z

Figure 49 – Cad module of an accelerometer placed in the defined X, Y, and Z coordinates

Source: The author (2021)

On the same example, let's suppose that in the test metadata we have the following vehicle ID and phase: PROTO_23:01. Based on this, it is possible to generate the test object DMU, shown in Figure 50 below.

Figure 50 – Generated digital mock-up of PROTO_23:01

Source: The author (2021)

The next step on the proposed model would be to create a new sub-version of the test object, in this case, we will use the index "T" (from test), so the vehicle's unique identification would become: PROTO_23:01_T. Where "PROTO_23" is its identification, "01" is its phase, and "T" is the created test sub-phase. In this new sub-phase created, our model proposes the assembly of all the created instrumentation cad modules and the vehicle's DMU, originating a digital mock-up of the truck and its instrumentation – as illustrated in Figure 51 and Figure 52 below.

Figure 51 – PROTO_23:01_T: digital mock-up of the instrumented test object

Source: The author (2021)

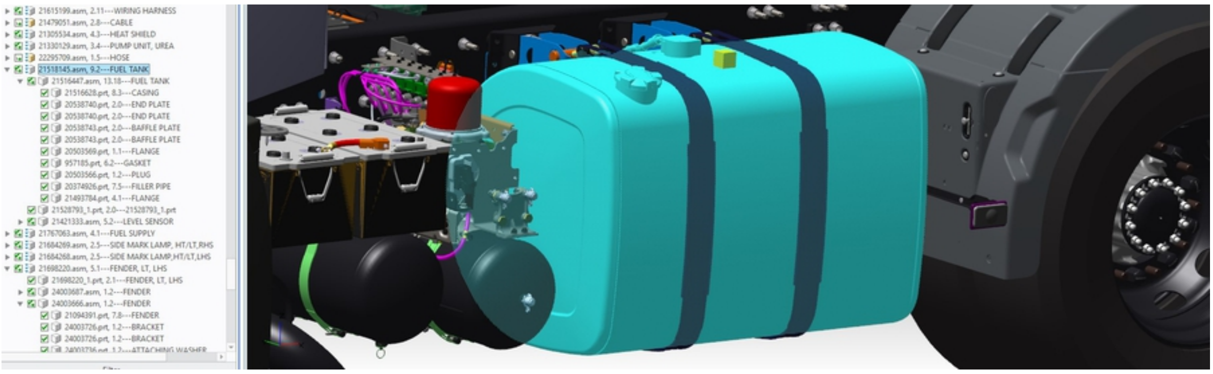
Figure 52 – PROTO_23:01_T: accelerometer placed on the test object's fuel tank

Source: The author (2021)

By that point, we have consolidated a virtual documentation of the test object and its instrumentation - directly connect to the respective measurement files. This would allow data viewers to have quick access and understand of the instrumentation and vehicle specifications – information that is currently partially available within test reports (as discussed in section 5.2.3).

Nonetheless, the model also proposes an additional step, which aims to enhance data search. For each assembled cad-module in the vehicle's DMU, a clustering algorithm shall be implemented to connect the instrumentation to a pre-defined vehicle sub-module. This must be done so each measurement channel can be textually connected to a specific system (or vehicle sub-module) – so users can search for data not only based on vehicle specification or coordinate position, but based on a precise system, such as fuel tank, battery box, front axle, steering arm, reaction-rod, and many others. This information is usually available through the PDM system that generates the DMU, as it is illustrated in the left panel shown in Figure 53 below.

Figure 53 – Product structure data in the generated DMU provides the vehicle sub-modules.



Source: The author (2021)

This "vehicle sub-module" information shall then be feedbacked as channel metadata. In our example, the new channel metadata would then be the following:

- X coordinate (mm): 5262.95
- Y coordinate (mm): -1024.97
- Z coordinate (mm): 1283.71
- Coordinate system: Main
- Axis direction: Z
- Vehicle sub-model: Fuel tank

We acknowledge the challenge for the physical modeling of this step. An alternative, and easier, option would be to have the vehicle sub-module data as a drop-down list in the channel metadata, so the data uploader would manually make the proper selection for each channel. Although we refer to “vehicle sub-module” in this research, this concept applies to any product type. Different companies can potently use any breakdown level of the product structure, the importance lies in following a standard naming and the right precision level that will fulfill the searching requirements. For example, “front axle” provides better precision than “front suspension”. However, the drawback of this precision is the exponential increase of the options available. Therefore, the right balance shall be assessed on a case-by-case basis.

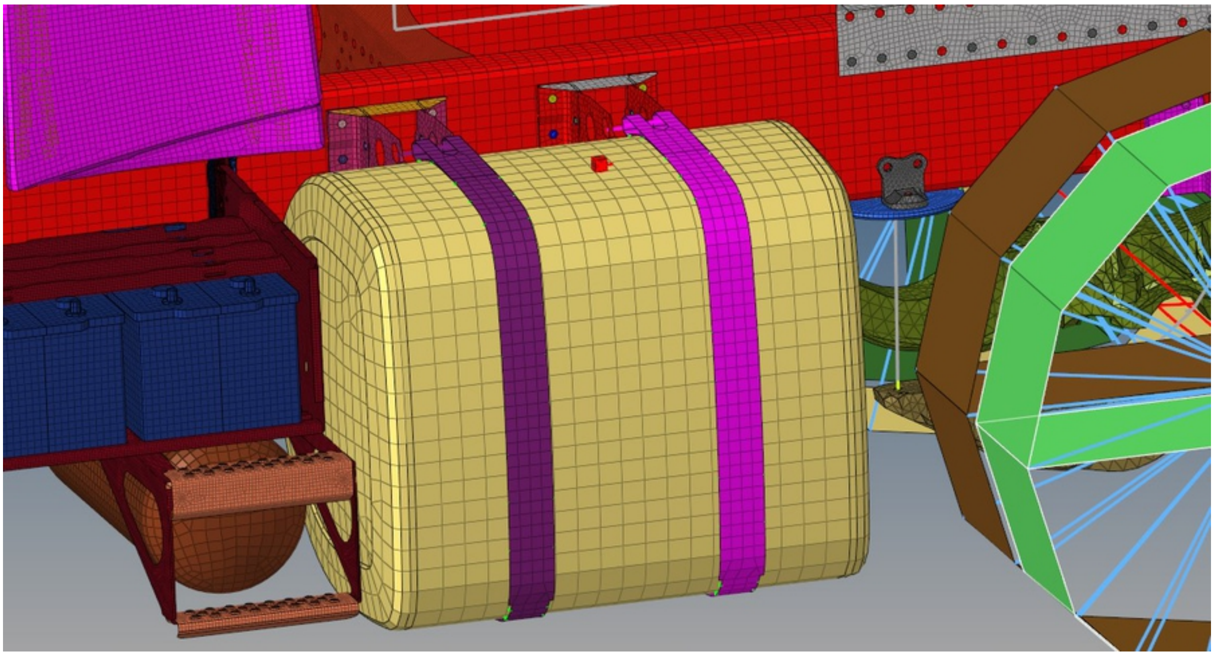
5.2.3.2.4 Sensors positioning – CAE integration

Moreover, by having a consolidated DMU of the instrumented test object, the integration between virtual and physical data could be significantly improved. Virtual models are generally build based on the product data, such as the bill of materials and variants specifications, connected to their respective Cartesian coordinates and coordinate systems. By having the instrumentation connected to the product structure (a similar approach to the one described by Toche et al. (2017)), complete vehicle models could potentially be automatically generated including the desired points for signal extraction. For instance, let us keep the example from the prototype identified as PROTO_23:01_T in section 5.2.3.2.3. Figure 54, Figure 55, and Figure 56 present an example of how a virtual (CAE) model generated based on PROTO_23:01_T’s product structure would look like.

Figure 54 – Complete vehicle model generated based on the product structure data for PROTO_23:01_T

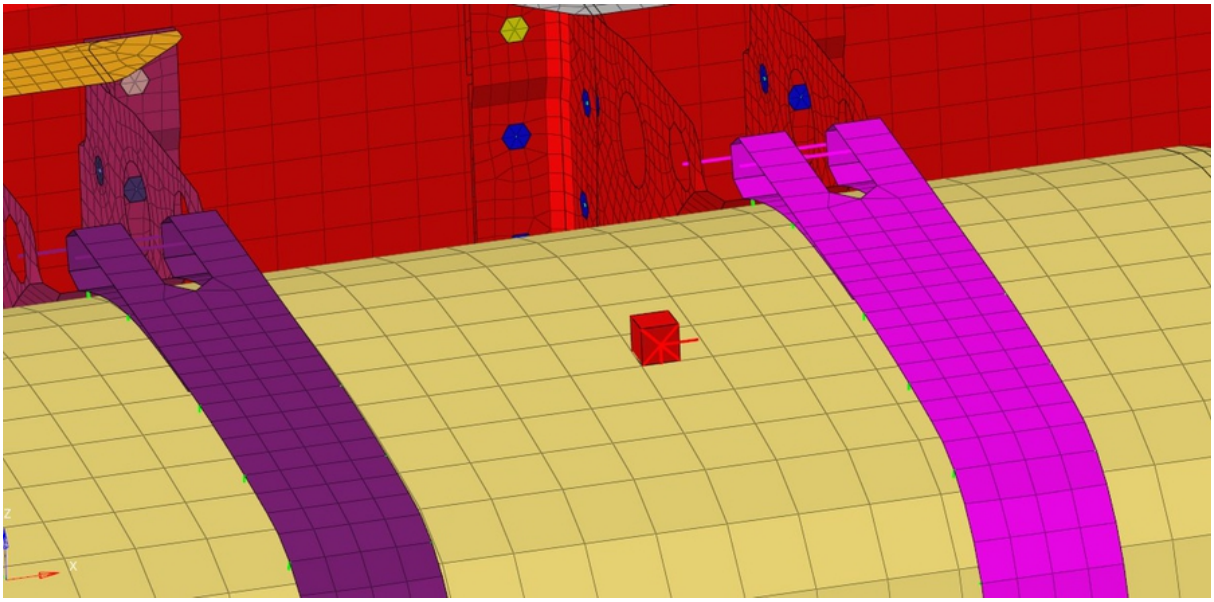
Source: The author (2021)

Figure 55 – Model of the fuel tank (system sub-module) in PROTO_23:01_T



Source: The author (2021)

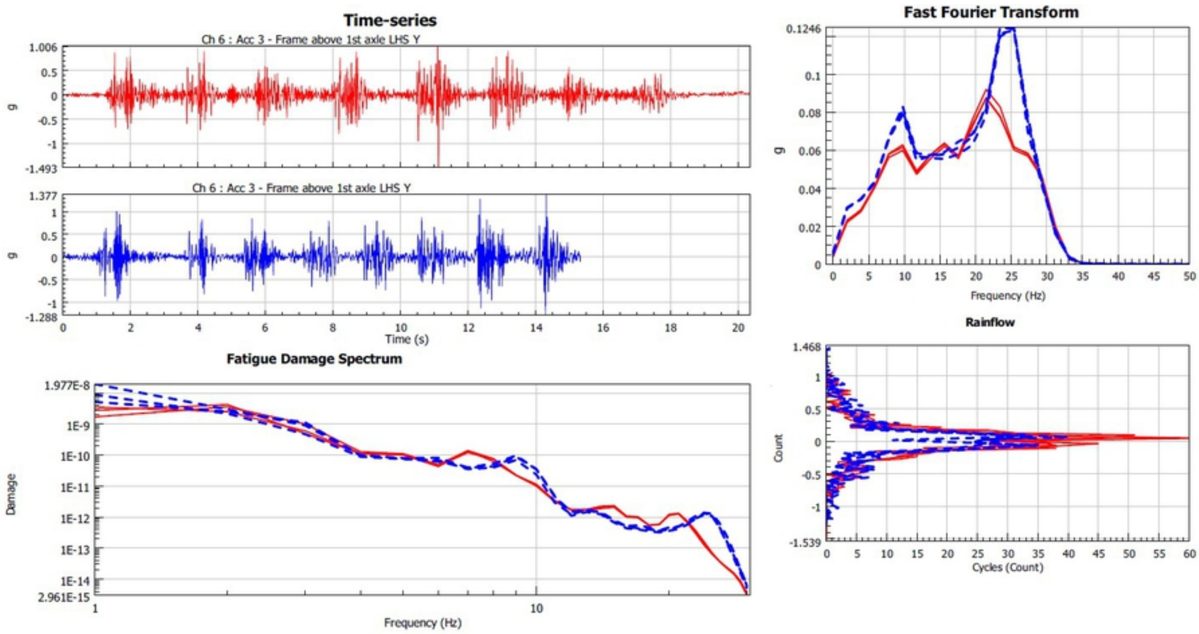
Figure 56 – Virtual accelerometer automatically placed on the fuel tank - similar to the DMU and the physical instrumentation



Source: The author (2021)

With the instrumentation points well defined in the virtual model, the comparison between measured and simulated signals (such as the one illustrated in Figure 57) could potentially be streamlined in a single process – which could save a significant amount of engineer hours.

Figure 57 – Comparison template of physical and virtual data



Marker	Test name	Max	Min	Unit
———	Measurement_repetition_1	1.006	-1.493	g
———	Measurement_repetition_2	1.015	-1.524	g
———	Measurement_repetition_3	1.06	-1.464	g
- - - -	Simulation_loop_1	1.377	-1.288	g
- - - -	Simulation_loop_2	1.33	-1.256	g
- - - -	Simulation_loop_3	1.454	-1.2	g

Source: The author (2021)

Although some effort would still be needed to operationalize and automate this process, the integration of the physical instrumentation into the product structure provides a fundamentally important part of the puzzle (TOCHE et al., 2017). The current process for comparing physical and virtual data is cumbersome: the CAE analyst has to go over all the instrumentation points (unstructured) described in the test report and manually replicate them one by one in the virtual model. Besides the amount of work for big acquisitions, this process is prone to errors because of the imprecision within the translation of instrumentation photos to precise positions in the model. Finally, the proposed concept goes in line with the previously discussed work from Tahera and Earl (2018) and Auweraer and Leuridan (2005a).

Obtaining sensors positioning data

At this point, a question that may arise is how to obtain the Cartesian coordinates from the measuring sensors. Although the model implementation is beyond the scope of this research, this section provides an outlook on possible solutions, techniques, and technologies.

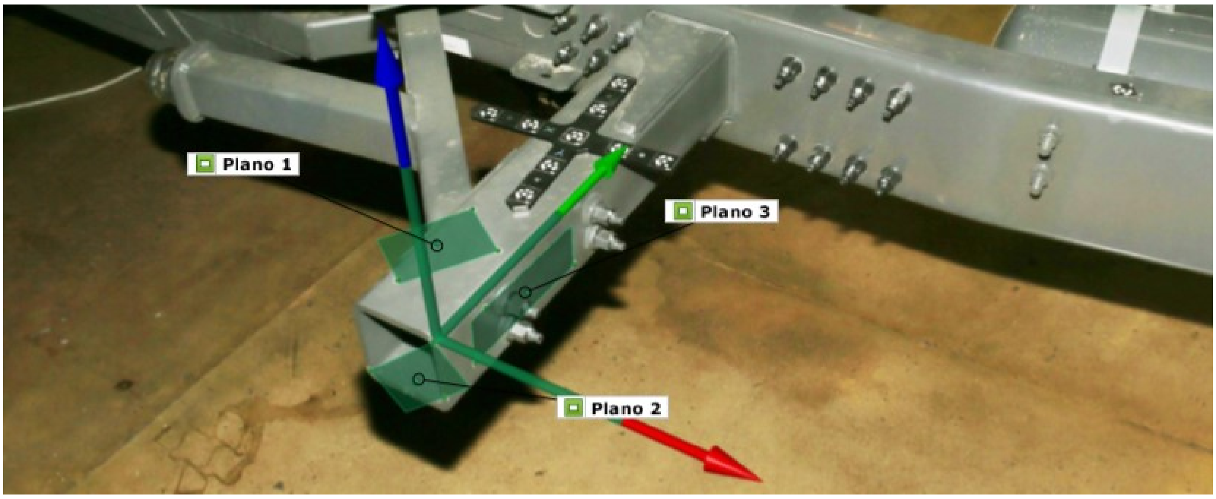
Regarding implementation, the easier way would be to require the test engineers to manually select the instrumentation points from the generated DMU, gather the coordinates for each measuring channel and add them as their respective channel metadata. While the advantage of this technique is its easiness of implementation, the main drawback is the extra work that would be required from the test engineers – especially for measurements with many channels.

A more advance (and technological) option would be to automate this process during the test object instrumentation. Ideally, the DAQ could be capable of automatically receiving the coordinates positions from each sensor and directly stamping this information as channel metadata. The coordinates could be collected using equipment like new 3D coordinate measuring machines (CMM).

As a proof-of-concept, we have tested an optical, and portable, photogrammetry system from GOM called TRITOP CMM, which “measures coordinates of three-dimensional objects quickly and precisely” (GOM, 2020). As with tactile coordinate measuring machines, this system can record the coordinates and their orientation in space for any feature of interest - which in our case are the instrumentation points.

As shown in Figure 58 below, the first step is to define a coordinate system, which is calculated based on the intersection of three plans. For demonstration purposes, we have defined an arbitrary coordinate system in the side frame of a transit bus. In a real use case, this coordinate system should be defined to match one of the standard coordinate systems used for the product structure.

Figure 58 – Example of the coordinate system defined in the side frame of a test vehicle



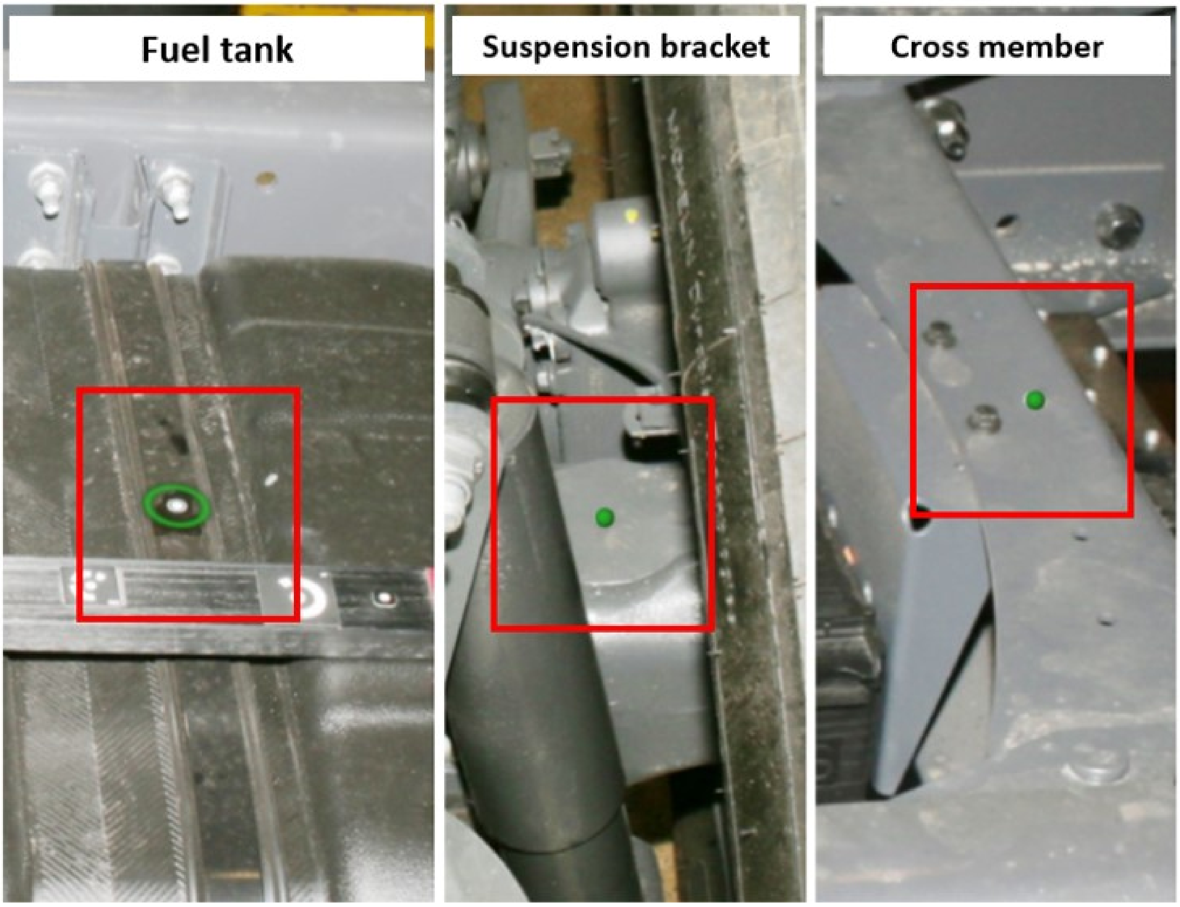
Source: The author (2021)

With the coordinate system defined, three (also arbitrary) instrumentation points were set:

1) In the middle of the vehicle's fuel tank.
2) Above the suspension bracket.
3) In the middle of a cross member.

Physically, those points were identified with coded points on adhesives or magnetic foils. Figure 59 below shows the processed photos generated by the equipment (hardware + software), which identifies the instrumentation points as green points.

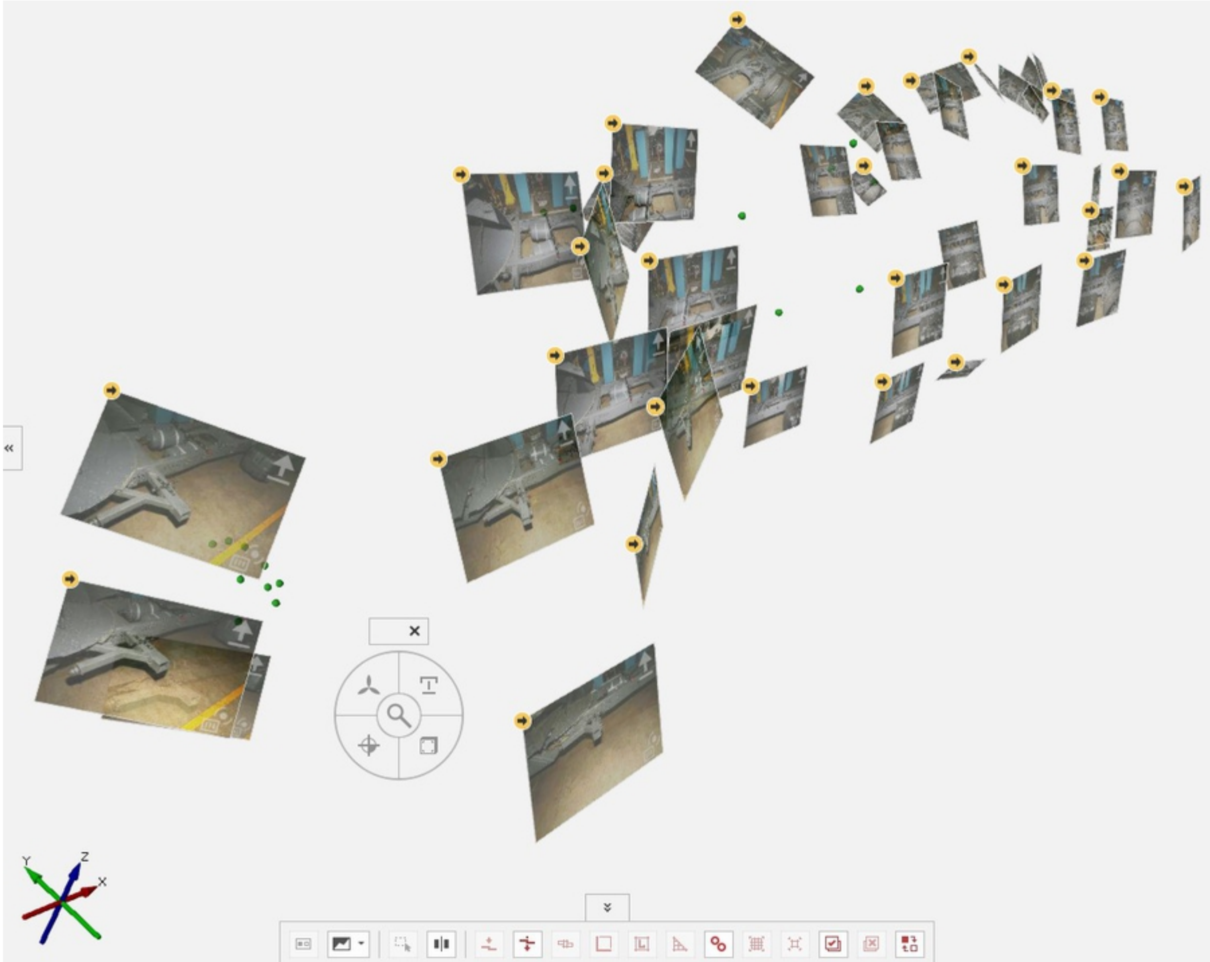
Figure 59 – Arbitrary instrumentation points - photos processed by the equipment



Source: The author (2021)

To gather the desired coordinate positions, several photos are taken by the equipment. Figure 60 presents the cloud points generated by the post-processing software based on the taken photos.

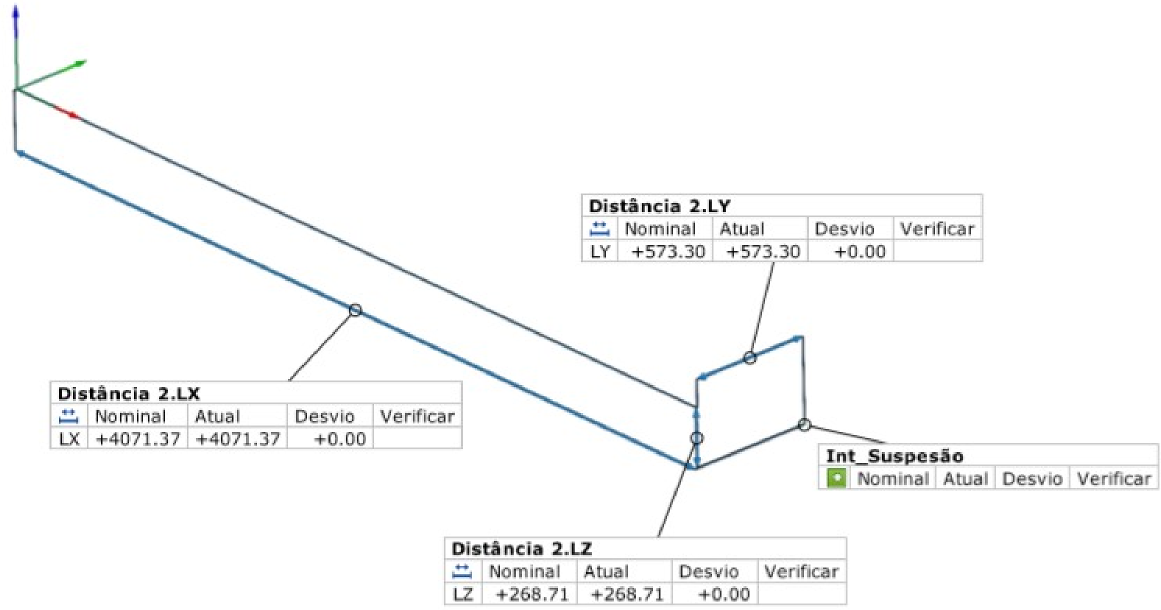
Figure 60 – Cloud points generated based on the taken photos by the photogrammetry system



Source: The author (2021)

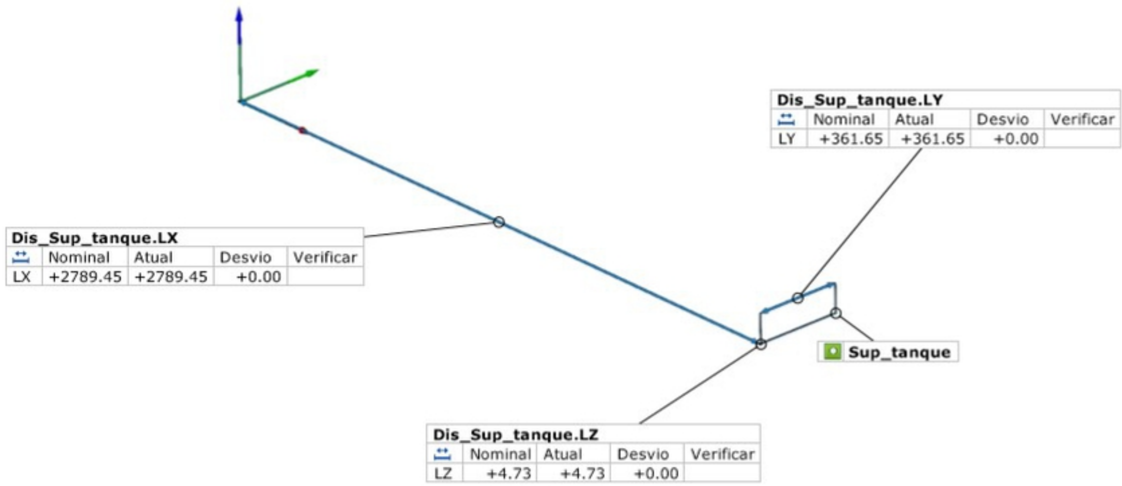
The X, Y, and Z coordinates are then possible to be extracted for each instrumentation point, as it is illustrated in Figure 61, Figure 62, and Figure 63 below.

Figure 61 – X, Y, and Z coordinates of the instrumentation point on the vehicle's fuel tank



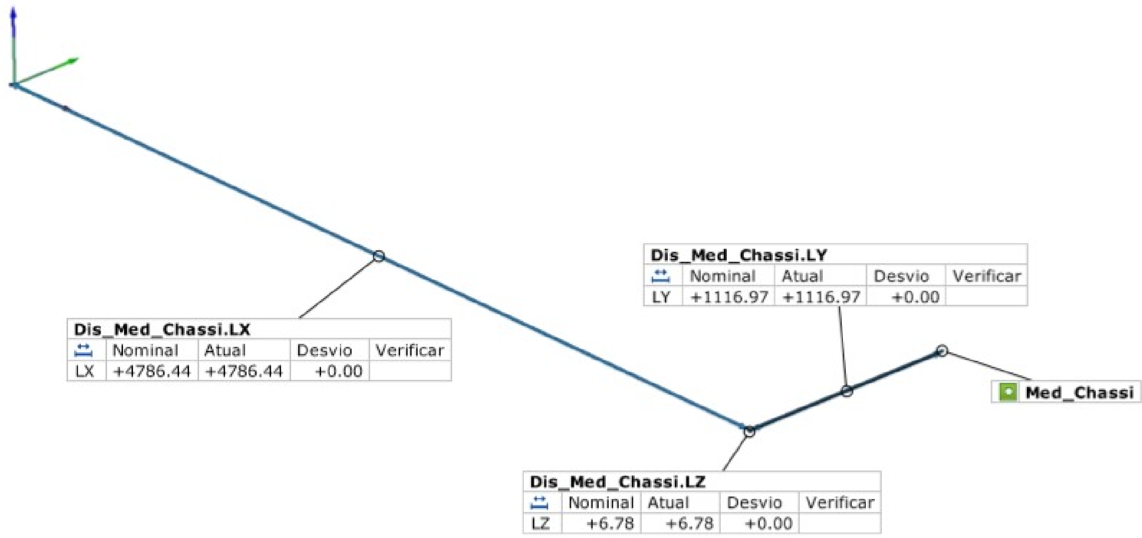
Source: The author (2021)

Figure 62 – X, Y, and Z coordinates of the instrumentation point on the vehicle's suspension bracket



Source: The author (2021)

Figure 63 – X, Y, and Z coordinates of the instrumentation point on the vehicle's cross member



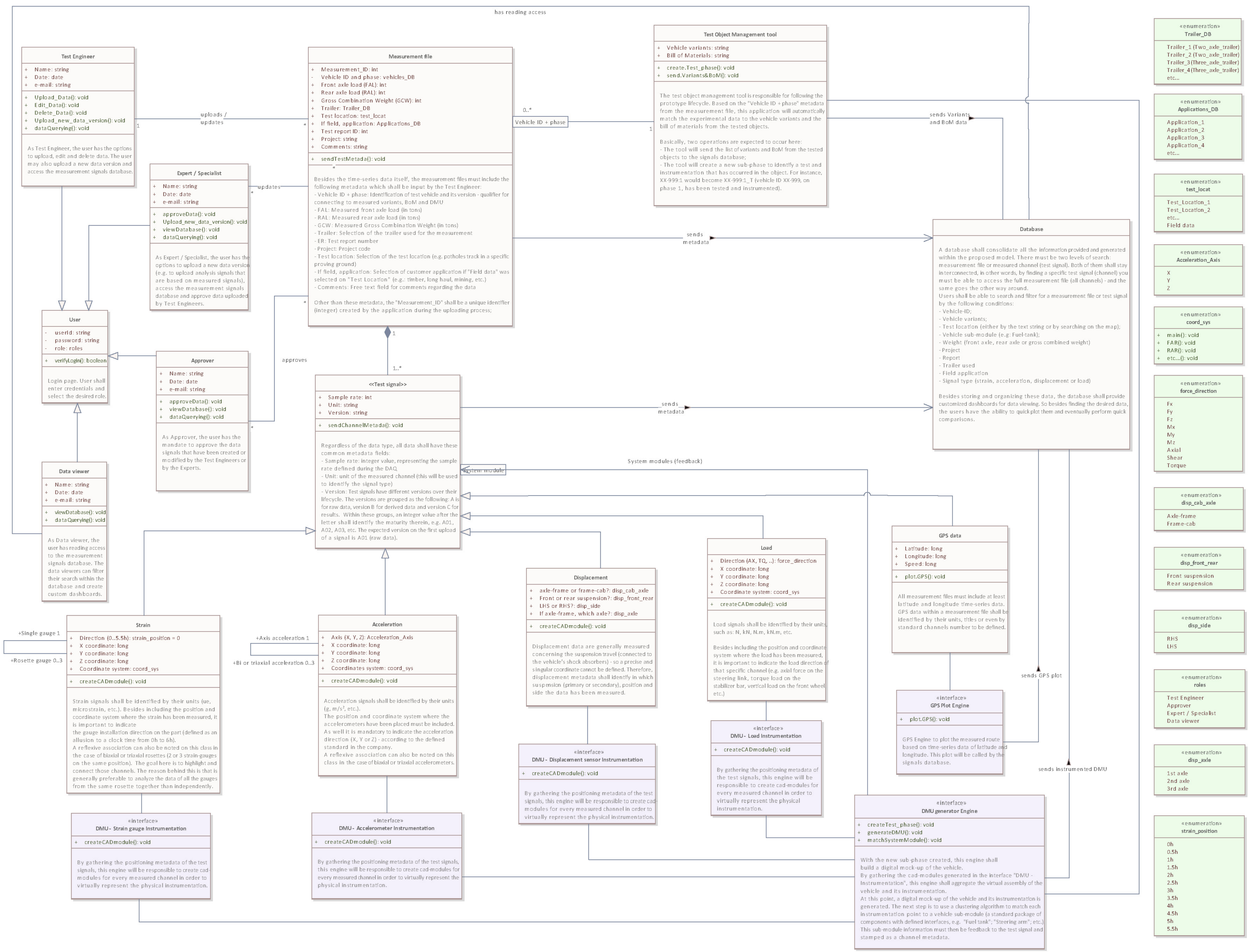
Source: The author (2021)

These coordinates values were verified with physical measurements and a good precision has been observed in this demonstration. Although the concept has been proved, further research and work shall assess the practical aspects of the implementation and integration of this type of tool/equipment.

5.3 LOGIC MODEL OVERVIEW IN UML

Figure 64 presents the representation of the proposed logical data model in UML notation. By providing a static structure of the application, in terms of concepts and relations between them, this can be seen as a summary for developers of the model described in detail in the previous sections. The works from Purchase et al. (2001) and Berardi, Calvanese and Giacomo (2005) are suggested as a reference for the interpretation of the UML notation.

Figure 64 – UML model for the management of durability measurement files

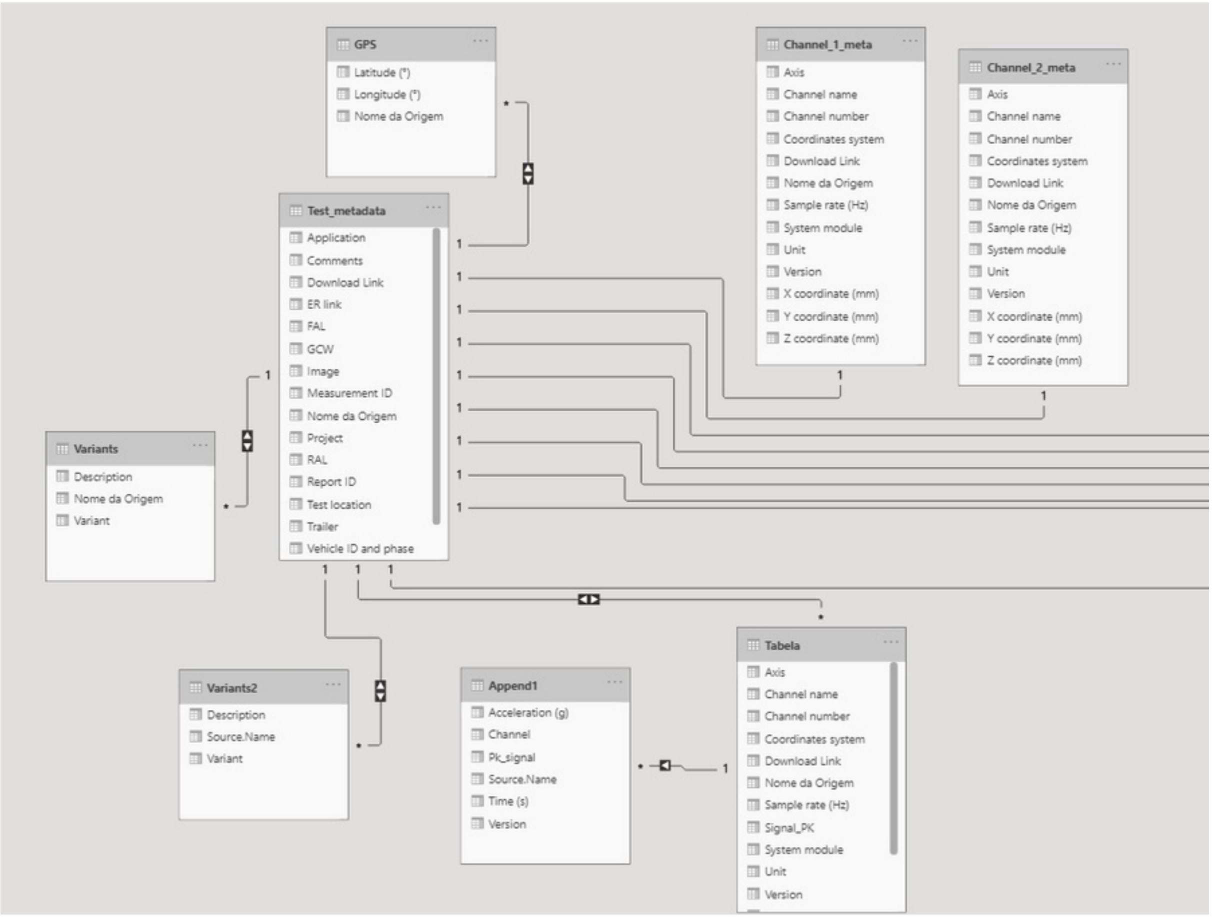


Source: The author (2021)

6 RESULTS: DEMONSTRATION

After a detailed walkthrough of the proposed model and concepts, this phase focused on the demonstration of how the database of measurement files would potentially work. In this regard, real durability measurement files have been merged with artificially created metadata (such as vehicle spec, test location, project, and others). As a proof-of-concept, a database and user interface have been modeled using Microsoft Power BI. Figure 65 below presents part of the data model developed in Power BI.

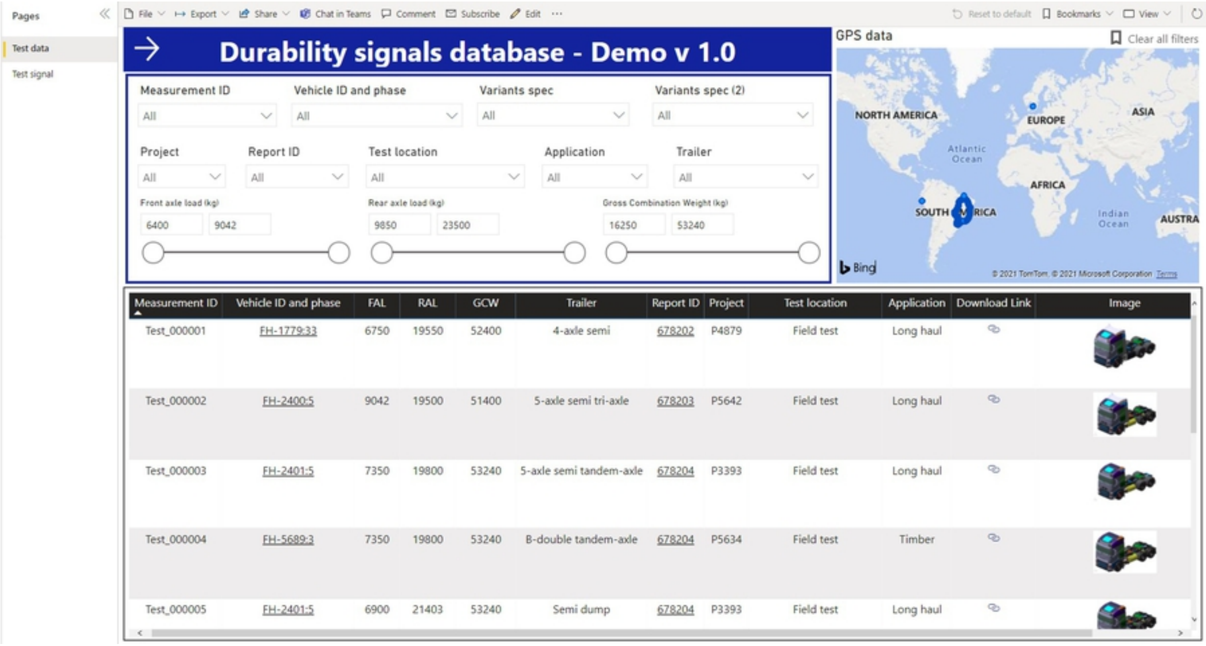
Figure 65 – Part of the data model developed in Microsoft Power BI



Source: The author (2021)

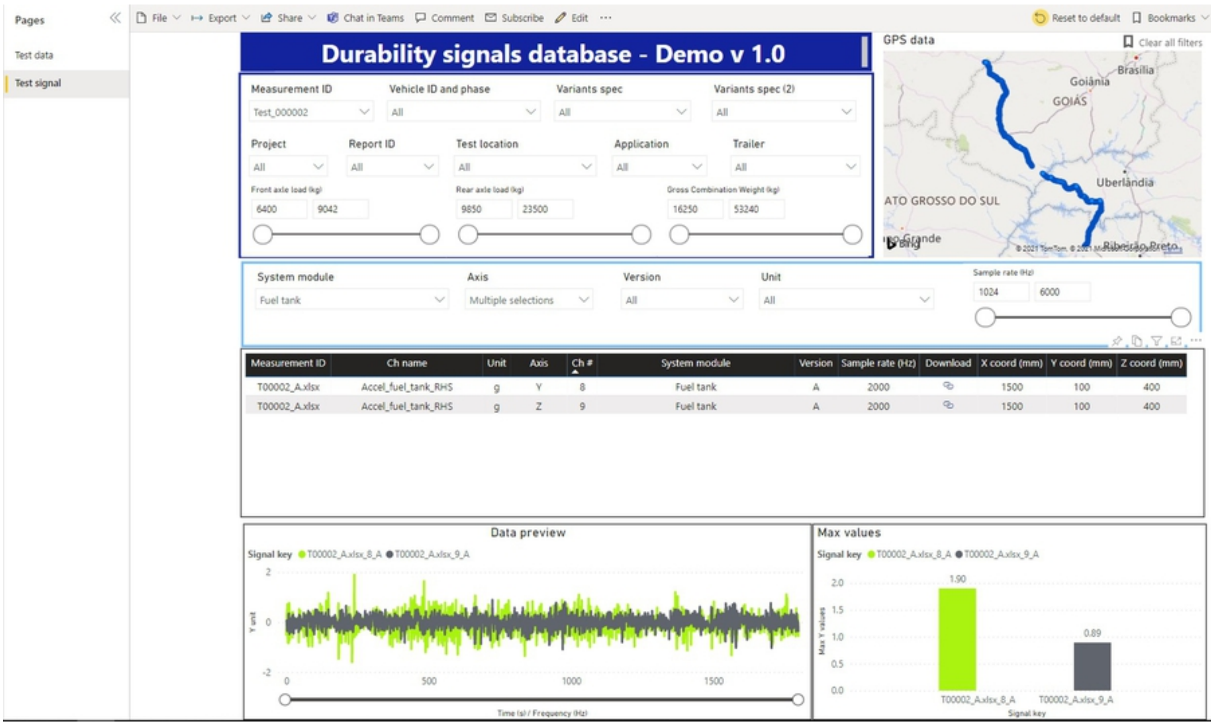
The user interface of the database has been split into two: test data (measurement files - Figure 66) and test signal (measurement channels - Figure 67).

Figure 66 – Measurement file search interface of proof-of-concept database



Source: The author (2021)

Figure 67 – Measured signal search interface of proof-of-concept database



Source: The author (2021)

So, in this proof-of-concept, the data viewer has to choose between searching for measurements (e.g. measurements performed on PROTO_23:01_T in the state of São Paulo) or a specific signal (e.g. acceleration measurements on the fuel tank of vehicles with air suspension). In both interfaces, the user is able to search and filter data by:

- GPS plot directly on the map.
- Measurement ID.
- Vehicle ID and phase.
- Vehicle specification (list of variants that describe the test object)
- Front axle load (FAL).
- Rear-axle load (RAL).
- Gross Combination Weight (GCW).
- Trailer.
- Report ID.
- Project code.
- Test location.
- Application.

Within the test signal interface, the data viewer has further options to narrow down his search, such as:

- Sample rate.
- System module.
- Unit.
- Axis.
- Signal version.

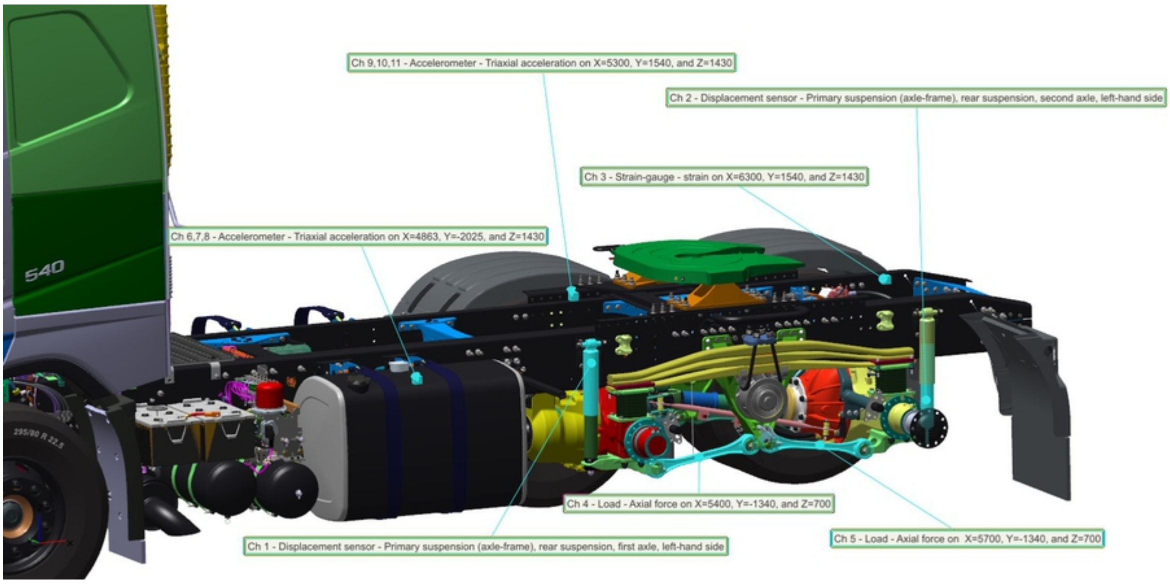
Figure 68 below illustrates an example of a comparison of the same signal on different versions: raw data (version A) and post-processed data (version B).

Figure 68 – Comparison of the same signal on different versions (A and B)

Source: The author (2021)

All these search criteria are connected to the defined requirements described in section 5.2 and the enhanced metadata detailed in section 5.2.3. In both interfaces, the users have access to download either the test data or test signal and also access to the 3-D view of the instrumented digital mock-up of the test object - as illustrated in Figure 69 below.

Figure 69 – Demonstration of an instrumented digital mock-up - available for data viewers



Source: The author (2021)

The model has been demonstrated for main stakeholders within the collaborating company on April 30th, 2021. A cross-functional team of nine employees has been selected for this phase:

- One group manager of the test department.
- One durability specialist from the company headquarters.
- Two durability test engineers.
- One project manager.
- One reliability specialist.
- Two product development engineers from different segments (products).
- One simulation engineer.

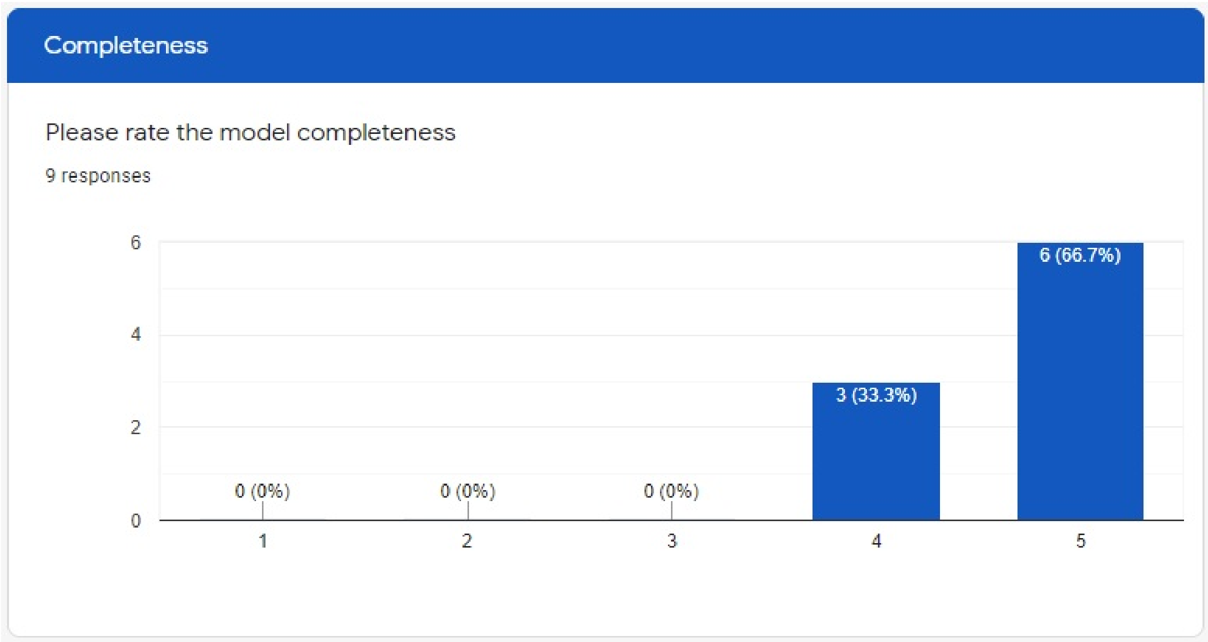
The demonstrated model has also been published as a report model to the company's cloud server. So besides watching the demonstration, the stakeholders had access to test the model by themselves.

7 RESULTS: EVALUATION

As detailed in section 4.4.3 and according to DSRM's framework, the same audience that took part in the demonstration has been asked to respond to a feedback survey. The nine attendees were asked to rate the model's "completeness", "fidelity with real-world phenomena", and "internal consistency" with a score from 1 (poor) to 5 (excellent). Regarding the model's "level of detail" and "robustness", they were asked to compare the proposed model with their current model/way of working for the management of durability measurement files.

Regarding the model's completeness, 66.7% of the respondents evaluated the proposed as excellent (score 5) and the remaining evaluated as good (score 4) - see Figure 70.

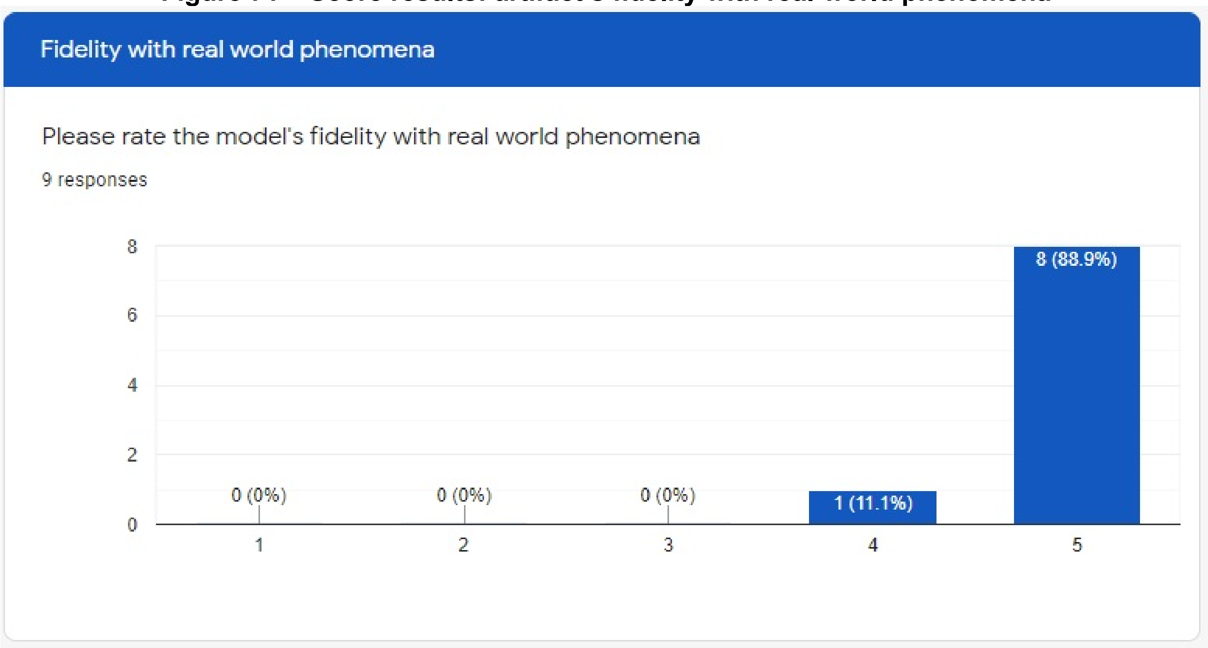
Figure 70 – Score results: artifact's completeness



Source: The author (2021)

The model's fidelity with real-world phenomena was evaluated with the maximum score by 88.9% of the stakeholders from the collaborating company - see Figure 71.

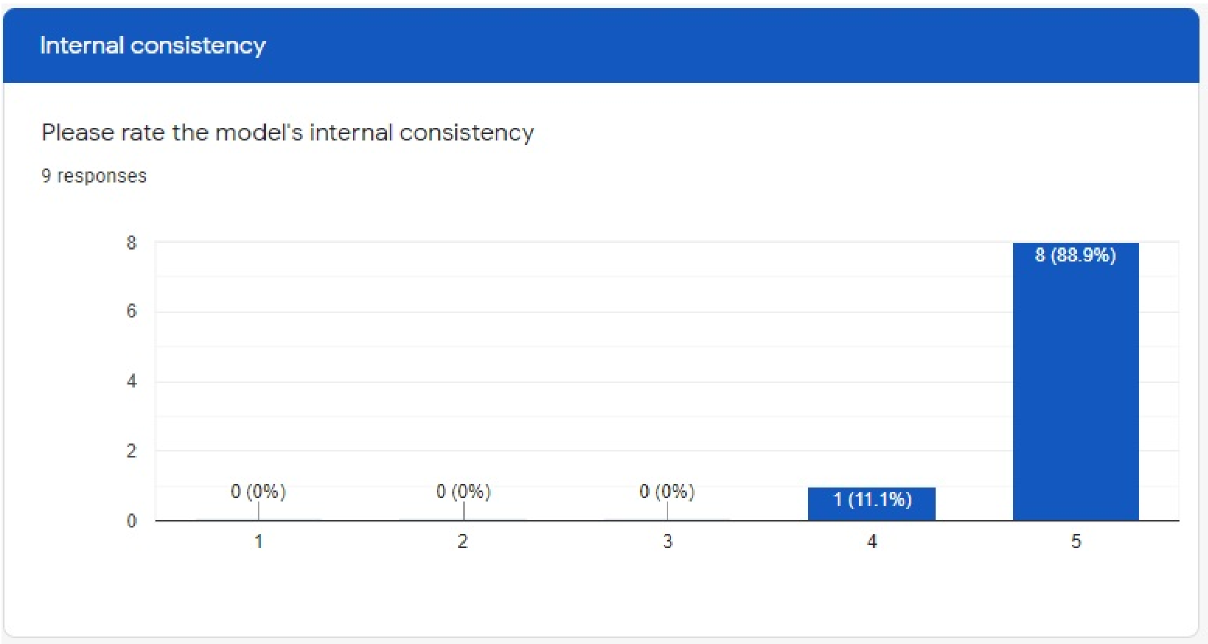
Figure 71 – Score results: artifact's fidelity with real-world phenomena



Source: The author (2021)

The model's internal consistency was deemed excellent for 88.9% of the respondents and good for one of the evaluators - see Figure 72.

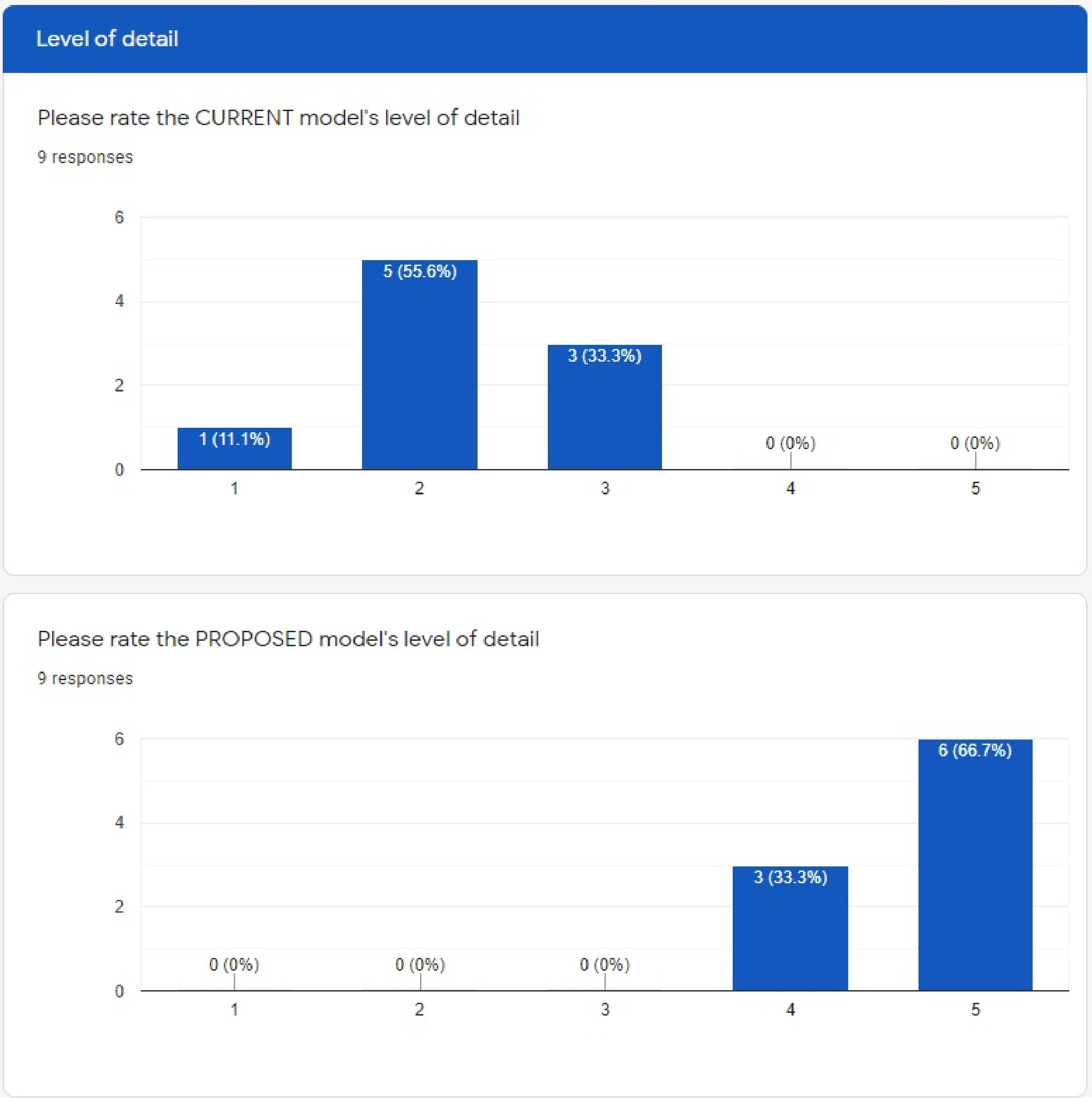
Figure 72 – Score results: artifact's internal consistency



Source: The author (2021)

While most (55.6%) of the stakeholders assessed the level of detail of the current model as fair (score 2), the level of detail of the proposed model was assessed as excellent by 66.7% of the respondents. The full results are shown in Figure 73.

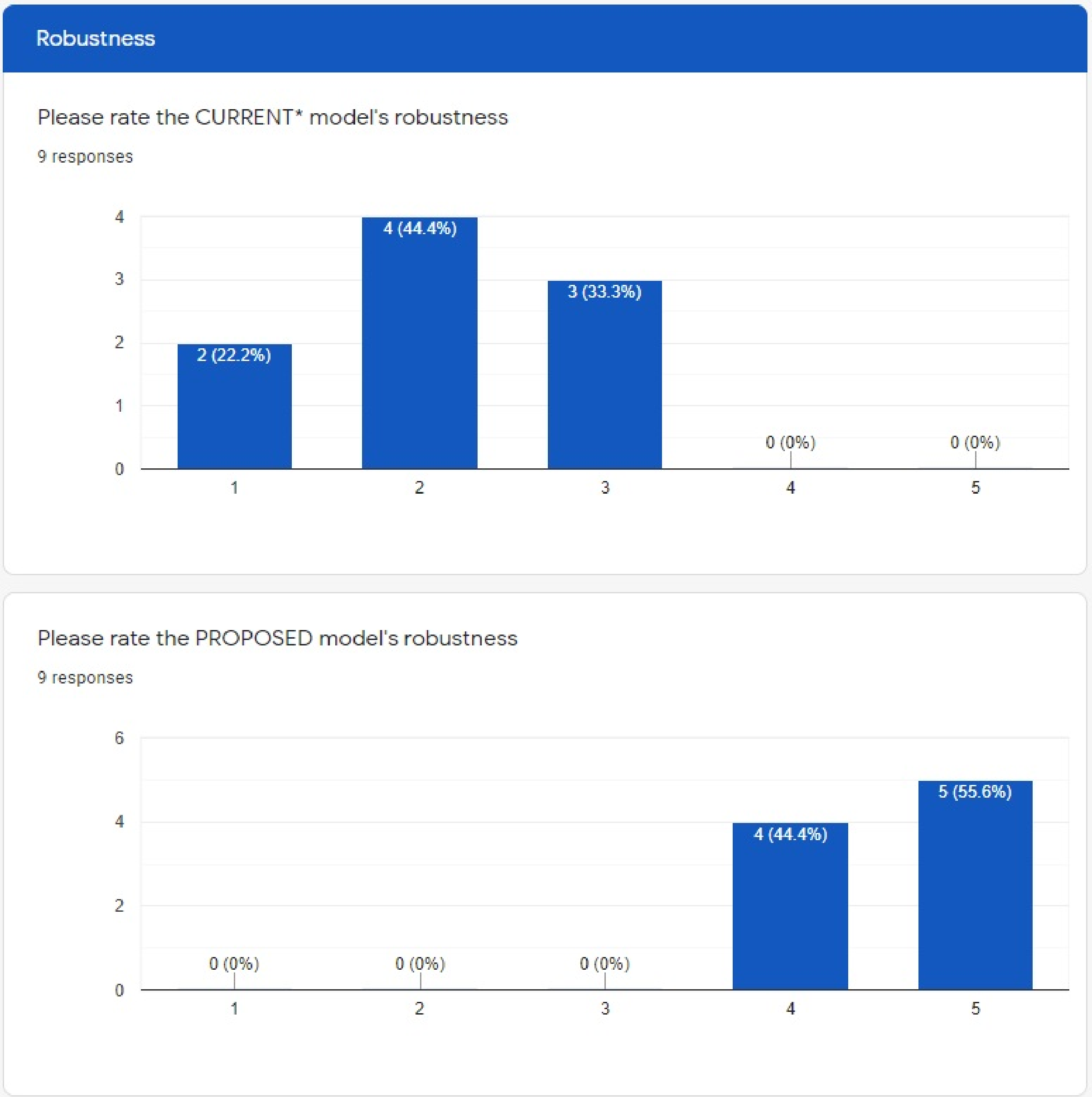
Figure 73 – Score results: comparative assessment between current and proposed models regarding their level of detail



Source: The author (2021)

From the perspective of models' robustness, the same trend was observed when comparing the assessments from the current and the proposed model, as shown in Figure 74.

Figure 74 – Score results: comparative assessment between current and proposed models regarding their robustness



Source: The author (2021)

Seven stakeholders added comments and/or suggestions in the free text field available in the survey. Overall, the proposed model received very good feedback. However, some pertinent suggestions and concerns were raised:

- Processing time for the DMU creation.
- Processing time for metadata inclusion.
- Database sizing to store measurement files and their different versions (A, B, and C).

These concerns are relevant and must be taken into account for the (potential) realization of this artifact in its environment (implementation), which is a suggestion for future work.

Table 7 below summarizes and consolidates the average results obtained in this evaluation phase.

Table 7 – Survey's average scores: 1 – Poor, 2 – Fair, 3 – Average, 4 – Good, 5 – Excellent

	Average score (1 to 5)	
Criterion	**Current model**	**Proposed model**
Completeness	-	4.67
Fidelity with real-world phenomena	-	4.89
Internal consistency	-	4.89
Level of detail	2.22	4.67
Robustness	2.11	4.56
Average (all evaluated criteria)	**2.17**	**4.73**

Source: The author (2021)

While the current model has been assessed with an average score of 2.17 regarding its level of detail and robustness, the proposed model has been assessed with an average score of 4.73 considering all criteria.

This section presented the results of the artifact's evaluation, the 5th activity in the DSRM. As discussed in the methodology, the last activity (communication) is expected to be completed after the defense of this dissertation work. Therefore, through an overview of the full research, the next section will present the conclusions and recommendations for future works.

8 CONCLUSION

Within the context of the product development process and collaborative engineering, the management of experimental data has been identified as a research gap. While many studies have tackled data and collaboration based on computer-aided technologies (CAx) environments, the management of the measured data has been observed as a scarce topic in the literature. Moreover, among all verifications and validations performed during development, further analyses in a multinational manufacturing company indicated a higher prevalence of durability tests. In this regard, a consultation in a collaborating company revealed some important perspectives:

- The measurement files are perceived as the foundation of many of the tests performed in a durability concern. If they were properly archived and identified, they would have the potential to eliminate (some of) the need for a new physical test for future demands/projects.
- The use of past measurements, when occurred, was only possible given previous experiences from the team members involved in the task. Simply put, this knowledge is embedded in the people that took part in the assignment, but not in the organization itself.

To address these shortcomings, the general objective of this research was to propose a logical model to manage the measurement files collected in durability tests within the PDP which would enable (and foster) their reuse across projects and throughout the product development organization.

Thereby, a qualitative applied research tackled this challenge within the framework of the DSRM. The main scope of this research included the development of a logical data model that could be demonstrated and evaluated. This was built upon previous researches and a participatory development methodology, which involved stakeholders and end-users during its development phase.

The resulting artifact from this research proposes three main foundations for the management of experimental data collected from durability tests:

1) Data versions: Measured data should have different versions (maturity levels) and these versions must be interconnected. Our model proposes the discretization of data into three different versions: Raw data (A), post-treated data (B), and results (C). By connecting the different stages of the data, we aim to enable the lifecycle tracking of the measured data.
2) Data approval: Given the concerns regarding signal quality and reliability, our model includes an approval process for the uploading of the test data into the database. This is not supposed to be different from the drawing checks that are performed in CAD data and, the even closer, approval of test reports.
3) Metadata enhancement: One key approach of the proposed model is the metadata enhancement within the measured data. Metadata provides context to raw data, therefore playing a vital role in data retrieving and reuse. For both test and channel metadata, the objective is the same: stamp and structure relevant information (that was normally concentrated in the test reports) to the measurement files. In other words, measurement files should be able to be interpreted and used without supporting documentation (such as the current test reports).

The test object specification (variants that describe the object configuration), and the test location (together with GPS data) were deemed as key aspects to foster data reuse within the PDP. Additional test metadata include information regarding the vehicle weight, the test report number, the trailer used, and other relevant data that support data searching and filtering.

Regarding channel metadata, the key proposed concept was to discretize each sensor position as X, Y and Z coordinates from a known Cartesian coordinate system and add this information as channel metadata. This information shall then be inputted into the product structure data, enabling two steps:

1) The creation (and documentation) of a digital mock-up of the test object with its instrumented sensors.
2) The feedback of a system module description (equivalent to the sensor positioning) to the channel metadata.

A proof-of-concept model has been demonstrated to experts and potential key users from the collaborating company. The concepts and approaches of the model have been presented and a controlled experiment has been carried out using a mix of real and artificial data in Power BI. Feedback and satisfaction surveys suggested that the proposed model is complete and faithful to real-world phenomena. They also suggested that the proposed artifact is better than the current model (or way of working) regarding its internal consistency, level of detail, and robustness. Considering all the evaluated criteria, the proposed model has been assessed with an average score of 4.73 on a scale from 1 (poor) to 5 (excellent). Nonetheless, some limitations of this work are inherent to its own nature. As suggested by Vaishnavi and Kuechler (2015, p. 13), models have to ignore things due to their abstract view of the world - and that is good that they do. The limitations of this work include the evaluation of the cloud storage needs to manage a large amount of measured data and digital mock-ups. The lack of some algorithms (methods) to operationalize some of the functions conceptualized by the model are also a limitation of this research. For instance, future research could work on a reverse geocoding methodology to automate the description of the measurement locations. Moreover, researchers are encouraged to develop a clustering algorithm to automatically categorize the so-called “system module” of a signal (based on the overlap of the Cartesian coordinates from the instrumentation and the digital mock-up of the test object). Finally, the realization of this artifact in its environment (implementation) is also suggested for future works.

REFERENCES

ABDUL-GHAFOUR, S. et al. Semantic interoperability of knowledge in feature-based CAD models. **CAD Computer Aided Design**, v. 56, p. 45 – 57, 2014.

AKERS, K. G.; DOTY, J. Disciplinary differences in faculty research data management practices and perspectives. **The International Journal of Digital Curation**, v. 8, n. 2, p. 5 – 26, 2013. ISSN 1746-8256.

ALAEI, N.; KURVINEN, E.; MIKKOLA, A. A methodology for product development in mobile machinery: case example of an excavator. **Machines**, v. 7, n. 4, 2019.

ALAVI, M.; TIWANA, A. Knowledge integration in virtual teams: The potential role of KMS. **Journal of the American Society for Information Science and Technology**, Wiley Online Library, v. 53, n. 12, p. 1029 – 1037, 2002. ISSN 1532-2882.

ALI, N. H.; SHUKUR, Z.; IDRIS, S. A Design of an Assessment System for UML Class Diagram. In: **International Conference on Computational Science and its Applications**. Kuala Lampur, Malaysia: [s.n.], 2007. p. 539 – 546.

ALLEN, T. J. Studies of the problem-solving process in engineering design. **IEEE Transactions on Engineering Management**, n. 2, p. 72 – 83, 1966. ISSN 0018-9391.

ANDRADE-VALBUENA, N. A.; MERIGO, J. M. Outlining new product development research through bibliometrics: analyzing journals, articles and researchers. **Journal of Strategy and Management**, v. 11, n. 3, p. 328 – 350, 2018. ISSN 1755425X.

ARDITO, L. et al. A bibliometric analysis of research on big data analytics for business and management. **Management Decision**, v. 57, n. 8, p. 1993 – 2009, 2019. ISSN 0025-1747.

AUWERAER, H. V. der; LEURIDAN, J. A new testing paradigm for today’s product development process - Part 1. **Sound and Vibration**, v. 39, n. 9, 2005a. ISSN 0038-1810.

AUWERAER, H. V. der; LEURIDAN, J. A new testing paradigm for today's product development process - Part 2. **Sound and Vibration**, v. 39, n. 11, 2005b. ISSN 0038-1810.

BARONE, L.; WILLIAMS, J.; MICKLOS, D. Unmet needs for analyzing biological big data: A survey of 704 NSF principal investigators. **PLoS computational biology**, v. 13, n. 10, p. e1005755, 2017. ISSN 1553-734X.

BARTELS, J.; ZIMMERMANN, J. Scheduling tests in automotive R&D projects. **European Journal of Operational Research**, v. 193, n. 3, p. 805 – 819, 2009. ISSN 0377-2217.

BENNEBACH, M.; CAWTE, R. Using a common Test and Simulation environment to optimize the Durability Process and verify the results. 2007. Available at: https://www:ncode:com/ images/DesignLife/Downloads/Whitepaper_nCode_CommonTestEnviroinment-CawteBennebach:pdf. Accessed on: 10/01/2021.

BERARDI, D.; CALVANESE, D.; GIACOMO, G. D. Reasoning on UML class diagrams. **Artificial Intelligence**, v. 168, n. 1, p. 70 – 118, 2005. ISSN 0004-3702. Available at: https://www:sciencedirect:com/science/article/pii/S0004370205000792.

BLONDET, G. et al. Towards a knowledge based framework for numerical design of experiment optimization and management. **Computer-Aided Design and Applications**, v. 13, n. 6, p. 872 – 884, 2016.

BLONDET, G. et al. Simulation data management for adaptive design of experiments: A literature review. **Mechanics & Industry**, v. 16, n. 6, 2015. ISSN 2257-7777.

BOBER, G. D. Method for reducing validation test sample size through sample manipulation, **SAE Technical Paper**, v. 1, n. 2005-01-1059, p. 1 – 8, 2005.

BOLLEN, K. A. Multiple indicators: Internal consistency or no necessary relationship? **Quality and Quantity**, Springer, v. 18, n. 4, p. 377 – 385, 1984. ISSN 1573-7845.

BORGHI, J. et al. Support your data: A research data management guide for researchers. **Research Ideas and Outcomes**, Pensoft Publishers, v. 4, 2018. ISSN 2367-7163.

BORSATO, M. et al. An ontology building approach for knowledge sharing in product lifecycle management. **International Journal of Business and Systems Research**, v. 4, n. 3, p. 278 – 292, 2010.

BORSATO, M.; PERUZZINI, M. Collaborative engineering. In: **Concurrent engineering in the 21st century**. Cham: Springer, 2015. p. 165 – 196.

BOSCH-MAUCHAND, M. et al. Knowledge-based assessment of manufacturing process performance: integration of product lifecycle management and value-chain simulation approaches. **International Journal of Computer Integrated Manufacturing**, v. 26, n. 5, p. 453 – 473, 2013. ISSN 0951-192X.

BRACKETT, M.; EARLEY, P. S. **The DAMA guide to the data management body of knowledge**. Bradley Beach: Technics Publications, 2009. ISBN 978-0-9771400-8-4.

BRADEN, D. R.; HARVEY, D. M. A prognostic and data fusion based approach to validating automotive electronics. **SAE Technical Paper**, v. 1, n. 2014-01-0724, p. 1 – 7, 2014.

BROTHERTON, B. **Researching hospitality and tourism: A student guide**, 2. Ed Croydon: Sage, 2008.

BRUUN, H. et al. PLM system support for modular product development. **Computers in Industry**, v. 67, p. 97 – 111, 2015.

BUI, Y. N. **How to write a master's thesis**. [S.l.]: Sage Publications, 2013. ISBN 1483321320.

CAMP, W. Formulating and evaluating theoretical frameworks for career and technical education research. **Journal of Vocational Education Research**, Association for Career and Technical Education Research, v. 26, n. 1, p. 4 – 25, 2001. ISSN 0739-3369.

CARDINAL, L. B.; ALESSANDRI, T. M.; TURNER, S. F. Knowledge codifiability, resources, and science-based innovation. **Journal of Knowledge Management**, MCB UP Ltd, v. 5, n. 2, p. 195 – 204, 2001.

CHANDRA, L.; SEIDEL, S.; GREGOR, S. Prescriptive knowledge in IS research: Conceptualizing design principles in terms of materiality, action, and boundary conditions. In: **48th Hawaii International Conference on System Sciences**. Hawaii: IEEE, 2015. p. 4039 – 4048. ISBN 147997367X.

CHANDRASEGARAN, S. K. et al. The evolution, challenges, and future of knowledge representation in product design systems. **Computer-Aided Design**, v. 45, n. 2, p. 204 – 228, 2013. ISSN 0010-4485.

CHELST, K. et al. Rightsizing and management of prototype vehicle testing at Ford Motor Company. **Interfaces**, v. 31, n. 1, p. 91 – 107, 2001. ISSN 0092-2102.

COLLINS, F. S.; TABAK, L. A. Policy: NIH plans to enhance reproducibility. **Nature News**, v. 505, n. 7485, p. 612, 2014.

COLLIS, J.; HUSSEY, R. **Business research**: A practical guide for undergraduate and postgraduate students. [S.l.]: Macmillan International Higher Education, 2013. ISBN 1137037482.

CRAGIN, M. H. et al. Data sharing, small science and institutional repositories. **Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences**, v. 368, n. 1926, p. 4023 – 4038, 2010. ISSN 1364-503X.

CUI, Y.; KARA, S.; CHAN, K. C. Manufacturing big data ecosystem: a systematic literature review. **Robotics and Computer-Integrated Manufacturing**, v. 62, 2020. ISSN 0736-5845. Available at: http://www.sciencedirect.com/science/article/pii/S0736584519300559.

CUMMINS, F. A. **Building the Agile Enterprise**: with SOA, BPM and MBM. Burlington: Elsevier, 2010. ISBN 0080560083.

DATRON TECHNOLOGY LTD. **6-component Wheel Force Transducer**. 2016. Available at: https://www.datrontechnology.co.uk/wheel-force-transducer/. Accessed on: 28/03/2021.

DEMOLY, F. et al. Product relationships management enabler for concurrent engineering and product lifecycle management. **Computers in Industry**, v. 64, n. 7, p. 833 – 848, 2013. ISSN 01663615.

DEMOLY, F. et al. Integrated product relationships management: A model to enable concurrent product design and assembly sequence planning. **Journal of Engineering Design**, v. 23, n. 7, p. 544 – 561, 2012. ISSN 09544828.

DO, N. Identifying experts for engineering changes using product data analytics. **Computers in Industry**, v. 95, p. 81 – 92, 2018. ISSN 01663615.

DO, N.; CHAE, G. A Product Data Management architecture for integrating hardware and software development. **Computers in Industry**, v. 62, n. 8-9, p. 854 – 863, 2011. ISSN 0166-3615.

DUDA, J. Modelling and implementation of product development strategy. **Concurrent Engineering Research and Applications**, v. 26, n. 2, p. 187 – 197, 2018.

EBRAHIMI, M. **Concurrent engineering approaches within product development processes for managing production start-up phase**. 2011. 99 p. Dissertation (Production System: Production Development and Management) — KTH Royal Institute of Technology. Available at: https://www:diva-portal:org/smash/get/diva2: 525648/FULLTEXT01:pdf. Accessed on: 10/06/2020.

EIGNER, M.; HANDSCHUH, S.; GERHARDT, F. Concept to enrichen lightweight, neutral data formats with CAD-based feature technology. **Computer-Aided Design and Applications**, v. 7, n. 1, p. 89 – 99, 2010.

ENSSLIN, L. et al. ProKnow-C, knowledge development process-constructivist. **Processo técnico com patente de registro pendente junto ao INPI. Brasil**, v. 10, n. 4, 2010.

ERNST, H. Success factors of new product development: a review of the empirical literature. **International journal of management reviews**, Wiley Online Library, v. 4, n. 1, p. 1 – 40, 2002. ISSN 1460-8545.

FEDERER, L. M. et al. Biomedical data sharing and reuse: Attitudes and practices of clinical and scientific research staff. **PloS one**, v. 10, n. 6, p. e0129506, 2015. ISSN 1932-6203.

FERNANDES, E. C. et al. Machine learning and process mining applied to process optimization: bibliometric and systemic analysis. **Procedia Manufacturing**, Elsevier, v. 38, p. 84 – 91, 2019. ISSN 2351-9789.

FERNANDEZ-REYES, F. C.; HERMOSILLO-VALADEZ, J.; MONTES-Y-GÓMEZ, M. A prospect-guided global query expansion strategy using word embeddings. **Information Processing & Management**, Elsevier, v. 54, n. 1, p. 1 – 13, 2018. ISSN 0306-4573.

FISHER, T. **The data asset**: How smart companies govern their data for business success. Hoboken: John Wiley & Sons, 2009. v. 24. ISBN 0470508027.

GAMMACK, J.; HOBBS, V.; PIGOTT, D. **The book of informatics**. Melbourne: Thomson Course Technology, 2007.

GEERTS, G. L. A design science research methodology and its application to accounting information systems research. **International Journal of Accounting Information Systems**, Elsevier, v. 12, n. 2, p. 142 – 151, 2011. ISSN 1467-0895.

GIL, A. C. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo: Atlas, 2010.

GOM. **TRITOP**. 2020. Available at: https://www.gom.com/en/products/3d-scanning/tritop. Accessed on: 03/04/2021.

GRILLENBERGER, A.; ROMEIKE, R. Key concepts of data management: an empirical approach. In: **Proceedings of the 17th Koli Calling International Conference on Computing Education Research**. [S.l.: s.n.], 2017. p. 30 – 39.

GUJARATHI, G.; MA, Y. Parametric CAD/CAE integration using a common data model. **Journal of Manufacturing Systems**, v. 30, n. 3, p. 118 – 132, 2011.

HAMBALI, A. et al. The important role of concurrent engineering in product development process. **Pertanika Journal of Sciences and Technology,** v. 17, n. 1, p. 9 – 20, 2009.

HAUG, E. J. **Concurrent engineering**: Tools and technologies for mechanical system design. Iowa: Springer Science & Business Media, 2012. v. 108. ISBN 3642781195.

HBM PRENSCIA INC. **GlyphWorks - signal processing and durability analysis**. 2020. Available at:

https://www.ncode.com/images/documents/GlyphWorks_2020.pdf. Accessed on: 05/11/2020.

HEVNER, A. R. et al. Design science in information systems research. **MIS quarterly**, p. 75 – 105, 2004. ISSN 0276-7783.

HIPPEL, E. V.; TYRE, M. J. How learning by doing is done: problem identification in novel process equipment. **Research policy**, Elsevier, v. 24, n. 1, p. 1 – 12, 1995. ISSN 0048-7333.

HOFFMANN, K. **An introduction to measurements using strain gages**. Alsbach: Hottinger Baldwin Messtechnik Darmstadt, 1989.

HU, Y.; ZHOU, X.; LI, C. Internet-based intelligent service-oriented system architecture for collaborative product development. **International Journal of Computer Integrated Manufacturing**, v. 23, n. 2, p. 113 – 125, 2010. ISSN 0951-192X.

JÄRVINEN, P. Action research is similar to design science. **Quality & Quantity**, Springer, v. 41, n. 1, p. 37 – 54, 2007. ISSN 0033-5177.

JI, S. et al. Intelligent test system for diesel engine based on digital collaborative technology. **Mechatronics and Manufacturing Technologies**. v. 20, n. 1, p. 39 – 45, 2017.

JOHANNESSON, P.; SPECKERT, M. **Guide to load analysis for durability in vehicle engineering**. [S.l.]: John Wiley & Sons, 2013.

JONES, S.; PRYOR, G.; WHYTE, A. **How to Develop Research Data Management Services-a guide for HEIs.** Edinburgh: Digital Curation Centre, 2013. Available at: http://www.dcc.ac.uk/resources/ how-guides.

KALANTARI, A. et al. A bibliometric approach to tracking big data research trends. **Journal of Big Data**, Springer, v. 4, n. 1, 2017. ISSN 2196-1115.

KRISTIAN, K. Tacit knowledge management: the role of artifacts. **Journal of Knowledge Management**, Emerald Group Publishing Limited, v. 6, n. 2, p. 112 – 123, 2002. ISSN 1367-3270.

LACERDA, R. T. de O. et al. Research perspectives on performance evaluation and project management. **Revista de Gestão e Secretariado**, Sindicato das Secretárias do Estado de São Paulo, v. 9, n. 2, 2018. ISSN 2178-9010.

LAROCCA, A. et al. Data management within new product development and collaborative engineering: a bibliometric and systemic analysis. **Journal of Information and Knowledge Management Systems**, v. ahead-of-print, n. ahead-of-print, 2021.

LAVALLE, S. et al. Big data, analytics and the path from insights to value. **MIT sloan management review**, v. 52, n. 2, p. 21 – 32, 2011.

LEE, J. Y. et al. Concurrent material flow analysis by P3R-driven modeling and simulation in PLM. **Computers in Industry**, v. 63, n. 5, p. 513 – 527, 2012. ISSN 01663615.

LEVANDOWSKI, C. E. et al. An integrated approach to technology platform and product platform development. **Concurrent Engineering-Research and Applications**, v. 21, n. 1, p. 65 – 83, mar 2013. ISSN 1063-293X.

LIU, Y. et al. A Semantic Feature Model in Concurrent Engineering. **IEEE Transactions on Automation Science and Engineering**, v. 7, n. 3, p. 659 – 665, 2010. ISSN 1558-3783.

MARCH, S. T.; SMITH, G. F. Design and natural science research on information technology. **Decision support systems**, North-Holland, v. 15, n. 4, p. 251 – 266, 1995. ISSN 0167-9236.

MARCHAND, D. A. Competing with intellectual capital. In: MARCHAND, D. A. (Ed.). **Knowing in Firms**: Understanding, managing and measuring knowledge. London: SAGE Publications Ltd, 1998. chap. 11, p. 253 – 268.

MARPLES, D. L. The decisions of engineering design. **IRE Transactions on Engineering Management**, n. 2, p. 55 – 71, 1961. ISSN 0096-2252.

MCSORLEY, G. **A novel approach to product lifecycle management and engineering based on product in-use information**. 2014. 142 p. Thesis (Mechanical Engineering) — École Polytechnique de Montréal.

MERMINOD, V.; ROWE, F. How does PLM technology support knowledge transfer and translation in new product development? Transparency and boundary spanners in an international context. **Information and Organization**, v. 22, n. 4, p. 295 – 322, 2012. ISSN 1471-7727.

MILBURN, T. J. The new product development paradigm led by simulation and testing. **SAE Technical Paper**, v. 1, n. 2004-01-2667, p. 1 – 12 , 2004.

MILLS, G. E.; GAY, L. R. **Educational research**: competencies for analysis and applications. Upper Saddle River: Pearson, 2019. ISBN 0134784227.

MINAVAND, H.; LORKOJOURI, Z. The linkage between strategic human resource management, innovation and firm performance. **Journal of Business and Management,** v. 11, n. 2, p. 85 – 90, 2013.

MISHRA, D. et al. Big Data and supply chain management: a review and bibliometric analysis. **Annals of Operations Research**, Springer, v. 270, n. 1-2, p. 313 – 336, 2018. ISSN 0254-5330.

MONTICOLO, D. et al. A meta-model for knowledge configuration management to support collaborative engineering. **Computers in Industry**, v. 66, p. 11 – 20, 2015.

MOURTZIS, D.; VLACHOU, E. Cloud-based cyber-physical systems and quality of services. **TQM Journal**, v. 28, n. 5, p. 704 – 733, 2016.

NALLUSAMY, S. et al. A review on valuable trends of product data management (PDM) occupied in new product development (NPD). In: **Applied mechanics and materials**. [S.l.]: Trans Tech Publ, 2015. v. 786, p. 262 – 268. ISBN 3038355496.

NGMTI. **Strategic Investment Plan for the Model-Based Enterprise**. Next Generation Manufacturing Technologies Initiative. [S.l.], 2005.

NONAKA, I.; TAKEUCHI, H. **The knowledge-creating company**: How Japanese companies create the dynamics of innovation. New York: Oxford university press, 1995. ISBN 0199879923.

NUNAMAKER, J. F.; CHEN, M.; PURDIN, T. D. M. Systems development in information systems research. **Journal of Management Information Systems**, Taylor & Francis, v. 7, n. 3, p. 89 – 106, 1990. ISSN 0742-1222.

PASCOAL, E. T.; SILVA, P. P. da. The competencies of Brazilian's experimental engineering in planning and realization validation's tests of the vehicles during the phase of project. **SAE Technical Paper**, v.1, n. 2010-36-0092, 2010.

PATON, N. W. Managing and sharing experimental data: standards, tools and pitfalls. **Biochemical Society Transactions**, Portland Press Ltd., v. 36, n. 1, p. 33 – 36, 2008.

PEFFERS, K. et al. A design science research methodology for information systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45 – 77, 2007. ISSN 0742-1222.

PENG, D. X.; HEIM, G. R.; MALLICK, D. N. Collaborative product development: the effect of project complexity on the use of information technology tools and new product development practices. **Production and Operations Management**, v. 23, n. 8, p. 1421 – 1438, 2014. ISSN 1059-1478.

PINFIELD, S.; COX, A. M.; SMITH, J. Research data management and libraries: Relationships, activities, drivers and influences. **PLoS One**, v. 9, n. 12, 2014. ISSN 1932-6203.

PITT, M.; MACVAUGH, J. Knowledge management for new product development. **Journal of Knowledge Management**, v. 12, n. 4, p. 101 – 116, 2008. ISSN 1367-3270.

PLESSIS, M. D. The role of knowledge management in innovation. **Journal of knowledge management**, v. 11, n. 4, p. 20 – 29, 2007. ISSN 1367-3270.

PRIES-HEJE, J.; BASKERVILLE, R.; VENABLE, J. R. Strategies for design science research evaluation. In: **European Conference on Information Systems 2008 proceedings**. [S.l.: s.n.], 2008.

PURCHASE, H. et al. UML class diagram syntax: An empirical study of comprehension. In: **Australian Symposium on Information Visualisation**. Sydney: Peter Eades and Tim Pattison, 2001. v. 9, p. 1 – 8.

RAHMANI, K.; THOMSON, V. Ontology based interface design and control methodology for collaborative product development. **Computer-Aided Design**, v. 44, n. 5, p. 432 – 444, 2012. ISSN 0010-4485.

RED, E. et al. Emerging design methods and tools in collaborative product development. **Journal of Computing and Information Science in Engineering (Transactions of the ASME)**, v. 13, n. 3, p. 31001 – 31014, 2013a. ISSN 1530-9827.

RED, E. et al. Considerations for multi-user decomposition of design spaces. **Computer-Aided Design and Applications**, v. 10, n. 5, p. 803 – 815, 2013b.

RIALTI, R. et al. Big data and dynamic capabilities: a bibliometric analysis and systematic literature review. **Management Decision**, v. 57, n. 8, p. 2052 – 2068, 2019. ISSN 0025-1747.

RICHTER, T. et al. Exploitation of potentials of additive manufacturing in ideation workshops. In: **The Fifth International Conference on Design Creativity (ICDC 2018)**. [S.l.: s.n.], 2018. p. 354 – 361. ISBN 1912254077.

ROZENFELD, H.; AMARAL, D. C. **Gestão de projetos em desenvolvimento de produtos**. São Paulo: Saraiva, 2006.

SAPUAN, S. M.; OSMAN, M. R.; NUKMAN, Y. State of the art of the concurrent engineering technique in the automotive industry. **Journal of Engineering Design**, Taylor & Francis, v. 17, n. 2, p. 143 – 157, 2006. ISSN 0954-4828.

SCARBROUGH, H. Knowledge management, HRM and the innovation process. **International Journal of Manpower**, v. 24, n. 5, p. 501 – 516, 2003.

SHAH, N. U.; NAEEM, S. B.; BHATTI, R. Emerging trends of data management and data analytical practices in academic libraries: a theoretical lens. **Journal of Information and Computational Science**, v. 10, n. 4, p. 545 – 556, 2020.

SHERMAN, R. **Business Intelligence Guidebook**: From Data Integration to Analytics. Waltham: Elsevier, 2015.

SIMON, H. A. **The sciences of the artificial**. Cambridge: MIT press, 1969. ISBN 0262537532.

SIVAKUMAR, S.; DHANALAKSHMI, V. An approach towards the integration of CAD/CAM/CAI through STEP file using feature extraction for cylindrical parts. **International Journal of Computer Integrated Manufacturing**, v. 26, n. 6, p. 561 – 570, 2013.

SONNENBERG, C.; BROCKE, J. V. Evaluation Patterns for Design Science Research Artefacts. In: **European Design Science Symposium (EDSS)**. Dublin: [s.n.], 2011. v. 286, p. 77 – 83.

STARK, J. **Product Lifecycle Management (Volume 1)**: 21st Century Paradigm for Product Realisation. Geneva: Springer, 2015. ISBN 978-3-319-17440-2.

STARK, J. **Products2019: A project to map and blueprint the flow and management of products across the product lifecycle**: Ideation; Definition; Realisation; Support of Use; Retirement and Recycling. [S.l.: s.n.], 2020. 359 p. ISBN 979-8-664-16844-0.

STEYER, G.; VOIGHT, M.; HERING, D. The future of NVH testing - an end user's perspective. **SAE Technical Paper**, v. 1, n. 2005-01-2270, 2005.

SU, H. A road load data processing technique for durability optimization of automotive products. **SAE International Journal of Passenger Cars-Mechanical Systems**, v. 7, n. 2014-01-0884, p. 244 – 259, 2014. ISSN 1946-3995.

TAHERA, K.; EARL, C. Testing and PLM: Connecting Process and Product Models in Product Development. In: **Product Lifecycle Management-Terminology and Applications**. [S.l.]: IntechOpen, 2018.

TAN, C. L.; VONDEREMBSE, M. A. Mediating effects of computer-aided design usage: From concurrent engineering to product development performance. **Journal of Operations Management**, v. 24, n. 5, p. 494 – 510, 2006. ISSN 0272-6963.

TENOPIR, C. et al. Data management education from the perspective of science educators. **International Journal of Digital Curation**, v. 11, n. 1, p. 232 – 251, 2016. ISSN 1746-8256.

THE UNIVERSITY OF SHEFFIELD. **What is research data management?** 2020. Available at: https://www:sheffield:ac:uk/library/rdm/whatisrdm. Accessed on: 06/11/2020.

THOMKE, S. Learning by experimentation: Prototyping and testing. In: THOMKE, S. (Ed.). **Handbook of New Product Development Management**. [S.l.]: Routledge, 2008. chap. 15, p. 401 – 420.

TOCHE, B. et al. A framework to support collaboration during prototyping and testing. **International Journal of Product Lifecycle Management**, v. 10, n. 4, p. 348 – 374, 2017.

TORRES, V. H. et al. Integration of design tools and knowledge capture into a CAD system: a case study. **Concurrent Engineering-Research and Applications**, v. 18, n. 4, p. 311 – 324, 2010. ISSN 1063-293X.

TORRES, V. H. et al. Approach to integrate product conceptual design information into a computer-aided design system. **Concurrent Engineering: Research and Applications**, v. 21, n. 1, p. 27 – 38, mar 2013. ISSN 1063-293X, 1063-293X.

TSOUKAS, H. The firm as a distributed knowledge system: A constructionist approach. **Strategic Management Journal**, v. 17, n. S2, p. 11 – 25, 1996. ISSN 0143-2095.

TUYL, S. V.; WHITMIRE, A. Investigation of Non-Academic Data Management Practices to Inform Academic Research Data Management. **Research Ideas and Outcomes**, v. 4, 2018. ISSN 2367-7163.

UNIVERSITY OF SOUTHERN CALIFORNIA. **Organizing Your Social Sciences Research Paper**. 2020. Available at: https://libguides:usc:edu/writingguide/theoreticalframework. Accessed on: 27/11/2020.

VAISHNAVI, V.; KUECHLER, W. **Design research in information systems**. 2004. Available at: http://www:desrist:org/design-research-in-information-systems/.

VAISHNAVI, V. K.; KUECHLER, W. **Design science research methods and patterns**: innovating information and communication technology. Boca Raton: Auerbach Publications, 2015.

VALILAI, O.; HOUSHMAND, M. INFELT STEP: An integrated and interoperable platform for collaborative CAD/CAPP/CAM/CNC machining systems based on STEP standard. **International Journal of Computer Integrated Manufacturing**, v. 23, n. 12, p. 1095 – 1117, 2010.

VEZZETTI, E.; MOOS, S.; KRETLI, S. A product lifecycle management methodology for supporting knowledge reuse in the consumer packaged goods domain. **Computer-Aided Design**, v. 43, n. 12, p. 1902 – 1911, 2011. ISSN 0010-4485, 0010-4485.

VIEIRA, E. L. et al. Application of the proknow-c methodology in the search of literature on performance indicators for energy management in manufacturing and industry 4.0. **Procedia Manufacturing**, v. 39, p. 1259 – 1269, 2019. ISSN 2351-9789.

VORNHOLT, S.; GEIST, I.; LI, Y. Categorisation of data management solutions for heterogeneous data in collaborative virtual engineering. In: **Proceedings of the First International Workshop on Digital Engineering**, 2010. p. 9 – 16.

WALLS, J. G.; WIDMEYER, G. R.; SAWY, O. A. E. Building an information system design theory for vigilant EIS. **Information systems research**, INFORMS, v. 3, n. 1, p. 36 – 59, 1992. ISSN 1047-7047.

YU, S.; YANG, D. The role of big data analysis in new product development. In: **2016 International Conference on Network and Information Systems for Computers (ICNISC)**, 2016. p. 279 – 283. ISBN 1467388386.

ZEHTABAN, L.; ELAZHARY, O.; ROLLER, D. A framework for similarity recognition of CAD models. **Journal of Computational Design and Engineering**, v. 3, n. 3, p. 274 – 285, 2016.

ZHAN, Y. et al. Unlocking the power of big data in new product development. **Annals of Operations Research**, v. 270, n. 1-2, p. 577 – 595, 2018. ISSN 0254-5330.

ZHAO, Z. V.; LOPEZ, C. E.; TUCKER, C. S. Evaluating the impact of idea dissemination methods on information loss. **Journal of Computing and Information Science In Engineering**, v. 19, n. 3, 2019. ISSN 1530-9827.