VTT Technical Research Centre of Finland

DigiBuzz-VTT – Towards digital twin's concrete commercial exploitation

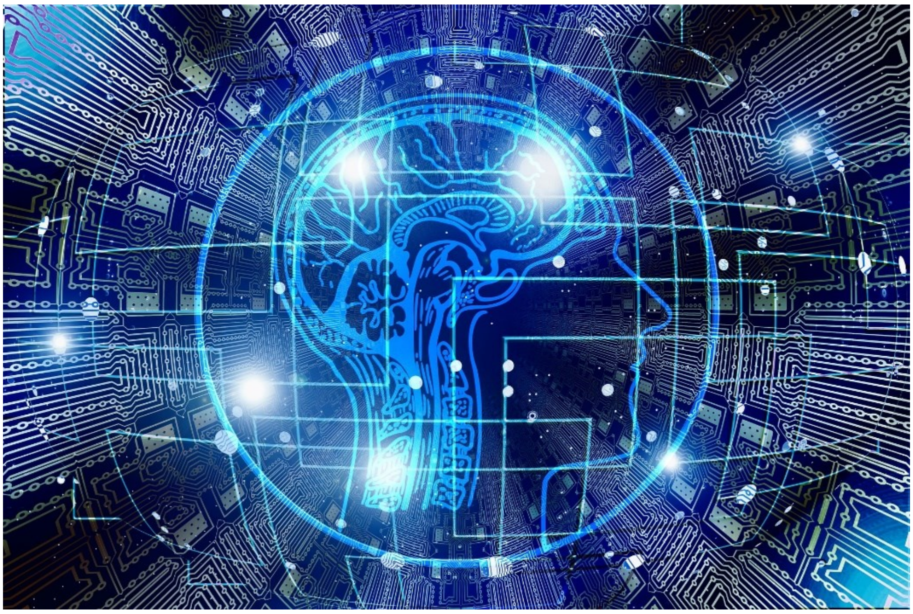
Kortelainen, Juha; Lämsä, Ville S.; Nieminen, Vesa; Rantala, Tuija; Virtanen, Juha

Published: 31/01/2022

Document Version
Publisher's final version

RESEARCH REPORT

VTT-R-00118-22



DigiBuzz-VTT
Towards digital twin’s concrete commercial exploitation

Authors: Juha Kortelainen
Ville Lämsä
Vesa Nieminen
Tuija Rantala
Juha Virtanen

Confidentiality: VTT Public

Version: 31.1.2021

VTT

beyond the obvious

Report's title	
DigiBuzz-VTT – Towards digital twin's concrete commercial exploitation	
Customer, contact person, address Business Finland, Kari Koskela / Kirjaamo, P.O. Box 69, FI-00101 Helsinki, Finland	**Order reference** Dnro 4437/31/2019
Project name Kohti digitaalisen kaksosen konkreettista kaupallista hyödyntämistä	**Project number/Short name** 125078 / BF_DigiBuzz
Author(s) Juha Kortelainen, Ville Lämsä, Vesa Nieminen, Tuija Rantala, Juha Virtanen	**Pages** 21/–
Keywords Added business value, business model, data management, digital twin, feature recognition, hybrid modelling, machine learning, modelling, lifecycle, simulation	**Report identification code** VTT-R-00118-22

Summary

The DigiBuzz-VTT project, a part of the DigiBuzz common effort, focused on the applications of digital twins in manufacturing industry ecosystems. The DigiBuzz-VTT project had two main focuses, 1) functional digital twins, or simulation-based digital twins, of machines and machine systems and their applications, and 2) the life cycle management of digital twins (the digital part of the twin), emphasising data modelling and data management. These themes were studied from the technical and from the business point of views. The detailed research topics were:

- Business opportunities and added value of digital twins for manufacturing industry
- Data-based digital twins, use of machine learning for feature recognition
- The status of standardisation for the lifecycle data management of digital twins, means for preserving model data
- Hybrid modelling with digital twins, combination of experimental and simulation data
- The optimisation of the measurement points location, method development
- The use of Kalman filters in estimating simulation data correlation with measured data
- The status of Industrial Internet of Things (IIoT) for digital twins

This report summarises the implementation of the DigiBuzz-VTT project and lists the main deliverables of the project. The project produced several scientific articles and research reports, which report the research results in detail.

Confidentiality	VTT Public

Espoo 8.2.2022

Written by	**Reviewed by**
Juha Kortelainen, Research Manager	Liisa Poussa, Manager, Operations Support

VTT's contact address

VTT Technical Research Centre of Finland Ltd, P.O. Box 1000, FI-02044 VTT, Finland

Distribution (customer and VTT)

Business Finland, Registry: 1 copy

VTT Registry: 1 copy

Approval

VTT TECHNICAL RESEARCH CENTRE OF FINLAND LTD

Date:	
Signature:	
Name:	Tarja Laitinen
Title:	Vice President, Knowledge-driven Design

Preface

The interest in digital twins and their applications is active in industry and in the research community. The concept itself and the enabling technologies introduce new and interesting challenges for basic and applied research. At the same time, the end-user companies are increasingly evaluating opportunities to apply digital twins in their products and services. The global interest in the topic has shown some features of a hype and the communication, especially for product and service marketing, has been confusing from time to time.

One of the main roles for research is to define and clarify concepts and to provide objective new information. This role has been fulfilled well in the DigiBuzz common effort and the DigiBuzz-VTT project. The concept itself has been discussed between the project parties, and the opportunities for new development and business have been analysed and identified. The dialog between the research parties and the industrial partners has been good, the feedback from the partners has been positive, and the project has been found to fulfil the expectations.

We would like to thank Business Finland for providing funding and support during the project, the LUT University for the great and active collaboration in research, and the consortium companies for good dialog and support. The provided financial framework has enabled the research community to share the latest research knowledge with the business and industry, and to gain valuable feedback and grounding of research visions to serve the society.

Espoo 8.2.2022

Authors

Contents

1. Introduction

The industrial interest in digital twins (DTs) has been growing steadily and there is increasing information and solutions available for their implementation. At the same time, the DT term has been used in communication and marketing for applications that do not necessarily fulfil the main features of the DT concept. This, together with a strong hype, has shown to be confusing to many companies and is causing additional efforts in communication also in the research community. The DT concept has no single widely accepted definition, but the concept has been defined by researchers and industry in numerous ways during the past decade. One could claim that based on earlier definitions a DT should have all the following features involved:

- a real system (or a target system) that can be physical or non-physical, e.g., a machine or an operational process, such as a manufacturing process,
- a virtual representation (DT) of the target system that contains those features that are relevant for the purpose of the DT, and
- online, real-time, and two directional data connection between the two, i.e., the target system and its DT.

The context and the purpose of the DT defines the concrete meaning for the represented features and the real-time communication. However, DT as its simplest format should include in some form all above-mentioned bulleted items. For example, by definition, a virtual representation of the target system solely, is not a DT representation.

1.1 DigiBuzz-VTT project

DigiBuzz was a Business Finland's common co-innovation research and development effort. It was financed by Business Finland and the participating organisations. DigiBuzz focused on the applications of DTs in manufacturing industry ecosystems. DigiBuzz contained two research projects:

- DigiBuzz-LUT, LUT University
- DigiBuzz-VTT, VTT Technical Research Centre of Finland Ltd

and six development projects of the following companies:

- Novatron Oy
- Mevea Oy
- Eurostep Oy
- Process Genius Oy
- Raute Oyj
- Wärtsilä Finland Oy

This report deals with the DigiBuzz-VTT research project (referred later as the project). The DigiBuzz-VTT project had two main research challenges in its scope:

1. Applications of functional DTs of machines and machines systems, and

2. The life cycle management of DTs.

The first research challenge focused on how to implement functional, simulation-based DT applications. The emphasis was on applications in which structural dynamics is in essential role. The second research challenge focused on how to manage the data for functional DTs throughout their life cycle, which can be even tens of years. One of the related questions is the development and evolvement of underlying technology that is necessary to keep the DTs in service.

In addition to technical challenges, business related questions were in essential focus. The main questions were "what are the benefits and the added value of DTs for the end-users, customers, and the service providers" and "what are the likely technical and financial investments in general level that are required and the expected benefits".

The project was organised in work packages and tasks as listed below (from the project plan). VTT did not have research activities in the tasks marked in grey colour.

- Work package 1, Service business requirements
    - Task 1.1 Current status and possible uses of the Digital Twin product / service (Digital Twin tuotteen / palvelun nykytila ja mahdolliset käyttökohteet)
    - Task 1.2 The desired effects and results of the Digital Twin (Digital Twinin tavoitellut vaikutukset ja tulokset)
    - Task 1.3 The formation of the Digital Twin ecosystem and the identification of the necessary properties to enable commercial exploitation (Digital Twinin ekosysteemin muodostuminen sekä tarvittavien ominaisuuksien selvittäminen kaupallisen hyödyntämisen mahdollistamiseksi)
- Work package 2, Technical means to increase the value of data
    - Task 2.1 Platforms for digital twins, artificial intelligence and IIoT (Alustat digitaalisille kaksosille, keinoälylle ja IIoT:lle)
    - Task 2.2 Data collection for digital twins (Digitaalisten kaksosten datan kerääminen)
    - Task 2.3 Lifecycle management of digital resources (Digitaalisten resurssien elinkaaren hallinta)
    - Task 2.4 Artificial intelligence and the digital twin (Keinoäly ja digitaalinen kaksonen)
    - Task 2.5 IIoT and digital twins (IIoT ja digitaaliset kaksoset)
- Work package 3, New business and business models based on visualisation and digital twins
    - Task 3.1 Development of business processes and business through visualisation and digital twinning (Toimintaprosessien ja liiketoiminnan kehittäminen visualisoinnin ja digitaalisella kaksosen avulla)
    - Task 3.2 Development of business processes (Toimintaprosessien kehittäminen)
    - Task 3.3 Life cycle model management of products and services (Tuotteiden ja palveluiden elinkaarimallinhallinta)
- Work package 4, Exploiting the information gathered in EU projects
    - Task 4.1 Creating links to European platforms and Digital Innovation Hubs (Linkkien luominen eurooppalaisiin alustoihin ja Digital Innovation Hubeihin)
    - Task 4.2 Systematic analysis of Horizon 2020 projects and data transfer from European communities to Finnish companies (Horisontti2020 projektien systemaattinen analyysi ja tiedonsiirto eurooppalaisista yhteisöistä suomalaisiin yrityksiin)
    - Task 4.3 Utilisation of DigiBuzz results (DigiBuzz-tulosten hyödyntäminen)
- Work package 5, Project management
    - Task 5.1 Project steering
    - Task 5.2 Dissemination

The DigiBuzz-VTT project started on 1 October 2019 and ended on 31 January 2022; the original project duration was extended by four months from the end of September 2021. The project implementation and results are discussed in more detail in the following sections.

2. Project implementation

2.1 Digital twin business perspectives

The business-related research in the DigiBuzz-VTT project was done in close co-operation with the LUT University.

2.1.1 Description

DT business perspectives were studied in two work packages:

- WP1 *Service business requirements* and
- WP 3 *New business and business models based on visualisation and digital twins*.

The goal was to study the current state of DT business in companies and to identify business opportunities, understanding the value creation and service development related to DTs. Selling DTs was seen challenging in companies and, therefore, the main topic studied from several perspectives.

2.1.2 Methods

A qualitative case study was employed as the research methodology in studying the business perspectives of DTs. The case study is suitable for situations and processes containing complex and multiple variables and, therefore, was suitable for this study as well (Yin, 2014). A case study can be used as an empirical study for examining a phenomenon, e.g., the DT concept, especially, with a volatile boundary of phenomenon and real-life contexts (Yin, 2014). In our research, six provider companies and eight utiliser companies were interviewed in order to study the concept of a complex, data-based solution, such as a DT, from the business and sales point of views. The research was done in close collaboration with the LUT University.

In the first round, 16 qualitative and semi-structured theme interviews from six different provider companies were collected during January–February 2020 (Table 1).

Table 1. Interviewees, first round.

Company	Main products and services	Number of interviewees
A	Product data sharing software and solutions for product lifecycle management collaboration	1
B	Solutions for digitalising earthmoving jobsites and robotising earthmoving machinery	4
C	Data visualisation and rationalising the operative hands-on work with different kinds of DT solutions	4
D	Tools and tooling systems for industrial metal cutting, stainless steels	4
E	Technologies and lifecycle solutions for the marine and energy markets	1
F	Software for DTs and real-time simulations	2

In the second round, the qualitative data were collected in March 2021 from nine semi-structured theme interviews held with 10 representatives from 8 different large, industrial companies from the DT utiliser perspectives (Table 2).

Table 2. Interviewees, second round.

Case company	Industry solutions	Number of interviewees
A	for conveying passengers	1
B	for the marine industry	1
C	for the refining industry	1
D	for automation and electric power technology	1
E	for the marine and energy industry	1
F	for production machinery	3
G	for lifting loads	1
H	for wood processing	1

The case companies all operate in B2B markets, and they can be both providers and utilisers of data-based solutions. They are not representatives of one specific business ecosystem but provide experiences from several different business environments and ecosystems. They were selected because they are actively developing and taking into use complex data-based solutions, such as DTs. In addition, they are all established, industrial companies, whose main offering is a big investment for their customers. The interviews were recorded, and comprehensive notes were taken during each interview by several interviewers. The duration of a typical interview was 1–1.5 hours, and each involved 1–4 interviewers. The main source of empirical material was semi-structured theme interviews because the study was partly explorative in nature. The area of industrial data-based solutions, including ecosystem development and complex data utilisation, is evolving, and the definition of concepts, such as complex data-based solutions, needed to be discussed with the interviewees. The interviews went beyond the sale of data-based components of solutions to cover a broad range of themes, such as the current utilisation of DTs, their advantages, and challenges, ecosystem development and customer value and understanding. The approach for analysing the data was content analysis. In several researcher meetings, the primary emerging themes were grouped and analysed.

2.1.3 Results

The main results of studying DT business perspectives were four scientific papers (three conference papers and one book chapter). In addition, one conference paper is submitted into a journal for review.

The main results indicate that companies are struggling with several challenges related to the DT business development. For example, the selling of complex DT solutions requires not only the solution understanding but also understanding of the customer's business. Therefore, salespersons need to focus on the value creation to people in different roles in the customer organisation as well as on the technological proficiency. DTs have different kind of impacts to the ecosystem – they can be only small improvements and process changes in the organisation, but also complex DTs which may change the market environment and require large ecosystem to succeed. However, the ecosystem development related to DT is quite challenging and not a customary manner to the companies. For example, data sharing, customer value creation, and shared goals were mentioned challenging by the interviewed companies. In the long run, a new kind of DTs and their ecosystems may create a significant turning point in the markets. The research group proposes that the DT's business perspectives need to be analysed early enough in order to create a suitable business model and an ecosystem for specific market need also in the future.

Publications:

- Rantala, T.; Valjakka, T.; Kokkonen, K.; Hannola, L.; Timperi, M. & Torvikoski, L. (2021). *Selling the Value of Complex Data-based Solution for Industrial Customers*. In L. M. Camarinha-Matos, X. Boucher, & H. Afsarmanesh (Eds.), Smart and Sustainable Collaborative Networks 4.0: 22nd IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2021, Saint-Étienne, France, November 22–24, 2021, Proceedings (pp. 332–342). Springer. IFIP Advances in Information and Communication Technology

  https://doi.org/10.1007/978-3-030-85969-5_31

  Abstract: S*elling complex data-based solutions is multifaceted as ecosystem actors perceive the value differently during the product lifecycle. The purpose of this paper is to study data-based solution sales in business ecosystems by presenting findings from nine recent interviews with manufacturing industry professionals. The results are demonstrated in two categories: 1) challenges in sharing, selling and buying data, and 2) the value of data for different actors in an ecosystem-like business environment. The managerial implications consist of clarifying the scattered viewpoints for selling data-based solutions and value formulation for different actors in an ecosystem. Theoretical contributions provide important aspects for the gap between business and sales research of data-based solutions, as current literature mainly focuses on the technical aspects.*

- Rantala, T.; Kokkonen, K. & Hannola, L. (2021). *Selling digital twins in business-to-business markets*. In Real-time Simulation for Sustainable Production: Enhancing User Experience and Creating Business Value (1st ed.) (51–62). Routledge.

  https://doi.org/10.4324/9781003054214-6

  Abstract: *Selling value generated by digital solutions for customers is different than selling traditional physical equipment. The purpose of this chapter is to study how digital twin sales can be promoted. The chapter combines the literature of business-to-business (B2B) sales with new information coming from a study on digital twinning in a manufacturing context. The study includes the results of ten semi-structured theme interviews that were conducted in January and February of 2020 with six digital twin provider companies from the areas of intelligent machine control systems, software solutions, real-time simulation, and machine engineering. The result is a model illustrating the levels of customer and data-based-solution understanding needed to successfully sell digital twins in the B2B market. The model was applied to the six case companies, each with a different digital-twin solution, and practical examples are presented that were revealed during the qualitative interviews. The empirical study identified different challenges related to the sale of differing levels of the digital-twin solution. The results indicate that companies that targeted more complex digital-twin developments required a higher level of technological proficiency from their salespeople and a stronger understanding of the value elements for the customers, their businesses, and their business ecosystems.*

- Rantala, T.; Kokkonen, K. & Hannola, L. (2020). *Impacts of Digital Twins: Significance from Sales Perspective*. In The ISPIM Innovation Conference: Innovating in Times of Crises Lappeenranta University of Technology. June 7–10, 2020, online, LUT Scientific and Expertise Publications.

  https://cris.vtt.fi/en/publications/impacts-of-digital-twins-significance-from-sales-perspective

  Abstract: *Disruptive innovations, data utilization and digital twins have received significant attention in scientific publications in recent years. Complex data-based solutions, for example digital twins, require understanding of the customer's business and the technological solution. This paper presents a framework for understanding the impacts of digital twins and discusses the sales perspective based on the interview findings. The qualitative interview data included altogether 16 semi-structured interviews from six case companies from the areas of intelligent machine control systems, software solutions, real-time simulation and machine engineering. As a result, the impact of digital twins can be classified in four categories: process change, competition environment* *change, role chain in value chain, and significant turning point in the market and emerging ecosystems. For example, when the impacts of complex digital twins are for the ecosystem, the importance of customer and ecosystem understanding and generated value is growing.*

- Kokkonen, K.; Hannola, L.; Rantala, T.; Ukko, J.; Saunila, M. & Rantala, T. (2020). *Digital twin business ecosystems: Preconditions and benefits for service business*. In Proceedings of the 21st CINET Conference: Practicing Continuous Innovation in Digital Ecosystems. September 20–22, 2020, Virtual and Milan, Italy.

  https://cris.vtt.fi/en/publications/digital-twin-business-ecosystems-preconditions-and-benefits-for-s

  Abstract: *In recent years, digital twins have been raised as enablers for several new service business opportunities for manufacturing companies as well as their service providers. However, these companies often need to consider an ecosystem-based approach in the development of their service business related to digital twins. By combining the examination of digital twin–based service opportunities in the manufacturing industry and recent perspectives on digital business ecosystems, we aim to clarify what kind of preconditions, challenges and benefits must be considered when designing digital twin business ecosystems. The empirical study is conducted as a multiple case study in six companies; digital twin utilizers in the manufacturing industry and companies representing their service and technology providers. The results indicate that at its best, a digital twin business ecosystem is based on open collaboration and provides inter-organizational value for its members. However, there are many preconditions and challenges to be tackled in ecosystem design, such as willingness to be open and share data, and the technological fit of solutions and systems within the ecosystem. In addition, formulating shared goals between participating companies and adopting a customer-oriented approach were identified as key preconditions for a successful digital twin business ecosystem.*

- Kokkonen, K.; Hannola, L.; Rantala, T.; Ukko, J.; Saunila, M. & Rantala, T. (2022). *Building service business in manufacturing via digital twin business ecosystems*. Submitted to the International Journal of Computer Integrated Manufacturing.

  Abstract: *This study focuses on the preconditions and benefits related to designing business ecosystems around digital twins. The empirical evidence builds on qualitative data gathered from six case companies. The results are studied on company and ecosystem levels. The results show that at the individual company level, the emergence of the ecosystems is determined by goal orientation and willingness to share information. The expected benefits of the ecosystems are related to effective resource utilization, possibilities for learning and development, as well as new business creation. At the ecosystem level, forming a digital twin business ecosystem requires technical receptivity and common interests and goals among the ecosystem members, as well as a strong customer-centric approach. At its best, a digital twin business ecosystem has the potential to provide inter-organizational value for its members.*

2.2 Digital twin technical and applications point of view

2.2.1 Description

The objective of the DT-based technical research and implementation can be divided into data-based and physics-based approaches. The overall emphasis in WP2 was on technical means to increase the value of data – whether it would be created by the physics-based simulator or by measurements.

The data-based technical perspective to the DTs can be explained by the combination of methods, procedures, and guidelines to perform data analyses, feature extraction, feature selection, model development, model training, and model testing. These tools were used and developed to improve the capabilities of the DT implementations in industrial usage. Separate tasks of A) data analytics & information generation procedures, B) feature engineering, and C) preparation of machine learning models for situational awareness and testing of them, were introduced as a practical development framework. Existence of good quality data is the main prerequisite of all data-based methods, procedures, and guidelines. In WP2 the measurement data of industrial-scale engine-generator set was used to demonstrate the progress of data-based approaches. Anyhow, the practical data delivery processes were out of the scope of the WP2.

To extend the data-based approach to also include physics-based model-driven approaches, the physics-based perspectives were addressed by the development of hybrid methods for the DTs. In essence, a hybrid method integrates physics-based model-driven methods and data-driven measurement-based methods. In WP2 a hybrid method for virtual sensing was utilised to gather data, which cannot be discovered by traditional approaches alone. Aim of this approach was to predict or measure virtually dynamic response of whole structure by minimum number of indirect measurements. Sensor placement optimisation methods were studied to find minimum number and optimal placement of the sensors. To illustrate the performance of hybrid modelling and the sensor placement optimisation, the method was applied for dynamic stress monitoring using only acceleration measurements. The method was studied with numerical models of different type of example structures. In addition to laboratory scale structures, industrial-scale engine-generator set was also used as a numerical case study. For validating the method, experimental studies were conducted with the laboratory scale structure.

2.2.2 Methods

Distinctive methodological backgrounds were applied to study the technical means to increase the value of data in the data-based and physics-based approaches. Hence, the studies in this part of the project were organised to subtasks as given in the following list:

- Subtask 1: Data analytics were applied to the available measurement data and used as a standardised technique to analyse the intrinsic nature of the data. Moreover, the information generation procedures (i.e., the joint usage of simulated and measured information) were investigated within the main concepts of knowledge discovery and data engineering. Additionally, the subtask included approach for data normalisation and harmonisation.
- Subtask 2: The data engineering procedures provide the normalised and cleaned data to be examined and modelled by machine learning methods and other advanced approaches. The feature engineering process is the key component to extract valuable information from the data. Thus, the subtask included stages for formulation of feature definitions for operational states, implementation of feature identification and extraction procedures of several different signal processing -based features, investigation and evaluation of feature comparison and selection methods, and initial assessment of feature classification.
- Subtask 3: The extracted features were used in the development process for machine learning model. The main objective of the developed models was to address the state recognition process. The subtask included separate stages for model preparation and implementation, implementation of the actual learning processes, determination of test cases, model testing and reporting.
- Subtask 4: State-of-the-art survey on hybrid modelling methods for the DTs were conducted. The potential hybrid methods of structural monitoring especially suitable for industrial type of applications were studied.
- Subtask 5: Development of best practices in hybrid modelling for engineering stage of lifecycle were conducted. Sensor placement optimisation methods were studied to find minimum number and optimal placement of the sensors for the hybrid modelling.
- Subtask 6: Analysis techniques and tools for hybrid modelling and virtual sensing for operative stage of lifecycle point of view were developed. Important technique to be developed and implemented was monitoring of complete system with limited number of indirect measurements utilising simulation model and sophisticated expansion techniques.
- Subtask 7: The developed methods were validated by experiments in laboratory. This task included validation experiments of hybrid modelling and sensor placement optimisation method. A new experimental research setup was designed and implemented for this research.

- Subtask 8: State-of-the-art study on IoT technologies and solutions for digital twins. A small-scale literature study was done to find out what kinds of services the Industrial Internet of Things (IIoT) solution providers have for DTs.

2.2.3 Results

Data-based approaches were investigated in a condition monitoring scenario, where continuous near real-time state recognition based on measurement data (acceleration data) of an industrial-scale engine-generator set was in the focus. The main outcomes and conclusions of the work were that the operational state of the engine-generator set can be accurately defined and identified based on measured vibration data by using traditional classifier algorithms. Additionally, an efficient two-step state recognition model can be built by combining the state classification, i.e., the identification of the known states, and novelty detection (the identification of the unknown states) models and processes. However, fast and accurate state recognition is also possible by using a classifier trained with the advanced neural network -based models.

A hybrid modelling method for physics-based approaches was developed for the industrial virtual sensing applications. A sensor placement optimisation method was developed to find minimum number and optimal placement of the sensors for the hybrid modelling. The developed method was validated with experimental studies, which also included quantification of actual modelling errors. The developed sensor placement optimisation method is especially applicable to large industrial-scale structures.

The main results are comprised in following publications.

Publications:

- Junttila, J. (2021). Operational State Recognition of a Rotating Machine Based on Measured Mechanical Vibration Data. Master's Thesis, Master of Engineering – Big Data Analytics, Arcada University of Applied Sciences, Helsinki.

  https://www.theseus.fi/bitstream/handle/10024/501834/Junttila_Jukka.pdf

  Abstract: *Digital twin is a relatively new concept. Also, it lacks a formal definition and can be applied in virtually any field of technology. Considering digital twins of rotating machines, and especially the in-service phase of their lifecycle, a digital twin should produce valuable information for the owner and operator of the application. The information produced by a digital twin should be accurate, up-to-date, and available anywhere. These requirements act as limiting factors for the complexity of the digital twin and promote the need for efficient data transfer, data acquisition and especially data processing methods at the source of information.*

  This study investigates how these requirements can be fulfilled in continuous, near real-time operational state recognition of a gas engine genset. Therefore, the objective of this study is to provide a data-based model for operational state recognition and detection of abnormal operation of a gas engine generating set in near real-time. Two different types of machine learning models for the state recognition of the generating set are presented. The first, a classification model, can identify the current power output level of the generating set using the measured mechanical vibration data. The second, a novelty detection model, can detect abnormal operation of the generating set, in fault situations, at a specific power output level. A two-step state recognition model can be built by combining the classification and novelty detection models.

- Junttila, J.; Lämsä, V. & Espinosa-Leal, L. (2021). *Extreme Learning Machine-Based Operational State Recognition: a Feasibility Study with Mechanical Vibration Data*. The 11th International Conference on Extreme Learning Machines (ELM2021), 15–16 December 2021, online, in publishing.

  Abstract: *The benefits of digital twins and accurate near real-time onsite condition monitoring of heavy machinery or load-bearing structures are undeniable. Both demand computationally light* *and accurate models based on continuously measured data. Extreme Learning Machine (ELM) algorithm provides the means for building accurate and fast predicting classification models. Therefore, the feasibility of the ELM algorithm for building models for near real-time operational state recognition of a rotating machine was studied. Three different models, called one, two, and six cycles, built using the ELM algorithm were compared with corresponding models trained using Support Vector Machine (SVM) and linear regression (LR) algorithms based on their accuracy and prediction times. The comparisons show that the SVM algorithm produces the best accuracy but with the cost of high prediction times. The LR models have the lowest prediction time. In contrast, the ELM model for the two cycles presents better performance than the corresponding LR and SVM models when the combination of accuracy and the prediction time is considered. The great benefit of the ELM method comes from its mathematical properties: new data can be added to the ELM model without the need to retrain the whole model, and the model is competent to take strong nonlinearities into account. Thus, the possibilities of the ELM algorithm to act as a novelty detector in operational state recognition shall be investigated.*

- Junttila, J. & Lämsä, V. *Near real-time state classification procedure for gas engine generating set based on feature engineering*. The article manuscript is in progress.

  Abstract: *This paper proposes general procedure and guideline to perform data standardization, feature extraction and feature selection for real-time state classification of a gas engine generating set in fault ride through tests. Methods are linked to the real-life industrial Big Data application supporting a hybrid-modelling -based Digital Twin approach. Faults and anomalies in structures and machines can be identified from characteristic features extracted from the vibration measurements. The first task in successful detection procedure is the elimination of distracting effects from the raw vibration data, i.e., data normalization and standardization, which is essential for reliable detection. The second task is the feature extraction done by defining and calculating relevant information from the normalized vibration data. In feature selection step, a relevant subset of the extracted sensitive features are used as input variables for a classifier model, so that the generalization error of the model is as small as possible. This paper proposes general procedure and guidelines with mathematical background to perform data standardization, feature extraction and selection in continuous and real-time high sampling rate application of a gas engine generating set state classification.*

- Zeb, A. & Kortelainen, J. (2021). *Industrial IoT solutions for digital twins: An overview*. VTT Research Report VTT-R-00782-21, Espoo.

  https://cris.vtt.fi/ws/portalfiles/portal/53821515/IIoT_solutions_for_DTs_AnOverview_.pdf

  Abstract: *A digital twin (DT) is a digital representation of a real-world entity, such as a device, machine, process, or complex system. The regular synchronisation between the DT and its physical counterpart offers better monitoring, improved performance, optimised maintenance, reduced downtime, and a network of connected products.*

  Industrial Internet of Things (IIoT) is considered as a subset of the Internet of Things (IoT) that connects various industrial assets (including machines and control systems) with the information systems and the business processes and can be utilised for optimal industrial operations. The IoT essentially connects 'things' to the Internet and to networks that use Internet technology. These things or items collect and share data about their internal state, the objects to which they are attached, and the environment they are in through gateways and edge computing devices. The data from such items, when fed to the DT models that combines modelling and analytics techniques using, e.g., artificial intelligence, provide information about the past and present operation and forecast the future of its physical counterpart, thus enabling the prevention of minor problems from turning into major ones and extending asset's lifecycle.

  This work introduces some of the prevailing IIoT platforms for developing DTs indicated by various studies. Information about the IIoT solutions are collected from websites and online documentation. These platforms have similar types of capabilities, with one IIoT platform performing better in one *area than another. The selection of a suitable platform could be challenging as there are very few examples documented in the literature and the available information is largely from marketing materials. The future work could be focused on exploring the practical applications and limitations or scope of the IIoT platforms introduced in this study.*

- Nieminen, V. & Sopanen, J. *Optimal sensor placement of triaxial accelerometers for modal expansion*. Submitted to the Journal Mechanical Systems and Signal Processing.

  Abstract: *Sensor placement is a vital factor affecting the quality and accuracy of virtual sensing. Modal expansion techniques are well-known methods to expand the measured displacements or accelerations to all unmeasured degrees of freedom. For this purpose, a two-phase sensor placement optimisation method is proposed for commonly used triaxial accelerometers. The method uses minimum variance criterion of an estimation error of structural responses. A measure of redundancy of information is introduced as an additional criterion for the placement of the triaxial sensors to minimise the redundancy between the sensors. This was addressed to avoid spatial correlation and clustering of the sensor locations. In addition, a proposal for modal displacement-based weighting is introduced to avoid potential selection of sensor locations with low vibration energy, which can be critical in noisy environments. The efficiency of the proposed method is verified with numerical models of different types of structures and finally with the laboratory scale experiments. This method is especially applicable to large finite element models of industrial scale structures with fine meshes.*

- Neisi, N.; Nieminen, V.; Kurvinen, E.; Lämsä, V. & Sopanen, J. *Estimation of unmeasurable vibration of rotating machine using Kalman filter*. The article manuscript is in progress.

  Abstract: *Typically, the rotor-bearing systems are equipped with vibration sensors near the bearing locations to assess the machine condition. The state change in a real mechanical system to be monitored over the lifetime may be e.g., due to change of the unbalance, stiffness of the support or the bearings, mechanical failure, etc. However, there are only limited measurement points where the vibration can be measured. Therefore, the evaluation of machine conditions might not be straightforward. In this work, an experimental and numerical study was conducted on a rotor-bearing system. The Kalman filter is used to estimate the condition of the machine in unmeasurable points for vibration. The real-life system behaviour is nonlinear. The idealised system dynamical model is linearized, and the model has a good accuracy with respect to the experiments. The Unscented Kalman Filter (UKF) was selected for the state estimation. The real-life main structure is mounted on a heavy pedestal, and the machine pedestal is supported by its weight. The experiments show that the machine pedestal has translational and rotational movement. Thus, the experimental setup resembles the complex real-life industrial multi-fault scenarios. The present study provides a simplified simulation model that accounts for the rotor unbalance and frame movement. The finite element model (FEM) is used to model the flexible rotor. The bearings are modelled with constant stiffness and damping, where the bearing properties are approximated based on the vibration measurements. The machine pedestal is modelled with an equivalent mass, where the stiffness and inertia properties of the machine pedestal are estimated by the natural frequencies of vibration measurements at the pedestal. The system is studied for two sweeps in the experiments. First, the machine is running with its current configuration, and then a specific unbalance mass was added to a disc of the rotor. In the experiment, the vibration at the bearing locations is measured by accelerometer sensors. The measured acceleration is used as an observation, and the Kalman filter is applied to estimate the response at other locations. The results indicate that the estimated vibration along the rotor nodes have a reasonable agreement with the rotor whirling.*

2.3 Digital twin lifecycle point of view

2.3.1 Description

DTs in industrial applications are often related to products or systems that are expected to have long lifecycle. Depending on the application, the role of the DT may emphasise the early phases of the product lifecycle, especially engineering and design, or it may be crucial for the whole lifecycle. The later one is the case, when the DT is directly affecting the operation of the system, e.g., as an element of the control system.

In this part of the project, the research focus was on so-called functional DTs, i.e., DTs that are based on numerical simulation. To operate properly, the simulation of a DT typically requires:

- A computing hardware to run the simulation programs and any required auxiliary processes,
- A computer operating system that provides the computing environment, such as the file system and process scheduling,
- A simulation software for running the actual simulation, and
- A simulation model of the DT.

Preserving the DT and the underlying computing infrastructure operational may require addition efforts if the expected lifecycle of the system is long. The fast evolvement in computing technology makes it challenging to preserve computing technology for as long time periods as decades. The focus of the research was on general means to preserve computing and simulation data for long time periods, such as standardisation and data modelling.

2.3.2 Methods

The study on DT data lifecycle management contained mainly literature studies, but also included experimentations with data modelling. The main effort within this project task was to study the means to preserve simulation model data for long lifecycles.

The research was organised as follows:

- Subtask 1: A study of the current state in standardisation and existing open specification of data models for functional DTs, including standardisation of structural finite element method (FEM), multibody system simulation, computational fluid dynamics (CFD), and general system simulation. The work started with a study of the ISO standard 10303-239 *Product life cycle support*[1]. It was soon found out that the standard did not provide tools for detailed description of simulation models.
- Subtask 2: A study of different approaches and strategies to explicitly define the necessary simulation model data for a functional DT. The identified strategies included standardisation, data modelling with common data modelling technologies, use of simulation modelling languages, use of software source code, and preservation of the whole computing infrastructure of the DT. The study also included an analysis of bottlenecks and risks in data management of DTs in long lifecycles.
- Subtask 3: A study on data modelling of the selected case examples in selected simulation domains, based on existing standards, open data specifications or industrial de facto standards. The objective was to study, if common data modelling approaches could be used for DT model data description. A small experimentation was done to test how to represent system simulation models with *Web Ontology Language* (OWL)[2]. The original system simulation model was presented in the

[1] ISO standard 10303-239:2005: https://www.iso.org/standard/38310.html
[2] Web Ontology Language, W3C recommendation: https://www.w3.org/TR/owl2-overview/

Modelica modelling language[3]. In addition, the detailed data features of system simulation, FEM, and CFD models were studied further.

2.3.3 Results

The main outcome and conclusion of the work was that significant efforts are needed to improve the preservability of model and simulation data for functional DTs for long time periods. The level of standardisation and the technical means for preserving critical simulation model data, especially for system simulation, are still modest. At the same time, the interest in applying the DT concept is increasing, which introduces the risk that the lifecycle aspects of the technology are not always taken into account. The research group did not find any novel solutions for the problem but proposes that the issue is in general analysed early enough and existing means are evaluated for each case in hand.

The outcome of the research was the set of publications listed below:

- Rantala, T.; Saunila, M.; Ukko, J.; Kortelainen, J. & Zeb, A. (2021). *Managing digital twin life cycle – recognition and handling of business risks*. In Real-time Simulation for Sustainable Production: Enhancing User Experience and Creating Business Value (1st ed.). Routledge.

  https://doi.org/10.4324/9781003054214-20

  Abstract: *While the adoption of digital twins among industrial companies is growing due to rapid developments in available solutions, the utilization of digital twins as a part of company operations and management is a new phenomenon. As a result, the long-term functionality of the digital-twin lifecycle is not well understood. Because of the small number of example cases, there is a lack of existing case data. This chapter offers both theoretical and empirical insights into the business-related risks that can arise during the different digital-twin lifecycle phases. The empirical information introduced here came from two Finnish multinational industrial companies that are planning to develop new service businesses based on digital twins.*

- Zeb, A. & Kortelainen, J. (2021). *Web Ontology Language data modelling of Modelica simulation models*. VTT Research Report VTT-R-01517-20, Espoo.

  https://cris.vtt.fi/ws/portalfiles/portal/44925967/Web_ontology_language_data_modelling_of_Modelica_simulation_models.pdf

  Abstract: *Digital twins are promising several benefits, including better monitoring and control, improved performance, and lower maintenance costs of their physical counterparts. However, there are still some challenges associated with the development and management of digital twins that need to be addressed before the actual realisation of digital twins' applications. In this study, we look at the data management issue of digital twins that are based on system simulation.*

  The problem with most of the commercial modelling and simulation software is the use of proprietary data models and specific formats to store these models, resulting in poor interoperability. Since digital twins of certain assets need to be operated for a couple of decades, there is a possibility that the underlying application software may not be available through the assets' life cycles. In order to avoid unknown risks or business disruption, it is advisable to preserve information about the simulation models of digital twins in long-lasting formats. This is to ensure that in situations where the original software is inappropriate/inaccessible for opening/running a digital twin' simulation model, one should be able to revive the model, using the preserved information, and adopt alternative tools.

  This work illustrates the long-term data preservation of digital twins by exploring two Modelica system simulation models. Modelica is an open-source language for modelling complex systems and

[3] Modelica modelling language, the Modelica Association: https://modelica.org/

provides detailed information about the models of components that can be pre-served in user-specific formats. The models were built in OMEdit environment, a free tool for Modelica implementation. Data models of the corresponding simulation models were developed in Web Ontology Language (OWL), which is used for standardised representation of in-formation on the Semantic Web. Protégé ontology editor was used to design four ontologies that provide the necessary concepts and relationships for describing the simulation models. Using these concepts and relationships, the OWL data models were generated. It was found that OWL lacks certain features and the visualisation of data models gets complex as the model size grows.

- Zeb, A.; Kortelainen, J.; Rantala, T.; Saunila, M. & Ukko, J. *On the alleviation of imminent technical and business challenges of long-lasting functional digital twins*. Submitted to the Computers in Industry journal.

  Abstract: In this article, we discuss the technical and business risks associated with long-lasting functional digital twins, and describe different strategies for their alleviation. Functional digital twins are based on physics-based simulation models and are operated alongside the life cycle of their physical counterparts. These simulation-based digital twins are built using a simulation software. The problems with most of the commercial modeling and simulation tools are their black box nature and storing data in protective formats, leading to poor interoperability. Since the digital twins of certain assets need to be operated for a long period, even for several decades, there is a possibility that the computing infrastructure, i.e., the computing hardware and software, may not remain the same throughout the product or system life cycle. The computer hardware and operating systems are usually third-party components with limited choices for their users, whereas the selection of simulation tools is more flexible and the designer can choose from, for example, commercial, open-source, or in-house solutions. To avoid substantial costs or business disruption, the digital twin providers must be able to reproduce the underlying simulation models with up-to-date tools and adopt alternative solutions whenever needed. The findings of the study are presented in the form of propositions throughout the article.

- Kortelainen, J.; Minav, T. & Tammi, K. *Digital twin – The dream and the reality*. In review in the *IEEE Access* journal.

  Abstract: *Digital twins (DTs) are under active research and development in research community, industry, and in the digital engineering solution business. The roots of the concept of DT are almost two decades old, but the fast progress in enabling technologies, especially in data analytics, artificial intelligence, and the Internet of Things, has accelerated the evolution of DT during the last five years. The growing interest, increasing development activities, and raising business opportunities of the concept are also feeding the hype in media. Consequently, this has led to the scattering and even misuse of the concept and its definition. In this article, we discuss different applications of DTs and what kind of solutions there are for DTs. We analyze some most cited definitions of DT in scientific literature and discuss the interpretation of the definitions through a hypothetical case example. Furthermore, we discuss different life cycle aspects of DTs and potential risks that there may raise. To further concretize the concept of DT, we introduce nine reported case examples of implemented DTs in scientific literature and analyze their features. Eventually, we discuss the future development directions of DTs and the aspects that are affecting the development trends.*

2.4 Experimental demonstrators and setups

VTT's existing rotating machine demonstrator system was used for experimental demonstrations and further developed in the project (Figure 1). The demonstration system was used to study the connection of a real system with a corresponding simulation model. The focus of the studies was on the use of Kalman filters with DTs to estimate unmeasurable vibration of rotating machine. In the experiments, the vibration at the bearing locations was measured by accelerometer sensors. Thereafter, the measured acceleration was used as an observation, and the Kalman filter was applied to estimate the response at other locations.

The applied process and its implementation provide a methodology and process for modelling and analysing the complicated real-life systems. The demonstrator system was used for research collaboration between VTT and the LUT University. The joint publication of the demonstrator is under progress.

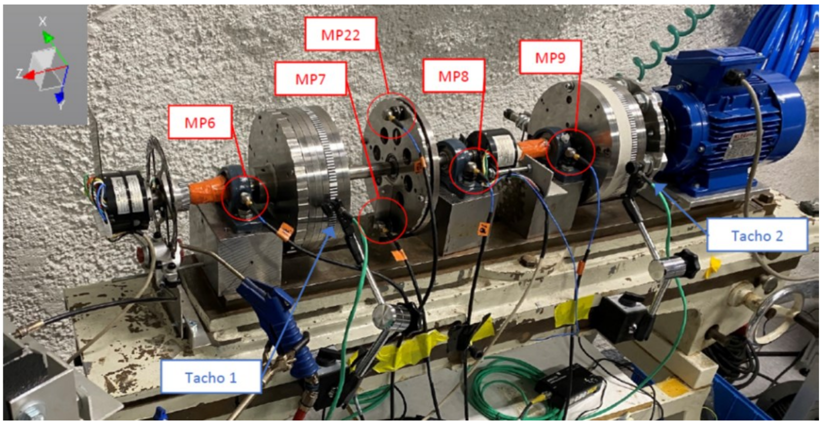


Figure 1: The rotating machine demonstrator environment. The rotational speed tachometers (Tacho 1 and Tacho 2), and some of the measurement points are marked in the picture.

A new experimental research setup was designed and implemented for the research on sensor placement optimisation (Figure 2). The test structure consisted of assembly of the steel plates, which were joined together by bolt joints introducing non-linearities into the structure. The setup was used for verification measurements for development of the hybrid modelling and sensor placement optimisation method. The work included implementing a detailed high-fidelity structural analysis model (FEM) of the test structure. Aim of the case study was predicting dynamic state and stress monitoring of whole structure by limited (minimum) number of indirect measurements.

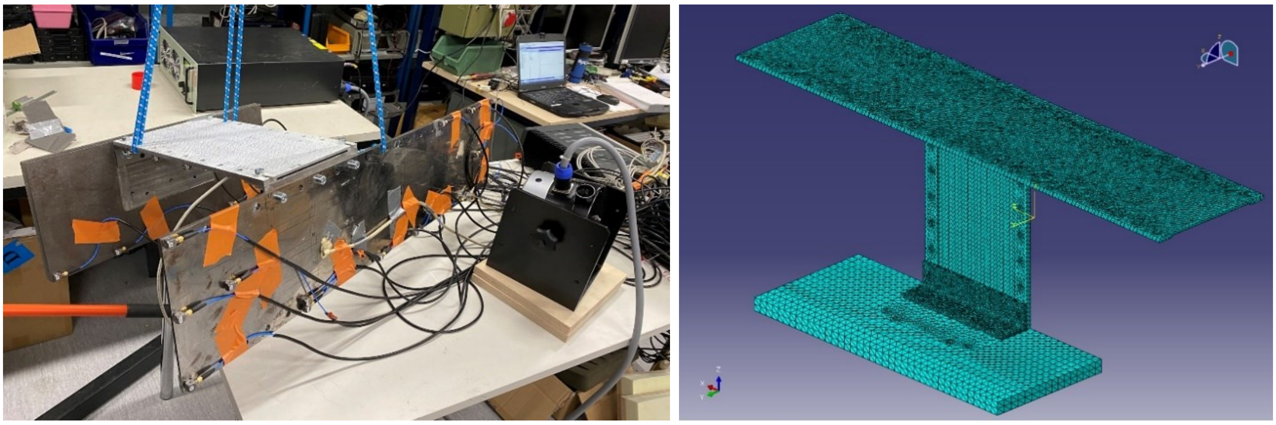


Figure 2: Sensor placement optimisation test setup and structural FE-model.

2.5 Project collaboration, events, and seminars

The whole DigiBuzz common effort had active collaboration between the participating organisations throughout the project time. The collaboration and knowledge sharing were implemented as regular project seminars combined with a project steering group meeting during the same day, regular monthly status meetings between all the parties, and targeted meetings focusing on certain topics with a smaller set of participating organisations. In addition, the project had a common Microsoft Teams workspace (organised by the LUT University) that was used for document management and sharing between all the project parties.

The DigiBuzz-VTT project participated the following scientific conferences, where and conference article was published and a presentation was given:

- PRO-VE 2021, Saint-Étienne, France, November 22–24, 2021.
  Rantala, T.; Valjakka, T.; Kokkonen, K.; Hannola, L.; Timperi, M. & Torvikoski, L. (2021). *Selling the Value of Complex Data-based Solution for Industrial Customers*.
- The ISPIM Innovation Conference, June 7–10, 2020, online.
  Rantala, T.; Kokkonen, K. & Hannola, L. (2020). *Impacts of Digital Twins: Significance from Sales Perspective*.
- The 11th International Conference on Extreme Learning Machines (ELM2021), December 15–16 2021, online.
  Junttila, J.; Lämsä, V. & Espinosa-Leal, L. (2021). *Extreme Learning Machine-Based Operational State Recognition: a Feasibility Study with Mechanical Vibration Data*.

The DigiBuzz common effort had the following internal seminars where VTT gave the following presentations:

- DigiBuzz researchers' seminar, 31 January 2020, LUT University Lahti campus
  - *DigiBuzz-VTT – Research topics and recent activities*;
    Tiina Valjakka, Tuija Rantala, Akhtar Zeb, Ville Lämsä, Vesa Nieminen, Jukka Junttila, Juha Kortelainen, and Juha Virtanen, VTT
- DigiBuzz project seminar 1, 10 August 2020, virtual event
  - *Selling digital twins in business-to-business markets*;
    Tuija Rantala and Tiina Valjakka, VTT; Kirsi Kokkonen and Lea Hannola, LUT
  - *WP2.2 – Data, information and decision-making*; Ville Lämsä and Jukka Junttila, VTT
  - *DigiBuzz-VTT Digital twin data life cycle management*;
    Juha Kortelainen and Akhtar Zeb, VTT
  - *WP2.4 & 2.5 – Virtual sensing for structural monitoring in engineering and operative stage of lifecycle*;
    Vesa Nieminen and Mikko Savolainen, VTT
  - *DigiBuzz-VTT – Introducing international consortiums*;
    Ville Lämsä and Juha Kortelainen, VTT
- DigiBuzz project seminar 2, 4 February 2021, virtual event
  - *Data-based identification and classification of events in industrial high sampling rate applications*;
    Ville Lämsä, VTT
  - *Digital twin data life cycle management*;
    Kortelainen Juha and Zeb Akhtar, VTT
  - *Sensor placement optimization for virtual sensing*;
    Vesa Nieminen, VTT
- DigiBuzz project seminar 3, 27 May 2021, virtual event
  - *Sensor placement optimization, results of full-scale engine-generator set*;
    Vesa Nieminen, VTT
- DigiBuzz project seminar 4, 24 September 2021, virtual event
  - *Operational state recognition of a rotating machine based on measured mechanical vibration data*;
    Jukka Junttila, VTT
  - *IIoT solutions for DT*;
    Akhtar Zeb, VTT
- DigiBuzz project final seminar, 20 January 2022, virtual event
  - *Extreme Learning Machine-Based Operational State Recognition: a Feasibility Study with Mechanical Vibration Data*;
    Jukka Junttila and Ville Lämsä, VTT
  - *Optimal sensor placement for modal expansion; experimental validation*;
    Vesa Nieminen, VTT

- *Selling the value of complex data-based solution for industrial customers*; Tuija Rantala, VTT

The project's steering group meetings were focusing on project administrational topics and progress, and they were rather short as the research work progress was already presented in the project seminar before the meeting the same day.

The project monthly status meetings were organised as virtual meetings. The project research work progress was presented in task level both for the LUT University and for VTT. In some of the meetings, also the participating companies presented their research and development activities in their own project. There was also reserved some time for discussion, questions, and answers.

In addition to regular status meetings, targeted meetings were organised based on the companies' interests and activity. These meetings had narrow scope and the emphasis was on topics that the companies had indicated. VTT had also dedicated meetings with Wärtsilä to synchronise the research work in the parallel projects that had interfacing topics, such as the ARTIE project (Wärtsilä).

3. Conclusions and summary

The research conducted in the DigiBuzz-VTT research project continued the work done in earlier research and development projects, such as the SIMPRO project and several VTT internal research projects, focusing on the technologies and applications of DTs. The emphasis of the research was clearly closer to the applications and business opportunities compared to the previous, technology focused research projects. The project had good collaboration between the common effort parties, both the LUT University and the participating companies.

Already in the beginning of the project, the concept of DT was found to be broad, fuzzy, and evolving. During the project, the concept and its definition was found to be even more vague, which may have challenged the related communication. The project did not put any additional efforts to clarify the definition or to formulate yet another definition of DT because it was not seen to solve the challenge.

The research on business opportunities showed in general that DTs are seen as an attractive technology for companies and that there are numerous applications for DTs. The current focus of DTs emphasises the data-based approach over the simulation-based approach. This may be due to the more mature offering of enabling technologies, such as data analytics and IoT.

The main conclusion regarding the data-based approaches was that continuous near real-time state recognition of an industrial-scale engine-generator set is feasible. Efficient and accurate recognition models (i.e., state classification and novelty detection) can be built upon traditional or advanced algorithms. Some of these models have enhanced practical capabilities to add new data without the need of retraining the whole model.

The life cycle management of DTs was studied especially from the simulation model data management point of view. The scope of research included both technical data management and the business aspects. In addition, the lifecycle risks and means to mitigate the risks were studied. The main conclusion of the study was that the means to preserve detailed simulation model data for long lifecycles are modest and more development is needed to improve the situation. Standardisation was seen as one of the major approaches.

Physics-based perspectives were combined with data-driven measurement-based methods by the development of hybrid methods for the DTs. Sensor placement optimisation methods were studied to find minimum number and optimal placement of the sensors. To illustrate the performance, the method for dynamic stress monitoring using only acceleration measurements was studied with numerical models of different type of structures. Finally, the developed methods were validated successfully in laboratory experiments.

The small-scale study on the IIoT solutions for DTs showed that there is already a rich set of solutions available. The IoT technology in general has been identified as one of the main enablers for the development of DTs. The big solution providers, such as Microsoft and Amazon, are actively developing their solutions, starting from the data-based DTs, and most likely extending the development to also support simulation-based DTs.

The research parties, the LUT University and VTT, collected the main outcomes and conclusions of the research into a common scientific article:

- Kurvinen, E.; Kutvonen, A.; Ukko, J.; Khadim, Q.; Jaiswal, S.; Neisi, N.; Zhidchenko, V.; Kortelainen, J.; Timperi, M.; Kokkonen, K.; Virtanen, J.; Zeb, A.; Lämsä, V.; Nieminen, V.; Junttila, J.; Savolainen, M.; Rantala, Tu.; Valjakka, T.; Donoghue, I.; Elfvengren, K.; Rantala, Te.; Kurinov, I.; Sikanen, E.; Pyrhönen, L.; Hannola, L.; Handroos, H.; Rantanen, H.; Saunila, M.; Sopanen, J.; and Mikkola, A. *Physics-Based Digital Twins Merging with Machines: Cases Mobile Log Crane and Rotating Machine*. In review in the *IEEE Access* journal.

  Abstract: *Real-world products and physics-based simulations are becoming interconnected. In particular, real-time capable dynamic simulation has made it possible for simulation models to run in parallel and simultaneously with operating machinery. This capability combined with state observer techniques such as Kalman filtering have enabled the synchronization between simulation and the real world. State estimator techniques can be applied to estimate unmeasured quantities, also referred as virtual sensing, or to enhance the quality of measured signals. Although synchronized models could be used in a number of ways, value creation and business model development are currently defining the most practical and beneficial use cases from a business perspective. The research reported here reveals the communication and collaboration methods that lead to economically relevant technology solutions. Two case examples are given that demonstrate the proposed methodology. The work benefited from the broad perspective of researchers from different backgrounds and the joint effort to drive the technology development towards business relevant cases.*

References

Yin R. K., (2014). Case Study Research: Design and Methods. 5th edition, Sage Publications, Thousand Oaks, p. 282.