



PDENG PROJECTS 2019

# Software Technology

**TU/e**

EINDHOVEN  
UNIVERSITY OF  
TECHNOLOGY

4TU.School for Technological Design, Stan Ackermans Institute offers 2-year postgraduate technological designer programmes and tracks. This institute is a joint initiative of the four technological universities of the Netherlands: Delft University of Technology, Eindhoven University of Technology, University of Twente and University of Wageningen. For more information please visit: [4tu.nl/en/](https://4tu.nl/en/)

# Software Technology - PDEng Projects 2019

- 4 **Giovanni de Almeida Calheiros PDEng;** Designing an Infrastructure for On-premise Software Deployment - A Case Study for Thermo Fisher Data Management Platform
- 6 **Pranav Bhatnagar PDEng;** Automatic Alignment of Microscope using Machine Learning
- 8 **Ganduulga Gankhuyag PDEng;** Model-Based System Engineering Design of Functional Modules for Configurable Topology
- 10 **Konstantinos Karmas PDEng;** VESSEL - Driving Behavior Monitoring System
- 12 **Gilda Khosravi PDEng;** Dependency Management and Scope Control
- 14 **Arne Lapõnin PDEng;** Distilling Insights from the Text Written by Field Technicians
- 16 **Dũng Lê PDEng;** Data Quality Dashboard
- 18 **Trupti Manik PDEng;** Increasing Coverage Through Coverage-based Test Selection
- 20 **Ani Megerdounian PDEng;** From Informal TWINSCAN Machine Diagnostics Data to Reliable Formal Customer Interface
- 22 **Munkhdalai Sainbileg PDEng;** Multi-domain Modeling Assistance (MMA) Tool: Design and Implementation
- 24 **Dimas Satria PDEng;** Data Storing and Structuring for Information Discovery in AgTech
- 26 **Konstantinos Smanis PDEng;** Intelligent Microscope II
- 28 **Hongyu Tan PDEng;** Integration of Modeling and Data Science
- 30 **Beza Getachew Tassew PDEng;** Sample Flow Modeling for Rapid Performance and Failure Analysis
- 32 **Wahyu Utomo PDEng;** Design Complexity Reduction Using Object-oriented Paradigm - Exemplified by Redesigning the Scanning Application of ASML's Wafer Positioning Module



**PROF.DR. M.G.J. VAN DEN BRAND**  
**DR. Y. DAJSUREN PDENG**

# Model-driven Software Engineering Meets Data Science

Software has been one of the driving factors for innovations in our society in the last 25 years. Software creates new ways of transportation, new ways of communication, new ways of production, etc. Whereas software was eating the world, data has become the fuel of the world. A lot of innovations are based on the fact that we have more data as ever before, this also changes the developments of high-tech systems. By means of performing analysis of these vast amounts of data we are able to identify subtle bugs, improve performance, predict maintenance, etc. It is remarkable that model-driven software engineering and data science evolve hand-in-hand and fit perfectly together. Based on results of machine learning existing models can be fine-tuned. We are on the very edge of deriving models from the data by means of advanced machine learning techniques.

Although that the importance of data and AI is already observable for a number of years, it becomes also more visible in the ST training program and final project assignments from industry. In this booklet you will find project examples of the current generation of PDEng trainees on how model-driven software development techniques are combined with data science and AI techniques. The applications range from using machine diagnostic data to deriving a reliable customer interface, and the use of data to identify problems in the paper flow of high-end printers, the development of a dashboard to analyse the quality of the vast of amounts of data itself, the development of a data management platform, the use of AI techniques to make microscopes more intelligent, and the use of AI techniques to analyse the logs of maintenance engineers, these are examples within the high-tech industry. Data plays also an important role in agriculture in order to predict and improve the harvesting of crop. Of course, a number of the graduation projects are closer to the traditional software design and model-driven software engineering, and deal with mastering the complexity of existing software, the refactoring of existing code to reduce complexity, improving the test coverage of code, all in the area of high-systems. Finally, a number of projects are related to the automotive domain, modelling the functional safety and analysing traffic data.

Our trainees have once again shown that they are able to tackle tough design problems and are able come up with innovative solutions that are eagerly adapted by our industrial partners. This booklet provides a source of inspiration and shows the challenges to advanced software development. We hope you enjoy reading the summaries of the projects. We would like to congratulate our trainees with their results and wish them a bright and successful career.

Mark van den Brand, Scientific Director  
Yanja Dajsuren, Program Director

PDEng Software Technology programme  
Eindhoven University of Technology





"Giovanni's contribution of system administration, automation and monitoring jumpstarted significant parts of the realization of our ambitions. He drove the realization of the entire automated build infrastructure, facilitated the deployment automation of build results, and built lightweight yet useful monitoring capabilities."

Ir. Egbert Algra PDEng  
Thermo Fisher Scientific

## CHALLENGES

The first challenge faced in this project comes from the fact that the Data Management Platform (DMP) is an already existing solution and any new process or solution should be carefully evaluated before being implemented. Once that the decision has been taken, sharing it with the other teammates that are geographically distributed also required some effort.

## RESULTS

The main results of this project are the redesign of the Kubernetes definitions in order to better accommodate customers' differences, the automated management of the delivering and installation for those kubernetes application, as well as the implementation of a Continuous Delivery pipeline in which different pieces of software from different sources can be packaged and deployed in a uniform and automated fashion.

## BENEFITS

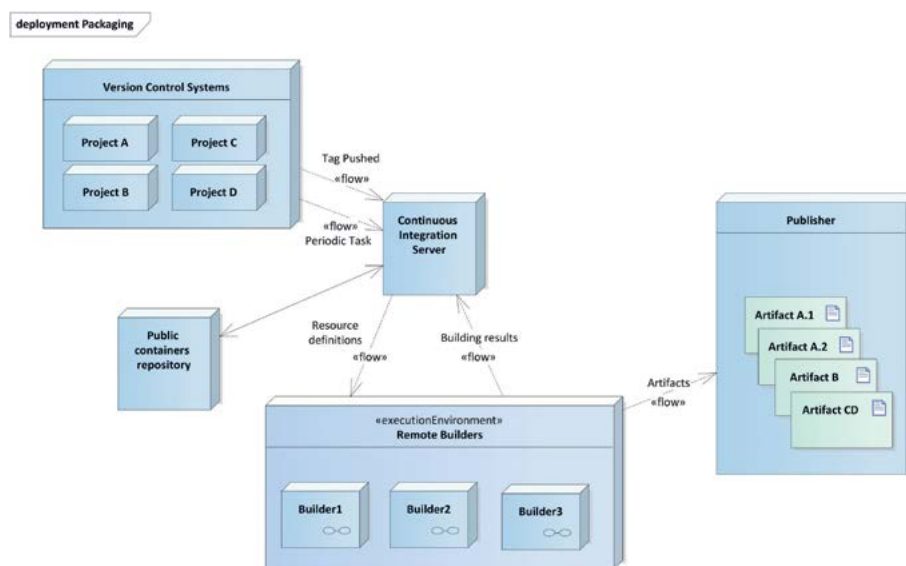
The emphasis on process automation will shorten the time required to roll-out software towards customers, reduce the possibility of human-error during the deployment of Thermo Fisher software, and create mechanisms to rollback in case of failure during update/upgrade process.

**GIOVANNI DE ALMEIDA CALHEIROS PDEng**

# Designing an Infrastructure for On-premise Software Deployment

## A CASE STUDY FOR THERMO FISHER DATA MANAGEMENT PLATFORM

Over the years, equipment for serving science has evolved. Its capability to generate the data as well as its needs for supporting software has increased. Due to the amount of data generated by the equipment, problems such as data storage, data analysis, and online collaboration have become more and more important. To address these problems, Thermo Fisher Scientific created the Data Management Platform (DMP).



DMP is a digital platform to enable easier and safe data management, efficient teamwork, online collaboration, and live image viewing. To implement these, the DMP platform requires automated application management and packaged software ready to use.

The proposed solution aims to improve the Application Lifecycle Management (ALM) in the DMP context. Tasks such as software packaging, software delivery, and software management were addressed in it.

## CHALLENGES


Due to the nature of the project, the challenge was to select and prioritize the beam shapes that had to be classified. Following that we needed an optimal amount of data that includes the great diversity of patterns, shapes, sizes and angles. Other challenge was to develop an automated machine learning toolchain that takes images as input and produces the trained algorithm with just one command.

## RESULTS

The work and outcome of this project prove that it is possible to use machine learning (ML) to increase the robustness of the traditional algorithms. From the training of ML classifiers using Transfer learning, we obtained 99.1% of total accuracy with 0.94% of false positives. This can be further reduced by applying appropriate filters, increasing the dataset size per class and opting for more suitable network architecture. As part of the solution, we produced an automated machine learning toolchain that can be used to further improve the model.

## BENEFITS

The project demonstrated that machine learning (ML) can help with the classification of different beam shapes. Being able to distinguish between diverse set of shapes with great precision helps in making the right decision on the corrective actions that the tuning algorithms must apply to properly tune the microscope. Additionally, the ML toolchain is extensible and can be used for different experiments later in the future.

A portrait of a man with a shaved head and a goatee, smiling. He is wearing a dark blue and red plaid button-down shirt. The background is a blurred indoor setting with warm lighting.

"Pranav did a great job gathering a large database of data that could be used to train the machine learning models, applied the right techniques to train the model and showed the advantages of using machine learning for classification of the beam shapes on demo on a real microscope. Pranav also created a machine learning toolchain that can be easily used by others to extend data and labels to further enhance the beam classification model.

The project showed that machine learning helps with the classification of the different shapes of the beam, this can be used to make the right decision on the corrective actions that the tuning algorithms must apply to properly tune the microscope. There is still plenty of work to do, but this was for sure a good start!"

Andrea Pasqualini MSc  
Thermo Fisher Scientific



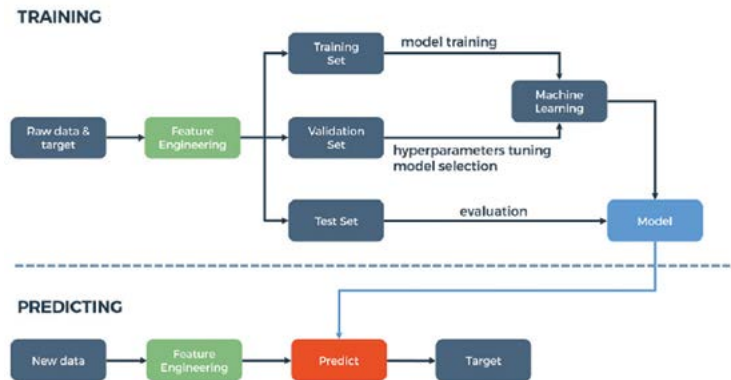
**PRANAV BHATNAGAR PDEng**

# Automatic Alignment of Microscope using Machine Learning

Thermo Fisher Scientific is a world leader in scientific instruments. In Eindhoven, the company designs and manufactures transmission-electron microscopes (TEM) used for a.o. analysis of materials with ultra-high resolution to sub-Ångström levels. Software plays a key role in controlling the microscope, creating a range of workflows and applications.

Aligning the Optics is essential for correctly shaping the beam and getting a good image quality. However, manually aligning a microscope is time consuming. Therefore, it is desirable to execute these procedures automatically. Furthermore, if possible, the alignment algorithms should run faster and produce better alignments compared to manual procedures.

The development of automated alignments can be approached from different angles. In this project, we aimed at using machine learning (ML) to create robust alignment algorithms. As the world moves further into the digital age, the use of data generated by machines is very important as it holds key insights into the behavior of the system. Similarly, data generated by Electron Microscopes (EM) can be used to further improve the precision and accuracy of these microscopes. This is what prompted the question for this research: Can ML help to make the auto alignment of the Microscope more robust and faster? If yes, how? These are the main questions that this project aimed to answer.





"The final report provides one specific architecture that abides by the vision of Vanderlande, with accompanied rationale. This result provides insights into our current way of working and provides feedback on our vision. An even greater contribution lies in the classification and clarification of the problem: Ganduulga provided insights on how the functional requirements lead to a particular model-driven design with accompanied design patterns. These new insights will be very valuable in the future to (re)classify our current vision and to derive the next steps and focus areas."

Dr.ir. Lennart Swartjes  
Vanderlande

## CHALLENGES

One of the main challenges in this project is to acquire the domain and the systems' knowledge of Vanderlande in a very short period of time. It is a challenge because the complexity and the diversity of the company systems are increasing due to the customer's market-specific requirements. The second challenge is a modeling technique choice that suits the project goal the best.

## RESULTS

To improve system development efficiency, reference architecture and reusable design of the modules are produced as the main results of the project. Besides, the results are accompanied by extended analysis, evaluation of design choices and decision rationales. Furthermore, simulation is facilitated to show the validity and feasibility of the proposed design and architecture.

## BENEFITS

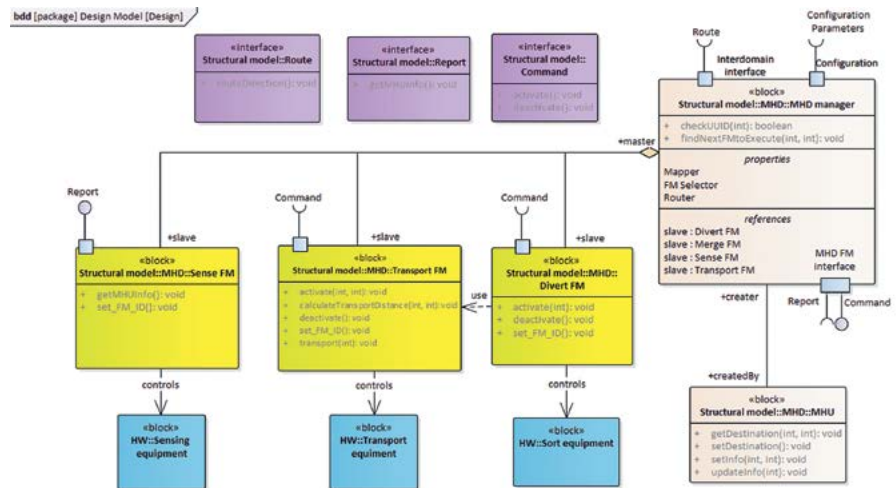
On one hand, the system developers save a great amount of time by using the reference architecture as a guideline for development and as a communication asset for knowledge transfer. On the other hand, various analyses to produce the architecture is fruitful to the company for realizing their future vision.

# Model-Based System Engineering Design of Functional Modules for Configurable Topology

Vanderlande is a global market leader in logistics automation systems. The company provides the systems in airport, warehouse, and parcel markets. Recently, the company has encountered a problem related to efficiency in development due to the diversity of the systems within the markets. To improve efficiency, the company has produced a modular system architecture.

This project specifies the system architecture further by addressing the most fundamental software modules that are directly responsible for moving physical items throughout the systems. As a result of the problem, domain and requirement analysis, a reference architecture is proposed using Model-Based System Engineering methodology. The proposed architecture and design models were developed by emphasizing modularity and (re)configurability for the reusability purpose of the modules.

We proposed a master-slave design pattern to organize the architecture of the software modules. For the flexibility of (re)configuration in terms of topology, a parameterized module mechanism is specified with support of the master component of the proposed design. Since modern industrial automation standards, namely IEC 61131-3 and IEC 61499, are compliant with the object-oriented concept, architectural specifications and designs are illustrated with the use of SysML diagrams.





## CHALLENGES

The challenge of this project was to understand the concept of correct driving behavior and translate it into a set of rules that can be interpreted by a computer. Additionally, it was challenging to design and develop a functional prototype system that can constantly monitor these rules while the car is driving in traffic. Finally, demonstrating the system in a real traffic situation was a very demanding task.

## RESULTS

The result of this project is a system called VESSEL, which retrieves driving data from vehicles deployed in traffic, checks whether they comply with the defined notion of correct driving behavior, and stores the data in a persistent storage to be used for offline analysis. In addition, the design of the system offers modifiability, allowing the maintainers to easily extend the set of rules within the system.

## BENEFITS

The prototyped system can be deployed in vehicles to monitor their driving behavior. With such a system we can ensure that both human-driven and autonomous vehicles comply with the same notion of correct driving behavior. Additionally, TNO researchers can use the driving data stored by the VESSEL System to further explore the notion of correct driving behavior and extend the set of rules deployed inside the system.



"Konstantinos embraced with grace and professionalism the challenging R&D task of scoping and demonstrating the potential of the VESSEL System to monitor the behavior of autonomous driving vehicles. In just 10 months he was able to engage all the relevant stakeholders and technology experts at TNO, learn from them, and develop and deploy a prototype of the VESSEL system in our Carlabs facilities. We are very pleased with his results, which we will leverage, together with the knowledge he gathered during his project, to develop the first generation of VESSEL systems."

Dr. Arturo Tejada Ruiz  
Senior Scientist, TNO



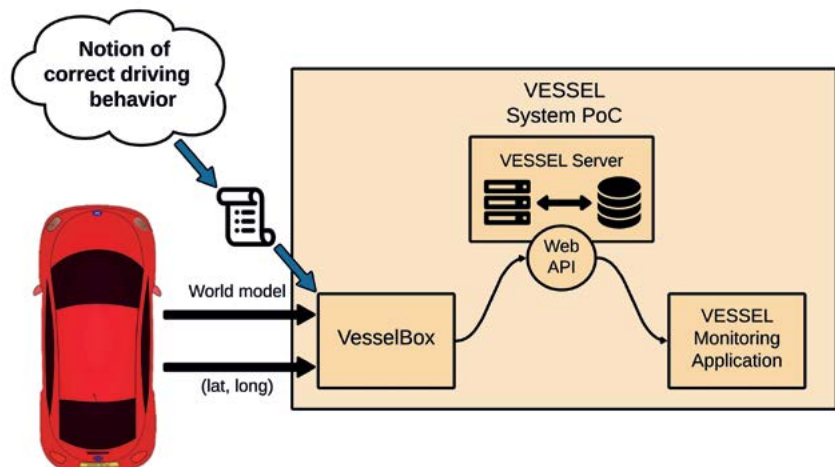
**KONSTANTINOS KARMAS PDEng**

# VESSEL - Driving Behavior Monitoring System

According to the World Health Organization, road traffic injuries are one of the leading causes of death in the modern era. The primary factors that lead to this issue are traced in faulty driving habits of humans, such as speeding, reckless driving, and distracted driving. The development of self-driving cars makes this problem even more complicated, considering that autonomous vehicles must exhibit human-like driving behavior to be harmonically integrated into public roads. In addition, the dependency of self-driving vehicles on their software raises the concern of whether their behavior will change unexpectedly with every new software update.

In response to that, TNO envisioned the Vehicular Safety and Security Evaluation and Logging (VESSEL) System. The purpose of the VESSEL System is to monitor the driving behavior of humans and autonomous vehicles to ensure that they comply with the same notion of correct driving behavior.

In this project, a functional Proof of Concept (POC) of the VESSEL System was designed and developed. The VESSEL System PoC consists of three subsystems: the VesselBox, which uses the vehicle's data to check its driving behavior, the VESSEL Server, which gathers up the driving data from vehicles around the world, and the VESSEL Monitoring Application, which can be used to remotely monitor the driving behavior of a vehicle. The implemented prototype demonstrates the value of the VESSEL System and sets the groundwork for TNO's StreetLive program, which aims to achieve safer accident-free roads.





"Gilda accepted the challenge to try to link both worlds. She managed to create a product she calls GrayDep, which generates enough output to do real-time code checking, while the dependency definition is language independent and allows multi-level & team input. The function architects can refine the architecture per domain, but only within the boundary of the reference architecture. I appreciate how all the individual ideas have been integrated into a product that shows flexibility beyond the individual features."

Mark Jaminon  
ASML

## CHALLENGES

The challenge was to model the reference architecture of the system and to link it to the code live checking. Reviewing the code was the approach that was used beforehand in ASML.

## RESULTS

The result of this project is a software tool named GrayDep. GrayDep consists of two major functionalities as follows:

- By reading the reference architecture that is written in a parsable DSL, GrayDep generates enough output for the YieldStar live dependency checker tool to enforce the logical dependency rules in the software.
- By mapping the reference architecture to source code, GrayDep analyzes the code and finds the real dependencies between the modules and interfaces. Furthermore, GrayDep monitors the change of dependencies over time.

## BENEFITS

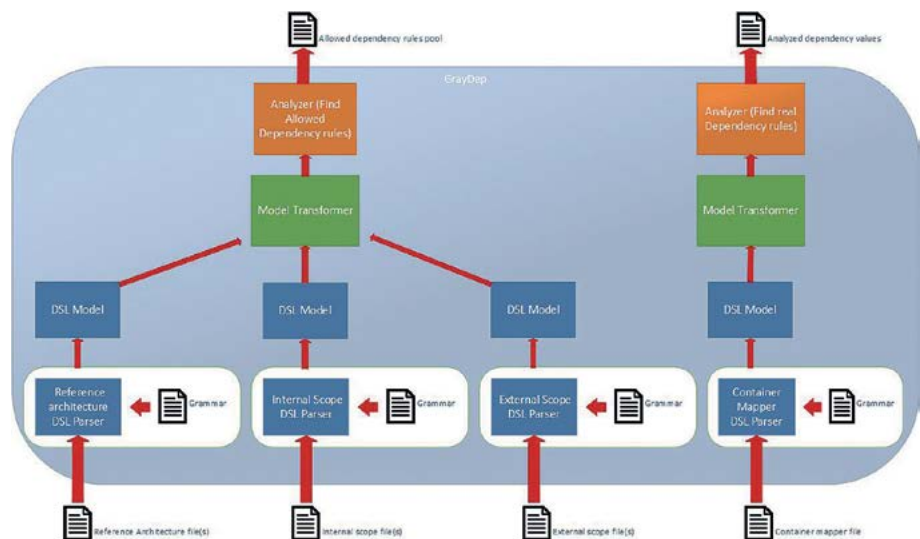
To ensure that code follows the reference architecture, architects and engineers spend much time. Making the process automated, by modeling the reference architecture and linking the reference architecture to the code live, enables ASML's architects to achieve more by using less effort. Moreover, GrayDep can be used in a team for flexibly evolving the architecture.

**GILDA KHOSRAVI PDEng**

# Dependency Management and Scope Control

As software grows in size, complexity increases. As a result, maintenance and change management become more complicated. One of the main reasons for this complication is an increasing number of dependencies, either internally between different parts of the software or externally with third-party software.

More modular architecture with desired dependencies among modules that preserve the non-functional requirements of the system, which is called reference architecture in ASML, can speed up the build and maintenance process of the software. Modular design is about breaking down a design into smaller parts (modules), creating them independently, and combining them into a more extensive system by defining the policies for their interconnections.



The main question is how to define modules and their dependency rules flexibly as well as to couple the design and code live in an automated fashion. The goal of the “Dependency Management and Scope Control” (DM&SC) project was the answer to this question by creating a tool for defining modules and the policies that have to be followed when one module is used by another one. The rules governing module dependencies preserve the principles of the modularity and ensure that dependencies in software comply with the allowed dependencies in the design, as defined in the reference architecture.

The designed and implemented tool, which realizes the solution, was accomplished within ten months in C# by creating a domain-specific language (DSL) using ANTLR parser.

## CHALLENGES

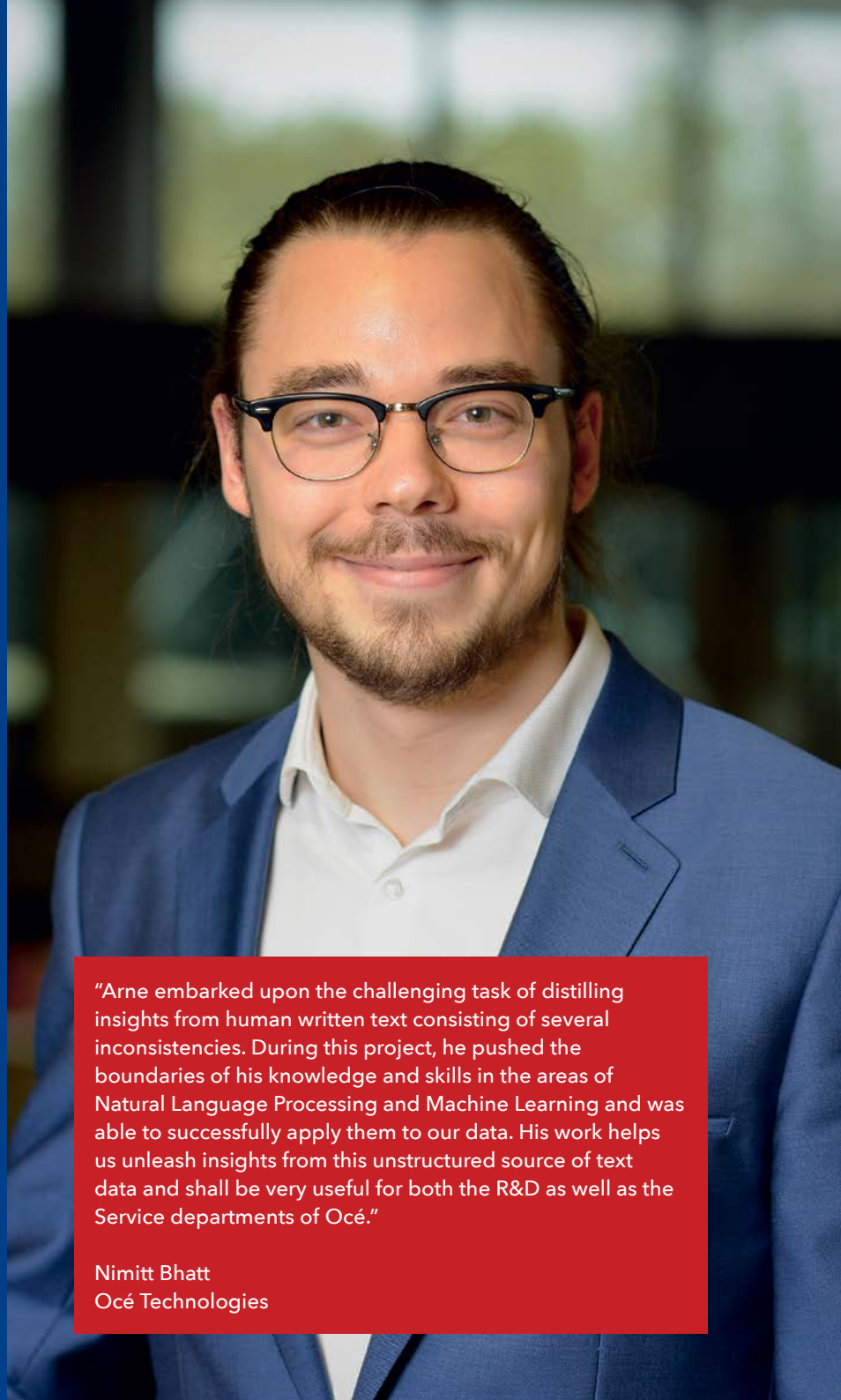
Field technicians concentrate on recording essential information in the service logbooks. They don't pay attention to following grammar rules. Spelling mistakes, disregard for sentence structure, and the heavy use of domain-specific terminology, makes automatically processing the text challenging. Another challenge is that different technicians write about similar concepts using different terms, introducing noise to the data.

## RESULTS

Cleaning and analyzing the text using the techniques of Natural Language Processing has given Océ an overview of how technicians fill in the logbooks. A proof of concept that links problem categories to their typical solutions results in insights into how technicians approach problems at an aggregate level.

## BENEFITS

The benefits of this project are two-fold. In the short term, the results help Océ improve the data gathering process from the field technicians and gain insights into the efficiency of the maintenance visits. In the long term, the results of the project can serve as a basis for a data-driven advisory tool that would complement the current service manuals offered to the field technicians.

A portrait of Arne Bhatt, a man with long hair tied back, wearing glasses and a blue blazer over a white shirt. He is smiling slightly and looking towards the camera. The background is a blurred indoor setting.

"Arne embarked upon the challenging task of distilling insights from human written text consisting of several inconsistencies. During this project, he pushed the boundaries of his knowledge and skills in the areas of Natural Language Processing and Machine Learning and was able to successfully apply them to our data. His work helps us unleash insights from this unstructured source of text data and shall be very useful for both the R&D as well as the Service departments of Océ."

Nimitt Bhatt  
Océ Technologies



**ARNE LAPÖNIN PDEng**

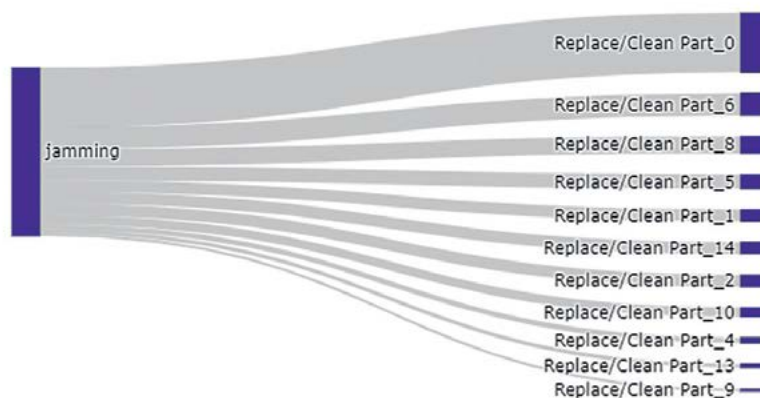
# Distilling Insights from the Text Written by Field Technicians

Océ, a Canon company, develops, manufactures, and sells high-tech printing devices. Océ supports its customers during the maintenance phase of their products' life-cycle. The company wants to move towards a more data-driven maintenance approach that allows Océ to anticipate and predict when the machines should be maintained.

One of the potentially useful data sources for data-driven maintenance is the text found in the service logbooks. Field technicians fill in these logbooks when they make service visits. In their native language, they write about the problems they observed and actions they took to resolve the problems. Currently, this data is sporadically analyzed by hand. Insights extracted from that data can be used by the Service department to improve their maintenance strategies and the R&D department to improve the quality of the products.

Techniques from the field of Natural Language Processing (NLP) and Machine Learning (ML) were used to structure problems and solutions written in the logbooks. Even though existing NLP tools support major languages, such as English, Spanish, and German; human input is still needed for analyzing, cleaning, and preparing the data before ML algorithms can be used on it. The results show that the text in the logbooks can be used as a source of insights into populations of machines.

Sankey Diagram connecting problems to solutions





"As supervisors it was very rewarding to see the personal growth of Dũng during the project. We saw that Dũng has great eagerness to understand new domains and learn new techniques in order to develop himself. Dũng is eager to work with stakeholders to come to a joint understanding of requirements. He applied both Behavior Driven Development and UI mock-ups to reach this goal, which will be a valuable experience in future projects. We are sure future project members will highly value Dũng's willingness to cooperate and his original thinking in designing solutions. We would like to wish Dũng all the best in the future."

Herman Roelfs  
Pieter Verduin  
Océ Technologies

## CHALLENGES

The first challenge is how to design a data quality dashboard that not only shows the quality quickly but also supports the root cause analysis of the data quality issues. Designing a dashboard that can show different data quality in different data projects is the second challenge. The execution challenges are prioritizing, scheduling and scoping the project to create the most valuable output and communicating the solution to the stakeholders.

## RESULTS

The proposed Data Quality Dashboard solution includes: a high-level design of a framework for Data Quality Dashboard, a model of data quality that supports causality analysis, and a process to use with this framework. A detailed design of a model processing pipeline demonstrates how this data quality model can be processed, visualized in the configurable Dashboard, and then analyzed for the data quality issue's root causes in the Analysis view.

## BENEFITS

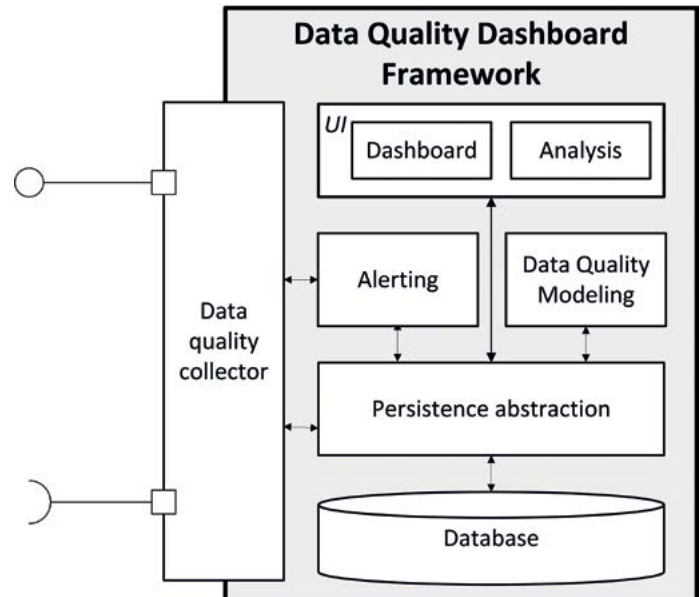
The proposed model can support various types of quality measurements from different contexts. The flexible data quality models can be designed for different systems. The mental model data experts use to find the root causes of the data quality issues can be extracted, stored, and reused. The framework design can be used to develop a Dashboard solution. Using the proposed model and framework, the data quality can be designed upfront.

**DŨNG LÊ PDEng**

# Data Quality Dashboard

Controlling the data quality is one of the main focuses of the data projects in Océ. One approach to support this goal is to have a data quality dashboard (DQD) which shows the overview of the data quality so that the operators can see the quality quickly. Moreover, the dashboard should support analyzing the causes of the data quality issues in order to figure out the actions needed to fix these issues.

This project proposes a solution that has a framework to realize the aforementioned DQD. This framework has a Dashboard that shows overview of data quality, an Analysis view that supports root cause analysis, a Data Quality Collector that collects the information of data quality, an Alerting component that automates the alerting tasks, and a component that is responsible for the data quality modelling tasks. A model of the data quality is also designed to capture the essence of the data quality domain knowledge. This model supports a causality model that was inspired by the practical knowledge of the data project Managed Print Services. The processing of this model from database to user interface is illustrated in a detailed design of a pipeline. This design comes with a proof of concept prototype using Dash framework, Python, and Neo4j database. The last piece of the solution is a process for designing the data quality upfront. The process, the model, and the framework are the foundation for the Data Quality Dashboard that can be used in different data projects in Océ.



## CHALLENGES


The exponential growth of test cases is a universal problem for (multiple levels of testing at) ASML, as well as for other companies dealing with a complex domain. The goal was to come up with a solution that is universally applicable to these domains. This was a challenge.

## RESULTS

The outcome of this project is a framework using which the users can define their own criteria for measuring test coverage, and automatically make test selection based on these criteria. The time required to do the test selection scales linearly with the number of tests and coverage criteria. This makes it a future-proof solution that is applicable to a wide variety of domains.

## BENEFITS

In the proposed solution, test selection is made automatically. The benefits of this solution compared to manual test selection are twofold. Firstly, it does not rely on a tester's experience and domain knowledge but instead is based on a uniform, formalized definition of coverage that can be reviewed and agreed upon a priori by internal and external stakeholders. Secondly, it is scalable: the test selection currently can work with 100,000 tests, and this limit can be increased further.

A portrait of a woman with long dark hair, wearing a dark blue patterned button-down shirt. She is looking slightly to the right of the camera with a neutral expression. The background is blurred, showing what appears to be an office or industrial setting with warm lighting.

"Trupti was challenged with the daunting task of increasing test efficiency at ASML. She excelled in making this broad problem concrete. Her solution - a framework in which a user can define her own criteria for measuring test coverage, and do test selection based on these criteria - resonates on different test levels, and will therefore undoubtedly be developed further within ASML".

Dr. Rick Smetsers  
ASML

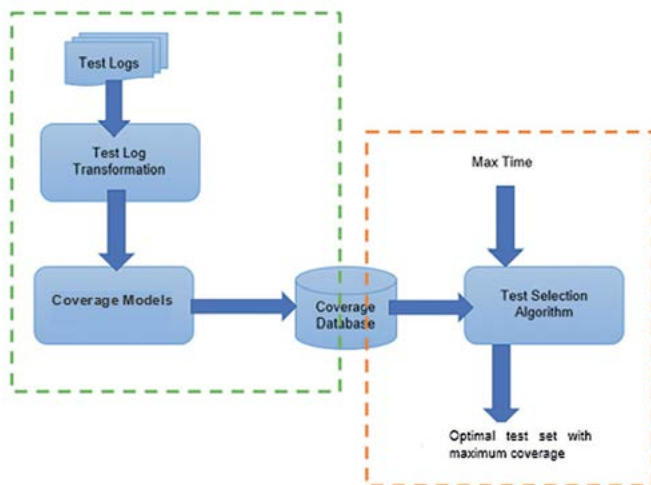


TRUPTI MANIK PDEng

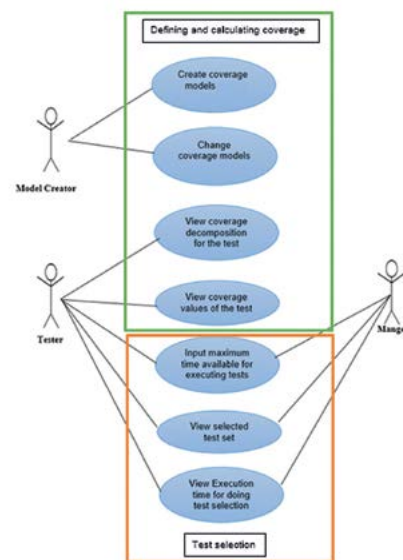
# Increasing Coverage Through Coverage-based Test Selection

ASML manufactures complex systems that are needed for the chip-making process. One of the consequences of the complex system is that the number of tests is growing exponentially. However, the resources to execute and maintain these tests are not growing exponentially. Moreover, the tests might contain a lot of duplication. Hence, it is not feasible and beneficial to execute all the tests at all times. Currently, at ASML, the tests to execute are selected manually. Manual test selection becomes non-scalable and inaccurate when the number of tests grows. Hence, this requires a solution that can enable efficient test selection.

In this project, a framework was developed that can be used to automatically select the tests to run within a given amount of test time. Tests are selected based on their coverage. The framework allows the user to define the notion of coverage that is used to select the tests.



Legends: 1. Defining and calculating coverage: - - -  
2. Test Selection: - - -



It is beneficial for ASML to do test selection using the proposed framework (as opposed to doing it manually) because it is based on a uniform, formalized definition of coverage (as opposed to a tester's experience and domain knowledge). The benefit of having a uniform and formalized definition of coverage is that it can be reviewed and agreed upon a priori by internal and external stakeholders. Previously this was not the case.



"Ani and her prototype guides Metrology in a faster transition towards independence from MDL in favor of the new format and infrastructure for diagnostics data. To support her investigation, Ani created a tool that reconstructs data context. A tool that is now used by data mining engineers."

Ir. Patrick Peeters PDEng  
ASML

## CHALLENGES

This project's main challenge was to understand the structure of the data which is logged in MDL files as well as the formal reports of ASML. Then proposing meaningful grouping of MDL content, as independent as possible. It was needed to express the confidence level for how much the containing data in any group which was corresponding to an XML-based report is correct.

## RESULTS

The tangible results of this project are firstly the procedure of transferring the context and meta information of an MDL log file to an XML-based hierarchical format. Secondly, the decomposition of the MDL data into suitable and meaningful XML-based reports. And finally, the method for converting an MDL file to the XML-based format.

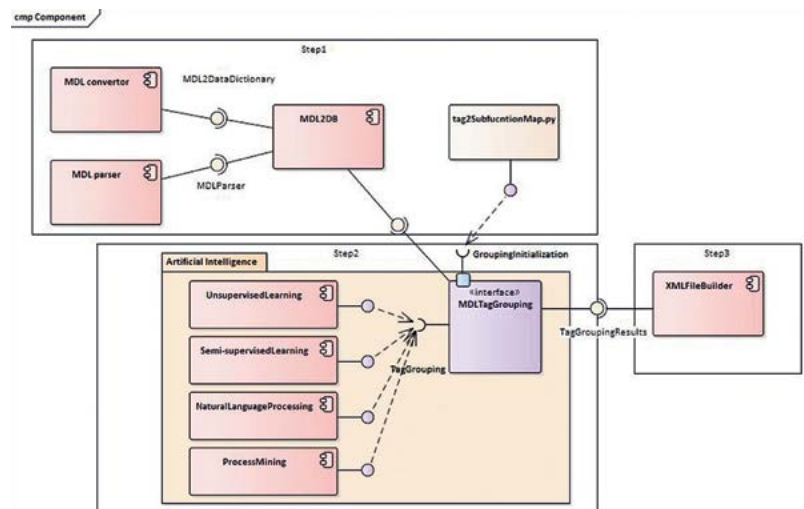
## BENEFITS

The contribution of this project is to transfer the MDL to a meaningful XML-based format. This helps to reduce the human effort spent on the manual task of XML-based report generation; in other words, in less costly manner. The developed tool is recommended to be helpful as a support tool for faster production of XML-based reports for each domain in the Metrology Department.

**ANI MEGERDOUMIAN PDEng**

# From Informal TWINSCAN Machine Diagnostics Data to Reliable Formal Customer Interface

As the complexity of any system grows, the need for diagnostics becomes essential. In this sense, the data produced as an input for any complex system, like the high-tech machine of ASML is critical. During the wafer production cycle in TWINSCAN, the measurement, modeling, and applied corrections are logged to a diagnostic file called MDL (Machine Diagnostic Log). MDL contains essential data that helps designers (as well as support engineers) to understand the behavior of the machine. Over the years, customers have started using informal MDL data as well. By design, MDL does not protect itself from incompatible changes.



This project is initiated to analyze the possibility of converting all needed data to official XML-based files at low cost. As a roadmap, the current intention is to provide formal data to the customers. This, however, comes at a high cost if conversion is done manually.

As a result of this project, a support tool, with mappings of the log file content to XML-based file is delivered. Besides, an iterable pipeline and corresponding prototype for producing the grouping of the tags is delivered based on application of artificial intelligence. Finally, a prototype which provides an XML-based report for aiding the human resources in designing the XML-based formal reports is submitted.



## CHALLENGES

The main challenge of this project is to decrease manual modeling work of the TNO IVS department by providing the automated solution for conceptual modeling.

## RESULTS

In this project, we designed and developed a software tool that provides a systematic process for conceptual modeling to improve the consistency of the resulting models and the traceability of the modeling process. Further, the tool separates conceptual modeling from identifying domain concepts. This separation allows the tool to accept various ISO standards. Since the tool can be applied to different domains it has been named Multi-domain Modeling Assistance (MMA) tool.

## BENEFITS

Added-value of this project is the following pair usages of the MMA tool. First, the tool contributes to centralizing models by coupling with in-house tooling, and the resulting models of the MMA tool can be used for TNO projects and products. Second, the tool can be useful for different ISO standards, different norms, and numerous European Union (EU) project requirements.



"Munkhdalai worked on a challenging project: developing, for the first time, a support system for developing conceptual model. Along the way he helped us understand what the requirements of such system should have been in the first place. We are happy with Munkhdalai's contributions and wish him the best in his life and career."

Arash Khabbaz Saberi  
TNO Automotive



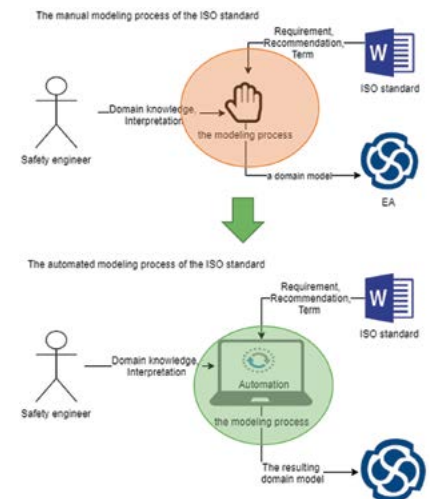
**MUNKHDALAI SAINBILEG PDEng**

# Multi-domain Modeling Assistance (MMA) Tool: Design and Implementation

Compliance with norms and standards is a vital part of quality assurance in the automotive domain. Compliance assurance is becoming challenging due to the rapid increase of system complexity. The primary reason for the growth is new features such as automated driving and smart mobility. These features enhance the boundary of the industrial standards continually thanks to the ever-changing standards such as ISO 26262:2018 on functional safety, ISO/SAE CD 21434: 2019 on cybersecurity, and ISO/PAS 21448 on the safety of the intended functionality.

The Integrated Vehicle Safety (IVS) department of TNO is developing an innovative tool for compliance assurance. This tool ensures that the quality of the department product meets the requirements of the ISO standard, and it detects design errors in the early stage of the system development process using conceptual models, which come from the ISO standard requirement. The essential part of the tool is the creation of the conceptual models.

Conceptual modeling is a labor-intensive task for the modeler (end-user) because of a large amount of manual work in the department. The goal of this project is to reduce the end user's manual work by supporting a software solution. In the scope of this project, we addressed two aspects of the manual work. The first aspect is the consistency of the resulting models, and it highly depended on the end user's carefulness. There was room for mistakes during the modeling process because of less-structured and low-controlled manual work. The mistakes led to inconsistent models. Therefore, a systematic modeling process is needed to support the end-user to increase the consistency of the resulting models. The second aspect is the traceability of modeling. It had several steps to produce the resulting models from the ISO standard requirements. The end-user needed to be aware of the ISO standard requirement specifications associated with the resulting models. The user manually maintains the association between the requirement specifications and the resulting models to trace the manual modeling process. This is a time-consuming task for the end-user. The manual traceability is not always sufficient for further usage. Hence, traceability can be automated.





"The future starts now with collecting data for others to feed into systems. Dimas Satria successfully researched the existing void in storage and structuring data and created a novel approach to help farmers and researchers share data. With the little time he had, the proof of concept is a working demo of the open-source deployment package for any future service provider on the planet interested in helping farmers forward."

Paul van Zoggel  
TU/e and the Precision Agriculture Trial Center Reusel

## CHALLENGES

The greatest challenges for this project were the exploration of the domain and data model standards for agriculture. An additional challenge was to ensure that the system has sufficient metadata for researchers. Finally, the process of getting access permission to farm clouds and harvest sensor data directly from the servers owned by sensor companies was time-consuming and complicated.

## RESULTS

This project investigated the possibility for farmers in storing and structuring precision agriculture data with different data schemas easily. A software system was designed and realized so that farmers can import their data to the system manually, harvest data from sensors placed at fields, or retrieve farm data from their cloud servers directly. Researchers may access the data with its metadata for performing analysis.

## BENEFITS

The developed system enables farmers to integrate data from various sources easily without vendor locked-in. They can store data from many sensor providers without asking permission to the system maintainer. In the system, researchers can search and find data from many farmers for data analysis.

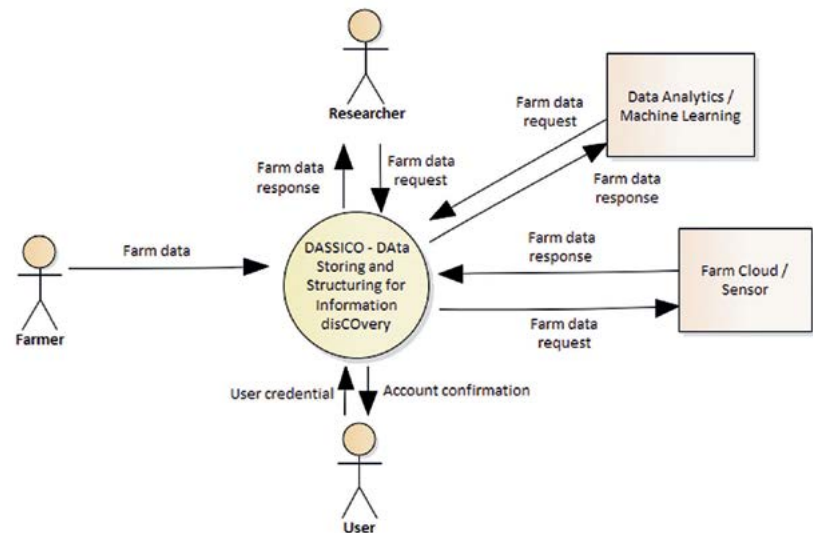
**DIMAS SATRIA PDEng**

# Data Storing and Structuring for Information Discovery in AgTech

Artificial Intelligence in food resilience begins with manually and automatic collecting data on the growth of plants and resources needed. Continuous and structured data is vital for discovering breakthroughs in sustainable food production. Independent and neutral data storage by a farmer is the first step to train future systems. For information discovery, researchers need to know from which fields around the world data is gathered and made available for scientific research. Researchers and third-party agronomical consulting systems can analyze the big data to discover useful knowledge for farmers to treat the crops and fields properly.

We aimed to design and implement a software system to integrate observation and measurement data from various sources with diverse formats. The system structures the data for researchers or analytics tools to find the required data efficiently for performing analysis. The system needs to ensure that researchers have sufficient metadata to understand the stored data easily.

To achieve that, a metadata-based Extract-Transform-Load (ETL) was designed for the system. With this design, farmers only need to provide farm data and its metadata to the system. Then, the system stores the data at the proper tables in the database based on the metadata. The system design allows the extension and deployment of new software components independently for software evolution with minimum impact on other existing components.



## CHALLENGES

The Intelligent Microscope is a research project, whose goal is to explore the future of electron microscopy software. Therefore, the two biggest challenges consist of identifying promising directions to follow and incorporating them into the existing design by refactoring the code base as necessary. Considering the associated technological stack, expertise in the fields of machine learning, web development, and software architecture is required in order to make sound and well-founded design decisions.


## RESULTS

The results of this project are three-fold:

1. An automated particle detection workflow that relies on convolutional neural networks in order to improve traditional solutions in terms of throughput.
2. A unified voice interface capable of controlling multiple microscopes of different types using high-level commands.
3. A smart image processing technique based on a deep learning architecture that can be retrained at will in order to optimize its performance according to the current operational settings of the microscope.

## BENEFITS

The Intelligent Microscope project highlights the benefits of including state-of-the-art technologies in microscopy systems. More specifically, it showcases how Artificial Intelligence (AI) can be used to solve formerly infeasible tasks, such as object detection and speech recognition, as well as how existing workflows can be improved with the inclusion of AI components. Finally, the system can be used both as a testbed for prototyping as well as a demonstrator of interesting thought experiments.

A portrait of Dr. Remco Schoenmakers, a man with dark hair and a beard, wearing a blue checkered shirt. He is smiling slightly and looking towards the camera. The background is a blurred indoor setting.

“Konstantinos had the challenge to quickly become acquainted with a whole technology stack, get familiar with the code and the architecture of the application, understand the domain, stakeholders and their challenges, and then also think of ways to add and improve! Since Konstantinos had to significantly expand the capabilities of the software, a lot of “plumbing” needed to be done. He successfully redesigned several parts of the code to be better prepared for the expansions needed and applied sound design patterns to those areas. Next to that, he broadened the AI and control part of the microscope.

All in all, Konstantinos created a super-cool piece of software that we will have a lot of fun playing with and learning from!”

Dr. Remco Schoenmakers  
Director Digital Science Technologies  
Thermo Fisher Scientific



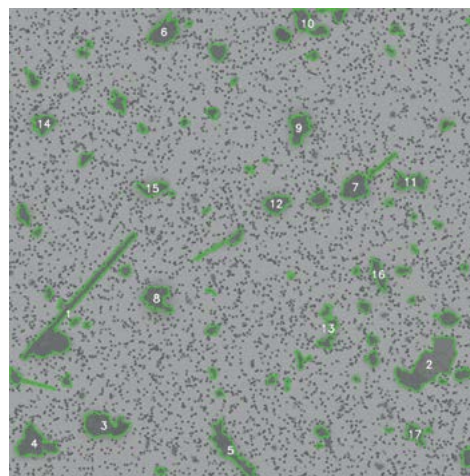
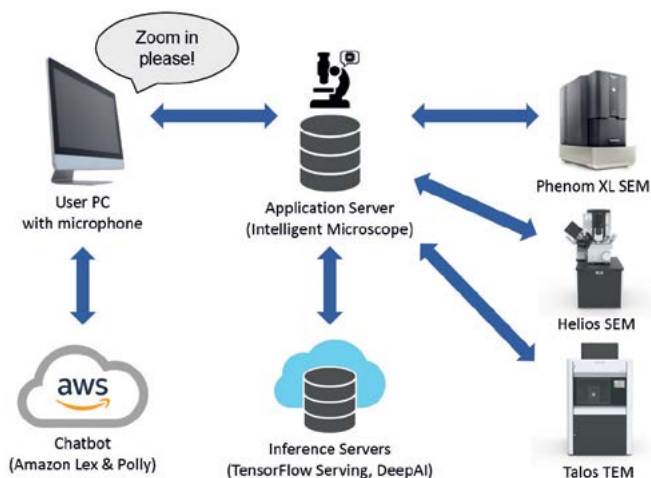
**KONSTANTINOS SMANIS PDEng**

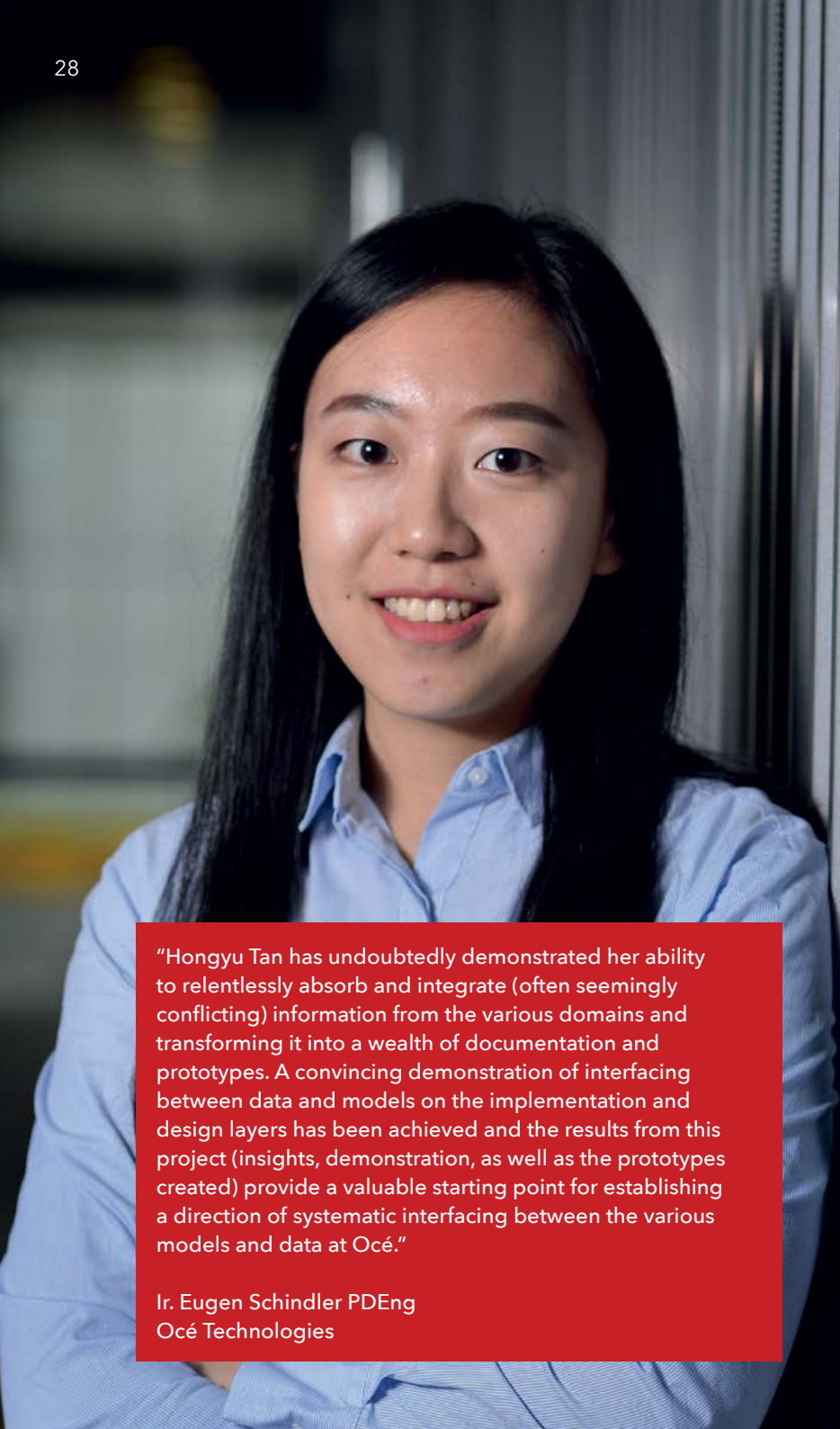
# Intelligent Microscope II

State-of-the-art technologies affect our everyday lives in ways previously only imagined. Self-driving cars, smooth-talking virtual assistants, and computer-aided healthcare diagnostics are just a few examples of already existing technologies that seem to have emerged out of science fiction tales. The driving force behind all these exciting applications is Artificial Intelligence (AI) and, in particular, its branches, Machine and Deep Learning. Of course, these developments also influence the way software is designed, as machine learning frameworks are slowly becoming standard components of an increasing number of software systems.

Considering these trends, Thermo Fisher Scientific is conducting research into the application of such cutting-edge technologies in its products and services. More specifically, the company's Advanced Technology department in Eindhoven is interested in exploring in what way these technologies will influence the way people use electron microscopes in the future and incorporate them into a prototype software system.

As a result, the Intelligent Microscope (IM) system was created in 2018. The IM is an AI-powered, web-based software system that can be controlled through a voice interface, perform microscope operations, detect cells and mitochondria in microscope images, as well as interpret information about the detected objects. Intelligent Microscope II constitutes the second development iteration of the IM project, during which the system was significantly extended in order to detect certain particles, control multiple microscopes of different types using high-level voice commands, as well as perform smart image processing tasks.





"Hongyu Tan has undoubtedly demonstrated her ability to relentlessly absorb and integrate (often seemingly conflicting) information from the various domains and transforming it into a wealth of documentation and prototypes. A convincing demonstration of interfacing between data and models on the implementation and design layers has been achieved and the results from this project (insights, demonstration, as well as the prototypes created) provide a valuable starting point for establishing a direction of systematic interfacing between the various models and data at Océ."

Ir. Eugen Schindler PDEng  
Océ Technologies

## CHALLENGES

One of the challenges in this project was to work in a multidisciplinary context, integrating different domains such as data-science, model-based development, and software engineering. Another challenge was to be a product owner for a team of 14 Bachelor students from TU/e, supervising them to develop a 3D visualization tool for this project.

## RESULTS

As a result, this project has established proof-of-concept for interfacing between data and models through a number of use cases in the media-handling domain. It shows the possibility of model verification and validation using the real data collected from the machines. Also, through these use cases, two connection patterns have been identified, showing the relationship between the modeling world and the data world.

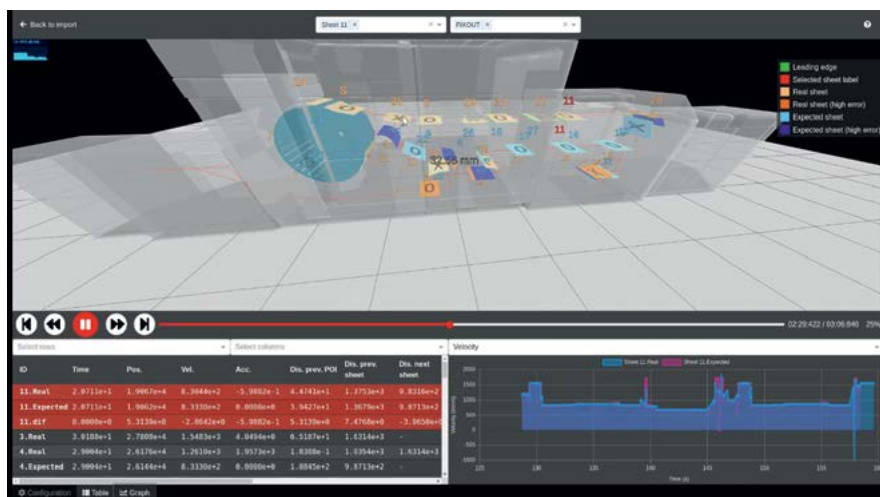
## BENEFITS

This project provides extra tools for Océ to investigate the difference between the modeling results and actual measurements from the machines, for example Exhibit, a web-based 3D visualization tool. This tool can be used by not only internal employees for helping them to analyze the root-cause of transporting errors, but also external customers for having a more intuitive way of understanding the products developed by Océ.

HONGYU TAN PDEng

# Integration of Modeling and Data Science

Océ develops products and services for markets in printing for professionals and digital document management. One of the challenges Océ faced is that the design of media-handling should avoid the occurrence of sheet collision so that the printing system can achieve high productivity. On the other side, Océ collects a lot of data from the machines but is not fully leveraging them yet. This data is valuable for Océ as it reflects the real behavior of the printing system.



Based on these challenges, Océ initiated this project with the idea of exploiting the machine data for improving the design of media-handling in printing systems. To be specific, the project goal is to check the validity of media-handling design models by comparison with real measurements through visualization and identify the possible causes of unexpected sheet behavior such as sheet collision.

This project selects three use cases, each of them related to one domain-specific model. These models are used in the media handling development of VPi300. A data processing and/or visualization pipeline was designed and implemented for each of the use case so that Océ engineers can get more intuitive insights about the difference between the design and the actual measurements on the media-handling.

## CHALLENGES

The challenge in this project was the fact that knowledge of the overall system was found in pieces among different individuals and teams across different locations. Being a multidisciplinary domain consisting of multiple devices and dedicated experts, this creates a challenge of putting everything together in a way understandable by everyone and capturing the most important aspects.

## RESULTS

The result consisted of process models that show the integrated model of the Sample Flow as well as detailed sub processes at device level. Furthermore, simulation artifacts were produced that represented different use cases. These simulations were designed in a configurable way so that further analysis can be made for different machines, process options or customer scenarios. In addition a deployable model of the system was developed and a proof of concept was provided to the client that integrated the microscope's desktop based metrology service through a command line interface, a RESTful alignment webservice through http and human tasks through form based inputs, using Camunda workflow engine.

## BENEFITS

Proper modeling and analysis through simulation provides a quantifiable way of suggesting workflow improvements that help optimize the customer's processes. Furthermore, by enabling the deployment of custom process models for different customers, ThermoFisher Scientific can benefit from a better understanding of microscope usage in the real environment and accelerate innovation by using such insights as feedback in the design of new hardware and software products and features. In the long term, the large amount of data that can be gathered from these deployed processes can also be used as input for different data driven approaches for usage pattern recognition, decision making, machine learning and process automation.



"Beza created the model that was not only correct and realistic, but also enabled us to simulate the different system configurations. That allowed us to find the optimal system configuration maximizing the throughput and minimizing the time required for getting the customer data. Beza accomplished another important goal, namely proved usability of the model for rapid system prototyping. She integrated the model execution engine with the microscope software and proved its usability. The concept is very promising."

Marcin Gramza  
Thermo Fisher Scientific



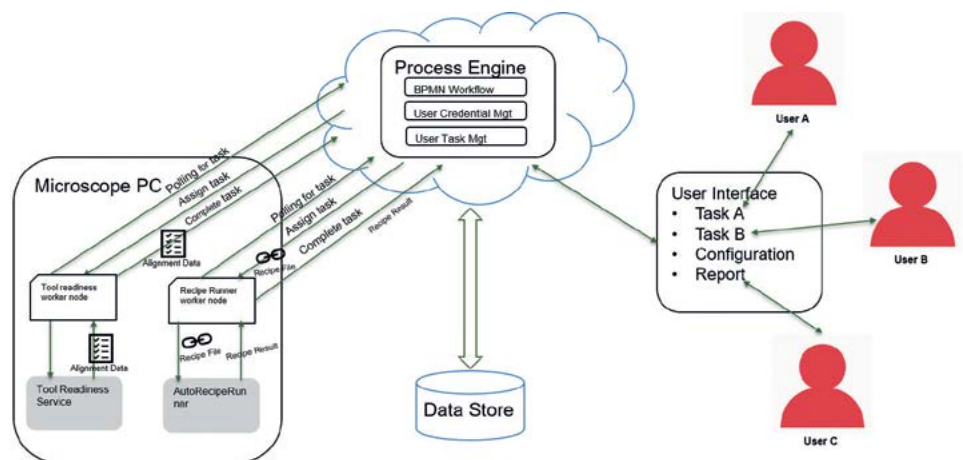
**BEZA GETACHEW TASSEW PDEng**

# Sample Flow Modeling for Rapid Performance and Failure Analysis

Thermofisher Scientific's Materials and Structural Analysis Division (MSD) Group at Eindhoven produces High End Transmission Electron Microscopes for Life Sciences, Material Sciences and Semiconductor Industries. This project deals with modeling of the sample flow in the context of semiconductor manufacturing process analysis. Sample flow refers to the process of taking a semiconductor wafer, taking thin slices of it called lamella and then measuring certain characteristics that are used for process control, yield estimation or defect analysis.

The goal of this project was to model the Sample Flow in order to analyze key performance metrics such as throughput (number of samples processed per second) and time-to-data (time it takes from a wafer being submitted for analysis until analysis or metrology data is ready) both in current systems as well as in future designs of hardware and software components.

As part of this project, process models were created that allow the documentation of current as well as desired process variants using BPMN. Subsequently these models were simulated using BPSim to provide a quantitative analysis that shows the relationship between various parameters and the overall performance of the overall system. And finally, a prototype is built that demonstrates the deployment of such process models in a real environment by integrating two interfaces of the Metrios TEM.





"The outcome of Wahyu's project allows us to develop future generations of wafer stages for our scanners meeting the high-quality standards our customers expect from us. It even has the potential to become the reference design for other scanning subsystems which have the same needs."

Ir. M. Hendriks  
ASML

## CHALLENGES

The challenge in this project was to reverse engineer a software component with outdated documentation. The software component itself is a complex component, which holds too many responsibilities and has a big size.

## RESULTS

The result was a new design of the software component in object-oriented style, which is less complex than the original design. This design was created using popular object-oriented design principles and proven design patterns. A prototype implementation was provided as a proof of concept. The prototype shows that the functionality and performance requirements are fulfilled while complexity is reduced.

## BENEFITS

The new object-oriented design provides a design that is less complex compared to the original design. With the complexity reduced, the cost to maintain the software component becomes lower. Furthermore, the proposed design opens the possibility for future improvement in which the design can be reused in other subsystems with the same needs.

WAHYU UTOMO PDEng

# Design Complexity Reduction Using Object-oriented Paradigm

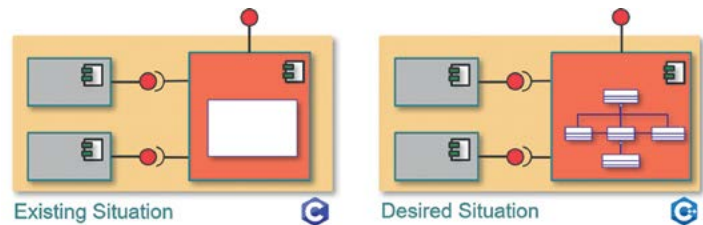
## EXEMPLIFIED BY REDESIGNING THE SCANNING APPLICATION OF ASML'S WAFER POSITIONING MODULE

ASML is the world's leading supplier of photolithography systems for the semiconductor industry. Inside the ASML machines, there are two Wafer Positioning modules, which transports silicon wafers to predefined locations (e.g., measure and expose position). This module is controlled by a set of software components with the Scanning Application being one of them.

Over the years, the Scanning Application has grown significantly bigger as a result of incorporating new requirements. As the software component gets bigger, it accumulates complexity making the component hard to analyze, to test, and to modify. Consequently, maintaining the Scanning Application becomes expensive.

A new object-oriented design of the Scanning Application is proposed. This design is a modular design, which distributes the responsibilities of the Scanning Application into smaller classes. This arrangement helps developers understand the component better. The new design prevents modification of a class to ripple to other classes as the coupling between classes is kept low. Also, dividing the component into small chunks makes the unit test applicable to this component. For a proof of concept, a prototype implementation of this design in the C++ language was provided. This implementation was tested on the testing environment that involves real hardware.

The design and implementation show that the proposed design of the Scanning Application manages to reduce the complexity of the component while maintaining the functionalities and performance. As a result, the cost of maintaining this component becomes lower.







tue.nl/softwaretechnology

Follow us:

 [linkedin.com/company/tue\\_pdengst](https://www.linkedin.com/company/tue_pdengst)

 [facebook.com/TUePDEngST](https://www.facebook.com/TUePDEngST)

 [twitter.com/tue\\_pdengst](https://twitter.com/tue_pdengst)

## CREDITS

**Edited by:** Dr. Y. Dajsuren PDEng and M.L.L. van Oorschot,  
Software Technology Programme, Eindhoven University of  
Technology

**Text:** Graduates of generation 2017 of the PDEng  
Software Technology Programme

**Photography:** Rien Meulman

**Production:** Communication Expertise Centre,  
Eindhoven University of Technology

**Design:** Grefo Prepress

**Printing:** De Digitale Drukker



**Giovanni de Almeida Calheiros PDEng;** Designing an Infrastructure for On-premise Software Deployment - A Case Study for Thermo Fisher Data Management Platform ■ **Pranav Bhatnagar PDEng;** Automatic Alignment of Microscope using Machine Learning ■ **Ganduulga Gankhuyag PDEng;** Model-Based System Engineering Design of Functional Modules for Configurable Topology ■ **Konstantinos Karmas PDEng;** VESSEL - Driving Behavior Monitoring System ■ **Gilda Khosravi PDEng;** Dependency Management and Scope Control ■ **Arne Lapõnin PDEng;** Distilling Insights from the Text Written by Field Technicians ■ **Dũng Lê PDEng;** Data Quality Dashboard ■ **Trupti Manik PDEng;** Increasing Coverage Through Coverage-based Test Selection ■ **Ani Megerdounian PDEng;** From Informal TWINSCAN Machine Diagnostics Data to Reliable Formal Customer Interface ■ **Munkhdalai Sainbileg PDEng;** Multi-domain Modeling Assistance (MMA) Tool: Design and Implementation ■ **Dimas Satria PDEng;** Data Storing and Structuring for Information Discovery in AgTech ■ **Konstantinos Smanis PDEng;** Intelligent Microscope II ■ **Hongyu Tan PDEng;** Integration of Modeling and Data Science ■ **Beza Getachew Tassew PDEng;** Sample Flow Modeling for Rapid Performance and Failure Analysis ■ **Wahyu Utomo PDEng;** Design Complexity Reduction Using Object-oriented Paradigm - Exemplified by Redesigning the Scanning Application of ASML's Wafer Positioning Module

The Software Technology PDEng (Professional Doctorate in Engineering) degree programme is an accredited and prestigious two-year doctorate-level engineering degree programme. During this programme trainees focus on strengthening their technical and non-technical competencies related to the effective and efficient design and development of software for resource constrained and intelligent software intensive systems in an industrial setting. During the programme our PDEng trainees focus on systems architecting and designing software for software intensive systems in multiple application domains for the High Tech Industry.

**The programme is provided by the Department of Mathematics and Computer Science of Eindhoven University of Technology in the context of the 4TU.School for Technological Design, Stan Ackermans Institute.**

For more information, visit our website:

[tue.nl/softwaretechnology](https://tue.nl/softwaretechnology)