Thesis projects at Data Ductus

We are IT consultants that believe technology plays a crucial role in solving many of the challenges that our world faces. Our passionate people work at the crossroads of tech and sustainability delivering amazing solutions for clients.

Write your thesis with us and you can contribut to this, and be a part of an organization that cares for society, employees and your future.

Our core values guide us in everything we do and we strive for a working culture that is inclusive, inspiring and collaborative.

Trustworthiness

- Do what's right for the customer and our company
- Be professional
- Practice honesty and transparency

Empathy

- Always help and support when you can
- Treat everyone with kindness and respect
- Learn the why, not just the what

Collaboration

- Exceed customer expectations by leveraging the experience and skills of our colleagues, partners and customers
- Share your knowledge
- Value your colleagues' knowledge

Innovation

- Never stop learning new skills
- Keep an open mind, value different perspectives
- Embrace change

Many of our employees started their careers at Data Ductus by completing their thesis projects here. It's a great way to get to know each other and we have a long tradition of mentoring tech students. If you choose to undertake your thesis with us, we promise:

- A thesis based on your skills, studies and interests.
- An allocated supervisor with domain knowledge in thesis topic.
- Regular meetings with supervisor and support from other experts.
- Support with thesis university application.
- Invitations to take part in office and/or team meetings during thesis production.
- Potential employment following graduation.
- 50% of thesis students are employed!

We hope that you will find an interesting topic in this document. If not, you're most welcome with your own ideas!

Available thesis topics

At Data Ductus we are always looking for bright minds to help us explore IT. These are the areas we would like to focus on right now. If you have amazing idea or project that's not listed, we are open to suggestion, but you need to wow our experts.

Beyond our listed projects we are particularly interested in sustainability related projects connected to the IT industry.

Submit your application on our website: dataductus.com/students/

The Data Ductus Thesis Program

Let's speak to Max Block to find out more about his role as a thesis supervisor and how the Data Ductus thesis program can help you succeed academically and in the workplace.

What is the Data Ductus thesis program?
It's a program that we have developed to provide the optimal support and working environment for talented students who want to write a thesis in computer science or similar fields. Students get a taste of working life at our company and our consultants get research support and ideas on how to solve certain problems. It is also a great way for students to meet a potential future employer – around 50% of thesis students join the company when they finish their studies.

How can you help students?
We offer all the prerequisites for a great thesis, providing students are dedicated. We have around 250 employees with expertise in a variety of topics, by tapping into that knowledge, we can help students overcome hurdles by answering questions that arise. We also support students with the necessary university formalities, such as application and thesis specifications, which can be difficult to understand if you haven't done them before.

> "It is a great way for students to meet a potential future employer – around 50% of thesis students join the company when they finish their studies."

What do you expect from a student?
Dedication is key. We are sharing our time and expertise and expect similar commitment from students. We want them to write the best thesis possible, that way we all benefit. We prefer it if students work from one of our offices, making it easier for them to meet and get to know our employees – having met someone in the lunchroom usually makes it easier for them to muster the courage to ask for help when they need it. We do not, however, expect students to write code that will be used as part of a delivery to a customer, instead, we want to get insights into specific problems.

What can a student expect from you as a supervisor?
As a supervisor, I have near-daily contact with a student over a 20-week period. They share their findings, discuss any issues/problems they need support with, and they often take part in our stand-ups. I want students to feel part of the team while they are at the company, so they get a feel for what it is like to work here. I also help them to meet the expectations of the university.

Why should someone apply to the Data Ductus thesis program?
It is a great opportunity for students to write a relevant thesis, learn about working life and prove themselves to a potential employer.

How do you decide on the right thesis?
We have a couple of theses in our catalogue that we have identified, but that's the tip of the iceberg. For us, it's much more important that we choose the right thesis for the right person, which is why we try to tune the thesis topics to a student's academic knowledge while giving them a real-life problem to solve. Before

deciding on the topic, we tend to have a couple of meetings to discuss potential ideas that interest both the student and the supervisor. Also, for the university to accept the topic, it must align with the courses that the student has completed.

> "We try to tune the thesis topics to a student's academic knowledge while giving them a real-life problem to solve."

Where can students do their thesis?
We work with most technical universities around Sweden and have offices in six locations from Kiruna in the north to Malmö in the south. We have even had students do their thesis at one office and start working at another office after they graduated. We are flexible, and of course, we work according to the restrictions enforced by the pandemic.

How can students apply?
Applications can be made through dataductus.com/students.

1. Structured GUI Testing Platform

Summary

Research and develop a web application that enables structured testing of graphical user interfaces both by dedicated testers but also during development by developers to ensure higher quality of delivery.

Problem statement

When developing graphical user interfaces testing becomes a time-consuming issue. Both in the hand over from design to development where details about the design can easily be missed and when end to end testing of a finished product takes place. If testing of a GUI could be performed in a structured manner such issues can be avoided. Developers will be less likely to miss details of a design if those details are part of a test, and end to end testing can ensure that all parts of a design will be tested, and a complete test report can be presented to the stakeholders

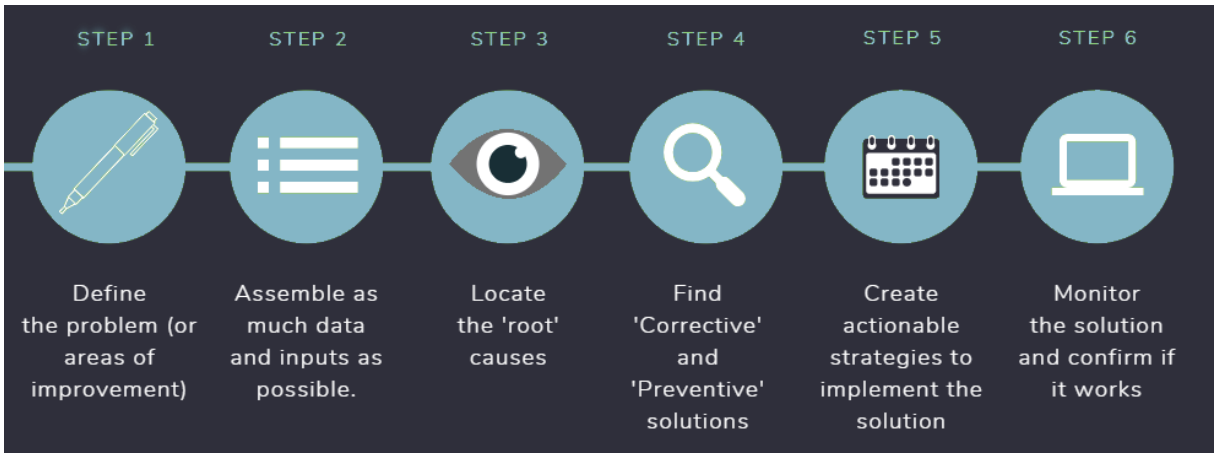
2. Machine learning enhanced Log Analytics

Summary

Research and develop machine learning techniques to automate log analytics and root cause analysis.

Problem statement

Logs are emitted by network devices, operating systems, applications and all manner of intelligent or programmable devices. Logs often comprise of streams of messages in time sequence that usually need to be interpreted and understood. Machine learning has shown huge potential to automate task previously performed exclusively by humans. In this project we will explore and evaluate existing frame works for log analytics and enhance these by injecting machine learning techniques.



Root cause analysis is the art and science of performing log analytics to interpret error messages and understand underlying causes and suggest solutions.

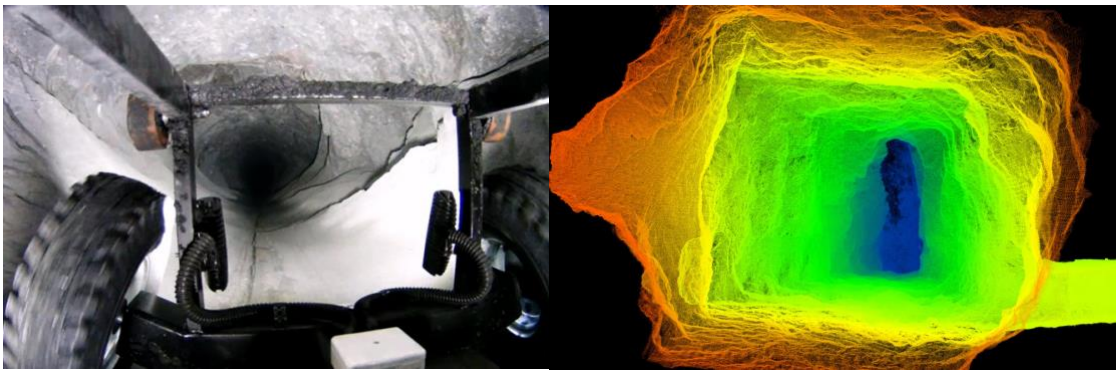
3. Visualization and interaction with 3D and 2D data from mineshafts

Summary

Research and develop user centric visualization of 2D and 3D data collected from underground mineshafts.

Problem statement

Data Ductus have designed and developed several generations of mineshaft-trolleys equipped with video cameras and LiDARs to scan and survey underground mineshafts. Data Ductus have been responsible for hardware, sensors and software. The latest version will have updated capability for high resolution 3D scanning and to explore and interact with the multimodal data collected the operators will need a new user interface.



Left, the mineshaft-trolley descending an underground mineshaft. Right, 3D point cloud data from an underground mineshaft.

4. Foreign object detection on conveyor belts

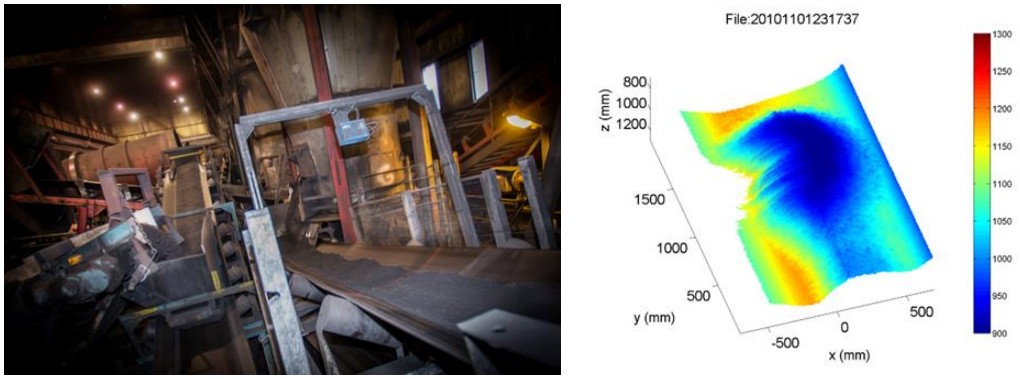
Summary

Research and develop algorithms for foreign object detection in 3D-data from Belt Conveyor Systems.

Problem statement

Data Ductus have designed and developed several generations of Belt Conveyor Systems (BCSs). The current generation is based on a sturdy LiDAR 3D scanner that measure the cross-section of the material transported on the conveyor belt. Data Ductus systems are mainly deployed in the mining industry to measure iron ore and the paper mills to measure wood chips.

Custom algorithms for detecting foreign objects are needed and must be stable and easy to maintain. The algorithms must work on different materials and settings. Increasingly advanced algorithms from Kalman filters to Deep Learning methods will be implemented and evaluated.



Left, the Belt Conveyor System (BCS) installed in an underground mine. Right, 3D point cloud data from the system.

5. Evaluate Time of Flight 3D sensor for Belt Conveyor Systems

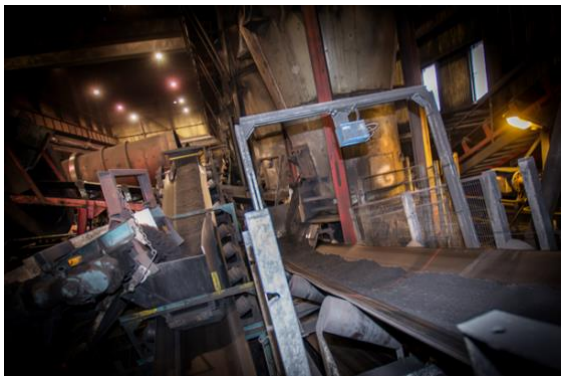
Summary

Integrate and evaluate a new Time of Flight (TOF) 3D sensor for the Belt Conveyor System.

Problem statement

Data Ductus have designed and developed several generations of Belt Conveyor Systems (BCSs). The current generation is based on a sturdy LiDAR 3D scanner that measure the cross-section of the material transported on the conveyor belt. Data Ductus systems are mainly deployed in the mining industry to measure iron ore and the paper mills to measure wood chips.

Time of Flight (TOF) 3D sensors have improved lately and seems to be a viable option to replace the LiDAR. The new sensor must be integrated into the existing system and new algorithms for volume measurements must be implemented. The new sensor must be evaluated in terms of price and performance.



Left, the Belt Conveyor System (BCS) installed in an underground mine. Right, Time of Flight camera that captures 3D point cloud data.

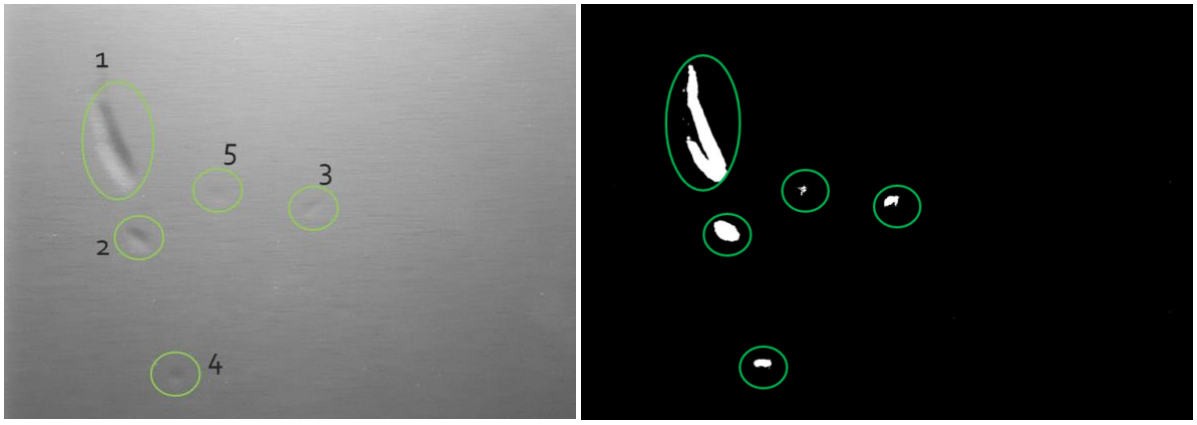
6. Anomaly detection on lacquered surfaces

Summary

Design, develop and evaluate methods for anomaly detection on lacquered surfaces.

Problem statement

Lacquered surfaces are often glossy and even small defects such as scratches and dents are usually reasons to reject a product within the manufacturing industry. Manual visual inspection is time consuming and prone to errors. In this project the student will explore all or some of the steps including imaging techniques, data collection and analysis, training and evaluating machine learning models, visualisation of results.



Left, defects on a lacquered surface. Right, anomaly detection based on adaptive filters and machine learning.

7. Absurdle Solver

Summary

While there is a lot of work done on Wordle solvers, not much exists for Absurdle. Various questions can be answered, for example: Which is the best starting word, assuming optimal play? What is the worst outcome, assuming optimal play?

Problem statement

The web-based word game Wordle gained popularity in late 2021. In the game, the user will try to guess a five-letter word within six tries. After each guess, the user gets an indication which letters are not in the word at all, are in the word but in the wrong place, or in the correct place, respectively.

Wordle's evil sibling, Absurdle, is a spin-off where the game itself changes the word to try to make the user guess as many wrong letters as possible.

In this project, the student will create a Absurdle solver. The input to the solver will be a game, with the previous tries. The output of the solver will be the best next guess.

Example game of Wordle solved after three guesses

Absurdle can be found at: https://qntm.org/files/absurdle/absurdle.html

8. Investigating the Value of ChatGPT in a Developer's Toolkit

Summary

Investigate the usability of OpenAI's ChatGPT and its place in a developer's toolkit by developing with and without chatGPT.

Problem statement

"In the rapidly evolving landscape of software development, the role of artificial intelligence (AI) and natural language processing (NLP) technologies has become increasingly prominent. ChatGPT, a language model developed by OpenAI, has gained significant attention for its ability to generate human-like text and engage in meaningful conversations. This master's thesis aims to explore the utility of ChatGPT as a valuable tool in a developer's toolkit. Through a comprehensive investigation, we will assess its potential benefits, limitations, and impact on software development processes.

Potential research questions:

Can the use of ChatGPT be correlated with improvements in code readability and maintainability, as measured by standardized code complexity metrics (e.g., cyclomatic complexity, code duplication)?

To what extent does ChatGPT contribute to the reduction of repetitive coding tasks, as reflected in the percentage decrease in lines of code written manually?

What is the average time saved by developers when using ChatGPT for code generation compared to traditional coding methods?

What are the best practices for effectively integrating ChatGPT into a developer's workflow, and how can developers maximize its potential benefits while mitigating risks?

" – ChatGPT

One approach to testing this thesis could involve the development of multiple applications. These applications may perform similar functions or be entirely different applications. One of these applications would primarily employ ChatGPT for development, while the second would be developed without any ChatGPT assistance. A potential third application could utilize a hybrid approach, combining ChatGPT and conventional development methods. During the development process, various metrics could be collected and analyzed. Subsequently, conclusions could be drawn based on the data derived from these metrics.