CHALMERS
UNIVERSITY OF TECHNOLOGY

Research and Development of Data Science and AI Tools for Project Management at Trafikverket

Through a Product Development Process

Master's thesis in Master Program Product Development

Franz Erik Weijmer

DEPARTMENT INDUSTRIAL AND MATERIAL SCIENCE

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2025
www.chalmers.se

Master's thesis 2025

Product Development Master's Thesis Research and Development of Data Science and AI Tools for Project Management at Trafikverket

Through a Product Development Process

Franz Erik Weijmer

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Industrial and Material Science
Product Development Master's Program
Chalmers University of Technology
Gothenburg, Sweden 2025

Final Report for Master's Thesis Research and Development of Data Science and AI Tools for Project Management at Trafikverket
From a Product Development Perspective.
Franz Erik Weijmer

Supervisor: Per Persson Schön, Trafikverket
Examiner: Dag Henrik Bergsjö, Department of Industrial and Material Science

Master's Thesis 2025
Department of Industrial and Material Science
Product Development Master's Program
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Image AI-generated with the prompt "An image of a

Acknowledgments, dedications, and similar personal statements in this thesis, reflects the author's own views.

Typeset in LaTeX
Printed by Chalmers Reproservice
Gothenburg, Sweden 2025

Abstract

This master's thesis explores the application of product development methodology to the design and prototyping of data science tools within the public sector. The work was conducted in collaboration with Trafikverket, Sweden's national transport administration, to support more efficient project management in large infrastructure projects.
The study investigates how AI and data-driven tools can be identified, evaluated, and developed at a public agency. The research follows a structured product development process comprising user needs analysis, opportunity identification and screening, concept development, prototyping, and deployment assessment.
Initial insights were gathered through interviews, meetings, and a survey conducted with Trafikverket staff. The data of which was collected and assembled into a needs list. Several potential AI tool concepts were proposed and screened throughout these conversations. Four high-potential concepts were screened through a ranking and comparison to the needs list.
One of these concepts, an AI-supported clustering tool for managing public consultation feedback (Samråd), was chosen for prototyping. Using Swedish-language text embeddings and the BERTopic framework, the tool was designed to automatically group and summarize incoming insights to reduce manual workload and increase response consistency. This opportunity also underwent a deployment assessment of economic, environmental, and social impacts. Based partly on the research knowledge from the project, but also based on a follow-up interview with a consultant at SWECO.
This project demonstrates that utilizing a product development process was not only appropriate, but could also have been beneficiary as opposed to traditional data science processes for an organization with limited knowledge of their own needs for data science and AI tools and how to implement them. The outcomes serve as a road map and proof of concept for Trafikverket's continued adoption of AI tools.

Keywords: Product Development, Data Science, Artificial Intelligence, Public Sector Innovation, Project Management, Opportunity Identification

Acknowledgements

This master's thesis was conducted at Chalmers University of Technology, in collaboration with Trafikverket, at the Department of Industrial and Materials Science during the 2025 spring semester.
I am very grateful to supervisor Per Persson Schön for creating the opportunity for this master's thesis project and the guidance and support throughout its course.
I am also grateful to all those who participated in this project. To those who attended meetings and Interviews at Trafikverket, to Amer Catic from Yolean, and especially Jesper Lunberg from SWECO, for the follow-up interview at the end of this project. Also appreciative of all those who participated in the larger Survey held at Trafikverket's Gothenburg-Borås office.
With this support, the project proceeded as planned and has become both an insightful learning experience and a meaningful conclusion to my time at Chalmers.
This work was supported by Formas- a Swedish Research Council for Sustainable Development, under grant decision number 2024-00087.

Franz Erik Weijmer, Gothenburg, June 2025

Contents

List of Figures

List of Tables

1
Introduction

Artificial Intelligence (AI) is revolutionizing the industry, adopted in almost every digitalized field, including governments and their agencies like the military [1]. Trafikverket, Sweden's traffic administration, is also investigating AI, with multiple different AI-driven projects [2].
This project is about researching and developing a data science and AI tool for Trafikverket from a product development perspective. Tools that will aid Trafikverket's project management, particularly within large railway projects. This means determining Trafikverket's needs for data science and AI tools and identifying the opportunities for such tools.

This chapter presents the background to Trafikverket and the process of large railway projects in Sweden. It also includes the aspects of data science and AI applicable to this project, alongside the project definition, objectives, and scope.

1.1 Background

Trafikverket is one of the larger agencies in Sweden, with around 10,000 employees [3]. As an agency responsible for the long-term planning and construction of railways, roads, and sea travel, Trafikverket is also security rated, increasing its responsibility for Sweden's safety [4]. The agency deals with huge projects that can cost taxpayers upward of 48 billion SEK, as in the case of Gothenburg-Borås [5]. This scale means that even slight improvements to work efficiency will save millions, hence Trafikverkets' interest in AI technology and other ways of improving their work.
These large railway projects need to proceed under Swedish law for the construction of new railways [6], which defines the construction process of new railways that Trafikverket needs to comply with.

1.1.1 Large Projects at Trafikverket

Described in detail at Trafikverket's site [7] and seen in Figure 1.1, the process for large railway follows a simple formula, but is extremely complex and time-consuming. Before initiating a new railway project, there are also pilot studies and considerations made based on predicted future needs. In the case of Gothenburg-Borås, the pilot studies and background before initiating the localization investigation started in the year 2000, before the localization process finally took place in 2019 [5].

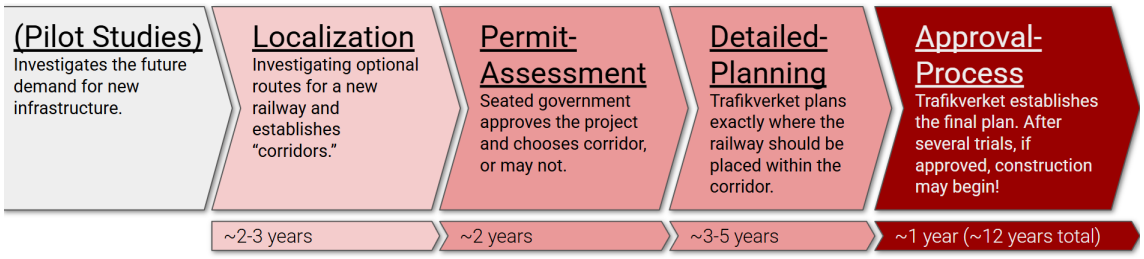


Figure 1.1: Flow chart over the process of large railway projects at Trafikverket.

Multiple "corridors" are established within the localization investigation, or wide stretches where the railway could take place. The corridor could be several hundred kilometers wide in some sections while narrowing down in others. There has to be more than one corridor to allow the current government to make a politically driven choice. The corridors act like suggestions to the government and are meant to include different perspectives, like differing amounts of cost, different branches and stops, and different environmental considerations as well. The seated government chooses the corridor and approves the project before Trafikverket may start with the detailed planning of the new railway, exactly how the railway will be placed within the corridor. The final plan goes through several trials before being approved, in which case construction may begin [7].

During the localization and detailed planning phases of this process, there are "Samråd", public consultations with affected municipalities, other agencies, property owners, and the general public. This allows those affected to provide their insights to any new project [7]. According to Swedish law, Miljöbalk 1998 [8], it is mandated for every project of larger impact to conduct a public consultation, like the construction of a new railway. However, for maintenance and other exceptions, consultations are not necessary. The collected insights have to be answered by Trafikverket and included in their decision-making as part of the democratic process [7].

These major phases- localization, permit assessment, planning proposal, and approval- all take upwards of 12 years before the construction may even begin, which in and of itself can take upwards of 15 years for large railway projects. A modern example of this process would be Västlänken in Gothenburg, which started the investigation in 2007 and got approval to build in 2017, with actual construction starting in 2018 [9]. Västlänken is estimated to be fully operational by 2030.
The completed railway will remain operational for upwards of 100 years, meaning the time-consuming planning and construction is only a small part of a railway's life cycle. This is the main reason for the detailed and careful process of a new railway, since its impact will be felt for multiple decades and generations later [7].

1.2 Data Science

Data science is the broader use of statistics, computing, and scientific methods to extract insights from structured and unstructured data [10]. AI, meanwhile, is the

broad ability of computers to perform tasks that typically require human intelligence, associated with chatbots and different content generators [11]. An example of an AI would be a Large Language Model (LLM), which can read and understand context, and generate text [12]. Or image generators like OpenAI's Sora, which is a combination of different models to achieve the desired result of a prompted image [13]. Notably, there is a significant difference between AI models. Broad AI models such as the general language model OpenAI's ChatGPT [14] or Google Gemini [15] are energy and hardware-demanding [16], while narrow AI, those that predict a machine's failure, for example [17], have considerably lower demands hardware and energy demands [18].

There are processes for developing data science and AI tools other than the product development process, notably CRISP-DM (Cross-Industry Standard Process for Data Mining), which has turned into an industry standard [19]. This process goes through the following steps:

1. Business understanding
2. Data understanding
3. Data preparation
4. Modeling
5. Evaluation
6. Deployment

Which is different from the product development process [19].

1.3 Project Definition

The precise project definition would be the research and development of data science and AI tools at Trafikverket. There was no prior knowledge of what the tool could be or how this topic should have been tackled, which made way for the product development process and the process of opportunity identification.

The process in this project is similar to that of literature from a previous project course from the product development master's program, that being *Product Design and Development* by Karl T. Ulrich, et al (2020) [20]. The process went as follows:

1. Identifying User Needs
2. Opportunity Identification
3. Screening of opportunities
4. Concept Generation/Selection for the screened, worthwhile opportunities
5. Prototyping of an exceptional opportunity
6. Deployment Considerations/Implications of prototyped opportunities

This methodology distinguishes this project as a product development master, as opposed to other fields like data science and AI, and can be read about in Chapter 2.

The broader research question would be whether a product development process is an effective way to research and develop data science and AI tools at an organization. As opposed to other common data science processes, such as CRISP-DM [19].

1.3.1 Objectives

The project objective was to establish the needs and potential use of data science and AI tools at Trafikverket. Alongside identifying opportunities for such tools and further developing some screened opportunities that are more beneficial to the agency.
Several deliverables come naturally from the product development process and become the natural objectives of the project as well:

- Research of Trafikverket's current needs for AI tools, including challenges, and creating a Trafikverket's needs list
- List of many opportunities from the opportunity identification, alongside reasoning and screening of those opportunities
- Some more detailed concepts for worthwhile opportunities for further development in the future
- Creating a prototype for communicating the usefulness and function of exceptional tools/opportunities
- Established deployment considerations for these prototyped AI tools

While there are several deliverables, it is important to mention that these are not meant to be overly comprehensive and detailed, instead acting as a start for further research if desired.

1.3.2 Scope

This project did not research Trafikverket's other responsibilities as Sweden's traffic administration, so the road and sea departments at Trafikverket were never of concern. However, tools that benefit all of Trafikverket are naturally more desirable, something that is taken into account.
The type of data science tool was not limited, and anything was acceptable. Since this is not a Master's in Data Science and AI, this project emphasizes the process of finding opportunities for AI and data science tools and researching the needs of such tools. The tool was not meant to be industry-ready or state-of-the-art within data science and AI, which would perhaps be the objective of other master's theses.

2
Methodology

The methodology is an important part of distinguishing this thesis as a Product Development master's thesis, as opposed to a thesis in Data Science and AI or some other field.
This chapter includes the processes of the approach to tackling this project and motivation, while the actual application and takeaways are found in the respective chapters 3, 4, 5, 6, and 7.

2.1 Product Development Process

The process is based on course literature from the courses "Product Planning - Needs and Opportunities" [21] and "Product Development Project" [22], being *Product Design and Development*[20]. The process was adapted for this project as follows:

1. Identifying User Needs
2. Opportunity Identification
3. Concept Generation
4. Prototyping
5. Deployment Considerations/Implications
   (a) Economic
   (b) Environmental
   (c) Social

Each of the processes has its subsection for further detail.

This process naturally encourages a funnel approach to deciding the product for further development. Instead of assuming the best solution from the start, this process encourages seeking out many opportunities and screening away those that seem less desirable. It also naturally includes several deliverables that could be of use to the agency, like the needs list seen in Table 3.2, and the larger list of opportunities seen in Table 4.1, which includes several ideas that could be of use in the future. The deployment considerations, meanwhile, will be material for future decision-making.

2.1.1 Identifying User Needs and Product Specification

Identifying customer/user needs is described in steps on page 79 *Product Design and Development*[20]. Adapted to this project as:

1. Gather raw data from the user.
   Through interviews

Meeting and discussions
Survey

2. Interpret the raw data in terms of user needs
3. Organize the needs into a hierarchy of needs
4. Establish the relative importance of needs
5. Reflect on the result and process

"Data" refers to the responses and statements made during meetings, which need to be interpreted into takeaways for this project. Meetings were held to allow us students to get a better idea of Trafikverket's tasks and responsibilities, particularly within railway projects. Throughout the obligatory course "Product Planning - Needs and Opportunities" at the product development master's program at Chalmers [21], material surrounding data collection motivated the following decisions for the user needs research.
Interviews would serve as the main method for gathering the data needed, as the semi-structured interviews would allow the interviewees to share their perspective on their work. Since there was no prior knowledge of the work tasks or responsibilities, it would have been difficult to perform structured interviews with effective questions, which is why there were fewer prepared questions and more of a semi-structured approach. Interviews were also recorded and later transcribed using an AI tool at Chalmers. That way, interview takeaways could be connected with an interviewee's quote.

The survey was a joint effort between this and another master's thesis project at Trafikverket, and was made in collaboration since the results could benefit both projects. The questions tackle different subjects surrounding AI at Trafikverket, such as employees' general experience and knowledge of AI development at the agency. Microsoft Forms was used with answers being collected in person at Trafikverket's "Gothenburg-Borås" office's lunch room. QR-codes were also set up in the same lunch room, by the microwaves and coffee machines at eye level. The survey can be found in its entirety in Appendix A.

The interpretation of raw data is rather subjective, but to mitigate that, additional follow-up emails were sent to interviewees to confirm whether interpretations aligned with their views or opinions.

2.1.2 Opportunity Identification

In product development, an opportunity can be an early product idea, a newly identified need, a technological discovery, or a potential solution to any problem, according to the authors of *Product Design and Development*[20].
In this case, the opportunities are in any way that data science tools could aid Trafikverket during their projects, particularly within railway projects. *Product Design and Development* page 41[20], provides a detailed process for opportunity identification:

1. Establish a Charter/detailed purpose for the opportunity
2. Generate and Sense Many Opportunities

3. Screen Opportunities
4. Develop Promising Opportunities
5. Select Exceptional Opportunities
6. Reflect on Result and Process

The charter is essentially created in the project definition:

> *A data science or AI tool that improves the quality or efficiency of project management at Trafikverket in any way, particularly for large railway projects.*

As for finding opportunities, *Product Design and Development* mentions researching existing ideas, locating dissatisfaction, and following personal passion [20]. This process was ongoing when Trafikverket's needs were researched during the meetings and interviews. Outside of finding opportunities in conversations with Trafikverket, other opportunities were considered through complete brainstorming and external research.

- Are there some opportunities that could be tweaked for a vastly different application still beneficial?
- What do other industries do regarding AI, and could there be some cross-industry applications suitable for Trafikverket?
- In what other ways does Trafikverket collect data (or could it collect data), and could it help with large railway projects at Trafikverket?

Taught throughout the "Product Development Project" at the product development master's program [22], the screening of concepts and designs is typically done through some comparison to the user needs. In this project, the screening was meant to capture Trafikverket's opinions surrounding the opportunities through a joint ranking with an accompanying presentation. With enough participants, their opinions surrounding each opportunity should mean a valid average ranking and placements for each, screening based on that alone.
Unfortunately, not enough participants could attend the presentation, and while it was recorded and shared with previous interviewees, it did not collect enough responses either.
To increase the validity of the placements and the screening, the opportunities were also compared to the Trafikverket's needs list (Table 3.2). Particularly, which needs had valid risks not being fulfilled by some opportunity.

2.1.3 Concept Generation

The book *Product Design and Development*[20] shares a five-step concept generation process described in the book as follows:

1. Clarify the problem needing solving (**what** the opportunity for a tool is and what the tool would do, **but not how** the tool would do it)
2. Search externally (**how have others** solved this problem)
3. Internally (**how could we** solve this problem)
4. Explore the system of the tool (all possible sub-objectives and sub-solutions)
5. Reflect on the result and process

Since there is less time for full conceptualization for each opportunity, concept generation is reserved for the screened opportunities. For those screened opportunities, there was more emphasis on clarifying the problem and searching for potential solutions, rather than combining solutions and exploring the concepts systematically. It would be more befitting of a data science and AI master to explore the design solutions for a data science and AI tool on an architectural level.

2.1.4 Prototyping

In this project, prototyping is done for communication and integration of subsystems [20], alongside increasing knowledge. Since the prototype is software and not physical, the prototyping process was to simply start programming the tool and discover as much as possible about the model. To make prototyping fast and effective, the exact functions and design elements are prototyped to answer some questions surrounding the product, or a data science and AI tool in this case.
Of course, it is also important to evaluate the prototype and develop methods of testing the functionality, but this naturally depends on the function being prototyped [20].

2.1.5 Deployment

From Chalmers' product development master's program and the compulsory course "Product Development Project", it was a crucial part to include deployment assessments to the new product, to aid the case companies and project owners in their decision-making [22]. Importantly, subjects such as:

- Economic impact
- Environmental impact
- Social impact

For this project, which expands to the following:

- How much time/money will this tool save Trafikverket (efficiency)?
  Estimating time and cost savings with help or through research
- What are the implementation costs (costs of deployment), if any?
  Will there need to be some computing infrastructure for the tool, or can it be used on existing hardware?
- What will this mean socially at Trafikverket (social response from employees)?
  Will the tool be an appreciated asset?
- What will the environmental impact be, if any?
  Considering the deployment, energy needs, and potential efficiency gains.

3

Research Over Trafikverket's Needs for Data Science and AI Tools

To effectively and accurately create opportunities that could benefit Trafikverket, the agency's tasks, goals, processes, and employees' own experience and considerations need to be understood. The collection of this data was done through meetings, interviews, and a survey that was created in collaboration with another Chalmers master's thesis group working with Trafikverket.
The collected information and takeaways led to Trafikverket's needs list, a key deliverable in this project.
This chapter summarizes the takeaways from meetings, interviews, and the survey at Trafikverket, alongside Trafikverket's needs list, the performance indices and considerations important to Trafikverket in regards to new data science and AI tools.

3.1 Interviews and Meetings

Several meetings were booked by the supervisor at the beginning of this project in early February 2025. Some of the meetings were held as presentations by personnel at Trafikverket, allowing for discussions and questions. Others were like interviews held by us students, with semi-structured formatting for broader discussion and follow-up questions. Alongside these interviews with Trafikverket, an additional interview was held with consultant Jepser Lundgren at SWECO to increase perspective on challenges with working alongside Trafikverket. There was also an interview with the CEO and co-founder of Yolean, Amer Catic, to discuss AI opportunities within the construction industry, but that is addressed in Chapter 4.

3.1.1 Meeting-Takeaways

The meetings were meant to provide introductory knowledge of the agency. The takeaways from this project were important information for the development of a data science and AI tool.

Trafikverket has been working with different AI solutions, presented in a YouTube video "AI satsning på Trafikverket" [2]. Important considerations regarding AI at Trafikverket were presented in a meeting, which are listed internally as follows:

- **Transparency** - That the tool's workings are transparent and understandable to the user.

- **Security** - That the tool does not leak critical information to those without proper authority.
- **Energy** - The energy demands for these tools.
- **Black Box** - Determining the cause of an opaque model's decision-making if needed (connected with transparency).
- **Ethics** - The AI tool does not harm ethical viewpoints, such as equality, nor does not prioritize money over people's well-being.
- **Democracy** - The AI tool should not spread misinformation or in other ways harm Sweden's democracy but should instead aid it, if anything.

Another important consideration is the need to retain knowledge in case an AI tool shuts down or stops being used. In those scenarios, Trafikverket should still be perfectly capable of working and continuing without the AI tool.
There is currently a large and widespread demand at Trafikverket to improve their data handling and classification internally, since data and information can be scattered over multiple databases, while another desire is to strengthen the agency's communication with the public and business partners. This is brought up over multiple meetings and interviews.
Alexander Spondung, program coordinator at Trafikverket, describes the need for assistance with multiple preparatory tasks within large projects. Multiple deliverables are required to initiate the different phases of a new large railroad project, meaning larger cost savings from small improvements.

3.1.2 Interview-Takeaways

The semi-structured interview takeaways are structured in the following table alongside the translated quote.

Statement	**Takeaway**
We have an obligation as an authority to respond to every point of view	Need to answer every entry in public consultations from the public and other actors.
There is a real one. I don't know if you should call it a risk or call it an opportunity [...] *How can it be used and how can it be used against us?*	Some considerations that AI could be used to flood Trafikverket with insights and questions that make it more difficult and time-consuming to handle them.
Certain routines must be done, but it takes quite a long time.	many administrative, repetitive tasks are not value-adding but necessary within the agency.
But we've discussed that quite a lot right here, as well as around automation, and I think it could be done	Automation has been discussed at Trafikverket and is deemed possible, particularly within the context of reporting and documentation.
Who can manage to go through as much material as we produce	A load of material that is not feasible for quality control without becoming costly.

Table 3.1: Table of interview takeaways alongside its correlated quote

So that today we have to do sample checks to be able to check things to get by and even get an idea	Currently, sufficing with sampling some reports for quality control, which can be vague and inappropriate.
Yes. As I said, we produce any number of documents. The EIA and the railway plan are several hundred pages thick. And then numerous drawings on it. Run it into an AI and ask: Hello, what contradictions are there in these documents? [...] There you write one thing, but over there you write something else. It doesn't add up. You didn't have a grasp of the situation here.	Afraid that AI could search through all of Trafikverket's public documents and discover discrepancies that impact Trafikverket's trustworthiness.
So some of these things should be driven by the market that supplies us. For example, a question like this about creating an EIA. It is not really in our interest to create[...] *We do not have the same incentive to drive that development. Of the tools they use in their work*	Trafikverket should not develop AI tools outside of their processes and tasks, other partners should instead make their tools as the privatized market has more of an incentive to do so and do it well.
We want to achieve more efficient information management, and in that, automation, advanced data analysis, and so on are the target picture.	A genuine desire to improve data handling and analysis through more efficient tools.
Our challenge here is that in addition to producing enormous amounts of information, the majority is produced externally	Challenge with classifying data and working with it when the majority of it is made from partnering consultants and contractors.
So it is being able to set the requirements. We can't leave that to some machine or automation. But, on the other hand, compliance with requirements is something we would like more of automated	While some tasks are not desired to be automated or even assisted, setting requirements, but finding compliance with those requirements could be automated.
Somewhere we need to have a human brain that is actually an engineer and can do this	Still a need for a human brain to overview and make decisions.
What we want to do is really for all of us to understand what this signal is and use the same term across the board.	Wants to format new data with unified terms for various objects across planning, construction, and maintenance, so that there are fewer discrepancies.
Even if it were something that saves a few minutes per document or so, it still adds up to a lot of time.	Saving even a little time for some repeated tasks would mean huge savings over a whole project.

Table 3.1: Table of interview takeaways alongside its correlated quote

There are a lot of governing documents at Trafikverket. There are various things, like routine descriptions and T-docks[...] *I find that it can be difficult to know what is available.*	It is particularly hard to know for what things there are standards and templates at Trafikverket, and hard to know where to find them. An example is a T-dock for making a public announcement, which Lundgren had no idea even existed before making the announcement.
Quite recently looking for one. A time document that I had seen in another report that I wanted to read and research further. And there is a database or a website that you can access externally where you can search for documents, where I searched for like this specific number, then and there was nothing. [...] *And it didn't say anything about yes, it exists, but it's no longer valid.* [...] *and then I get a little. But oh, wait, am I on the right page, or is it that this one is not updated with the latest? So, yes, it gets very uncertain.*	Lacking communication surrounding the state of existing T-docs as well, which decreases the confidence of external search.
Then it becomes less independent.	The lack of confidence means asking Trafikverket more often and making the work less independent.
Great to kind of add a chatbot, but it is also like this. What other investigations have we done in a place or on a stretch or something like that before?	For a future Trafikverket's chatbot (which has been researched prior), there should also be information and background for different projects and products in the past. That would be extremely useful.
I've had insights that are... well, 30 pages long sometimes.	Public consultations can be upwards of 30 pages with images, statements, and text. Making it more difficult and time-consuming to understand the real meaning and main insight, and just process the insight.
Personally, I've started experimenting with it now, like [...] *use some type of GTP-bot to summarize it so that it goes quickly*	Is experimenting with using SWECO's closed AI-chatbot to summarize insights from the public consultations. Both to speed things up and to better understand the core material of the insight.

Table 3.1: Table of interview takeaways alongside its correlated quote

In the consultation phase, so you can sort of summarize thematically. So we don't have to respond to each one like that.	With "Samråd", the insights are already grouped thematically to respond to similar insights simultaneously, not having to respond individually to each insight.

3.2 Survey-Takeaways

The survey results were collected fourth of April 2025 with 60 respondents.
Below are the questions that had a meaningful takeaway for this project, while the rest can be found in Appendix A. Each of the sections includes the respective part of the Appendix where they are placed.

3.2.1 Do you have prior experience using AI tools? A.1

While the majority had some experience, particularly with OpenAI's ChatGPT, there was a substantial portion, 19% that had never used any AI tool nor attended courses about the subject. The takeaway would be that most employees at Trafikverket have limited experience using AI tools, which means that there is an extra need for a simple tool that does not expect the user to have extensive knowledge of AI.

3.2.2 How familiar are you with Trafikverket's work on AI? A.2

With an average rating of a low 1.79 on a scale of 1 - 5 in this question, most employees have little to no idea of Trafikverket's AI work. This could be because of communication issues within the agency. It could also be that a lack of experience with AI is because to a lack of interest. The question is, how would Trafikverket encourage the use of a data science and AI tool if the users are not interested? Perhaps through creating a simple and intuitive user interface like that of OpenAI's ChatGPT. The tool should take this fact into account in its development, if possible.

3.2.3 What do you believe are the largest challenges with AI at Trafikverket? A.4

There was also a question surrounding the respondent's believed advantages of AI at the agency, but the challenges prove more important requirement of any data science and AI tool.
While data security was the highest regarded challenge, the lack of AI competence and simple accuracy questions were not far behind. Interestingly, the fear of AI taking jobs was not of concern. This shows the importance of data security at the agency, with leadership and employees. It also shows the importance of simple AI

tools, since the knowledge and experience are lacking, and that the respondents themselves seem well aware of it.

3.2.4 What factors do you believe could create backlash against AI at Trafikverket? A.7

The lacking AI competence is shown in this question as well, with 33% including it as a potential reason for backlash against AI, alongside lacking trust in the AI's decision being at close second. Alongside previous takeaways, this also shows the importance of transparency with an AI tool, for the users to feel comfortable using it.

3.2.5 How do you think AI will affect your work tasks and responsibilities in the future? A.1

These responses can be found in Table A.1 in Appendix A. While a large portion of respondents make statements of increased productivity and efficiency, through supported searching for knowledge, language checking, and document control, there is also a larger part of skepticism and indifference. Some respondents think this will have a marginal or no change on the work tasks and responsibilities.

Broadly speaking, the survey results strengthen the importance of careful, secure, and well-motivated AI tools at Trafikverket. There is a need for internal communication surrounding AI at Trafikverket, its purpose, and the justification for its use. It is important to design future data science and AI tools with doubt and opposition in mind to ensure a successful deployment of such tools. Otherwise, there could be a risk of the tool not being fully utilized at the agency by its employees, but also a risk of dissatisfaction with the implementation of AI in general.

3.3 Trafikverket's Needs List

The needs list becomes a simple yet comprehensive way to summarize the needs at Trafikverket for data science and AI tools, while also ranking the needs on importance.

Need/Wishes/Demands	**Importance/10**	**Code**
Follow Laws ! *	10	D01
Absolute Data Security ! **	10	D02
Transparent/Black Box	9	N01
Ethical***	7	N02
Retaining Competence****	6	N03
Accuracy	8	N04
Improve Communication with Partners	8	W01
Improve Communication with the Public	7	W02

Continued on next page

Need	Importance	Code
Improve Internal Data Handling	9	W03
Lower Energy Consumption	4	W04
Increase Efficiency	6	W05
Increasing Consistency and Quality	7	W06
Simple User Interface	8	W07

Table 3.2: Table over needs at Trafikverket, including their description, importance, and code for later referencing.

'!' Refers to critical need/demand
* Follow the laws for the construction of new railroads and newer AI-related laws in Sweden, presented in Section 1.1
** Ensuring no leaking of safety-critical data and information.
*** The model shall not be biased against groups of people and should not prioritize cost savings over human needs.
**** Retaining the competence to work sufficiently independently from the tool if it shuts down.

Following Swedish laws is paramount for a government agency, and as a security-critical agency, leaking any of Trafikverket's classified information and data could jeopardize Sweden's safety.
Both accuracy and transparency are important factors for a data science and AI tool to be trustworthy at Trafikverket. High transparency is also important to share the competence and retain it with the user.
That the tool is ethical and not biased is important, much like how any employee at a Swedish agency should be non-biased and ethical in their work.
Improving communication with partners, consultants, and contractors is very important, as this communication makes up a considerable amount of Trafikverket's work. Communicating with the public is also very important to create infrastructure that is not entirely out of touch.
Data handling is a concern at Trafikverket, there is simply too much being produced to realistically handle, which has made data science a promising solution.
While improving the speed or efficiency at Trafikverket is naturally important, there cannot be any cutting corners. A tool helping Trafikverket to do better and be more comprehensive, and have higher quality results, is more important.
Other considerations and needs stem from general consequences of using demanding AI, alongside the user's experience surrounding AI.

4
Opportunity Identification

Since there was no predetermined idea of what a data science and AI tool could be in this project, identifying multiple opportunities for such tools became an integral step in the research and development of any such tool. Opportunities were identified predominantly through meetings and interviews with Trafikverket, with some other "external" opportunities being identified through different methods. External in that they do not stem from conversations with Trafikverket.
The first list of opportunities was screened through a ranking performed by members at Trafikverket, alongside the considerations listed in Trafikverket's user needs list 3.2.
This chapter addresses both the conclusive list of opportunities 4.1, including a brief description of the opportunities, alongside the elimination process, and briefly why opportunities were eliminated.

4.1 Listing All Opportunities

The list in Table 4.1 contains a brief description of the opportunity alongside its designated code for later referencing. The following subsections explore background and other details surrounding the individual opportunities.
The code also represents where the idea came from. "OTVK" for ideas addressed in meetings and interviews with Trafikverket, while "OEX" stands for external ideas from research and brainstorming.

Opportunity	**Code**
A clustering model for "Samråd" (public consultations) that clusters public questions and insights.	OTVK1
An automated model for quality control replacing sample control.	OTVK2
The model creates an entire railway project plan for a new railway based on all prior data from previous projects.	OTVK3
A converting model that changes all governing documents into the same format (into TDOK or some already existing one).	OTVK4
Automated EIA (environmental impact assessment) document creator, using new measured data and drawing conclusions based on previously made EIA.	OTVK5
Image generator for the first appearance of new projects, taking the view within the detailed plan program (not rendered) and creating a rendered version based on previous Trafikverket concept renders.	OTVK6

Table 4.1: List of brief descriptions of all opportunities, and designated code for later referencing.

An administrative assistant tool that would bring external services and tools securely to Trafikverket.	OTVK7
Consultant and contractor ordering aid, helping with added insight based on previous contracts/orders made.	OEX1
A public chatbot that could answer questions and point individuals to pages on a certain topic, like the stretch Gothenburg-Borås.	OEX2
Stakeholder and review generator, using previous reviews and feedback data from the public to predict reactions to new products.	OEX3
Automated pilot study for new railway projects based on monitoring travel times, budgets, railway capacity, expected city growth, and more.	OEX4

4.1.1 From Trafikverket's Employees

Most of the ideas in this project were brought up in the conversations with Trafikverket. Since meetings and interviews were less structured, there was an opportunity to discuss potential ideas and opportunities naturally when they came up. The opportunities that came from Trafikverket are designated with the code "OTVK" in the list of opportunities 4.1.

4.1.1.1 Samråds Clustering Model [OTVK1]

This idea came from hearing the process of collecting data on public insights surrounding new projects at Trafikverket, like the construction of new railways. Takeaways from the interviews surrounding public consultations or "Samråd", found in Table 3.1, could be summarized as follows:

- There is an obligation to respond to every insight surrounding a new project.
- Trafikverket does not need to give a unique answer to each insight; answering the thematically similar questions with one comprehensive answer is often done instead.
- Insights can be both concise and long-winded.
- It can be difficult to understand the core meaning of an insight.
- Some summarization tools are already being experimented with.
- Believed that AI could also flood Trafikverket with questions and insights that can increase the difficulty of properly executing the consultation.

The solution could be some model clustering of all insights that are sent to Trafikverket, simply clustering based on word choice within the insights, with the thinking that if the same words are used, the insights are thematically similar. This means larger quantities of insights, like those for larger projects or AI-based insights, could all be handled considerably easier through this simple clustering. Additionally, the data could go through additional models that summarize the clustered insights into one comprehensive yet concise question that, when answered, essentially addresses all

insights that were originally clustered. It could also showcase a summarized version of any insight.

4.1.1.2 Automated Document Control [OTVK2]

From the interview takeaway (Table 3.1), there are too many documents being produced in the projects, making it unfeasible to control the quality by hand. Instead of settling for a sample control approach. Automating this quality control process would simply be a solid increase in feasibility, thus improving quality and consistency at Trafikverket. This model could be a secure alternative as opposed to existing solutions, and would ideally be used by the writer of the document.

4.1.1.3 Railway Plan Generator [OTVK3]

This is simply a theoretical opportunity that was also addressed in an interview with Trafikverket. With all the detailed data surrounding previous railway projects and all their phases, it should theoretically be possible to create a railway plan generator for the various phases of a project. This could then be used as inspiration for the new railway project, with the considerations made based on previous projects.

4.1.1.4 Governing Documents Converter [OTVK4]

One problem addressed over meetings and in several interviews is the many governing documents and their formats. Since there are different formats of these governing documents, and they are not all found in the same database, it would aid workers and external consultants if the governing documents were all unified in the same format and database. With the sheer amount of data, however, it would be difficult to make this change by hand. Creating a tool that would aid with this transition could be necessary for that purpose.

4.1.1.5 EIA Document Generator [OTVK5]

Due to the sheer number of EIAs done through all of Trafikverket's projects, it would greatly aid the consultants if the template for reporting an EIA were partly or entirely automated by simply giving a model the new measured data. Conclusions could also be automatically drawn based on previous EIA conclusions, which could also benefit the user in drawing more comprehensive conclusions.

4.1.1.6 Concept Image Generator [OTVK6]

In a meeting, software for detailed railway planning was showcased. It visualized certain boundaries and many elements that the railway had to consider in the real world. This software has simple graphics meant to be easy to work with, which would not be announcement-ready for the public. To create publicly ready, rendered images of the concept behind the plan, the consultants take time to create a stylized and detailed scene to then render.
To allow for earlier conceptualization, there should be potential for an image-generating model to create the rendered concepts automatically from just seeing the work view

of the original software, based on such a view and previously rendered concept images, or even images from real life and the finished construction.

4.1.1.7 Administration Assistant Model [OTVK8]

In one interview, repetitive tasks were discussed, with administrative tasks being one in particular. Creating an AI assistant that can automatically book meetings and create agendas based on previous scheduling could be interesting and would, according to the interviewee, save some time.
With further thought, this idea became more of a simple concept: what existing AI services would be helpful to recreate for Trafikverket's needs? Locally on a device without connection, ensuring absolute data security. This could be anything from a secure grammar checker, a language model for summarization, etc.

4.1.2 External Ideas

These refer to the opportunities that are coded with "OEX" in the list of opportunities 4.1 for added distinction.

4.1.2.1 Assistant Ordering Model [OEX1]

This idea stems from an interview with the CEO and co-founder of Yolean, Amer Catic, a project management service [23]. Various ideas were discussed surrounding the ways AI could be used in project management and within the construction industry, with the one sticking in this project being that of a contract ordering aid. This contract model would be loaded with previous contracts and their requirements, and suggest new requirements based on those already set for a new contract and project. The previous contracts could also be updated with requirements that should have been included, considering how the project went, which would result in the tool suggesting requirements that could avoid future difficulties hard to predict.

4.1.2.2 Public TVK Chatbot [OEX2]

Since Trafikverket has already researched a chatbot application, but for employees and consultants, it could be interesting to adapt such a model to also address public interests. Answering questions with public information from the public web pages. Also pointing the user to what page where the information could be found in its entirety for further reading.

4.1.2.3 Predictive Review Generator[OEX3]

This generative model would use prior reactions and insights regarding railway plans to predict reactions to new plans. This could mean that the plan could effortlessly include the public perspective into any new railway project by including generated, predictable insights that the public could lift.

4.1.2.4 Automated Pilot Study Model [OEX4]

There is a whole lot of data that could be used for an automated pilot study:

- Data over trains, including their punctuality, occupancy rates, and punctuality on all railway stretches
- Data on project population growth in different municipalities
- Old project documentation, etc

A model could track all this information, taking as many things into account as possible to suggest a possible solution, if anything should be done at all. This supporting tool could perhaps aid pilot studies with the data needed for decisions, all automated within the model's application.

4.2 Elimination of Opportunities

With the list of opportunities, Table 4.1, there needed to be a screening for what ideas are more or less desirable to Trafikverket. In this project, this screening was performed by presenting the opportunities to Trafikverket, letting the participants rank the opportunities. The rankings were performed in a survey, with rankings being made based on the participants' ideas of what is most desirable to the agency. This ranking included all opportunities but the railway plan generator [OTVK3]. This is due to a limitation with Microsoft Forms: Rankings only allow for 10 entries. OTVK3 was eliminated since it does not align with Trafikverket's current needs. It is too energy and hardware-demanding while simultaneously being too difficult to realize, collect the training data, and ensure even somewhat dependable results.

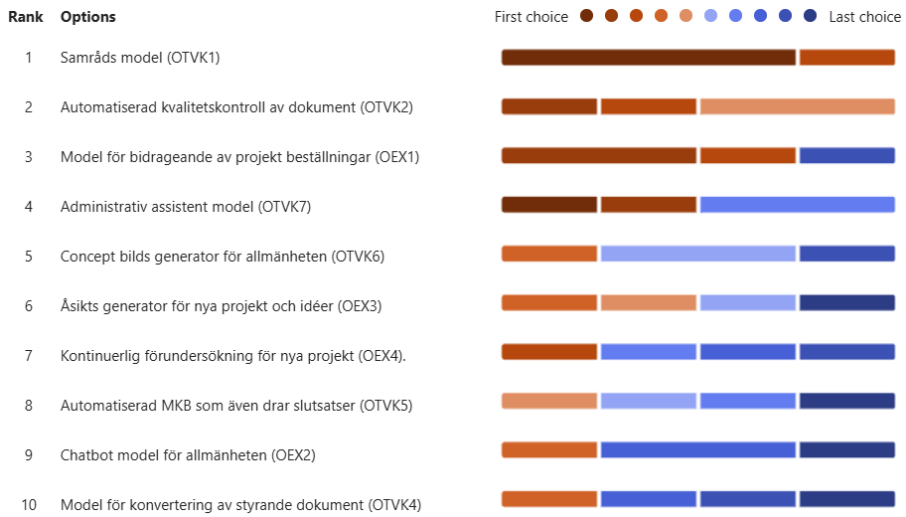


Figure 4.1: Resulting placements for the opportunities in the ranking. Names translated to Swedish for the rankings and presentation.

The ranking shows a particular preference in opportunities OTVK1, OTVK2, and OEX1, with OTVK7 being close behind. This screening means no more consideration for the eliminated opportunities due to time constraints, focusing development on these four.

Especially OTVK1, which became the opportunity to undergo prototyping and deployment assessment as well.

Additional considerations for screening would be comparing the needs from Table 3.2. Below, in Table 4.2, are the needs that are unlikely to be met by any given opportunity in this project. This affirms the ranking results from Figure 4.1, with fewer wishes and demands being fulfilled by the lower-ranked opportunities.

Table 4.2: Opportunities and their respective codes from Table 4.1, and the needs codes from Table 3.2. Development should be reserved for external private actors with more incentives.

Opportunity	**Needs Potentially Not Fulfilled**
OTVK3	N01, N03, N04, W01, W02, W03, W04
OTVK4	W02, W05, W06
OEX2	W01, N05
OTVK5	Should be reserved for private actors.
OEX4	N03, N04, N05, W01, W02, W04, W06
OEX3	N01, W01, W02, W03, W04, W06
OTVK6	N01, N02, N03, N04, W01, W03
OTVK7	N01, W01, W02, W03, W07
OEX1	W02, W03
OTVK2	W01, W02, W05, W07
OTVK1	W01

5
Concept Generation

With undesirable opportunities screened from further development, the remaining opportunities can continue their development and research. This chapter further develops the opportunities with function-block diagrams.

5.1 Samråds Clustering Model [OTVK1]

As seen from the screening process (Section 4.2), this opportunity is highly ranked and matches the needs from Trafikverket's needs list perfectly (Table 3.2). The idea is for some model to cluster the collected insights from any public consultation (Samråd) and simplify the earliest stages of resolving the consultation and responding to all insights.
Notably, Pluto already exists as the public consultation data center [24]. The idea is for this tool to be integrated into this existing service by sending the results to it.
The function block diagram in Figure 5.1 shows the data flow.

1. Entries are sent to Trafikverket, which is sent to this data science and AI tool.
2. The data gets preprocessed/prepared for the coming models. Likely reformatting and cleaning, or eliminating unnecessary data within the entry for clustering.
3. The prepared data is sent to a clustering model that is based on word usage and creates labels for the entries, with later entries falling into those labels/clusters if similar words have been used. These labels reach Pluto, allowing for filtering insights for any thematically similar questions.
4. Entries also go to a separate language-based model.
5. Out of it comes a summarization of individual entries. The summarization will join the entry in Pluto.
6. From the clustered group, and their respective summarization, they should be able to go into the same language model for summarization over the whole cluster.
7. That model should then comprehensively summarize the entire cluster. Using the already summarized and shortened entries '6b' seen in the diagram 5.1.

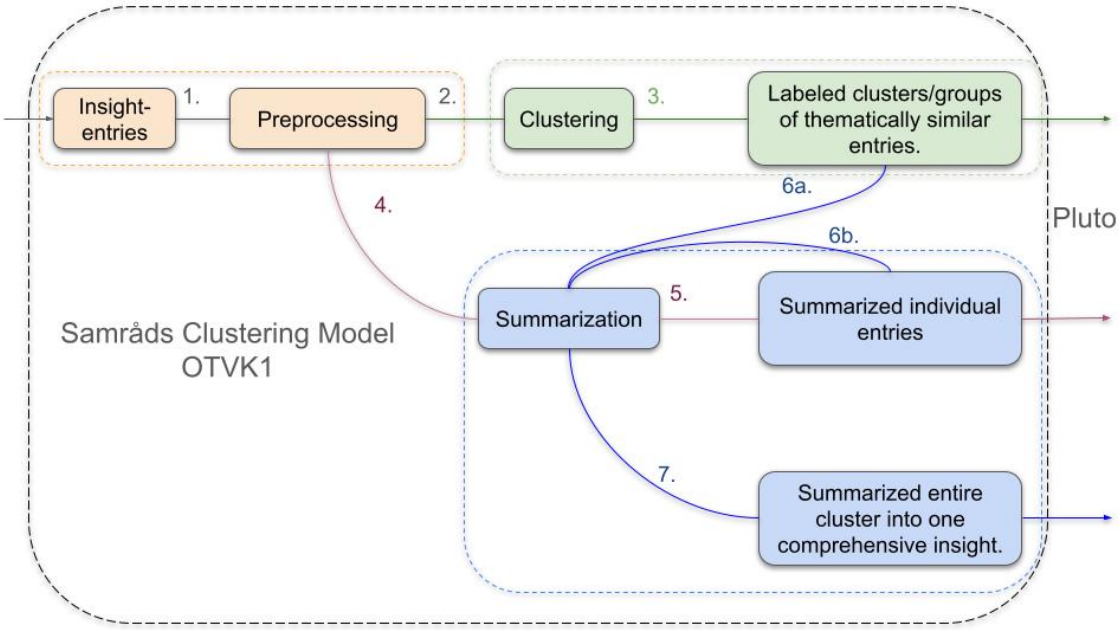


Figure 5.1: Figure over the function block diagram, with numbers representing the paths data goes between functions, and the colors representing the separate models or pieces in the tool.

5.2 Automated Document Control [OTVK2]

This tool would help improve quality and consistency across the agency. With the amount of documents being produced, this tool would aid Trafikverket when controlling for quality. Another function expressed at Trafikverket when presenting the opportunities would be the function of some filtered reading. Allowing the reader to remain focused on the contents of the document rather than the language and other faults.

The function block diagram, seen in Figure 5.2, is connected to the following points below:

1. Any documents are fed to the tool. The document goes through any necessary preprocessing, likely removing images that this tool is not meant to understand or read.
2. The document is checked for appropriate language and grammar, including using the same language as Trafikverket.
3. These changes are presented and have to be accepted or changed, perhaps allowing the model to learn from either its prediction or if the user made a different change.
4. The changes get saved into a new document.
5. Another model addresses formatting without controlling the language, ensuring the right content sits under the right header, which is meant to align with the standards Trafikverket has put in place.
6. The layout is once again suggested and needs to be accepted by the user.
7. Once again, save the changes before finalizing them.

8. From the two separate functions, the suggested changes could reach a different interface.
9. This different interface could allow for a "filtered reading" view, meaning someone not in charge of the document, but still interested in its contents, can read through it without bothering with language faults.
10. If the original user is happy with the edited and controlled document and the changes made, they simply save it as the controlled document.

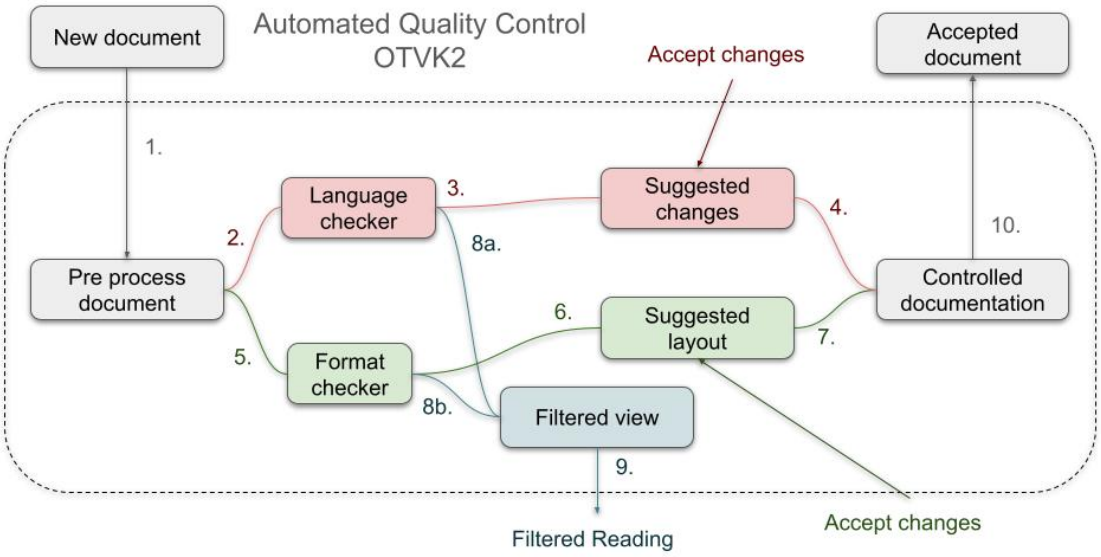


Figure 5.2: Figure over the function block diagram, with numbers representing the paths data goes between functions, and the colors distinguishing the separate pieces within the tool.

5.3 AI Assistant [OTVK7]

The "AI Assistant" is meant to aid with anything repetitive, simple, yet time-consuming. Many services are meant to tackle those tasks, such as Microsoft's Copilot [25], but there is no guarantee of that information being secure, with a considerable risk of the models being exploited to leak sensitive user data [26].
This tool should explore how these services could be implemented with assured data security. Either through having limited access or perhaps through having the entire software locally, ensuring that there is no data leakage.

- Meetings:
  - Speech recognition and creating transcriptions.
  - Process those transcriptions to create notes over a specific point.
  - Include a model that could answer questions behind this meeting through the transcriptions.
- Language checker/add-on:
  Allows for secured language checking even with security-sensitive text.

- Reading aid:
  Summarize documents and allow for asking questions about the document.

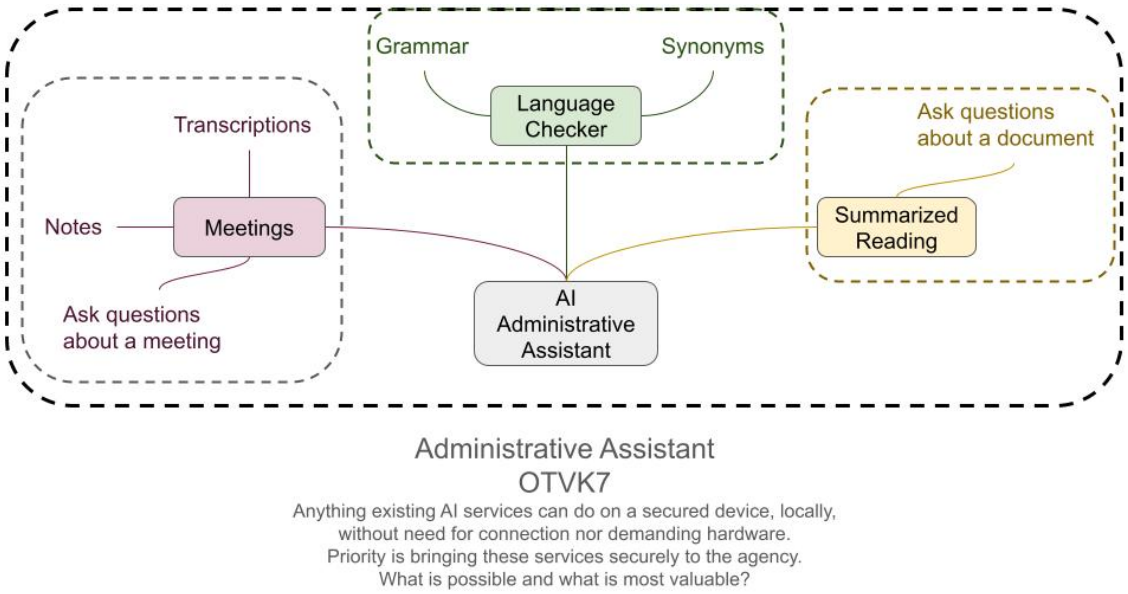


Figure 5.3: Figure over the function block diagram, with colors representing different pieces of the tool. No flow of information or data, instead showcasing the various administrative tasks that could be interesting or beneficial to Trafikverket, without being overly complicated or requiring a connection to the internet.

5.4 AI Contract Aid [OEX1]

As explained in Chapter 1, Trafikverket is a contracting authority and hires consultants and contractors to perform most of the actual work, like EIAs, detailed planning, construction, etc. This means that there is a load of contracts being made constantly. During this project, a team at Trafikverket had researched, developed, and performed a proof of concept (POC) of an AI-assisted tool for requirement specification. A tool that controls the language and grammar of a new requirement, and whether the requirement contradicts an existing law or standard.
This AI contract aid would essentially suggest new requirements based on previous contracts and the requirements already set for this new contract. The function block diagram, seen in Figure 5.4, showcases the flow of data and functions, and how this opportunity should be included alongside the existing POC.

1. When a new contract is meant to be made, a new tool POC at Trafikverket may check the requirement.
2. POC checks for formatting like ending with a period, but also checks so that the requirement is not contradictory to actual rules or standards. All of this information and fixes are suggested to the user.
3. The beginning of this new AI contracting tool lets the user add the controlled requirement to a profile.

4. Old requirements, meanwhile, have been loaded into this tool as well.
5. The requirements are also reformatted to be those of a profile that reaches the actual ordering model.
6. The model takes the current profile into account and compares it to previous contract requirement profiles. Based on these profiles, the model suggests requirements that should be included in this new project contract, considering the already included requirements and the old contracts.

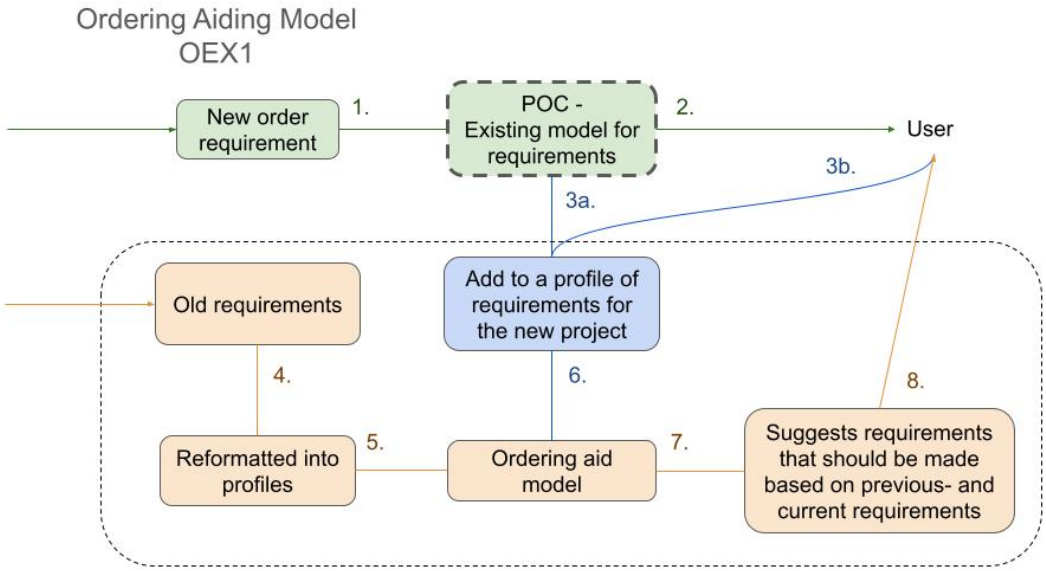


Figure 5.4: Figure over the function block diagram, with numbers representing the paths data goes between functions, and the colors representing the separate models or pieces in the tool.

6
Prototyping

Prototyping in this project is used to increase knowledge of the data science and AI tool OTVK1 while communicating its function and usefulness to the case company. It also extensively increased the knowledge of the Samråds clustering model OTVK1, which performed best in the screening of opportunities (Section 4.2).

6.1 Prototyping OTVK1

The prototype encompasses the functions of preprocessing Swedish insights and clustering, alongside retrieving information about the cluster of thematically similar insights, best showcased in the function block diagram 5.1.
BERTopic, a topic modeling technique [27], became a clear solution for clustering the insights, simultaneously saving important information, including keywords within a topic. The BERTopic model includes the other models necessary for preprocessing and clustering, such as CountVectorizer [28], LanguageTranslater [29], UMAP (Uniform Manifold Approximation and Projection) [30], and HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise). These components serve the following functions:

- **CountVectorizer** – Performs basic text vectorization by counting word occurrences, which may result in some words being removed or others weighted more heavily [28].
- **LanguageTranslater** – Converts sentences into high-dimensional numerical representations that capture semantic context, including synonyms and nuances of meaning. This enables more meaningful clustering [29].
- **UMAP** – Reduces the dimensionality of the embeddings, making clustering computationally efficient and more interpretable, particularly for large datasets [30].
- **HDBSCAN** – A hierarchical extension of DBSCAN that performs clustering across multiple parameter values to identify the most stable and meaningful groupings [31].

When testing the code, ChatGPT generated a handful of insights in Swedish, and an additional set was made by hand. The insights are meant to cover obvious subjects surrounding the planning and construction of a new railway, with some also acting as noise for the model, and should not be clustered with other entries. Alongside topics like noise pollution, parking, and the effect on local traffic, some entries aim to mention a place or location. Those insights should be clustered with the location, in that case.

6.1.1 Preprocessing OTVK1

The insights are already placed in an appropriate list of strings, which makes any such reformatting unnecessary. First step is instead to remove Swedish stop-words, such as "inget", "bli", "på", etc. [32]. This is done using a CountVectorizer to identify the stopwords and remove them [28]. This is done through the code shown in Listing 6.1.

```
# --- REMOVING STOPWORDS / COUNT VECTORIZER ---
from sklearn.feature_extraction.text import CountVectorizer
import nltk
from nltk.corpus import stopwords
swedish_stopwords = stopwords.words("swedish")
# print(swedish_stopwords)
vectorizer = CountVectorizer(stop_words=swedish_stopwords)
```

Listing 6.1: Removing swedish stop-words using a vectorizer.

The vectorizer is later included in the BERTopic, allowing it to remove the stop-words.

On top of removing stop words, the insights need to be encoded into embeddings that translate the words into their contextual value, with sentences closer in embedding space being about the same topic. This is what clustering is based on, as opposed to the words without context [33]. While there are multilingual embedding models, a Swedish-only embedding model would be more efficient and likely more accurate as a result [29]. This code can be seen in listing 6.2.

```
# --- EMBEDDINGS ---
from sentence_transformers import SentenceTransformer
swe = "KBLab/sentence-bert-swedish-cased"
embedder = SentenceTransformer(swe)
```

Listing 6.2: Embedding model to extract contextual meaning of the insights.

The embedded also gets included in the final BERTopic model.

6.1.2 Clustering

Clustering the insights is a custom HDBSCAN model [31] that is sent to the BERTopic.

```
# --- HDBSCAN ---
from hdbscan import HDBSCAN
hdbscan_model = HDBSCAN(
    min_cluster_size=2,
    min_samples=1,
    cluster_selection_method='leaf',
    cluster_selection_epsilon= eps
```

Listing 6.3: HDSCAN model customized with a smaller min_cluster_size = 2 and min_samples = 1.

This is the model to customize to improve the model's results, with min_samples being crucial to clustering alongside cluster_selection_epsiolon [34]. HDBSCAN is not as susceptible to noise, declaring entries with separate topics as -1 instead of other clustering methods that try to cluster every entry, even those that are unique or distinctly about something entirely different.
Finally, for clustering, all of the customized models are included in the BERTopic model [27], seen in Listing 6.4.

```
# --- Bertopic---
from bertopic import BERTopic
topic_model = BERTopic(
    embedding_model = embedder,
    vectorizer_model=vectorizer,
    min_topic_size = 2,
    umap_model=None,
    hdbscan_model=hdbscan_model
)
```

Listing 6.4: BERTopic model with the custom swedish embedded, vectorizer, UMAP removed, and HDBSCAN also custom for our needs.

Notably, BERTopics comes with all the models baked in for easy and early testing of the function through simply fitting the model seen in Listing 6.5.

```
topics, probs = topic_model.fit_transform(insights_Hand)
```

Listing 6.5: Fitting- and clustering the data using fit_transform() for BERTopics

The "topics, probs" results represent the clusters called topics, and the confidence in the insight's topic placement for the probs.

6.1.3 Evaluation and Testing

Using PANDAS, a common data analysis library for Python [35], the assigned topic, insight, and certainty can be displayed, as seen in Listing 6.6.

```
# --- PANDAS ---
detailed_df = detailed_df.sort_values(by=["Assigned_Topic", "Probability"],
    ascending=[True, False])
print("\n Detailed Assignment Overview:")
print(detailed_df.to_string(index=False))
)
```

Listing 6.6: Using pandas to display the clustering information.

BERTopic comprises multiple visualization aids, such as get_topic_info() and get_representative_docs(). With topic_info = get_topic_info(), the code in Listing 6.7 goes through each topic and presents its generated name, keywords, and represented document.

```
print("\n Topic Breakdown:")
for topic_id in topic_info['Topic']:
    if topic_id == -1:
```

```
        continue  # Skip outliers
    topic_keywords = topic_model.get_topic(topic_id)
    name_series = topic_info[
    topic_info['Topic'] == topic_id]['Name']

    print(f"\n Topic {topic_id}: {topic_name}")
    print("  Top Keywords:")
    for word, weight in topic_keywords:
        print(f"   - {word} ({weight:.4f})")

    print("  Representative Insights:")
    reps = topic_model.get_representative_docs(topic_id)
    for rep in reps[:3]:  # Limit to 3 examples for brevity
        print(f" {rep}")
```

Listing 6.7: Printing information about each topic including keywords.

Alongside the resulting table and topic overview seen in Listings 6.6 and 6.7, the list of clustering values is printed alongside the desired "true clusters" to compare with.

7
Deployment Assessment

Assessing a new product, service, or tool is meant to yield more information and considerations for further development or deployment. As an agency, Trafikverket has a responsibility not only to consider economic costs and savings, but also environmental and social impacts. The prototyped opportunity of OTVK1 is the one also being assessed in this chapter.

7.1 Economic Impact

To aid in establishing the economic impact of the Samråds clustering model (OTVK1), Jesper Lundberg, a consultant at SWECO who works actively with Trafikverket and their public consultations, was contacted for a follow-up interview.
The interview started with showcasing the prototype and its function, and further details of the concept of OTVK1 were discussed. Then, the following prepared questions were considered:

- How much time is spent on any given public consultation and its various phases, like handling the insights?
  - How many man-hours and consultant hours are there generally within any consultation?
  - How many of those hours are spent handling the insights?
- What would be the impacts of implementing this Samråds clustering model (OTVK1)?
  - Would the tool be appreciated?
  - Could the tool save considerable/measurable time within any given consultation?

These were the talking points and learnings from this interview:

No consultation is the same

Every project at Trafikverket is different in scale and environmental impact, so too are the consultations. Depending on the impact of any project, the consultation can get more detailed and complicated as more people are concerned or affected by it, including whether or not the project impacts other agencies as well.

Example Ostlänken

Ostlänken is another ongoing railway project in Sweden increasing the railway capacity between Linköping and Södertälje on the east coast [36]. In November

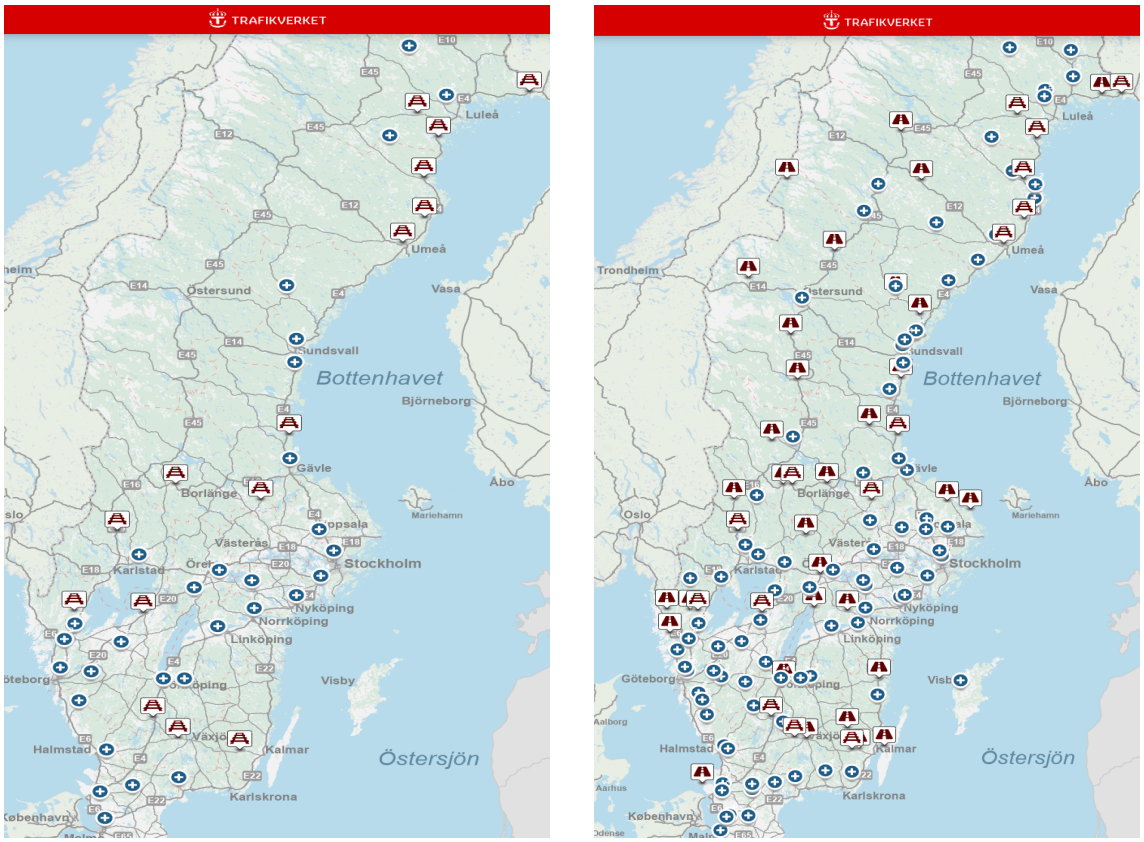
of 2024, the stretch Loddby - Butängen in Norrköping had its public consultation, which Jesper Lunberg worked with [37], wrapping up in April of 2025. Lundgren counted 154 work-hours in this consultation, approximately 310 hours for him and his colleague, including others who are involved through internal meetings, reviews, and technical insights surrounding specific areas, such as noise barriers.
Likely comparable to the public consultation Almedal - Mölndal that is being held in May 2025 [38].

7.1.0.1 Impact of implementing an insight clustering model (OTVK1)

Jesper Lundberg's impression of the functionality was positive. Like predicted, it was agreed that the tool would simplify planning and first sorting of insights, something particularly important for larger, more complicated projects and their public consultations. However, it was also deemed a solid aid to the work in general, and according to Lundgren, it could also be helpful in communication, with labels to relate to during meetings and across emails.

While the tool may not result in measurable time or cost savings for an individual consultation, its value becomes more apparent when considering the large number of consultations conducted each year. According to Swedish environmental legislation (Miljöbalk 1998 [8]), any project with potential environmental or social impact is legally required to include public consultation. This requirement does not apply to routine maintenance of existing infrastructure such as roads and railways. A review of Trafikverket's project map [39], shown in Figure 7.1, reveals a total of 428 ongoing road and railway projects. A search for the term "Underhåll" (maintenance) indicates that 53 of these are strictly maintenance projects and therefore exempt from the public consultation requirement [40].
Even though the time or cost savings of using the tool may not be easily measurable for a single consultation, the cumulative effect across many projects is potentially significant. For example, assuming a conservative time saving of 30 minutes per project for the 375 non-maintenance projects identified, this would amount to approximately 187.5 hours in total. At an average consultancy rate of 974 SEK per hour (as reported in 2020 [41]), this translates to a potential cost saving of around 182,625 SEK for taxpayers, with it increasing over time with more consultations.

As for deployment, the current prototype employs traditional clustering methods that are both computationally efficient and scalable, meaning costs remain low. In practice, larger projects typically involve around 60 insights— far fewer than the hundreds seen in more data-heavy domains, further reducing the hardware demands for this tool. Also discussed with Yolean CEO Amer Catic [23], rented server infrastructure is an option with an estimated cost of 1,500 SEK per month, making it a valid infrastructure choice as well.
This means that the tool should effectively pay for itself, as even conservative estimates of the average savings would easily pay for the development and deployment of this tool.



(a) Map of railway projects.

(b) Map of rail- and road projects.

Figure 7.1: Map over all ongoing infrastructure projects planned and ongoing at Trafikverket for road and railway [39]. The '+' is a group of geographically close projects, seen when zooming in enough.

7.2 Environmental Impact

Broader AI and its overwhelming energy demand are a modern concern according to United Nations [16]. On average, data center energy demands are those of 100,000 households, with the largest centers consuming 20 times more than that [42]. However, the prototype of OTVK1 is not a general AI or even an LLM or similar advanced model.

Comparably, most of the energy is spent on training for a traditional clustering method. However, using embeddings massively decreases the energy demands, assuming appropriate modern hardware and Graphical Processing Units (GPU) [18], and it is negligible like other regular household appliances being used at the agency.

7.3 Social Impact

The survey results showcase a general acceptance of AI at Trafikverket, seen in Appendix 3.2. Most respondents already consider the effects of implementing more AI tools as improved collaboration and data handling, seen in Question A.5, and

most had positive thoughts surrounding AI's future impact on their work, in question A.7. Few respondents also considered the AI's job-stealing controversy as a challenge at the agency seen in Question A.3.
Jesper Lundberg was also positive regarding the tool and would have appreciated the tool if developed and deployed.

As an agency, it is also important to consider Sweden citizens. Through a study performed at Gothenburg University, they established that the Swedish population is rather split when it comes to whether AI is a risk or a potential benefit, and there is a lack of knowledge in general, with several entries answering "I don't know" for several of the questions [43].
This further shows how narrow, simpler AI tools such as OTVK1 are a better move to begin with, before deploying broader AIs. It would be beneficial to both employees at Trafikverket alongside the public that the agency serves.

8
Discussion

A key component of the product development process described in Section 2.1 is reflecting on the results in every part of the process. The reflection is meant to address the execution of the methodology and the result itself, and is done after the sub-process to catch any faults in the approach and solve problems early, while they are more convenient and usually cheaper than later down the line [20].
This chapter also includes general discussion topics, future work/research, alongside discussing the research question introduced in Section 1.3.

8.1 Identifying User Needs

While meetings and interviews were intended methods of collecting qualitative data on Trafikverket's user needs, it ended up being very insightful to also acquire the quantitative opinions of several agency employees through the survey.
The survey was not intended because of concerns about time constraints. Considering the ease of collecting responses and the speed of creating the survey in collaboration with the other master's thesis group at Trafikverket, it certainly should not have been underestimated or excluded from the planning, something to consider for future projects.

The resulting Trafikverket needs list, seen in Table 3.2, did not go through any review. This makes the relative importance entirely subjective, based on the perceived importance from the meetings, interviews, and surveys. This could have been improved through some further follow-up with Trafikverket. As for now, that relative importance should be taken with a grain of salt, even if it has some merit.

8.2 Opportunity Identification

With the bulk of the opportunities being brought up during meetings and interviews, there was less room for external brainstorming and idea generation. Perhaps a more planned and structured approach would be preferable, creating more opportunities. But considering the limited knowledge of the agency's exact tasks and responsibilities, the approach of addressing opportunities throughout meetings and interviews as they appeared felt more effective.
As for the opportunities themselves, there could have been an entire project dedicating more time to researching Trafikverket's tasks and employees' daily challenges across the agency, and identifying significantly more data science and AI tool opportunities.

As opposed to this project, dedicating time to the continued development of some screened opportunities. The list of opportunities, seen in Table 4.1, can thus work as groundwork for such a comprehensive study of more potential tools and opportunities in the future.

8.3 Screening of Opportunities

Letting interviewees be involved with the screening through the ranking was intended to be an effective way to get an objective screening while also including the users in the selection process. Unfortunately, fewer than intended were able to attend the accompanying presentation on the opportunities for the rankings. Sending a recorded presentation did not work for everyone as well, meaning less that there were fewer entries than desired.
Luckily, the rankings were quite unified, giving the placements some merit, especially for the most appreciated OTVK1, Samråds clustering model. To increase the validity of the rankings, the opportunities were compared to the needs list (Table 3.2) to showcase the likely unfilled needs and justify their placements somewhat, as seen in Table 4.2. The results of which were enough to move on with the project.

8.4 Prototyping

Unlike physical prototyping, which requires careful planning and designing, software is easier to start coding and learning right away. Some time was spent testing different clustering models, like agglomerative clustering, but unfortunately, it was not tested on more advanced clustering LLMs.
LLMs like OpenAI's ChatGPT can cluster with higher accuracy than traditional clustering methods, whilst retaining context. However, clustering requires a significant number of calls on the model and ends up being very inefficient as described in Section 7.2. However, unknown at the start of prototyping, there are now LLMs specifically meant for clustering that have the same benefits as other LLMs, whilst still being efficient through reducing the number of calls. These models should have been considered for this prototype, especially since the scale of OTVK1 is not that large to begin with, justifying the choice of even normal LLMs.

8.5 Future Work

For the Samråds clustering model (OTVK1), there are still some functions from the function block diagram (Figure 5.1) that are not yet prototyped/developed:

- How the tool interacts and integrates into Pluto, Trafikverket's database for public consultation.
- Creating a summarized reading for an entire topic/cluster.

Importantly, the current prototype clusters handmade insights as opposed to actual insights from public consultations. A potential challenge when clustering real insights would be that they could include several topics within them, addressing noise and

public transport access, for instance. The solution to which could be dissecting the insights into their different topics, or using a more advanced LLM that can recognize the multiple topics within a single insight and deal with it that way.
While OTKV1 was selected for prototyping and deployment assessment, the other opportunities that went through screening would all be excellent considerations for similar prototyping. Especially the AI Ordering Aid (OEX1), as Trafikverket is a contracting authority.
For the general deployment of any data science and AI tool at the agency, there is a wish for improved internal communication surrounding the use and purpose of these tools at Trafikverket. This would, according to survey results seen in Appendix A, improve trust of these tools and would mean a more effective deployment overall. Something Trafikverket should be well aware of.

8.6 Research Question

As presented in the project definition, Section 1.3, there is somewhat of a research question. Is the product development process a valid approach to researching and developing data science and AI tools at an organization, compared to CRISP-DM [19] and other traditional data science processes?
CRISP-DM is an industry standard for new data science projects. However, it starts with a business objective derived from a domain expert that establishes what the tool should inevitably accomplish and what is deemed a successful project. This requires extensive knowledge of the business domain and application, but also some knowledge of what data science is capable of, so as not to set unrealistic objectives or expectations. Unless an organization invests in that sort of expertise, it would mean having extensive conversations to create the necessary knowledge and derive an appropriate business objective.
In this project, there was no such business objective, and creating one would have been difficult. Treating the "domain experts" (employees at Trafikverket) as users and researching their needs, developing a tool fit to meet them, felt more appropriate for this scenario. Considering the results of this project, there is now the necessary groundwork to consider further development of some of the opportunities, utilizing the CRISP-DM process for those projects.
With computational hardware becoming both cheaper and more powerful, any organization needs to consider how data science could benefit them. Performing an opportunity identification and investigating the needs has been proven to be an effective way in doing so, at least in this project. Then, screening and pursuing great opportunities should yield both some effective data science tools, alongside some more groundwork for more work down the line.
Important to mention is the possibility of a hybrid process between product development and traditional data science processes, perhaps treating the user needs research and opportunity identification as the business objective, moving on with the best opportunity and CRISP-DM.

9
Conclusion

AI is undeniably revolutionizing modern industry, and that Trafikverket is investigating different ways of implementing it is not surprising. As a master's thesis in product development, this project was meant to research and develop data science and AI tools at Trafikverket through a product development process. The research question was whether this process is appropriate for data science and AI tools, compared to the typical data science and AI processes, such as CRISP-DM.
Through interviews, meetings, and a survey, data was collected over Trafikverket's needs for data science and AI tools. The data was interpreted into takeaways and later assembled into a Trafikverket needs list (Table 3.2, a key deliverable of the project that was used for screening in and could be used by Trafikverket in the future when considering data science and AI tools.
The conversations, alongside some further brainstorming, led to a list of opportunities (Table 4.1). These opportunities had vastly different purposes and were screened through both the needs list and a ranking. The screened opportunities were those of an AI contract aid for Trafikverkets many contracts with consultant companies and contractors (OEX1), a secured AI assistant with no possible leakage of sensitive data (OTVK7), an AI document control replacing sample control for quality assurance over the many documents produced (OTVK2), and the Samråds clustering model (OTVK1). The latter was most appreciated and went on with continued development and prototyping.
The prototyped clustering model is meant to showcase the potential function and increase the knowledge of this concept. BERTopics was used as the model while handwritten insights were tested for clustering results. This opportunity also had a deployment assessment of its economic, environmental, and social impact, which all point toward an overall positive result.
The research question was answered with the project results: the product development process was not only appropriate, but had some advantages over traditional data science processes as well. In this case of a new technology becoming accessible, having knowledge of what to do with it or what needs there are for it is important before creating a business objective, something traditional data science processes lack.
While this is the end for this thesis, there is still future work for the Samråds clustering model, alongside the other screened opportunities like the ordering aid. And for an effective implementation of future data science and AI tools at Trafikverket, the agency should also improve internal communication surrounding AI's role in their work.

Bibliography

[1] A. B. Rashid and M. A. K. Kausik, "Ai revolutionizing industries worldwide: A comprehensive overview of its diverse applications," *Hybrid Advances*, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2773207X24001386.

[2] Trafikverket. "Ai-satsning på trafikverket." (Dec. 2024), [Online]. Available: https://www.youtube.com/watch?v=mitX23oh_8Y.

[3] Trafikverket, "Work With Us". [Online]. Available: https://www.trafikverket.se/om-oss/jobba-hos-oss/.

[4] Trafikverket, "About Us". [Online]. Available: https://www.trafikverket.se/om-oss/.

[5] Trafikverket, "New Railway Gothenburg-Borås and Branch-Line Mölnycke". [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-i-vastra-gotalands-lan/Goteborg-Boras/.

[6] S. Riksdag, *Lag (1995:1649) om byggande av järnväg*, Accessed: 3-Feb-2025, 1995. [Online]. Available: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/lag-19951649-om-byggande-av-jarnvag_sfs-1995-1649.

[7] Trafikverket, "That's why planning a railway takes time – planning is a democratic process. [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-i-vastra-gotalands-lan/Goteborg-Boras/darfor-tar-det-tid-att-planera-en-jarnvag/.

[8] Swedish Parliament, *Environmental Code (Miljöbalken, SFS 1998:808)*, https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/miljobalk-1998808_sfs-1998-808/, Available at https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/miljobalk-1998808_sfs-1998-808/, 1998.

[9] Trafikverket, "Västlänken". [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-i-vastra-gotalands-lan/vastlanken/.

[10] Wikipedia contributors, *Data science*, [Online; accessed 2-March-2025], 2025. [Online]. Available: https://en.wikipedia.org/wiki/Data_science.

[11] Wikipedia contributors, *Artificial Intelligence*, [Online; accessed 2-March-2025], 2025. [Online]. Available: https://en.wikipedia.org/wiki/Artificial_intelligence.

[12] Wikipedia contributors, *Large language model – Wikipedia, the free encyclopedia*, https://en.wikipedia.org/wiki/Large_language_model, Accessed: 2025-06-05, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Large_language_model.

[13] OpenAI, *Sora: Text-to-video generation model*, https://openai.com/sora, Accessed: 2025-06-05, 2024.

[14] OpenAI, *Chatgpt: Optimizing language models for dialogue*, Accessed: 2025-06-10, 2022. [Online]. Available: https://openai.com/blog/chatgpt.

[15] G. DeepMind, *Introducing gemini: Our most capable and general ai models yet*, Accessed: 2025-06-10, 2023. [Online]. Available: https://deepmind.google/technologies/gemini/.

[16] United Nations Environment Programme, "Ai has an environmental problem. here's what the world can do about that," *UNEP News and Stories*, Sep. 2024, Available: https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.

[17] D. K. Yadav, A. Kaushik, and N. Yadav, "Predicting machine failures using machine learning and deep learning algorithms," *Sustainable Manufacturing and Service Economics*, vol. 3, p. 100 029, 2024, Accessed: 2025-06-10. DOI: 10.1016/j.smse.2024.100029. [Online]. Available: https://doi.org/10.1016/j.smse.2024.100029.

[18] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, 2019, pp. 3645–3650. DOI: 10.18653/v1/P19-1355. [Online]. Available: https://aclanthology.org/P19-1355/.

[19] Data Science PM, *What is crisp dm?* Accessed: 2025-05-19, 2025. [Online]. Available: https://www.datascience-pm.com/crisp-dm-2/.

[20] K. T. Ulrich, S. D. Eppinger, and M. C. Yang, *Product Design and Development*, 7th ed. New York: McGraw-Hill, 2020.

[21] Chalmers University of Technology, *Produktplanering – behov och möjligheter (ppu085)*, https://www.chalmers.se/utbildning/dina-studier/hitta-kurs-och-programplaner/kursplaner/PPU085/?acYear=2024/2025, Accessed: 2025-06-05, 2024.

[22] Chalmers University of Technology, *Projekt i produktutveckling (mpp126) – product development project*, https://www.chalmers.se/utbildning/dina-studier/hitta-kurs-och-programplaner/kursplaner/MPP126/?acYear=2024/2025, Accessed: 2025-06-05, 2024.

[23] Yolean, *Yolean - visual planning and lean collaboration*, Accessed: 2025-03-10, 2025. [Online]. Available: https://www.yolean.com/start.

[24] Trafikverket, *Handläggningssystemen pluto och geoparden*, Accessed: 2025-03-20, 2025. [Online]. Available: https://bransch.trafikverket.se/tjanster/system-och-verktyg/projekthantering/handlaggningssystemen-pluto-och-geoparden/.

[25] Microsoft, *Microsoft copilot*, Accessed: 2025-06-08, 2024. [Online]. Available: https://www.microsoft.com/en-us/microsoft-copilot.

[26] M. Bargury, *Black hat 2024: Microsoft's ai will soon start leaking user data*, Cybernews article based on Black Hat presentation, Highlights probability of corporate data leakage via Copilot Studio, Aug. 2024.

[27] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based tf-idf procedure," *arXiv preprint arXiv:2203.05794*, 2022.

[28] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[29] KBLab at the National Library of Sweden, *Kblab/sentence-bert-swedish-cased*, https://huggingface.co/KBLab/sentence-bert-swedish-cased, Accessed: 2025-04-17, 2021. [Online]. Available: https://huggingface.co/KBLab/sentence-bert-swedish-cased.

[30] PAIR– Google Research, *Understanding umap*, https://pair-code.github.io/understanding-umap/, Accessed: 2025-06-10, 2024.

[31] scikit-learn developers, *Sklearn.cluster.hdbscan - scikit-learn documentation*, https://scikit-learn.org/stable/modules/generated/sklearn.cluster.HDBSCAN.html, Accessed: 2025-04-17, 2025. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.HDBSCAN.html.

[32] P. M. Dahlgren, *Swedish-stopwords.txt*, Accessed: 2025-04-17.

[33] M. Grootendorst, *Using embeddings in bertopic*, https://maartengr.github.io/BERTopic/getting_started/embeddings/embeddings.html, Accessed: 2025-04-17, 2022. [Online]. Available: https://maartengr.github.io/BERTopic/getting_started/embeddings/embeddings.html.

[34] S. IML, *Understanding the effect of hdbscan parameters*, https://stataiml.com/posts/hdbscan_parameter_differences/, Accessed: 2025-04-17, 2023. [Online]. Available: https://stataiml.com/posts/hdbscan_parameter_differences/.

[35] NVIDIA, *Pandas (python)*, https://www.nvidia.com/en-us/glossary/pandas-python/, Accessed: 2025-06-10, 2025.

[36] Trafikverket, *Ostlänken, en ny dubbelspårig järnväg*, Accessed 2025-05-12, May 2025. [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-som-stracker-sig-over-flera-lan/ostlanken-en-ny-dubbelsparig-jarnvag/#tidsplan.

[37] Trafikverket, *Samråd om Ostlänken, delsträckan Loddby–Butängen, Norrköpings kommun*, Accessed: 2025-05-12, Nov. 2024. [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-som-stracker-sig-over-flera-lan/ostlanken-en-ny-dubbelsparig-jarnvag/nyheter-for-ostlanken/2024/samrad-om-planeringen-av-ostlanken-delstrackan-loddbybutangen-norrkopings-kommun/.

[38] Trafikverket, *Delsträcka Almedal–Mölndal*, Accessed 2025-05-12, May 2025. [Online]. Available: https://www.trafikverket.se/vara-projekt/projekt-i-vastra-gotalands-lan/Goteborg-Boras/delstracka-almedalmolndal/.

[39] Swedish Transport Administration. "Project map – ongoing and planned infrastructure projects in sweden." Accessed: 2025-05-15. (2025), [Online]. Available: https://www.trafikverket.se/karta/?map_x=589280.07467&map_y=6782517.69419&map_z=2&map_l=1000100&Map=24.

[40] Swedish Transport Administration. "Infrastructure projects in sweden – swedish transport administration." Accessed: 2025-05-15. (2025), [Online]. Available: https://www.trafikverket.se/vara-projekt/?project=28%2C25.

[41] K. Thorsell. “Så höga (låga?) är konsultpriserna.” Accessed: 2025-05-15. (2021), [Online]. Available: https://ingenjoren.se/2021/05/20/vissa-konsultpriser-okade-trots-pandemin/.

[42] International Energy Agency, *Energy and ai: Executive summary*, https://www.iea.org/reports/energy-and-ai/executive-summary, Accessed: 2025-05-15, 2025.

[43] M. Weissenbilder, F. Wessman, and A. Bergström, “Den svenska opinionen om artificiell intelligens (ai),” in *Inferno*, U. Andersson, B. Rönnerstrand, and A. Carlander, Eds., Available: SOM-institutet, Göteborgs universitet, Feb. 2024.

A

Appendix - Survey Results

In collaboration with another Chalmers master's thesis group working with AI at Trafikverket, a survey was conducted over employees' thoughts surrounding AI at the agency. The survey was made using Microsoft Forms and was shared in person during lunch breaks with additional QR codes hung up by the coffee machines and the general lunch area. This appendix only includes the questions and responses, with takeaways being addressed in Section 3.2.
The first question was about the position the respondent has at the agency, but since every respondent is anonymous anyway, it does not really matter and is thus not included in this appendix.

Figure A.1: Do you have prior experience using AI tools?

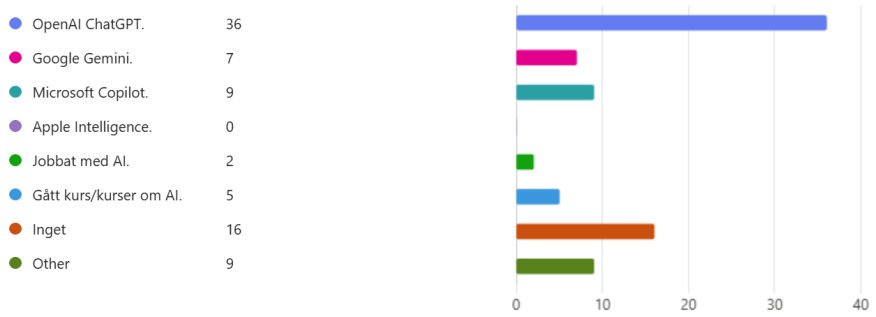
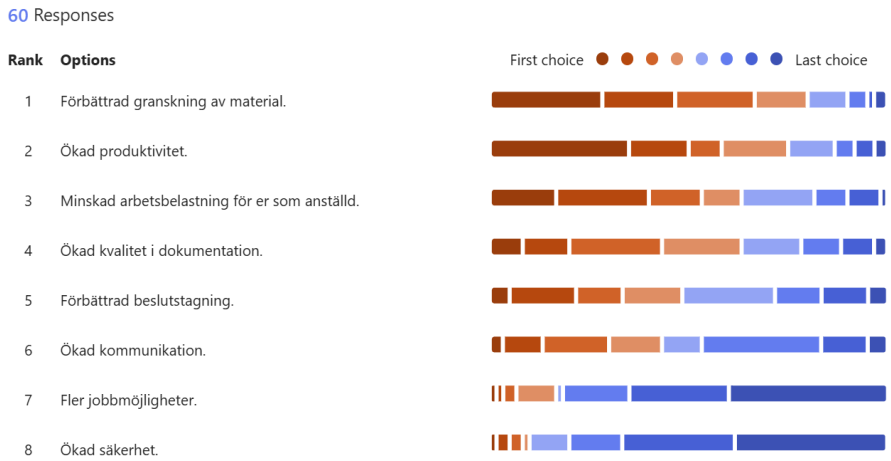


Figure A.2: How familiar are you with Trafikverket's work on AI (1 = have no idea of it, 6 = well aware of it)?

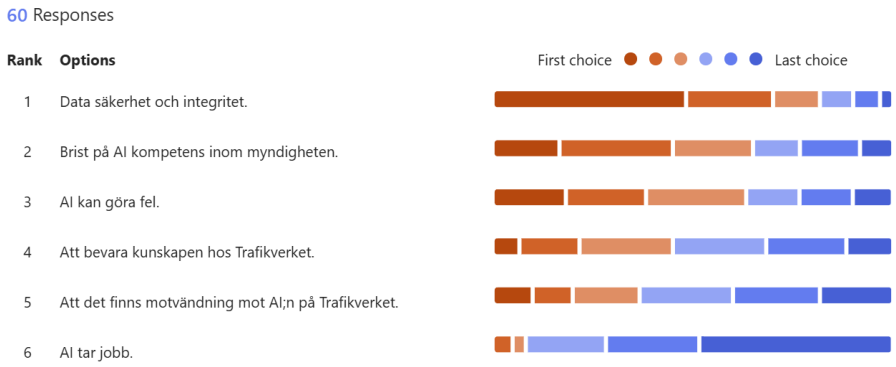
Figure A.3: What do you believe are the largest advantages of AI at Trafikverket?



Options translated as:

1. Improved interviewing of material.
2. Increased productivity.
3. Decreased workload for you as an employee.
4. Increased quality in documentation.
5. Improved decision making.
6. Increased communication.
7. More job opportunities.
8. Increased security.

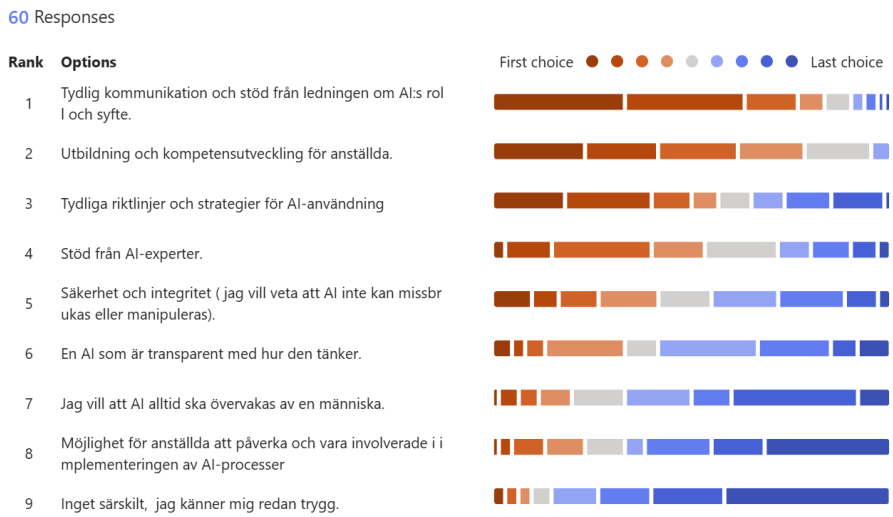
Figure A.4: What do you believe are the largest challenges with AI at Trafikverket?



Options translated as:

1. Data security and integrity.
2. Lacking AI competence at the agency.
3. AI can make mistakes.
4. Retaining the knowledge at Trafikverket.
5. That AI is not largely accepted at Trafikverket.
6. That AI will take jobs.

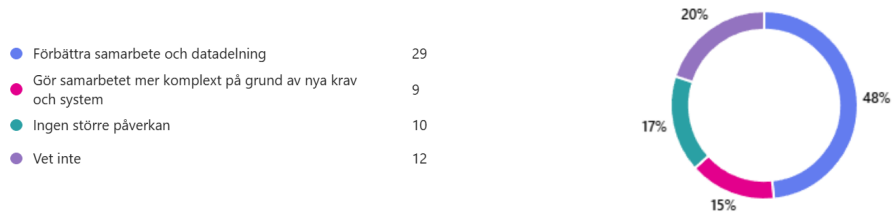
Figure A.5: What do you believe Trafikverket should do as an organization to implement AI effectively and for you to feel comfortable using it?



Options translated as:

1. Clear communication of AI's role and purpose from upper management.
2. Education and competence development for employees.
3. Clear guidelines and strategies for the usage of AI.
4. Security and integrity (I want to know that AI cannot be abused or misused).
5. The AI should be transparent about how it got to its result.
6. I want the AI to always be monitored by a human operator.
7. Opportunity for employees to influence and be involved in AI's application.
8. I already feel comfortable using AI.

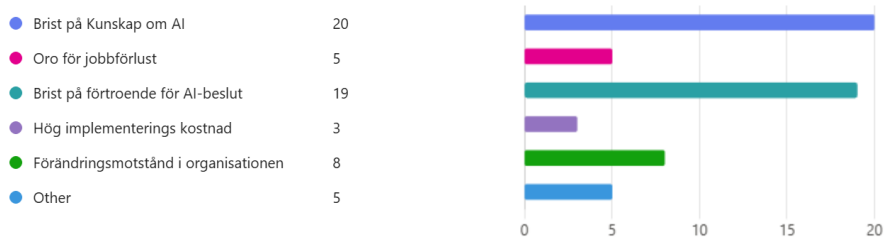
Figure A.6: How do you believe AI will affect cooperation with external partners (consultants, contractors, suppliers)?



Options translated as:

1. Improve cooperation and data sharing.
2. Increase the complexity of cooperation with added requirements and systems.
3. No larger impact on cooperation.
4. Don't know.

Figure A.7: What factors do you believe could create backlash against AI at Trafikverket?



Options translated as:

1. Lacking knowledge of AI.
2. Worry of job losses due to AI.
3. Lacking trust in AI decisions.
4. High deployment costs.
5. Resistance to change at the agency.
6. Other:
    - All of the above, and that AI would risk making us "dumber", since it would further reduce the need of actual knowledge among employees.
    - Risk and Security
    - Security and the management of sensitive data.
    - Security questions.
    - Relevance.

Table A.1: How do you think AI will affect your work tasks and responsibilities in the future?

ID	Answer
1	Reduced administration and time spent on analysis
2	Simplify and improve
3	Provide support, but not replace
4	Hopefully simplify/remove certain tasks
5	Relieve with meeting notes, summaries, and simple text analysis
6	Suggest texts, review for contradictions
7	I hope for a unified set of requirements
8	For the better
9	Hopefully help with time-consuming tasks and serve as a sounding board in specific complex issues
10	I believe I will be able to do parts of my work faster
11	Not much. Hopefully simplify certain aspects, like reviewing, meeting notes, etc.
12	Don't know
13	Increased productivity and reduced dependence on consultants
14	Streamline searching for information (e.g., requirements, documents)
15	I don't think it will significantly affect things
16	Don't know
17	It can streamline information searching and reviewing
18	Simplify data coordination/handling and help verify requirements and documentation
19	Facilitate work if I use it correctly
20	Higher quality in my deliveries
21	Might take over some of the tasks
22	Would simplify most of the work and allow more to be done in less time
23	Streamline things like meeting notes
24	Function as a sounding board to speed up the startup process
25	Positive. Simplify reviewing texts, generate text suggestions, etc. AI would be good support in my role
26	I think it can help and simplify, for example, in reviews
27	Not really sure
28	It will be easier to collect and compile data, which provides a better basis for decisions. It could also simplify and save time in tasks like protocols and report writing. Plus countless other areas
29	It entirely depends on the type of LLM implemented. Current LLMs are very limited in interpreting documents that aren't clearly standardized and accessible to the LLM
30	I think AI will make my work tasks easier and help me become more efficient

ID	Answer
31	Easier to search for information and facts. Easier to get suggestions for solutions. Better visualizations of what is to be built. Tasks become more automated
32	Help with document handling, etc.
33	Facilitates information gathering
34	Can create more facts faster before making decisions
35	Not at all, will not exist in the Swedish Transport Administration when implemented
36	-
37	Don't know
38	Contribute to increased productivity
39	Increase productivity
40	Not very familiar with AI or how it would affect things. Maybe AI can help compile information flows, give better summaries for decision-making
41	Marginally
42	Quick access to knowledge
43	Don't know
44	No, my job is about collaboration between contracting parties, i.e., a lot of psychology
45	Eliminates the initial review of deliveries
46	Risk of broader and more scattered areas of responsibility when fewer people are expected to do more due to AI
47	Standardize away the simple and time-consuming tasks, let us focus on work that increases quality
48	If a secure system can be developed where information from the Swedish Transport Administration cannot leak, I would benefit greatly from AI. E.g., text suggestions, help summarizing meetings into minutes (still needs to be checked), etc.
49	Streamline and improve!
50	Drastically streamline! So many tasks could be made more efficient
51	More than I currently believe. (Right now I don't think it will affect my work)
52	Faster and more reviews with fewer errors
53	More source verification
54	Support in work
55	Streamline e.g., the "question-answer" work when AI can handle it. Streamline the review process immensely. Also suggest technical solutions. I think this will be a shift for my role, becoming more of an overseer where AI is almost like a resource in itself, producing materials and relieving many tasks, which are then reviewed by humans
56	I believe AI can be a helpful tool if used in an organizational and controlled way. Using AI in combination with one's own knowledge gives a positive result

ID	Answer
57	Hopefully, AI can take over the review of designers' and contractors' deliveries against our requirements and regulations, and also answer questions about our rules. I also hope AI can go through all our regulations to find duplicates and contradictions
58	Positive
59	AI will be able to draft reports, suggest template texts, automate reference collection, check links. Possibly even assess whether content complies with laws, regulations, and standards
60	Will have a big impact as support and sounding board in decision-making, as well as handling some admin and communication

Table A.2: Do you have any further thoughts on how AI could impact Trafikverket as an organization?

ID	Answer
1	See great potential if done correctly. Smaller pilot infrastructure parts could serve as good examples
2	More efficient ways of working
3	No
4	No
5	Don't know
6	Streamline work and reduce workload
7	No
8	Hope it will be implemented
9	I think it can help and simplify for many, but I'm concerned that the use of AI might reduce transparency and traceability in decisions
10	No, haven't really thought about it
11	Everything depends on the strategy the Swedish Transport Administration chooses for data handling going forward. Will all data be controlled and managed by the agency on fully-owned servers, or will it happen on external servers? If the AI tool is an external service, can we then ensure that data doesn't leak if the chosen strategy is internal-only handling?
12	AI usage must be done with critical skepticism due to security concerns
13	Risk that independent thinking, innovation, etc., will be eroded
14	Increase efficiency
15	In the Transport Administration, all intelligence is probably welcome – artificial or not
16	-
17	A large budget item with little real impact that takes funding from other initiatives
18	Given the generally low IT competence in the part of the organization I work in, a lot of training will probably be needed. Some people are currently struggling just to convert a PDF file

ID	Answer
19	Again: streamline our long and drawn-out processes
20	can be used to simplify processes and dependencies. Scan contradictory technical documents
21	At present, I think many of the Swedish Transport Administration's suppliers are already using AI, for example in tendering and when producing documentation for us. We must keep up to avoid falling behind – there is no alternative. The administration must issue clear guidelines so employees don't start experimenting outside approved systems. I also believe AI could massively streamline our approval and review processes
22	Hopefully it can reduce the scope of our organization
23	-
24	Hopefully, AI can reduce administration and thereby reduce the need for non-technical staff (and associated costs), allowing a return to the agency's core mission – executing, maintaining, and managing infrastructure!
25	I believe the Swedish Transport Administration will be slower than the industry in implementation, and that security concerns will play a major role in how/when/if the implementation happens

Department of Industrial and Material Science
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY