Bachelor's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology

Emil Farstad Bjerknes
Paul Røkke Bjørneng
Torje Grimstad Bliksvær

AI Legal Assistant: Enhancing GDPR Compliance with Large Language Models

Bachelor's thesis in Digital Infrastructure and Cyber Security
Supervisor: Shao-Fang Wen
Co-supervisor: Muhammad Mudassar Yamin

May 2024

NTNU
Norwegian University of
Science and Technology

Emil Farstad Bjerknes
Paul Røkke Bjørneng
Torje Grimstad Bliksvær

AI Legal Assistant: Enhancing GDPR Compliance with Large Language Models

Bachelor’s thesis in Digital Infrastructure and Cyber Security
Supervisor: Shao-Fang Wen
Co-supervisor: Muhammad Mudassar Yamin
May 2024

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Dept. of Information Security and Communication Technology

NTNU
Norwegian University of
Science and Technology

AI Legal Assistant: Enhancing GDPR Compliance with Large Language Models

Emil Farstad Bjerknes, Paul Røkke Bjørneng, Torje Grimstad Bliksvær

May, 2024

Abstract

Title: AI Legal Assistant: Enhancing GDPR Compliance with Large Language Models

Date: 19th May 2024

Participants: Emil Farstad Bjerknes
Paul Røkke Bjørneng
Torje Grimstad Bliksvær

Supervisors: Shao-Fang Wen
Muhammad Mudassar Yamin

Employer: Futurize AS
Keywords: AI, LLM, GDPR, RAG
Number of Pages: 72
Number of Appendices: 9
Availability: Closed for two years

Abstract:
Futurize AS is a startup company specializing in developing innovative solutions for the legal business by integrating AI into their workflow. They are looking for new insight and perspective on the application of AI within the legal field. In response, this project builds on developing an AI legal assistant specialized in the General Data Protection Regulation (GDPR). This tool is designed to provide tailored support for GDPR compliance, offering solutions that address the specific needs of legal professionals. The AI assistant use a pre-trained Large Language Model for Question Answering, with the aid of Retrieval-Augmented Generation to ensure that the generated advise is correct. The findings provide Futurize with insight on how to develop an AI GDPR assistant for the legal field.

Sammendrag

Tittel: AI Legal Assistant: Enhancing GDPR
Compliance with Large Language Models

Dato: 19th May 2024

Deltakere: Emil Farstad Bjerknes
Paul Røkke Bjørneng
Torje Grimstad Bliksvær

Veiledere: Shao-Fang Wen
Muhammad Mudassar Yamin

Oppdragsgiver: Futurize AS
Nøkkelord: AI, LLM, GDPR, RAG
Antall Sider: 72
Antall Vedlegg: 9
Tilgjengelighet: Lukket i to år

Sammendrag:
Futurize AS er et oppstartsselskap som spesialiserer seg i utviklingen av innovative løsninger for den juridiske bransjen gjennom integrasjon av KI i arbeidsflyten. De søker ny innsikt og perspektiv på utviklingen av AI innen jus. I respons, bygger dette prosjektet på å utvikle en KI juridisk assistent spesialisert på personvernforordningen (GDPR). Dette verktøyet er laget for å gi skreddersydd støtte for GDPR-samsvar, og tilbyr løsninger som adresserer de spesifikke behovene for juridiske fagfolk. KI-assistenten bruker en forhåndstrent Large Language Model for spørsmålbesvarelse, med hjelp av Retrieval-Augmented Generation for å sikre korrekt rådgivning. Funnene gir Futurize innsikt i hvordan man kan utvikle en KI GDPR assistent for den juridiske bransjen.

Preface

We would like to thank our supervisors, Shao-Fang Wen and Muhammad Mudassar Yamin, for their guidance and insightful feedback throughout this project. Their knowledge and constant feedback has been invaluable and much appreciated.

We also extend our gratitude to Daniel Arlien and Futurize AS for their cooperation and assistance over the semester. Daniel's flexibility and willingness to provide information and answer questions have greatly helped us.

Finally, a special thanks to Dr. Mohamed Abomhara for his assistance with the evaluation part of our thesis and for sharing his GDPR expertise.

Acronyms

AI Artificial Intelligence
LLM(s) Large Language Model(s)
SLM(s) Small Language Model(s)
RAG Retrieval-Augmented Generation
MQR Multi-Query RAG
HyDE Hypothetical Document Embedding
LoRA Low-Rank Adaptation
GDPR General Data Protection Regulation
QA Question Answering
GPT Generative Pre-trained Transformer
DSR Design Science Research
DB Database
VM Virtual Machine
IaaS Infrastructure as a Service

Contents

Abstract . . . . **iii**

Sammendrag . . . . **v**

Preface . . . . **vii**

Acronyms . . . . **ix**

Contents . . . . **xi**

Figures . . . . **xvii**

Tables . . . . **xix**

Code Listings . . . . **xxi**

1 Introduction **1**

- 1.1 Background 1
  - 1.1.1 Futurize AS 1
  - 1.1.2 Application of AI in Legal Work 2
- 1.2 Problem Area 2
- 1.3 Scope 3
- 1.4 Limitation 3
- 1.5 Target Group 4
- 1.6 Goals 5
  - 1.6.1 Result Goals 5
  - 1.6.2 Learning Goals 5

1.6.3 Impact Goals 6
1.7 Group Background 6
1.7.1 What had to be learned 6
1.7.2 Why the task was chosen 6
1.8 Framework 6
1.9 Thesis Structure 7
2 Theory 9
2.1 Artificial Intelligence (AI) 9
2.1.1 Machine Learning vs Deep Learning 9
2.1.2 Generative AI 12
2.1.3 The Transformer Architecture 12
2.1.4 Natural Language Processing 14
2.2 Large Language Models (LLMs) 14
2.2.1 Conversational AI: QA-models vs Chatbots 16
2.3 Legal Informatics 16
2.3.1 Legal Language 17
2.4 General Data Protection Regulation (GDPR) 17
2.4.1 Legal vs Technical Perspective 18
3 Methodology 19
3.1 Development Methodology 19
3.1.1 Design Science Research (DSR) 19
3.1.2 Agile Software Development 21
3.1.3 Integration of DSR and Scrum 21
3.1.4 Using Scrum 22
3.2 Large Language Model Selection 25

3.3 Development Technologies 26
3.3.1 Retrieval-Augmented Generation (RAG) 26
3.3.2 Vector Database 28
3.3.3 Virtual Environments 29
3.3.4 Other Technologies 29
3.4 Development Environment 30

4 Implementation **31**
4.1 Setting Up the Virtual Infrastructure 31
4.2 Virtual Environment for Development 31
4.3 Program Architecture 32
4.4 RAG Configuration 33
4.4.1 Database 33
4.4.2 Retrieval 34
4.4.3 Overcoming Shortcomings 34
4.5 Prompting the Model 36

5 Results and Evaluation **39**
5.1 Evaluation Strategy 39
5.2 Preliminary Evaluation 40
5.2.1 Test Results and Evaluation 41
5.3 Expert Evaluation 44
5.3.1 Expert Evaluation Process 44
5.3.2 Scoring System 44
5.3.3 Evaluation Method 47
5.3.4 Evaluation Results 47

6 Discussion **51**

6.1 Results 51
6.1.1 Preliminary Evaluation Discussion 51
6.1.2 Expert Evaluation Discussion 52
6.2 Theory and Implementation 54
6.2.1 Constraints 54
6.2.2 Why RAG and not Fine-tune? 55
6.2.3 Potential Fine-tuning Methods 56
6.2.4 Why Only One PDF Document? 57
6.2.5 Practical Limitations 60
6.3 The Workflow of the Group 60
6.3.1 Delegation of Responsibilities 60
6.3.2 Gantt and Project Development 61
6.3.3 Meetings 61
6.3.4 Decision Making Process 62
6.3.5 Submission of Drafts 62
6.4 Project Reflections 63
6.4.1 Sustainability 63
6.4.2 Ethical Considerations 64
6.4.3 Use of AI 67
6.4.4 Critique of the Thesis 67
7 Conclusion 69
7.1 Achievement of Goals 69
7.2 Contribution of the Study 71
7.3 Future Work 71
Bibliography 73

A Project Plan **79**

B Standard Agreement **101**

C ChatGPT Answers **109**

C.1 ChatGPT Answers Used In Evaluation 112

D Evaluation Form **121**

E Program Answers for Evaluation **135**

F GDPR QA Dataset for Testing **153**

G Timetables **157**

H Meeting Minutes **165**

I GDPR Expert Scenarios **183**

Figures

2.1 Deep neural network used by Deep Learning algorithms [3] 10

2.2 Generative AI and the AI subcategories [7] 12

2.3 The Encoder-Decoder structure of the transformer model architecture [8] 13

2.4 The structure of encoder-decoder language models [15] 15

3.1 Model of the DSR process [25] 20

3.2 Example of using the Scrum board in Jira 22

3.3 The three paradigms of RAG [29] 27

4.1 Figure of the program architecture 32

5.1 Example of documents retrieved from the context 41

5.2 Query and response from the AI assistant 41

5.3 Where the AI assistant used the context 42

5.4 AI assistants response to Q1 42

5.5 AI assistants response to Q2 43

5.6 Scenario 1 from the expert evaluation 48

6.1 Taxonomy of parameter efficient fine-tuning methods. [42] 57

6.2 Comparisons of SLM performance on multiple benchmarks[49] . . 58

6.3 The Gantt chart that was made during the planning of the project . 62

6.4 Environmental impact of LLM training compared to other activities. [48] 64

6.5 Environmental impact of ChatGPT interaction. [56] 65

C.1 Initial GPT provided QA-dataset (in Norwegian) 110

C.2 Scoring system example for the program evaluation 111

C.3 ChatGPT Expert evaluation scenario 1 answer pt. 1 112

C.4 ChatGPT Expert evaluation scenario 1 answer pt. 2 113

C.5 ChatGPT Expert evaluation scenario 1 answer pt. 3 114

C.6 ChatGPT Expert evaluation scenario 2 answer pt. 1 115

C.7 ChatGPT Expert evaluation scenario 2 answer pt. 2 115

C.8 ChatGPT Expert evaluation scenario 3 answer pt. 1 116

C.9 ChatGPT Expert evaluation scenario 3 answer pt. 2 117

C.10 ChatGPT Expert evaluation scenario 4 answer pt. 1 118

C.11 ChatGPT Expert evaluation scenario 4 answer pt. 2 119

C.12 ChatGPT Expert evaluation scenario 4 answer pt. 3 119

Tables

3.1 Two-week Sprint Schedule 23
3.2 Technologies Used in Our Program 29

5.1 Scoring for Accuracy 45
5.2 Scoring for Relevance 46
5.3 Scoring for Completeness 46
5.4 Scoring for Clarity 46
5.5 Expert evaluation of our program's different versions, compared to GPT-4. 50

Code Listings

4.1 Setting up the vector database 33
4.2 Basic RAG implementation in LangChain 34
4.3 Multi-query RAG implementation 35
4.4 Hypothetical Document Embedder implementation 36
4.5 Prompting the model 37

Chapter 1

Introduction

This chapter introduces broadly how, and what, time has been spent on in the duration of this project. Specifically, the problem and scope is presented, as well as the group's background, goals and project structure.

1.1 Background

1.1.1 Futurize AS

Futurize AS is a startup company specializing in the development of innovative solutions for the Norwegian legal business. Their primary focus is to explore the use of artificial intelligence (AI) to streamline and enhance the workflow of lawyers, with a particular emphasis on legal research as well as the preparation of documents and texts [1]. By integrating AI into their daily tasks, Futurize has the potential to not only revolutionize traditional legal practices but also introduce a transformative shift within the whole legal sector. They aim to redefine the conventional approaches of legal work, especially by making processes such as legal research and document preparation more efficient and accurate. Through this convergence of AI and legal expertise, Futurize strives to improve both the efficiency and accuracy of lawyers' workflows, ultimately benefiting the entire legal landscape.

1.1.2 Application of AI in Legal Work

In pursuit of this transformation, Futurize focuses on the use of generative AI, specifically utilizing Large Language Models (LLMs) capable of generating high-quality text, images, and other content from large datasets [2]. This technology aims to significantly automate and refine the management of legal documents, and boost productivity as well as ensuring greater precision in legal practice. Generative AI could swiftly process and analyze vast amounts of legal data, and by the use of an AI legal assistant, lawyers could reduce their time spent on routine tasks regarding legal documents, and rather focus on client interaction and legal reasoning.

1.2 Problem Area

The legal business is one of many that stand to gain from introducing AI and LLMs specifically into their workflow, not only for streamlining purposes, but also to ensure compliance with extensive and complex legal documents and regulations such as the General Data Protection Regulation (GDPR). As regulations like GDPR grows in complexity, law firms and lawyers might face challenges in keeping up with it and maintaining compliance. By developing an AI GDPR assistant, Futurize could help the legal environment and lawyers to handle these challenges by integrating AI into their workflow, providing precise and contextual support when working with clients.

For us to be able to help Futurize offer such a service for lawyers, there are a few problems that needs to be addressed:

- The AI assistant needs to provide accurate and precise output and responses to avoid any ambiguities that could lead to missteps considering the complex nature of the GDPR.

- The AI assistant should be able to provide GDPR assistance using and understanding the Norwegian language. This is an issue given that most LLMs are trained in English.

- Ethical considerations and potential biases in AI systems must be paid close attention to, when integrating AI into the legal field and GDPR.

1.3 Scope

This thesis uses recent advancements of LLMs and Retrieval-Augmented Generation (RAG) as a baseline to develop an AI legal assistant for the legal industry, with requisite knowledge of GDPR as well as a basic understanding of Norwegian legal language. The objective of this thesis is to write a report on the process of producing this assistant. Through comparing the performances of various models, and model settings, a single program has been proposed. The optimization of this program is mainly achieved through the use of RAG, and other information-retrieval related technologies. The finished program is able to handle relatively simple close-ended questions, as well as more complex tasks, all related to GDPR specific cases or scenarios.

- **Domain scope:** The AI legal assistant is specifically designed for the legal business, more specifically for lawyers requiring assistance in navigating and complying with GDPR regulations. This tool aims to increase lawyers' productivity by efficiently assisting in GDPR-related cases.

- **Technology scope:** The core technology used by the AI assistant involves utilizing LLMs, further enhanced with RAG and advanced RAG technologies for better information retrieval and understanding. The deployment platform as of now is on a virtual machine in Openstack provided by NTNU (SkyHiGh).

- **Functional scope:** The functionality of the AI assistant includes responding to straightforward, close-ended questions as well as handling more complex tasks that involve deeper legal analysis related to GDPR scenarios. It will handle queries related to the whole GDPR document, chapter 1 - 11, and dynamically retrieve and generate relevant information based on this.

1.4 Limitation

With optimization of a domain specific LLM in mind, the group uses pre-trained models as a baseline for further experimentation. No novel technologies has been developed, as the group have rather sought to implement pre-existing technologies. This has been done for the purpose of providing guidance and insight into the possibilities of developing a working legal assistant. Furthermore, the thesis is mainly focused on the technical aspects of producing an LLM, so the exploration of user design, user interaction methodologies and user satisfaction is very limited.

The aim was for the final LLM to be as refined a legal assistant as possible, and therefore most of the allocated time and resources was spent experimenting on, and evaluating, the various implementations. The group also utilized other frameworks and tools to help with implementation, of which are further explained in Chapter 4.

The proposed program is restrained to the specific legal domain of GDPR and legal questions related to data privacy. It is produced to provide accurate interpretations and guidance on GDPR-related queries. Initially the program was meant to handle a much broader scope, but due to time restriction, and for performance evaluation purposes, the scope was decreased. External constraints are discussed further in Section 6.2.1.

The capabilities of the AI legal assistant are confined by:

- **Functionality:** The assistant is not designed as a user-interactive chatbot. Its functionality is limited to processing and responding to specific legal queries related to GDPR one at a time, without engaging in any broader conversational contexts or other areas of law.
- **Resources and budget:** The performance of the AI assistant is very dependent on the resources available during the project. Limited computational power and time will restrict its capabilities regarding accuracy, efficiency and overall performance. We did not spend any money on external services or additional resources that could have enhanced our programs functionality.
- **Text generation limits:** The amount of text that the AI assistant can generate responding to queries is capped by predetermined limits from the LLM used, to ensure relevance of the output.
- **Use cases:** The AI assistant will be used by legal professionals specifically for GDPR compliance tasks, and will not be suitable for other legal applications or for individuals without a legal background.

1.5 Target Group

The principal beneficiary for this thesis is the client Futurize AS. Our findings are mostly pertinent to Futurize as their primary goal is to use the technology of generative AI to streamline the workflow of lawyers, and this study's contributors might assist in providing insights and effective solutions. For this reason, the specific target group is lawyers.

The group's finished report could however also be of potential benefit to researchers and students of generative AI in other fields - not necessarily limited to the specific discipline of law and legal practice - as many other business environments can draw from the benefits of AI as well.

1.6 Goals

For this thesis, the goals are separated into result goals, learning goals, and impact goals. The group's own goals are presented under result goals and learning goals, while some of the client's goals and the potential impact this thesis has, are also included under impact goals.

The overall goal of the project is to develop and evaluate an AI-powered legal assistant that specializes in GDPR. This AI assistant should be able to understand and respond accurately in Norwegian.

1.6.1 Result Goals

1. To produce a comprehensive report on the process of developing a functional GDPR specialized AI assistant.
2. To produce results and ideas that could potentially be iterated by us or others in the future.
3. To assist Futurize AS with insights into LLM modifications, as they aim to develop an easy-to-use solution for future customers.

1.6.2 Learning Goals

1. To acquire knowledge of AI. Specifically about correlated theories.
2. To practice common workflows and methodologies that future employers may expect us to be acquainted with.
3. To work in a highly structured and independent manner.
4. To discern the impact that the continued development of AI will have on the workplace.

1.6.3 Impact Goals

1. To provide new insights in regards to using techniques for specializing LLMs.
2. To provide an easy to use solution for future customers of Futurize.

1.7 Group Background

The group is comprised of three students at the study program Digital Infrastructure and Cyber Security at NTNU in Gjøvik. The program covers a rather broad area of IT, and as a result the group has garnered highly relevant experience via courses including, but not limited to: "Cyber security and Teamwork", "Object-oriented Programming" and "Statistics".

1.7.1 What had to be learned

Specific topics related to Artificial Intelligence and LLMs have not at any point during our three years at NTNU been part of the syllabus. Except for general topics related to the courses listed above, any topic directly related to concepts, and development, of LLMs has been researched independently.
Furthermore, as the world of LLMs is in a state of constant enhancement the group has had to research and keep up with potentially relevant ideas as soon as they are presented.

1.7.2 Why the task was chosen

Already at the initial briefing, where the university presented the different options of tasks, our group singled out this one specifically. Naturally we were aware of the prevalent concepts of AI and LLMs of today, but we were drawn to fact that we nevertheless find ourselves in the midst an AI revolution. Because of its high relevance already, and because of the certain impact that AI will have on the workplace of tomorrow, this task was chosen. Our group wanted to explore this subject as we had little prior hands-on knowledge of it, and we could see AI being a required prerequisite in the future.

1.8 Framework

The report was delivered on the 19th of May 2024. The group had made an agreement with the supervisors to present them with a working draft for them to be able to give feedback. We received continuous feedback on the thesis during the last two weeks of the project.

A thesis defence is to be held on the 5th and 6th of June.

1.9 Thesis Structure

Chapter 1 - Introduction
Contains the introduction to the thesis. Includes the scope, thesis- and group background, as well as the goals.

Chapter 2 - Theory
Contains the theoretical concepts and frameworks that are necessary to grasp the other content of the thesis.

Chapter 3 - Development Process
Contains a description of the methodology that has been used to conduct the study, as well as how the methodology has been used together with the project development model: Scrum.

Chapter 4 - Implementation
Contains an overview of the different components of the program. Includes descriptions of how the technologies work and how they were utilized.

Chapter 5 - Results and Evaluation
Contains both results and evaluations of the AI assistant. Includes an evaluation from a GDPR expert, as well as an internal preliminary evaluation of expected functionality.

Chapter 6 - Discussion
Contains discussions of four aspects of the thesis. The first includes a discussion of the results of the study. The second relates to the theory and implementation, and why some techniques were utilized over others. The third relates to the process of working on the thesis, and challenges that was faced along the way. The fourth is a reflective review of factors influencing AI practice, as well as a critique of the thesis.

Chapter 7 - Conclusion
Contains a review of the goals, how the study contributes to the existing body of knowledge and its implications for the field, and suggestions for future work.

Chapter 2

Theory

The field of LLM development can be complicated, and as the group also realized, the need for sufficient theoretical knowledge on the topic is crucial. This chapter contains the core theoretical concepts essential for understanding the content of this thesis, and the foundational theories that underpin the development and evaluation of the program.

2.1 Artificial Intelligence (AI)

Artificial Intelligence (AI) has become a transformative force across various industries offering unique capabilities in data processing, analysis and decision-making. It is a technology that enables computers and machines to simulate human intelligence and problem-solving capabilities [3]. This includes understanding natural language, recognizing patterns, making decisions, and learning from experience. There are a few core components or subsets of AI that needs to be understood.

2.1.1 Machine Learning vs Deep Learning

Understanding the field of AI and the different layers of complexity it contains, requires knowledge of Machine Learning, which is a sub-set of artificial intelligence. Machine Learning algorithms focuses on building systems that learn from huge amounts of data using neural networks. These neural networks are programmatic structures modeled after the decision-making process of the human brain, and consist of layers of interconnected nodes that extract features from the data [3]. They are designed to recognize patterns and predict what the data signifies.

Machine Learning algorithms use neural networks with an input layer, one or two 'hidden' layers, and an output layer. These algorithms are typically limited to supervised learning which means that the data needs to be structured or labeled by human experts to enable the algorithm to extract features from the data [3].

Deep Learning is a specialized subset of Machine Learning that also employs neural networks with multiple layers, but in a different way than Machine Learning. It could be described as using 'deep' neural networks, as the networks are composed of an input layer, three or more (usually hundreds) hidden layers, and an output layer [3], as illustrated in Figure 2.1. The design of these neural networks is inspired by the neural network of the human brain, and makes the learning process for Deep Learning models far more efficient and able than standard Machine Learning models. These multiple layers enable unsupervised learning: they automate extraction of features from large, unlabeled and unstructured data sets [3]. Therefore, one of the key distinctions of Deep Learning from traditional Machine Learning is its ability to perform this extraction automatically, determining which features are most significant for making predictions.

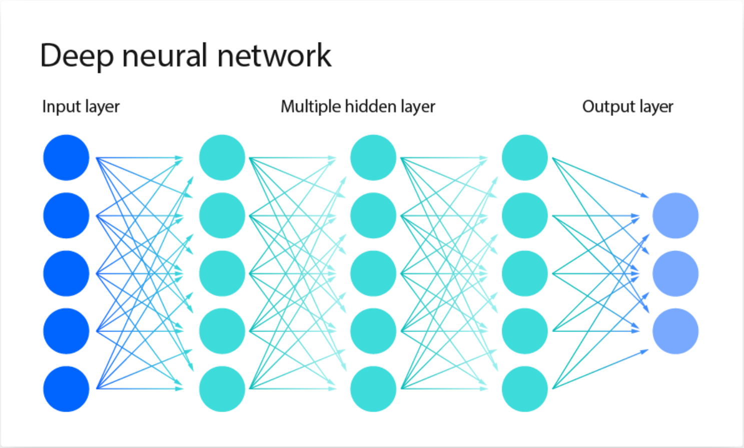


Figure 2.1: Deep neural network used by Deep Learning algorithms [3]

Continuing the exploration of neural networks, there are two specific types of neural networks designed to perform a very different set of tasks: Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).

The main difference between CNNs and RNNs is the ability to process temporal information - data that comes in sequences [4]. Recurrent Neural Networks are designed to handle sequential data such as text and speech, and unlike standard neural networks, they use other data points in a sequence to make better predictions [4]. RNNs maintain internal memory of previous inputs, which helps it use these past data points to influence the predictions it makes about future data. This makes RNNs useful for tasks like translation, spelling correction and speech analysis [4].

Convolutional Neural Networks, on the other hand, are not able to effectively process temporal information, but are known for using filters within convolutional layers to transform data [4]. Primarily used for analyzing visual data, such as images, CNNs excel at recognizing objects and patterns where the features are static and non-temporal [4]. Common use cases for CNNs include image classification and facial recognition [5].

2.1.2 Generative AI

Generative AI refers to deep-learning models that can take raw data and "learn" to generate statistically probable outputs when prompted [3]. It can generate new content, and works by learning from a large dataset and then using this learned information to generate new, but not identical, work similar to the original data [3]. A critical breakthrough in Generative AI has been the implementation of the **transformer architecture**. The transformer model is a type of neural network that can translate text and speech in near-real-time, and is used for prediction, summarizing and question answering [6]. The transformer model does not have to rely on RNNs or CNNs, which have both significant drawbacks, and can process input sequences in parallel, which makes it very efficient for training and inference [6]. Generative AIs' position in the world of AI is illustrated in Figure 2.2.

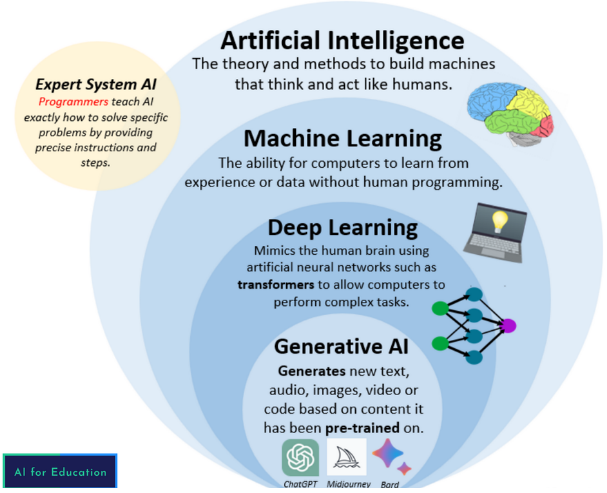

Artificial Intelligence
The theory and methods to build machines that think and act like humans.
Expert System AI
Programmers teach AI exactly how to solve specific problems by providing precise instructions and steps.
Machine Learning
The ability for computers to learn from experience or data without human programming.
Deep Learning
Mimics the human brain using artificial neural networks such as transformers to allow computers to perform complex tasks.
Generative AI
Generates new text, audio, images, video or code based on content it has been pre-trained on.
ChatGPT
Midjourney
Bard
AI for Education

Figure 2.2: Generative AI and the AI subcategories [7]

2.1.3 The Transformer Architecture

The transformer model consists of mainly two parts: the encoder and the decoder, where each part is made up of layers that repeat the same structure. The task of the *encoder* is to take an input sequence of symbols (x1, ..., xn), then mapping them to a continuous sequence of representations z = (z1, ..., zn), which is then fed into the decoder [8] [9]. The *decoder*, in turn, processes this information to generate the output text. Specifically, the decoder works in an auto-regressive manner, using the sequence of continuous representations from the encoder along with previously generated output symbols to produce the final output sequence (y1, ..., ym) one symbol at a time [8]. This ensures that each step of the output generation is "informed" by all preceding steps, thereby enhancing the context relevance of the produced text.

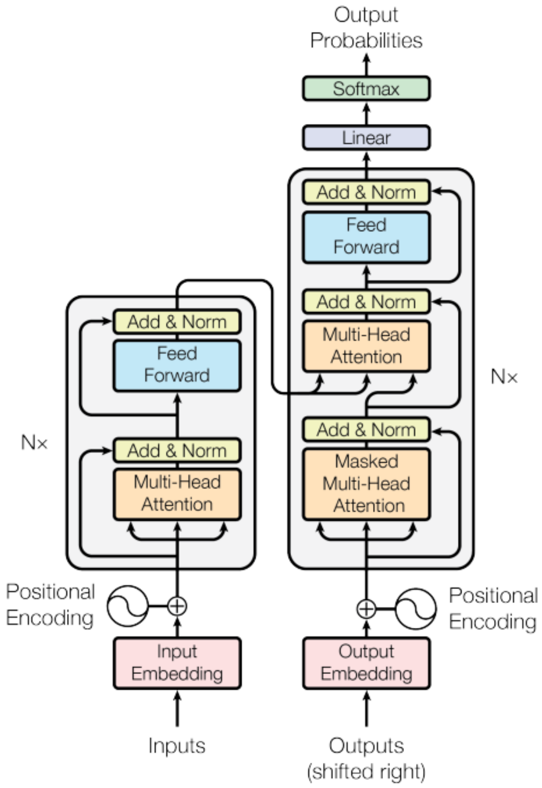


Figure 2.3: The Encoder-Decoder structure of the transformer model architecture [8]

From Figure 2.3, the architecture of the transformer model is shown, using stacked self-attention and point-wise fully connected layers for the encoder and decoder [8]. Self-attention and positional encoding are two key innovations introduced by the transformer model that significantly improve their ability to predict text:

- **Positional encoding:** Assigns a unique number to each word in a sequence, providing information about each token's position, which allows the model to consider the sequence's sequential information [6].

- **Self-attention:** Calculates the relational weights for every word in a sentence, enabling the model to predict words likely to appear in sequence. This mechanism allows each word to attend to every other word in parallel, crucial for learning grammatical rules based on statistical probabilities of how words are typically used in language [6].

2.1.4 Natural Language Processing

Natural Language Processing (NLP) combines computational linguistics - which involves using computational techniques to process and analyze human language - with statistical and machine learning models to enable computers to understand and generate text and speech [10]. In simpler terms, NLP can be described as "teaching computers to understand how humans write and speak" [11]. NLP can further be described as an engineering discipline that seeks to build technology to accomplish useful tasks, rather than just developing theoretical frameworks [12]. The NLP models work by finding connections between the basic elements of language such as letters, words and sentences within a text dataset [12], and, according to IBM [10], enables computers to perform a variety of these tasks in real-time, including:

- Translating text from one language to another
- Responding to typed or spoken commands
- Recognizing or authenticating users based on voice
- Summarizing large texts
- Assessing the intent or sentiment of text or speech
- Generating text, graphics or other content on demand

2.2 Large Language Models (LLMs)

"Large Language Models (LLMs) are a category of foundation models trained on immense amounts of data making them capable of understanding and generating natural language and other types of content to perform a wide range of tasks" [13].

These models operate by using deep learning techniques and can handle various NLP tasks. LLMs are primarily based on the transformer architecture, which uses the self-attention mechanisms to weigh the importance of different words in a sentence [13], see Section 2.1.2. Transformer LLMs are therefore capable of unsupervised learning, although it would be more accurate to say that transformers can perform self-learning. Through this self-learning process, and the enormous and diverse datasets, they acquire an understanding of basic grammar, spelling and knowledge [14]. They subsequently develop the ability to generate text responses that are contextually relevant and coherent, as well as translating languages, summarizing texts, answering questions and generating code [13].

Since LLMs are trained on vast amounts of data to provide foundational capabilities and to be useful in multiple use cases, this goes against the idea of building domain-specific models for each use case [13]. Instead, LLMs offers a more generalized approach that can adapt to various domains through fine-tuning and other techniques. This is a more cost-effective way of using AI for organizations when it comes to infrastructure and resources needed [13].

There are a few different classes of Large Language Models that are suited for different types of use-cases, according to Nvidia [15]:

- **Encoder only:** Typically suited for tasks that can understand language, such as classification and sentiment analysis.
- **Decoder only:** Extremely good at generating language and content. Some use cases include story writing and blog generation.
- **Encoder-decoder:** These models combine the encoder and decoder components of the transformers architecture, illustrated in Figure 2.4, to both understand and generate content. This architecture shines for use-cases including translation and summarizing.

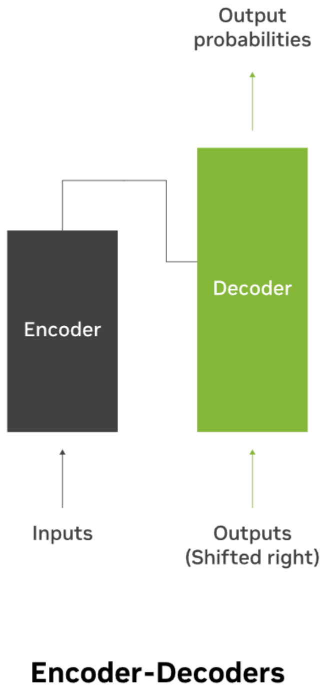


Figure 2.4: The structure of encoder-decoder language models [15]

2.2.1 Conversational AI: QA-models vs Chatbots

When it comes to conversational AI, and employing Large Language Models and advanced language processing techniques like NLP, there are two significant entities that stand out: Question-Answering Models and Chatbots. These serve distinctly different purposes, and are optimized for different functionalities.

A Chatbot is a conversational agent built to simulate human conversation with an end user [16]. They are commonly used for customer service, information or entertainment, where maintaining conversational flow and memory of previous questions is essential. They often include pre-defined scripts and responses, making Chatbots equipped to handle various conversational scenarios where engaging interaction is required.

Question Answering models are, on the other hand, designed primarily to answer specific questions with high accuracy and efficiency. These models can retrieve the answer to a question or user-query from a given text, which would be useful for searching for an answer in a document [17]. This makes QA-models ideal for use-cases where precise information retrieval is critical, and where the answers to user-queries need to be factual and concise. A popular technique involving information retrieval and text generation in a Question-Answering context is known as Retrieval Augmented Generation (RAG).

2.3 Legal Informatics

The integration of AI technologies in legal contexts provides efficiency and automation of routine tasks that would take a lot of time and efforts [18]. For lawyers to integrate AI tools based on machine learning and NLP, it could firstly help them analyze and search through legal documents quickly and efficiently. When working in the legal field, you often have clients to deal with, where you would want to spend as much time with them as possible, to provide assistance and solutions. By the assistance of an AI, lawyers would increase their time available for direct communication and client interaction by reducing the time spent on document analysis[19]. AI tools could extract the relevant information needed from documents based on the personal case of a client, allowing lawyers to focus more on coming up with strategic solutions instead of doing the information retrieval themselves. Furthermore, this integration of AI does not only make lawyers work more efficiently, but ensures that the information they receive is up-to-date and accurate [20], and that their client interactions results in them being able to maintain legal standards and comply to certain regulations. An even more advanced AI integration used in the legal field is customer service QA- or chat-bots, specialized in addressing legal cases directly and handling client issues and questions, only leaving the in-depth legal strategies to the lawyers. These assistants would be specialized and trained on large amounts of legal data and scenarios to be able to comply with certain standards or regulations and give direct legal advice.

2.3.1 Legal Language

When it comes to AI assistants in the legal field, they should also have some understanding of legal language. Legal language can be explained as a sort of a specialized form of communication used within legal documents and in relation to law, and differs from the everyday language. Legal documents and regulations can be challenging to understand sometimes, which is why it is important to use words in a logical and consistent way, when using them in various laws and regulations[21]. This is why the use of technical terms is important; so the information is precise and does not hinder the understanding of the texts[22]. These technical terms in legal language have precise explanations to avoid vague, unclear and ambiguous understandings of laws and regulations. It would therefore be important for an AI GDPR assistant to be aware of the technical terms and legal language used in the GDPR document.

2.4 General Data Protection Regulation (GDPR)

The General Data Protection Regulation (GDPR) is the toughest privacy and security law in the world, which is designed to strengthen and unify data protection for all individuals in the EU [23]. It gives people control over their personal data and imposes strict rules on those hosting and processing this data, anywhere in the world. It is built around seven protection and accountability principles that dictate how personal data should be processed, according to gdpr.eu [23]:

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitation
- Integrity and confidentiality
- Accountability

With this regulation, Europe is making a firm stance on data privacy and security, and will levy harsh fines up tens of millions of euros against those who violate these standards [23]. Therefore, it is important for lawyers to have a good understanding of these regulations, to effectively guide their clients through the complexity of GDPR compliance. By doing so, lawyers can help organizations implement the means necessary to protect personal data, achieve the highest levels of accountability and transparency.

2.4.1 Legal vs Technical Perspective

When working with GDPR, there are two different perspectives one must consider: the legal and the technical perspective. It is logical that some aspects of these perspectives intersect, but it is important for the differences to be clearly stated for the sake of understanding the capabilities of our program. Refer to Chapter 1 for a holistic outline of the program's capabilities.

In short, working with GDPR from a legal perspective means to focus on the importance of compliance. It is about seeking sufficient understanding of the regulations so that the basis of interpretation makes for fair judgement of infringements. For instance, if a judge passes a ruling based on their understanding of the GDPR documentation, it is important that subsequent rulings do not contradict this initial interpretation. A characteristic of the legal perspective is that it is very subjective, and the wording of the regulations for instance can have much to say. For the technical perspective however, the main focus lie on the practical implementations and suggestions based in the legal texts. It is here not necessarily important to understand the articles, but rather that technical systems etc. are in accordance with the regulations. For instance, a system developer may want to find out if encryption of data-in-transit at some point in a architecture is necessary based on paragraphs of the GDPR legal texts.

Given that the target audience for this study is lawyers, the focus of our program lies mostly in the legal aspect. Meaning that the expected interaction between the end-user and the program consists of the user asking research questions, and the program generating answers based on the official GDPR documentation. The program has not been developed specifically to handle requests regarding suggestion on technical implementations. Given that the base model we used (NorskGPT-Llama-2) is pre-trained on 13 billion parameters, we can expect it to be able to handle technical questions to some degree, but for enhanced output, further fine-tuning would be necessary.

Chapter 3

Methodology

This chapter describes the methodologies that has been used to conduct our study. One overarching methodology has been used, namely the Design Science Research methodology (DSR). For the project development model the group has used (Scrum) see also Section 4 in Appendix A. Other methodologies relating to LLM selection, development technologies and development environment are described in the following sections.

3.1 Development Methodology

3.1.1 Design Science Research (DSR)

"Design Science Research aims to generate prescriptive knowledge about the design of artifacts, such as software, methods, models or concepts. This design knowledge helps research and practice to design artifacts systematically and scientifically in future projects" [24]. DSR is a highly practical methodology that is tailored specifically for the development of new artifacts and solutions. For this reason it is often desirable to involve the client/customer in the process as well, as they might have specifications as to how the program should work. For our study, DSR has mainly been used for the benefit of the group, and so the client has not had much involvement in the methodology. An illustration of the full model can be seen in Figure 3.1.

Typically the six steps of DSR are:

1. **Problem identification and Motivation:** This phase focuses on gaining an understanding of and clearly defining the problem at hand, as well as ensuring its relevance in the appropriate field.
2. **Objectives of a Solution:** Here the focus is on setting clear goals and expectations of the artifact/program.
3. **Design and Development**: The actual development of the program happens in this phase. For each cycle you iterate in an attempt to meet the goals in step 2.
4. **Demonstration:** This phase is about testing if the program is able to handle problems and scenarios of which it might encounter.
5. **Evaluation:** This phase involves assessing both qualitative and quantitative performance against preset criteria. How the evaluation of this study in particular was set up can be found in Chapter 5.
6. **Communication:** Finally, the results of the study are presented to the appropriate channels.

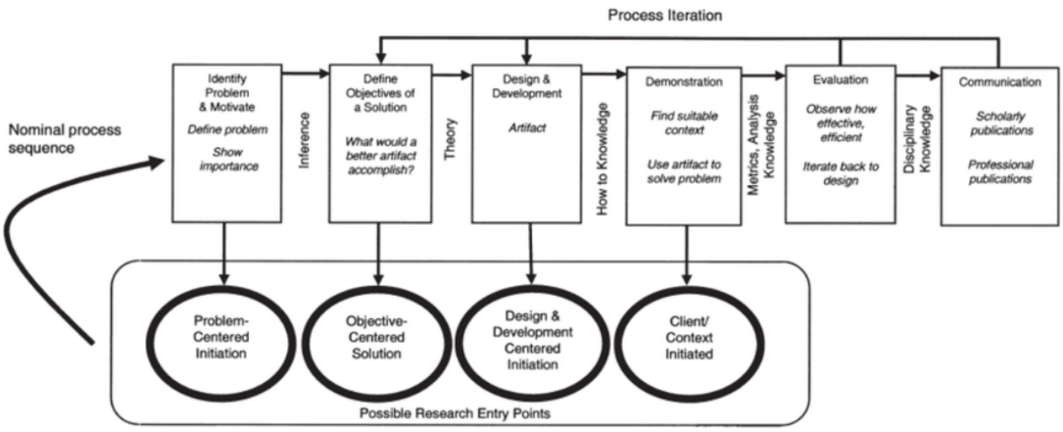

Process Iteration
Nominal process sequence
Identify Problem & Motivate
Define problem
Show importance
Inference
Define Objectives of a Solution
What would a better artifact accomplish?
Theory
Design & Development
Artifact
How to Knowledge
Demonstration
Find suitable context
Use artifact to solve problem
Metrics, Analysis Knowledge
Evaluation
Observe how effective, efficient
Iterate back to design
Disciplinary Knowledge
Communication
Scholarly publications
Professional publications
Problem-Centered Initiation
Objective-Centered Solution
Design & Development Centered Initiation
Client/ Context Initiated
Possible Research Entry Points

Figure 3.1: Model of the DSR process [25]

3.1.2 Agile Software Development

Agile software development is a fairly new methodology that arose as a counterpart to the commonly used, but also rigid, waterfall methodology. The idea of agility in the way one develops software relates to its focus on the work being highly iterative. It facilitates progress through self-organization and accountability, by assigning fixed roles, responsibilities and meetings. There are multiple Agile methodologies such as, Lean Software Development, Kanban, Extreme Programming, and perhaps the most popular, Scrum.

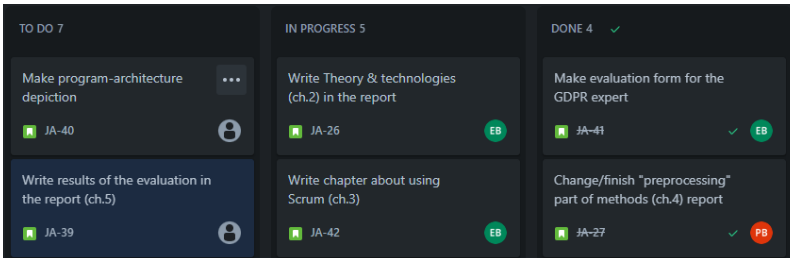
Scrum is most commonly used for software development in particular, although it could also easily be applied to other fields. Perhaps the main benefit of Scrum is when the scope doesn't necessarily remain the same throughout a project. Its focus on continuous collaboration and visibility makes it easier to maneuver often unpredictable and dynamic project conditions. The roles of Scrum, for instance Scrum master and developers, work together on tasks as they are needed, and not necessarily in a chronological manner such as for the waterfall method. The fixed-length iteration periods, with retrospectives and planning meetings before and after, makes for a highly effective development process. The specific Scrum process for our project is described in Section 3.1.4.

3.1.3 Integration of DSR and Scrum

Considering the group was fairly new to the world of AI, it was anticipated that the technologies used and the requirements of the project might change as the project evolved. The client, while not requiring a specific methodology for the project management, also still wanted to be involved and helpful throughout the project. Thus, the dynamic methodologies of both DSR and Scrum, seemed the obvious choice. The somewhat fluid process of DSR makes for a natural way towards progress, although utilizing DSR on its own could mean that the process of reiteration would take longer. Integrating Scrum with the DSR framework in the project would facilitate iterative development and frequent real-time adaptations to our heavily research-driven project, as well as frequent evaluation from program testing and client feedback. Essentially each DSR step would instead be incorporated into each Scrum sprint, thus providing more frequent retrospectives and options for external feedback. Specifically, the way in which we merged the two methodologies was as follows. Step 1 and 2 of DSR would effectively coincide with the sprint planning phase of Scrum. Anytime a new sprint was to begin, the group would come up with goals and tasks, and log them in the backlog in Jira. Within the execution of each sprint the tasks would then be performed, and this was connected with step 3 in DSR. The demonstration and evaluation (step 4 and 5) would typically come naturally at the end of a sprint, although constant evaluations would be performed while in the process of development as well. An evaluation at the end of the sprint would however also facilitate external feedback from the client to a greater degree. Likewise, step 6 of DSR would also become more of an ongoing measure, as feedback from the supervisors would take place on a more frequent basis. The final results however, in other words the report and the findings of the study, would be delivered for appraisal at the conclusion of the project period.

3.1.4 Using Scrum

Given the iterative nature of software development, and the potential requirement changes in the future, the group decided to structure the Scrum workflow into bi-weekly sprints. At the beginning of each sprint, on Mondays, the group would have a sprint planning meeting, client meeting, and a supervisor meeting. The goal for the sprint planning meetings would be to initiate the sprint and lay out the work to be done for the sprint [26]. Sprint goals would be defined to show the client and supervisors why the sprint is valuable, and the different tasks to be done throughout the sprint is put in the product backlog. This backlog was maintained using Jira, allowing the backlog items to be prioritized and managed based on the needs of the constantly evolving project, and potentially the clients feedback. The group also utilized the Scrum board in Jira to easily visualize the progress of each sprint task, as well as effectively distributing them in the group. An example of the structuring and some tasks in the group's Scrum board can be seen in Figure 3.2. The columns on the Scrum board would represent the different stages of each task: "To Do", "In Progress", and "Done".



TO DO 7	IN PROGRESS 5	DONE 4 ✓
Make program-architecture depiction JA-40	Write Theory & technologies (ch.2) in the report JA-26 EB	Make evaluation form for the GDPR expert ~~JA-41~~ ✓ EB
Write results of the evaluation in the report (ch.5) JA-39	Write chapter about using Scrum (ch.3) JA-42 EB	Change/finish "preprocessing" part of methods (ch.4) report ~~JA-27~~ ✓ PB

Figure 3.2: Example of using the Scrum board in Jira

Following the sprint planning meeting, the group would have a client meeting where the plan and goals of the sprint would be presented. The client would then provide their feedback, as well as tips and relevant resources on how to proceed with the tasks, considering the subject and academic material is relatively new for all the group members. The client would also present their requirements and needs for the project along the way, if any. After receiving feedback and insight from the client, the group would have finalized the sprint planning and filled up the backlog in Jira with tasks.

The group would also have a meeting with the supervisors where both what was achieved in the previous week, as well as the plan for the week would be presented. This would include the sprint plan and client feedback every other week, before we officially start the sprint.

In addition, a sprint retrospective meeting would take place after every sprint, right before the sprint planning meeting for the new sprint. The purpose of the sprint retrospective would be for the group to reflect on the sprint process, what was done and potentially not finished, and what could be improved for the next sprint. Tasks that were not finished in the previous sprint would be transferred to the next sprint.

Every other day, which would be every day the group would be together and have a work session, the group held a Daily Scrum meeting. Daily Scrum meetings were held to quickly update the team on what each member is working on, to make sure everyone is synchronized, as well as to kick start the day and the work to be done. In the beginning of the project the group held this every day from Monday to Friday, but eventually found this to not be necessary as the individual tasks the group were working on often lasted two days or longer. Every other Monday, when in the middle of a sprint, the group would also hold a group meeting where we discussed more in detail how we stood with the progress and if we were on track with reaching the sprint tasks and goals. An example of a full sprint schedule is illustrated in Table 3.1 below.

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
Week 1	Sprint retrospective, Sprint planning	Daily Scrum	Self Study	Daily Scrum	Self Study
	Client Meeting				
	Supervisor Meeting				
Week 2	Group Meeting	Daily Scrum	Self Study	Daily Scrum	Self Study
	Supervisor Meeting				

Table 3.1: Two-week Sprint Schedule

Summary of sprints

Sprint 1: 06.02 - 19.02
The first sprint included research around environments for running LLMs, and which LLM models to consider for the AI assistant. We decided to run AI LLMs from Hugging Face, using Hugging Face transformers. We also created a sample dataset.

Sprint 2: 19.02 - 04.03
The second sprint was used to research various fine-tuning strategies and technologies that would be useful for developing the AI program. Mainly, we looked into RAG and LoRA, as well as how to implement these techniques. The group was given access to a VM in SkyHiGh with significant CPU power, to be able to run LLMs. For the LLM, we decided to try the Llama2 model from Meta in Hugging Face, and to use the LangChain framework. We also created some test questions to test the model on some GDPR scenarios as we tried out the models. Eventually we tried out RAG with our AI program and created a dataset including the whole GDPR in PDF format. Additionally, we started the writing of the report, more specifically outlining the structure and starting on the introduction and methodology.

Sprint 3: 04.03 - 18.03
Sprint 3 included running the first version of the AI program with RAG, where we tried out Llama2 and PHI-2. The results proved that Llama2 was the best option, but ran into some issues about providing answers in Norwegian. This lead us to Norwegian versions of Llama2 in Hugging Face, called Ruter-Llama2-13B and Bineric-NorskGPT-Llama2-13B. Furthermore, we implemented Multi-Query-Retrieval along with RAG, for better results.

Sprint 4: 18.03 - 01.04
During the fourth sprint, we moved forward with the Bineric version of Llama2. The focus for this sprint was to create an evaluation method for the program to evaluate the program. In addition, we looked into implementing more advanced RAG methods: HyDE and prompt engineering/optimizing the prompt template.

Sprint 5: 01.04 - 15.04
The goal of sprint 5 was to finish up the AI Assistant and have a finalized product, so we could move on to the evaluation part and finalize the evaluation method and strategy.

Sprint 6: 15.04 - 29.04
In this sprint, the AI program was finished, and a GDPR expert was contacted to evaluate the GDPR knowledge of the program. Along with the evaluation, the focus shifted towards writing the report.

Sprint 7: 29.04 - 20.05
In this long and final sprint, the focus was on finishing the evaluation, and finishing the thesis through continuous iterations with the supervisors. Assistance and feedback was provided by doing this, making sure we included all necessary elements in the thesis.

3.2 Large Language Model Selection

As will be discussed in Section 6.2.4, picking the optimal size model can be challenging. Especially considering external constraints (Section 6.2.1), that may be an impediment. There were many LLMs to choose from, each with its advantages and disadvantages. We tried some different ones early in the development to find the best suited one for our use case. The following were various considerations we had to weigh when choosing the base model of optimal type and parameter size for our project.

- **Llama-2**
  Llama-2 is a very popular LLM developed by Meta. It comes in three different parameter sizes, 7 billion, 13 billion, and 70 billion. Llama-2 7b was the first model we tried. It struggled generating cohesive answers in Norwegian. Likewise did the 13b version not manage to generate Norwegian text consistently, but the overall quality of the answers were higher.
- **Phi-2**
  Phi-2 is a very small LLM developed by Microsoft. It has a total of 2.7 billion parameters. However it struggles with languages other than English, and we found that its ability to follow instructions was limited.
- **Ruter Llama-2**
  To try to solve the issues we had with Llama-2 we tried a version that's fine-tuned to be better at Norwegian. The fine-tuning was performed by a Norwegian public transport company called Ruter. We got the model from Hugging Face, where they have uploaded multiple versions. We used the base 13 billion parameter one, as our previous experimentation clearly pointed towards bigger models giving better answers.
- **NorskGPT-Llama-2**
  As with the Ruter Llama-2 fine-tuned model, the Bineric developed NorskGPT-Llama-2 model is fine-tuned for Norwegian. They have both a fine-tuned version of the 7 billion version and the 13 billion one on Hugging Face. We used the 13 billion parameter one for the same reason as for Ruter.

NorskGPT and Ruter had very similar performance for our tasks, but ultimately we decided to use the NorskGPT-Llama-2-13B version going forward. This was mainly on the basis of an advice from the client. Given that this model was already fine-tuned on some Norwegian data, this would be of benefit for our final program considering the scope mentioned in Section 1.3.

3.3 Development Technologies

This section describes the technologies that was utilized for the development of our program. It includes a general description of the main component we used, namely RAG, as well as a critical problem that the inclusion of RAG seeks to mitigate. Other key technologies for our study are presented in the following subsections, as well as in Table 3.2.

3.3.1 Retrieval-Augmented Generation (RAG)

"RAG's true strength lies in its flexibility. Changing what a pre-trained language model knows entails retraining the entire model with new documents. With RAG, we control what it knows simply by swapping out the documents it uses for knowledge retrieval" [27]. In other words, RAG was introduced to easily "expand and revise their memory", without having to retrain or fine-tune the model with updated data [28]. To grasp the theory behind this expansion of knowledge, one needs to be a little familiar with parametric and non-parametric memory. Neither of these are novel ideas by themselves, but through combining and using them to improve generative AI, there has been many improvements over the last few years [28] [29] [30]. In short, parametric memory in relation to AI means: models learning, storing and employing information directly within their parameters/weights. Nonparametric memory, sometimes called retrieval-based memory [31], relates to models being able to dynamically integrate information from an external dataset/knowledge base as needed.

The continuing interest for RAG in recent years therefore came as a result of the evident performance increase in comparison with only fine-tuning [32]. The ability to swiftly change out the knowledge on which the LLM should base its answers has been shown to be especially beneficial in domains where facts evolve often, and one is dependent on the information generated by LLMs being factual [30]. In the legal domain of GDPR compliance, this would be relevant as one could easily feed new rulings and legal cases to the model's context as needed.

RAG Process

As stated above, RAG works by introducing an information retrieval component with the typical transformer-based generator/reader. Thereby the architecture is

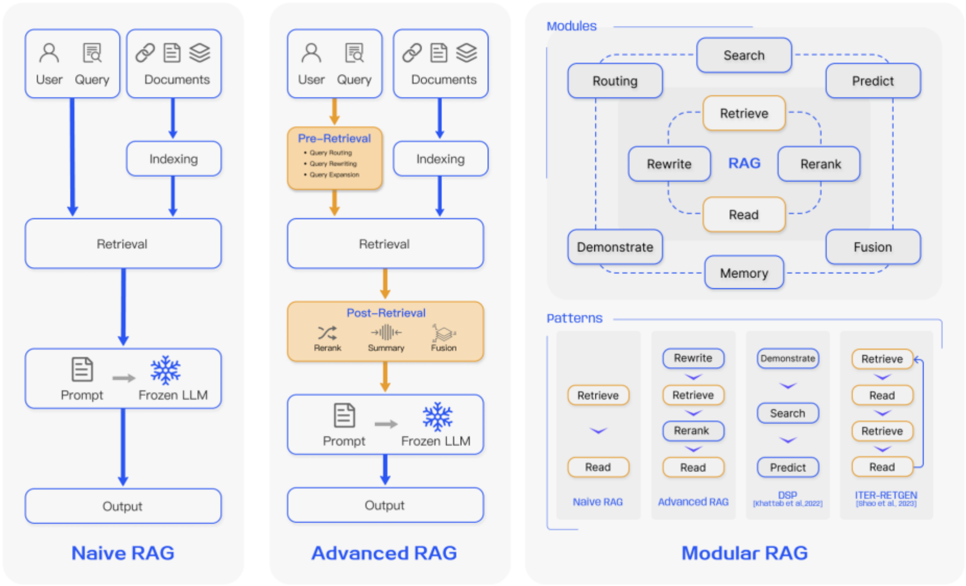


Figure 3.3: The three paradigms of RAG [29]

then consisting of a retriever and a generator. The role of the generator remains the same (as without RAG), but it uses the additional context that has been identified and retrieved by the retriever. In other words, the retriever fetches relevant documents or passages/chunks from an external knowledge source, for instance a large PDF-file. This can be achieved via first embedding of the prompt and then performing a similarity search on vectors in a vector database. When the relevant chunks have been fetched, this context (including the original user prompt), is used by the generator to generate the final output. In a naive RAG implementation this is essentially the entire process, but as can be seen in Figure 3.3 there are other more sophisticated ways that RAG can be implemented. The details of our implementation can be found in Chapter 4.

Hallucinations

"[...] it's possible to build a language model-based system that accesses external knowledge sources to complete tasks. This enables more factual consistency, improves reliability of the generated responses, and helps to mitigate the problem of "hallucination"" [33]. As already described: the RAG framework continue to severely improve the performance of LLM models on specificity, diversity and factuality [28], and this last point especially is of severe importance when accuracy is imperative (e.g. in LLM models that offer regulatory guidance) [30]. When LLMs generate entirely incorrect information, seemingly concocted out of nowhere, this is commonly referred to as "hallucination". Sources of hallucination in LLMs are:

"a lack of relevant knowledge or internalization of false knowledge", that "LLMs overestimate their capacities", a "problematic alignment process" and the LLM's "generation strategy" [34].

To elaborate, the lack of relevant knowledge occurs when LLMs have not been trained on enough correct and relevant data, leading them to fill gaps with fabricated content. This is one of the most significant issues among LLMs. Hallucination can be categorized into three types, according to Zhang et al. [34]: input-conflicting, context-conflicting and fact-conflicting hallucinations. Addressing these types of hallucination is crucial, particularly when dealing with the complexities and precise requirements of GDPR compliance.

3.3.2 Vector Database

A vector database is a way of storing data where the identifier of the data is a vector. This vector is generated by an embedding model. The way of retrieving data from this kind of database is through a similarity search, in which the input is embedded in a similar manner to the data in the database. The resulting vector is therefore compatible to the ones in the database, in the sense that if the input is identical to a piece of data in the database, the input and data will have the same resulting vector after embedding. Likewise, inputs and data with similar semantic meaning will have similar embedding vectors. This fact is used to calculate the similarities between data in the database and the input through various means, such as comparing their magnitudes or measuring their euclidean distance. The advantage of a vector database is that it is specifically made to store and operate on vectors, making it more efficient in comparison to other storage solutions for use with vectors.

Embedding Model

An embedding model is a transformer model that converts human readable information into a list of numbers, or a vector, called embeddings. These embeddings keep the semantic meaning of the original text when interpreted by that same embedding model.

3.3.3 Virtual Environments

A virtual environment is a piece of software that makes an abstract environment on top of the normal machine operating system. The purpose of this is for the development of different pieces of code using different versions of the same dependencies, as the libraries are not shared across environments. It also makes it easy to list out a programs dependencies, as redundant libraries will not be imported into the system.

3.3.4 Other Technologies

Technology	**Description**
Python	A versatile programming language used for software development.
Miniconda	Light version of Anaconda. An environment and package manager used with Python.
Hugging Face Transformers	A library providing a wide range of pre-trained models for Natural Language Processing tasks, facilitating state-of-the-art model implementation.
LangChain	A framework designed to integrate language models into applications, focusing on efficient and scalable language model use. Has many easy integrations for different frameworks and technologies in the AI field that allow them to be used together under one roof.
ChromaDB	A vector database solution for managing large-scale data efficiently, used specifically in data-intensive applications.
Git	Version control system that enables developers to manage and track changes to files, as well as coordinating work in software development.
SkyHiGh	NTNU implementation of Openstack. Platform providing virtualization resources like virtual machines to NTNU users in Gjøvik.

Table 3.2: Technologies Used in Our Program

3.4 Development Environment

The environment we used for development consisted of some previously mentioned libraries and technologies. Firstly we utilized Infrastructure as a Service (IaaS): NTNU's OpenStack installation "SkyHiGh" to run the VM, upon which all testing of LLMs occured. For the process of setting up, connecting and managing the virtual machine, we used the SSH protocol via Powershell. And for the virtual environment itself we applied Miniconda, which is a more lightweight version of Anaconda.

The framework that was used to manage the specific LLM implementations (e.g. RAG) was LangChain. Utilizing this framework allowed for efficient development for instance via the various integrations that are included. One such source of efficiency was the framework's compatibility with Hugging Face, which is where the base LLM was imported from. Once all of this was set up, we needed a couple of other technologies necessary for RAG to function. Specifically an embedding model to translate natural language to machine-readable data, and a database for this data to be stored. The embedding model was imported from Hugging Face as well, while for the vector database we utilized the open-source ChromaDB.

When it comes to the IDE that was used to for the programming itself, Visual Studio Code was chosen based on it being the preferred development environment from previous courses. It is open-source and lightweight as well, while still being highly powerful for instance because of the corpora of available extensions.

Due to the flexible nature of both Miniconda and the utilization of a VM, the group were able to effectively take down, manipulate and re-deploy as needed. As will be discussed in Section 6.4.4 this ultimately limited the need for Git, especially in regards to coordination and teamwork. We did however use Git throughout the project for backup purposes. For more detailed descriptions of the functioning of the mentioned technologies, concerning our use case, please refer to Chapter 4.

Chapter 4

Implementation

This chapter describes in detail how the technologies and methods were used for the final version of the program. For insights into the technologies in particular, please refer to Chapter 3.

4.1 Setting Up the Virtual Infrastructure

The development, testing, and deployment of the program was done on a virtual machine provided by NTNU on their Openstack installation for education and research activities at IIK in Gjøvik, called SkyHiGh [35]. We utilized the web UI to manage and set up the project. For the installation we followed the initial setup guide for command-line clients as provided here by the NTNU wiki [36]. The virtual machine we use within the infrastructure has 32 CPU cores, 128GB of RAM, and runs off a volume with 300GB of storage. The operating system is 'Ubuntu Server 22.04 LTS (Jammy Jellyfish)'. Initially we used Windows Subsystem for Linux (WSL) to connect to our VM, but due to an incompatibility with a newer version of the VPN given to us to connect to the NTNU network remotely, Cisco AnyConnect Secure Mobility Client, we had to switch to the use of PowerShell.

4.2 Virtual Environment for Development

We use Miniconda to keep track of program dependencies. This is a lightweight version of Anaconda with less packages pre-installed, which allows us more control and a better overview of what packages the current program iteration has access to.

4.3 Program Architecture

Figure 4.1 describes the dataflow between the different components of our program.
We start by setting up the database. The data is read from a file and split into smaller chunks. These chunks are then embedded by an embedding model, before they are inserted into the vector database upon initialisation. The user input is then used to perform RAG. This is done in one or more ways between normal RAG, MQR, and HyDE. For regular RAG, the input gets embedded directly, while in MQR and HyDE there is an intermediate step. MQR generates multiple new questions that are similar to the original one from the user. HyDE generates a hypothetical answer to the input question. Both the list of questions from MQR, and the hypothetical answer from HyDE, get embedded by the embedding model. Regardless of the methods used, the embeddings then get used in a similarity search to retrieve relevant documents from the database. All the retrieved documents get put together in a list. Then duplicates are removed, and the list is shortened down to a pre-defined maximum of documents. Finally, the list of documents and the user input query are used to generate the prompt used to query the model for the answer.

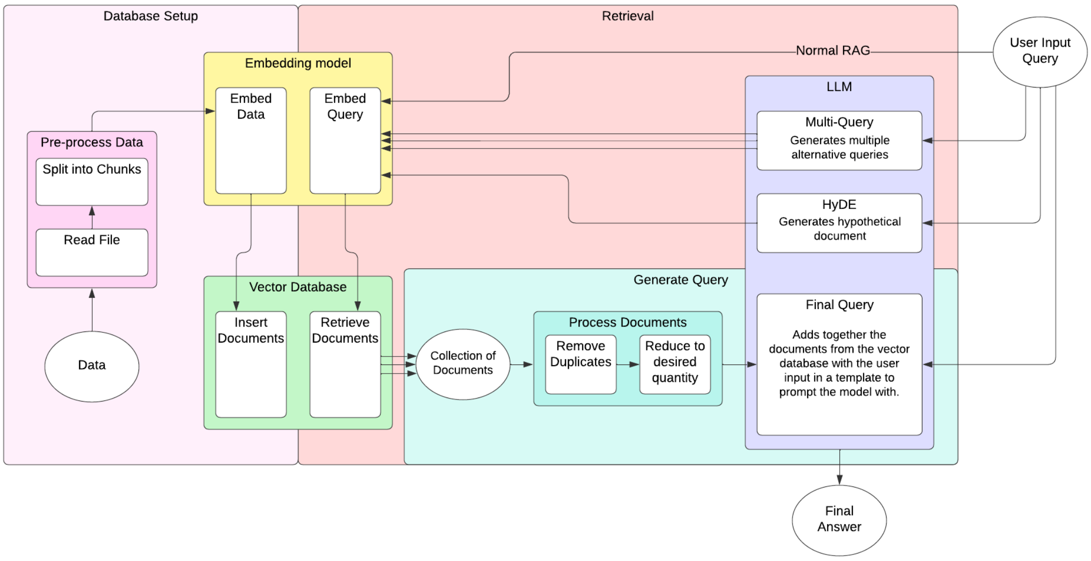


Figure 4.1: Figure of the program architecture

4.4 RAG Configuration

To achieve the desired precision and quality of content for the model answers we decided to use RAG. The method is implemented through the LangChain framework. LangChain has many helpful functions that streamline the process. Firstly we must store the data that we want our model's answers to be based off of in a vector database. After that we can retrieve the data.

4.4.1 Database

For the use of RAG in LangChain we need a vector database. We use ChromaDB as it's free, open source, and runs locally on our VM. Before the data can be stored in the database, it needs to be processed and embedded. The dataset we use is comprised of a pdf of the Norwegian version of the official GDPR document, "The personal data act". This version includes state specific implementation of the regulation (chapter 1-9).

Code listing 4.1: Setting up the vector database

```
 1  # Embedding model to be used
 2  self.embedding_model = HuggingFaceEmbeddings(model_name =
 3  "intfloat/multilingual-e5-large")
 4
 5  # Set up vector database
 6  splits = [] # Contains chunks of the loaded data
 7  text_splitter = RecursiveCharacterTextSplitter(chunk_size=512, chunk_overlap=256)
 8  # Loads data
 9  for source in pdf_data:
10      loaded = PyPDFLoader(source).load()
11      splits.extend(text_splitter.split_documents(loaded))
12  # Reformat texts in splits to work with embedding model requirements
13  for text in splits:
14      text.page_content = "passage: " + text.page_content
15  # Initiates vector database
16  self.vectorstore = Chroma.from_documents(documents=splits, embedding =
17  self.embedding_model)
```

Document Preprocessing

To preprocess our data we first load it through LangChain's PyPDFLoader class (line 10). This gives us a list of documents containing the content of each page alongside pertaining metadata. Then this list is further processed into smaller chunks with overlap using the LangChain RecursiveCharacterTextSplitter class (line 11). This allows for splitting the text in the documents on different characters. We use the default criteria as that suits our purposes.

Embedding Model

After the text has been loaded and formatted it must be embedded. We import a pre-existing embedding model from Hugging Face called multilingual-e5-large (line 2). This model is developed by Liang Wang and uploaded to Hugging Face under the alias 'intfloat'. We used this model because it's multilingual, meaning it should be better at embedding Norwegian text than a purely English model.
The model requires all documents that are to be embedded for insertion into the database to be prefixed with "passage: ", and all query text with "query: ". If this is not included, the model will not perform as expected. We add this part after text loading, but before embedding for passages (line 14), and remove it afterwards before the text is used as context. The queries have the text added right before being embedded. In LangChain one can initiate the vector database and embed the documents to be included directly in one line of code (line 16).

4.4.2 Retrieval

We retrieve documents from the vector database through LangChain's Retriever interface. This is a wrapper that gets put around the vector database that, amongst other things, allows for easy retrieval through similarity search. How this is done is shown in Code listing 4.2. The parameter in line 1 decides how many documents are to be retrieved. In line 3, we added the 'query: ' part in front of the question from the user for it to be in order with the requirements of the embedding model.

Code listing 4.2: Basic RAG implementation in LangChain

```
1 retriever = self.vectorstore.as_retriever(search_kwargs = {'k': 7})
2
3 rag_docs = retriever.get_relevant_documents(f"query: {question}")
```

4.4.3 Overcoming Shortcomings

As the name implies, a similarity search searches a vector database for vectors that are similar to a given one. This then naturally depends on there being similarities between the embedded document in the database and the embedded question. This may not necessarily be the case. To increase the likelihood of this not happening, we implement two methods; Multi-query RAG and Hypothetical Document Embedder.

Multi-Query RAG

Multi-query RAG is simple in concept. We pass the original question to an LLM, and have that LLM re-write the question in a variety of different manners. This then diversifies the word use and slightly alters the semantic meaning, to help broaden the scope of vectors found in the similarity search. The code implementation of this method is described in Code listing 4.3. We use the default multi-query retriever prompt from LangChain, but formatted the prompt for use with the NorskGPT model. We also changed it to generate four different answers, instead of the default five, and used a parameter to pass along the original question alongside the generated ones. We prompt the model as described in Section 4.5, with the exemption of the context section.

Code listing 4.3: Multi-query RAG implementation

```
question = f"query: {question}"

QUERY_PROMPT = PromptTemplate(
    input_variables=["question"],
    template=
"""
### Instruction:
You are an AI language model assistant.
Your task is to generate four different versions of the given user question to retrieve relevant documents from a vector database.
By generating multiple perspectives on the user question, your goal is to help the user overcome some of the limitations of the distance-based similarity search.
Each question must start with the text "query: " to be interpreted correctly by the embedding model, so make sure that it is included at the beginning of each question.
Provide these alternative questions separated by newlines.
Original question: {question}
### Response:
""",
)

multi_retriever = MultiQueryRetriever.from_llm(
    retriever = self.vectorstore.as_retriever(search_kwargs={"k": 2}), llm = model, prompt = QUERY_PROMPT, include_original = True
)

multi_docs = multi_retriever.get_relevant_documents(question)
```

Hypothetical Document Embedder

HyDE is similar to Multi-query in that it also utilizes an LLM to enhance the retrieval of documents. With this method, however, we ask the LLM to answer the question as best it can. This is the ’hypothetical document’. We then embed said answer and utilize that document in a similarity search. The implementation is described in Code listing 4.4.

Code listing 4.4: Hypothetical Document Embedder implementation

```
HYDE_PROMPT = PromptTemplate(
    input_variables=["question"],
    template=
"""
### Instruction:
Du er en ekspert i Norsk jurdistikk og spesialiserer deg innenfor
     personopplysningsloven (GDPR). Besvar så godt du kan følgene spørsmål:
{question}
### Response:
""",
)

hyde_retriever = self.vectorstore.as_retriever(search_kwargs={"k": 4})

hyde_chain = HYDE_PROMPT | model | StrOutputParser()
hyde_ans = hyde_chain.invoke({"question": question})

hyde_docs = hyde_retriever.get_relevant_documents(f"query: {hyde_ans}")
```

Culling the Retrieved Documents

When an LLM generates text it takes into account all previous text. This previous text includes the entire prompt that it was given, as well as what the model itself has generated until then. This is described in Section 2.1.3. Therefore we run the risk of having too much text to where the model can't handle more if there is too much context, or if the answer the model generates is too long. To combat this, we put a hard limit on the amount of documents that can be included. The amount we allowed is 10. This gives the model enough data to go off of, while also allowing it to generate enough text for a reasonable answer. To assure no bias in which documents are kept and which are not, we first randomize the list before cutting it off at the desired length. The list is then sorted again on the number of the page the document was gotten from. A better way of doing this would be by implementing a re-ranker, but due to practical limitations this is not feasible for us.

4.5 Prompting the Model

When prompting the model we use LangChain's PromptTemplate class. It works much like an f-string in python as you can dynamically insert text into a string where you want it by placing a variable there. We structure the prompt to follow NorskGPT's requirements. Code listing 4.5 describes how we query the model the final time to get answer to our question.

Code listing 4.5: Prompting the model

```
1  # Generate prompt from template
2  prompt = PromptTemplate(
3      input_variables = ["question", "context"],
4      template =
5  """
6  ### Instruction:
7  Du er en Norsk juridisk assistent og GDPR ekspert. Du skal hjelpe noen som jobber innenfor lov i å finne raske, presise svar. Sørg for at du ikke repeterer deg selv i svaret ditt. Bruk følgende kontekst til å svare på spørsmålet under. Om du ikke finner relevant informasjon i konteksten, så bare si at du ikke har nok informasjon til å svare på spørsmålet. Du skal svare på Norsk.
8  Spørsmål: {question}
9  ### Input:
10 {context}
11 ### Response:
12 """,
13 )
14 chain = prompt | model.model
15 result = chain.invoke({"question": question, "context": context})
```

The question provided by the user is put in line 8. There is a designated area for context (line 10). This is where the documents gathered through RAG are implemented. We tell the model to base its answers off of this context, and refrain from answering if it does not deem the information gathered to be relevant. The model will generate all new text after the ":" on line 11.

Chapter 5

Results and Evaluation

This chapter combines the results and evaluation of the AI assistant and its functionality, including both a group-internal evaluation and a GDPR expert evaluation.

5.1 Evaluation Strategy

There were various ways in which testing of program performance was possible, as will be discussed in Section 6.4.4. The ways one can evaluate, range from highly technical and automated, to more subjective and manual. The task of evaluating LLM performance does however prove to be tough [37]. Thus it is very common to involve humans in the evaluation process [30], often because of the presence of open-ended and subjective queries. For the evaluation of our program, we have utilized this concept of human evaluation, both in a preliminary testing phase, as well as an evaluation that was performed by a GDPR expert: Dr. Mohamed Abomhara. Including an external expert assessment was highly beneficial for the integrity of our findings, and provided perspective on our AI assistants capabilities, considering the group's lack of GDPR knowledge and experience.

The overall strategy leading up to the preliminary evaluation was first to conduct general testings, based on the group's impression of the overall promise of base model performance. At the beginning this testing was limited to differences solely based on model size and parameters. This expanded into testing of the increase or decrease in performance based on different technologies that we sought to implement. When we had decided on the "NorskGPT" version of Llama-2-13B, we then experimented with various settings such as chunk overlap and the content of the prompt instruction. These evaluations were conducted solely by the group members, and the specific factors that influenced our decisions were: the program's general language apprehension, relevancy of the output, inference time, and the completeness of the output. To test and do the preliminary evaluation of the AI assistants final version we will use the dataset in Appendix C, including ChatGPT created GDPR queries.

The strategy for the evaluation performed by Dr. Abomhara was for him to come up with some GDPR scenarios which would simulate real world scenarios. We would then run them through four variations of our program, presented in Section 5.3.3, and then he would assess the responses based on multiple evaluation criteria and a scoring system. We also introduced an extra feature by having Dr. Abomhara assess the performance of ChatGPT-4, fed with the same scenarios as for the the other four. This was unbeknownst to him, and the idea behind this was to provide some more context to the four original evaluations. Thus it is easier to relate the scoring of our program to the performance of an already established and robust application.

5.2 Preliminary Evaluation

Before conducting the expert evaluation, the group conducted internal testing and evaluation of the AI assistant's functionality. For the testing of functionality, the focus lied around what was expected for the program to do based on both the goals and some of the requirements from the client. Mainly, the tests and results described in this section will focus on the following overarching functionalities and expectations:

1. The program uses an LLM, and is able to use the GDPR document as context as well as provide answers using this context.

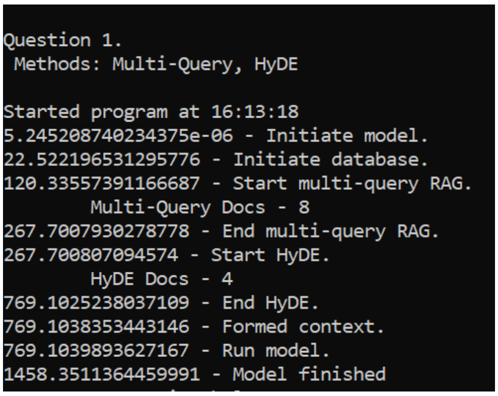
2. The program is able to provide legal advice for lawyers regarding GDPR and work as an "assistant". It should make lawyers work more efficiently for their clients, by providing simple access to GDPR related information.

3. For the program to provide this advice and information in sound Norwegian language.

5.2.1 Test Results and Evaluation

Functionality 1

Firstly, the program's ability to utilize the GDPR documents as context was achieved through RAG, and further improved with the advanced RAG techniques: MQR and HyDE. The fact that the AI assistant utilizes the context to generate a response is implemented in the code, by displaying how many relevant "documents" has been retrieved from the context based on the input query, displayed under "Multi-Query Docs" and "HyDE Docs" in Figure 5.1.



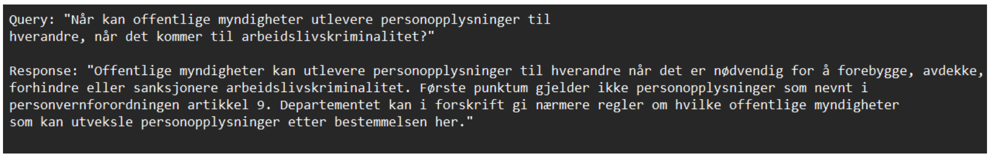
```
Question 1.
 Methods: Multi-Query, HyDE

Started program at 16:13:18
5.245208740234375e-06 - Initiate model.
22.522196531295776 - Initiate database.
120.33557391166687 - Start multi-query RAG.
        Multi-Query Docs - 8
267.7007930278778 - End multi-query RAG.
267.700807094574 - Start HyDE.
        HyDE Docs - 4
769.1025238037109 - End HyDE.
769.1038353443146 - Formed context.
769.1039893627167 - Run model.
1458.3511364459991 - Model finished
```

Figure 5.1: Example of documents retrieved from the context

To further make sure that the program utilizes the context, we tell the AI in the prompt to not provide an answer if it does not find relevant information in the context, based on the query. This is described in Section 4.5.

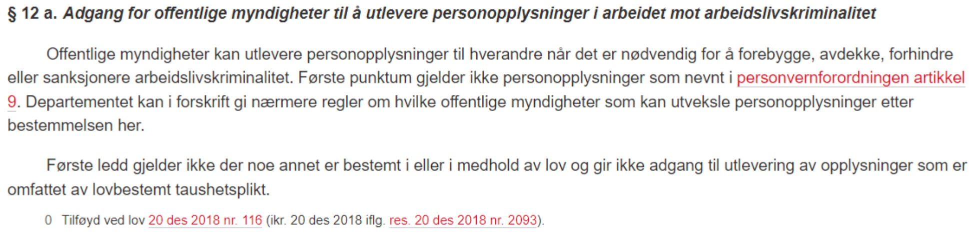
In addition, we can compare the GDPR related answer from the AI assistant with the actual GDPR document to prove the answer is related to the context. In this case, from Figure 5.2 the response to this query from the AI assistant is very similar and almost word for word, compared to the GDPR document, as shown in Figure 5.3.



```
Query: "Når kan offentlige myndigheter utlevere personopplysninger til
hverandre, når det kommer til arbeidslivskriminalitet?"

Response: "Offentlige myndigheter kan utlevere personopplysninger til hverandre når det er nødvendig for å forebygge, avdekke,
forhindre eller sanksjonere arbeidslivskriminalitet. Første punktum gjelder ikke personopplysninger som nevnt i
personvernforordningen artikkel 9. Departementet kan i forskrift gi nærmere regler om hvilke offentlige myndigheter
som kan utveksle personopplysninger etter bestemmelsen her."
```

Figure 5.2: Query and response from the AI assistant



§ 12 a. ***Adgang for offentlige myndigheter til å utlevere personopplysninger i arbeidet mot arbeidslivskriminalitet***

Offentlige myndigheter kan utlevere personopplysninger til hverandre når det er nødvendig for å forebygge, avdekke, forhindre eller sanksjonere arbeidslivskriminalitet. Første punktum gjelder ikke personopplysninger som nevnt i personvernforordningen artikkel 9. Departementet kan i forskrift gi nærmere regler om hvilke offentlige myndigheter som kan utveksle personopplysninger etter bestemmelsen her.

Første ledd gjelder ikke der noe annet er bestemt i eller i medhold av lov og gir ikke adgang til utlevering av opplysninger som er omfattet av lovbestemt taushetsplikt.

0 Tilføyd ved lov 20 des 2018 nr. 116 (ikr. 20 des 2018 iflg. res. 20 des 2018 nr. 2093).

Figure 5.3: Where the AI assistant used the context

Functionality 2

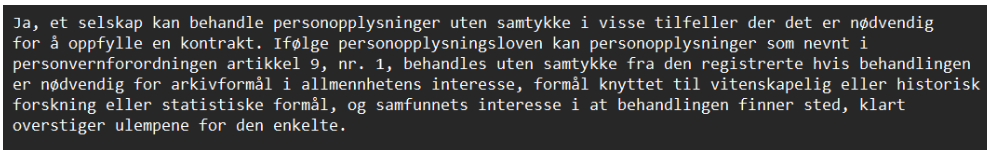
To evaluate the program's ability to work as an AI assistant and provide simple access to GDPR related information, we used some of the GPT generated queries and responses from Appendix C. These queries could represent GDPR questions or cases with which a lawyer would need assistance when dealing with potential clients. We will evaluate the response of these queries based on the correct answer:

- Q1: "Kan et selskap behandle personopplysninger uten samtykke i tilfeller der det er nødvendig for å oppfylle en kontrakt?"
- Q2: "Er det lovlig for et markedsføringsfirma å dele kundedata med en tredjepart uten kundens samtykke?"

From the first query, the response based on GPT and the GDPR should be:

"**Ja, hvis det er nødvendig for å oppfylle en kontrakt.**".

The AI assistant provided the response shown in Figure 5.4. It provides the correct answer, and also explains in which article this is mentioned, and more in detail in which cases a company can do this.



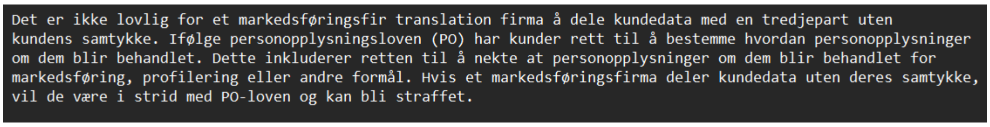
```
Ja, et selskap kan behandle personopplysninger uten samtykke i visse tilfeller der det er nødvendig
for å oppfylle en kontrakt. Ifølge personopplysningsloven kan personopplysninger som nevnt i
personvernforordningen artikkel 9, nr. 1, behandles uten samtykke fra den registrerte hvis behandlingen
er nødvendig for arkivformål i allmennhetens interesse, formål knyttet til vitenskapelig eller historisk
forskning eller statistiske formål, og samfunnets interesse i at behandlingen finner sted, klart
overstiger ulempene for den enkelte.
```

Figure 5.4: AI assistants response to Q1

From the second query, the response based on GPT and the GDPR should be:

"**Nei, med mindre det er uttrykkelig samtykke fra personen**".

The AI assistant then provided the response shown in Figure 5.5. In this case the response is also correct, where it mentions that sharing customer data with third-party companies is not allowed without the customers consent. It then goes on with explaining the rights of the customer more in detail about how their data should be treated.



```
Det er ikke lovlig for et markedsføringsfir translation firma å dele kundedata med en tredjepart uten
kundens samtykke. Ifølge personopplysningsloven (PO) har kunder rett til å bestemme hvordan personopplysninger
om dem blir behandlet. Dette inkluderer retten til å nekte at personopplysninger om dem blir behandlet for
markedsføring, profilering eller andre formål. Hvis et markedsføringsfirma deler kundedata uten deres samtykke,
vil de være i strid med PO-loven og kan bli straffet.
```

Figure 5.5: AI assistants response to Q2

This evaluation would prove the program's capabilities to serve as an assistant by providing advice and information to lawyers regarding GDPR. It is able to provide correct answers, to guide them to the correct articles, and to even provide more details and explanations when answering the question at hand. However it is not designed- or able to replace lawyers, and independently guarantee 100% accurate and detailed answers or strategies, regarding GDPR compliance on its own. And when it comes to providing 'simple' access, in our development environment the AI assistant uses 1458 seconds, approximately 24 minutes, to provide a response seen in Figure 5.1, which is not very efficient.

Functionality 3

The program being able to provide answers in, and understanding Norwegian language, was solved by utilizing a LLM already fine-tuned in Norwegian. Addressing the response in Figure 5.5, considering all the group members speak and understand Norwegian fluently, the language provided by the program is perfectly understandable for everyone. Therefore, the AI assistant does indeed provide information in sound Norwegian language.

When it comes to the AI assistant writing and understanding Norwegian "legal language", it includes specific references to articles within the GDPR, which is a legal framework, and uses terminology typical for legislation and regulations related to data privacy and security such as "data deletion" and "data correction". It may not specifically use more traditional legal terms typically intended for general judicial use, but is clearly within the scope of the legal language specific to the GDPR sector.

5.3 Expert Evaluation

To evaluate the AI assistant's abilities to handle GDPR related issues, a local GDPR expert from NTNU, Mohamed Abomhara, was contacted to provide scenarios and questions for the program to handle, as well as evaluating these responses. Dr. Abomhara is an experienced cyber security researcher with a strong background in data protection, privacy, and data ethics. He is also a certified EU GDPR Foundation and Practitioner, as well as a Certified Data Ethics Professional and Data Ethics Facilitator.

5.3.1 Expert Evaluation Process

During the evaluation process with Dr. Abomhara, the group initially had an introductory meeting with him, where we presented our project, what our AI GDPR assistant was intended for, and what we needed from him. We informed him that we needed GDPR related scenarios, including questions, for our AI assistant to handle. We then told him that after our AI assistant responded to the scenarios we would like him to evaluate these responses based on some criteria in an evaluation form we would provide. Dr. Abomhara then provided us four complex GDPR scenarios for our program. When the program's responses and the evaluation form was ready, we had an additional meeting where we explained how he would conduct the evaluation and how to fill out the evaluation form, as well as discussed some of the responses.

5.3.2 Scoring System

The following scoring system was used for the expert evaluation, and was inspired and partially generated by ChatGPT [38] (see Figure C.2), as well as by feedback and input from the supervisors.

Explanation of the evaluation criteria:

- **Accuracy**
  Measures how correct and legally accurate the answers provided by the AI are. This ensures that the information given is precise and can be reliably used for GDPR compliance without leading to potential legal risks. Scoring metrics are illustrated in Table 5.1.

- **Relevance**
  Assesses whether the answers directly address the questions asked, focusing

on the core issues without digressing. This ensures that the AI's responses are pertinent to the specific GDPR queries posed, avoiding unnecessary or unrelated information that could confuse users. Scoring metrics are illustrated in Table 5.2.

- **Completeness**
  Evaluates whether all parts of the question have been answered and whether the response includes all necessary details. This is important for providing thorough advice or information that covers all aspects of the query. Scoring metrics are illustrated in Table 5.3.

- **Clarity**
  Looks at how easily understandable the answers are, considering structure, language, and presentation. This is useful for ensuring that users can easily comprehend and apply the AI's advice, which is particularly important in complex regulatory environments like GDPR. Scoring metrics are illustrated in Table 5.4.

Score	Definition
1-3	The answers are mostly incorrect or legally inaccurate, failing to align with current GDPR regulations.
4-6	The answers are partially correct but include significant legal inaccuracies or outdated interpretations.
7-8	The answers are mostly correct, with minor legal inaccuracies or slight misalignment with current regulations.
9-10	The answers are fully correct, demonstrating precise legal accuracy and full alignment with current GDPR interpretations.

Table 5.1: Scoring for Accuracy

Score	Definition
1-3	The answers are largely irrelevant, failing to address the core questions or include excessive unrelated information.
4-6	The answers address the core question but include unnecessary or somewhat irrelevant information.
7-8	The answers are relevant with slight digressions but remain focused on the question.
9-10	The answers are highly relevant, directly addressing the question comprehensively without any irrelevant content.

Table 5.2: Scoring for Relevance

Score	Definition
1-3	The answers cover less than one-third of the necessary content, missing key points and details required for full understanding.
4-6	The answers cover more than one-third but misses important details or aspects needed to comprehensively answer the query.
7-8	The answers are largely complete, lacking some finer details or examples that could provide full perspective.
9-10	The answers are comprehensive and detailed, covering all aspects and necessary details of the query.

Table 5.3: Scoring for Completeness

Score	Definition
1-3	The explanations are confusing, poorly structured, and use inappropriate or incorrect terminology, making understanding difficult.
4-6	The explanations are understandable but could be clearer or better structured; uses mostly appropriate terminology but occasionally falters.
7-8	The explanations are clear and well-structured, with minor issues in terminology or presentation.
9-10	The explanations are exceptionally clear, well-structured, using appropriate legal and technical terminology accurately and efficiently.

Table 5.4: Scoring for Clarity

5.3.3 Evaluation Method

The GDPR expert from NTNU provided four different GDPR scenarios (see Appendix I), that included both a description of the scenario, as well as GDPR related issues and questions. Each scenario was fed into four versions of our AI assistant, which included the different technologies and combinations. The different versions used the Bineric/NorskGPT-Llama-2-13B LLM, with the following technologies:

- Only Retrieval-Augmentetd Generation (RAG)
- Only Hypothetical Document Embeddings (HyDE)
- RAG with Multi-Query Retriever (MQR)
- RAG with MQR + HyDE

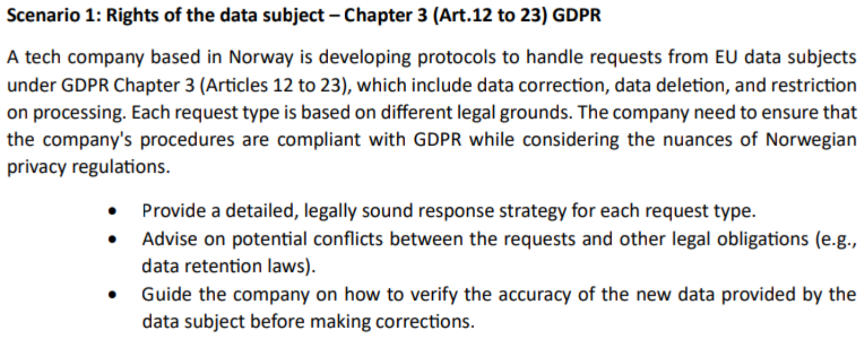
The scenarios were also fed into the mighty GPT-4 by OpenAI, for comparison.

The GDPR expert then evaluated the results of the five different AI programs, where all of them provided their answers to the four different scenarios using the scoring system. It is worth mentioning that the evaluator did not know which of the versions included which technologies, or that GPT-4 was one of the five programs.

5.3.4 Evaluation Results

This section presents the results of our expert evaluation of four different versions of our program, as well as GPT-4. The expert provided both overall comments about the AI assistants different responses, as well as scores. The specific evaluation form used by the GDPR expert can be found in Appendix D.

Note that this evaluation is performed only by Dr. Abomhara and although he is considered a local "GDPR expert", the assessment should be interpreted with caution. It represents the views and understandings of a single individual, who, despite his expertise and certifications, offers perspectives that may not encompass all possible interpretations or applications of the GDPR. Therefore, his evaluation should not be seen as a legal validation of the program, but rather as the perspective of a single expert's opinion. It does however provide valuable insights and guidance about the program's capabilities when it comes to responding to GDPR related scenarios, which in theory should give us a good evaluation of whether our AI assistant works well or not. The evaluator gave overall comments and thoughts about the answers from the different versions of the AI assistant, to further describe the reasoning behind the evaluation scores. We will compare Dr. Abomharas thoughts and comments about the responses from Scenario 1, shown in Figure 5.6.



Scenario 1: Rights of the data subject – Chapter 3 (Art.12 to 23) GDPR

A tech company based in Norway is developing protocols to handle requests from EU data subjects under GDPR Chapter 3 (Articles 12 to 23), which include data correction, data deletion, and restriction on processing. Each request type is based on different legal grounds. The company need to ensure that the company's procedures are compliant with GDPR while considering the nuances of Norwegian privacy regulations.

- Provide a detailed, legally sound response strategy for each request type.
- Advise on potential conflicts between the requests and other legal obligations (e.g., data retention laws).
- Guide the company on how to verify the accuracy of the new data provided by the data subject before making corrections.

Figure 5.6: Scenario 1 from the expert evaluation

Only RAG

When only using RAG, the evaluator thought the results were not entirely optimal, as they were lacking details and completeness. Using scenario 1 about "Rights of the data subject" as the example (Figure 5.6), the AI assistants response, shown under Group 1 in Appendix E, "*did not include any detailed protocols for each type of request and also did not include any strategy for handling requests. Moreover, no discussion or consideration about how to handle Conflicts with Norwegian Law or how to how to verify the accuracy of the new data provided by the data subject before making corrections.*", according to Dr. Abomhara.

Only HyDE

When using only HyDE, the response from Group 4 in Appendix E also lacks details when advising on handling GDPR data subject requests, but addresses some aspects of it. More specifically Dr. Abomhara thought "*the response addresses aspects of handling GDPR data subject requests. However, not in detailed steps. Specifically, the advice on handling conflict between GDPR and other laws, the response did not provide detailed strategies for resolving such conflicts*". This would lower this program versions abilities to provide very accurate and complete responses.

RAG with MQR

Using RAG with MQR, the response proves to be more accurate and detailed then the previous versions. From the answer from Group 3 in Appendix E, Dr. Abomharas overall comment was that "*the response accurately references GDPR rights related to data subject requests, including the right to correction, deletion, restriction etc... It outlines steps to ensure compliance with GDPR for various data subject requests which is exactly what is needed in the scenario. However, it lacks specific implementation details such as how to verify the identity of the data subject or how to handle cases where the data cannot be deleted due to other legal obligations.*". Using RAG with MQR would therefore give higher accuracy in the response, as well as clarity, but still lacks details for it to be an optimal response.

RAG with MQR + HyDE

The complete version of our AI assistant proved to provide the best answers, and received the best overall score in Table 5.5. The response from Group 2 in Appendix E, proved that the relevance of the response has clearly improved. Specifically, Dr. Abomhara thought that "*the response is relevant as it addresses the key issues outlined in the scenario. It correctly identifies the relevant GDPR articles related to each type of request and discusses potential conflicts. However, lacks specific details (e.g., reference to Norwegian laws) that would enhance its completeness. Also, there are a few areas where clarity could be improved such as explanation of legal terms*". The response is still not optimal, but not terrible considering the accuracy and relevance.

GPT-4

Comparing the responses with GPT-4 (Group 5 in Appendix E), Dr. Abomhara provided this comment: "*The response outlines steps and considerations necessary for complying with GDPR Arts. 16, 17, 18. It also considers the interplay between GDPR compliance and other legal obligations, such as data retention laws, which is critical for practical implementation. However, the response could be enhanced by including more examples or guidelines on handling specific scenarios, such as dealing with particularly sensitive data or more complex third-party data sharing agreements.*".

Evaluation Scores

The final scores of the evaluation are displayed in Table 5.5, where each of the final scores have set by taking the average of the four similar criteria scores (e.g. the four Accuracy scores), for each version of the program.

Model Version	Accuracy	Relevance	Completeness	Clarity
Only RAG	3.8	4.5	3.0	4.0
Only HyDE	3.0	4.5	3.0	3.0
RAG with MQR	4.0	4.0	3.3	4.0
RAG with MQR + HyDE	4.5	5.7	3.8	3.8
GPT-4	6.5	6.5	4.5	5.2

Table 5.5: Expert evaluation of our program's different versions, compared to GPT-4.

From the results of the expert evaluation, both considering the comments and scores from Dr. Abomhara, we see that the program version using RAG with both HyDE and MQR, which is our final product, performed the best out of the four which was also expected. But even though it performed the best, the resulting scores are not optimal considering the maximum score is 10. Comparing our programs to GPT-4, it scores a little bit better especially for its accuracy and relevance, but proves that the answers provided by the GPT-4 model are not optimal either. One might therefor argue that the scenarios provided by the expert were very complicated. He also mentioned that two of the scenarios would be particularly tricky, combined with the fact that the expert does not have a legal background. This means that he could evaluate the answers more based on if they are "technically sound", rather than if they are able to give "legally sound" advice.

Chapter 6

Discussion

This thesis has been conducted for a total of 18 weeks, or the full duration of a semester. During such a large amount of time, and such a large project, there are bound to be challenges along the way. Some of these naturally result in changes of perspective and focus, while others have simpler workarounds. This chapter contains discussions of just such challenges, and how they have been overcome.

The chapter is divided into four distinct sections. The first contains a discussion of the results from both the group's preliminary evaluation and the expert evaluation. The second has discussions of theory and possible implementations. The third section discusses the process of working on the thesis, while the fourth is a reflective review of factors influencing AI practice.

6.1 Results

6.1.1 Preliminary Evaluation Discussion

The results of the preliminary evaluation were presented based on three expected functionalities. These broadly encompassed: our program's ability to use a GDPR document as context, to work as an assistant by providing legal advice, and for it to do so in Norwegian. The basic ideas behind these are relatively straightforward, but the results can be interpreted variously based on several factors. Please refer to Chapter 5 for the exact results.

We used the idea of RAG and implementation of similar information-retrieval based technology to achieve the results of functionality 1. Our program is able to utilize the external GDPR document and it does so accurately, as can be seen in Figure 5.2 and Figure 5.3. We could however achieve these results through other means, as will be discussed later. Although the group focused on RAG technologies in particular for this study, we found that there were more substantial previous research on fine-tuning. Moreover, the fact that the field of AI technologies is in constant development, further complicates the process of choosing the best approach. For our use-case we based our approach specifically on research regarding the introduction of LLMs for regulatory guidance. Kim and Min [30] for instance states the importance of accuracy in LLMs for compliance, despite the highly detailed data that this would require. Incorporating fine-tuning techniques, for instance those mentioned in Section 6.2.3, could achieve similar or perhaps even better results than those we have found. The topic of the benefits of RAG vs. fine-tuning will be discussed further in Section 6.2.2.

Functionality 2 was based on the program being able to work as an assistant and thus streamline the access to GDPR related information. This was also achieved, although the robustness of the program could be improved. For instance there is a limit to the internal fact checking of multi-query RAG, and it could be of benefit to introduce another component more tailored for this purpose. One option could be to introduce another LLM in the architecture which has the specific purpose of making sure that the generated answer is sensible and complete. This solution is also mentioned in Section 7.3. It is however important to reiterate, that having our program function as an *assistant* means that the overall process should still require human oversight. While human oversight is needed, there will always be an element of bias introduced.

Similarly to the results of functionality 2, the result of functionality 3 could be upgraded through fine-tuning. It is able to answer in comprehensible Norwegian while also utilizing the legal language of the provided GDPR document, but fine-tuning would probably be necessary for further significant improvement. By the legal language that one can find in detailed regulations such as GDPR being highly specific, it is crucial for an AI assistant to be precise in its choice of words.

6.1.2 Expert Evaluation Discussion

The results of the expert evaluation provided us with valuable insight and a deeper understanding of the capabilities of our program. This subsection will mainly discuss the scores of the full version of our program (RAG with MQR + HyDE).

Although Dr. Abomhara is considered a local GDPR expert, we must restate that he might not have the necessary prior knowledge in regards to the purely legal aspect to GDPR, as he does not come from a legal background. Rather he generally focuses more on the technical aspects. Thus we found that the scenarios we were handed were somewhat tough for our program to handle. The overall comments he provided on the performance also suggests that he may have favored responses aligning more closely with technical accuracy rather than legal precision. But, rather than simplifying the scenarios we kept with them, as we nevertheless deemed them suitable to find the upper limit to the program's capabilities. These factors might together be some explanation to the somewhat low scores across the board, including for GPT-4. Moreover, the consequence of using human evaluation, which we did for both the preliminary testing and the expert evaluation, brings with it the factor of bias and subjectivity. To deal with this, we did as previously stated not disclose the presence of GPT-4 in the program answers before evaluation. If there were some bias present on the context of the group's inexperience in the field, the expert evaluating an established chatbot alongside our program could help reduce this. The lowest score that was given to the full program was 3.8 for clarity. This suggests that the explanations and use of correct terminology in the output were not consistently sufficient. Similarly the relatively low score for completeness suggest that the program is not always able to hit all the key points and details that Dr. Abomhara expected. Some of these faults might be attributed to him expecting more technicality in the output, but there could also be other reasons.

Besides this, based on the expert review we found some interesting results. Firstly, our prediction as to what RAG technologies would perform the best was pretty much realized, as the full program outperformed every other version, on every criterion except clarity. The score for relevance was the highest and this might suggest that RAG first and foremost affects the program's ability to find up-to-date, and correct data. Relevance was early on in the project deemed to be one of the two most important criteria for our use-case, with the other being accuracy. The reason for this was the aspect of our program being able to enhance GDPR compliance. Research suggest that utilizing the most precise data is crucial in domains that depend on reliable information[30], and we believed that focusing on RAG rather than fine-tuning would serve this purpose. As to why the score for clarity was lower, which was mentioned earlier, there could be several explanations. Apart from the potential bias, there is an argument as to the role that the size of the model has on improving LLM functionality. Likewise, the optimal amount of data that one includes in the training dataset and/or in an external database for RAG, is debated. These are essential perspectives one must consider when developing a QA-model, and both will be discussed further in Section 6.2.4.

All in all our program does function proportionally to what was expected, especially considering the limitations to our program, as well as the external constraints. We do however recognize the need for more objectivity in both evaluations, as this would put our current scores in greater perspective. There were several ways in which more objectivity could be achieved, and this will be discussed further in Section 6.4.4.

6.2 Theory and Implementation

6.2.1 Constraints

In the process of developing the AI legal assistant there were various external constraints that directly impacted the degree of which the program could be optimized. This was of course to be expected based on the already short allotted time frame of just one semester. Nevertheless, the group also found that the constant scaling of LLM technologies, due to the field still being relatively new, further narrowed the realistic time we could spend experimenting. Although frameworks and guidelines are available, there is not really any common understanding of an optimal baseline for QA-models. Just picking the optimal size model, or figuring out how to modify datasets to your program's benefit, can potentially be very time-consuming.

Besides the time limits, perhaps the greatest constraint to the thesis was the lack of available computational power. There were no on-campus physical PC's available, and we found that the thesis' contributors' personal PC's were sub-optimal due to them all having Windows operating systems. While the operating system would not necessarily impact the performance, many AI-tools and techniques are tailored for Linux. Moreover, the PC we had hoped to use had a AMD GPU, and we found that AI frameworks that supported AMD often did not support Windows. For these reasons we found it most suitable to get access to a VM through NTNU's OpenStack installation 'SkyHiGh'. In the end we ran the models on CPU instead of GPU, on the suggestion of a recommendation of a supervisor and because we did not utilize fine-tuning. The VM also ran Linux to ensure compatibility with possible frameworks. Refer to Section 4.1 for more details surrounding the VM's specifications.

Furthermore, it has been made clear that further drastic improvements of the program would most certainly require fine-tuning of some sort. As is discussed in Section 6.2.2, fine-tuning was the initial idea for the scope of the thesis, but was ultimately replaced with the RAG framework. Although RAG typically require less time spent on manipulating and modifying datasets compared to fine-tuning, to fully tailor datasets from the bottom-up for RAG would compare in complexity. At this time our program gets all its external context from a PDF-file, but there is an argument that collecting and labeling unstructured data manually would prove beneficial to the ability to retrieve as relevant data as possible [39]. This would however require high precision. By building your own dataset you would also naturally have more control over the privacy and security of the information. The ethical aspects of working with AI is discussed further in Section 6.4.2. Additionally, if we were to build our own datasets we would also have to consider preprocessing steps such as stemming and lemmatization more closely. These are techniques for structuring unstructured data, and are as of now handled automatically in the pipeline via Hugging Face's tokenizer "AutoTokenizer" [40]. Tailoring exactly how data is structured and ultimately embedded into computer readable information could enhance the matching of user queries and optimize the retrieval of external context.

6.2.2 Why RAG and not Fine-tune?

Lewis et al. states: "Large pre-trained language models have been shown to store factual knowledge in their parameters, and achieve state-of-the-art results when fine-tuned on downstream NLP tasks. However, their ability to access and precisely manipulate knowledge is still limited [...]" [28]. For this study, the group has only utilized technology related to RAG, even though a significant amount of hours was spent researching fine-tuning. The initial idea for the thesis was to apply technology and findings from the fine-tuning research, rather than focusing on the information retrieval aspect. Generally those two are the broad methods in which you can shape a LLM to your needs, but given that the term "fine-tuning" contains lots of more specific technologies [41][42], this was considered to have the most promise for our initial scope. There are also more ways that fine-tuning showed promise for our project. This is further explored in Section 6.2.3.

Perhaps the greatest benefit of fine tuning for our project in particular would be "deep integration of knowledge". Specifically for it to gain a deeper knowledge of the Norwegian language. This relates to the theory behind "deep learning" which is presented in Section 2.1.1. While the RAG framework helps our program with manipulating up-to-date knowledge, fine-tuning would certainly help our program with being able to maneuver the deep linguistic nuances of the Norwegian written languages in general, but also the highly specific legal language that is often found in legal texts. So utilizing some fine-tuning technique, for instance LoRA which is mentioned later, would probably be the realistic way to enhance the program's language capabilities. Furthermore, through enhanced knowledge integration during a fine-tuning training process, the program could also enhance its ability to "see" more nuanced connections in the GDPR document, perhaps through incorporating some data pertaining to related legal domains. For our program's use-case this would not have mattered much, but it could potentially be of more use if the program was meant to generate more practical implementation suggestions based on the user input. Considering that in the end no fine-tuning technique was applied, our program still performs reasonably well when it comes to its language abilities, much because of the base model used (bineric/NorskGPT-Llama-13B) already being fine-tuned on "a carefully selected mix of Norwegian instruction pairs" [43]. Please refer to Section 3.2 to see the reasoning behind the choosing of this model. Ultimately the group found after discussing with the supervisors, as well as the client, that incorporating fine-tuning would have taken a lot of time, with less of a guarantee of actually producing considerable enhancement of output.

The fact that LLMs that come pre-trained, or fine-tuned on some specific task, often have trouble accessing specific data could be particularly unwanted in domains that require high reliability [30], such as the legal industry. This ultimately was the main reason for our choosing to prioritize the RAG framework. As Gao et al. state: models that do not implement nonparametric memory faces challenges such as "[...] hallucination, outdated knowledge, and non-transparent, untraceable reasoning processes" [29]. All of these directly relate to this problem of accessing and maneuvering specific data. The problem of hallucination in particular, together with an apparent bias that is either inherited from the training data or through an "amplification of the hallucination behaviour" [34], must be eradicated for LLMs to be fully trusted in official legal rulings. Should an end-user for instance inquire on earlier given GDPR case-rulings for comparison, and through some reasoning process not apparent to the human eye the program hallucinates, that could have dire consequences in the legal industry.

Furthermore, RAG was chosen partly for its ease of use. Even though the program wasn't specifically developed with user experience and user design in mind, the functionality of RAG already makes it easy to link many new documents to the vector database as needed. For instance, other legal regulations could be applied, or even actual earlier GDPR related case-rulings for the program to use as extra, highly relevant context. The ease of implementing RAG also helped us spend our time on actual experimenting, rather than technical detailing.

6.2.3 Potential Fine-tuning Methods

If we were to perform fine-tuning to some extent, either that be to enhance the programs language capabilities or for it to gain a broader understanding of the law domain, there are several techniques which could be relevant.

What we initially realised was the resource demand it would take for any fine-tuning, let alone pre-training, to actually have an effect. Based on this, we recognized the need for any fine-tuning to be highly efficient. We looked at many different techniques and implementations, and the following are some examples. An overview of common efficient fine-tuning methods can be seen in Figure 6.1.

- **LoRA**
    - LoRA originally was introduced to work with the transformer model architecture , targeting its self-attention layer in particular. It essentially works by freezing the pre-trained, original parameters, and adding new more compact sets of parameters to be tuned instead. This can reduce the required GPU memory i.a.[44]. Another variation of LoRA is QLoRA [45].

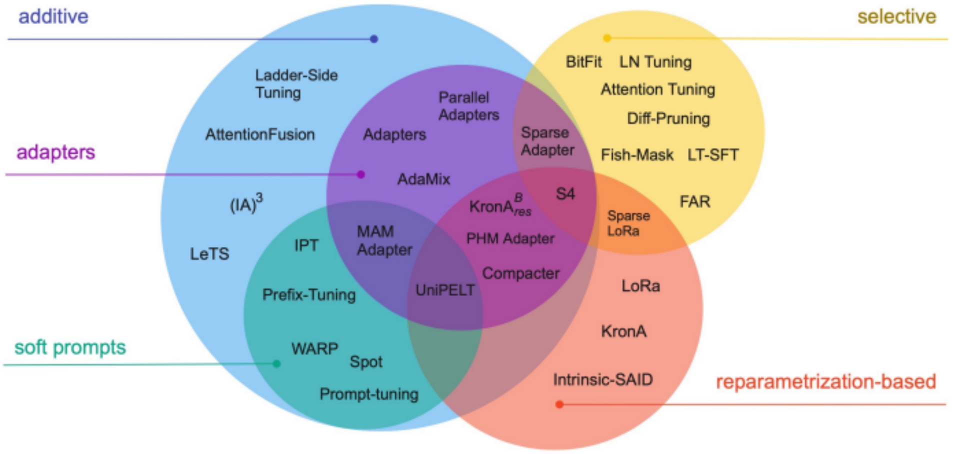


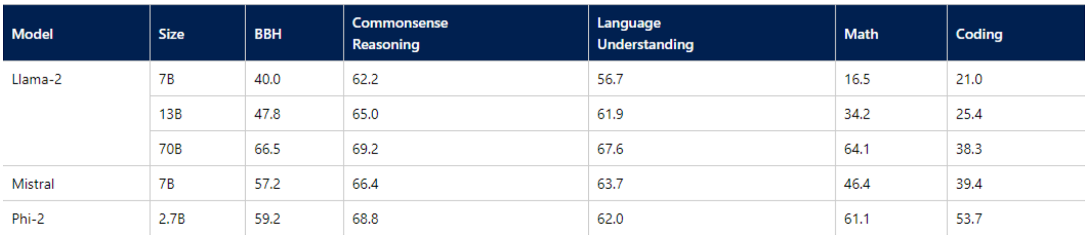
Figure 6.1: Taxonomy of parameter efficient fine-tuning methods. [42]

- **Adapters**
    - Functions similarly to LoRA, but has been shown to "require additional inference latency" [44]. Works by introducing small neural network modules between layers of a pre-trained model. It has been shown to produce similar results to full fine-tuning[42].
- **Attention Tuning**
    - Can work by fine-tuning the attention layer parameters themselves to make them focus on certain parts of the input. Thereby only minimal changes to the pre-trained parameters are required.

A common denominator for all of these are that they require minimal modification of the model itself, "[...] by only training a small set of parameters"[42]". Any of the techniques mentioned here could be a viable next-step for our program.

6.2.4 Why Only One PDF Document?

As stated previously, through implementing RAG in a generative AI system the problem of the LLM generating fictitious content largely dissipates. Nevertheless, the question of what, and how much, data that should be included in a knowledge base arises. The group found this to be a question with rather varied answers. For fine-tuning at least, the general trend has been that including an extensive amount of training-data directly relates to the performance of the model [46][47][48]. Keeping in mind that RAG is a newer, but nevertheless related technology, one might assume that the same logic could be applied here. Moreover, not only the amount of data, but the importance of the size of the base model itself has also



Model	Size	BBH	Commonsense Reasoning	Language Understanding	Math	Coding
Llama-2	7B	40.0	62.2	56.7	16.5	21.0
	13B	47.8	65.0	61.9	34.2	25.4
	70B	66.5	69.2	67.6	64.1	38.3
Mistral	7B	57.2	66.4	63.7	46.4	39.4
Phi-2	2.7B	59.2	68.8	62.0	61.1	53.7

Figure 6.2: Comparisons of SLM performance on multiple benchmarks[49]

recently been in some dispute. These two aspects will be discussed further in the following subsections.

What we once again see is the challenge of finding the right balance between various factors in developing a generative AI system. Even the degree of which only relevant documents should, or should not, be included in the knowledge base is not entirely clear. We find that the details of one's implementation differs depending on the use-case that the program is meant to handle, and what domain it affects.

Model Size

Liu et al states: "The Transformer architecture is exceptionally well-suited for scaling up models, and research analysis has revealed that increasing the model's scale or training data size can significantly enhance its performance" [46]. A similar understanding is common across other relevant research papers: Minaee et al. [47] for instance discusses how larger models both understand user input better, and also generates better output than small models. This correlation between size and performance was initially a guiding principle for the group. We figured that to encapsulate a broader range of nuances in understanding GDPR regulations, we would have to utilize the most substantial model possible.

The fact that small models were inferior also thus became the group's own understanding after a significant time of research into the subject. Not long after the thesis started however, Microsoft released the model Phi-2 [49]. They had earlier released Phi-1 and Phi-1.5, both with 1.3 Billion parameters and both making an argument for the potential of Small Language Models (SLMs). Phi-2 itself has a "mere" 2.7 billion parameters, which is a drastic drop from some of the largest models at up to a trillion. In Figure 6.2 the performance of Phi-2 on different factors can be seen. Largely because of the release of Phi-2, but also the performance results of the still quite small Llama-2 variations [50], we became increasingly curious about utilizing smaller models. We did also experiment some with the Phi-2 models, but ultimately we saw some drop in its general ability to handle other languages than English, including of course Norwegian.

The group experimented with different models of various sizes, which is further detailed in Section 3.2. In the end we settled on "bineric/NorskGPT-Llama-13B". This model ultimately has a more optimal balance between performance and practicality, and thus our program could remain both effective and more sustainable. Moreover, the practical constraints especially in regards to available computational resources also had a big influence on the decision to focus on models in this range of parameters.

Optimal Amount of Data

When it comes to the specific amount of data to include with a QA-program that utilizes RAG, the group faced similar challenges. It is important to reiterate that our program was always meant to be characterized as a QA-program rather than a chatbot. If it were instead to function as a chatbot the reasoning behind choosing the optimal amount data would potentially change. Having a broad amount of data could be advantageous to increase the program's ability to provide more comprehensive output. The way one interacts with OpenAI's ChatGPT models[38] for instance, relies heavily on it being able to remember, and utilize, previous chat context in answers to subsequent user-input. Such models would have to handle a lot of extra context, on top of its already extensive knowledge base. The subject of incorporating memory to our program is discussed further in Section 7.3.

Alternatively, a program with an extensive knowledge base could have challenges in finding the most applicable information. Not to mention the extra potential inference time that may be a consequence of the program having to sift through more irrelevant data. For highly specialized applications that must comply with regulations and be able to give legally sound advice, having to much data could reduce the program's effectiveness.

Given these considerations, the amount of data to incorporate would rely heavily on the requirements of the application. For our project we used just one highly relevant and authoritative PDF document, namely the "Personal Data Act" (Section 4.4.1). By limiting the volume of data, we were instead able to focus on what extra RAG technologies to include to enhance the program's ability to extract appropriate information. Thus we deemed including one document to be a more simple set up to experiment with. Furthermore, we believed that the specificity of said document would allow for the program to be more tailored to the particular domain of GDPR compliance. However, for our program to be able to generate more nuanced and technical suggestions, more data would certainly have to be included in the vector DB. Although recently Cuconasu et al. have produced rather surprising results in regards to the inclusion of seemingly irrelevant documents[51]. Contrary to common belief they suggest that introducing more "data noise", with documents of no immediate relation to the query, actually improves

LLM accuracy. They mention that the mechanics of these result are not yet understood. If these findings turn out to be factual, this would again bring up the question of hallucination in LLMs. Both enhanced complexity in the retrieval components, as well as the prevalence of irrelevant data could lead to increased hallucination in LLM output.

6.2.5 Practical Limitations

Our main limitation for the development of the program is a lack of computing resources. This limits the size of model we can use, and the time in which the models we do run take to complete.
A larger model would be able to process more text in the query, allowing for both more context to be included and a larger answer in return. Alternatively more processing power would also have allowed us to run the same size model we currently do, but at a faster pace, allowing for either more advanced methods to be used in the program or quicker iteration for testing.
There are some methods we would have liked to implement, but due to the added time it would have taken to run the program had we implemented them, we deemed it necessary to leave those out for now. These methods use LLM calls to evaluate some part of the process.
One is to implement a re-ranker. The re-ranker would have to run an LLM call for each gathered piece of context to evaluate the relevance of the retrieved document, which would increase the time spent per iteration of the program massively. In addition, to get the full effect of the re-ranker we would have gathered more context for it to rank, further adding to the time needed to run it.

6.3 The Workflow of the Group

This section contains an assessment of the workflow of the group and the process of working on the project. Some suggestions as to what improvements could be made during the work are also presented.

6.3.1 Delegation of Responsibilities

When starting the project, the group immediately decided it would be best to divide and delegate different tasks and working areas among the group. We identified our individual strengths and motivations, which gave us the following distribution of responsibilities early in the project:

- **Torje:** Took the lead on the coding aspects and developing the program, considering he had the stronger coding and python knowledge among the group members.
- **Emil:** Focused on the program development and testing, as well as researching AI theory and which methods and technologies that would be best applicable for our project.
- **Paul:** Focused on doing research and information gathering, looking into relevant research papers, as well as the starting the early writing of the thesis. He was also responsible for creating the GDPR dataset.

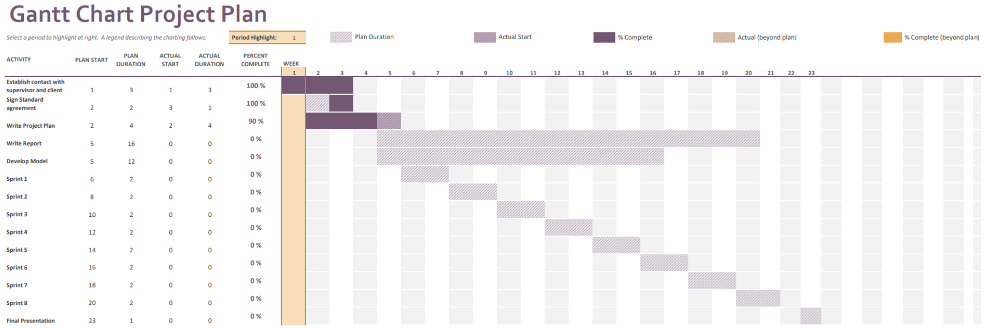
The delegation of responsibilities was a good idea, but in hindsight there were some changes that could have been made. The research phase took longer than anticipated so it would have been beneficial to divide more concrete areas of responsibility. The phase of implementation could then be more effective, while work on the actual writing of the thesis also could be prioritized earlier.

6.3.2 Gantt and Project Development

To maintain a consistent workflow throughout the project, a Gantt chart was made in the planning phase, which roughly planned our activities from start to finish over the semester. This, along with the sprints, structured the project around a series of phases and allowed us to adapt and respond to challenges dynamically. However, as will be discussed in Section 6.4.4 there was some troubles that ultimately hindered optimal use of Gantt. To have a rough structure to the proceedings of the project was helpful due to the otherwise flexible nature of developing a LLM, as well as using Scrum as the project development model. The specifics of each sprint in the chart was however rather hard to predict before beginning the actual sprint segment, although the overall deadlines of the sprints did ensure continuous progress. As the delimitation of the scope developed, each segment was filled in with concrete tasks and goals and added to the Scrum board in Jira. This was helpful because of the added responsibility of completing one's assigned tasks, while easily visualising the progression. An image of the Gantt chart can be seen in Figure 6.3.

6.3.3 Meetings

Much because of the usage of Scrum, the process of having meetings were quite extensive, which is further detailed in Section 3.1.4. The group found this process to be of substantial help. The bi-weekly sprint setup, where the first day of each new sprint contained meetings with both the supervisors and the client, did provide natural milestones to the project. This provided the group with time to



Gantt Chart Project Plan

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight: 1

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Establish contact with supervisor and client	1	3	1	3	100 %
Sign Standard agreement	2	2	3	1	100 %
Write Project Plan	2	4	2	4	90 %
Write Report	5	16	0	0	0 %
Develop Model	5	12	0	0	0 %
Sprint 1	6	2	0	0	0 %
Sprint 2	8	2	0	0	0 %
Sprint 3	10	2	0	0	0 %
Sprint 4	12	2	0	0	0 %
Sprint 5	14	2	0	0	0 %
Sprint 6	16	2	0	0	0 %
Sprint 7	18	2	0	0	0 %
Sprint 8	20	2	0	0	0 %
Final Presentation	23	1	0	0	0 %

Figure 6.3: The Gantt chart that was made during the planning of the project

recuperate and plan for the sprint ahead based on the feedback we got. However, some small alterations could be made. We could for instance have made a more complete agenda before each meeting, especially with the supervisors. The meetings were also quite informal, so there were times where productiveness could be increased. On the other hand, the feedback we did get were mostly quite helpful, and it helped to ground the group a bit in-between sprints. Furthermore, the meetings with the client did not always serve a particular purpose. This was largely due to the lack of specific requirements from the client. The meetings within the group were mostly good, although at the later phases they naturally became a little scarce. This was due to the often individual work that could be done in regards to the writing of the thesis.

6.3.4 Decision Making Process

The group's decision making process was influenced by ongoing testing results of the program in the early phases, using the initial GPT provided QA-dataset Figure C.1, which assessed the language capabilities and its ability to provide answers. Feedback and advice from the supervisors and the client also played a big role in the decision making and steering the project in the right direction. As the project progressed, considering supervisor feedback and available computing resources, the scope underwent several changes. For instance, it was steered from creating a general Norwegian AI legal assistant, to making it specialized in a specific area: GDPR. This was on the basis of a suggestion from the supervisors.

6.3.5 Submission of Drafts

The process of submitting drafts to the supervisors could be improved. The first draft of the thesis was delivered pretty late, and that resulted in less time to ameliorate. Had we instead, together with the supervisor, specified an exact date for draft submission, more time could be spent on quality assurance. The ideal submission date for the first draft would have been some time in mid-April.

6.4 Project Reflections

6.4.1 Sustainability

A common way to define sustainable development was first introduced by the World Commission on Environment and Development, otherwise known as the Brundtland Commission: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [52]. One could also say that "sustainability is the balance between the environment, equity and economy" [53]. As the prevalence of AI keeps growing it is natural we should also assess how this new technology, through development and use, could affect the world's values in regards to sustainability. Moreover, the speed of which AI is being developed has been of some concern, and it has long been argued that we need to put in place strong guidelines and regulations of AI development [54] [55]. The challenge then becomes balancing the speed of development and innovation against the potential need for regulation, without "compromising future generations".

An area that AI might impact is the environment. Although the actual impact of AI on the climate differs depending on the perspective. The difference in emission whether the LLM is being trained (Figure 6.4) or the LLM post-training (Figure 6.5) is substantial. Of course it would be appropriate to look at the source of electricity for these processes to fully understand the numbers. In the grand perspective of things, considering AI is still a small part of the IT sector, the overall environmental impact of LLMs still remain quite small. Smith et al. [56] suggest that as long as AI vendors fine-tune only based on specific program-demands, for instance by modifying only a small amount of parameters, this should help ensuring a sustainable future for AI.

Except for the climate in particular there are, as mentioned, other aspects of sustainability. Perhaps the most appropriate way to consider the effect of AI on sustainability is via UN's Sustainable Development Goals [57]. There are 17 of them in total, and many could relate to AI and our project specifically. One goal for instance, is goal 16 - "Peace, justice and strong institutions". One potential outcome that an AI legal assistant might have is that it could promote and make objective rulings, hopefully free of any human biases etc. It would also make legal texts more accessible for lawyers, which in turn would strengthen the institutions. Furthermore, provided AI could help inform the public of their rights, and help evening out differences and discrimination, this might also have an effect on goal

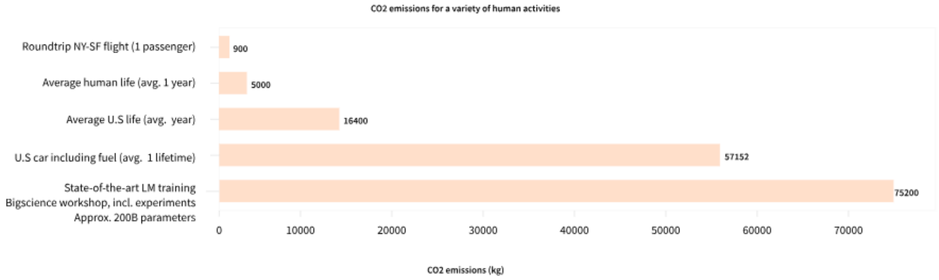


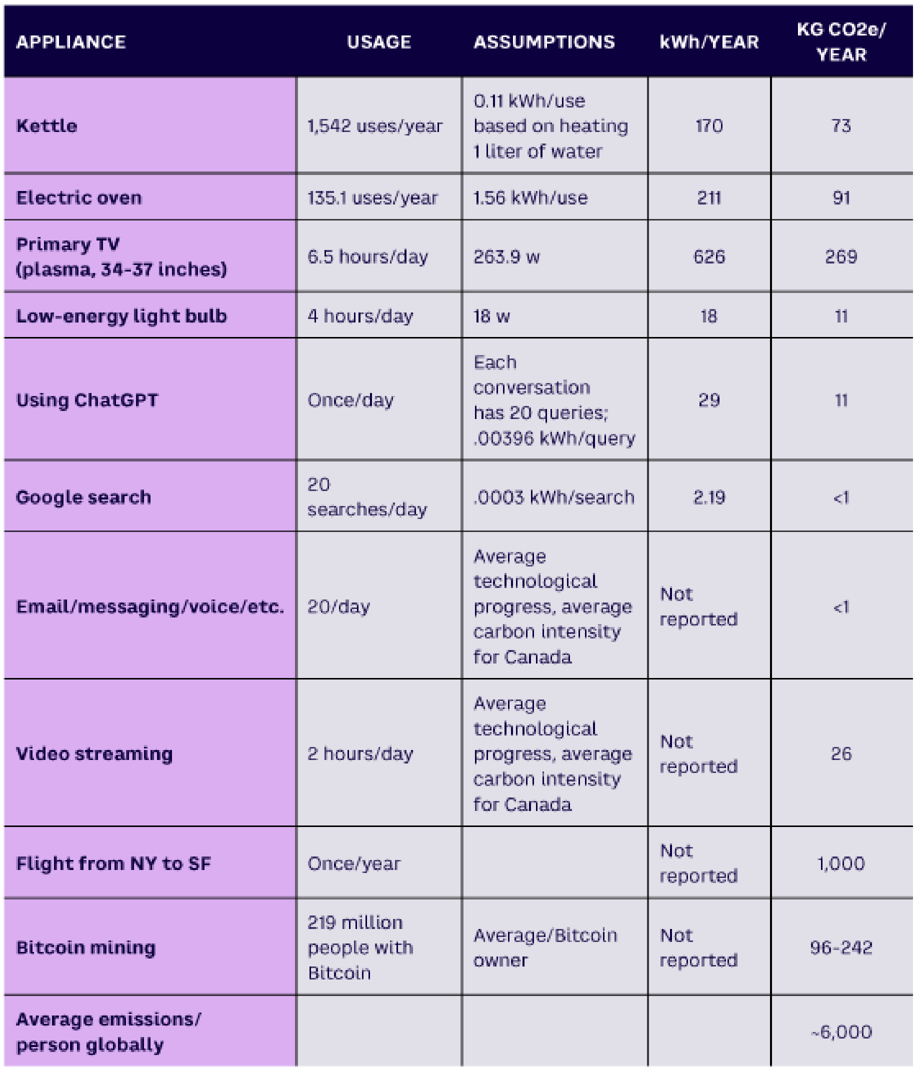
Figure 6.4: Environmental impact of LLM training compared to other activities. [48]

10 - "Reduce inequalities".

The potential negative effects of AI, in regards to the climate, relates to goal 7 - "Affordable and clean energy", 12 - "Responsible consumption and production" and 13 - "Climate action" especially. It is important that the AI industry minimize its participation in overconsumption, which is a significant factor contributing to climate change. Also important to note is the discussion of AI putting people out of jobs. On one side there are sure to be excess jobs once AI fully enters the workplace, which of course would immediately decrease the availability of decent work. On the other hand AI might also promote safer, more forward-looking employment in the long run. This last point relates to goal 8 - "Decent work and economic growth". All in all, whether or not regulations should be put on the development of AI, one must definitely encourage forward-thinking innovation, while still adhering to the many goals of sustainability.

6.4.2 Ethical Considerations

As discussed earlier, with the rise of AI one must also consider how to guide and potentially regulate its development. Thereby there is also a question as to the ethics of working with AI. For the duration of this project the group has attempted to adhere strictly to ethical AI development practices via for instance UNESCO's "Recommendation on the Ethics of Artificial Intelligence" [58]. Furthermore, while it is only natural to consider the ethics of AI in this project, we must also note the importance of legal considerations. Specifically aspects related to GDPR compliance. The two perspectives will we discussed further in this section.



APPLIANCE	USAGE	ASSUMPTIONS	kWh/YEAR	KG CO2e/ YEAR
Kettle	1,542 uses/year	0.11 kWh/use based on heating 1 liter of water	170	73
Electric oven	135.1 uses/year	1.56 kWh/use	211	91
Primary TV (plasma, 34-37 inches)	6.5 hours/day	263.9 w	626	269
Low-energy light bulb	4 hours/day	18 w	18	11
Using ChatGPT	Once/day	Each conversation has 20 queries; .00396 kWh/query	29	11
Google search	20 searches/day	.0003 kWh/search	2.19	<1
Email/messaging/voice/etc.	20/day	Average technological progress, average carbon intensity for Canada	Not reported	<1
Video streaming	2 hours/day	Average technological progress, average carbon intensity for Canada	Not reported	26
Flight from NY to SF	Once/year		Not reported	1,000
Bitcoin mining	219 million people with Bitcoin	Average/Bitcoin owner	Not reported	96-242
Average emissions/ person globally				~6,000

Sources: Carbon Footprint, Medium, Full Fact, Luciano Rodrigues et al., *The Guardian*, Crypto News, Our World in Data

Figure 6.5: Environmental impact of ChatGPT interaction. [56]

AI Development

There are many aspect to ethical development of AI, but the overarching goal must perhaps be to ensure that AI is developed and used in accordance with human rights and dignity, while also remembering to consider its impact on the environment. For instance, the goal of our program has not been for it to replace humans in their work, but rather it is meant as a supplement to streamline the workday of lawyers. Another important ethical consideration is the prevalence of bias in AI and LLMs, for instance gender bias. This was one of the main reasons that the group implemented RAG as opposed to using fine-tuning, as previously discussed. There were however little measures that could be performed to deal with the inherent bias one can expect to find in (pre)training datasets. To thoroughly prevent our program from generating explicit and unwanted text, we probably would have to closely curate and clean the training datasets. Other possibilities could be to implement, blocklists, filters or other mechanisms to detect and hinder inappropriate content. Another point the group considered for the project is the importance of human oversight, as well as the effects that training and using AI has on the environment. If an application of our program were to be launched at some time in the future, it is clear that human oversight would be essential to comply with point 5 ("Responsibility and Accountability") of UNESCO’s 10 point human rights approach to AI [58].

Legal

Although there are multiple bodies of work that is meant as guidance for the development of AI, there are not yet any official legally binding "rules". That is, except when they intersect with enforceable laws like GDPR. It has been natural that the group have considered the need to comply with GDPR, although the AI legal assistant has not been designed to process user input containing personal data directly. Through working with AI, there is a considerable risk in the way it may handle personal data, perhaps by it being present, to some degree, in the training data. As previously stated, the training data was not closely regulated for the purpose of this project and so there is an inherent risk in the program handling said potential personal data. For a responsible future scaling of the program appropriate measures and mechanisms would have to be explored and implemented. Nevertheless, the group has adhered to the point of data minimization in regards to the retrieval component of the program. "In accordance with Article 5(1)(c) of the General Data Protection Regulation (GDPR), personal data must be ’adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed’ (data minimization)." [59]. Only a highly specific document, namely the Personal Data Act, is currently included in the nonparameteric memory, and thereby it is the only retrievable document for extended context.

This was discussed further in Section 6.2.4.

6.4.3 Use of AI

The utilization of artificial intelligence was permitted as a supplementary tool for the preparation of this thesis, contingent upon the completion of an AI declaration form in accordance with the university's requirements. The use of AI tools has been used merely as a ways to enhance effectiveness, inspire or otherwise assist the group in various aspects of the project. The specific tools that have been used are OpenAI's ChatGPT 3.5 and 4 [38]. The following are cases where AI has been utilized:

- The scoring system, detailed in Section 5.3.2, was partially generated and adopted with some modifications.
- For spell checking or suggestions as to how to formulate sentences in English, given that the thesis' contributors' first language is Norwegian.
- Group 5 of the sample answers, found in Appendix E, that was given to the GDPR expert for evaluation/comparison of the programs' performance was fully generated. The exact answer from ChatGPT 3.5 can be seen in Appendix C.1.
- A simple dataset containing suggested questions and answers (found in Appendix F). We used this for experimentation in regards to which model configurations and RAG technologies we should utilize.

6.4.4 Critique of the Thesis

Structuring of Project Phases

During the later phases of the project we felt the apparent consequences of not having defined more concrete milestones. We did propose a Gantt chart for the project plan in an attempt to gain some structure, and the chart was generally helpful, although some troubles arose. For instance, the time we had to spend initially on acquiring enough knowledge about the field delayed proceedings to some extent. We should also have switched overall focus to the writing of the report earlier, as not much else than writing could take precedence in the last phase of the project. For this same reason, We should have ended the implementation phase at an earlier stage as well. Moreover, there was some unexpected delays to the GDPR expert-evaluation of our program that ultimately made implementing other evaluation metrics hard.

Multiple Evaluation Metrics

For the evaluation of the program's performance there could have been several improvements. Considering that besides our own testing when deciding on what technologies to implement, we did only have one point of evaluation: the scoring provided by the GDPR expert. We are aware that this would not typically suffice for a true performance evaluation. Even though the expert provided much valued extra competence, this way of evaluation is rather subjective. There are multiple ways we could have enhanced the evaluation aspects of the thesis. First, there are standardized frameworks such as: BLEU[60], for evaluating text quality after natural language translations, and F1[61] for scoring the balance between recall and precision. Another possibility could be to implement "LLMs-as-judges"[62] for simulating human evaluations.

Git

The general idea from the start was to utilize Git for easy cooperation, version control and backup of our code. Using Git would have also been highly compatible with the flexible project development model that the group utilized: Scrum. In the end this idea was rather neglected. The main reason for this was the defined roles that each group member were assigned. Although a couple of members worked on research of technologies, a single member had the code as a main responsibility. Moreover, the fact that that we had access to a VM and that we were using Miniconda environments further decreased the need for Git. This is due to the rapidly changing experimental setups, together with the use of other flexible libraries and frameworks. The group did however utilize Git for the purpose of backing up the source code of the program.

Realistic Scope

From the start of the project the group was given a lot of flexibility in choosing the scope, as the client did not have specific program requirements. This, together with the fact that we had little previous hands-on experience with AI technologies, ultimately made settling on a scope hard. We spent a lot of time on the initial research, and had several discussions with the supervisors, just to be able to agree on a practical scope for the thesis. Ideally this accumulation of time would instead have been spent on actual experimentation and evaluation.

Chapter 7

Conclusion

7.1 Achievement of Goals

In this section we judge to which degree we believe the goals of the project have been met. These goals are described in Section 1.6.

Main goal: *To develop and evaluate an AI-powered legal assistant that specializes in GDPR. This AI assistant should be able to understand and respond accurately in Norwegian.*

> This was achieved to a degree. We have a functioning program, albeit without a ready-to-use application. We did perform evaluations, although several improvements to the process could have been made, as discussed in Section 6.4.4. The capabilities and concrete results are described in Chapter 5. It handles the Norwegian language reasonably well.

The points below refer to the project goals as documented in Section 1.6. The letter refers to the type of goal, and the digit which number in the sequence it is. The goal text is provided below as it is in the introduction. First are the result goals (R), followed by the learning goals (L), then lastly the impact goals (I).

R1: *To produce a comprehensive report on the process of developing a functional GDPR specialized AI assistant.*

> This was achieved by the creation of this thesis.

R2: *To produce results and ideas that could potentially be iterated by us or others in the future.*

We believe our results have laid a foundation for further research and development of our product in the future.

R3: *To assist Futurize AS with insights into LLM modifications, as they aim to develop an easy-to-use solution for future customers.*

We have provided useful insights and techniques focused around RAG, for LLM modifications. But when it comes to the solution being easy-to-use, this has not been completely fulfilled, as we did not focus on the user interface aspect, and since it takes a lot of time for the program to provide answers in its current form on the current hardware.

L1: *To acquire knowledge of AI. Specifically about correlated theories.*

This was achieved. The group has gained extensive knowledge on the subject and related fields and theories.

L2: *To practice common workflows and methodologies that future employers may expect us to be acquainted with.*

The group has used Scrum and DSR, which are widely recognized methodologies used in both academic and professional settings.

L3: *To work in a highly structured and independent manner.*

This has been the experience of all group members.

L4: *To discern the impact that the continued development of AI will have on the workplace.*

This was achieved. Many possible consequences of incorporating AI has been explored.

I1: *To provide new insights in regards to using techniques for specializing LLMs.*

We have not provided any new insights in particular, as we have mostly used and combined pre-existing techniques for specializing an LLM.

I2: *To provide an easy to use solution for future customers of Futurize.*

This has not necessarily been achieved considering the lack of a functioning user interface and the substantial inference time of the program.

7.2 Contribution of the Study

The continuing expansion in complexity in regards to the massive corpora of regulations, results in a dire need for simplified methods for insight. This is especially true for lawyers, in view of the responsibility they hold for their clients. There are multiple ways in which introducing a QA generative AI might be of benefit to lawyers. For instance, routine work and other GDPR related inquires will be highly streamlined. This makes it possible for lawyers to spend their resources where they are most needed. One would also achieve less dependency in human competence and subjectivity, presuming that the QA program is able to provide unbiased answers. Hopefully this could result in greater consistency in legal advice, research and document preparation. Furthermore, AI might be more cost-effective in that it could make legal services more accessible and also reduce labor cost in law firms.

As stated in Chapter 1, the objective of this thesis was to produce an AI legal assistant for the benefit of lawyers. Even though a complete ready-to-use application have not yet been made, the current findings can still be useful for lawyers with little GDPR experience by providing them simple information and guidance on GDPR related cases. Furthermore, the group has used recent advancements in the field to provide insights into the possibilities of optimizing a working assistant. The findings of this thesis lay some groundwork in regards to the effectiveness of various model configurations and implementations. Our progress facilitate future work in that there are multiple ways of continued exploration, and new research in the field is being conducted every day.

7.3 Future Work

There are some key unanswered questions that would be beneficial to explore. For instance, for more exact performance evaluation of our program, as is mentioned in Section 6.4.4 we should have implemented more standardized and objective evaluation metrics. Doing so could substantiate our findings. It would have also been easier to compare the performance against findings from other related research.

Given that the field of LLM development is in constant expansion one can imagine that new and improved technologies can be applied to our program in the future. However, when it comes to specific areas of improvement with the technologies of today, there are several suggestions:

- Utilize models with a greater amount of parameters, or newly launched models such as Llama-3-8B[63]. This model has a double context length compared to LLama-2-13B[64].

- Make a functioning application so that the program could be tested by end-users. This would also make for direct input as to ways to improve the program.
- Incorporate memory, or chat history, to our program. This would be useful given that users could have follow-up queries. This would however make a program that functions more as a chatbot rather than a QA-system.
- Implement another LLM in the architecture that takes the suggested output from our program as input, to ensure sensible answers every time. Our program by itself does not guarantee this. This would introduce some more robustness to our implementation.
- Explore fine-tuning techniques. This could improve the program's Norwegian abilities, as well for gaining a more nuanced understanding of the law domain. Some concrete techniques are described in Section 6.2.3.

Moreover, as has been discussed in Section 6.2.4, there is also room for scaling in regards to the amount of data that is included in the external knowledge base. For instance, it could be of interest to include GDPR related case-rulings, so that the program would be able to draw from actual interpretations of the regulations. Incorporating similarly related law documents could improve the program's ability to see indirect connections, that is not immediately apparent to the human eye.

Bibliography

[1] ‘About futurize.’ Message received on Teams; translated from Norwegian. (2023).

[2] K. Martineau. ‘What is generative AI?’ IBM Research Blog. (9th Feb. 2021), [Online]. Available: `https://research.ibm.com/blog/what-is-generative-AI` (visited on 28/02/2024).

[3] ‘What is artificial intelligence (AI)? | IBM,’ What is artificial intelligence (AI)? (), [Online]. Available: `https://www.ibm.com/topics/artificial-intelligence` (visited on 05/04/2024).

[4] ‘What’s the difference between CNN and RNN?’ (), [Online]. Available: `https://www.telusinternational.com/insights/ai-data/article/difference-between-cnn-and-rnn` (visited on 26/04/2024).

[5] ‘CNN vs. RNN: How are they different? | TechTarget,’ Enterprise AI. (), [Online]. Available: `https://www.techtarget.com/searchenterpriseai/feature/CNN-vs-RNN-How-they-differ-and-where-they-overlap` (visited on 27/04/2024).

[6] ‘Transformer model | IBM.’ (9th Apr. 2024), [Online]. Available: `https://www.ibm.com/topics/transformer-model` (visited on 26/04/2024).

[7] ‘Generative AI explainer,’ AI for Education. (), [Online]. Available: `https://www.aiforeducation.io/ai-resources/generative-ai-explainer` (visited on 27/04/2024).

[8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, *Attention is all you need*, 1st Aug. 2023. arXiv: `1706.03762[cs]`. [Online]. Available: `http://arxiv.org/abs/1706.03762` (visited on 30/04/2024).

[9] S. Cristina. ‘The transformer model,’ MachineLearningMastery.com. (17th Sep. 2022), [Online]. Available: `https://machinelearningmastery.com/the-transformer-model/` (visited on 30/04/2024).

[10] ‘What is natural language processing? | IBM.’ (), [Online]. Available: `https://www.ibm.com/topics/natural-language-processing` (visited on 17/04/2024).

[11] 'Retrieval augmented generation (RAG) for LLMs – nextra.' (), [Online]. Available: `https://www.promptingguide.ai/research/rag#naive-rag` (visited on 19/04/2024).

[12] 'Natural language processing (NLP) - a complete guide.' (11th Jan. 2023), [Online]. Available: `https://www.deeplearning.ai/resources/natural-language-processing/` (visited on 30/04/2024).

[13] 'What are large language models (LLMs)? | IBM.' (), [Online]. Available: `https://www.ibm.com/topics/large-language-models` (visited on 05/04/2024).

[14] 'What are large language models? - LLM AI explained - AWS,' Amazon Web Services, Inc. (), [Online]. Available: `https://aws.amazon.com/what-is/large-language-model/` (visited on 17/04/2024).

[15] 'Large language models explained,' NVIDIA. (), [Online]. Available: `https://www.nvidia.com/en-us/glossary/large-language-models/` (visited on 29/04/2024).

[16] 'What is a chatbot? | IBM.' (5th Apr. 2024), [Online]. Available: `https://www.ibm.com/topics/chatbots` (visited on 29/04/2024).

[17] 'Question answering - hugging face.' (18th Mar. 2024), [Online]. Available: `https://huggingface.co/tasks/question-answering` (visited on 29/04/2024).

[18] L. Eliot, *AI and legal argumentation: Aligning the autonomous levels of AI legal reasoning*, 11th Sep. 2020. DOI: `10.48550/arXiv.2009.11180`. arXiv: `2009.11180[cs]`. [Online]. Available: `http://arxiv.org/abs/2009.11180` (visited on 16/05/2024).

[19] -. 'How artificial intelligence is used in legal practice,' Bloomberg Law. Section: Technology. (1st Aug. 2023), [Online]. Available: `https://pro.bloomberglaw.com/insights/technology/ai-in-legal-practice-explained/` (visited on 16/05/2024).

[20] M. De'Shazer, *Advancing legal reasoning: The integration of AI to navigate complexities and biases in global jurisprudence with semi-automated arbitration processes (SAAPs)*, version: 3, 29th Feb. 2024. DOI: `10.48550/arXiv.2402.04140`. arXiv: `2402.04140[cs]`. [Online]. Available: `http://arxiv.org/abs/2402.04140` (visited on 16/05/2024).

[21] D. M. Katz, D. Hartung, L. Gerlach, A. Jana and M. J. Bommarito II, *Natural language processing in the legal domain*, 23rd Feb. 2023. DOI: `10.48550/arXiv.2302.12039`. arXiv: `2302.12039[cs]`. [Online]. Available: `http://arxiv.org/abs/2302.12039` (visited on 16/05/2024).

[22] 'Ordvalg i lovtekster.' (), [Online]. Available: `http://www.sprakradet.no/klarsprak/om-skriving/sprak-i-lover-og-forskrifter/skriverad/ordvalg/` (visited on 06/05/2024).

[23] 'What is GDPR, the EU's new data protection law?' GDPR.eu. Section: GDPR Overview. (7th Nov. 2018), [Online]. Available: `https://gdpr.eu/what-is-gdpr/` (visited on 08/04/2024).

[24] D. Siemon. 'Methods in design science research,' Design Science Research. (14th Jan. 2021), [Online]. Available: `https://design-science-research.de/en/post/methods-in-dsr/` (visited on 06/04/2024).

[25] K. Peffers, T. Tuunanen, M. A. Rothenberger and S. Chatterjee, 'A design science research methodology for information systems research,' *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, Dec. 2007, ISSN: 0742-1222, 1557-928X. DOI: `10.2753/MIS0742-1222240302`. [Online]. Available: `https://www.tandfonline.com/doi/full/10.2753/MIS0742-1222240302` (visited on 12/04/2024).

[26] 'Scrum guide | scrum guides.' (), [Online]. Available: `https://scrumguides.org/scrum-guide.html#sprint-planning` (visited on 24/04/2024).

[27] 'Retrieval augmented generation: Streamlining the creation of intelligent natural language processing models.' (), [Online]. Available: `https://ai.meta.com/blog/retrieval-augmented-generation-streamlining-the-creation-of-intelligent-natural-language-processing-models/` (visited on 17/04/2024).

[28] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel and D. Kiela, *Retrieval-augmented generation for knowledge-intensive NLP tasks*, 12th Apr. 2021. DOI: `10.48550/arXiv.2005.11401`. arXiv: `2005.11401[cs]`. [Online]. Available: `http://arxiv.org/abs/2005.11401` (visited on 14/03/2024).

[29] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang and H. Wang, *Retrieval-augmented generation for large language models: A survey*, 27th Mar. 2024. arXiv: `2312.10997[cs]`. [Online]. Available: `http://arxiv.org/abs/2312.10997` (visited on 17/04/2024).

[30] J. Kim and M. Min, *From RAG to QA-RAG: Integrating generative AI for pharmaceutical regulatory compliance process*, 26th Jan. 2024. arXiv: `2402.01717[cs]`. [Online]. Available: `http://arxiv.org/abs/2402.01717` (visited on 14/02/2024).

[31] K. Guu, K. Lee, Z. Tung, P. Pasupat and M.-W. Chang, *REALM: Retrieval-augmented language model pre-training*, 10th Feb. 2020. arXiv: `2002.08909[cs]`. [Online]. Available: `http://arxiv.org/abs/2002.08909` (visited on 19/04/2024).

[32] A. Balaguer, V. Benara, R. L. d. F. Cunha, R. d. M. E. Filho, T. Hendry, D. Holstein, J. Marsman, N. Mecklenburg, S. Malvar, L. O. Nunes, R. Padilha, M. Sharp, B. Silva, S. Sharma, V. Aski and R. Chandra, *RAG vs fine-tuning: Pipelines, tradeoffs, and a case study on agriculture*, 30th Jan. 2024. arXiv: `2401.08406[cs]`. [Online]. Available: `http://arxiv.org/abs/2401.08406` (visited on 14/02/2024).

[33] 'Retrieval augmented generation (RAG) – nextra.' (), [Online]. Available: `https://www.promptingguide.ai/techniques/rag` (visited on 17/04/2024).

[34] Y. Zhang, Y. Li, L. Cui, D. Cai, L. Liu, T. Fu, X. Huang, E. Zhao, Y. Zhang, Y. Chen, L. Wang, A. T. Luu, W. Bi, F. Shi and S. Shi, *Siren's song in the AI ocean: A survey on hallucination in large language models*, 24th Sep. 2023. arXiv: `2309.01219[cs]`. [Online]. Available: `http://arxiv.org/abs/2309.01219` (visited on 20/04/2024).

[35] 'Openstack at NTNU - SkyHigh - NTNU wiki.' (), [Online]. Available: `https://www.ntnu.no/wiki/display/skyhigh/Openstack+at+NTNU` (visited on 06/05/2024).

[36] 'Using the commandline clients - SkyHigh - NTNU wiki.' (), [Online]. Available: `https://www.ntnu.no/wiki/display/skyhigh/Using+the+commandline+clients` (visited on 06/05/2024).

[37] R. Awasthi, S. Mishra, D. Mahapatra, A. Khanna, K. Maheshwari, J. Cywinski, F. Papay and P. Mathur, *HumanELY: Human evaluation of LLM yield, using a novel web-based evaluation tool*, 27th Dec. 2023. DOI: `10.1101/2023.12.22.23300458`. [Online]. Available: `http://medrxiv.org/lookup/doi/10.1101/2023.12.22.23300458` (visited on 16/05/2024).

[38] *ChatGPT*, version April 2023, 2024. [Online]. Available: `https://chat.openai.com`.

[39] . 'How to create a chatbot using your own data using RAG - lettria.' (), [Online]. Available: `https://www.lettria.com/blogpost/how-to-create-a-chatbot-using-your-own-data-using-rag` (visited on 03/05/2024).

[40] 'Processing the data - hugging face NLP course.' (), [Online]. Available: `https://huggingface.co/learn/nlp-course/chapter3/2` (visited on 06/05/2024).

[41] Z. Han, C. Gao, J. Liu, J. Zhang and S. Q. Zhang, *Parameter-efficient fine-tuning for large models: A comprehensive survey*, 17th Apr. 2024. DOI: `10.48550/arXiv.2403.14608`. arXiv: `2403.14608[cs]`. [Online]. Available: `http://arxiv.org/abs/2403.14608` (visited on 26/04/2024).

[42] V. Lialin, V. Deshpande and A. Rumshisky, *Scaling down to scale up: A guide to parameter-efficient fine-tuning*, 27th Mar. 2023. DOI: `10.48550/arXiv.2303.15647`. arXiv: `2303.15647[cs]`. [Online]. Available: `http://arxiv.org/abs/2303.15647` (visited on 26/04/2024).

[43] 'Bineric/NorskGPT-llama-13b-v0.1 · hugging face.' (), [Online]. Available: `https://huggingface.co/bineric/NorskGPT-Llama-13B-v0.1` (visited on 25/04/2024).

[44] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang and W. Chen, *LoRA: Low-rank adaptation of large language models*, 16th Oct. 2021. DOI: `10.48550/arXiv.2106.09685`. arXiv: `2106.09685[cs]`. [Online]. Available: `http://arxiv.org/abs/2106.09685` (visited on 26/04/2024).

[45] T. Dettmers, A. Pagnoni, A. Holtzman and L. Zettlemoyer, *QLoRA: Efficient finetuning of quantized LLMs*, 23rd May 2023. DOI: `10.48550/arXiv.2305.14314`. arXiv: `2305.14314[cs]`. [Online]. Available: `http://arxiv.org/abs/2305.14314` (visited on 26/04/2024).

[46] Y. Liu, H. He, T. Han, X. Zhang, M. Liu, J. Tian, Y. Zhang, J. Wang, X. Gao, T. Zhong, Y. Pan, S. Xu, Z. Wu, Z. Liu, X. Zhang, S. Zhang, X. Hu, T. Zhang, N. Qiang, T. Liu and B. Ge, *Understanding LLMs: A comprehensive overview from training to inference*, 5th Jan. 2024. arXiv: `2401.02038[cs]`. [Online]. Available: `http://arxiv.org/abs/2401.02038` (visited on 08/05/2024).

[47] S. Minaee, T. Mikolov, N. Nikzad, M. Chenaghlu, R. Socher, X. Amatriain and J. Gao, *Large language models: A survey*, 20th Feb. 2024. arXiv: `2402.06196[cs]`. [Online]. Available: `http://arxiv.org/abs/2402.06196` (visited on 08/05/2024).

[48] 'How do transformers work? - hugging face NLP course.' (), [Online]. Available: `https://huggingface.co/learn/nlp-course/chapter1/4` (visited on 28/04/2024).

[49] M. Javaheripi and S. Bubeck. 'Phi-2: The surprising power of small language models,' Microsoft Research. (12th Dec. 2023), [Online]. Available: `https://www.microsoft.com/en-us/research/blog/phi-2-the-surprising-power-of-small-language-models/` (visited on 14/02/2024).

[50] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave and G. Lample, *LLaMA: Open and efficient foundation language models*, 27th Feb. 2023. arXiv: `2302.13971[cs]`. [Online]. Available: `http://arxiv.org/abs/2302.13971` (visited on 08/05/2024).

[51] F. Cuconasu, G. Trappolini, F. Siciliano, S. Filice, C. Campagnano, Y. Maarek, N. Tonellotto and F. Silvestri, *The power of noise: Redefining retrieval for RAG systems*, 12th Feb. 2024. DOI: `10.48550/arXiv.2401.14887`. arXiv: `2401.14887[cs]`. [Online]. Available: `http://arxiv.org/abs/2401.14887` (visited on 19/04/2024).

[52] W. C. on Environment {and} Development, *Our Common Future*. Oxford University Press, 416 pp. [Online]. Available: `https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf`.

[53] a. 'What is sustainability?' UCLA Sustainability. (), [Online]. Available: `https://www.sustain.ucla.edu/what-is-sustainability/` (visited on 27/04/2024).

[54] C. Domonoske, 'Elon musk warns governors: Artificial intelligence poses 'existential risk',' *NPR*, 17th Jul. 2017. [Online]. Available: `https://www.npr.org/sections/thetwo-way/2017/07/17/537686649/elon-musk-warns-governors-artificial-intelligence-poses-existential-risk` (visited on 27/04/2024).

[55] S. Andriole. ‘Generative AI: Love, hate, ignore, or just regulate? | cutter consortium.’ (), [Online]. Available: `https://www.cutter.com/article/generative-ai-love-hate-ignore-or-just-regulate` (visited on 27/04/2024).

[56] G. Smith, B. Michael, R. Gillet and E. Thanisch. ‘Environmental impact of large language models.’ (24th Aug. 2023), [Online]. Available: `https://www.cutter.com/article/environmental-impact-large-language-models` (visited on 28/04/2024).

[57] ‘FNs bærekraftsmål.’ (), [Online]. Available: `https://fn.no/om-fn/fns-baerekraftsmaal` (visited on 27/04/2024).

[58] ‘Ethics of artificial intelligence | UNESCO.’ (), [Online]. Available: `https://www.unesco.org/en/artificial-intelligence/recommendation-ethics` (visited on 01/05/2024).

[59] *Regulation (EU) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (general data protection regulation) (text with EEA relevance)*, May 2016. [Online]. Available: `http://data.europa.eu/eli/reg/2016/679/oj`.

[60] ‘BLEU - a hugging face space by evaluate-metric.’ (), [Online]. Available: `https://huggingface.co/spaces/evaluate-metric/bleu` (visited on 10/05/2024).

[61] T. Hu and X.-H. Zhou, *Unveiling LLM evaluation focused on metrics: Challenges and solutions*, 13th Apr. 2024. arXiv: `2404.09135[cs]`. [Online]. Available: `http://arxiv.org/abs/2404.09135` (visited on 10/05/2024).

[62] L. Zheng, W.-L. Chiang, Y. Sheng, S. Zhuang, Z. Wu, Y. Zhuang, Z. Lin, Z. Li, D. Li, E. P. Xing, H. Zhang, J. E. Gonzalez and I. Stoica, *Judging LLM-as-a-judge with MT-bench and chatbot arena*, 23rd Dec. 2023. arXiv: `2306.05685[cs]`. [Online]. Available: `http://arxiv.org/abs/2306.05685` (visited on 10/05/2024).

[63] ‘Meta llama 3,’ Meta Llama. (), [Online]. Available: `https://llama.meta.com/llama3/` (visited on 11/05/2024).

[64] ‘Meta-llama/meta-llama-3-8b · hugging face.’ (18th Apr. 2024), [Online]. Available: `https://huggingface.co/meta-llama/Meta-Llama-3-8B` (visited on 11/05/2024).

Appendix A

Project Plan

DEPARTMENT OF INFORMATION SECURITY AND COMMUNICATION TECHNOLOGY

DCSG2900
BACHELOR-ASSIGNMENT BACHELOR IN DIGITAL INFRASTRUCTURE AND CYBERSECURITY

Project Plan

Authors:
Emil Bjerknes
Paul Bjørneng
Torje Bliksvær

Date: 01.02.24

Table of Contents

1 Goals and restrictions **1**

1.1 Background 1

1.2 Project goals 1

1.2.1 Result goals 1

1.2.2 Learning goals 1

1.2.3 Impact goals 2

1.3 Framework 2

2 Scope **3**

2.1 Subject area 3

2.2 Task description 3

2.3 Delimitation 3

3 Project organization **4**

3.1 Member Roles 4

3.2 Scrum Team roles 4

3.3 Group Rules 5

3.4 Routines 5

4 Planning, follow-up and reporting **6**

4.1 Project Development Model 6

4.2 Plan for status meetings and decision points during the period 6

5 Organization of quality assurance **7**

5.1 Documentation 7

5.2 Standards 7

5.3 Tools 8

5.4 Plan for testing 9

5.5 Risk analysis 9

6 Plan for execution **12**

6.1 Gantt chart 12

Bibliography **13**

Appendix **14**

A Group Contract **15**

A.1 Roles 15

A.2 Group rules 16

A.3 Consequences for violating the rules 16

1 Goals and restrictions

1.1 Background

Futurize AS is a startup company specializing in the development of innovative solutions for the legal business. Their primary focus is to explore the use of artificial intelligence to streamline and enhance the workflow of lawyers, with a particular emphasis on legal research as well as the preparation of documents and texts [1]. By integrating AI into their daily tasks, Futurize could revolutionize how lawyers do legal work and lead to significant improvements in efficiency and accuracy. To achieve this goal they want to use generative AI, which refers to deep-learning models that can generate high-quality text, images, and other content based on the data they are trained on [5]. In our case and for this assignment, the model will be trained on GDPR. This may be supplemented with Norwegian law sources to become better at Norwegian legal language.

1.2 Project goals

For this thesis the goals have been separated into effect goals, result goals and learning goals. The group's own goals are mainly presented under result goals and learning goals, while some of the client's goals and the potential effects of the thesis, are also included under impact goals.

The overall goal of the project is to develop and evaluate an AI-powered legal assistant that specializes in GDPR. This LLM should be able to understand and respond accurately in Norwegian.

1.2.1 Result goals

The result goals of this thesis revolve broadly around producing a report consisting of aspects related to developing an AI legal assistant, with specific focus on its performance in regards to GDPR. The concrete goals are:

1. To produce a comprehensive report.
2. To have a functional LLM with knowledge of GDPR, able to comprehend and use Norwegian legal language.
3. To produce results and ideas that can be iterated by us or others in the future.
4. To attain a respectable grade.

1.2.2 Learning goals

As AI is growing in popularity and becoming more and more relevant for the future, it is clear that we will benefit from a deeper understanding of related concepts. Our specific learning goals are:

1. To acquire knowledge of AI. Specifically about LLMs and correlated theories.

2. To practice common workflows and methodologies that future employers may expect us to be acquainted with.
3. To work in a highly structured and independent manner.
4. To discern the impact that the continued development of AI will have on the workplace.

1.2.3 Impact goals

The impact goals include:

1. The goal of Futurize is to streamline the way lawyers interact with the intricacies of law and legal sources.
2. To provide new insights in regards to fine-tuning LLMs.
3. To provide a easy to use solution for future customers of Futurize.

1.3 Framework

- Finish and deliver the Project Plan by 1st of February.
- Finish and deliver the Bachelor thesis by 21st of May.
- Final presentations are to be held on the 5th and 6th of June.

2 Scope

2.1 Subject area

The subject area for this thesis are boiled down to these main points:

- Legal study, especially GDPR in this project.
- Use of AI in legal work (the use of AI in legal content search, or predication of legal outcome).
- AI model training.
- Evaluation of AI models.

2.2 Task description

We will be creating an AI legal assistant with requisite knowledge of the General Data Protection Regulation (GDPR). The objective of this thesis is to write a report on the process of developing this AI. We will compare and use pre-existing and pre-trained LLMs, leveraging publicly available sources to further specialize it. This will be done through fine-tuning and other AI enhancing techniques. The finished model should be able to receive questions and answer in Norwegian. Furthermore, as the law is a complicated subject, it is also important to consider the highly specific language that's used in these contexts. The finalized model will be evaluated as described under '5.4 Plan for testing'.

2.3 Delimitation

- When developing the AI legal assistant the focus will primarily be on the technical aspects of the model, and less on the exploration of user design, user interaction methodologies, or user satisfaction metrics.
- An existing LLM will also be utilized rather than developing a new model form scratch, which constrains modifications to fine-tuning processes.
- The model's performance will be evaluated on a selected set of legal issues or scenarios, rather than attempting a comprehensive assessment across all legal domains.

3 Project organization

3.1 Member Roles

- **Group Leader**
  The group leader is responsible for making sure every group member has something to work on, and will delegate tasks to optimize team performance. He is also responsible for pushing the other members when they are behind on work, as well as motivate them.
  - Assignee(s): Emil Bjerknes
- **Communications Responsible**
  This role functions as a bridge between the group and external parties such as the client and the supervisors. Specific tasks include creating the agenda for formal meetings, communicating said agenda to the appropriate parties, and reflecting the external parties' thoughts back to the group.
  - Assignee(s): Paul Bjørneng
- **Documentation Responsible**
  This members task is to know of, keep and organize all documents used and developed by the group.
  - Assignee(s): Torje Bliksvær
- **Secretary**
  This person is responsible for taking notes during meetings and writing meeting minutes. [1]
  - Assignee(s): Emil Bjerknes
- **Sources Responsible**
  This members task is to know of, keep and organize all sources used by the group.
  - Assignee(s): Torje Bliksvær
- **Quality Assurance**
  The task of quality assurance is to make sure written documents have proper spelling and grammar, and that it is well structured.
  - Assignee(s): Emil Bjerknes, Paul Bjørneng

3.2 Scrum Team roles

- **Product Owner**
  The product owner is the one accountable for what is built during the process, and is supposed to create a vision of the product that they wish to develop and communicate this to the Scrum team.
  - Assignee(s): Futurize AS

- **Scrum Master**
  The Scrum master is the one who ensures that the scrum framework is followed, and serves to facilitate Scrum to a team. The Scrum master is committed to the Scrum methodology, and acts like a coach to the rest of the team.
    - Assignee(s): Torje Bliksvær

- **Developers**
  Developers are the people in the Scrum team that are committed to planning and doing the work that is necessary in each Sprint.
    - Assignee(s): Paul Bjørneng, Emil Bjerknes, Torje Bliksvær

3.3 Group Rules

- Each member should aim to work a total of 30 hours per week.
- Each member is responsible for logging their respective hours in the proper Excel document.
- Unless otherwise agreed upon, each member have to be present in planned group activities whether it be in-person or digitally.
- If a member realises he will be late to an activity he must notify the rest of the group as fast as possible.
- A group conflict will initially be resolved without external intervention.

3.4 Routines

- Weekly meetings with at least one supervisor present will occur each Monday at 11:30-12:00, either on Microsoft Teams or physically on campus.
    - A meeting can be unscheduled if it is deemed unnecessary that week.
- Microsoft Teams meetings with the client are scheduled every other week, on Mondays at 10:30-11:00.
    - A meeting can be unscheduled if it is deemed unnecessary, but we must send notice of this as soon as possible.
    - A meeting can be rescheduled with at least five days notice.
- We will use Discord as a centralized platform for communication within the group. Here we will share resources and otherwise communicate when not together physically.
- We will use Outlook and distinct Microsoft Teams channels for external communication.
- We will use Office 365 for various document drafts (e.g. meeting agendas and minutes).
- The group has regular mandatory work sessions each Monday, Tuesday and Thursday on campus (minimum 12 hours total).

4 Planning, follow-up and reporting

4.1 Project Development Model

Scrum is a widely used agile development method. The workflow is similar to other agile models such as the Kanban method, mostly in that it is highly flexible. What differs between these two aforementioned models is that Scrum has the idea of "Sprints". These Sprints could be viewed as isolated periods of work, in that ideally the workload (which is planned at the start of each Sprint) is the only work to be done during said period. With Kanban there is no such fixed periods. Rather, new tasks are pulled in continuously. With both these models it is common to visualize the workload with either Scrum boards or Kanban boards. This visualization makes it easy to follow what tasks are planned, underway, done or blocked at any point in time.

What made us decide on Scrum was initially because of its relevance in the workplace. As it is such a common approach to project development it would therefore certainly be of benefit to us to practice this model. We discussed different options such as a more pure Kanban approach, or perhaps a Waterfall approach, but these options were deemed not entirely fitting to our project. We felt that we needed smaller intervals of work, with some checkpoints along the way, so Kanban was deemed too unstructured. The Waterfall model would, on the contrary, be too rigid. As the nature of our project calls for iterations and changes of perspective as we proceed, this model was discarded as well.

The Scrum team will be working in Sprints of two weeks starting and ending on Mondays. In this way, after every Sprint, the team is able to receive feedback from the Product Owner in a Sprint Review. The purpose of the Sprint Review is to inspect the outcome of the Sprint and determine future adaptations [4]. The team will present the results of the work to the owner and discuss the progress towards the Project Goal. After the Sprint Review, the team will do Sprint Planning for the next two-week sprint period. This consists of the work that is to be done during the sprint, as well as a few other topics. It will address why the sprint is valuable, what can be done during the sprint, and how the chosen work will be done [4].

Since our group is so small, having a designated person be the Scrum Master alone wouldn't be feasible. Therefore, all members will be Developers. However one person will have the Scrum Master role in addition. This person will be Torje Bliksvær. The Product Owner won't have as central of a role as in typical Scrum uses. Instead he will give feedback and recommendations at each Sprint Planning based on what has been found out during the last Sprint.

(This Scrum setup was inspired by the Project Plan of Urne et al. [2])

4.2 Plan for status meetings and decision points during the period

We will have regular status meetings with both the client and the supervisors.

- Client meeting every other week on Monday.
- Supervisor meetings every Monday.

Follow-up meetings within the group will typically happen every week, directly after the supervisor meetings. This is a logical day to recuperate and report, as it's follows nicely with our Scrum Sprints. If more meetings are required, time will be allotted from pre-planned work sessions. The agenda for these meetings will consist of:

- A review of our timetables.
- Resolution of any important decisions that has come up during the past week.
- Make a rough plan for the week ahead.

5 Organization of quality assurance

5.1 Documentation

During each meeting the assigned secretary will take notes. Afterwards we will go through the notes and elaborate where needed, and fix potential grammatical mistakes. These notes are then written down in the collective meeting minute document on Overleaf.

Documents used will be notified of and given to the documentation responsible member of the group. It is their responsibility to keep the documents in a safe and organized manner. These documents can be shared freely otherwise where the group may see fit, but it is important for the documentation responsible to have these safely stored for future reference.

Each group member will take note of their individual time spent on the project in a designated shared online Excel spreadsheet. This will encompass everything related to the project, be it actively working or learning new material. There it will also be written a brief description of what the time was spent on, and given a general classification.

5.2 Standards

For the duration of the project any changes to the code must have descriptive comments, as well as a brief overview of the changes in the Git commit message. Specifically, any commit message will be written in this format:

type(scope): subject [3].

The *type* describes the purpose of the commit, the *scope* describes the area that the commit affects, and the *subject* contains a short description of the changes [3].

5.3 Tools

Tool	Description	Our use
Huggingface	huggingface.co is a website where people share AI related resources, such as fully trained models and datasets.	We will be using it to discover and research different potential LLM's.
Visual Studio Code	IDE with support for many languages through extensions.	This will be our IDE of choice, since it's very popular and we already have experience using it.
Teams	Teams is a service by Microsoft that allows for digital meetings and provides written communication alongside file sharing.	We will be using Teams as our way of communication with both the client and our supervisors through their own respective Teams. This is because Teams is a very standard platform for communication in a professional setting where it is easy to manage different groups.
Microsoft 365	Cloud solution for storing, sharing and collaboratively working on the Microsoft catalogue of services.	We use Excel to describe our time usage, and Word for 'quick and dirty' notes that come to mind for meetings or future reference. Excel is used for its grid structure which makes it easy to get an overview, alongside its capability to automatically calculate how much time is used. Word is chosen for how easy it is to use.
Overleaf	Online service for storing, sharing and collaboratively working on LaTeX documents.	Overleaf will be used to write any materials that may be included in the final report. With overleaf it is easy to collaboratively make a good looking and cohesive document.
Jira	Jira is an online service for enhancing efficiency in code development with integration possibilities for many other development tools such as GitHub.	We will be using Jira as a centralised place to keep our issue board. Jira was chosen for this as it is not inherently connected to any other service and is used by many big names in the industry.
Discord	Popular social media platform.	Discord will be used for all communication within the group, because it is easy to use. We will not be disclosing any sensitive information here, so potential security issues won't be any issue.

GitLab	Service for streamlining the code development process.	We will be using NTNUs own GitLab to develop our code. This is so we can collaboratively work on the code, whilst also keeping it safe within NTNUs domain.

Table 1: Tools to be used in the project.

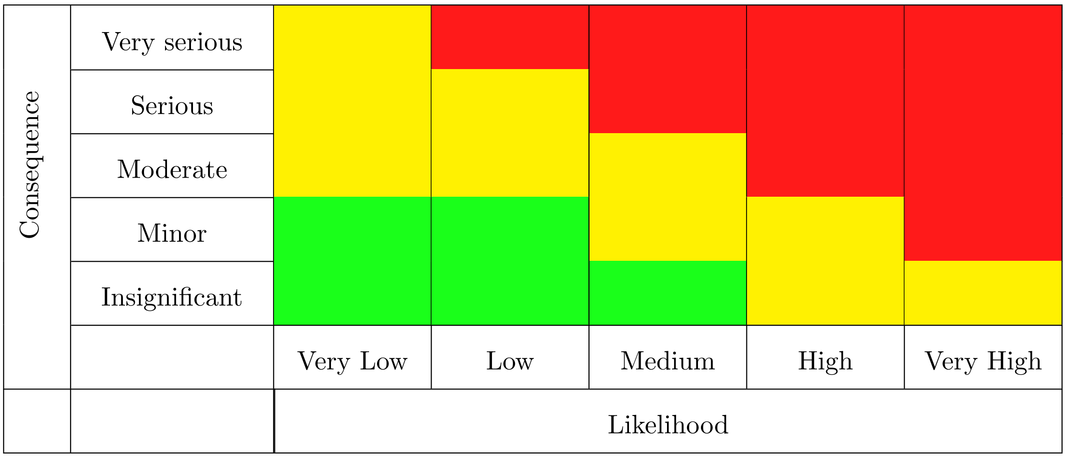
5.4 Plan for testing

For each new iteration of our model, it will be tested by asking pre-defined questions which we already have answers to. We will then evaluate how well the model answers the questions in comparison to earlier iterations to see if it has improved. Each iteration will be tested multiple times to compensate for variation in the answers given by the model. All answers will be taken into consideration when evaluating it. The evaluations will have to be done manually as there's no good way of automating the process. The models will be tested for correctness and accuracy of language. They will be given a subjective general score by us between 0 and 10, for the sake of comparison. 0 being an unrelated answer, and 10 being as though a lawyer had answered.

5.5 Risk analysis

To make work on the project more robust the group has completed a risk analysis covering the risks regarding our thesis. By identifying and evaluating potential risks that could occur throughout the project, these risks can be both foreseen and mitigated. The group has used NTNU's Risk Matrix, rating the potential risk scenarios from insignificant to very serious, with a likelihood between very low and very high. Each risk scenario is given a risk level based on its consequence and likelihood, and is also given a measure that would mitigate the risk.

(*This risk analysis setup is inspired by the Project Plan of Urne et al. [2].*)



Consequence						
	Very serious	🟨	🟥	🟥	🟥	🟥
	Serious	🟨	🟨	🟥	🟥	🟥
	Moderate	🟨	🟨	🟨	🟥	🟥
	Minor	🟩	🟩	🟨	🟨	🟥
	Insignificant	🟩	🟩	🟩	🟨	🟨
		Very Low	Low	Medium	High	Very High
		Likelihood				

Table 2: Risk matrix from NTNU. [6]

Colour	**Description**
Red	Unacceptable risk. Measures must be taken to reduce the risk.
Yellow	Assessment risk. Measures must be considered.
Green	Acceptable risk. Measures can be considered based on other considerations.

Table 3: NTNU Risk matrix colour explanation. [6]

Risk	Description	Likelihood	Consequence	Risk Level
Client is unavailable	The group is unable to have a meeting with the client due to the client being unavailable.	Very Low	Serious	Yellow
Sick group member	One of the group members are sick, and is unable to work for a short period of time.	Low	Moderate	Yellow
Scope does not fit	Scope is either too large and leads to too much work in the given time-frame, or it is too small and we have a lot of time left when we are finished.	Medium	Serious	Red
Loss of data	Lose access to our data, which could include our project on Overleaf, our code or other important documents.	Low	Very Serious	Red
Hard deadlines not met	Delivery deadlines such as final thesis deadline and project plan deadline are not met due to unfinished work.	Very Low	Very Serious	Yellow
Soft deadlines not met	Internal deadlines are not met for individual or group work. These are not crucial to be finished within a certain time-frame, but highly recommended to be on schedule.	Low	Minor	Green
Group disagreement	Internal arguments and disagreement on certain decisions. This would lead to ineffective work.	Low	Moderate	Yellow

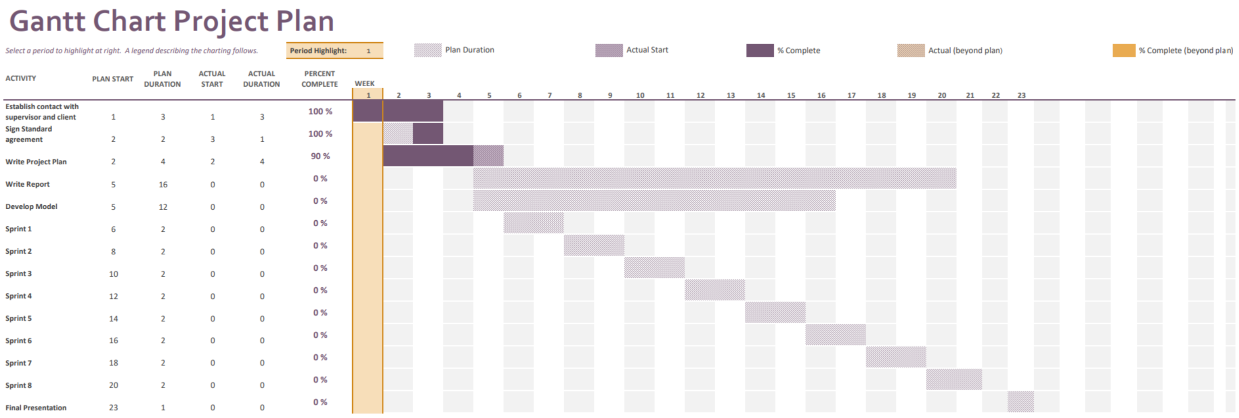
Table 4: Risk scenarios

Risk	**Measure**
Client is unavailable	Set regular meetings in Teams with the client early in the process. In this way the client has enough time to plan and make room in their own schedule. If for some reason we need to reschedule a meeting, this will be done with at least five days notice.
Sick group member	The member who is sick will make up for the missed hours when they are well again. In this way the group does not fall behind on time.
Scope does not fit	Create a scope with our client that suits both parties, and have this approved by our supervisors before starting the work. Weekly updates with supervisors will also make sure that our scope fits. If our scope appears to be unsuitable, this will be dealt with early in the process.
Loss of data	Create a local backup of the report, as well as the code and other documents. Everyone will save an updated copy of the reports and the code locally on their computer after every work session. This will be in addition to having the report on Overleaf, and our code on Git.
Hard deadlines not met	Make sure that we stay on schedule by following our soft deadlines, as well as having status/progress meetings every Monday both with supervisors and internally in the group.
Soft deadlines not met	Make sure we follow our Gantt chart, and use sprints. In a worst case scenario we prioritize the work we have missed before moving on, and we prepare to put in extra hours to finish this as soon as possible.
Group disagreement	Have the group leader decide, unless the others still are really dissatisfied, in which case a supervisor will be contacted.

Table 5: Risk measures

6 Plan for execution

6.1 Gantt chart



Gantt Chart Project Plan

Select a period to highlight at right. A legend describing the charting follows.

Period Highlight: 1

ACTIVITY	PLAN START	PLAN DURATION	ACTUAL START	ACTUAL DURATION	PERCENT COMPLETE
Establish contact with supervisor and client	1	3	1	3	100 %
Sign Standard agreement	2	2	3	1	100 %
Write Project Plan	2	4	2	4	90 %
Write Report	5	16	0	0	0 %
Develop Model	5	12	0	0	0 %
Sprint 1	6	2	0	0	0 %
Sprint 2	8	2	0	0	0 %
Sprint 3	10	2	0	0	0 %
Sprint 4	12	2	0	0	0 %
Sprint 5	14	2	0	0	0 %
Sprint 6	16	2	0	0	0 %
Sprint 7	18	2	0	0	0 %
Sprint 8	20	2	0	0	0 %
Final Presentation	23	1	0	0	0 %

Bibliography

[1] *Aboute Futurize*. Message received on Teams; translated from Norwegian. 2023.

[2] Urne et al. *Project Plan*. Accessed on DCSG2900's Blackboard page: Prosjektplan - Eksempel 2.

[3] *European Commission, Git Commit Guidelines*. Accessed: January 30, 2024. URL: https://ec.europa.eu/component-library/v1.15.0/eu/docs/conventions/git/.

[4] Scrum Guides. *The 2020 Scrum Guide*. Accessed: January 25, 2024. 2020. URL: https://scrumguides.org/scrum-guide.html.

[5] Kim Martineau. *IBM. What is generative AI?* Apr. 2023. URL: https://research.ibm.com/blog/what-is-generative-AI.

[6] *NTNU Risk matrix*. https://i.ntnu.no/documents/1306938287/1306984199/Matrix+risk+assessments+-+eng.pdf/0d037956-7ea5-49db-94c3-512dc2ffdaff?t=1443197424892&status=0. Accessed: 24.01.24. Mar. 2010.

Appendix

A Group Contract

Group members:
Emil Bjerknes
Paul Bjørneng
Torje Bliksvær

A.1 Roles

Group Leader
The group leader is responsible for making sure every group member has something to work on, and will delegate tasks to optimize team performance. He is also responsible for pushing the other members when they are behind on work, as well as motivate them.

- Assignee: Emil Bjerknes

Communications Responsible
This role functions as a bridge between the group and external parties such as the client and the supervisors. Specific tasks include creating the agenda for formal meetings, communicating said agenda to the appropriate parties, and reflecting the external parties' thoughts back to the group.

- Assignee: Paul Bjørneng

Documentation Responsible
This members task is to know of, keep and organize all documents used and developed by the group.

- Assignee: Torje Bliksvær

Secretary
This person is responsible for taking notes during meetings and writing meeting minutes. [1]

- Assignee: Emil Bjerknes

Sources Responsible
This members task is to know of, keep and organize all sources used by the group.

- Assignee: Torje Bliksvær

Quality Assurance
The task of quality assurance is to make sure written documents have proper spelling and grammar, and that it is well structured.

- Assignee: Emil Bjerknes, Paul Bjørneng

A.2 Group rules

- Each member is required to work a total of 30 hours per week.
  - Failing to meet the expected number at the weekly checkup will result in a warning. You are then required to catch up by the end of next week.
- Each member is responsible for logging their respective hours in the proper Excel document.
- If a member is not able to participate in a planned group activity, they are required to notify the group at least 24 hours beforehand. The exception is if they fall sick on the day. It is then acceptable to notify the group 2 hours before the activity starts.
  - If a member is absent from a meeting or work session, and does not notify the group, they will be given two warnings.
- If a member finds out they will be late to an assigned activity, be it a work session or any meeting, they are to inform the rest of the group as soon as possible.
  - Failing to do so and coming more than 5 minutes late, they will be given a warning.
- If a disagreement occurs within the group, and internal problem resolution has failed, the supervisor(s) will be contacted for advice.

A.3 Consequences for violating the rules

A violation of the group rules will result in a number of warnings, dependant on the severity of the violation. For instance while a late arrival of 5 minutes will result in only one warning, a member being absent without notice will result in two. The different scenarios are specified in section A.2.

If a member acquires a total of three warnings, they are required to buy the other members an energy drink of their choosing. These will be distributed between the appropriate parties at the beginning of the next group activity.

Emil F. Bjerknes 30.01.2024

Emil Farstad Bjerknes Date

30.01.2024

Paul Røkke Bjørneng Date

Torje G. Bliksvær 30.01.2024

Torje Grimstad Bliksvær Date

Appendix B

Standard Agreement

NTNU
Norges teknisk-naturvitenskapelige universitet

Fastsatt av prorektor for utdanning 10.12.2020

STANDARDAVTALE

om utføring av studentoppgave i samarbeid med ekstern virksomhet

Avtalen er ufravikelig for studentoppgaver (heretter oppgave) ved NTNU som utføres i samarbeid med ekstern virksomhet.

Forklaring av begrep

Opphavsrett

Er den rett som den som skaper et åndsverk har til å fremstille eksemplar av åndsverket og gjøre det tilgjengelig for allmennheten. Et åndsverk kan være et litterært, vitenskapelig eller kunstnerisk verk. En studentoppgave vil være et åndsverk.

Eiendomsrett til resultater

Betyr at den som eier resultatene bestemmer over disse. Utgangspunktet er at studenten eier resultatene fra sitt studentarbeid. Studenten kan også overføre eiendomsretten til den eksterne virksomheten.

Bruksrett til resultater

Den som eier resultatene kan gi andre en rett til å bruke resultatene, f.eks. at studenten gir NTNU og den eksterne virksomheten rett til å bruke resultatene fra studentoppgaven i deres virksomhet.

Prosjektbakgrunn

Det partene i avtalen har med seg inn i prosjektet, dvs. som vedkommende eier eller har rettigheter til fra før og som brukes i det videre arbeidet med studentoppgaven. Dette kan også være materiale som tredjepersoner (som ikke er part i avtalen) har rettigheter til.

Utsatt offentliggjøring

Betyr at oppgaven ikke blir tilgjengelig for allmennheten før etter en viss tid, f.eks. før etter tre år. Da vil det kun være veileder ved NTNU, sensorene og den eksterne virksomheten som har tilgang til studentarbeidet de tre første årene etter at studentarbeidet er innlevert.

1. Avtaleparter

Norges teknisk-naturvitenskapelige universitet (NTNU) Institutt: IIK (Institutt for informasjonssikkerhet og kommunikasjonsteknologi)
Veileder ved NTNU: Shao-Fang Wen. e-post og tlf: shao-fang.wen@ntnu.no, 406 77 280. Veileder ved NTNU: Muhammad Mudassar Yamin. e-post og tlf: muhammad.m.yamin@ntnu.no, 969 99 968.
Ekstern virksomhet: Futurize AS Ekstern virksomhet sin kontaktperson, e-post og tlf.: Daniel Arlien, daniel.arlien@gmail.com , 90089188.
Student: Emil Farstad Bjerknes. Fødselsdato: 06.11.2000.
Ev. flere studenter[1] Student: Paul Røkke Bjørneng. Fødselsdato: 13.07.2000.
Student: Torje Grimstad Bliksvær. Fødselsdato: 05.06.2002

Partene har ansvar for å klarere eventuelle immaterielle rettigheter som studenten, NTNU, den eksterne eller tredjeperson (som ikke er part i avtalen) har til prosjektbakgrunn før bruk i forbindelse med utførelse av oppgaven. Eierskap til prosjektbakgrunn skal fremgå av eget vedlegg til avtalen der dette kan ha betydning for utførelse av oppgaven.

2. Utførelse av oppgave

Studenten skal utføre: (sett kryss)

Masteroppgave	
Bacheloroppgave	x
Prosjektoppgave	
Annen oppgave	

Startdato: 08.01.24
Sluttdato: 21.05.24

Oppgavens arbeidstittel er: **AI-drevet juridisk assistent: Språkmodell med forståelse av det norske rettssystemet**.

---

[1] Dersom flere studenter skriver oppgave i fellesskap, kan alle føres opp her. Rettigheter ligger da i fellesskap mellom studentene. Dersom ekstern virksomhet i stedet ønsker at det skal inngås egen avtale med hver enkelt student, gjøres dette.

Ansvarlig veileder ved NTNU har det overordnede faglige ansvaret for utforming og godkjenning av prosjektbeskrivelse og studentens læring.

3. Ekstern virksomhet sine plikter

Ekstern virksomhet skal stille med en kontaktperson som har nødvendig faglig kompetanse til å gi studenten tilstrekkelig veiledning i samarbeid med veileder ved NTNU. Ekstern kontaktperson fremgår i punkt 1.

Formålet med oppgaven er studentarbeid. Oppgaven utføres som ledd i studiet. Studenten skal ikke motta lønn eller lignende godtgjørelse fra den eksterne for studentarbeidet. Utgifter knyttet til gjennomføring av oppgaven skal dekkes av den eksterne. Aktuelle utgifter kan for eksempel være reiser, materialer for bygging av prototyp, innkjøp av prøver, tester på lab, kjemikalier. Studenten skal klarere dekning av utgifter med ekstern virksomhet på forhånd.

Ekstern virksomhet skal dekke følgende utgifter til utførelse av oppgaven:

Dekning av utgifter til annet enn det som er oppført her avgjøres av den eksterne underveis i arbeidet.

4. Studentens rettigheter

Studenten har opphavsrett til oppgaven[2]. Alle resultater av oppgaven, skapt av studenten alene gjennom arbeidet med oppgaven, eies av studenten med de begrensninger som følger av punkt 5, 6 og 7 nedenfor. Eiendomsretten til resultatene overføres til ekstern virksomhet hvis punkt 5 b er avkrysset eller for tilfelle som i punkt 6 (overføring ved patenterbare oppfinnelser).

I henhold til lov om opphavsrett til åndsverk beholder alltid studenten de ideelle rettigheter til eget åndsverk, dvs. retten til navngivelse og vern mot krenkende bruk.

Studenten har rett til å inngå egen avtale med NTNU om publisering av sin oppgave i NTNUs institusjonelle arkiv på Internett (NTNU Open). Studenten har også rett til å publisere oppgaven eller deler av den i andre sammenhenger dersom det ikke i denne avtalen er avtalt begrensninger i adgangen til å publisere, jf. punkt 8.

5. Den eksterne virksomheten sine rettigheter

Der oppgaven bygger på, eller videreutvikler materiale og/eller metoder (prosjektbakgrunn) som eies av den eksterne, eies prosjektbakgrunnen fortsatt av den eksterne. Hvis studenten

---

[2] Jf. Lov om opphavsrett til åndsverk mv. av 15.06.2018 § 1

skal utnytte resultater som inkluderer den eksterne sin prosjektbakgrunn, forutsetter dette at det er inngått egen avtale om dette mellom studenten og den eksterne virksomheten.

Alternativ a) (sett kryss) Hovedregel

	Ekstern virksomhet skal ha bruksrett til resultatene av oppgaven

Dette innebærer at ekstern virksomhet skal ha rett til å benytte resultatene av oppgaven i egen virksomhet. Retten er ikke-eksklusiv.

Alternativ b) (sett kryss) Unntak

x	Ekstern virksomhet skal ha eiendomsretten til resultatene av oppgaven og studentens bidrag i ekstern virksomhet sitt prosjekt

Begrunnelse for at ekstern virksomhet har behov for å få overført eiendomsrett til resultatene: **For å beskytte mot konkurranse og forhindre at andre kopierer våre innsikter på kort sikt.**

6. Godtgjøring ved patenterbare oppfinnelser

Dersom studenten i forbindelse med utførelsen av oppgaven har nådd frem til en patenterbar oppfinnelse, enten alene eller sammen med andre, kan den eksterne kreve retten til oppfinnelsen overført til seg. Dette forutsetter at utnyttelsen av oppfinnelsen faller inn under den eksterne sitt virksomhetsområde. I så fall har studenten krav på rimelig godtgjøring. Godtgjøringen skal fastsettes i samsvar med arbeidstakeroppfinnelsesloven § 7. Fristbestemmelsene i § 7 gis tilsvarende anvendelse.

7. NTNU sine rettigheter

De innleverte filer av oppgaven med vedlegg, som er nødvendig for sensur og arkivering ved NTNU, tilhører NTNU. NTNU får en vederlagsfri bruksrett til resultatene av oppgaven, inkludert vedlegg til denne, og kan benytte dette til undervisnings- og forskningsformål med de eventuelle begrensninger som fremgår i punkt 8.

8. Utsatt offentliggjøring

Hovedregelen er at studentoppgaver skal være offentlige.

Sett kryss

x	Oppgaven skal være offentlig

I særlige tilfeller kan partene bli enige om at hele eller deler av oppgaven skal være undergitt utsatt offentliggjøring i maksimalt tre år. Hvis oppgaven unntas fra offentliggjøring, vil den kun være tilgjengelig for student, ekstern virksomhet og veileder i denne perioden. Sensurkomiteen vil ha tilgang til oppgaven i forbindelse med sensur. Student, veileder og sensorer har taushetsplikt om innhold som er unntatt offentliggjøring.

Oppgaven skal være underlagt utsatt offentliggjøring i (sett kryss hvis dette er aktuelt):

Sett kryss		Sett dato
	ett år	
x	to år	**21.05.26**
	tre år	

Behovet for utsatt offentliggjøring er begrunnet ut fra følgende:

For at ekstern virksomhet skal kunne beskytte seg mot konkurranse og forhindre at andre kopierer våre innsikter på kort sikt.

Dersom partene, etter at oppgaven er ferdig, blir enig om at det ikke er behov for utsatt offentliggjøring, kan dette endres. I så fall skal dette avtales skriftlig.

Vedlegg til oppgaven kan unntas ut over tre år etter forespørsel fra ekstern virksomhet. NTNU (ved instituttet) og student skal godta dette hvis den eksterne har saklig grunn for å be om at et eller flere vedlegg unntas. Ekstern virksomhet må sende forespørsel før oppgaven leveres.

De delene av oppgaven som ikke er undergitt utsatt offentliggjøring, kan publiseres i NTNUs institusjonelle arkiv, jf. punkt 4, siste avsnitt. Selv om oppgaven er undergitt utsatt offentliggjøring, skal ekstern virksomhet legge til rette for at studenten kan benytte hele eller deler av oppgaven i forbindelse med jobbsøknader samt videreføring i et master- eller doktorgradsarbeid.

9. Generelt

Denne avtalen skal ha gyldighet foran andre avtaler som er eller blir opprettet mellom to av partene som er nevnt ovenfor. Dersom student og ekstern virksomhet skal inngå avtale om konfidensialitet om det som studenten får kjennskap til i eller gjennom den eksterne virksomheten, kan NTNUs standardmal for konfidensialitetsavtale benyttes.

Den eksterne sin egen konfidensialitetsavtale, eventuell konfidensialitetsavtale den eksterne har inngått i samarbeidprosjekter, kan også brukes forutsatt at den ikke inneholder punkter i motstrid med denne avtalen (om rettigheter, offentliggjøring mm). Dersom det likevel viser seg at det er motstrid, skal NTNUs standardavtale om utføring av studentoppgave gå foran. Eventuell avtale om konfidensialitet skal vedlegges denne avtalen.

Eventuell uenighet som følge av denne avtalen skal søkes løst ved forhandlinger. Hvis dette ikke fører frem, er partene enige om at tvisten avgjøres ved voldgift i henhold til norsk lov. Tvisten avgjøres av sorenskriveren ved Sør-Trøndelag tingrett eller den han/hun oppnevner.

Denne avtale er signert i fire eksemplarer hvor partene skal ha hvert sitt eksemplar. Avtalen er gyldig når den er underskrevet av NTNU v/instituttleder.

Signaturer:

Instituttleder:
Dato:

Veileder ved NTNU:
Dato: 31.01.2024

Veileder ved NTNU:
Dato: 31.01.24

Ekstern virksomhet: Daniel Arlien
Dato: 30/01/2024

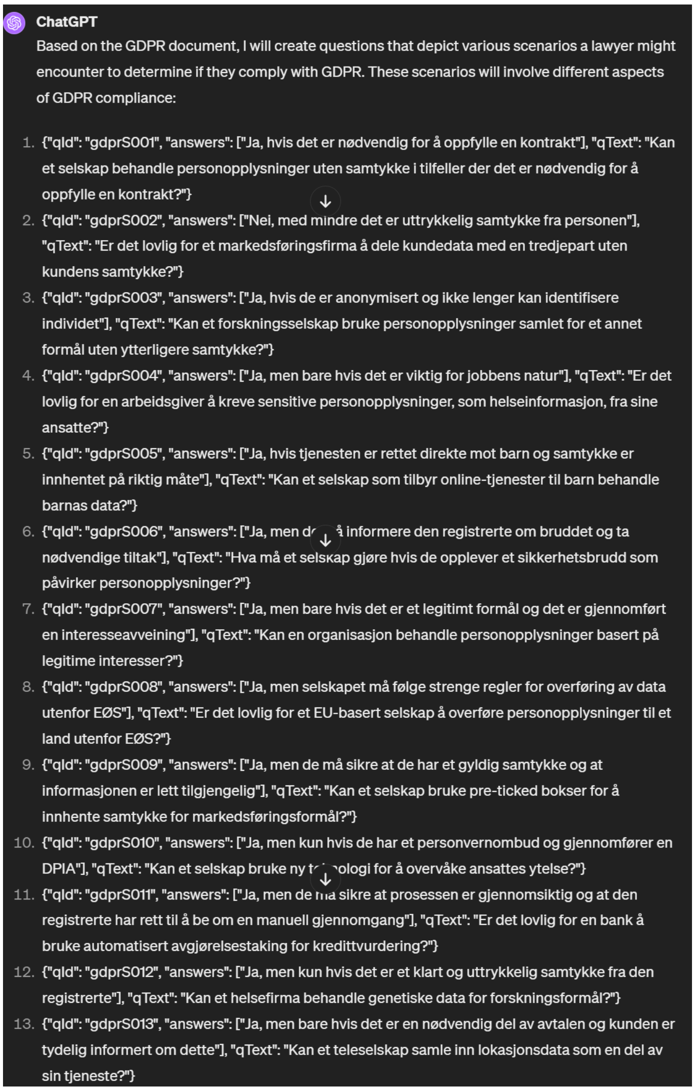
Student: Emil F. Bjerknes
Dato: 01.02.2024

Student:
Dato: 01.02.2024

Student: Tonje G. Bliksvær
Dato: 01.02.2024

Appendix C

ChatGPT Answers

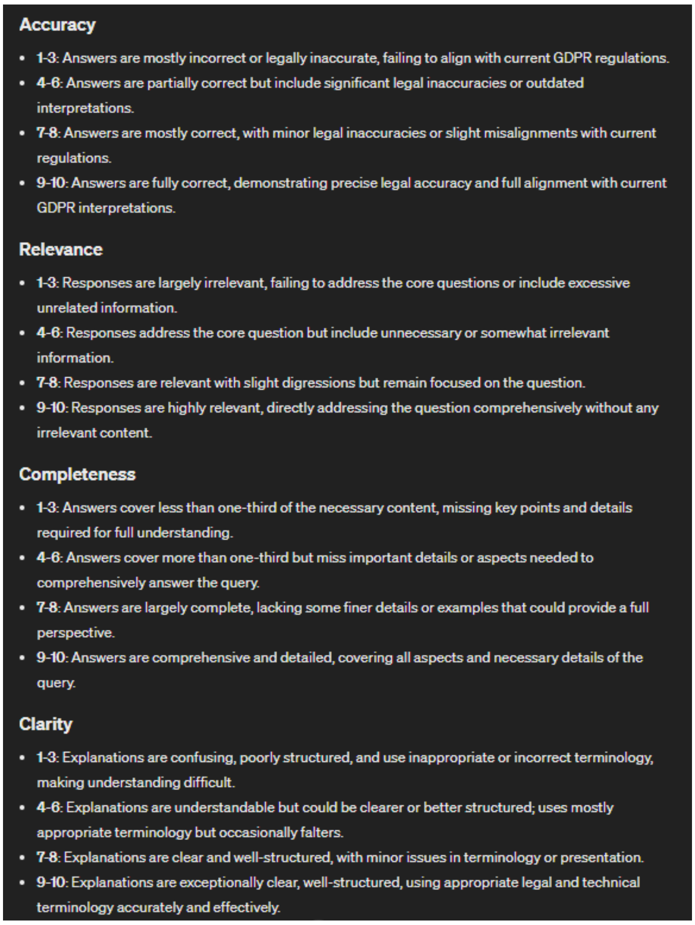


ChatGPT

Based on the GDPR document, I will create questions that depict various scenarios a lawyer might encounter to determine if they comply with GDPR. These scenarios will involve different aspects of GDPR compliance:

1. {"qId": "gdprS001", "answers": ["Ja, hvis det er nødvendig for å oppfylle en kontrakt"], "qText": "Kan et selskap behandle personopplysninger uten samtykke i tilfeller der det er nødvendig for å oppfylle en kontrakt?"}
2. {"qId": "gdprS002", "answers": ["Nei, med mindre det er uttrykkelig samtykke fra personen"], "qText": "Er det lovlig for et markedsføringsfirma å dele kundedata med en tredjepart uten kundens samtykke?"}
3. {"qId": "gdprS003", "answers": ["Ja, hvis de er anonymisert og ikke lenger kan identifisere individet"], "qText": "Kan et forskningsselskap bruke personopplysninger samlet for et annet formål uten ytterligere samtykke?"}
4. {"qId": "gdprS004", "answers": ["Ja, men bare hvis det er viktig for jobbens natur"], "qText": "Er det lovlig for en arbeidsgiver å kreve sensitive personopplysninger, som helseinformasjon, fra sine ansatte?"}
5. {"qId": "gdprS005", "answers": ["Ja, hvis tjenesten er rettet direkte mot barn og samtykke er innhentet på riktig måte"], "qText": "Kan et selskap som tilbyr online-tjenester til barn behandle barnas data?"}
6. {"qId": "gdprS006", "answers": ["Ja, men d[illegible]å informere den registrerte om bruddet og ta nødvendige tiltak"], "qText": "Hva må et selskap gjøre hvis de opplever et sikkerhetsbrudd som påvirker personopplysninger?"}
7. {"qId": "gdprS007", "answers": ["Ja, men bare hvis det er et legitimt formål og det er gjennomført en interesseavveining"], "qText": "Kan en organisasjon behandle personopplysninger basert på legitime interesser?"}
8. {"qId": "gdprS008", "answers": ["Ja, men selskapet må følge strenge regler for overføring av data utenfor EØS"], "qText": "Er det lovlig for et EU-basert selskap å overføre personopplysninger til et land utenfor EØS?"}
9. {"qId": "gdprS009", "answers": ["Ja, men de må sikre at de har et gyldig samtykke og at informasjonen er lett tilgjengelig"], "qText": "Kan et selskap bruke pre-ticked bokser for å innhente samtykke for markedsføringsformål?"}
10. {"qId": "gdprS010", "answers": ["Ja, men kun hvis de har et personvernombud og gjennomfører en DPIA"], "qText": "Kan et selskap bruke ny t[illegible]ologi for å overvåke ansattes ytelse?"}
11. {"qId": "gdprS011", "answers": ["Ja, men de må sikre at prosessen er gjennomsiktig og at den registrerte har rett til å be om en manuell gjennomgang"], "qText": "Er det lovlig for en bank å bruke automatisert avgjørelsestaking for kredittvurdering?"}
12. {"qId": "gdprS012", "answers": ["Ja, men kun hvis det er et klart og uttrykkelig samtykke fra den registrerte"], "qText": "Kan et helsefirma behandle genetiske data for forskningsformål?"}
13. {"qId": "gdprS013", "answers": ["Ja, men bare hvis det er en nødvendig del av avtalen og kunden er tydelig informert om dette"], "qText": "Kan et teleselskap samle inn lokasjonsdata som en del av sin tjeneste?"}

Figure C.1: Initial GPT provided QA-dataset (in Norwegian)



Accuracy

- **1-3**: Answers are mostly incorrect or legally inaccurate, failing to align with current GDPR regulations.
- **4-6**: Answers are partially correct but include significant legal inaccuracies or outdated interpretations.
- **7-8**: Answers are mostly correct, with minor legal inaccuracies or slight misalignments with current regulations.
- **9-10**: Answers are fully correct, demonstrating precise legal accuracy and full alignment with current GDPR interpretations.

Relevance

- **1-3**: Responses are largely irrelevant, failing to address the core questions or include excessive unrelated information.
- **4-6**: Responses address the core question but include unnecessary or somewhat irrelevant information.
- **7-8**: Responses are relevant with slight digressions but remain focused on the question.
- **9-10**: Responses are highly relevant, directly addressing the question comprehensively without any irrelevant content.

Completeness

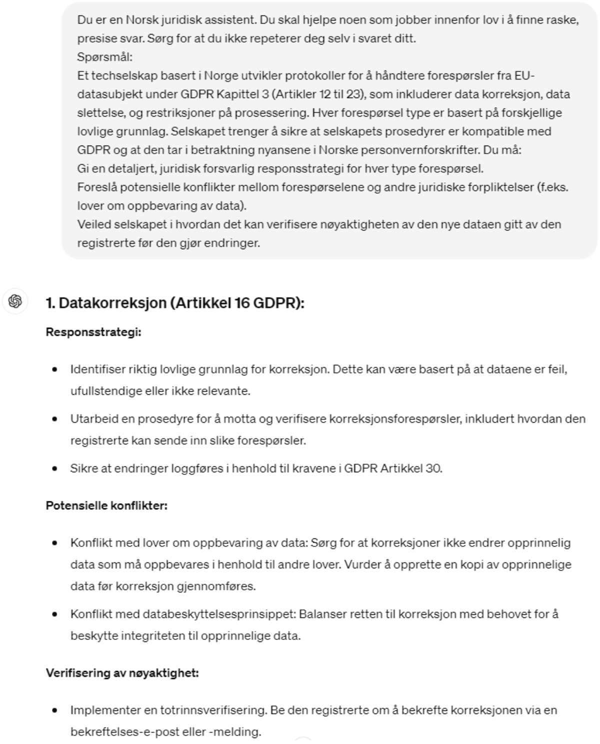
- **1-3**: Answers cover less than one-third of the necessary content, missing key points and details required for full understanding.
- **4-6**: Answers cover more than one-third but miss important details or aspects needed to comprehensively answer the query.
- **7-8**: Answers are largely complete, lacking some finer details or examples that could provide a full perspective.
- **9-10**: Answers are comprehensive and detailed, covering all aspects and necessary details of the query.

Clarity

- **1-3**: Explanations are confusing, poorly structured, and use inappropriate or incorrect terminology, making understanding difficult.
- **4-6**: Explanations are understandable but could be clearer or better structured; uses mostly appropriate terminology but occasionally falters.
- **7-8**: Explanations are clear and well-structured, with minor issues in terminology or presentation.
- **9-10**: Explanations are exceptionally clear, well-structured, using appropriate legal and technical terminology accurately and effectively.

Figure C.2: Scoring system example for the program evaluation

C.1 ChatGPT Answers Used In Evaluation



Du er en Norsk juridisk assistent. Du skal hjelpe noen som jobber innenfor lov i å finne raske, presise svar. Sørg for at du ikke repeterer deg selv i svaret ditt.
Spørsmål:
Et techselskap basert i Norge utvikler protokoller for å håndtere forespørsler fra EU-datasubjekt under GDPR Kapittel 3 (Artikler 12 til 23), som inkluderer data korreksjon, data slettelse, og restriksjoner på prosessering. Hver forespørsel type er basert på forskjellige lovlige grunnlag. Selskapet trenger å sikre at selskapets prosedyrer er kompatible med GDPR og at den tar i betraktning nyansene i Norske personvernforskrifter. Du må:
Gi en detaljert, juridisk forsvarlig responsstrategi for hver type forespørsel.
Foreslå potensielle konflikter mellom forespørselene og andre juridiske forpliktelser (f.eks. lover om oppbevaring av data).
Veiled selskapet i hvordan det kan verifisere nøyaktigheten av den nye dataen gitt av den registrerte før den gjør endringer.

1. Datakorreksjon (Artikkel 16 GDPR):

Responsstrategi:

- Identifiser riktig lovlige grunnlag for korreksjon. Dette kan være basert på at dataene er feil, ufullstendige eller ikke relevante.
- Utarbeid en prosedyre for å motta og verifisere korreksjonsforespørsler, inkludert hvordan den registrerte kan sende inn slike forespørsler.
- Sikre at endringer loggføres i henhold til kravene i GDPR Artikkel 30.

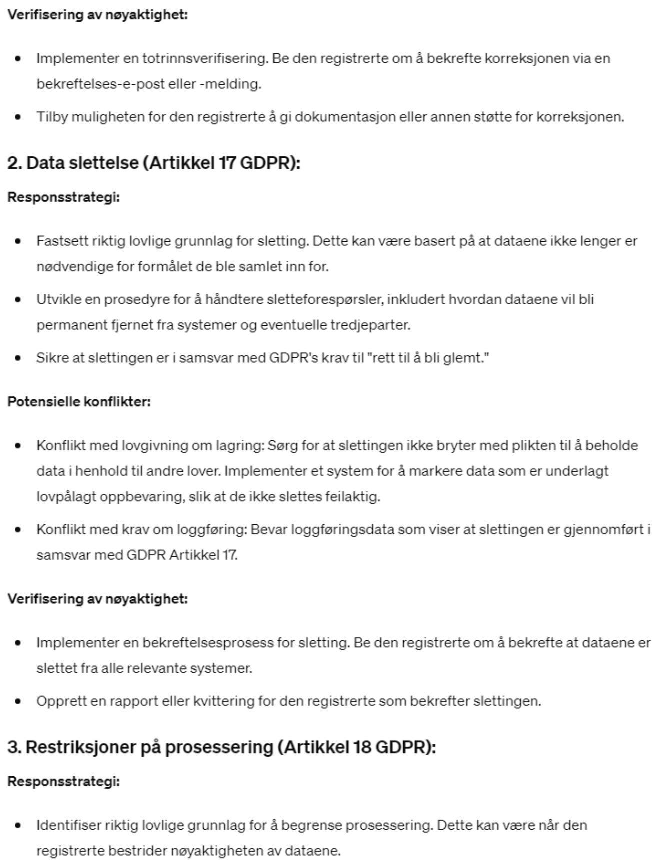
Potensielle konflikter:

- Konflikt med lover om oppbevaring av data: Sørg for at korreksjoner ikke endrer opprinnelig data som må oppbevares i henhold til andre lover. Vurder å opprette en kopi av opprinnelige data før korreksjon gjennomføres.
- Konflikt med databeskyttelsesprinsippet: Balanser retten til korreksjon med behovet for å beskytte integriteten til opprinnelige data.

Verifisering av nøyaktighet:

- Implementer en totrinnsverifisering. Be den registrerte om å bekrefte korreksjonen via en bekreftelses-e-post eller -melding.

Figure C.3: ChatGPT Expert evaluation scenario 1 answer pt. 1



Verifisering av nøyaktighet:

- Implementer en totrinnsverifisering. Be den registrerte om å bekrefte korreksjonen via en bekreftelses-e-post eller -melding.
- Tilby muligheten for den registrerte å gi dokumentasjon eller annen støtte for korreksjonen.

2. Data slettelse (Artikkel 17 GDPR):

Responsstrategi:

- Fastsett riktig lovlige grunnlag for sletting. Dette kan være basert på at dataene ikke lenger er nødvendige for formålet de ble samlet inn for.
- Utvikle en prosedyre for å håndtere sletteforespørsler, inkludert hvordan dataene vil bli permanent fjernet fra systemer og eventuelle tredjeparter.
- Sikre at slettingen er i samsvar med GDPR's krav til "rett til å bli glemt."

Potensielle konflikter:

- Konflikt med lovgivning om lagring: Sørg for at slettingen ikke bryter med plikten til å beholde data i henhold til andre lover. Implementer et system for å markere data som er underlagt lovpålagt oppbevaring, slik at de ikke slettes feilaktig.
- Konflikt med krav om loggføring: Bevar loggføringsdata som viser at slettingen er gjennomført i samsvar med GDPR Artikkel 17.

Verifisering av nøyaktighet:

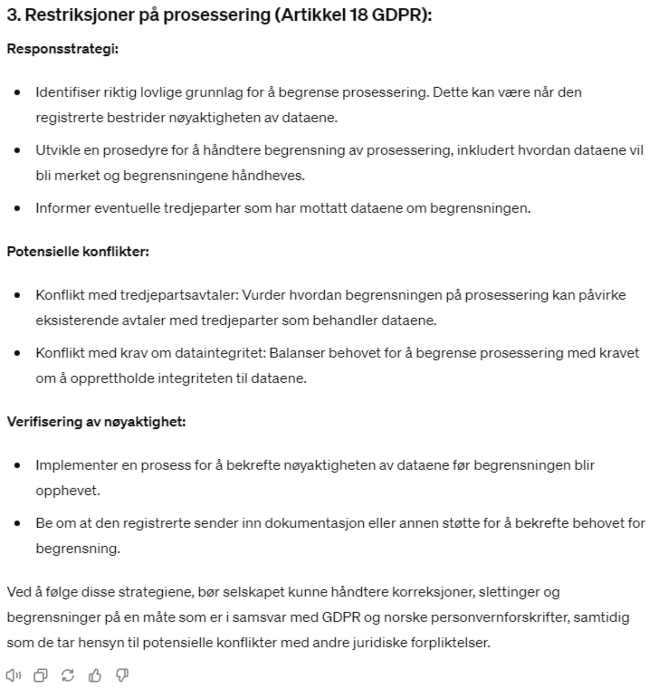
- Implementer en bekreftelsesprosess for sletting. Be den registrerte om å bekrefte at dataene er slettet fra alle relevante systemer.
- Opprett en rapport eller kvittering for den registrerte som bekrefter slettingen.

3. Restriksjoner på prosessering (Artikkel 18 GDPR):

Responsstrategi:

- Identifiser riktig lovlige grunnlag for å begrense prosessering. Dette kan være når den registrerte bestrider nøyaktigheten av dataene.

Figure C.4: ChatGPT Expert evaluation scenario 1 answer pt. 2



3. Restriksjoner på prosessering (Artikkel 18 GDPR):

Responsstrategi:

- Identifiser riktig lovlige grunnlag for å begrense prosessering. Dette kan være når den registrerte bestrider nøyaktigheten av dataene.
- Utvikle en prosedyre for å håndtere begrensning av prosessering, inkludert hvordan dataene vil bli merket og begrensningene håndheves.
- Informer eventuelle tredjeparter som har mottatt dataene om begrensningen.

Potensielle konflikter:

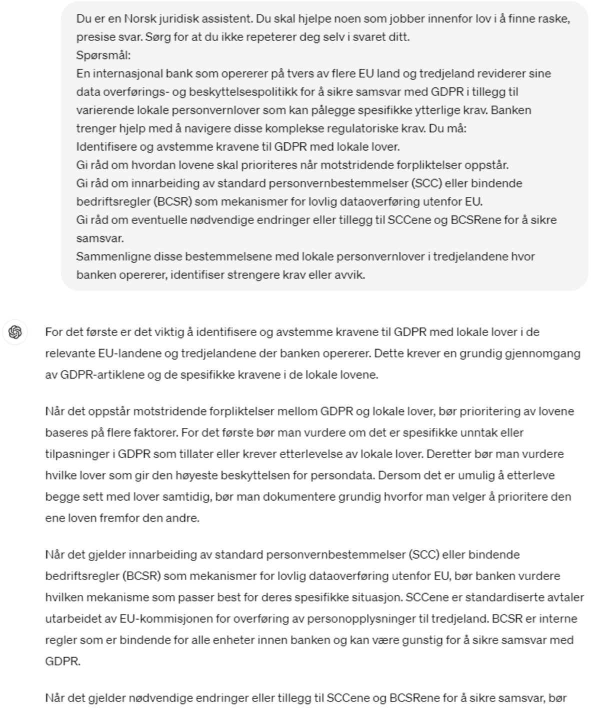
- Konflikt med tredjepartsavtaler: Vurder hvordan begrensningen på prosessering kan påvirke eksisterende avtaler med tredjeparter som behandler dataene.
- Konflikt med krav om dataintegritet: Balanser behovet for å begrense prosessering med kravet om å opprettholde integriteten til dataene.

Verifisering av nøyaktighet:

- Implementer en prosess for å bekrefte nøyaktigheten av dataene før begrensningen blir opphevet.
- Be om at den registrerte sender inn dokumentasjon eller annen støtte for å bekrefte behovet for begrensning.

Ved å følge disse strategiene, bør selskapet kunne håndtere korreksjoner, slettinger og begrensninger på en måte som er i samsvar med GDPR og norske personvernforskrifter, samtidig som de tar hensyn til potensielle konflikter med andre juridiske forpliktelser.

Figure C.5: ChatGPT Expert evaluation scenario 1 answer pt. 3



Du er en Norsk juridisk assistent. Du skal hjelpe noen som jobber innenfor lov i å finne raske, presise svar. Sørg for at du ikke repeterer deg selv i svaret ditt.
Spørsmål:
En internasjonal bank som opererer på tvers av flere EU land og tredjeland reviderer sine data overførings- og beskyttelsespolitikk for å sikre samsvar med GDPR i tillegg til varierende lokale personvernlover som kan pålegge spesifikke ytterlige krav. Banken trenger hjelp med å navigere disse komplekse regulatoriske krav. Du må:
Identifisere og avstemme kravene til GDPR med lokale lover.
Gi råd om hvordan lovene skal prioriteres når motstridende forpliktelser oppstår.
Gi råd om innarbeiding av standard personvernbestemmelser (SCC) eller bindende bedriftsregler (BCSR) som mekanismer for lovlig dataoverføring utenfor EU.
Gi råd om eventuelle nødvendige endringer eller tillegg til SCCene og BCSRene for å sikre samsvar.
Sammenligne disse bestemmelsene med lokale personvernlover i tredjelandene hvor banken opererer, identifiser strengere krav eller avvik.

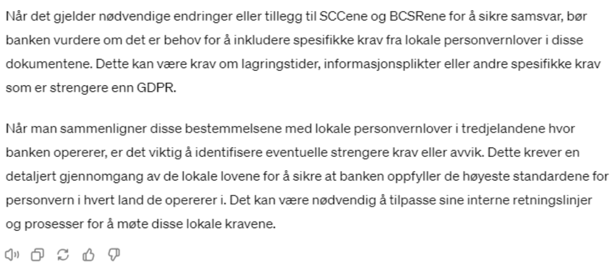
For det første er det viktig å identifisere og avstemme kravene til GDPR med lokale lover i de relevante EU-landene og tredjelandene der banken opererer. Dette krever en grundig gjennomgang av GDPR-artiklene og de spesifikke kravene i de lokale lovene.

Når det oppstår motstridende forpliktelser mellom GDPR og lokale lover, bør prioritering av lovene baseres på flere faktorer. For det første bør man vurdere om det er spesifikke unntak eller tilpasninger i GDPR som tillater eller krever etterlevelse av lokale lover. Deretter bør man vurdere hvilke lover som gir den høyeste beskyttelsen for persondata. Dersom det er umulig å etterleve begge sett med lover samtidig, bør man dokumentere grundig hvorfor man velger å prioritere den ene loven fremfor den andre.

Når det gjelder innarbeiding av standard personvernbestemmelser (SCC) eller bindende bedriftsregler (BCSR) som mekanismer for lovlig dataoverføring utenfor EU, bør banken vurdere hvilken mekanisme som passer best for deres spesifikke situasjon. SCCene er standardiserte avtaler utarbeidet av EU-kommisjonen for overføring av personopplysninger til tredjeland. BCSR er interne regler som er bindende for alle enheter innen banken og kan være gunstig for å sikre samsvar med GDPR.

Når det gjelder nødvendige endringer eller tillegg til SCCene og BCSRene for å sikre samsvar, bør

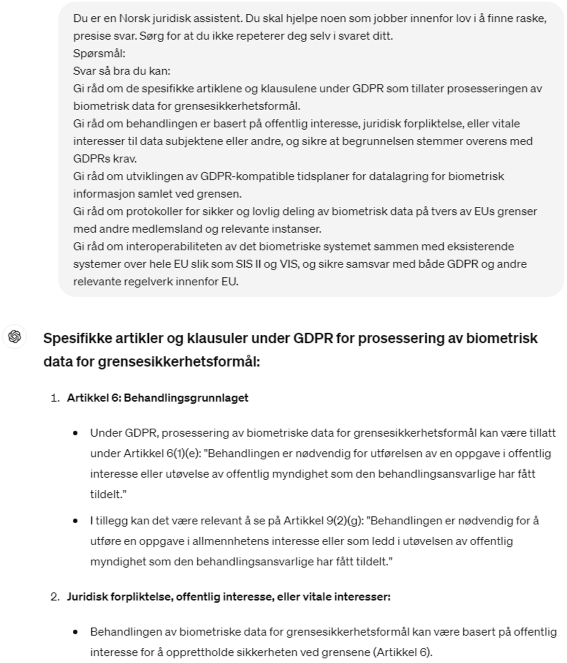
Figure C.6: ChatGPT Expert evaluation scenario 2 answer pt. 1



Når det gjelder nødvendige endringer eller tillegg til SCCene og BCSRene for å sikre samsvar, bør banken vurdere om det er behov for å inkludere spesifikke krav fra lokale personvernlover i disse dokumentene. Dette kan være krav om lagringstider, informasjonsplikter eller andre spesifikke krav som er strengere enn GDPR.

Når man sammenligner disse bestemmelsene med lokale personvernlover i tredjelandene hvor banken opererer, er det viktig å identifisere eventuelle strengere krav eller avvik. Dette krever en detaljert gjennomgang av de lokale lovene for å sikre at banken oppfyller de høyeste standardene for personvern i hvert land de opererer i. Det kan være nødvendig å tilpasse sine interne retningslinjer og prosesser for å møte disse lokale kravene.

Figure C.7: ChatGPT Expert evaluation scenario 2 answer pt. 2

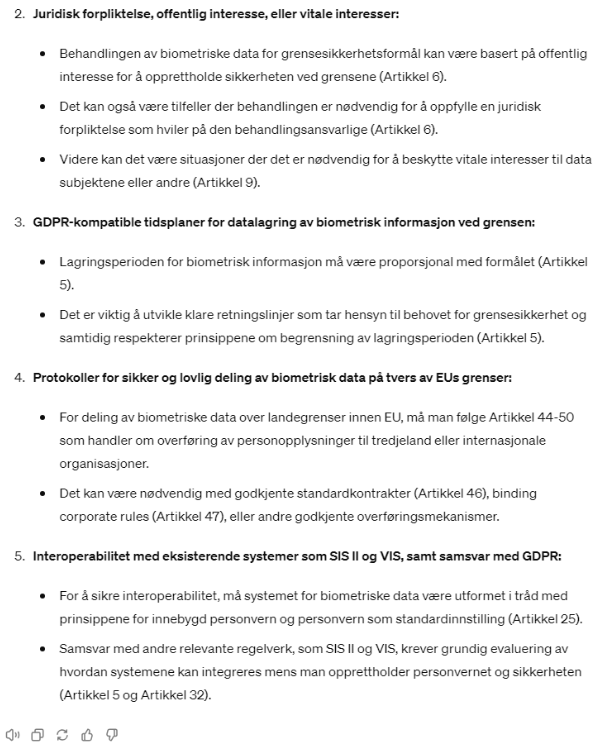


Du er en Norsk juridisk assistent. Du skal hjelpe noen som jobber innenfor lov i å finne raske, presise svar. Sørg for at du ikke repeterer deg selv i svaret ditt.
Spørsmål:
Svar så bra du kan:
Gi råd om de spesifikke artiklene og klausulene under GDPR som tillater prosesseringen av biometrisk data for grensesikkerhetsformål.
Gi råd om behandlingen er basert på offentlig interesse, juridisk forpliktelse, eller vitale interesser til data subjektene eller andre, og sikre at begrunnelsen stemmer overens med GDPRs krav.
Gi råd om utviklingen av GDPR-kompatible tidsplaner for datalagring for biometrisk informasjon samlet ved grensen.
Gi råd om protokoller for sikker og lovlig deling av biometrisk data på tvers av EUs grenser med andre medlemsland og relevante instanser.
Gi råd om interoperabiliteten av det biometriske systemet sammen med eksisterende systemer over hele EU slik som SIS II og VIS, og sikre samsvar med både GDPR og andre relevante regelverk innenfor EU.

Spesifikke artikler og klausuler under GDPR for prosessering av biometrisk data for grensesikkerhetsformål:

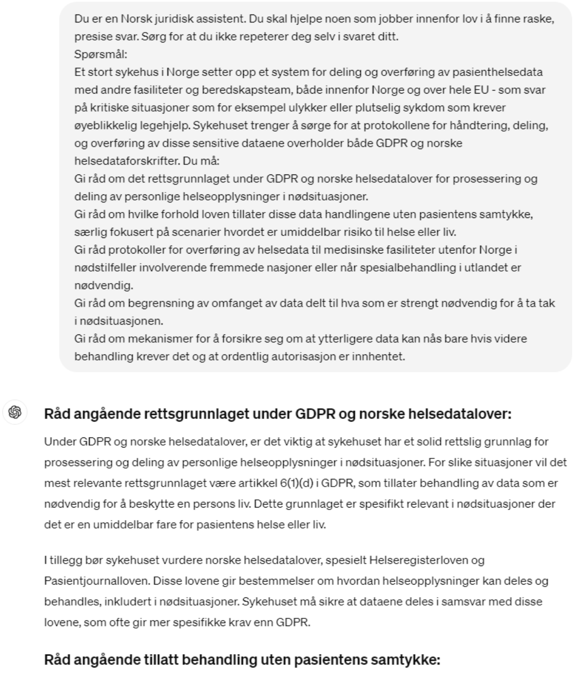
1. **Artikkel 6: Behandlingsgrunnlaget**
    - Under GDPR, prosessering av biometriske data for grensesikkerhetsformål kan være tillatt under Artikkel 6(1)(e): "Behandlingen er nødvendig for utførelsen av en oppgave i offentlig interesse eller utøvelse av offentlig myndighet som den behandlingsansvarlige har fått tildelt."
    - I tillegg kan det være relevant å se på Artikkel 9(2)(g): "Behandlingen er nødvendig for å utføre en oppgave i allmennhetens interesse eller som ledd i utøvelsen av offentlig myndighet som den behandlingsansvarlige har fått tildelt."
2. **Juridisk forpliktelse, offentlig interesse, eller vitale interesser:**
    - Behandlingen av biometriske data for grensesikkerhetsformål kan være basert på offentlig interesse for å opprettholde sikkerheten ved grensene (Artikkel 6).

Figure C.8: ChatGPT Expert evaluation scenario 3 answer pt. 1



2. **Juridisk forpliktelse, offentlig interesse, eller vitale interesser:**
    - Behandlingen av biometriske data for grensesikkerhetsformål kan være basert på offentlig interesse for å opprettholde sikkerheten ved grensene (Artikkel 6).
    - Det kan også være tilfeller der behandlingen er nødvendig for å oppfylle en juridisk forpliktelse som hviler på den behandlingsansvarlige (Artikkel 6).
    - Videre kan det være situasjoner der det er nødvendig for å beskytte vitale interesser til data subjektene eller andre (Artikkel 9).
3. **GDPR-kompatible tidsplaner for datalagring av biometrisk informasjon ved grensen:**
    - Lagringsperioden for biometrisk informasjon må være proporsjonal med formålet (Artikkel 5).
    - Det er viktig å utvikle klare retningslinjer som tar hensyn til behovet for grensesikkerhet og samtidig respekterer prinsippene om begrensning av lagringsperioden (Artikkel 5).
4. **Protokoller for sikker og lovlig deling av biometrisk data på tvers av EUs grenser:**
    - For deling av biometriske data over landegrenser innen EU, må man følge Artikkel 44-50 som handler om overføring av personopplysninger til tredjeland eller internasjonale organisasjoner.
    - Det kan være nødvendig med godkjente standardkontrakter (Artikkel 46), binding corporate rules (Artikkel 47), eller andre godkjente overføringsmekanismer.
5. **Interoperabilitet med eksisterende systemer som SIS II og VIS, samt samsvar med GDPR:**
    - For å sikre interoperabilitet, må systemet for biometriske data være utformet i tråd med prinsippene for innebygd personvern og personvern som standardinnstilling (Artikkel 25).
    - Samsvar med andre relevante regelverk, som SIS II og VIS, krever grundig evaluering av hvordan systemene kan integreres mens man opprettholder personvernet og sikkerheten (Artikkel 5 og Artikkel 32).

Figure C.9: ChatGPT Expert evaluation scenario 3 answer pt. 2



Du er en Norsk juridisk assistent. Du skal hjelpe noen som jobber innenfor lov i å finne raske, presise svar. Sørg for at du ikke repeterer deg selv i svaret ditt.
Spørsmål:
Et stort sykehus i Norge setter opp et system for deling og overføring av pasienthelsedata med andre fasiliteter og beredskapsteam, både innenfor Norge og over hele EU - som svar på kritiske situasjoner som for eksempel ulykker eller plutselig sykdom som krever øyeblikkelig legehjelp. Sykehuset trenger å sørge for at protokollene for håndtering, deling, og overføring av disse sensitive dataene overholder både GDPR og norske helsedataforskrifter. Du må:
Gi råd om det rettsgrunnlaget under GDPR og norske helsedatalover for prosessering og deling av personlige helseopplysninger i nødsituasjoner.
Gi råd om hvilke forhold loven tillater disse data handlingene uten pasientens samtykke, særlig fokusert på scenarier hvordet er umiddelbar risiko til helse eller liv.
Gi råd protokoller for overføring av helsedata til medisinske fasiliteter utenfor Norge i nødstilfeller involverende fremmede nasjoner eller når spesialbehandling i utlandet er nødvendig.
Gi råd om begrensning av omfanget av data delt til hva som er strengt nødvendig for å ta tak i nødsituasjonen.
Gi råd om mekanismer for å forsikre seg om at ytterligere data kan nås bare hvis videre behandling krever det og at ordentlig autorisasjon er innhentet.

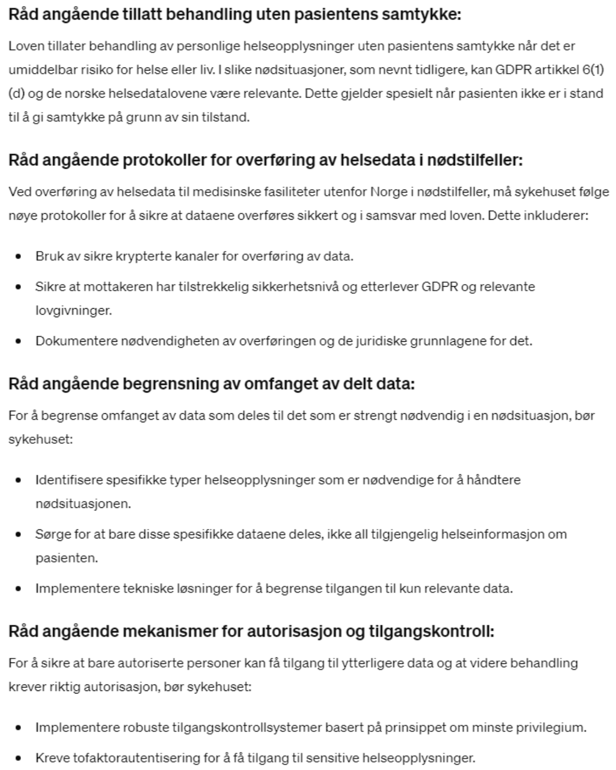
Råd angående rettsgrunnlaget under GDPR og norske helsedatalover:

Under GDPR og norske helsedatalover, er det viktig at sykehuset har et solid rettslig grunnlag for prosessering og deling av personlige helseopplysninger i nødsituasjoner. For slike situasjoner vil det mest relevante rettsgrunnlaget være artikkel 6(1)(d) i GDPR, som tillater behandling av data som er nødvendig for å beskytte en persons liv. Dette grunnlaget er spesifikt relevant i nødsituasjoner der det er en umiddelbar fare for pasientens helse eller liv.

I tillegg bør sykehuset vurdere norske helsedatalover, spesielt Helseregisterloven og Pasientjournalloven. Disse lovene gir bestemmelser om hvordan helseopplysninger kan deles og behandles, inkludert i nødsituasjoner. Sykehuset må sikre at dataene deles i samsvar med disse lovene, som ofte gir mer spesifikke krav enn GDPR.

Råd angående tillatt behandling uten pasientens samtykke:

Figure C.10: ChatGPT Expert evaluation scenario 4 answer pt. 1



Råd angående tillatt behandling uten pasientens samtykke:

Loven tillater behandling av personlige helseopplysninger uten pasientens samtykke når det er umiddelbar risiko for helse eller liv. I slike nødsituasjoner, som nevnt tidligere, kan GDPR artikkel 6(1)(d) og de norske helsedatalovene være relevante. Dette gjelder spesielt når pasienten ikke er i stand til å gi samtykke på grunn av sin tilstand.

Råd angående protokoller for overføring av helsedata i nødstilfeller:

Ved overføring av helsedata til medisinske fasiliteter utenfor Norge i nødstilfeller, må sykehuset følge nøye protokoller for å sikre at dataene overføres sikkert og i samsvar med loven. Dette inkluderer:

- Bruk av sikre krypterte kanaler for overføring av data.
- Sikre at mottakeren har tilstrekkelig sikkerhetsnivå og etterlever GDPR og relevante lovgivninger.
- Dokumentere nødvendigheten av overføringen og de juridiske grunnlagene for det.

Råd angående begrensning av omfanget av delt data:

For å begrense omfanget av data som deles til det som er strengt nødvendig i en nødsituasjon, bør sykehuset:

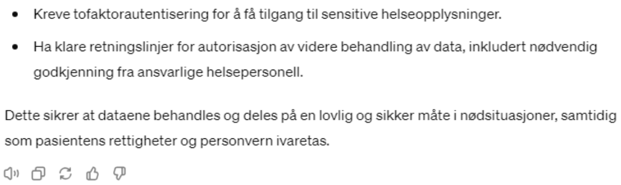
- Identifisere spesifikke typer helseopplysninger som er nødvendige for å håndtere nødsituasjonen.
- Sørge for at bare disse spesifikke dataene deles, ikke all tilgjengelig helseinformasjon om pasienten.
- Implementere tekniske løsninger for å begrense tilgangen til kun relevante data.

Råd angående mekanismer for autorisasjon og tilgangskontroll:

For å sikre at bare autoriserte personer kan få tilgang til ytterligere data og at videre behandling krever riktig autorisasjon, bør sykehuset:

- Implementere robuste tilgangskontrollsystemer basert på prinsippet om minste privilegium.
- Kreve tofaktorautentisering for å få tilgang til sensitive helseopplysninger.

Figure C.11: ChatGPT Expert evaluation scenario 4 answer pt. 2



- Kreve tofaktorautentisering for å få tilgang til sensitive helseopplysninger.
- Ha klare retningslinjer for autorisasjon av videre behandling av data, inkludert nødvendig godkjenning fra ansvarlige helsepersonell.

Dette sikrer at dataene behandles og deles på en lovlig og sikker måte i nødsituasjoner, samtidig som pasientens rettigheter og personvern ivaretas.

Figure C.12: ChatGPT Expert evaluation scenario 4 answer pt. 3

Appendix D

Evaluation Form

Evaluation form

Please evaluate the answers from the five different groups based on:

Accuracy

- Whether the answers provided by the AI are legally accurate and if they align with current GDPR regulations and interpretations.

Relevance

- The responses should be directly relevant to the questions asked, providing clear and pertinent information without digression.

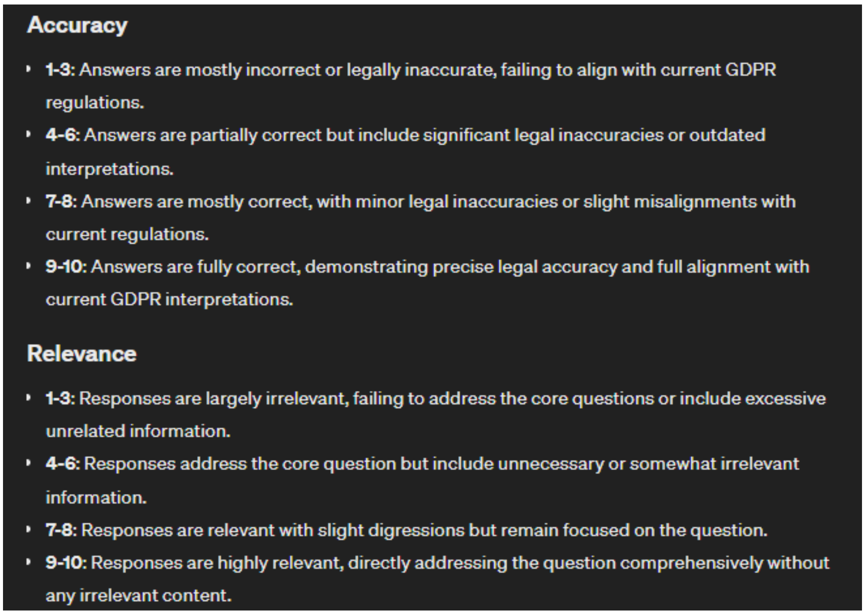
Completeness

- Whether all parts of the question have been answered, and if the response includes all necessary details to cover all aspects of the query.

Clarity

- To assess if the explanations are straightforward, structured, and understandable, and to evaluate the use of appropriate legal and technical terminology.

And give each criteria a **score** from 1 – 10, based on these scoring criteria definitions:

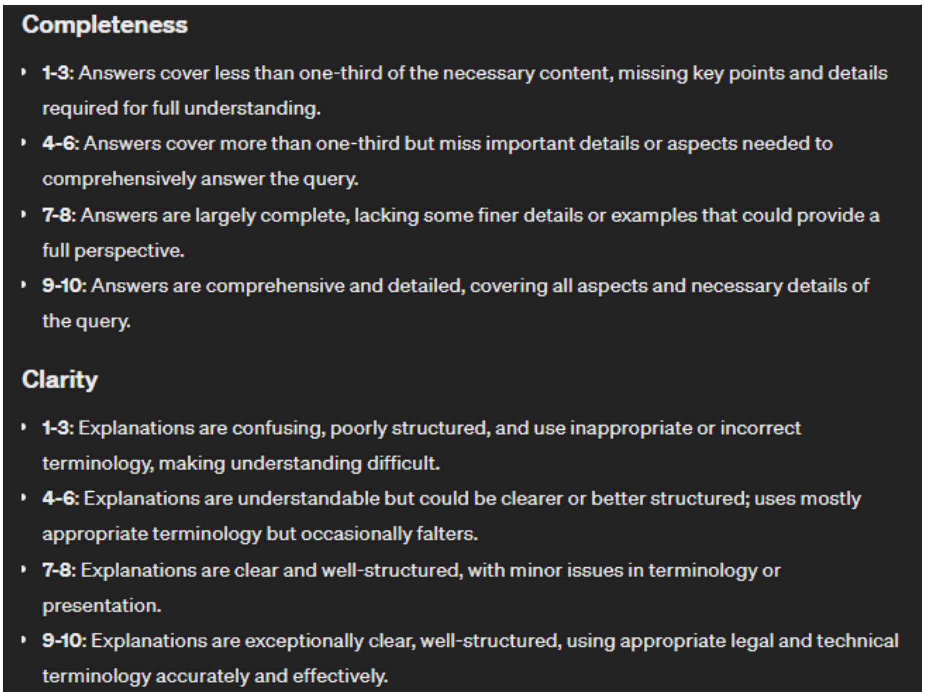


Accuracy

- **1-3:** Answers are mostly incorrect or legally inaccurate, failing to align with current GDPR regulations.
- **4-6:** Answers are partially correct but include significant legal inaccuracies or outdated interpretations.
- **7-8:** Answers are mostly correct, with minor legal inaccuracies or slight misalignments with current regulations.
- **9-10:** Answers are fully correct, demonstrating precise legal accuracy and full alignment with current GDPR interpretations.

Relevance

- **1-3:** Responses are largely irrelevant, failing to address the core questions or include excessive unrelated information.
- **4-6:** Responses address the core question but include unnecessary or somewhat irrelevant information.
- **7-8:** Responses are relevant with slight digressions but remain focused on the question.
- **9-10:** Responses are highly relevant, directly addressing the question comprehensively without any irrelevant content.



Completeness

- **1-3:** Answers cover less than one-third of the necessary content, missing key points and details required for full understanding.
- **4-6:** Answers cover more than one-third but miss important details or aspects needed to comprehensively answer the query.
- **7-8:** Answers are largely complete, lacking some finer details or examples that could provide a full perspective.
- **9-10:** Answers are comprehensive and detailed, covering all aspects and necessary details of the query.

Clarity

- **1-3:** Explanations are confusing, poorly structured, and use inappropriate or incorrect terminology, making understanding difficult.
- **4-6:** Explanations are understandable but could be clearer or better structured; uses mostly appropriate terminology but occasionally falters.
- **7-8:** Explanations are clear and well-structured, with minor issues in terminology or presentation.
- **9-10:** Explanations are exceptionally clear, well-structured, using appropriate legal and technical terminology accurately and effectively.

Also, please give a short **overall comment** of the answers where you give your own feedback and thoughts on the answer.

Group 1

Question 1

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 2

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 3

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 4

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Group 2

Question 1

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 2

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 3

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 4

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Group 3

Question 1

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 2

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 3

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 4

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Group 4

Question 1

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 2

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 3

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 4

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Group 5

Question 1

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 2

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 3

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Question 4

Accuracy:

Relevance:

Completeness:

Clarity:

Overall comment:

Appendix E

Program Answers for Evaluation

The questions have been grouped together on which methods were used to get the answers to better evaluate which ones help more, if any.

Some places in the answers it says ‘prossesering’ instead of the correct ‘prosessering’. This is (most likely) a reflection of a spelling mistake in the prompts (inputs). That is my bad.

Some answers terminate mid-sentence. This is due to a technical limitation where we must set a boundary on how much text the model can generate as it only supports a certain amount of text. A bigger model would support the generation of more text, but due to resource constraints we can only run medium to low end models size wise. If you want a more detailed explanation, feel free to ask.

All the answers are from one ‘run through’ of the program of each method, so none of them were handpicked. As such some answers are more helpful than others.

Wherever it says ‘#*Note:* […]’, that is a comment from me relating to the answer.

Group 1

Question 1.

For å gi en detaljert, juridisk forsvarlig responsstrategi for hver type forespørsel, må vi først vurdere de nødvendige lovlige grunnlagene og konfliktene som kan oppstå mellom forespørslene.

1. Data korreksjon:
   a. Identifiser de nøyaktige dataene som skal endres.
   b. Sikre at endringene er i samsvar med personvernlovgivningen.
   c. Gi den registrerte muligheten til å overprøve endringene og kreve rett til saksomfattende.

2. Data slettelse:
   a. Identifiser de spesifikke dataene som skal slettes.
   b. Sørg for at slettingen skjer i samsvar med personvernlovgivningen.
   c. Gi den registrerte muligheten til å overprøve slettingen og kreve rett til saksomfattende.

3. Restriksjoner på prossesering:
   a. Identifiser de spesifikke prosesseringene som skal begrenses.
   b. Sørg for at begrensningene er i samsvar med personvernlovgivningen.
   c. Gi den registrerte muligheten til å overprøve begrensningene og kreve rett til saksomfattende.

Når det gjelder potensielle konflikter mellom forespørslene og andre juridiske forpliktelser, må vi ta hensyn til de spesifikke lover og regler som er nevnt i forordningen. For eksempel, kan det være konflikter mellom forespørslene og lover om oppbevaring av data.

For å verifisere nøyaktigheten av den nye dataen gitt av den registrerte, bør selskapet ta følgende skritt:

1. Vurder den nye dataen i forhold til personvernlovgivningen.
2. Sørg for at den nye dataen er i samsvar med den opprinnelige dataen som ble oppbevart.
3. Gi den registrerte muligheten til å overprøve den nye dataen og kreve rett til saksomfattende.

Ved å ta disse skrittene kan selskapet sikre at det har et solid og juridisk forsvarlig forhold for hver type forespørsel og unngå potensielle konflikter mellom forespørslene og andre juridiske forpliktelser.

Question 2.

For å navigere disse komplekse regulatoriske kravene, bør banken følge disse trinnene:

1. Identifisere og avstemme kravene til GDPR med lokale lover: Banken bør undersøke de spesifikke kravene som pålegges av GDPR og sammenligne dem med de lokale personvernlovene som gjelder i de tredjelandene banken opererer i. Dette vil sikre at banken er i samsvar med de strengeste kravene som pålegges.

2. Gi råd om hvordan lovene skal prioriteres når motstridende forpliktelser oppstår: Banken bør vurdere de ulike lovene som pålegges og prioritere de som gir størst beskyttelse for personvernet til de involverte individene.

3. Gi råd om innarbeiding av standard personvernbestemmelser (SCC) eller bindende bedriftsregler (BCSR) som mekanismer for lovlig dataoverføring utenfor EU: Banken bør undersøke og forstå SCC og BCSR for å bestemme hvilken av disse som er mest passende for å sikre lovlig dataoverføring og oppfylle GDPR-kravene.

4. Gi råd om eventuelle nødvendige endringer eller tillegg til SCCene og BCSRene for å sikre samsvar: Hvis det er nødvendig, bør banken gjennomgå og endre SCC og BCSR for å sikre at de er i samsvar med GDPR og de lokale personvernlovene som gjelder.

5. Sammenligne disse bestemmelsene med lokale personvernlover i tredjelandene hvor banken opererer, identifiser strengere krav eller avvik: Banken bør sammen billig med lokale personvernlover i tredjelandene for å identifisere eventuelle strengere krav eller avvik som må tas hensyn til.

Ved å følge disse trinnene vil banken være i stand til å navigere de komplekse regulatoriske kravene og sikre at de er i samsvar med GDPR og de lokale personvernlovene som gjelder.

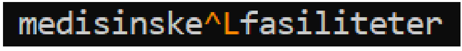
Question 3.

GDPR tillater prossesering av biometrisk data for grensesikkerhetsformål basert på artikkel 6, som spesifikt nevner lovens grunnlag for behandling av personopplysninger. Grunnene som er oppført i artikkel 6 inkluderer for eksempel behovet for å opprettholde en rettferdig og åpen behandling av personopplysninger (artikkel 6.1), oppfyllelse av en juridisk forpliktelse (artikkel 6.2), og beskyttelse av vitale interesser til data subjekter (artikkel 6.3).

GDPR krever også at behandlingsansvarlige og databehandlere utvikler GDPR-kompatible tidsplaner for lagring av biometrisk informasjon samlet ved grensen (artikkel 6.1.a), og opprettholder en tidsplan for sikker og lovlig deling av biometrisk data på tvers av EUs grenser (artikkel 6.1.b). Videre krever GDPR at databehandlere sikrer interoperabilitet av biometriske systemer med eksisterende systemer som SIS II og VIS, samt sikre at samsvaret oppfyller både GDPR og andre relevante regelverk innenfor EU (artikkel 6.1.c).

I tillegg krever GDPR at behandlingsansvarlige og databehandlere opprettholder en klar og konsis forklaring av grunnene for behandlingen av biometriske data, og sikrer at denne informasjonen blir gitt til de registrerte på en tydelig og kortfattet måte (artikkel 6.1.d).

Question 4.

#Note: 'Medisinskefasiliteter' is meant to be two words; 'medisinske fasiliteter'. The model output it like this  `medisinske^Lfasiliteter`*. I don't know what it means, but from Google it seems to be related to VIM (which I don't use). The model also wrote the word 'finally', which is a mistake.*

Det rettsgrunnlaget under GDPR og norske helsedatalover for prossesering og deling av personlige helseopplysninger i nødsituasjoner er som følger:

1. GDPR - Artikkel 9 og 10: Disse artiklene gir retningslinjer for nødvendige og berettigede grunner for behandling av personopplysninger i nødstilfeller.
2. Norsk helsedatalov - Lovens artikler 46, 47, 49, 50 og 51: De gir retningslinjer for behandling, deling og overføring av personopplysninger i nødstilfeller, samt sikring av personopplysningers vern.

Forhold som tillater disse data handlingene uten pasientens samtykke i nødstilfeller:

1. Kritiske situasjoner: For eksempel ulykker eller plutselig sykdom som krever øyeblikkelig legehjelp.
2. Fremmede nasjoner: Overføring av helsedata til medisinskefasiliteter i utlandet for å imøtekomme nødsituasjoner.
3. Spesialbehandling: Overføring av helsedata til spesialister eller fagfolk i utlandet for å imøtekomme spesialbehandlingsbehovet.

Protokoller for overføring av helsedata til medisinske fasiliteter utenfor Norge i nødstilfeller: finally, det er viktig å følge de spesifikke retningslinjene og forskriftene som er fastsatt i GDPR og norske helsedatalover for behandling, deling og overføring av personopplysninger i nødstilfeller.

Group 2

Question 1.

#Note: I wrote the word 'forespørslene' wrong in the prompt, and that spelling mistake is reflected in this answer. The model did however manage to overlook this in Question 1 for Group 1.

For å gi en detaljert, juridisk forsvarlig responsstrategi for hver type forespørsel, må vi først analysere de spesifikke lovgivningene som er relevante for hver type forespørsel. Deretter vil vi kunne foreslå potensielle konflikter og veilede selskapet om hvordan de kan verifisere nøyaktigheten av den nye dataen.

1. Data korreksjon:
For å håndtere forespørsler om data korreksjon, må selskapet følge GDPRs artikkel 12, som innebærer at de må opprettholde en register over personopplysninger og oppdatere den når det er nødvendig. Selskapet må også ha en mekanisme for å ta tak i bekymringer fra registrerte personer.

2. Data sletting:

For å håndtere forespørsler om data sletting, må selskapet følge GDPR Irritasjon 17, som krever at de fjerner personopplysninger når det er nødvendig, som for eksempel når det er en bekymring om oppbevaring av personopplysninger eller når en person har gitt en tydelig og uttrykkelig begrunnelse.

3. Restriksjoner på prosessering:
For å håndtere forespørsler om restriksjoner på prosessering, må selskapet følge GDPRs artikkel 18, som innebærer at de må ta en beslutning innen en rask periode og ikke behandle personopplysninger som er omfattet av en restriksjon.

4. Verifisering av nye data:
For å verifisere nøyaktigheten av den nye dataen, bør selskapet følge GDPRs artikkel 4, som krever at de samler nøyaktige og oppdatert personopplysninger. De bør også sikre at de har en riktig prosess for å opprettholde og oppdatere personopplysninger, og at de følger GDPRs andre bestemmelser for oppbevaring av data.

Potensielle konflikter:
- Konflikter mellom forespørselene og andre juridiske forpliktelser (f.eks. lover om oppbevaring av data).
- Konflikter mellom forespørselene og selskapets eksisterende prosedyrer og praksis.
- Konflikter mellom forespørselene og selskapets internasjonale forpliktelser og praksis.

Veiledning:
- Sørg for at selskapet vurderer alle relevante lover og bestemmelser når de behandler forespørsler.
- Veiled selskapet om hvordan de kan verifisere nøyaktigheten av den nye dataen, og sørg for at de følger GDPRs bestemmelser for oppbevaring av data.
- Sørg for at selskapet er klar over sine internasjonale forpliktelser og praksis og hvordan de kan påvirke deres forpliktelser og praksis.

Question 2.

For å identifisere og avstemme kravene til GDPR med lokale lover, bør du først undersøke GDPR og de påfølgende forskriftene som er gitt av Norge. Deretter bør du analysere de relevante delene av personopplysningsloven og forordningen for å identifisere eventuelle konflikter eller ulikheter mellom GDPR og lokale lover. I tillegg bør du sammenligne de relevante bestemmelsene i personopplysningsloven og forordningen med de påfølgende forskriftene og lovene i tredjelandene der banken opererer. Dette vil gi deg en bedre forståelse av de potensielle konfliktene og mulige endringer som kan kreves for å sikre samsvar.

Question 3.

#Note: ‘Sletteing’ and ‘behandlingsansvarligens’ are both spelling mistakes. ‘biångulære’ is not a word, and so is not ‘børɸpu’. This answer also terminates pre-maturely. All in all, not the programs greatest work.

Gjennomgang av GDPR-kompatible tidsplaner for datalagring for biometrisk informasjon samlet ved grensen:
- Datalagring av biometrisk informasjon bør være i samsvar med GDPR-kravene og den lokale datalagringsloven.

- Databeskyttelsesmyndigheten bør fastsette tidsplaner for datalagring for biometrisk informasjon samlet ved grensen.
- Tidsplanene bør ta hensyn til behovet for personvern og sikkerhet for grenseoverskridende informasjon.

Behandling av personopplysninger på grunnlag av offentlig interesse, juridisk forpliktelse eller vitale interesser til data subjekter eller andre:
- Offentlig interesse, juridisk forpliktelse eller vitale interesser til data subjekter eller andre kan være grunnlaget for behandling av personopplysninger.
- Behandlingen bør være nødvendig og forsvarlig, og den skal ikke oppgis uten tilstrekkelig informasjon.
- Den registrerte bør ha rett til å få informasjon om behandlingen og gjøre krav om sletteing av personopplysninger.

Kravet om grunnlagsbevis og behandlingsansvarligens ansvar:
- Behandlingsansvarlige bør gjennomgå og godkjenne behandlingsgrunnlaget og sikre at det oppfyller GDPR-kravene.
- Behandlingsansvarlige bør ha en forpliktelse til å gjennomgå behandling av personopplysninger og sikre at de er i samsvar med GDPR-kravene.
- Behandlingsansvarlige bør ha en plan for å oppdage og håndtere feil i behandlingen av personopplysninger.

Utviklingen av GDPR-kompatible tidsplaner for datalagring for biometrisk informasjon samlet ved grensen:
- Databeskyttelsesmyndigheten bør utarbeide tidsplaner for datalagring for biometrisk informasjon samlet ved grensen.
- Tidsplanene bør ta hensyn til behovet for personvern og sikkerhet for grenseoverskridende informasjon.
- Tidsplanene bør være i samsvar med GDPR-kravene og den lokale datalagringsloven.

Prosessering av biångulære personopplysninger i henhold til GDPR:
- Biångulære personopplysninger kan prosesseres i henhold til GDPR-kravene, men det bør være i samsvar med de spesifikke kravene i GDPR.
- Behandlingsansvarlige bør vurdere om behandlingen er nødvendig og forsvarlig, og de bør ha en forpliktelse til å gjennomgå og sikre at behandlingen er i samsvar med GDPR-kravene.
- Den registrerte bør ha rett til å få informasjon om behandlingen og gjøre krav om sletteing av personopplysninger.

Samsvar mellom GDPR og den lokale datalagringsloven:
- Databeskyttelsesmyndigheten børфри GDPR-kravene og den lokale datalagringsloven for å sikre at behandlingen av personopplysninger er i samsvar med GDPR-kravene.
- Behandlingsansvarlige bør vurdere om behandlingen er nødvendig og forsvarlig, og de bør ha en forpliktelse til å gjennomg

Question 4.

Det er ikke nok informasjon i konteksten til å svare på spørsmålet ditt. Kan du gi mer informasjon eller kontekst om hvilket spørsmål du har?

Group 3

Question 1.

#Note: This time it added ‘ĭ’ at the end of the word ‘personopplysning’. Also it should be ‘et register’. Also also terminated pre-maturely.

For å gi en detaljert, juridisk forsvarlig responsstrategi for hver type forespørsel, bør du følge disse stegene:

1. Forespørsel om data korreksjon:
   - Sørg for at den registrerte har rett til å kontrollere og forbedre personopplysningene sine, slik det er fastsatt i GDPR.
   - Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og oppdatere personopplysningĭ.
   - Oppretthold en register over endringer og endringer i personopplysningene.
   - Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en tredjestat under GDPR.

2. Forespørsel om data slettelse:
   - Sørg for at den registrerte har rett til å få personopplysningene sine slettet når det er nødvendig, slik det er fastsatt i GDPR.
   - Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og be om sletting av personopplysningĭ.
   - Oppretthold en register over endringer og endringer i personopplysningene.
   - Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en tredjestat under GDPR.

3. Forespørsel om data begrensning:
   - Sørg for at den registrerte har rett til å begrense bruken av personopplysningĭ, slik det er fastsatt i GDPR.
   - Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og begrense bruken av personopplysningĭ.
   - Oppretthold en register over endringer og endringer i personopplysningĭ.
   - Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en tredjestat under GDPR.

4. Forespørsel om bekreftelse av data:
   - Sørg for at den registrerte har rett til å be om bekreftelse av personopplysningĭ, slik det er fastsatt i GDPR.
   - Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og be om bekreftelse av personopplysningĭ.
   - Oppretthold en register over bekreftede personopplysningĭ.
   - Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en tredjestat under GDPR.

5. Forespørsel om opplysninger om data:
   - Sørg for at den registrerte har rett til å få opplysninger om personopplysningĭ, slik det er fastsatt i GDPR.
   - Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og be om opplysninger om personopplysningĭ.

- Oppretthold en register over opplysninger om personopplysningǐ.
- Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en tredjestat under GDPR.

6. Forespørsel om overføring av data:

- Sørg for at den registrerte har rett til å be om at personopplysningǐ overføres til en annen person eller tjeneste, slik det er fastsatt i GDPR.
- Sørg for at den registrerte har en enkel og effektiv prosess for å kontakte og be om overføring av personopplysningǐ.
- Oppretthold en register over overføringer av personopplysningǐ.
- Sørg for at den registrerte har rett til å få slik informasjon som er gitt til en

Question 2.

For å kunne gi en detaljert analyse og identifisere eventuelle motstridende forpliktelser, må jeg først vite hvilke lover og forordninger som er relevante for denne saken. Kan du gi mer informasjon om konteksten i spørsmålet, som for eksempel de relevante EU-forordningene og personvernlovene i tredjelandene?

Question 3.

Basert på informasjonen som er gitt, er det ikke mulig å gi spesifikk advarsel om de spesifikke artiklene og klausulene under GDPR som tillater prosseseringen av biometrisk data for grensesikkerhetsformål. Imidlertid er det nevnt i artikkel 87 at medlemsstatene kan fastsette nærmere særlige vilkår for behandling av et nasjonalt identifikasjonsnummer eller andre generelle identifikatorer. Videre er det nevnt i artikkel 28 at databehandlere bør gi tilstrekkelige garantier for at de vil gjennomføre egnede tekniske og organisatoriske tiltak for å sikre at behandlingen oppfyller kravene i denne forordningen og vern av den registrerte sine rettigheter. Det er også nevnt i artikkel 39 at personopplysninger for arkivformål i allmennhetens interesse, for formål knyttet til vitenskapelig eller historisk forskning eller for statistiske formål bør være omfattet av nødvendige garantier som sikrer den registrertes rettigheter og friheter i henhold til denne forordningen.

Question 4.

Gjennom GDPR og norske helsedataforskrifter er det tillatt for pasienter å få personopplysninger delt uten deres samtykke i nødsituasjoner som involverer kritisk helse eller livsfare. Disse situasjonene omfatter ulykker, plutselig sykdom eller andre situasjoner som krever øyeblikkelig medisinsk hjelp. Dataene som kan deles i slike situasjoner inkluderer informasjon om pasientens helse, medisinske historie og medisinske behandlinger.

For å sikre at disse dataene deles på en lovlig og ansvarlig måte, er det viktig å følge bestemte retningslinjer og protokoller. For eksempel kan pasientens samtykke samles inn, og de kan ha muligheten til å rette opp en slik deling hvis de ønsker det. I tillegg er det viktig å begrense den opplyste data som deles til det som er strengt nødvendig for å håndtere nødsituasjonen.

I tilfeller der data deles utenfor Norge, er det viktig å følge norske helsedataforskrifter og GDPR-krav om personvern og databeskyttelse. Dette kan innebære å sikre at databehandleren har tillatelse til å behandle personopplysninger, og at de er klar over de norske lovgivningene som gjelder. Det er også viktig å sikre at ytterligere data kan nås bare hvis det er nødvendig for videre behandling og at det er innhentet riktig autorisasjon.

Group 4

Question 1.

#Note: The model put part of the question in the answer at the beginning.

Et techselskap basert i Norge utvikler protokoller for å håndtere forespørsler fra EU-datasubjekt under GDPR Kapittel 3 (Artikler 12 til 23), som inkluderer data korreksjon, data slettelse, og restriksjoner på prossesering. Hver forespørsel type er basert på forskjellige lovlige grunnlag. Selskapet trenger å sikre at selskapets prosedyrer er kompatible med GDPR og at den tar i betraktning nyansene i Norske personvernforskrifter.

For å gi en detaljert, juridisk forsvarlig responsstrategi for hver type forespørsel:

1. Data Korreksjon:
   a. Sørg for at prosedyren er i samsvar med GDPR Artikkel 12 og Norske personvernforskrifter.
   b. Identifiser og rette eventuelle feil i personopplysningene.
   c. Bekreft endringene i personopplysningene med den registrerte.
   d. Håndtere eventuelle bekymringer knyttet til den registrerte sine rettigheter.

2. Data Slettelse:
   a. Sørg for at prosedyren er i samsvar med GDPR Artikkel 17 og Norske personvernforskrifter.
   b. Identifiser og slette personopplysningene i henhold til den registrerte sine rettigheter.
   c. Håndtere eventuelle bekymringer knyttet til den registrerte sine rettigheter.
   d. Dokumenter prosedyren for å gi bevis på at personopplysningene er slettet.

3. Restriksjoner på Prossesering:
   a. Sørg for at prosedyren er i samsvar med GDPR Artikkel 18 og Norske personvernforskrifter.
   b. Identifiser og implementere restriksjoner på prosseseringen av personopplysningene.
   c. Håndtere eventuelle bekymringer knyttet til den registrerte sine rettigheter.
   d. Dokumenter prosedyren for å gi bevis på at restriksjonene er implementert.

For å forutsi potensielle konflikter mellom forespørslene og andre juridiske forpliktelser, bør selskapet analysere de norske personvernforskriftene og GDPR, samt eventuelle andre relevante lover og forordninger som kan påvirke prosedyrene. Det er viktig å vurdere eventuelle konflikter og rette opp i dem, og sikre at selskapets prosedyrer er i samsvar med alle påfølgende forpliktelser.

For å verifisere nøyaktigheten av den nye dataen gitt av den registrerte før det gjøres endringer, bør selskapet:

1. Vurdere og verifisere den registrerte sine rettigheter og bekreft at de har rett til å gi personopplysningene.
2. Bekreft identiteten til den registrerte og sammenligne den gitte informasjonen med deres opprinnelige opplysninger.
3. Sjekke den gitte informasjonen for eventuelle feil eller ufullstendigheter.

4. Bekreft at den nye dataen er i samsvar med de opprinnelige opplysningene og Norske personvernforskrifter.
5. Dokumenter prosedyren for å gi bevis på at den nye dataen er nøyaktig og i samsvar med de påfølgende forpliktelsene.

Question 2.

Basert på informasjonen som er gitt, er det ikke mulig å identifisere og avstemme kravene til GDPR med lokale lover eller gi råd om innarbeiding av standard personvernbestemmelser (SCC) eller bindende bedriftsregler (BCSR) for å sikre lovlig dataoverføring utenfor EU. Dette dokumentet diskuterer primært personopplysningsloven i Norge og dens relevans for GDPR og personvernforordningen.

Question 3.

#Note: Spelling mistake, ‘offentilg’ should be ‘offentlig’. This was again a mistake from the prompt reflected in the answer.

For å svare på spørsmålet ditt, la oss først se på relevant informasjon som er gitt i konteksten:

1. Artikkel 33: Melding til tilsynsmyndigheten om brudd på personopplysningssikkerheten
2. Artikkel 20: Rett til dataportabilitet
3. Artikkel 46: Overføring av personopplysninger til tredjestater eller internasjonale organisasjoner
4. Artikkel 9: Særlige kategorier av opplysninger
5. Artikkel 10: Straffedommer og lovovertredelser

Nå kan vi gi råd om hver av de spesifikke artiklene og klausulene under GDPR som tillater prosseseringen av biometrisk data for grensesikkerhetsformål:

1. Artikkel 33: Melding til tilsynsmyndigheten om brudd på personopplysningssikkerheten
2. Artikkel 20: Rett til dataportabilitet
3. Artikkel 46: Overføring av personopplysninger til tredjestater eller internasjonale organisasjoner
4. Artikkel 9: Særlige kategorier av opplysninger
5. Artikkel 10: Straffedommer og lovovertredelser

For behandlingen som er basert på offentilg interesse, juridisk forpliktelse, eller vitale interesser til data subjektene eller andre, og sikre at begrunnelsen stemmer overens med GDPRs krav, kan vi gi råd om at det er nødvendig å vurdere artiklene og klausulene som er gitt i konteksten, og sikre at de er i tråd med GDPRs krav.

For utviklingen av GDPR-kompatible tidsplaner for datalagring for biometrisk informasjon samlet ved grensen, kan vi gi råd om at det er nødvendig å sikre at disse tidsplanene er i tråd med GDPRs krav og retningslinjer for datalagring og dataproteksjon.

For protokoller for sikker og lovlig deling av biometrisk data på tvers av EUs grenser med andre medlemsland og relevante instanser, kan vi gi råd om at det er nødvendig å sikre at disse protokollerne er i tråd med GDPRs krav og retningslinjer for dataproteksjon og dataskyting.

Til slutt, for interoperabiliteten av det biometriske systemet sammen med eksisterende systemer over hele EU slik som SIS II og VIS, og sikre samsvar med både GDPR og andre relevante regelverk innenfor EU, kan vi gi råd om at det er nødvendig å sikre at det Area biometriske systemet er designet for å samsvare med eksisterende systemer og retningslinjer for dataproteksjon og dataskyting.

Konklusjonen er at det er viktig å vurdere GDPRs krav og retningslinjer for dataproteksjon og dataskyting når man utvikler og implementerer systemer for prossesering av biometrisk data for grensesikkerhetsformål, og sikre at behandlingen er basert på offentilg interesse, juridisk forpliktelse, eller vitale interesser til data subjektene eller andre. Det er også viktig å utvikle GDPR-kompatible tidsplaner for datalagring, sikre samsvar med GDPRs krav og retningslinjer for dataproteksjon og dataskyting, og sikre interoperabilitet med eksisterende systemer over hele EU.

Question 4.

#Note: Grammatical mistake, 'å delte' should be 'å dele'.

Det rettsgrunnlaget under GDPR og norske helsedatalover for prossesering og deling av personlige helseopplysninger i nødsituasjoner er forpliktelsen til å beskytte de registrerte personenes rettigheter og friheter. Dette innebærer at helsepersonell må ta beslutninger som er i samsvar med de gjeldende reglene om vern av personopplysninger og de forpliktelsene som er knyttet til behandling av personopplysninger.

For å tillate disse data handlingene uten pasientens samtykke, er det nødvendig å følge nødsituasjoner som involverer umiddelbar risiko til helse eller liv. For eksempel kan det være nødvendig å skaffe medisinsk hjelp til en pasient som er i kritisk tilstand eller som er involvert i en plutselig ulykke. I slike situasjoner kan helsepersonell, som leger, sykepleiere og andre helsepersonell, delte pasientens personopplysninger med andre helsepersonell, helsefasiliteter og nødsituasjoner uten pasientens samtykke.

For å sikre at protokollene for håndtering, deling og overføring av disse sensitive dataene overholder GDPR og norske helsedatalover, bør helsepersonell ta følgende forholdsregler:

1. Sikre at de er klar over de nødvendige reglene og forskrifter om personopplysninger og nødsituasjoner.
2. Samle pasientens samtykke før de deler eller overfører personopplysninger til andre helsepersonell, helsefasiliteter eller nødsituasjoner.
3. Vurdere om det er nødvendig å delte eller overføre personopplysninger til andre helsepersonell, helsefasiliteter eller nødsituasjoner for å unngå umiddelbar risiko til helse eller liv.
4. Håndtere personopplysninger sikkerhetstiltak, som kryptering og sikker overføring av data.
5. Sikre at pasientens personopplysninger bare blir delt eller overført til nødvendig og strengt nødvendig informasjon for å ta tak i nødsituasjonen.
6. Sikre at ytterligere data kan nås bare hvis videre behandling krever det og at ordentlig autorisasjon er innhentet.
7. Gi registrerte personer informasjon om behandlingen av deres personopplysninger og deres rettigheter og friheter i henhold til personopplysningsloven.

Ved å følge disse forholdsreglene kan helsepersonell sikre at protokollene for håndtering, deling og overføring av personopplysninger i nødsituasjoner overholder GDPR og norske helsedatalover, og dermed beskytte de registrerte personenes rettigheter og friheter.

Group 5

Question 1.

1. Datakorreksjon (Artikkel 16 GDPR):

Responsstrategi:

1. Identifiser riktig lovlige grunnlag for korreksjon. Dette kan være basert på at dataene er feil, ufullstendige eller ikke relevante.

2. Utarbeid en prosedyre for å motta og verifisere korreksjonsforespørsler, inkludert hvordan den registrerte kan sende inn slike forespørsler.

3. Sikre at endringer loggføres i henhold til kravene i GDPR Artikkel 30.

Potensielle konflikter:

1. Konflikt med lover om oppbevaring av data: Sørg for at korreksjoner ikke endrer opprinnelig data som må oppbevares i henhold til andre lover. Vurder å opprette en kopi av opprinnelige data før korreksjon gjennomføres.

2. Konflikt med databeskyttelsesprinsippet: Balanser retten til korreksjon med behovet for å beskytte integriteten til opprinnelige data.

Verifisering av nøyaktighet:

1. Implementer en totrinnsverifisering. Be den registrerte om å bekrefte korreksjonen via en bekreftelses-e-post eller -melding.

2. Tilby muligheten for den registrerte å gi dokumentasjon eller annen støtte for korreksjonen.

2. Data slettelse (Artikkel 17 GDPR):

Responsstrategi:

1. Fastsett riktig lovlige grunnlag for sletting. Dette kan være basert på at dataene ikke lenger er nødvendige for formålet de ble samlet inn for.

2. Utvikle en prosedyre for å håndtere sletteforespørsler, inkludert hvordan dataene vil bli permanent fjernet fra systemer og eventuelle tredjeparter.

3. Sikre at slettingen er i samsvar med GDPR's krav til "rett til å bli glemt."

Potensielle konflikter:

1. Konflikt med lovgivning om lagring: Sørg for at slettingen ikke bryter med plikten til å beholde data i henhold til andre lover. Implementer et system for å markere data som er underlagt lovpålagt oppbevaring, slik at de ikke slettes feilaktig.

2. Konflikt med krav om loggføring: Bevar loggføringsdata som viser at slettingen er gjennomført i samsvar med GDPR Artikkel 17.

Verifisering av nøyaktighet:

1. Implementer en bekreftelsesprosess for sletting. Be den registrerte om å bekrefte at dataene er slettet fra alle relevante systemer.

2. Opprett en rapport eller kvittering for den registrerte som bekrefter slettingen.

3. Restriksjoner på prosessering (Artikkel 18 GDPR):

Responsstrategi:

1. Identifiser riktig lovlige grunnlag for å begrense prosessering. Dette kan være når den registrerte bestrider nøyaktigheten av dataene.

2. Utvikle en prosedyre for å håndtere begrensning av prosessering, inkludert hvordan dataene vil bli merket og begrensningene håndheves.

3. Informer eventuelle tredjeparter som har mottatt dataene om begrensningen.

Potensielle konflikter:

1. Konflikt med tredjepartsavtaler: Vurder hvordan begrensningen på prosessering kan påvirke eksisterende avtaler med tredjeparter som behandler dataene.

2. Konflikt med krav om dataintegritet: Balanser behovet for å begrense prosessering med kravet om å opprettholde integriteten til dataene.

Verifisering av nøyaktighet:

1. Implementer en prosess for å bekrefte nøyaktigheten av dataene før begrensningen blir opphevet.

2. Be om at den registrerte sender inn dokumentasjon eller annen støtte for å bekrefte behovet for begrensning.

Ved å følge disse strategiene, bør selskapet kunne håndtere korreksjoner, slettinger og begrensninger på en måte som er i samsvar med GDPR og norske personvernforskrifter, samtidig som de tar hensyn til potensielle konflikter med andre juridiske forpliktelser.

Question 2.

For det første er det viktig å identifisere og avstemme kravene til GDPR med lokale lover i de relevante EU-landene og tredjelandene der banken opererer. Dette krever en grundig gjennomgang av GDPR-artiklene og de spesifikke kravene i de lokale lovene.

Når det oppstår motstridende forpliktelser mellom GDPR og lokale lover, bør prioritering av lovene baseres på flere faktorer. For det første bør man vurdere om det er spesifikke unntak eller tilpasninger i GDPR som tillater eller krever etterlevelse av lokale lover. Deretter bør man vurdere hvilke lover som gir den høyeste beskyttelsen for persondata. Dersom det er umulig å etterleve begge sett med lover samtidig, bør man dokumentere grundig hvorfor man velger å prioritere den ene loven fremfor den andre.

Når det gjelder innarbeiding av standard personvernbestemmelser (SCC) eller bindende bedriftsregler (BCSR) som mekanismer for lovlig dataoverføring utenfor EU, bør banken vurdere hvilken mekanisme som passer best for deres spesifikke situasjon. SCCene er standardiserte avtaler utarbeidet av EU-kommisjonen for overføring av personopplysninger til tredjeland. BCSR er interne regler som er bindende for alle enheter innen banken og kan være gunstig for å sikre samsvar med GDPR.

Når det gjelder nødvendige endringer eller tillegg til SCCene og BCSRene for å sikre samsvar, bør banken vurdere om det er behov for å inkludere spesifikke krav fra lokale personvernlover i disse dokumentene. Dette kan være krav om lagringstider, informasjonsplikter eller andre spesifikke krav som er strengere enn GDPR.

Når man sammenligner disse bestemmelsene med lokale personvernlover i tredjelandene hvor banken opererer, er det viktig å identifisere eventuelle strengere krav eller avvik. Dette krever en detaljert gjennomgang av de lokale lovene for å sikre at banken oppfyller de høyeste standardene for personvern i hvert land de opererer i. Det kan være nødvendig å tilpasse sine interne retningslinjer og prosesser for å møte disse lokale kravene.

Question 3.

Spesifikke artikler og klausuler under GDPR for prosessering av biometrisk data for grensesikkerhetsformål:

1. Artikkel 6: Behandlingsgrunnlaget

a. Under GDPR, prosessering av biometriske data for grensesikkerhetsformål kan være tillatt under Artikkel 6(1)(e): "Behandlingen er nødvendig for utførelsen av en oppgave i offentlig interesse eller utøvelse av offentlig myndighet som den behandlingsansvarlige har fått tildelt."

b. I tillegg kan det være relevant å se på Artikkel 9(2)(g): "Behandlingen er nødvendig for å utføre en oppgave i allmennhetens interesse eller som ledd i utøvelsen av offentlig myndighet som den behandlingsansvarlige har fått tildelt."

2. Juridisk forpliktelse, offentlig interesse, eller vitale interesser:

a. Behandlingen av biometriske data for grensesikkerhetsformål kan være basert på offentlig interesse for å opprettholde sikkerheten ved grensene (Artikkel 6).

b. Det kan også være tilfeller der behandlingen er nødvendig for å oppfylle en juridisk forpliktelse som hviler på den behandlingsansvarlige (Artikkel 6).

c. Videre kan det være situasjoner der det er nødvendig for å beskytte vitale interesser til data subjektene eller andre (Artikkel 9).

3. GDPR-kompatible tidsplaner for datalagring av biometrisk informasjon ved grensen:

a. Lagringsperioden for biometrisk informasjon må være proporsjonal med formålet (Artikkel 5).

b. Det er viktig å utvikle klare retningslinjer som tar hensyn til behovet for grensesikkerhet og samtidig respekterer prinsippene om begrensning av lagringsperioden (Artikkel 5).

4. Protokoller for sikker og lovlig deling av biometrisk data på tvers av EUs grenser:

a. For deling av biometriske data over landegrenser innen EU, må man følge Artikkel 44-50 som handler om overføring av personopplysninger til tredjeland eller internasjonale organisasjoner.

b. Det kan være nødvendig med godkjente standardkontrakter (Artikkel 46), binding corporate rules (Artikkel 47), eller andre godkjente overføringsmekanismer.

5. Interoperabilitet med eksisterende systemer som SIS II og VIS, samt samsvar med GDPR:

a. For å sikre interoperabilitet, må systemet for biometriske data være utformet i tråd med prinsippene for innebygd personvern og personvern som standardinnstilling (Artikkel 25).

b. Samsvar med andre relevante regelverk, som SIS II og VIS, krever grundig evaluering av hvordan systemene kan integreres mens man opprettholder personvernet og sikkerheten (Artikkel 5 og Artikkel 32).

Question 4.

Råd angående rettsgrunnlaget under GDPR og norske helsedatalover:

Under GDPR og norske helsedatalover, er det viktig at sykehuset har et solid rettslig grunnlag for prosessering og deling av personlige helseopplysninger i nødsituasjoner. For slike situasjoner vil det mest relevante rettsgrunnlaget være artikkel 6(1)(d) i GDPR, som tillater behandling av data som er nødvendig for å beskytte en persons liv. Dette grunnlaget er spesifikt relevant i nødsituasjoner der det er en umiddelbar fare for pasientens helse eller liv.

I tillegg bør sykehuset vurdere norske helsedatalover, spesielt Helseregisterloven og Pasientjournalloven. Disse lovene gir bestemmelser om hvordan helseopplysninger kan deles og behandles, inkludert i nødsituasjoner. Sykehuset må sikre at dataene deles i samsvar med disse lovene, som ofte gir mer spesifikke krav enn GDPR.

Råd angående tillatt behandling uten pasientens samtykke:

Loven tillater behandling av personlige helseopplysninger uten pasientens samtykke når det er umiddelbar risiko for helse eller liv. I slike nødsituasjoner, som nevnt tidligere, kan GDPR artikkel 6(1)(d) og de norske helsedatalovene være relevante. Dette gjelder spesielt når pasienten ikke er i stand til å gi samtykke på grunn av sin tilstand.

Råd angående protokoller for overføring av helsedata i nødstilfeller:

Ved overføring av helsedata til medisinske fasiliteter utenfor Norge i nødstilfeller, må sykehuset følge nøye protokoller for å sikre at dataene overføres sikkert og i samsvar med loven. Dette inkluderer:

1. Bruk av sikre krypterte kanaler for overføring av data.

2. Sikre at mottakeren har tilstrekkelig sikkerhetsnivå og etterlever GDPR og relevante lovgivninger.

3. Dokumentere nødvendigheten av overføringen og de juridiske grunnlagene for det.

Råd angående begrensning av omfanget av delt data:

For å begrense omfanget av data som deles til det som er strengt nødvendig i en nødsituasjon, bør sykehuset:

1. Identifisere spesifikke typer helseopplysninger som er nødvendige for å håndtere nødsituasjonen.

2. Sørge for at bare disse spesifikke dataene deles, ikke all tilgjengelig helseinformasjon om pasienten.

3. Implementere tekniske løsninger for å begrense tilgangen til kun relevante data.

Råd angående mekanismer for autorisasjon og tilgangskontroll:

For å sikre at bare autoriserte personer kan få tilgang til ytterligere data og at videre behandling krever riktig autorisasjon, bør sykehuset:

1. Implementere robuste tilgangskontrollsystemer basert på prinsippet om minste privilegium.

2. Kreve tofaktorautentisering for å få tilgang til sensitive helseopplysninger.

3. Ha klare retningslinjer for autorisasjon av videre behandling av data, inkludert nødvendig godkjenning fra ansvarlige helsepersonell.

Dette sikrer at dataene behandles og deles på en lovlig og sikker måte i nødsituasjoner, samtidig som pasientens rettigheter og personvern ivaretas.

Appendix F

GDPR QA Dataset for Testing

1. {"qId": "gdprS001", "answers": ["Ja, hvis det er nødvendig for å oppfylle en kontrakt"], "qText": "Kan et selskap behandle personopplysninger uten samtykke i tilfeller der det er nødvendig for å oppfylle en kontrakt?"}
2. {"qId": "gdprS002", "answers": ["Nei, med mindre det er uttrykkelig samtykke fra personen"], "qText": "Er det lovlig for et markedsføringsfirma å dele kundedata med en tredjepart uten kundens samtykke?"}
3. {"qId": "gdprS003", "answers": ["Ja, hvis de er anonymisert og ikke lenger kan identifisere individet"], "qText": "Kan et forskningsselskap bruke personopplysninger samlet for et annet formål uten ytterligere samtykke?"}
4. {"qId": "gdprS004", "answers": ["Ja, men bare hvis det er viktig for jobbens natur"], "qText": "Er det lovlig for en arbeidsgiver å kreve sensitive personopplysninger, som helseinformasjon, fra sine ansatte?"}
5. {"qId": "gdprS005", "answers": ["Ja, hvis tjenesten er rettet direkte mot barn og samtykke er innhentet på riktig måte"], "qText": "Kan et selskap som tilbyr online-tjenester til barn behandle barnas data?"}
6. {"qId": "gdprS006", "answers": ["Ja, men de må informere den registrerte om bruddet og ta nødvendige tiltak"], "qText": "Hva må et selskap gjøre hvis de opplever et sikkerhetsbrudd som påvirker personopplysninger?"}
7. {"qId": "gdprS007", "answers": ["Ja, men bare hvis det er et legitimt formål og det er gjennomført en interesseavveining"], "qText": "Kan en organisasjon behandle personopplysninger basert på legitime interesser?"}
8. {"qId": "gdprS008", "answers": ["Ja, men selskapet må følge strenge regler for overføring av data utenfor EØS"], "qText": "Er det lovlig for et EU-basert selskap å overføre personopplysninger til et land utenfor EØS?"}
9. {"qId": "gdprS009", "answers": ["Ja, men de må sikre at de har et gyldig samtykke og at informasjonen er lett tilgjengelig"], "qText": "Kan et selskap bruke pre-ticked bokser for å innhente samtykke for markedsføringsformål?"}
10. {"qId": "gdprS010", "answers": ["Ja, men kun hvis de har et personvernombud og gjennomfører en DPIA"], "qText": "Kan et selskap bruke ny teknologi for å overvåke ansattes ytelse?"}
11. {"qId": "gdprS011", "answers": ["Ja, men de må sikre at prosessen er gjennomsiktig og at den registrerte har rett til å be om en manuell gjennomgang"], "qText": "Er det lovlig for en bank å bruke automatisert avgjørelsestaking for kredittvurdering?"}
12. {"qId": "gdprS012", "answers": ["Ja, men kun hvis det er et klart og uttrykkelig samtykke fra den registrerte"], "qText": "Kan et helsefirma behandle genetiske data for forskningsformål?"}
13. {"qId": "gdprS013", "answers": ["Ja, men bare hvis det er en nødvendig del av avtalen og kunden er tydelig informert om dette"], "qText": "Kan et teleselskap samle inn lokasjonsdata som en del av sin tjeneste?"}
14. {"qId": "gdprS014", "answers": ["En konsekvensvurdering for personvern (DPIA) kreves når typen behandling, spesielt ved bruk av ny teknologi, sannsynligvis vil medføre høy risiko for fysiske personers rettigheter og friheter. Dette inkluderer systematisk og omfattende profilering med betydelige effekter, storstilt behandling av spesielle kategorier av data, eller systematisk overvåking av offentlig tilgjengelige områder i stor skala."], "qText": "Når kreves en konsekvensvurdering for personvern under GDPR?"}
15. {"qId": "gdprS015", "answers": ["Under GDPR må et personvernombud (DPO) utnevnes når behandlingen utføres av en offentlig myndighet eller organ, når kjerneaktivitetene til behandlingsansvarlig eller databehandler består av operasjoner som krever regelmessig og systematisk overvåking av registrerte på stor skala, eller ved storstilt behandling av spesielle kategorier av data eller personopplysninger relatert til straffedommer og lovbrudd."], "qText": "Når er det obligatorisk å utnevne et personvernombud under GDPR?"}

16. {"qId": "gdprS016", "answers": ["Under GDPR har den registrerte rett til å be om at behandlingsansvarlig begrenser behandling hvis nøyaktigheten av personopplysningene bestrides, noe som gir behandlingsansvarlig tid til å verifisere nøyaktigheten. Den registrerte kan også be om begrensning i stedet for sletting hvis behandlingen er ulovlig, hvis dataene ikke lenger er nødvendige for behandlingsansvarlig, men kreves av den registrerte for å etablere juridiske krav, eller hvis den registrerte har protestert mot behandling i påvente av verifisering av behandlingsansvarliges legitime grunner."], "qText": "En klient ønsker å begrense behandlingen av sine personopplysninger av et selskap på grunn av unøyaktigheter i dataene. Hvilke rettigheter har de under GDPR?"}
17. {"qId": "gdprS017", "answers": ["GDPR gir den registrerte rett til å protestere mot behandling av sine personopplysninger, spesielt for direkte markedsføringsformål, inkludert profilering. Behandlingsansvarlig må opphøre behandlingen med mindre de kan demonstrere tvingende legitime grunner for behandlingen som overstyrer den registrertes interesser, rettigheter og friheter, eller for å etablere, utøve eller forsvare juridiske krav."], "qText": "En klient ønsker å protestere mot behandlingen av sine personopplysninger av et markedsføringsselskap. Hvordan støtter GDPR deres beslutning?"}
18. {"qId": "gdprS018", "answers": ["Ja, GDPR tillater begrensninger på forpliktelser og rettigheter fastsatt i forskriften, forutsatt at disse begrensningene er et nødvendig og forholdsmessig tiltak i et demokratisk samfunn. Slike begrensninger kan være for å beskytte nasjonal sikkerhet, offentlig sikkerhet, forebygging eller etterforskning av kriminelle aktiviteter, beskytte folkehelse, eller sikre uavhengigheten til rettsvesenet blant andre."], "qText": "Kan forpliktelser og rettigheter under GDPR begrenses under visse omstendigheter?"}
19. {"qId": "gdprS019", "answers": ["Behandlingsansvarlige er ansvarlige for å implementere passende tekniske og organisatoriske tiltak for å sikre og demonstrere at databehandling utføres i samsvar med GDPR. Disse tiltakene må gjennomgås og oppdateres ved behov."], "qText": "Hva er ansvarene til behandlingsansvarlige under GDPR?"}
20. {"qId": "gdprS020", "answers": ["GDPR krever at både behandlingsansvarlige og databehandlere implementerer passende tekniske og organisatoriske tiltak for å sikre et sikkerhetsnivå som er egnet med hensyn til risikoen. Dette inkluderer tiltak som pseudonymisering og kryptering av personopplysninger, sikring av pågående konfidensialitet, integritet, tilgjengelighet og robusthet i behandlingssystemer og -tjenester, og regelmessig testing og evaluering av effektiviteten av disse sikkerhetstiltakene."], "qText": "Hvilke sikkerhetstiltak må iverksettes for å beskytte personopplysninger under GDPR?"}
21. {"qId": "gdprS021", "answers": ["Ved brudd på personopplysningssikkerheten skal behandlingsansvarlig uten ugrunnet opphold og, om mulig, senest 72 timer etter å ha fått kjennskap til det, melde bruddet til den relevante tilsynsmyndigheten, med mindre bruddet sannsynligvis ikke vil medføre en risiko for fysiske personers rettigheter og friheter. Meldingen må beskrive arten av datainnbruddet, kategoriene og omtrentlig antall berørte individer, og de sannsynlige konsekvensene og tiltakene som er tatt for å håndtere bruddet."], "qText": "Hva er kravene for rapportering av et datainnbrudd under GDPR?"}

Appendix G

Timetables

Emil Farstad Bjerknes	
Tot. hours:	528,96

Day	Start-time	End-time	Description	Category	Hours
10.jan	10:00	10:30	Preperation for client meeting	Preparation	0,50
	10:30	11:33	Meeting with client	Meeting	1,05
	12:15	14:00	Seminar	Education	1,75
11.jan			Sick day		0,00
13.jan	12:00	17:00	Read up on Latex, and created a template for the project plan	Self study	5,00
14.jan	13:00	18:00	Read up and watched youtube video about LLMs (Llama 2) and RAG	Self study	5,00
15.jan	12:30	16:45	Preparation for the project plan,and for upcoming meetings	Work Session	4,25
16.jan	11:00	15:15	Worked on the project plan and created group contract	Work Session	4,25
17.jan	11:05	15:05	Project plan; scope and development models. Prepared client meeting	Work Session	4,00
18.jan	12:00	12:50	Supervisor meeting	Meeting	0,83
	12:50	16:45	Project plan.	Work Session	4,00
22.jan	09:00	10:30	Preparation for client meeting	Meeting	1,50
	10:30	11:15	Client meeting	Meeting	0,75
	11:00	13:00	Project plan; risk analysis	Work Session	2,00
23.jan	09:00	13:00	Project plan; risk analysis	Work Session	4,00
24.jan	11:00	19:00	Project plan; risk analysis, references, other research.	Self study	8,00
25.jan	09:00	13:30	Project plan; Methodology (Scrum), Gantt chart	Work Session	4,50
26.jan	15:00	16:30	Filled out standardagreement, sent to client	Self study	1,50
29.jan	09:00	15:00	Project plan	Work Session	6,00
	13:00	13:30	Supervisor meeting	Meeting	0,50
30.jan	09:00	15:00	Project plan	Work Session	6,00
31.jan	12:00	14:00	Sent standard agreement to supervisors. Final touches on project plan	Self study	2,00
	12:00	15:00	Research on huggingface	Self study	3,00
01.feb	09:00	13:00	Deliver project plan, made a 10 week plan	Work Session	4,00
02.feb	10:00	12:00	Finished 10 week plan	Work Session	2,00
	10:00	16:00	Research, how to use huggingface	Self study	6,00
04.feb	17:00	19:00	Research, LLM's and RAG (youtube)	Self study	2,00
05.feb	09:00	11:30	Sprint 1 work	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	15:00	Sprint 1 work	Work Session	3,00
	16:00	20:00	Tried to run LLM locally	Self study	4,00
06.feb	09:00	13:00	Research and Client Meeting	Meeting	4,00
08.feb	10:00	17:00	Research on running LLM (and how to do it with AMD GPU)	Self study	7,00
11.feb	12:00	17:00	Research on pytorch + rocm	Self study	5,00
12.feb	11:30	11:50	Client meeting	Meeting	0,33
	13:15	14:15	Group meeting on discord	Meeting	1,00
	15:00	16:30	Research on web scraping with python	Self study	1,50
13.feb	14:00	18:15	Sprint 1 work + setting up skyhigh VM	Work Session	4,25
14.feb	10:00	16:00	Research on running models with with huggingface and transformers	Self study	6,00
15.feb	09:00	13:00	Sprint 1 work. Trying to run llama2 through huggingface in our VM	Work Session	4,00
16.feb	13:00	16:00	Solved VPN SkyHiGh problem	Self study	3,00
19.feb	09:00	11:30	Sprint 2 work, preparation for supervisor meeting	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	16:00	Sprint 2 work	Work Session	4,00
20.feb	09:00	16:00	Sprint 2 work	Work Session	6,00
22.feb	09:00	16:00	Sprint 2 work	Work Session	6,00
25.feb	10:00	18:00	Research on running llama2 w/langchain	Self study	8,00
26.feb	09:00	11:30	Sprint 2 work, preparation for supervisor meeting	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	13:00	Sprint 2 work	Work Session	1,00
27.feb	09:00	13:00	Sprint 2 work, RAG research	Work Session	4,00
	13:00	17:00	Sprint 2 work, RAG, ChromaDB, Langchain. Testing Llama2 w/RAG	Self study	4,00
28.feb	10:00	18:00	Testing Llama2 with RAG (PDF)	Self study	8,00
29.feb	09:00	17:00	Sprint 2. RAG	Work Session	8,00
01.mar	12:00	17:30	Coding/running Llama2 with RAG (PDF). Testing with questions	Self study	5,50
04.mar	09:00	17:00	Sprint 3	Work Session	8,00
05.mar	12:00	15:00	Updated meeting minutes and the 10 week progress plan	Self study	3,00
	15:00	16:00	Course in Report-writing	Education	1,00
06.mar	11:00	19:00	Research and testing RAG with Ruter Llama2	Self study	8,00
07.mar	09:00	17:00	Sprint 3 work	Work Session	8,00
11.mar	09:00	11:30	Sprint 3 work	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Post meeting discussion, sprint 3 work	Work Session	5,00
12.mar	09:00	17:00	Sprint 3 work	Work Session	8,00
13.mar	10:00	18:00	Multi-query retriever research, sprint 3	Self study	8,00
14.mar	09:00	17:00	Sprint 3 work	Work Session	8,00
18.mar	09:00	10:30	Sprint 4 start + group meeting	Work Session	1,50
	10:30	11:00	Client meeting	Meeting	0,50
	11:00	11:30	Sprint 4 work	Work Session	0,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:30	16:30	Sprint 4 work	Work Session	4,00
20.mar	10:00	16:00	Sprint 4 work	Self study	6,00
26.mar	10:00	18:00	Sprint 4 work	Self study	8,00
05.apr	10:00	16:00	Sprint 5, report writing	Work Session	6,00
07.apr	10:00	17:00	Report writing	Self study	7,00

08.apr	09:00	11:30	Sprint 5 work	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Report writing/research	Self study	5,00
09.apr	10:00	17:00	Report writing	Self study	7,00
11.apr	10:00	17:00	Report writing	Self study	7,00
14.apr	10:00	17:00	Report writing	Self study	7,00
15.apr	09:00	11:30	Sprint 6, preparation for evaluation meeting	Work Session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Meeting with GDPR expert, sprint 6 work	Work Session	5,00
16.apr	09:00	16:00	Sprint 6 work, client meeting, evaluation planning	Work Session	7,00
17.apr	11:00	16:00	Sprint 6, writing	Work Session	5,00
	16:00	18:00	Research, writing report	Self study	2,00
18.apr	11:00	19:00	Report writing	Self study	8,00
19.apr	12:00	17:00	Planning evaluation and writing	Work Session	5,00
20.apr	10:00	17:00	Report writing	Self study	7,00
22.apr	10:30	13:00	Planning meeting and evaluation of program	Meeting	2,50
23.apr	10:00	17:00	Finishing evluation form, writing report	Self study	7,00
24.apr	10:00	18:00	Report writing, sprint 6	Self study	8,00
25.apr	10:00	18:30	Report writing, sprint 6	Self study	8,50
26.apr	10:00	18:00	Report writing, sprint 6	Self study	8,00
27.apr	10:00	18:00	Report writing, sprint 6	Self study	8,00
29.apr	11:00	18:00	Sprint 7: Report writing	Self study	7,00
30.apr	09:00	17:00	Report writing	Work Session	8,00
01.mai	11:00	17:00	Report writing	Self study	6,00
02.mai	10:00	18:00	Report writing	Self study	8,00
03.mai	12:30	18:30	Report writing and group discussion	Work Session	6,00
04.mai	11:00	18:00	Report writing	Self study	7,00
06.mai	09:00	17:00	Writing and meetings	Work Session	8,00
	12:00	19:00	Report writing	Self study	7,00
07.mai	10:00	18:00	Report writing	Self study	8,00
08.mai	12:00	19:00	Report writing, supervisor feedback ch. 1-3 + 6	Self study	7,00
09.mai	11:00	19:00	Report writing	Self study	8,00
10.mai	11:00	17:00	Report writing	Work Session	6,00
11.mai	12:00	20:00	Report writing - complete draft	Work Session	8,00
13.mai	09:00	14:00	Report writing and supervisor meeing	Work Session	5,00
14.mai	12:00	19:00	Report writing - feedback from supervisor	Self study	7,00
15.mai	10:00	18:00	Report writing, feedback on evaluation chapter	Self study	8,00
16.mai	12:00	18:00	Report writing	Work Session	6,00
17.mai	14:00	18:00	Report writing, feedback on ch.6 and 7	Self study	4,00
18.mai	11:00	19:00	Final touches and delivery	Work Session	7,00

Paul Røkke Bjørneng

Tot. hours: 520,10

Day	Start-time	End-time	Description	Category	Hours
10.jan	10:00	10:30	Preperation for client meeting	Preparation	0,50
	10:30	11:33	Meeting with client	Meeting	1,05
	12:15	14:00	Seminar	Education	1,75
11.jan	09:30	09:57	Preperation for supervisor meeting	Preparation	0,45
	09:57	10:30	Meeting with supervisors	Meeting	0,55
	11:30	13:00	Writing of meeting-minutes	Administrative	1,50
	13:10	15:00	Read up on Overleaf/latex and copied the minutes to a shared latex template	Self-study	1,83
	15:00	16:08	Made to-do list for upcoming week	Administrative	1,13
	16:20	16:45	Improved excel timesheet	Administrative	0,42
15.jan	09:45	10:12	Read up on biblatex and Zotero	Self-study	0,45
	12:25	16:43	Preparation for the project plan,and for upcoming meetings	Work session	4,30
16.jan	09:29	11:47	Read up on RAG, LangChain, LLMChain	Self-study	2,30
	12:00	16:15	Writing of project plan/group contract	Work session	4,25
	17:00	18:30	Read up on LangChain, NLP	Self-study	1,50
17.jan	12:00	15:05	Writing of project plan	Work session	3,08
18.jan	08:00	11:03	Read up on types of AI, huggingface, rettskilder	Self-study	3,05
	11:30	11:46	Read up on huggingface	Self-study	0,27
	12:00	12:50	Meeting with supervisors	Meeting	0,83
	13:50	20:10	Writing of project plan	Work session	6,33
22.jan	08:00	10:29	Preperation for client meeting, GDPR, personal data	Preparation	2,48
	10:29	11:11	Meeting with client	Meeting	0,70
	11:11	13:00	Writing of project plan	Work session	1,82
	13:10	16:16	Read up on transformer models, ML, LoRA/QLoRA, pytorch/tensorflow	Self-study	3,10
23.jan	07:12	11:15	Read up on NLP, ML/DL, vectors(databases), weights(NLP), transformers	Self-study	4,05
	11:20	14:30	Writing of project plan	Work session	3,17
	15:30	18:30	Read up on Scrum, DSR, word embeddings	Self-study	3,00
25.jan	08:00	09:11	Read up on Scrum	Self-study	1,18
	09:11	14:00	Writing of project plan	Work session	4,82
26.jan	07:22	07:42	Read up on ethical consideration AI	Self-study	0,33
	11:19	12:48	Read up on encoder/decoder, tokens	Self-study	1,48
29.jan	08:54	12:45	Writing of project plan	Work session	3,85
	13:00	13:30	Meeting with supervisors	Meeting	0,50
30.jan	09:00	17:30	Writing of project plan	Work session	8,50
01.feb	08:53	13:01	Planning for the next 10 weeks	Administrative	4,13
02.feb	06:45	07:45	Read up on neural machine translation (arxiv paper)	Self-study	1,00
	08:54	10:15	Read up on neural machine translation (arxiv paper), gathered sources	Self-study	1,35
	10:15	11:00	Planning for the next 10 weeks	Administrative	0,75
05.feb	06:00	07:30	Read up on the transformer model architecture(arxiv paper)	Self-study	1,50
	09:00	11:30	Sprint 1 work	Work session	2,50
	11:30	12:00	Meeting with supervisors	Meeting	0,50
	12:00	17:30	Sprint 1 work	Work session	5,50
06.feb	09:00	10:30	Planning sprint, start issueboard	Work session	1,50
	10:30	11:00	Meeting with client	Meeting	0,50
	11:00	17:00	Sprint 1 work	Work session	6,00
08.feb	10:00	18:00	Research, running LLMs, frameworks	Self-study	8,00
11.feb	11:15	19:20	Research, planning for meeting	Self-study	8,08
12.feb	13:15	14:15	Group meeting	Meeting	1,00
13.feb	05:00	09:09	Research, gathered sources, dataset/RAG	Self-study	4,15
	14:00	18:15	Sprint 1 work	Work session	4,25
14.feb	04:15	10:00	Research, RAG, document pre-processing	Self-study	5,75
	10:55	13:47	Research	Self-study	2,87
15.feb	04:00	08:15	Research, made dataset	Self-study	4,25
	09:00	13:00	Sprint 1 work	Work session	4,00
16.feb	08:00	16:20	Sprint 1 work	Self-study	8,33
19.feb	09:00	11:30	Sprint 2 work	Work session	2,50
	11:30	12:00	Meeting with supervisors	Meeting	0,50
	12:00	13:00	Sprint 2 work	Work session	1,00
20.feb	06:20	07:20	Research	Self-study	1,00
	09:00	15:00	Sprint 2 work, dataset	Work session	6,00
	16:00	17:00	Research: llm-embedder, langchain	Self-study	1,00
22.feb	08:00	09:00	Research: langchain, chunking/textsplitting	Self-study	1,00
	09:00	16:00	Sprint 2 work, writing of thesis	Work session	7,00
24.feb	08:30	12:45	Writing of thesis	Self-study	4,25
	14:00	18:30	Writing of thesis	Self-study	4,50
26.feb	08:00	09:00	Preparation for supervisor meeting, planning the week ahead	Administrative	1,00
	09:00	11:30	Sprint 2 work	Work session	2,50
	11:30	12:00	Meeting with supervisors	Meeting	0,50
	12:00	13:00	Sprint 2 work	Work session	1,00
27.feb	08:00	09:00	Sprint 2 work, Evaluation QA	Self-study	1,00
	09:00	13:19	Sprint 2 work	Work session	4,32
28.feb	08:00	13:24	Writing of thesis	Self-study	5,40
29.feb	09:00	17:10	Writing of thesis	Self-study	8,17
04.mar	09:00	17:00	Writing of thesis	Self-study	8,00
07.mar	08:25	13:05	Writing of thesis	Self-study	4,67

11.mar	08:55	11:30	Sprint 3 work	Work session	2,58
	11:30	12:00	Meeting with supervisors	Meeting	0,50
	12:00	13:05	Writing of thesis	Work session	1,08
12.mar	08:59	17:05	Writing of thesis	Work session	8,10
13.mar	08:00	16:20	Writing of thesis, sprint 3	Work session	8,33
14.mar	08:55	17:25	Writing of thesis	Work session	8,50
17.mar	09:05	13:30	Writing of thesis	Work session	4,42
20.mar	09:00	17:00	Research, Writing of thesis	Self-study	8,00
22.mar	09:30	17:30	Research, Writing of thesis	Self-study	8,00
25.mar	10:00	18:00	Sprint 4 work	Self-study	8,00
27.mar	12:00	20:15	Writing of thesis	Self-study	8,25
29.mar	10:15	14:20	Writing of thesis	Self-study	4,08
01.apr	11:30	16:40	Research, Writing of thesis	Self-study	5,17
05.apr	09:00	16:00	Writing of thesis	Work session	7,00
06.apr	07:35	15:40	Writing of thesis	Self-study	8,08
08.apr	09:00	17:00	Writing of thesis	Work session	8,00
12.apr	09:00	13:00	Writing of thesis	Self-study	4,00
15.apr	09:00	17:20	Meetings, planning, sprint 6	Work session	8,33
16.apr	09:00	14:30	Evaluation planning, writing	Work session	5,50
17.apr	08:00	11:00	Writing of thesis	Self-study	3,00
	12:00	18:00	Writing of thesis	Self-study	6,00
19.apr	09:50	12:30	Writing of thesis	Self-study	2,67
	13:55	19:30	Planning, writing	Work session	5,58
20.apr	10:00	18:00	Research, Writing of thesis	Self-study	8,00
22.apr	10:45	19:00	Research, Writing of thesis	Work session	8,25
23.apr	10:50	12:15	Research	Self-study	1,42
24.apr	07:15	11:30	Research, Writing of thesis	Self-study	4,25
25.apr	08:45	13:00	Writing of thesis	Self-study	4,25
26.apr	07:20	11:30	Writing of thesis	Self-study	4,17
27.apr	08:20	11:40	Writing of thesis	Self-study	3,33
28.apr	07:00	11:20	Writing of thesis	Self-study	4,33
30.apr	09:00	17:00	Writing of thesis	Work session	8,00
01.mai	07:30	12:00	Research, Writing of thesis	Self-study	4,50
02.mai	08:55	13:00	Writing of thesis	Self-study	4,08
	14:00	19:00	Writing of thesis	Self-study	5,00
03.mai	07:45	14:00	Writing of thesis	Self-study	6,25
	14:36	15:46	Meeting	Work session	1,17
04.mai	09:00	17:20	Writing of thesis	Work session	8,33
06.mai	07:40	08:40	Writing of thesis	Work session	1,00
	09:00	15:15	Writing, planning, meetings	Work session	6,25
	16:00	17:18	Writing of thesis	Self-study	1,30
07.mai	08:30	13:00	Writing of thesis	Self-study	4,50
	14:30	19:00	Research, Writing of thesis	Self-study	4,50
08.mai	08:40	12:40	Writing of thesis	Self-study	4,00
	12:50	17:00	Writing of thesis	Self-study	4,17
09.mai	09:40	17:30	Writing of thesis	Self-study	7,83
10.mai	07:20	09:45	Writing of thesis	Self-study	2,42
	11:00	19:10	Writing of thesis	Work session	8,17
11.mai	12:00	20:15	Writing of thesis	Work session	8,25
13.mai	08:55	15:05	Writing of thesis	Work session	6,17
14.mai	09:50	13:45	Writing of thesis	Self-study	3,92
15.mai	08:40	10:20	Writing of thesis	Self-study	1,67
	11:00	17:00	Writing of thesis	Work session	6,00
16.mai	12:00	18:00	Writing of thesis	Work session	6,00
17.mai	10:00	18:00	Writing of thesis, restructuring	Work session	8,00
18.mai	11:00	19:00	Writing of thesis, restructuring, delivery	Work session	8,00

Torje Grimstad Bliksvær	
Tot. Hours:	518,17

Day	Start-Time	End-Time	Description	Category	Hours
10.jan	10:00	10:30	Perperation for client meeting	Work session	0,50
	10:30	11:33	Meeting with client	Meeting	1,05
	12:15	14:00	Seminar	Education	1,75
11/01	09:57	10:30	Meeting with supervisor	Meeting	0,55
	11:00	16:00	Initialization of this document and meeting notes document	Work session	5,00
15/01	12:38	17:30	Prep for the project plan and prep for upcoming meetings	Work session	4,87
16/01	11:05	15:05	Project plan, group contract	Work session	4,00
17/01	10:59	19:00	Scope and planning, self study	Work session	8,02
18/01	11:00	12:00	Prep for supervisor meeting	Work session	1,00
	12:00	12:50	Supervisor meeting	Meeting	0,83
	12:50	18:00	Work on the project plan	Work session	5,17
19/01	10:00	17:30	Work on the project plan	Self-study	7,50
22/01	09:00	10:29	Work on the project plan	Work session	1,48
	10:29	11:11	Client meeting	Meeting	0,70
	11:11	18:00	Work on the project plan	Work session	6,82
23/01	09:05	16:00	Work on the project plan	Work session	6,92
25/01	09:06	14:30	Work on the project plan	Work session	5,40
29/01	09:04	16:30	Work on the project plan	Work session	7,43
	12:58	13:30	Supervisor meeting	Meeting	0,53
30/01	09:04	13:06	Work on the project plan	Work session	4,03
01/02	09:00	13:30	Work on the project plan	Work session	4,50
02/02	10:00	18:00	Created 10 week plan, research	Work session	8,00
05/02	09:00	11:30	Sprint 1	Work session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	15:00	Sprint 1 work	Work session	3,00
	16:00	18:30	Research ML on AMD GPU solution	Work session	2,50
06/02	09:12	10:30	Research	Work session	1,30
	10:30	11:00	Client meeting	Meeting	0,50
	11:00	18:30	Sprint 1 work	Work session	7,50
08/02	11:00	19:00	Research ML on AMD GPU solution	Self-study	8,00
12/03	11:30	11:52	Supervisor meeting	Meeting	0,37
	13:15	14:16	Discord group meeting	Meeting	1,02
13/02	11:00	19:00	Setup SkyHigh VM	Work session	8,00
15/02	09:15	17:00	Sprint 1 work, run models in VM	Work session	7,75
19/02	09:00	17:00	Sprint 2	Work session	8,00
20/02	09:11	17:00	Sprint 2 work	Work session	7,82
22/02	09:02	17:00	Sprint 2 work	Work session	7,97
26/02	09:00	11:30	Sprint 2 work	Work session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Sprint 2 work, discussion	Work session	5,00
27/02	09:09	17:00	Rebuild VM with CPU focus and Volume Boot drive. Read up on LangChain.	Work session	7,85
29/02	09:08	17:00	Sprint 2, RAG research	Work session	7,87
04/03	08:56	17:00	Sprint 3	Work session	8,07
07/03	09:04	17:10	Sprint 3 work	Work session	8,10
11/03	09:04	11:32	Sprint 3 work	Work session	2,47
	11:32	12:01	Supervisor meeting	Meeting	0,48
	12:01	18:00	Sprint 3 work, discussion	Work session	5,98
12/03	09:10	17:20	Sprint 3 work	Work session	8,17
13/03	09:06	18:00	Sprint 3, Multi-Query Ret. research	Work session	8,90
18/03	09:04	10:30	Sprint 4 + group meeting	Work session	1,43
	10:30	11:00	Client meeting	Meeting	0,50
	11:00	11:30	Sprint 4	Work session	0,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:30	Sprint 4 work	Work session	5,50
04/04	09:00	17:00	Work on glossary and acronym list + report	Self-study	8,00
05/04	10:00	17:00	Sprint 5	Work session	7,00
06/04	13:00	18:00	Sprint 5	Work session	5,00
07/04	15:30	23:38	HyDE	Self-study	8,13
08/04	09:30	11:30	Sprint 5	Work session	2,00
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	18:00	Sprint 5	Work session	6,00
09/04	10:00	18:00	Sprint 5 - coding	Self-study	8,00
10/04	11:32	19:11	Program testing	Self-study	7,65
12/04	10:00	18:06	Program testing	Self-study	8,10
15/04	08:58	11:30	Sprint 6	Work session	2,53
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Sprint 6 + evaluation meeting	Work session	5,00
16/04	09:02	12:30	Sprint 6	Work session	3,47
	12:30	13:02	Client meeting	Meeting	0,53
17/04	11:00	17:30	Sprint 6	Work session	6,50
18/04	12:00	20:00	Report writing	Self-study	8,00
19/04	12:00	17:00	Program development, report writing	Work session	5,00
20/04	11:00	13:35	Program testing, running	Self-study	2,58
	15:00	22:30	Program testing, running	Self-study	7,50
21/04	16:00	22:14	Program development, running	Self-study	6,23
22/04	10:00	17:30	Program testing, running	Self-study	7,50

23/04	11:05	19:00	Program testing, report writing	Self-study	7,92
24/04	10:54	18:30	Report writing	Self-study	7,60
	20:32	21:04	Report writing	Self-study	0,53
25/04	11:00	18:30	Program development	Self-study	7,50
	19:23	21:38	Report writing	Self-study	2,25
26/04	09:01	17:00	Report writing	Self-study	7,98
27/04	09:10	18:32	Report writing	Self-study	9,37
29/04	10:05	17:34	Program development, report writing	Self-study	7,48
30/04	09:00	14:13	Report writing	Work session	5,22
01/05	11:00	18:00	Report writing	Self-study	7,00
02/05	10:00	18:00	Report writing	Self-study	8,00
03/05	11:00	18:00	Report writing, group meeting	Work session	7,00
04/05	10:00	18:00	Report writing	Self-study	8,00
06/05	09:00	11:30	Report writing	Work session	2,50
	11:30	12:00	Supervisor meeting	Meeting	0,50
	12:00	17:00	Report writing	Work session	8,00
07/05	10:00	17:30	Report writing	Self-study	7,50
08/05	12:00	19:00	Report writing	Self-study	7,00
09/05	11:00	19:00	Report writing	Self-study	8,00
10/05	11:00	17:00	Report writing	Work session	6,00
11/05	12:00	20:00	Report writing - complete draft	Work session	8,00
13/05	09:00	16:00	Supervisor meeting + report writing	Work session	7,00
14/05	12:00	19:00	Report writing - fixing feedback	Self-study	7,00
15/05	11:00	19:00	Report writing, code commentary	Self-study	7,00
16/05	12:00	18:00	Report writing	Work session	6,00
17/05	13:00	20:00	Report writing	Self-study	7,00
18/05	11:00	19:00	Final touches and delivery of thesis	Work session	8,00

Appendix H

Meeting Minutes

Meeting Minutes

Emil Farstad Bjerknes, Paul Røkke Bjørneng, Torje Grimstad Bliksvær

January 2024

1 Introduction

This document includes notes from all of the meetings throughout the project, including client meetings, supervisor meetings, and internal group meetings.

2 Client meetings

2.1 10/01 - Client meeting 1

10:30 - 11:33
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.
Not present: n/a.

- The meeting was conducted to get to know the client and his company.
- We discussed briefly what his idea behind the proposed thesis was.
- We got an initial overview of different aspects of the proposed thesis.
    - For instance, how we could validate our fine-tuned model's answers with answers from ChatGPT when that time comes.
- The client presented us with potential resources we can use (e.g. huggingface.co etc.) and other relevant sources/ideas we should read up on.
- We discussed what programming language we should use.
    - Python is the most used and has the most resources online.

2.2 22/01 - Client meeting 2

10:30 - 11:15
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.
Not present: n/a.

- Discussed the scope. Asked about GDPR, was ok to add this and train the model on this.
- Could be interesting to test the finished model on language, English vs Norwegian.
- We should make predictions of what we think will happen and explain why we think this.
- Other notes from Daniel
  - LORA and QLORA. (Low Rank Adaptation).
  - Balance how many weights that needs to be fine-tuned.
  - The model might be to big based on dataset. Chose a small model?
  - Specific benchmarks?
  - First, run LLM from huggingface locally to test
  - The best is to use both RAG and fine-tuning to train the model
- "Ruter" has models on huggingface.

2.3 06/02 - Client meeting 3

10:30 - 11:00
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.

Discussed methods and thoughts when it comes to running the models, and also RAG and the dataset:

- Daniel used Chat-GPT, "pay as you go". Has not ran anything locally.
- We could try to do it privately in huggingface?
- Look into Langsmith. would have started with this.
- Important to have a good quality dataset, not necessarily a big one.

2.4 04/03 - Client meeting 4

10:30 - 11:00
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.

Discussed and presented our sprint progress.

- That we have successfully ran Llama2 13B with RAG using our GDPR dataset. We received some good answers, but did not consistently receive answers in Norwegian.
- We are shifting our focus to mainly focusing on RAG when it comes to fine-tuning tools.
- Furthermore, we discussed how to receive answers in Norwegian.

Thoughts and tips from the client.

- We can try the Ruter Llama2 model from Huggingface, which is fine-tuned on norwegian datasets.
- Look into Bineric AI, a Norwegian LLM "NorskGPT"
- LM Studio: Desktop app for downloading and running LLM's

2.5 18/03 - Client meeting 5

10:30 - 11:00
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.

Presented our progress from the previous sprint. Discussed our plan going forward.

- Talked about HyDE and re-ranking and how we want to implement these methods.
- Discussed how we use the dataset (PDF-loader) and other methods that potentially can be used.
- Gave us some tips for the report writing. In the discussion part we could write what we would do if we were to further develop the model, and with what techniques (that we did not already use in this current model).

2.6 16/04 - Client meeting 6

12:30 - 13:00
Attendees: Emil Bjerknes, Paul Bjørneng, Torje Bliksvær, Daniel Arlien.

Discussed our progress and that we are more or less finished with the technical aspect and that we are shifting focus towards the writing of the thesis. We discussed the evaluation/results of our program when it comes to language.

- About the legal language, how precise is it when it comes to what words to use, compared to ”normal” language. Does the AI know the difference?
- Discussed the difference and importance of legal language, and examples/use cases in legal scenarios.
- This should be mentioned in the thesis.

3 Supervisor meetings

3.1 11/01 - Supervisor meeting 1

09:57 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Shao-Fang Wen, Muhammad Yamin
Not present: Emil Bjerknes

- Priority number 1 is to discuss what the final issue for our thesis should be with the client.
  - We need to scope our thesis. Especially in regards to the specific types of legal questions and laws we want the LLM to assist with, and base its answers upon.
  - We need to find out what information/data our group should gather, and what the client can provide (if anything).
  - We must find out about the client's background (i.e. what his/the company's reasons are for proposing this thesis etc.).
- We must get familiar with and study the ideas related to our thesis.
  - Such as reading up on potential resources we can use (e.g. huggingface.co)
- We discussed when the weekly meeting should happen each week.
  - Decided on every Thursday at 12:00 – 12:30.
  - Can be de-scheduled if a meeting is not necessary that week.
- Decided on a template we should prepare and follow for each weekly meeting
  1. What we have done the past week (5min).
  2. Issues that arise, and how the supervisors can help.
  3. Tasks for upcoming week.
  4. Other.
- We discussed how we must divide our group into different roles, so that we will efficiently conquer upcoming tasks without overlapping.
  - Also that we should always document who has done what task, and at what time.
- We were reminded that we should always take minutes from each meeting, whether that be between ourselves (the group) or with the supervisors/client.
- We must get familiar with Overleaf/Latex, and get templates started for different documents that will be shared.

3.2 18/01 - Supervisor meeting 2

12:00 - 12:50
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes, Shao-Fang Wen

- New weekly meeting time on Mondays at 13:00.
- Discussed the security concept in this assignment.
  - We should try to focus on security. Maybe train it on GDPR as well. Privacy. Less law, more security. Since this is relevant for our course.
  - For example: How can i protect the client's information/data/privacy?
  - Ask the client what he wants from the law, and ask him if we can make an assistant that knows GDPR. Maybe have both law and GDPR? (Steven will find a good source for us on GDPR if we end up doing this).
- We need to choose a development method, Scrum is common. But when writing it should be more scientific, and to pursue problems in a scientific way. Explain WHY we use something, instead of something else. This goes for everything.
- The assignment is application based, not a proof of concept.
- Risk analysis should focus on the technical things, as well as the client. Not about risks regarding the supervisor.

3.3 29/01 - Supervisor meeting 3

13:00 - 13:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes, Shao-Fang Wen, Muhammad Mudassar Yamin

- Models / comparison
  - We need to write how we do the comparison and how we test/compare the different models. Pick for example three different models with different characteristics.
  - We need to understand the steps to be able to develop a model.
  - Practice how to use (small) models in huggingface. Try it out on LLAMA 2.
- Make a 10 week (Feb - March) plan for the tasks and the goals we are going to do moving forward. Make a small "presentation" on our plan. Detailed step by step.

- Ethical considerations/constraints.
  - We can mention it in the project, but since we are doing this for educational purposes we don't have to address it too seriously. we do not have to mention it in the project plan.

3.4 05/02 - Supervisor meeting 4

11:30 - 12:00
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Farstad Bjerknes, Shao-Fang Wen, Muhammad Yamin

Started the meeting by showing them our 10-week plan. Feedback and discussion after the presentation:

- Write specifically our goals for each week/sprint in the presentation. Share it with the supervisors.
- Do not use LoRA.
- Could be a good idea to create a "live demo" version for running the models (for the presentation).
- Important to base our work on research papers. Use good argumentation when choosing what we do, what models/techniques we use, and why.
- "pip3 install ollama".
- We will go through whether or not we reached our goals for each week/sprint. Let them know what we learned and what we did this week.

- Take notes of what sources we use. We can share this with the supervisors.

3.5 12/02 - Supervisor meeting 5

11:30 - 11:50
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Shao-Fang Wen, Muhammad Yamin
Not present: Paul Bjørneng

Discussion about last week, and plan for this week:

- We could utilize sky high. Contact Lars Erik / fill out form for access. Try to get at least 2 instances?
- We should use RAG.
- Look into web scraping, python libraries. This to collect massive amount of data.
- We should divide the tasks this week. Paul will do data gathering (dataset, GDPR), Torje will continue to try to run LLMs, Emil will look into web scraping with python.
- Remember to update the plan for the weeks, and explain in more detail what we have done each week. What we accomplished.

3.6 19/02 - Supervisor meeting 6

11:30 - 11:50
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin
Not Present: Shao-Fang Wen

Started by presenting what we achieved last week/sprint, and if we reached our goals. We also presented the plan for this week, and that we had to make some adjustments.

Pointers and discussion with the supervisor:

- We do not specifically have to write a "problem" we are going to solve, but we should write a specific theme in a paragraph at the start of the thesis about what the report and assignment is all about.
- For dataset we should look into NLTK (Natural Language Toolkit).
- We should look into TAG.
- Try ollama for running models as an alternative.

3.7 26/02 - Supervisor meeting 7

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin, Shao-Fang Wen

Discussed our progress the past week. We have used langchain to run Llama2. Mudassar has also tried langchain, and used RAG with success. He used a PDF-loader in this case, with GDPR dataset. We will try something similar this week. He also used ChromaDB.

Other discussion points:

- We should "divide and conquer" the tasks. Some do research and write on the report, and others do "hands on" and coding work.
- About the research papers we use, we can mention the research papers we base our thesis on early on in the report, after we write the project background and introduce the problem.
- We can use ChatGPT to create the test/evaluation questions for the model testing.
- We should ask for a CPU heavy SkyHiGh flavour, instead of GPU as we do not need GPU for our use.

3.8 11/03 - Supervisor meeting 8

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Shao-Fang Wen
Not present: Muhammad Yamin

Started the meeting by presenting our progress, and our goals for the week/sprint. Then, we showed a quick demo of the model we have created and the answers we get from the model.
Other talking points and suggestions:

- We are getting ok answers from the model when using RAG, but it can definitely be better. We should explore further methods to improve the RAG with the model.
- We are also been writing quite a bit on the report.
- The most important part is to create a good report, not a perfect model/product. We should document everything that we do, and what we try to do, and explain why we do it.

3.9 18/03 - Supervisor meeting 9

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin
Not present: Shao-Fang Wen

Presented out progress from last week and our plan for this week. Discussed the evaluation part of the program:

- Create a proper evaluation method and maybe ask the customer to create questions. The evaluation should always be done from an external part (e.g. the client). Is the information correct or false? Do both our own evaluation, and as well the clients evaluation.
- Shift focus to evaluation of the model when the technical aspect is finished.

3.10 08/04 - Supervisor meeting 10

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin, Shao-Fang Wen

Presented out progress from last week and the Easter break, and our plan for this week. We will now have finished our program, and start evaluating it. Comments from the supervisors:

- Instead of asking the client to make evaluation questions, we should contact a GDPR expert. Steven and Muhammad will get us in contact with one from NTNU.
- We now have 4 weeks until the first draft of the report should be finished. We should add the supervisors to the Overleaf project as well so they can supervise the report.
- We should improve the code a little bit by making it more ”our own”.

3.11 15/04 - Supervisor meeting 11

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin, Shao-Fang Wen

Presented our plan for the week, as well as what we were going to present for the GDPR expert. We also presented our evaluation plan. Other talking points:

- For the introduction, the ”problem” is very important and needs to be exact and to 100 percent connect with the context and the thesis.
- Back up every decision we do. Explain everything and why we do something, and also why we do not do something.
- Our audience is legal people, define what legal people will use this problem for! What are the requirements for legal people? Target audience, who will use this solution? Maybe change our target audience to system devs?, or stay with lawyers.
- Define legal language.
- Be sure to thoroughly explain to the GDPR expert what we need from him and who the target audience is.

3.12 22/04 - Supervisor meeting 12

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Muhammad Yamin, Shao-Fang Wen

Meeting was mainly focused on the program answers based on the scenarios provided by the expert, and the evaluation.

- How can you explain the results, especially the incomplete questions. We should be appending the questions/answers in the appendix. we can discuss issues and technical difficulties if we want and if we find it worth discussing.

- crewAI, maybe we can implement this to not include ”garbage” and fix our incomplete answers.

- Evaluation form the evaluator - we will have a scoring system.

3.13 13/05 - Supervisor meeting 13

11:30 - 12:00
Attendees: Torje Bliksvær, Emil Farstad Bjerknes, Paul Bjørneng, Shao-Fang Wen

Last supervisor meeting. Discussed the draft of the thesis and its structure and content. Tips and feedback from the supervisors:

- Introduction: Problem area - first paragraph needs to be more concrete. Use bullet-points to explain what are the problems, and try to connect them to GDPR. Alsom how can we help Futurize and connect this to GDPR.

- Theory: Include only sections and information that are used later in the thesis.

- Methodology: New section: "Agile Software Development", instead of "Scrum Methodology".

- In Development Environment: Talk about infrastructure, deployment, the environment we created the program in.

- The "Technologies" are system dependent and cannot be "replaced".

- We have to mention tables or figures used in the text, and the meaning of them (e.g. a paragraph where we mention the table or figure and what they mean).

4 Group meetings

4.1 15/01 - Group meeting 1

12:30 - 16:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

- Did a quick recap of the supervisor meeting since Emil did not attend it last week.
- Started filling out the "standardavtale mellom bedrift og studenter".
- Wrote down and sent questions to Daniel regarding the different points in the project plan.
- Discussed our problem and scope, as well as group rules and roles.

4.2 06/02 - Group meeting 2

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint planning for sprint 1, preparation for meeting with client.

- Discussed what is to be done in the sprint.
- Created Backlog and issues "to do" in Jira. Sprint goal. Delegated issues.
- Prepared for meeting with client.

4.3 19/02 - Group meeting 3

09:00 - 11:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes
Goal for meeting: Sprint retrospective, and sprint planning for sprint 2

- Also discussed what we accomplished and if we reached our goals or not.
- Sprint planning for sprint 2. Created goals, as well as a backlog and issues in Jira. Divided and delegated the work among the team members.
- Prepared for supervisor-meeting.

4.4 04/03 - Group meeting 4

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint retrospective of sprint 2, and sprint planning for sprint 3. Prepare for client meeting.

- We discussed our progress. In this sprint we managed to create a demo of Llama2 running with RAG containing a PDF of the GPDR document. We also tested it with some queries. The only challenge is that the model does not answer in Norwegian consistently. It also relies heavily on finding keywords that match the query, resulting in that it often cannot answer the question.
- We discussed what needs to be done, set our goals, and planned sprint 3 in Jira where we created issues for the team to work on.

4.5 18/03 - Group meeting 5

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint retrospective of sprint 3, and sprint planning for sprint 4 in Jira. Prepare for client meeting, and to present progress for supervisors.

- We have managed to successfully run Ruter's and Bineric's version of the 13B Llama2 model with RAG, including multi-query retrieval. Moving forward we will look into HyDE and Re-ranking, as well as utilizing prompt engineering in the best way possible.
- We also discussed how to make the evaluation/testing methods better. We will use a scoring/point system.
- The writing of the report will also be done continuously throughout the sprint. Paul will mainly work on the report, and Emil and Torje will be working mainly on the model and the different RAG methods.

4.6 01/04 - Group meeting 6

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint retrospective of sprint 4 and planned sprint 5. Overview and planning in Jira.

- We want to finish up the technical aspect by the end of this sprint, so we can focus on writing the thesis next sprint.
- The next sprint will also hopefully include an evaluation of the program.

4.7 15/04 - Group meeting 7

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint retrospective of sprint 5, plan sprint 6. Also prepare for meeting with GDPR expert regarding the evaluation of the AI assistant.

- Be ready to present our project, and what we need from him regarding the evaluation

- We want him to provide questions or scenarios for our program to answer, and then evaluate these answers based on GDPR understanding and strategies.

4.8 29/04 - Group meeting 8

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Goal for meeting: Sprint retrospective of sprint 6, planning the last sprint, which focuses mainly on the writing the thesis and finishing it up.

- Discussed the evaluation, and moved our focus to the writing. Delegated working areas in the thesis based on motivation and experience.
- Set some deadlines for things to be finished, and discussed first draft delivery. The supervisors wants a first draft 2 weeks before the deadline.

4.9 13/05 - Group meeting 9

09:00 - 10:30
Attendees: Paul Bjørneng, Torje Bliksvær, Emil Bjerknes

Last group meeting where we discussed the progress in the thesis. We have gone through an iteration of feedback from the supervisors, and will have our last supervisor meeting after this meeting, where we will receive even more feedback.

- Still some work needed mainly in chapter 3 and 4. We will most likely receive feedback and improvements mainly on chapter 3-6.
- The plan is to deliver the thesis by the end of the week.

Appendix I

GDPR Expert Scenarios

To evaluate the tool from legal perspective, we need to ensure that the tool provides accurate, reliable, and legally sound advice. Keep in mind that I do not have a legal background, so it might be hard for me to provide/ evaluate the “legally sound advice.”

Scenario 1: Rights of the data subject – Chapter 3 (Art.12 to 23) GDPR

A tech company based in Norway is developing protocols to handle requests from EU data subjects under GDPR Chapter 3 (Articles 12 to 23), which include data correction, data deletion, and restriction on processing. Each request type is based on different legal grounds. The company need to ensure that the company's procedures are compliant with GDPR while considering the nuances of Norwegian privacy regulations.

- Provide a detailed, legally sound response strategy for each request type.
- Advise on potential conflicts between the requests and other legal obligations (e.g., data retention laws).
- Guide the company on how to verify the accuracy of the new data provided by the data subject before making corrections.

Scenario 2: **International organizations and transfer of data to third countries – Chapter 5 (Art. 44 to 50) GDPR**

An international bank operates across several EU countries and third countries is revising its data transfer and protection policies to ensure compliance with GDPR as well as varying local privacy laws that might impose specific additional requirements. The bank need a help with navigating these complex regulatory requirements.

- Identify and reconcile the requirements of GDPR with local laws.
- Advise on how to prioritize the laws when conflicting obligations occur.
- Advise on incorporating Standard Contractual Clauses (SCCs) or Binding Corporate Rules (BCRs) as mechanisms for legal data transfer outside the EU.
- Advise on any necessary amendments or additions to SCCs and BCRs to ensure compliance.
- Compare these provisions with local privacy laws in the third countries where the bank operates, identifying any stricter requirements or discrepancies.

Scenario 3: Border crossing and use of biometric data

An international laws.

- Advise on the specific articles and clauses under GDPR that permit the processing of biometric data for border security purposes.
- Advise whether the processing is based on public interest, legal obligation, or vital interests of the data subjects or others, ensuring the justification aligns with GDPR requirements.
- Advise on the development of GDPR-compliant data retention schedules for biometric information collected at the border.

- Advise on protocols for the secure and lawful sharing of biometric data across EU borders with other member states and relevant agencies.
- Advise on the interoperability of the biometric system with existing EU-wide systems such as SIS II and VIS, ensuring compliance with both GDPR and other relevant EU regulations.

Scenario 4: Healthcare data sharing and transfer in case of emergency

A major hospital in Norway is setting up a system for sharing and transferring patient health data with other medical facilities and emergency response teams, both within Norway and across the EU -- in response to critical situations such as accidents or sudden illnesses that require immediate medical attention. The hospital needs to ensure that the protocols for handling, sharing, and transferring this sensitive data comply with both GDPR and Norwegian health data regulations.

- Advise on the legal grounds under GDPR and Norwegian health data laws for processing and sharing personal health data in emergency situations.
- Advise on circumstances under which the law allows for these data activities without patient consent, particularly focusing on scenarios where there is an immediate risk to health or life.
- Advise on protocols for the transfer of health data to medical facilities outside Norway in emergency situations involving foreign nationals or when specialized treatment abroad is necessary.
- Advise on limiting the scope of data shared to what is strictly necessary to address the emergency.
- Advise on mechanisms to ensure that additional data can be accessed only if further treatment requires it and proper authorization is obtained.

NTNU
Norwegian University of
Science and Technology