# Disposition

## Frontpage (Deadline: October 9th)

* Title: “An investigation into using LLMs for GIS analysis”

## Forewords (Deadline December 11th)

* **1 page**
* Context (result of subject TBA4560, etc.)
* Personal motivation: why did I select this task?
* Thanks to mom and dad, Norkart, supervisor, etc.

## Abstract (Deadline December 8th)

* **1 page**
* 4-6 words per section
* Purpose (1/4)
* Methods (1/4)
* Results (1/4)
* Implications (1/4)

## Table of contents

## Table of figures

## Introduction (Deadline: October 9th)

* **4-6 pages**
* Motivation
  + LLM is a booming field of research, etc.
  + Actor map
    - Different language models
      * GPT-4 (OpenAI)
      * BERT (Google)
      * ELMo (Allen AI)
      * XLNet (Google/CMU)
    - LLM Providers
      * OpenAI
      * Google Cloud
      * Microsoft Azure
      * Amazon AWS
    - User groups
      * GIS professionals
      * City planners
      * Academics and researchers
      * Emergency services
      * Business analysts
      * General public
    - API providers
      * OGC API
      * NGIS Open API
      * STAC API
    - Regulatory bodies
      * European Union (GDPR)
      * Local municipal regulations
    - Data sources
      * Government databases
      * Private databases
      * Crowdsources data
* Problem formulation
  + Research questions (delmål?)
    - Main challenges
    - Repeat them in the discussion
    - 3 RQs maybe?
    - Suggestions
      1. How can OGC API Features be used in an overlay analysis using ChatGPT-4?
      2. How can ChatGPT-4 be used to fetch relevant datasets from OGC Records API using a natural language (NL) prompt?
      3. How would you test the accuracy of your GIS-Chat?

## Theory (Deadline: October 23rd)

* **12-16 pages**
* GIS analysis
  + Examples of usefulness
  + Who needs GIS analysis?
  + How is classical GIS work performed?
* LLMs
  + What are they?
  + Use-cases
    - General use-cases
    - Not specific to GIS analysis
  + Different models
    - Models for specific purposes (trained to produce code, etc.)
  + Strengths/weaknesses
    - Prompting vs. fine-tuning
    - How does language used affect the results (Norwegian vs. English)
    - Access to up-to-date information
  + Fine-tuning
  + Embeddings
  + Prompt engineering
* Geospatial technology
  + State of the art in machine readable standards
  + APIs
    - Geonorge
    - NGIS Open API
  + OGC Records API
  + STAC API
  + Cloud native

## Method (Deadline November 6th)

* **8-12 pages**
* 2-3 different, small technical proofs-of-concepts/experiments. Examples:
  + Test what is possible without any modifications (much like Arild’s demonstration)
    - Chat GPT4 with code completion
    - Emphasise limitations in results chapter
  + Make a web application that takes a prompt, gets relevant datasets from data catalogues, and displays them on the client
    - No spatial queries (try to limit the task)
  + Make web application that connects to a PostGIS database, accepts natural language queries, converts them to SQL and retrieves result to client
* What are the “lab settings”
  + What provider is used?
    - ChatGPT4 API
    - Azure ML studio
* How were the experiments performed?

## Results (Deadline November 13th)

* **8-12 pages**
* Literature that answers you research questions
* Present findings from experiments
  + Elaborate

## Discussion (Deadline December 4th)

* **10-14 pages**
* Strengths/weaknesses of results from experiments
* Path forward
  + How to improve results from weaknesses
* Privacy/legal/regulatory concerns
* Ethical concerns
  + Unintentional bias or inaccuracies
* Comparison with traditional methods
* Integration challenges
  + Want to avoid too much hard-coding – trying to utilize the LLM’s flexibility
* Scalability
  + Handling more complex tasks
  + Combining multiple datasets
* Financial aspects
  + Cost-effectiveness compared to traditional methods
* Adaptability to different languages and regions
  + Focus on Norway at first?
  + Parallel tests using Norwegian and English
    - Same queries, different results?
* Data security
  + Can we mindlessly give ChatGPT access to cadastral data?
  + Azure ML Studio better option in terms of data security?

## Conclusion (Deadline December 6th)

* **3-4 pages**

## Appendix (Deadline December 8th)

* Supplementary material
* Additional data
* Detailed information

## Bibliography (Deadline December 8th)

* List of cited sources