



Geospatial AI-Agent Workflows for Land Management Applications

Alexander Adiyasa

Master's Thesis
February 2025 – May 2025
Surveying, Planning, & Land Management



Abstract

This thesis investigates how Natural Language Processing and geospatial analysis can augment the GIS-based conflict assessment procedure carried out by Danish planning authorities, providing tailored data-driven insights for dispensation case work. Through an iterative methodology informed by stakeholder interviews and analysis of past dispensation cases, a geospatial AI-agent is proposed. This agent integrates an interactive chatbot for spatial query execution with a novel report generation capability, wherein a task-instructed Large Language Model (LLM), guided by a domain-specific instruction-set, synthesizes deterministic geospatial tool outputs into augmented conflict assessment reports. Evaluation demonstrated high factual accuracy in geospatial screening and a 97% success rate in the LLM's application of conditional logic for deriving case-specific regulatory implications. While challenges related to data quality and the LLM's interpretation of uncaptured spatial nuances were identified, the study points to the transformative potential of such AI-driven architecture to facilitate case work and augment existing conflict assessment practices, suggesting co-design of LLM instructions with domain experts as a key methodology for future development.

Total words: 17.229

Preface

My studies in the Landinspektør programme has given me a deep appreciation for land management in Denmark. My time at Aalborg University has been a period of significant personal and professional development, for which I am very thankful. This thesis marks a culmination of that journey.

The project would not have been possible without the contributions and engagement of key individuals. I would like to thank my professors **Jamal Jokar Arsanjani** and **Michael Tophøj Sørensen** for providing foundational insights that helped orient and elevate the project.

Special thanks also to **Vejdirektoratet** for their collaboration and for facilitating dialogue with stakeholders from **Ejlskov A/S** and **Hedensted Municipality**. I am grateful for their willingness to share practical domain knowledge and valuable perspectives, which were essential to the project.

The responsibility for the content, interpretations, and results presented lies entirely with me.

A handwritten signature in dark grey ink that reads "A. Adiyasa". The script is cursive and fluid, with the first letter 'A' being particularly large and stylized.

Alexander Adiyasa
Aalborg University
May 28, 2025

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | Research Questions | 3 |
| 2 | Domain Context | 4 |
| 2.1 | Statutory Basis for Basin Dispensations | 4 |
| 2.2 | Municipal Decision-Making Process | 7 |
| 2.3 | Municipal Dispensations | 8 |
| 3 | Natural Language Processing | 10 |
| 3.1 | Tokenization | 10 |
| 3.2 | Transformer Architectures and Large Language Models | 11 |
| 3.3 | Embedding Representations and Vector Search | 13 |
| 3.4 | Domain-Specific Adaptation | 14 |
| 3.5 | Retrieval-Augmented Generation | 16 |
| 3.6 | Enabling Action and Automation: AI-Agents | 17 |
| 3.7 | NLP Application in Land Management | 20 |
| 4 | Methodology | 22 |
| 4.1 | Bridging Domain Context and NLP workflow | 22 |
| 4.2 | Data Collection | 23 |
| 4.2.1 | Interviews | 23 |
| 4.2.2 | Official Document Cases | 24 |
| 4.2.3 | Geospatial Data | 25 |
| 4.3 | Software | 25 |
| 4.4 | Core Agent Architecture | 25 |
| 4.4.1 | Modularity | 26 |
| 4.4.2 | Agent-Tool Connection | 27 |
| 4.4.3 | Workflow Control | 27 |
| 4.5 | Agent Tools | 28 |
| 4.5.1 | Information Retrieval | 28 |
| 4.5.2 | Geospatial Screenings | 31 |
| 4.5.3 | Document Creation | 35 |
| 4.6 | Use of Generative AI in the Project | 39 |
| 5 | Results: Model Assessment | 40 |
| 5.1 | Chatbot Interactions | 40 |
| 5.1.1 | Simple Request | 40 |
| 5.1.2 | Complex Request | 41 |
| 5.2 | Basin Report Assessments | 43 |
| 5.2.1 | Spatial Accuracy | 43 |
| 5.2.2 | Descriptive Summaries | 44 |
| 5.2.3 | Adherence to Conditional Logic | 46 |
| 6 | Lessons Learned for Future Work | 47 |
| 7 | Conclusion | 50 |

| | |
|---|-----------|
| References | 51 |
| Appendices | 55 |
| A Case Document Examples | 55 |
| A.1 Dispensation Application from Ejlskov | 55 |
| A.2 Environmental Conflict Assessment Report | 59 |
| A.3 Faxe Municipality Decision | 62 |
| B Interview Data | 68 |
| B.1 Interview 1 Questions for Vejdirektoratet and Ejlskov A/S | 68 |
| B.2 Interview 2 Questions for Hedensted Municipality | 68 |
| B.3 Interview 1 Transcription | 68 |
| B.4 Interview 2 Transcription | 68 |
| C Technical Documentation | 69 |
| C.1 Agent Tool Descriptions | 69 |
| C.2 System Instructions for LLM | 71 |
| C.3 Payload Prompt for LLM | 75 |
| D Basin Reports | 77 |
| D.1 Hvidovre | 78 |
| D.2 Ringsted | 80 |
| D.3 København | 82 |
| D.4 Middelfart | 84 |
| D.5 Vejen | 86 |
| D.6 Silkeborg | 88 |
| D.7 Kalundborg | 90 |
| D.8 Horsens 1A | 92 |
| D.9 Horsens 1B | 94 |
| D.10 Horsens 1C | 96 |
| D.11 Køge 1A | 98 |
| D.12 Køge 1B | 100 |

1 Introduction

Vejdirektoratet, the Danish Road Directorate, manages a vast network of rainwater basins supporting the road networks across Denmark [1]. These basins play an important role in collecting runoff water and sediment from roads and highways, and are designed to ensure that potentially polluted sediment is deposited in the basin before lead into downstream creeks and marine environments [1]. Most of the basins are old, with varying levels of documented maintenance history, and as illustrated in Figure 1, internal data from Vejdirektoratet reveals that historic screening and cleaning efforts have been inconsistent, with an often unbalanced emphasis across different periods.

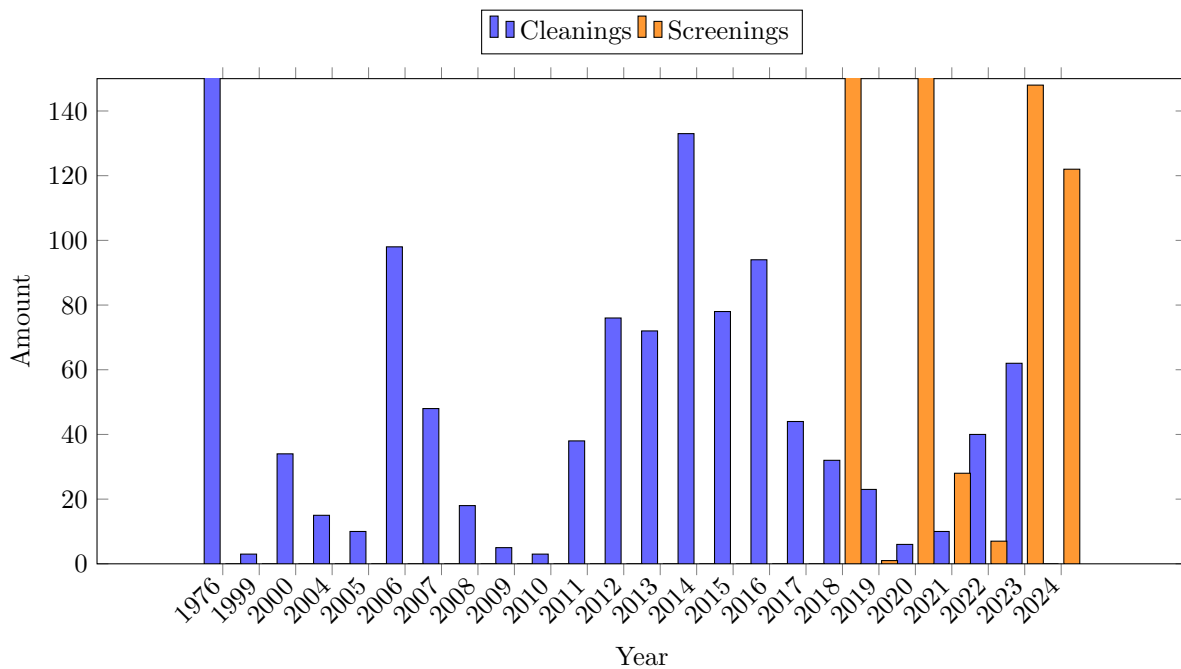


Figure 1: Internal data from Vejdirektoratet; Number of registered basins that have been cleaned and screened between 1976 and 2024. Note: 1976 represents placeholder registrations as that is when the registration first began. Notably, 783 and 246 screenings were conducted in 2018 and 2020.

Emerging as a response to the EU Water Directive [2], Denmark’s third-generation Water Management Plans [3] has made Vejdirektoratet adopt renewed focus to adhere to the plans and ensure the basins do not negatively impact hydrological environments. Through new plans to systematically monitor and further maintain their basins, their developed operational strategy is structured in three main stages: (1.) conducting an initial sediment analysis to determine pollution levels; (2.1) accessing the site, (2.2) draining the basin, (2.3) excavating accumulated sediment, and (2.4) refilling the basin; and afterwards (3) disposing of the sediment through a licensed soil handler equipped to manage the identified pollution levels [1].

To highlight the magnitude of the task, Vejdirektoratet have applied for 500 million DKK, with an additional 25.000 DKK annually per basin to cover costs for environmental studies, operational logistics, and the significant expense of the physical cleaning activities [1]. However, a primary component of the task lies in gaining permission, as the natural areas and species around the basins are often protected by

natural and environmental legislation, prohibiting activities such as the outlined basin cleaning procedure. Central to this is the Danish Nature Protection Act [4], with §3 prohibiting any alterations to the state of natural lakes exceeding 100 m² and other specified nature types like meadows and grasslands over 2500 m².

Given that a significant portion of Vejdirektoratet’s basins meet these criteria due to them originally being natural lakes, the outlined activity, carried out by accessing the site with heavy machinery, necessitates a formal dispensation from the respective municipality. With around 1519 basins, this translates into a notable upcoming workload for Danish municipalities, which must process a large volume of applications. Each application necessitates screening of the site to ensure compliance with the Nature Protection Act and other environmental protection laws, most notably being internationally protected habitats (Natura 2000 sites) and species (Annex IV species) [5, 6].

Carrying out the processing of these cases are the municipal case handlers, which, as a common initial step, performs a *conflict assessment*, where they generate tailored reports outlining spatial relationships to various regulatory spatial data [6]. These standardized factual reports are based on official data and support case handlers in achieving an initial understanding of cases. An example of a basin situation is shown on Figure 2, showing proximity to regulated areas which have to be considered in the case. Case handlers often follow the assessment with site visits to inspect and identify site-specific conditions that shape potential requirements which the work carried out has to follow.



Figure 2: Example of a basin situation with multiple regulatory considerations: as the basin is registered as §3 protected, connected to a stream, with an Annex IV species observation in the near vicinity. Such scenarios are typical of what municipal case handlers assess.

This significant task for both Vejdirektoratet and municipalities has lead Vejdirektoratet to explore ways to support and streamline these processes [1]. Through an iterative research phase anchored by interviews with stakeholders, the thesis identifies the conflict assessment step as a technical intersection where AI, through a combination of geospatial analysis and Natural Language Processing techniques, can augment existing GIS-based methods of generating conflict assessment reports by introducing a layer of reasoning ability to generate tailored preliminary insights.

This research investigates how to incorporate LLM-driven interpretation of geospatial findings to derive initial case facts and potential legal implications from the underlying spatial data. Throughout the

development phase, the project designs and assesses a geospatial AI-agent that answers spatial queries and generates augmented conflict assessment reports. Using task-specific instructions, LLMs can extend their capability to interpret geospatial findings and translate them into case-specific recommendations. The solution is evaluated for factuality, consistency, and alignment with the practical needs of municipal decision-makers.

1.1 Research Questions

To explore this potential and guide the development of the AI-agent, the thesis is structured around addressing the below primary research question and its associated sub-questions. The overall project structure, outlining the progression throughout the project, is shown in Figure 3.

How can natural language processing and geospatial analysis be designed and evaluated to effectively support municipalities in the factual screening and processing of §3 basin dispensation applications within the Danish regulatory context?

To address this overarching question, the following sub-questions have been identified:

- *What specific legal, procedural, and geospatial information is essential for municipal decision-making in §3 basin dispensations, and what are the key variations in current municipal practice?*
- *How can an NLP workflow, incorporating deterministic geospatial tools and LLM-driven interpretation, be used to reliably extract and synthesize the information from spatial data to support the case?*
- *How can the factual accuracy, reliability, and practical utility of the developed AI workflow's outputs be measured and assessed in the context of supporting municipal case handling?*

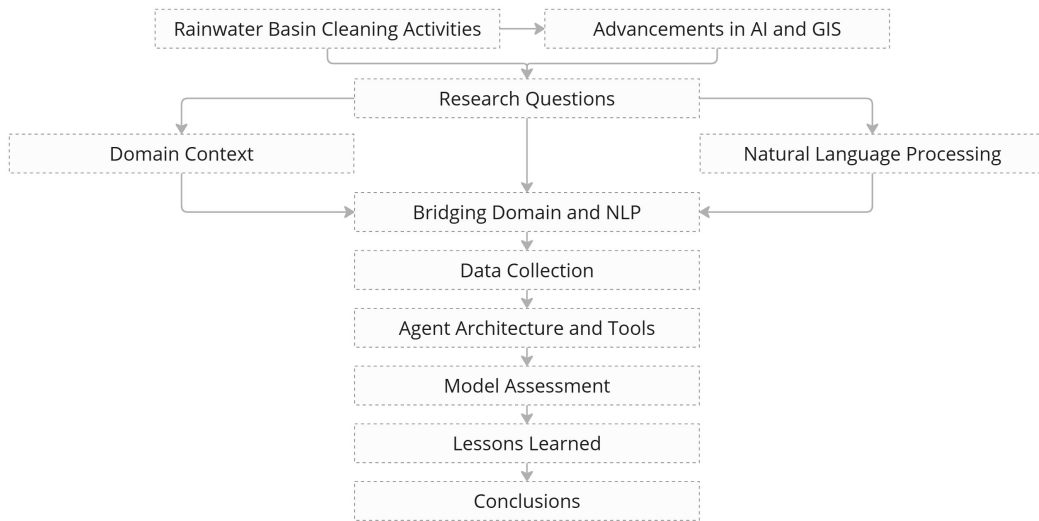


Figure 3: Thesis structure.

The subsequent chapters will; detail the domain context and regulatory framework (Chapter 2), explore Natural Language Processing techniques (Chapter 3), describe the iterative methodology employed in designing the agent and its tools (Chapter 4), present and assess the results (Chapter 5), and discuss key lessons learned and future work (Chapter 6).

2 Domain Context

This chapter examines the regulatory framework and procedural processes governing basin cleaning in Denmark. It draws on insights from interviews with Vejdirektoratet and Hedensted Municipality, as well as analyses of past dispensation cases from municipalities and legislation. The basin cleaning request is governed by an interplay of environmental regulations and administrative procedures, with municipalities having to balance activities with statutory requirements to protect nature, natural habitats, nearby species, and watercourses. By investigating relevant legislation, the municipal case-handling process, and past cases, the chapter aims to establish how an AI agent can be aligned with the realities of municipal workflows.

2.1 Statutory Basis for Basin Dispensations

The cleaning of basins managed by Vejdirektoratet takes place within a regulatory framework shaped by nature conservation and environmental protection laws. Although the primary function of these basins is to control water discharge and road pollution, cleaning activities often require formal approval due to the ecological characteristics that develop over time. Navigating this legal landscape involves complying with key regulations, including the Nature Protection Act, EU Habitat Directives, Environmental Protection Act, and specific rules governing sediment management.

Nature Protection Act

The primary legal basis for municipal involvement is the Danish Nature Protection Act (*Naturbeskyttelsesloven*, NBL) [4], specifically NBL §3, which protects various nature types, including lakes, from alteration. If a rainwater basin within a natural lake exceeds 100 m², it falls under §3 protection. Consequently, any activity altering its state, such as using heavy machinery to empty, remove sediment, or refill the basin, is prohibited. However, municipalities can grant permission through NBL §65, pt. 2, which allows dispensation from §3 under 'special circumstances', balancing cleaning needs against the protection interests of plant and animal life outlined in the regulation.

Natura 2000

If the cleaning activity takes place within a designated Natura 2000 site, the EU Habitats Directive, implemented through the Habitat Act (*Habitatbekendtgørelsen*, HB) [7], imposes additional assessment requirements before any potential NBL §65 dispensation. According to HB §6, the municipality must conduct a 'significant assessment' to determine whether the activity will negatively impact the site's conservation objectives (§6, pt. 1). Generally, if the assessment concludes that the activity will harm the Natura 2000 site's integrity, the dispensation cannot be granted (§6, pt. 2). However, §9 allows for a narrow derogation pathway, permitting dispensation only under the following conditions:

- Has no alternative solution, **and**
- Is necessary for imperative reasons of overriding public interest, including social or economic reasons (§9, pt. 1), **or**
- Is necessary for reasons relating to human health, public safety, or ensuring beneficial environmental impacts (§9, pt. 2).

Importantly, any decision invoking this §9 derogation requires prior consultation with the Minister for Environment (§9, pt. 5), who is obliged to inform the European Commission about the dispensation. This high procedural threshold underscores that overriding harm to Natura 2000 site integrity is a rare and exceptional measure, involving higher administrative levels. As mandated by HB §6, pt. 4, the significance assessment must be documented in the final §65 decision.

Annex IV species

Irrespective of Natura 2000 status, the Habitat Directive imposes protection requirements for Annex IV species throughout their natural ranges, reinforced by the direct prohibition in NBL §29a, which is punishable by fine if violated (NBL §89, pt. 1). If an activity negatively impacts breeding or resting sites for Annex IV animal species, HB §10 states that permits under HB §§ 7-8 (including NBL §65 dispensations) generally cannot be granted. In such cases, the municipality can only grant permission if no satisfactory alternative exists, as provided by the derogation in HB §11, which requires that:

- There is no satisfactory alternative solution, **and**
- The derogation does not hinder the maintenance of the species' conservation status in its natural range, **and**
- The activity serves one of the specific imperative purposes listed in §11, pt. 1, which includes ensuring environmental interests.

Before granting a dispensation based on this derogation, the public authority (the municipality) must first consult the Minister for Environment to obtain a statement, as dictated by §11, pt. 4. Separately, NBL §65, pt. 3 allows the Minister for Environment to directly grant dispensation from the NBL §29a prohibition itself, following similar strict conditions. However, for activities requiring a municipal permit, such as the §65, pt. 2 dispensation, the integrated procedure via HB §11 typically applies when handling Annex IV conflicts¹.

Environmental Protection Act

Beyond the regulation covering protected nature and species, the Environmental Protection Act (*Miljøbeskyttelsesloven*, MBL) [8] addressees potential pollution impacts resulting from the cleaning activity. Primarily, MBL §27 prohibits the discharge of substances that could pollute water bodies. Therefore, discharging potentially polluting water from the basin during cleaning is prohibited under MBL §27 unless permitted by the municipality through §28. Consideration of the receiving water body's sensitivity, informed by Water Framework Directive classifications, may influence permit conditions.

Sediment Handling

The physical handling and disposal of sediment from the basin constitute a parallel regulatory consideration. The excavated sediment is subject to the Soil Contamination Act (*Jordforureningsloven*) [9], necessitating chemical analysis before excavation to classify contamination levels. This classification determines the legally required methods for transport, disposal, and the license of soil handlers in accordance with the Waste Act

¹Importantly, the formal derogation procedure only applies if the activity will negatively impact Annex IV species or Natura 2000 sites.

(*Affaldsbekendtgørelsen*) [10]. Additionally, the Museum Act (*Museumsloven*) [11] mandates that excavation work stop and be reported if archaeological findings are encountered.

Synthesis of Framework

In summary, basin cleaning activities are governed by intersecting regulations. A primary NBL §65 dispensation, required due to a basin’s potential §3 status, may be constrained by mandatory assessments for Natura 2000 impacts and Annex IV species protection. If negative impacts are identified, derogation rules apply, involving ministry consultation and notification to the EU committee. Concurrently, separate processes under the MBL regulate water discharge, while soil contamination and waste laws govern sediment handling, alongside archaeological safeguards under the Museum Act. Table 1 summarizes the relevant regulatory implications.

| Legal Area | Key Trigger / Issue | Permit / Procedure / Implication | Constraint on NBL §65 Dispensation? |
|-------------------------|--|---|--|
| §3 Protected Nature | Basin meets §3 criteria (natural lake, $> 100 m^2$, natural life) | Dispensation Required (NBL §65 pt. 2) for alteration of state. | N/A (Core permit) |
| Natura 2000 | Activity affecting site | Mandatory Assessment Procedure (HB §6): Significance/impact assessment. <i>Precedes & informs §65 decision.</i> | Yes. Generally <i>CANNOT</i> grant if site integrity is harmed (HB §6 stk 2). <i>UNLESS:</i> derogation via HB §9 |
| Annex IV Species | Activity potentially affecting breed/rest site or plant | Mandatory Assessment Procedure (HB §11): Significance/impact assessment. <i>Precedes & informs §65 decision.</i> | Yes. Generally <i>CANNOT</i> grant if site integrity is harmed (HB §10). <i>UNLESS:</i> derogation via HB §11 |
| Water Discharge | Pumping water out, discharge to water body | Potential Discharge Permit requirement per MBL §28. Must avoid pollution (MBL §27). | Indirectly (Permit conditions affect work). |
| Sediment Handling | Removal of potentially contaminated sediment | Analysis, classification, regulated transport/disposal. <i>Parallel process.</i> | Indirectly (Permit conditions affect work). |
| Archaeological Findings | Discovery during excavation | Stop work, report finding. | No. |

Table 1: Overview of key statutory considerations affecting basin cleaning activities.

2.2 Municipal Decision-Making Process

The Municipal handling of dispensation applications follows established principles of judicial theory, encompassing fact-finding, legal assessment, and a justified decision [12]. The process formally commences when the municipality receives an application, typically submitted by Ejlskov A/S, acting on behalf of Vejdirektoratet as environmental consultants. These applications have been developed to follow a standard format and frequently include a set of suggested conditions based on previous successful dispensations [5]. As outlined in an example of a standard application shown in Appendix A.1, the document specifies the basin, cleaning methodology, justification, and work conditions, along including sediment handling.

Upon receiving the application, the initial municipal step often involves a factual screening referred to as a '*conflict assessment*' [6], where standardized data retrieval procedures generate factual overviews to derive potential regulatory implications (see shortened example from Vejdirektoratet's equivalent environmental conflict assessment in Appendix A.2). Key objectives of the basin screening include verifying the basin's NBL §3 protected status; checking for known observations of Annex IV species; and identifying nearby watercourses and potential Natura 2000 areas [6]. While Vejdirektoratet possesses internal operational data regarding basin conditions and history, this information is not consistently included in initial applications, potentially leaving gaps to be uncovered by the municipality [6].

While digital screening provides an essential baseline, a site visit is frequently deemed necessary by municipalities for assessing the sites [6]. Although some past decisions indicate occasional reliance solely on remote data, interviews suggest site visits are crucial for addressing uncertainties unresolved by digital checks [5, 6]. Such site visits allow case handlers to assess actual ground conditions, including practical site access limitations, actual vegetation state, unrecorded features potentially relevant to sensitive species, and other things relevant to the case, hereby enabling more informed decision permits than those based on digital data alone, with closer alignment to legislative intent.

Following the information gathering from the application, digital screening, and potential site visit, the municipality undertakes the legal assessment [6]. This involves synthesizing the established facts and evaluating them against the specific legal requirements. This stage represents the core administrative task, where case handlers balance the applicant's documented needs against legal obligations to protect the specific interests. These considerations influence the decision-making process and determine the procedural pathways used by municipalities.

The decision-making process, shown in Figure 4, culminates in the municipality issuing a formal decision. While a rejection is possible, the outcome for the necessary cleaning often involves granting the NBL §65 dispensation under specific conditions. Legally, the final decision must be accompanied by a justification referencing the factual basis and the considered legislation, potentially including documentation of the mandatory assessments under the Habitat Act if applicable. An example of a standard decision is provided in Appendix A.3 alongside the related application.

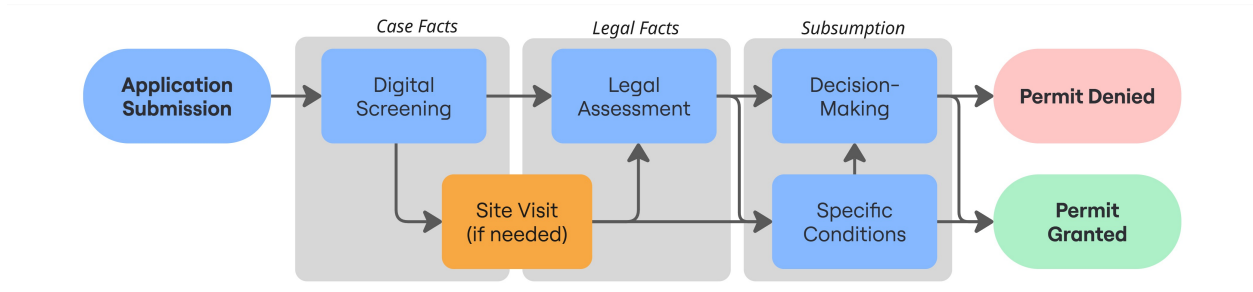


Figure 4: Illustrates the general process municipalities follow when handling basin applications.

2.3 Municipal Dispensations

While the legal framework provides pathways for basin cleaning, practical applications by municipalities show both patterns and variations. Analysis of past cases from 56 municipalities reveals a common starting point: municipalities generally recognize basin cleaning as a valid "special circumstance" under NBL §65, justifying dispensation from §3 protection [13, 14, 15, 16, 17]. Typically, municipalities support the cleaning activity due to its potential to remove accumulated, nutrient-rich sediment and control dense vegetation, thereby restoring open water surfaces, increasing light penetration, and enhancing habitat quality for fauna and amphibians [13, 14, 16].

Despite differences in the final permits, the underlying legal screening process appears standardized across municipalities. Case handlers consistently verify the basin's §3 status and conduct the mandatory screenings stipulated by the Habitat Act. This involves checking for proximity with Natura 2000 sites and screening for the presence of potential habitats of Annex IV species. A significant degree of standardization is evident in the conditions attached to the dispensations. A core set of recurring requirements is present in the vast majority of permits reviewed, mirroring those proposed in the standardized applications. Conditions typically include:

- Strict timing limitations confining work to the period outside bird and amphibian breeding seasons (e.g., September to March/April).
- Methodological constraints such as only removing sediment until a firm bottom (*fast bund*) is met, avoiding deepening, and prohibiting the steepening of banks.
- Requirements for managing water removal to minimize sediment suspension and ensuring controlled outflow.
- The preservation of an undisturbed vegetated buffer zone (e.g., 10-15% of the bank area, or a 10-15m distance) to facilitate natural recolonization.
- Prohibition of seeding or planting non-native species, favoring natural succession post-activity.
- Prohibiting the introduction of non-native animals (e.g., fish, ducks) and installing artificial structures unrelated to drainage (e.g., duck houses, feeding platforms, jetties).
- Prohibition regarding chemical use (pesticides/fertilizers) near the basin.
- Inclusion of standard clauses regarding the Museum Act (§§25-27) procedure for handling potential archaeological finds.

One of the most prominent areas of variation lies in the decisions around handling and disposal of the excavated sediment. This is a decision influenced by sediment quality, surrounding land sensitivity, and local municipal practice. While direct removal to a licensed waste/soil facility is frequently mandated, especially if analyses show contamination or if the surrounding area is protected [14, 15, 17], some municipalities permit on-site spreading under strict conditions. In the reviewed applications, this typically requires;

- The land on which the sediment is placed is owned by the applicant (Vejdirektoratet).
- The land must not be §3 protected.
- The sediment is documented as minimally contaminated (e.g., meeting specific thresholds).
- Adherence to specific limitations on layer thickness (e.g., < 50 cm) and distance from water bodies (e.g., > 10 m).

The approach can also evolve, as seen where one municipality shifted from allowing on-site reuse in 2018 to requiring off-site disposal in a later 2022 permit for the same site [13, 18]. Temporary on-site storage is generally avoided due to practical limitations and potential permit needs [5].

Variation arises from site-specific conditions identified during municipal site visits [6]. These often involve vegetation management, such as preserving valuable trees while clearing others for light exposure, or retaining specific vegetation to support local ecology [19]. Sometimes, cut material is repurposed as "insect hotels" for biodiversity [16]. Physical modifications vary, with some permits specifying exact bank slopes, while others focus on maintaining natural contours [20, 21]. Procedural demands range from basic work notifications to requiring photo documentation or drainage plan approval [14, 19, 22]. The permissibility of future maintenance also differs, with some permits clearly stating actions not needing new dispensation, while others implicitly follow Vejdirektoratet's long-term plan [14, 16, 15, 19].

While Annex IV species screening is standard, mitigating actions beyond timing restrictions are central to municipal management of these cases. Instead of invoking formal derogation procedures under Habitatbekendtgørelsen §11 (requiring Ministerial consultation), municipalities typically impose specific conditions within the NBL §65 dispensation itself to minimize harm to Annex IV species and their habitats, thereby avoiding violations of NBL §29a or HB §10 [6]. Examples include daily pre-work checks for amphibians and reptiles, safe relocation procedures, defining strict work boundaries, or mandating non-intrusive methods [22]. By proactively setting these protective conditions, municipalities manage the activity without triggering higher-level derogation or EU/Ministry involvement [6].

In sum, Danish municipalities work under a common statutory framework and follow broadly consistent assessment routines and decision outcomes. However, final permits often show notable variation due to each basin's unique on-site conditions, which must be addressed through tailored requirements set through the permit conditions. This combination of standardized rules and context-specific adjustments defines the practical limits and opportunities that any digital support tool must respect. To align an AI support workflow with these realities, the next chapter will explore relevant natural language processing techniques and models.

3 Natural Language Processing

Natural Language Processing (NLP) techniques offer new possibilities for automating parts of land management, particularly in handling the vast amounts of data that characterize the field. This chapter explores foundational and advanced NLP techniques, focusing on how they can be leveraged to support geospatial AI-agent workflows. By covering methods such as tokenization, transformer architectures, embedding representations, and retrieval-augmented generation, this chapter aims to bridge the gap between NLP theory and practical application. Additionally, the chapter examines AI-agents and prompt engineering to better understand how they can be applied to decision-support within specific tasks.

3.1 Tokenization

Tokenization serves as a fundamental component of nearly all NLP applications as it is the initial step in text processing [23, 24]. Tokenizing text breaks unstructured data and natural language text into chunks of information that can be processed. These token sequences can be directly translated into vector representations, turning text into numerical meaning suitable for machine learning. By representing text in a manner which is meaningful for machines, tokenization allows NLP algorithms to identify semantic relationships and patterns in natural language [25, 26, 27]. Modern large language models predominantly employ subword tokenization techniques such as Byte-Pair Encoding [26] and SentencePiece [27] due to their ability to understand morphological variation², hereby balancing vocabulary size with representational completeness, as showcased in Figure 5

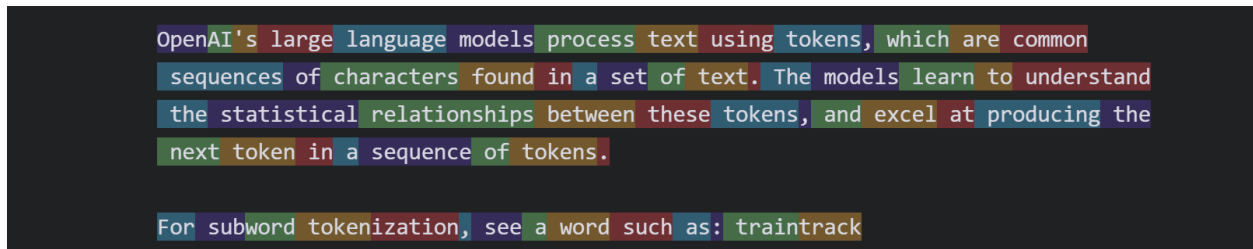


Figure 5: Example of a tokenized text for GPT 4o using a mix of subword- and character level tokenization. 326 characters, 63 tokens.

Tokenization effectiveness varies significantly across languages and domains, presenting challenges for universal tokenizers. Research comparing tokenization between Turkish and English shows that morphologically rich languages like Turkish may require 2-3 times more tokens than English for semantically equivalent sentences [29]. Similarly, domain-specific corpora often contain specialized terminology underrepresented in general tokenizer training data, leading to suboptimal token segmentation and limiting domain understanding [30]. Inefficient tokenization can manifest as rare domain terms being fragmented into unrelated subwords, incorrect segmentation of morphological features, or minor spelling variations causing drastically different tokenization patterns. Poor tokenization quality reduces accuracy by misplacing boundaries, while longer sequences impair computational efficiency and contextual modeling.

²In linguistics, *morphology* refers to the internal construction of words [28].

3.2 Transformer Architectures and Large Language Models

Large Language Models (LLMs) rely on the *Transformer* architecture, introduced by Vaswani et al. [31], which represents a major advancement in *sequence transduction models*. These models are designed to map one sequence to another, making them suitable for tasks like translation, summarization, and text generation. Earlier models, such as *recurrent neural networks* (RNN), process data sequentially, which makes *capturing long-range dependencies* challenging. The dependencies refer to the ability to capture relationships between words or phrases that are far apart in a sentence or document³. Traditional models like RNNs often struggle to capture these distant connections because they process text sequentially, which can lead to information loss over long sequences [32].

In contrast, Transformers use *multi-headed self-attention* to process all tokens simultaneously, allowing the model to dynamically weigh the relevance of each word in the context of the entire sequence. This approach overcomes a major limitation of RNNs by processing all tokens in parallel. By doing this, Transformers significantly enhance computational efficiency and the ability to understand relationships between distant words or concepts within a text. The self-attention mechanism allows each word to attend to every other word in the input, calculating attention scores that reflect their relative importance, making Transformers more scalable and accurate for complex language tasks. Additionally, some modern Transformers extend to multi-modal applications, combining text with audio, images, or video, enabling cross-modal reasoning and content generation [33].

Two important sub-types of Transformer-based architectures are **BERT** [34] and **GPT** [35]. BERT is bidirectional, meaning it reads a sentence by considering both preceding and following tokens, which makes it excel in context-aware tasks such as question answering or zoning-clause classification. This bidirectional nature allows BERT to capture semantic relationships from both directions, making it suitable for embedding-based models (discussed in later sections). In contrast, GPT is unidirectional and auto-regressive, predicting each next token from left to right, making it commonly used in state-of-the-art LLMs. Its capacity for text generation also supports *zero-, one-, and few-shot* prompting [35], where the model performs a task with no examples, a single example, or a few examples. This ability allows for task specialization without further training, enhancing its flexibility.

Modern large language models are trained on vast and diverse datasets, often comprising billions of tokens from sources such as books, websites, academic papers, and code repositories [35]. During training, the model learns by analyzing patterns within the data, using techniques like *masked language modeling* [31] or *next-word prediction* [36] (both expanded upon in later sections) to build statistical associations between words and phrases. These learned patterns are stored as numerical weights within the model's neural network represents *parametric knowledge*, which allow it to generalize and generate language that resembles human text [31]. The extensive training enables LLMs to encode statistical patterns and semantic relationships across various contexts, making them highly versatile for tasks ranging from text generation to complex reasoning.

³As an example, in the sentence "The servitude obliged the owner of the land to preserve access for the municipality." The word "servitude" and "access" are related, despite being separated by several words.

Despite being able to fluently generate linguistic responses to queries, these models do not store factual knowledge explicitly, but rather encode associations between tokens [37]. When the input deviates from the model’s training data, especially when the data contains biases, misinformation, or lacks domain-specific content, the model may produce text that sounds plausible but is factually incorrect - a phenomenon known as *hallucination*. According to Huang et al. [37], hallucinations arise from three main sources: data issues, training misalignment, and inference errors. Data-related causes include biases and misinformation embedded in the training corpus or a lack of representation for niche topics. Training-related causes often stem from processes where the model is encouraged to respond confidently, even when uncertain. Inference errors typically involve decoding strategies that favor fluency over factual accuracy.

Model capacity, defined by the number of parameters and the size of the context window, plays a crucial role in the ability of large language models to handle different levels of reasoning. LLMs with larger capacities, such as GPT-4o with 175 billion parameters and a context window of 128k tokens, can manage more extensive input sequences compared to smaller models with limited context handling [38]. This ability to process longer contexts is essential for reasoning over multi-step tasks or analyzing comprehensive documents. To objectively assess model capabilities, a range of benchmarks have been developed, each targeting different aspects of language comprehension and reasoning. These benchmarks include Humanity’s Last Exam (HLE) [39], Graduate-Level Google-Proof Q&A (GPQA) [40], and LiveCodeBench (LCB) [41], which evaluate reasoning, scientific knowledge, and coding skills respectively. The results of these evaluations are presented in Table 2.

Humanity’s Last Exam (HLE): *Reasoning & Knowledge*

Designed as a close-ended multi-modal academic test of human knowledge, this benchmark evaluates AI systems on expert-level reasoning and knowledge across 2,700 questions spanning physics, mathematics, biology, chemistry, history, literature, and philosophy. The questions are significantly more difficult than existing academic benchmarks, requiring deep subject understanding rather than answers based on memorization or internet lookup [39].

Graduate-Level Google-Proof Q&A Benchmark (GPQA): *Scientific Reasoning*

This is a benchmark of 448 expert-authored multiple-choice questions in biology, physics, and chemistry, designed to evaluate deep scientific reasoning. The questions are intentionally difficult, with PhD-level experts achieving 65-74% accuracy. Non-experts with full web access perform significantly worse at 30-34% accuracy, highlighting a resistance to surface-level search strategies and a focus on reasoning over recall [40].

LiveCodeBench (LCB): *Coding*

LCB is a benchmark for evaluating code-related capabilities of LLMs across 400 problems sourced from competitive programming contests. Unlike earlier benchmarks, it tests code generation as well as execution, self-repair, and other types of coding-related tasks, offering a more comprehensive and up-to-date assessment of coding proficiency. The problems are selected from recent contests to reduce the chance that models have seen them during training [41].

As mentioned, LLMs can produce hallucinations, where they generate plausible-sounding but false information [37]. This can occur for various reasons, such as when the model encounters ambiguous input,

| Developer | Model | Context Window | Parameters | HLE (%) | GPQA (%) | LCB (%) |
|------------|----------------|-------------------|------------|---------|----------|---------|
| Google | Gemini 2.5 Pro | 2m (7800 pages) | n/a | 17.70 | 83 | 73 |
| OpenAI | o3-mini high | 200k (781 pages) | n/a | 12.30 | 77 | 73 |
| Deepseek | Deepseek R1 | 128k (500 pages) | 671B | 9.30 | 71 | 62 |
| OpenAI | GPT-4o | 128k | 175B | 3.30 | 54 | 30 |
| Mistral AI | Mistral Large | 128k | 123B | n/a | 48 | 29 |
| Mistral AI | Mistral Small | 128k | 24B | n/a | 45 | 21 |

Table 2: Comparison of state-of-the-art LLMs across model specifications and benchmark results [38].

lacks domain-specific knowledge, or attempts complex multi-step reasoning [37, 42]. Additionally, models may rely too heavily on statistical patterns learned from their training data, leading to incorrect outputs when handling novel or nuanced topics. To mitigate hallucinations, several strategies have been proposed, including *prompt engineering*, which encompasses a range of techniques aimed at guiding the model [43], *chain-of-thought* (CoT) prompting [44] to guide reasoning processes (subset of prompt engineering), and *reinforcement learning from human feedback* to align outputs with expert judgment [45]. In high-stakes applications, incorporating a *human-in-the-loop* verification process ensures that critical decisions are cross-checked by experts, balancing automation with transparency [46].

3.3 Embedding Representations and Vector Search

Embedding models transform textual data into high-dimensional numerical vectors that capture semantic relationships between words and phrases. In the context of LLMs, these embeddings can function as bridges to external knowledge bases by enabling meaningful comparisons between the query text and the knowledge base [47]. Once text is embedded as numeric vectors, it can be stored in vector databases and retrieved through semantic similarity searches [34, 48]. In an example where the query *building height requirements* is embedded, it is converted into a vector and matched against stored data to retrieve relevant documents - ideally from local plans, servitudes, and other regulatory texts in a Danish context.

Approaches to generating embeddings range from traditional methods like *Word2Vec* [48] and *GloVe* [25], to more advanced techniques like *BERT* [34]. Earlier methods like Word2Vec generate word embeddings independently of their context, while newer models like BERT produce dynamic contextual embeddings that capture word meanings based on surrounding text. This difference allows contextual embeddings to perform better in tasks where word meaning varies. Embeddings are typically high-dimensional, often with hundreds of dimensions, enabling them to capture deeper contextual nuances. As illustrated in Figure 6, the embedding process clusters semantically related objects together, allowing the model to find contextually relevant information even when the wording of a query differs from the stored data.

Embedding models trained on general web data often struggle to capture the specialized terminology and semantic relationships inherent in specific domains [34]. In the earlier *building height example*, a generic model might incorrectly retrieve information from the Building Registry (BBR), providing insufficient context to the LLM. If the retrieval process is of poor quality, the LLM’s performance will be compromised, requiring careful model adaptation [34]. Such limitations require consideration of domain adaptation

techniques to improve model understanding of domain-specific concepts and nuances. The following section will explore strategies for aligning embedding models with domain-specialized applications.

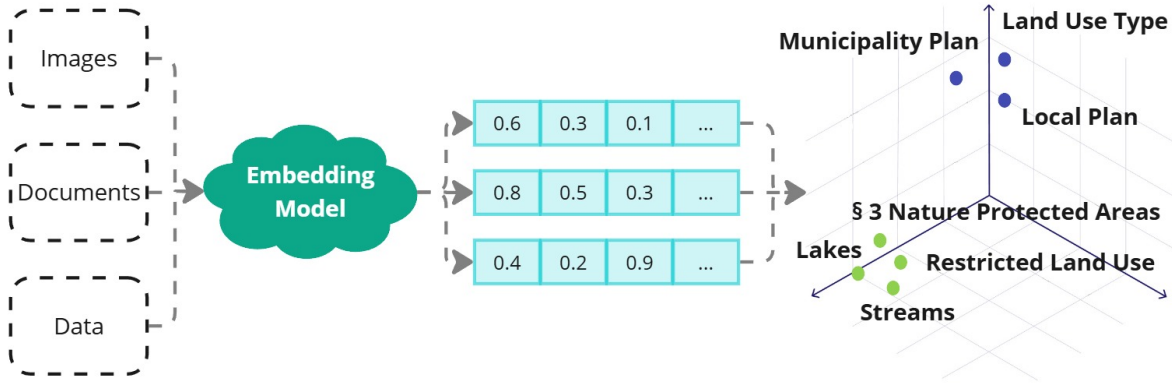


Figure 6: Embedding process, where data are transformed into high-dimensional vectors using an embedding model, allowing semantically related objects to be clustered together in a vector database for efficient retrieval.

3.4 Domain-Specific Adaptation

Domain-specific adaptation is an essential consideration when LLMs and embedding models are to be employed in specialized fields such as land management. General-purpose models, typically trained on vast and diverse datasets from the web, often fail to capture the nuanced terminology and contextual relationships present in specific domains. For example, the term *parcel* in land management should be closely related to *property* and *servitudes*, but a generic model might incorrectly associate it with unrelated contexts, such as logistics or shipping. This mismatch can lead to imprecise outputs, making domain adaptation critical for ensuring accurate retrieval and interpretation of regulatory texts and zoning laws. The effect on semantic clustering is shown in Figure 7, illustrating the differences between a pre-trained (general) and a fine-tuned (specialized) model.

Adapting LLMs and Embedding Models

Adapting LLMs and embedding models to domain-specific tasks involves fine-tuning them to better understand specialized terminology and context. One approach is Broad Domain Learning (Pre-training), where the model is exposed to large volumes of domain-relevant text without targeting a specific task. This method builds a foundational understanding of the terminology used in regulatory and legal contexts. In contrast, Complete Adaptation (Full Fine-Tuning) updates all model parameters using a carefully curated dataset, ensuring that the model captures nuanced language features, but requiring significant computational resources. An intermediate approach is Partial Adaptation (Task-Specific Fine-Tuning), where only selected layers are fine-tuned to balance efficiency and performance for a specific task [34]. The three following levels represent common training techniques:

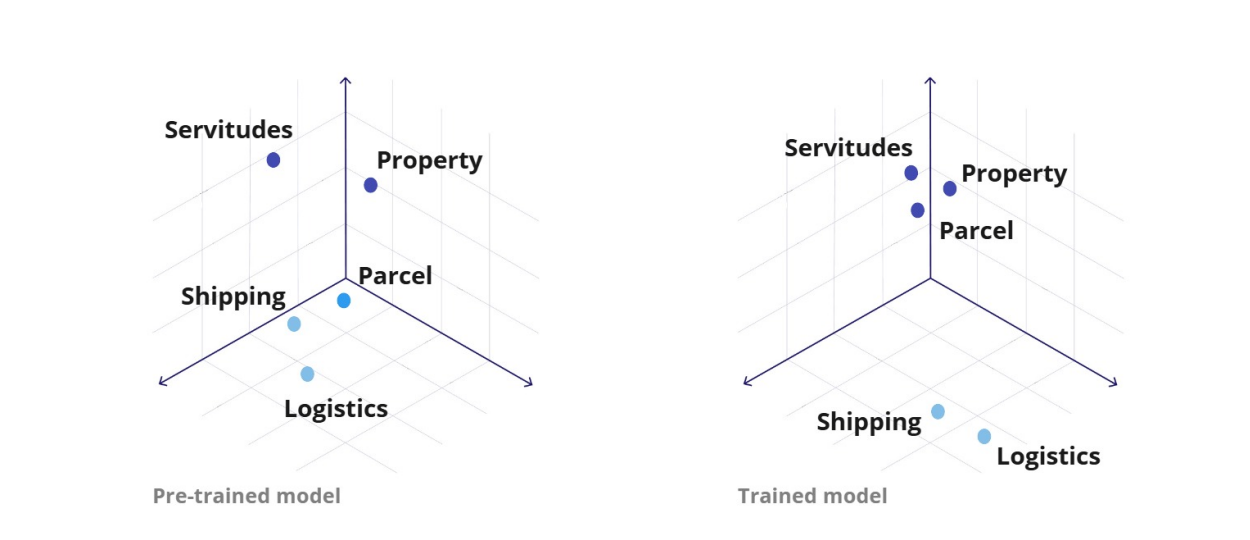


Figure 7: Simplified visualization of embedded data in a vector database. The pre-trained model (left) shows a more dispersed representation of terms across domains (property law and logistics). After fine-tuning (right), domain-specific words cluster more closely (it should be noted that embedding models often utilize hundreds of dimensions to compare and contrast).

- **Broad Domain Learning (Pre-training):** The model is exposed to large quantities of domain-relevant text, developing a foundational understanding of regulatory terminology, without optimizing for a specific retrieval task. This approach provides generalizable knowledge for the domain with moderate computational requirements [34, 36].
- **Complete Adaptation (Full Fine-Tuning):** The entire embedding model undergoes further training on specialized datasets to maximize alignment with specific regulatory and legal terminology. This ensures high precision but requires significant computational resources [34, 36].
- **Partial Adaptation (Task-Specific Fine-Tuning):** Only selected layers of the model are fine-tuned for specific retrieval or classification tasks. This approach balances computational efficiency and performance, being cost-effective for adaption to specific use cases within the broader domain [34, 49].

The effectiveness of these training techniques depends on the availability of well-structured training data. Structuring this data is critical, as it determines how well the internal parameters of an embedding model understands semantic relationships and how well it performs in retrieval tasks [50].

Training Data for Domain Adaptation

High-quality training data is essential for optimizing embedding models [34]. The training data should accurately reflect semantic relationships and knowledge structures within the domain, ensuring that the model understands how legal and regulatory concepts interconnect in land management contexts. Well-structured training data enables embedding models to distinguish between closely related concepts, interpret variations in terminology, and retrieve the most relevant documents with high precision [51]. To effectively train embeddings for domain-specific tasks, training data must be structured in ways that expose the model to variations in legal terminology, phrasing, and conceptual relationships [51]. Different structuring techniques reinforce

different aspects of language understanding. While some ensure consistency in terminology, others help recognize equivalent meanings across expressions. The right combination of methods enables embeddings to learn nuanced variations in legal concepts, improving their performance in applications within the domain.

One fundamental technique is next-word prediction, where the model is trained to predict the next token given a sequence of preceding words [45]. This approach is particularly useful for generating contextually appropriate text and maintaining coherence during reasoning tasks [45]. Another widely used method is masked learning, where a portion of the data is hidden by a [MASK] token, training the model to predict the missing words within a sentence [34]. This technique enhances the model’s ability to understand context and make accurate inferences from partial information [34]. A complementary technique is contrastive learning, which trains embeddings on triplets: an anchor (concept, term, or definition), a positive (related information), and a negative (unrelated information). This method ensures that similar meanings are positioned closer in the vector space while unrelated ones are pushed further apart. In land management, this helps distinguish between terms like "parcel" (as a land unit) and "parcel" (as a package), allowing more accurate semantic retrieval [48, 52].

3.5 Retrieval-Augmented Generation

Retrieval-Augmented Generation (RAG) is a technique that enhances the reasoning abilities of LLMs by grounding their outputs in reliable, external data sources [50]. While LLMs generate responses based purely on *parametric knowledge*, RAG dynamically retrieves relevant information from databases or knowledge bases during inference⁴ and injects it as context to the LLM. This approach not only improves factual accuracy but also ensures that responses are contextually relevant and traceable to its sources. A typical RAG pipeline has three components. First, an embedding model converts both user queries and corpus documents (such as Danish planning acts, local plans, or past rulings) into numerical vectors. Second, those document vectors are stored in a vector database that supports semantic similarity searches. Third, at inference time, a query vector retrieves the top- k relevant passages; with these passages supplementing the original query passed to the LLM. This then produces a context-aware, source-backed answer. Figure 8 illustrates this workflow.

Despite its potential, RAG faces challenges related to retrieval precision, where the system fails to retrieve relevant results [50]. This is not a trivial issue, and often arises when query phrasing differs from how the information is stored, leading the model to overlook relationships between the query and the stored information [50]. The problem can be compounded by poorly suited embedding models that fail to capture domain specific semantic nuances, and ensuring retrieval quality is therefore critical, especially in regulated domains where precise information is essential. Addressing these challenges involves techniques like query reformulation, where the query is reformulated before being used to search to improve match quality, or re-ranking, where retrieved results are prioritized based on contextual relevance. By refining how the query and stored vectors are aligned, these techniques prove very useful in reducing the risk of irrelevant or misleading outputs.

⁴ "the process where [an LLM] uses its learned knowledge to make predictions or decisions on new, unseen data."

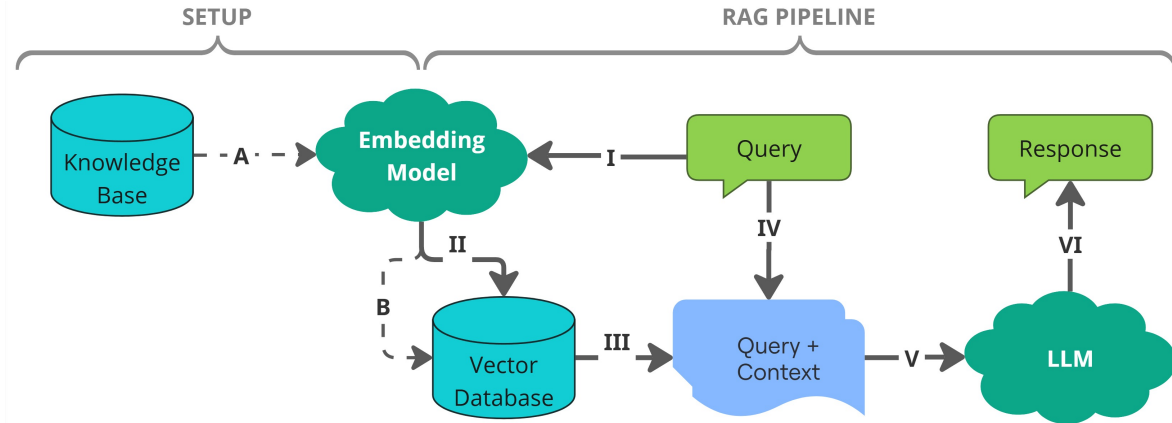


Figure 8: Illustration of the RAG setup and pipeline workflow. **Database creation:** using an embedding model, a knowledge base is (A) embedded and (B) stored in a vector database. **RAG pipeline:** during retrieval, when a query is passed, it is (I) converted to a vector and (II) semantically matched to relevant documents. Matched information is then (III) combined with (IV) the original query and (V) passed to the LLM, which (VI) produces a contextually informed response.

3.6 Enabling Action and Automation: AI-Agents

While LLMs are only capable of generating text, agents extend the capabilities of LLMs by being able to act and accomplish tasks [53, 54]. As OpenAI describes, “*while conventional software enables users to streamline and automate workflows, agents are able to perform the same workflows on the users’ behalf*” [53]. These enhancement allows LLMs to execute custom scripts, perform API calls, and take other actions as part of a larger reasoning process. Unlike traditional software that follows pre-defined rules, an AI agent can make decisions, adapt to new information, and handle complex workflows without direct human intervention. An agent typically consists of three core components: (1) an LLM powering the agent’s reasoning process, (2) external tools and functions to take action (tool calling), and (3) an instruction set guiding behavior and action, with memory management maintaining context throughout multi-step processes.

Rule-Based and Autonomous Agents

Developing multi-agent workflows falls on a spectrum between rule-based and autonomous systems [55]. The primary difference lies in the degree of freedom granted to autonomous agents compared to the sequential, deterministic nature of rule-based workflows [55]. Rule-based systems produce predictable, transparent outcomes, making them suitable for tasks that require accuracy and consistency within regulated domains [56, 57]. These systems integrate structured or unstructured data into LLMs, which function as nodes within a predefined workflow [55, 56]. However, the main limitation of rule-based systems is their lack of adaptability when faced with scenarios beyond the predefined rules, making them less effective for handling complex or evolving tasks [55]. Additionally, maintaining and expanding a rule base is labor-intensive, especially as tasks become more diverse or complex [55, 56].

In contrast, autonomous systems enable dynamic approaches by integrating explicit toolsets and relying on learned knowledge and reasoning capabilities, rather than being constricted by a fixed, sequential

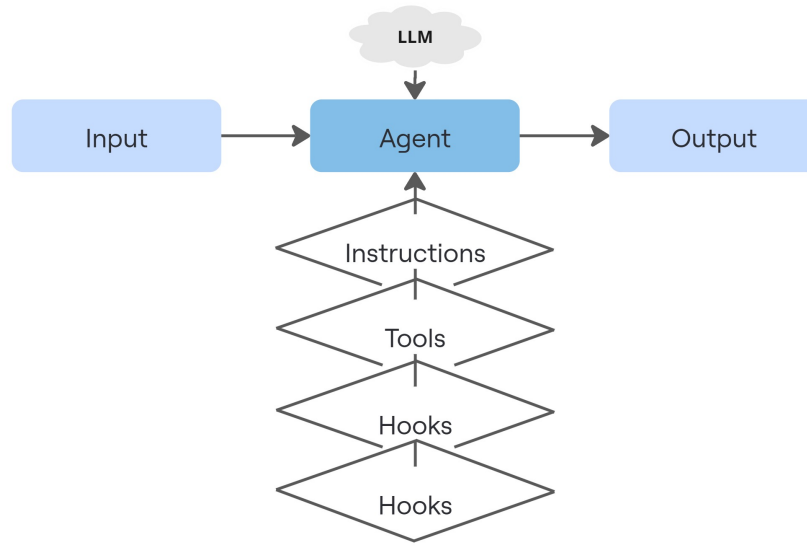


Figure 9: Foundational agent architecture, inspired by OpenAI [[53]]. Illustrates how agents can be augmented through multiple layers: instructions that guide behavior, tools that enable action, hooks that allow custom logic such as logging, validation, or intervention, and guardrails to impose constraints to prevent undesired behavior.

processes. This flexibility makes them well-suited for handling open-ended or complex, nuanced queries [54, 58]. The autonomy of these agents stems from their ability to access external tools, memory, and plan action. In these systems, multiple agents often collaborate to solve a task, navigating novel situations by utilizing their reasoning capabilities and tools to achieve specified objectives [54, 55, 58]. However, this dynamic nature also introduces challenges related to reliability and control, as their decision-making processes remain opaque, making it unclear why a particular output was generated [55, 58].

An autonomous AI agent operates by orchestrating a continuous loop where it assesses the user query, selects tools, processes data, and evaluates whether the task is complete [53, 54]. This process involves three main steps: (1.) tool selection, (2.) action execution, and (3.) result verification. These steps are performed iteratively until an exit condition is met, such as obtaining a satisfactory answer [53]. Depending on the complexity of the task, the agent may utilize a single tool, chain multiple tools in sequence, or interact with other agents. Agents can function as single-agent systems, where one LLM handles all processing, or as multi-agent systems, where specialized agents parallelize processes and collaborate on complex tasks [53, 56, 58]. In practice, maximizing the capabilities of each individual agent before adding more agents improves efficiency and reduces complexity [53].

Tool Calling, Action Planning, & Memory

For autonomous agents, the capabilities through tool calling, memory, and action planning enable agent to adapt to complex tasks by leveraging external data, maintaining context over interactions, and planning multi-step actions. Tool calling allows the agent to invoke external functions, such as accessing databases or APIs, to gather relevant information during decision-making. Memory management ensures that information from prior steps is retained and applied as the task progresses, allowing the agent to make context-aware

decisions. Action planning involves selecting and sequencing tasks based on the agent’s current understanding, optimizing the workflow to achieve specified objectives. These capabilities collectively enable agents to perform complex reasoning and adaptive decision-making, enhancing their autonomy when handling diverse and unpredictable tasks. Integrating these features into an LLM-powered agent system as illustrated in Figure 10 increases allows an agent to respond to different requests without strictly following a pre-defined machine-workflow.

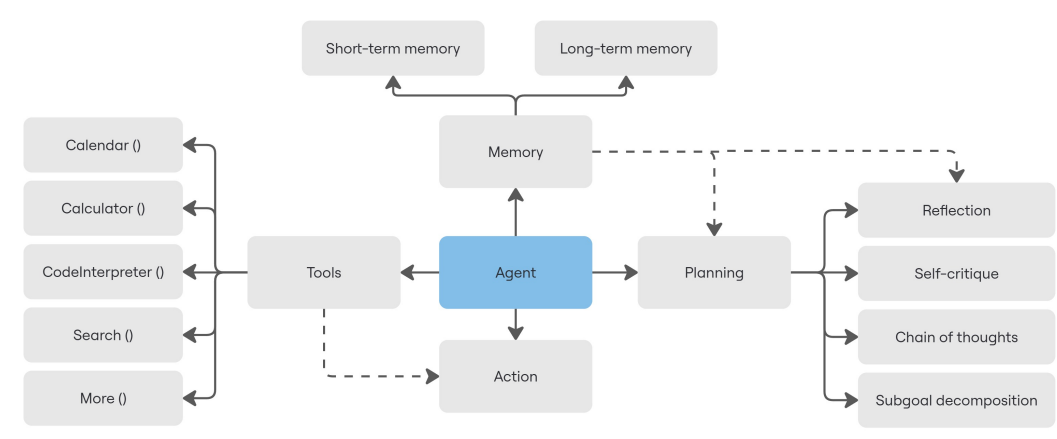


Figure 10: Example of individual LLM-powered autonomous agent system capable of action planning, and with access to memory and tools. Inspired from Weng et al. [54].

Prompt Engineering for Enhancing LLMs

Prompt engineering is the foundational method of interacting with LLMs to make them perform specific tasks without fine-tuning their internal parameters [43, 59]. Carefully curated prompts extend the capabilities of LLMs and enable them perform a range of tasks purely through input manipulation. Sahoo et al. [43] mapped the taxonomy of prompting techniques demonstrated in literature, finding a large amount of simple and advanced methods to facilitate structured decision-making in AI agent systems. While a range of methods exist, prompting can generally be done through the *system instructions* or through the *input query*, in which additional language-based techniques can be used within. Some of these techniques include providing context, defining a role, setting an objective, or describing the reasoning process to use [43, 59].

Prompt engineering is essential because it dictates *control*, *quality*, *efficiency*, and *adaptability* when using LLMs [59]. In practical applications, the quality of the prompt directly impacts the accuracy and reliability of the output, while also influencing how well the model adheres to desired tone and format (control). Efficient prompting not only minimizes errors and ensures accurate responses but also helps maintain conciseness and relevance, reducing verbosity and unnecessary information in the output (efficiency). Additionally, well-structured prompts enable the model to adapt its responses to new or domain-specific tasks, enhancing its flexibility (adaptability). Achieving these benefits often require considering the configuration of core parameters such as temperature, max tokens, and top-K/top-P, which directly influence how creative, concise, or diverse the model’s outputs are [59].

Choosing the appropriate prompting technique is important for achieving consistent and accurate outputs, especially when using LLMs in structured decision-making contexts [43, 59]. While techniques like zero-shot, few-shot, chain-of-thought (CoT), and structured output prompting, covered in Table 3, each serve distinct purposes. In practice, combining multiple techniques can enhance performance, such as using CoT for complex reasoning while employing structured output formatting to ensure machine-readability [43, 54, 59]. Selecting the right approach not only improves control over the model’s behavior but also reduces errors and inconsistencies [43, 59].

| Technique | Definition | Use Case |
|----------------------------------|--|---|
| Zero-Shot Prompting | Direct task and instruction without examples. | Suitable for simple tasks where the model can reason effectively from a single prompt and its reasoning capability. |
| Few-Shot Prompting | Provides a small number of examples within the prompt | Useful for enhancing output consistency by explicitly defining the desired response. |
| Chain-of-Thought (CoT) Prompting | Guides the model through step-by-step reasoning to improve accuracy and transparency. | Suitable for complex tasks requiring logical progression. |
| Structured Output Prompting | Ensures the response follows a predefined format, such as JSON, for machine-readability. | Applied when outputs must be easily parsed or integrated into machine-workflows. |

Table 3: Core Prompting Techniques [43, 59]

Prompt engineering is essential for leveraging LLMs effectively in structured decision-making contexts. By carefully selecting and combining techniques like zero-shot, few-shot, chain-of-thought, and structured output prompting, it becomes possible to balance control, accuracy, and efficiency when guiding AI behavior. While the choice of technique largely depends on the complexity of the task and the desired output format, adopting a systematic approach to prompting enhances both the reliability and practical utility of AI-driven workflows.

3.7 NLP Application in Land Management

The land management field is a regulated domain characterized by extensive documentation that can be challenging to navigate. Decisions are often grounded in regulatory statutes, from which related documents and registered information emerge. When a decision needs to be made, it often involves multiple related data sources, such as municipality plans, local plans, and property information. LLMs offer a way to interpret such texts by leveraging their natural language understanding. However, since decisions carry high stakes and require factual accuracy, the inherent tendency to produce plausible-but-wrong statements must be addressed by using techniques that augments and enhances the reasoning capabilities of LLMs.

While domain-specific adaptation is often considered optimal, it requires large, high-quality, and domain-specific datasets, which can be challenging to collect, curate, and process. However, studies have demonstrated that prompt engineering can effectively extend the capabilities of LLMs, enabling them to perform specific tasks efficiently without fine-tuning. By clearly defining a task and its boundaries, prompt engineering can guide LLMs to focus on specific information while following a structured reasoning process, which can notably enhance the accuracy of its outputs. Moreover, tailored prompts allows for adjusting the model’s behavior, ensuring that outputs align with domain-specific requirements and maintain consistency.

Agent systems make it possible to combine natural language understanding with traditional software workflows, bridging the gap between human-like interpretation and structured, rule-based processing. By leveraging agents, LLMs can interact seamlessly with external tools, databases, and APIs, enabling dynamic information retrieval and validation. This modular setup not only enhances the accuracy of outputs but also ensures that LLM-driven systems can be effectively integrated into existing land management processes. For instance, an agent can query cadastral databases to verify property data and accurately determine ownership. This combination of natural language capabilities with structured tool usage significantly improves the robustness and precision of land management tasks.

While each of the NLP and AI techniques discussed above offers specific functionalities, their relevance to land management depends heavily on the nature of the task. Setting up these AI workflows require careful tailoring to adapt to the specific application, whether it is for answering questions, querying cadastral databases, or relating a case to past cases, to name a few. The challenge lies in defining the domain territory effectively and identifying the most suitable combination of AI techniques to match the scope of the given task. Table 4 synthesizes how RAG, embeddings, and multi-agent workflows align with typical challenges in land management.

| Method | Strengths | Limitations | Land Management Use Cases Examples |
|-------------------------------|---|--|--|
| Prompt Engineering | Extends model capabilities, allows task-specific adaptation, reduces error rates | Requires careful design to avoid ambiguous or misleading prompts | Summarizing legal texts, structuring responses, etc. |
| Tool-Calling with LLMs | Connects NLP with conventional data processing | Requires setup and coordination with agent and LLM interpretation | Accessing cadastral databases, calling on data sources, running geospatial scripts, etc. |
| RAG | Ensures factual accuracy, dynamically retrieves verified information, and adapts well to changing data contexts | Retrieval quality is highly dependent on embedding quality and knowledge base structure | Retrieving past cases, finding legal clauses, etc. |
| Embeddings | Enables semantic search and contextual matching | Requires careful domain adaptation; cannot perform reasoning like LLMs | Performing similarity searches for past cases, textual classification of terms, etc. |
| Agent Systems | Serve as the orchestration layer for embedding models, RAG pipelines, and external tools; enables modular reasoning and scalable decision workflows | Require custom development and coordination of agents; can be computationally expensive at scale | Complex processing, rule-based decision making, system-level automation, etc. |

Table 4: Synthesis of AI techniques and their alignment with land management tasks as found throughout the chapter.

4 Methodology

This chapter describes the iterative methodology used to design and develop a geospatial AI-agent workflow aimed at supporting Danish municipalities in handling §3 basin dispensation applications and streamlining processes for applicants like Vejdirektoratet. The development was shaped by ongoing insights that culminated in a decision-support tool designed for factual screening. This tool orchestrates a sequence of sub-tools for information retrieval and geospatial analysis, presenting findings in a consolidated report where an instructed LLM summarizes results and interprets practical implications. To facilitate user interaction, these tools were integrated into an AI agent capable of executing analyses based on user queries and reasoning upon the results. The iterative approach ensured that the project was aligned with practical municipal needs. The subsequent sections will elaborate on each phase of this methodology.

4.1 Bridging Domain Context and NLP workflow

With the primary goal to reduce the burden on municipalities and streamline the process, initial research explored the potential for AI in automating decision-making aspects of basin dispensation application process. The development of the geospatial AI agent workflow followed an iterative design process (see Figure 11), which was shaped by; (1.) an initial understanding of the domain context, (2.) past dispensation cases, and (3.) NLP techniques, and (4.) three sets of interviews. Early interviews with Vejdirektoratet and their consultant Ejlskov A/S confirmed the value of standardized application formats, suggesting that an AI capable of generating structured information would align well with existing practices and be beneficial by easing the processing burden on municipalities.

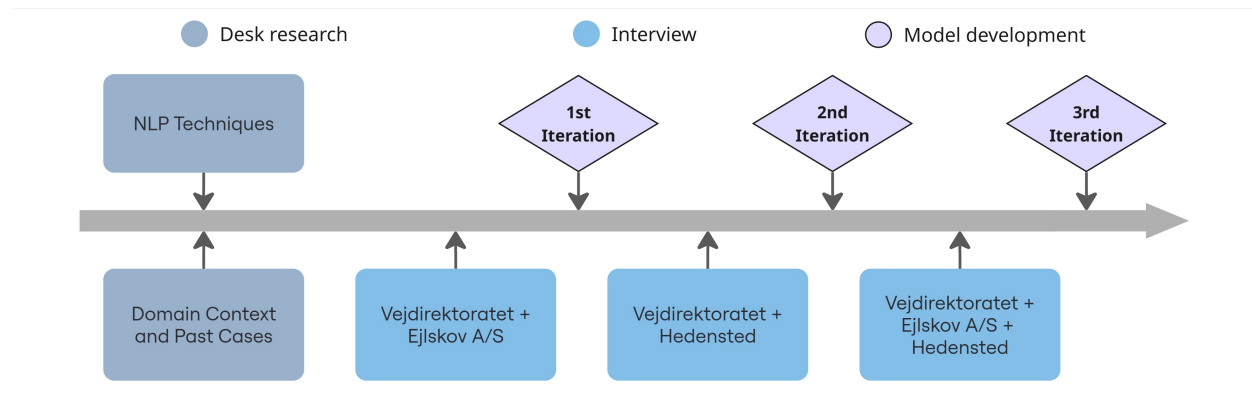


Figure 11: Timeline of the model development process throughout the project period.

However, interviews with Hedensted case workers emphasized the critical role of site visits for assessing site-specific nuances (e.g., current vegetation, unrecorded species), making a fully automated decision-making agent impractical. These discussions revealed the common municipal practice of GIS-based conflict assessment. Consequently, the project's focus shifted to augmenting this procedure with LLM-driven insights. Building on Hedensted's indication that municipalities would benefit from Vejdirektoratet's internal basin data, the model was refocused as a factual screening- and decision-support tool, presenting basin information and its regulatory relationships, using LLMs to identify preliminary legal and practical implications as derived from the data.

This core decision-support was primarily realized through a script capable of generating a report by executing a sequence of geospatial analyses of basins against key regulatory layers, which mimicked existing methods. Going beyond the existing format of conflict assessment reports, the factual findings from these tools were then synthesized by an LLM instructed to summarize the results and offer practical takeaways by highlighting regulatory and practical implications. The overall agent architecture (detailed in Section subsection 4.4) facilitated interaction with the model by enabling a user to invoke these individual analysis tools or request the full report through a chatbot interface, with the agent managing the tools and answering based on the results of those.

4.2 Data Collection

4.2.1 Interviews

To gain a practical understanding of the §3 dispensation process for rainwater basins and to inform the agent’s design, three sets of semi-structured interviews were conducted. The overarching purpose was to clarify existing workflows, identify essential information requirements for decision-making, understand the varying considerations of each involved actor, and gather insights for refining the generated report. Participants included Vejdirektoratet, Ejlskov A/S, and environmental and nature case workers from Hedensted Municipality.

The interviews were conducted using a semi-structured approach, allowing for exploration of emergent themes while covering predetermined topics. Guiding questions were prepared beforehand for the first two interviews (see Appendix B.1 and Appendix B.2 for interview questions for the first and second interview), which with participant consent were audio-recorded and transcribed to enable qualitative analysis (see Appendix B.3 and Appendix B.4 for transcriptions). A third, untranscribed follow-up interview served as a discussion focused on refining the generated report and its proposed takeaways. Each interview session had a specific focus:

1. **Interview 1 (with Vejdirektoratet & Ejlskov A/S):** The primary focus of this session was to understand the applicant’s perspective. This included detailing their standard procedures for preparing §3 applications, the typical content and structure of these submissions, their data sources (including internal Vejdirektoratet data), common challenges, and their interaction and experience with municipalities.
2. **Interview 2 (with Hedensted Kommune & Vejdirektoratet):** This session centered on the municipal decision-making process. Key discussion points included their workflow upon receiving an application, essential information requirements for their assessment (including their use of GIS for conflict screening), the role and necessity of site visits in their evaluation, common complexities encountered, and their opinions on how AI could potentially support or streamline their tasks.
3. **Interview 3 (Follow-up discussion with key stakeholders):** This untranscribed session was aimed at iteratively refining the agent’s proposed outputs. The discussion focused on the structure and content of the AI-generated summary insights and the overall PDF report format to increase its practical utility for municipal case handlers, building upon earlier feedback and model development progress.

Several key findings from these interviews directly shaped the methodological approach and system design. The insights gained provided practical guidance on data selection, defining the AI's role, and refining outputs to align with municipal needs. The four most significant impacts are outlined below:

1. **Validation of Data Sources:** The interviews, particularly with Hedensted Kommune, confirmed the relevance and municipal use of the specific digital geospatial datasets (from Danmarks Miljøportal, including §3 protected areas, Natura 2000 sites, Annex IV species occurrences, and VP3 streams) intended for the project's conflict screening tools. This validated their selection as primary inputs.
2. **Value of Standardized Application Formats:** Insights from both Vejdirektoratet/Ejlskov A/S and Hedensted highlighted the value of standardized application formats, with the former noting their submissions often streamline municipal processing and the latter confirming the utility of receiving Vejdirektoratet's internal basin data. This reinforced the design goal for the AI to produce similarly structured and informative outputs.
3. **Defining the Role of AI:** A key insight from the municipal perspective emphasized the need for site visits to complement digital data, addressing factors like vegetation state, possible species, and access issues. While Vejdirektoratet desired process streamlining, it became evident that purely remote assessment had limitations. Consequently, the AI's role was narrowed from full automation to consolidating digital data to aid human case handlers.
4. **Iterative Refinement of Model Results:** Feedback from Hedensted Municipality, especially regarding conflict screening needs, guided improvements to the agent's outputs. The focus was on enhancing usability within municipal workflows by refining both the overall model and the LLM's instruction set. These adjustments aimed to increase the relevance of the PDF report, ensuring it directly supports municipal decision-making processes.

In summary, the interviews were central to validating the project's premise, refining its scope towards realistic and valuable automation (factual screening), confirming data requirements, and guiding the functional design of the agent and its specialized tools to effectively support the specific dispensation workflow.

4.2.2 Official Document Cases

To gain insight into real-world practices, a set of data containing past dispensation cases was obtained from Vejdirektoratet. This data comprised both the initial applications submitted (primarily by the external consultant, Ejlskov A/S) and the resulting permits issued by 56 different municipalities across Denmark. A systematic manual review and qualitative analysis of these documents were conducted, with the primary objectives being to:

- Understand the typical components and structure of dispensation applications and permits.
- Determine the core information requirements typically needed by municipalities.
- Identify common legislative references and procedures.
- Characterize patterns and the variations in the final dispensation conditions.

This qualitative analysis uncovered the formal reasoning of municipalities, revealing wide standardization of specific conditions of final municipal dispensations, but also often diverging in regards to localized factors and site-specific assessments. This finding reinforced the decision, also informed by stakeholder interviews, to focus the AI agent on factual screening rather than attempting to automate the drafting of final decision documents. The understanding derived from past cases was therefore foundational in defining the agent’s requirements for performing objective screening, thereby providing a consistent information basis to support municipal decision-making.

4.2.3 Geospatial Data

To enable spatial reasoning, several key geospatial datasets representing various spatial legal considerations were retrieved from Danmarks Miljøportal, including layers for §3-protected nature areas, Natura 2000 sites, recorded locations of Habitat Directive Annex IV species, streams classified under the third Water Framework Directive plan (VP3 streams), and coastal water catchment boundaries. The core basin location and attribute data, containing Vejdirektoratet’s specific information on basin status, type, condition, and access notes, was provided directly by Vejdirektoratet. These specific layers were selected as they were directly relevant to the municipal assessments and spatial checks regarding environmental impacts and regulatory constraints when evaluating basin dispensation applications.

This collected geospatial data formed the foundation for the agent’s spatial analysis capabilities. Importantly, beyond geographic location and shape, these datasets contained valuable attribute information that provided essential context. For instance, the regulatory layers include details such as protection type names, observed species, or stream characteristics, while the basin layer holds Vejdirektoratet’s operational data. This integration of geometric and attribute data allows the agent, through its specialized tools (detailed in Section subsection 4.5), to perform spatial and attribute-related assessments, including checks for overlap and proximity analysis as well as retrieval of specific attributes from relevant features, hereby informing and supporting the overall dispensation handling process.

4.3 Software

The development of the geospatial AI agent leveraged a combination of established and specialized Python libraries. The core agent framework was constructed using LangChain, which provided functionalities for agent creation, tool integration, and managing interaction between the LLM and the custom tools. Access to the underlying LLMs was facilitated via the OpenAI API. Geospatial data processing and analysis relied on GeoPandas and Pandas. The user-facing web interface was developed with Streamlit. For output generation, Matplotlib was used for creating map visualizations, OWSLib for accessing web map services (e.g., orthophoto basemaps), Pillow for image processing tasks, and ReportLab for compiling the final PDF reports.

4.4 Core Agent Architecture

The model was built by integrating an *Agent Executor* from LangChain with GPT-4. The LLM was chosen based on its appropriate reasoning capabilities to interact with the user and the custom toolset developed and explained in the following sections. The *Agent Executor* was made to act as the central orchestrator, capable

of planning as well as calling and receiving information from the specialized tools. This orchestrator used the LLM, which acted as the brain within this framework, capable of interpreting user input, choosing appropriate tools, and generating natural language outputs for decision-making support. Figure 12 illustrates this overall system architecture, depicting the relationships and flow of information between the user interface, the agent executor, the LLM, and the tools.

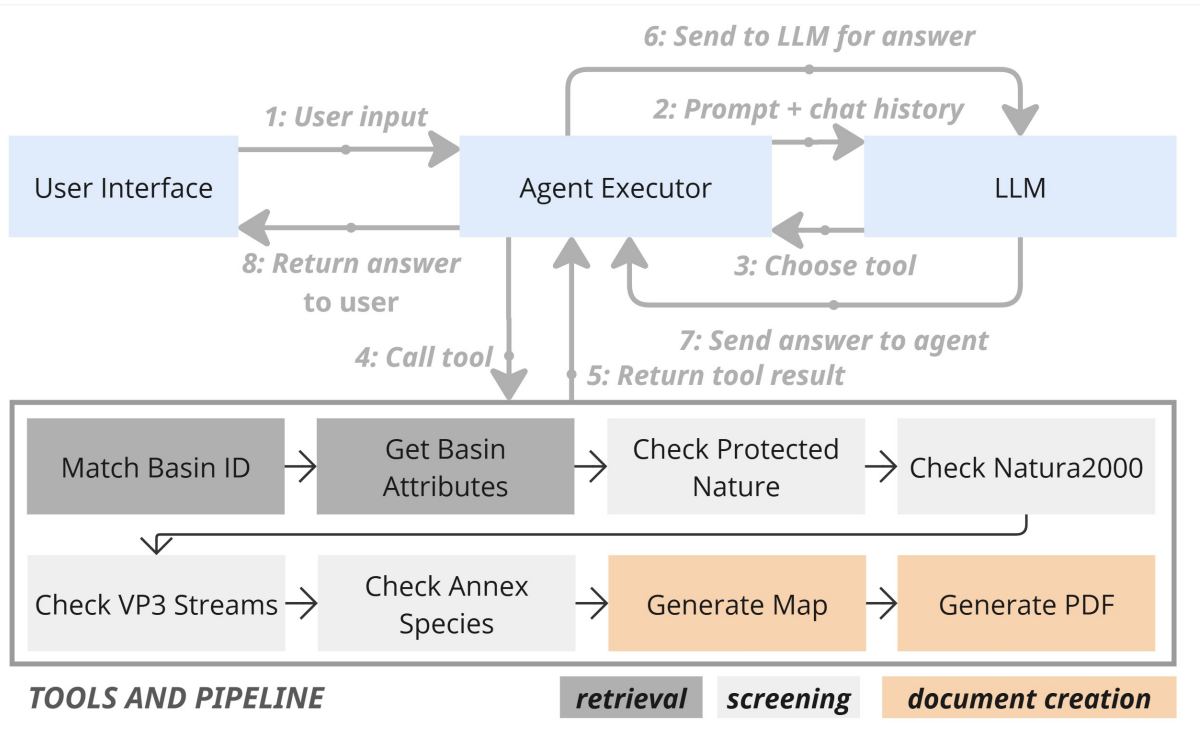


Figure 12: System architecture diagram showing data sources, agent tools, and the core agent system with key interaction steps numbered.

4.4.1 Modularity

The system architecture was explicitly designed with modularity as a core principle. It consisted of eight individual scripts as tools, which were able to be run in sequence as a pipeline while also able to be used by the agent. A separate script defined these tools along with descriptions to enable the agent to use them, and a final script served as the main application, able to run a web-interface interacting with the agent capable of executing the tools. The modular approach informed and supported development, testing, iterative extension, and overall maintainability of the code. The primary motivation was to segment the application logic into manageable, distinct components, which allowed for easier troubleshooting and a "plug and play" approach to adding or modifying functionality.

Within this modular framework, a dedicated directory was created for the *agent tools*, where each Python script encapsulated the logic for a specific, self-contained task (e.g., matching basin identifiers, checking Natura 2000 overlap, generating a map, and so on). While individual scripts might contain helper functions, each centered around a main callable function representing the distinct tool capability. This

architectural choice effectively decoupled the specific implementation logic of each task from the overarching agent framework. Consequently, individual tool scripts could be developed, tested, debugged, and refined largely independently, enhancing the system’s robustness and simplifying maintenance.

4.4.2 Agent-Tool Connection

The connection between the tools and the central LangChain agent was defined in a separate Python file from the individual tools and the main application. In this script, each tool’s main function was wrapped using LangChain’s `BaseTool` class. The class for each tool contained tool names and natural language descriptions of the tools. Importantly, the expected input arguments to execute each tool were also described in Pydantic `BaseModel` schemas, which were linked to the `BaseTool` descriptions. To illustrate these definitions, the following box provides an example of the `BaseTool` class and associated `BaseModel` input schema for the tool checking the status of protected nature (see Appendix C.1 for exact tool descriptions).

Check Protected Nature Tool

*Checks whether a specific basin is registered as §3 protected nature and whether it is nearby other nearby §3 nature within a given buffer distance. Returns a JSON result with the overlap status, identified nature types, and the total area. **Input Arguments:***

- **basin_identifier** (required): The exact basin identifier.
- **kommune** (required): The exact municipality name.
- **buffer_distance** (optional): Distance in meters (default is 100 m).

This structured approach, combining natural language tool descriptions with formal input schema, ensured that the LLM-powered agent could precisely understand each available tool, capabilities, and how to use it. This, in turn, enabled the agent to select the appropriate tool(s) and determine the correct arguments to pass for successful execution during a task. Finally, all defined tools were aggregated into a list, which was subsequently passed to the `AgentExecutor` in the main application, making them available for the agent to utilize.

4.4.3 Workflow Control

While the agent tools were developed to run in sequence, the agent web-interface was set up to allow for dynamic interaction with the tools and the information contained in the spatial datasets. The dynamic operation and control of the analysis workflow was managed by the LangChain `AgentExecutor`, which integrated the capabilities of the GPT-4 LLM. While more and less advanced models were tested (GPT 3.5-Turbo and GPT-4o), the GPT-4 model was selected due to its sufficient reasoning capability for the task and its proficiency in interpreting user requests and relate them to tool execution. A temperature setting of 0.8 was configured for the LLM, which slightly restrained the creative deviations while still allowing for natural variation in conversational responses.

The LLM was provided with a set of system instructions, which established the agent’s overall role

as a specialized assistant for basin analysis. These instructions in the prompt below allowed the LLM considerable autonomy in selecting the appropriate tool based on its interpretation of the user's query and the detailed descriptions of each tool's function and arguments. While the system instructions for the agent provided an opportunity to explicitly define and explain its role in detail, control was primarily guided by the agent's understanding derived from the tool descriptions defined within the `BaseTool` wrappers. The omitted tool elaboration within the system instructions was done to avoid excessive instructions and token counts, simplifying the instruction-set as a minimal.

System Instructions for Agent

You are a specialized assistant for the analysis of Danish rainwater basins. Your goal is to answer the user's questions using the available tools. Users may speak Danish or English, and you should always respond in the same language as the user.

To enable the agent to maintain the conversational context, LangChain was used to provide a method for handling chat history. This enabled multi-turn interactions, allowing users, for example, to ask for an analysis of specific basins identified in a previous turn (e.g., "analyze basin number 3 from the list") without needing to restate the full technical identifier. The agent leveraged this memory to ensure the correct, validated identifiers from the 'Match Basin Tool'. Initial testing revealed instances where the LLM attempted to use shortened identifiers, passing invalid basin specifications to other tools, leading to errors. For this reason, the descriptions in the input arguments as defined in the `BaseModel` schemas were tightened, as exemplified in the example above.

4.5 Agent Tools

A range of specialized tools was developed to support municipal screening by enabling the LLM agent to analyze basin status and regulatory relationships, produce maps, and generate PDF reports with overviews and insights. These tools provided deterministic data retrieval and analysis capabilities rooted in geospatial and attribute data to ground the LLM in factual outputs. Each tool's core logic is contained within separate Python scripts, each integrated into the agent framework via the standardized `BaseTool` and `BaseModel` definitions. This interface provided the agent with the necessary understanding of each tool's function and required inputs. The developed toolset encompassed three main categories of functionality, detailed in the following subsections: information retrieval, geospatial screenings, and document creation.

4.5.1 Information Retrieval

The first step in a basin analysis was to correctly identify the specific basin in question (approximately 1600 basins) and then gather the registered information about it. This initial set of tools was therefore designed to handle these foundational tasks. The first tool in this category helped the user and the AI agent pinpoint the exact basin even with incomplete or slightly varied input (like a partial name or ID). Once a basin was accurately identified, the second tool retrieved descriptive information about it from the basin data, such as its visit history, registered condition, and access notes. This information provided essential context for all subsequent analyses and for the final summary report.

Match Basin Identifier

A primary challenge was ensuring precise basin identification due to the technical nature and format variations of official Vejdirektoratet identifiers⁵. The Match Basin Identifier tool was therefore developed to bridge potentially ambiguous user input, such as partial basin IDs or municipality names, with the exact technical identifiers required by other tools. This tool validated user queries against an authoritative master list (a CSV file), which was an essential first step for any reliable downstream processing in the workflow, as illustrated in Figure 13.

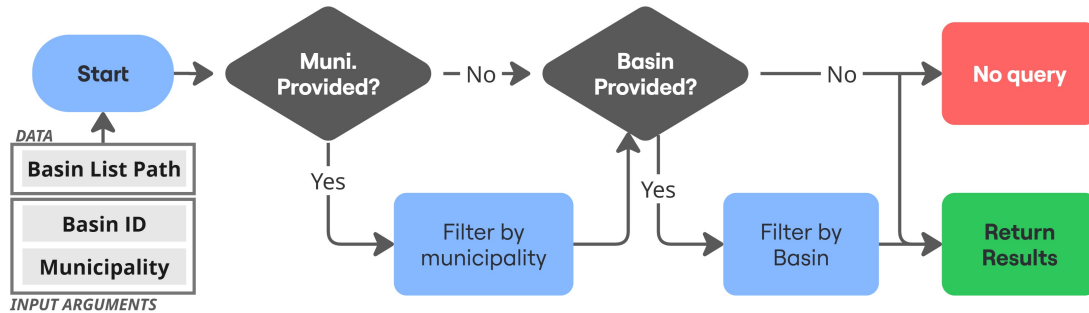


Figure 13: Simplified visualization of match basin logic to match and retrieve basin identifiers.

Functionally, the tool employed a normalization process on user query terms and master list data (e.g., converting to lowercase, stripping whitespace, transliterating Danish special characters) to ensure consistent comparison. It then utilized Pandas filtering for case-insensitive substring matching to identify potential basins. The tool was designed to handle various outcomes, such as exact matches, multiple possibilities, or no matches, returning a structured JSON dictionary detailing the status and a list of precise identifiers for the found basins. The tool returned a list of basins found as shown in the example below, illustrating what the LLM sees.

Match Basin Result: Used for Targeting Specific Basins (Runtime: 0.11 seconds)

```

"data": {
  "status": "found_multiple",
  "results": [
    {"id": "20-0 43/0563 Højre", "municipality": "Køge"},
    {"id": "20-0 44/0753 Højre", "municipality": "Køge"},
    {"id": "20-0 45/0665 Højre", "municipality": "Køge"},
    {"id": "10-0 37/0140 Højre", "municipality": "Køge"},
    {"id": "...", "municipality": "Køge"}
  ],
  "message": "Found 21 possible basins. Please specify."
}

```

⁵Examples include: 'x-x xx/xxxx H/V', 'xx-xx xxx/xxxx H/V', 'xxxxxx-x x/xxxx H/V', and occasionally names.

Get Basin Attributes

Following precise basin identification, the next tool was developed to retrieve relevant non-spatial meta-data. Its primary purpose was to gather contextual information regarding the basin's history and registered condition, utilizing data recorded by Vejdirektoratet during previous assessments or maintenance. This information, often not included in initial municipal applications but highlighted as valuable during interviews, included details like maintenance dates, registered comments, basin type, and vegetation levels. The tool's workflow, illustrated in Figure 14, required a correct basin identifier and municipality name to ensure successful execution.

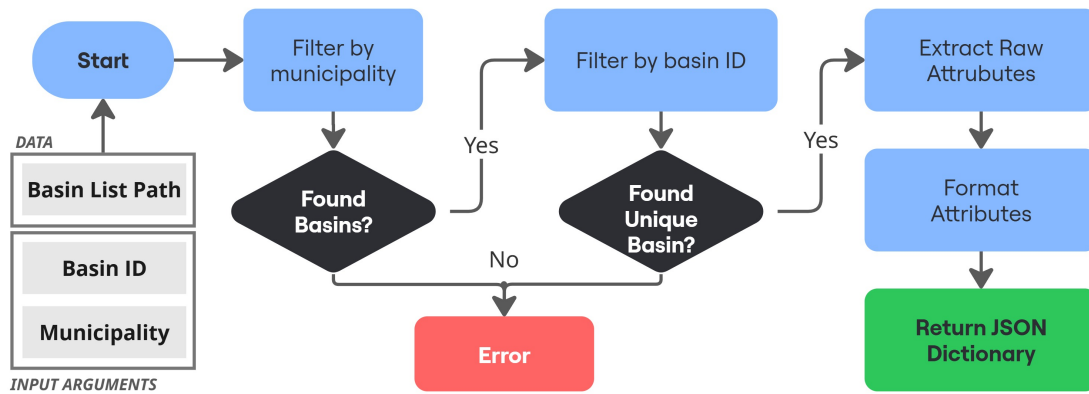


Figure 14: Visualization of the logic used for retrieving attributes from basin data.

The tool used GeoPandas to read Vejdirektoratet's basin GeoPackage file and filter it to isolate the specific basin record. From this record, a predefined list of attributes considered useful for dispensation cases was extracted, despite some of the data being potentially or poorly maintained. Formatting was applied for consistency (e.g., dates, missing data represented as "No data registered"), and the extracted and formatted attributes were then compiled into a structured JSON dictionary, which presented the data in the below example output that is shown to the LLM.

Basin Attribute Results: Retrieved Operational Basin Details & History (Runtime: 0.19)

```

"data": {
  "last_cleaning": "3/1/1976",
  "basin_type": "Lake",
  "general_assessment": "Functioning",
  "access": "Direct access",
  "inlet": "Ok",
  "bank_covered_with_bushes": "50%",
  "bottom_covered_with_trees": "10%",
  // more data (shortened for brevity) ...

```

4.5.2 Geospatial Screenings

Following basin identification and attribute retrieval, a set of specialized tools was developed to perform automated geospatial analyses to assess the basin’s spatial relationship with key environmental and nature protection regulatory layers. These screenings utilized Python with the GeoPandas library to execute standard spatial query operations, such as intersection, buffering, connectivity analysis, and proximity calculations. The analyses were conducted against authoritative national datasets, with each tool designed to return structured, objective findings in a JSON format, intended to form a crucial part of the factual basis required for municipal assessment of the dispensation application. This output provided a consistent, location-specific factual baseline intended to inform the municipal assessment, support and expedite the case handling process by providing easily interpretable data suitable for further processing or summarization by the LLM agent.

Protected Nature Tool

To screen for §3 protected nature conflicts, this tool assessed the spatial relationship between a target basin and registered §3 areas. Its primary objectives was to determine if the basin directly overlaps a §3 area (implying a need for dispensation under the Nature Protection Act) and to evaluate its connectivity to nearby §3 areas as an indicator of the surrounding ecological context relevant to the basin’s status. The analysis relied on Vejdirektoratet’s basin dataset and the national §3 protected nature layer, and implemented spatial filtering, loading data only in a 5 km radius for performance. Key analytical steps, illustrated in Figure 15, include direct overlap checks and connectivity assessments.

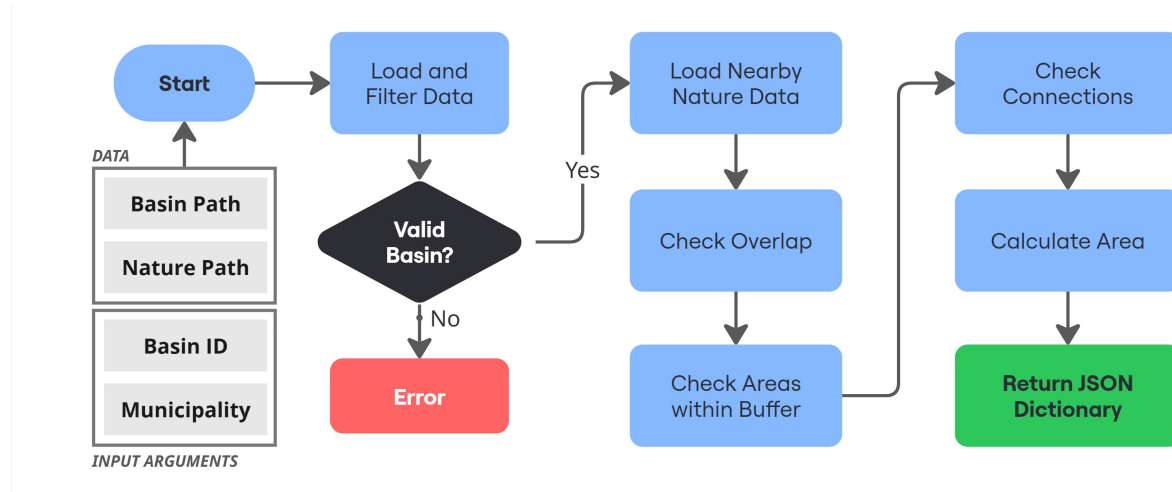


Figure 15: Workflow for determining basin relationship to protected nature.

The tool’s findings were summarized in a structured JSON dictionary for the agent. This included a boolean status ('yes'/'no') for direct §3 protection of the basin itself and a determination of its spatial connection to nearby protected nature. If any §3 polygons were found, key attributes such as unique nature type names and the total connected area was extracted and passed in the result. This provided a factual basis regarding the basin’s §3 status, ecological connectivity, and the nature types involved, supporting subsequent reasoning or report inclusion. An example of the format is shown below of what the LLM sees.

Protected Nature Result: Determined Protection Status & Ecological Connectivity (Runtime: 0.21 seconds)

```
"data": {
  "registered_as_protected": "yes",
  "in_conjunction_with_protected_nature": "yes",
  "nature_type": [
    "Lake",
    "Meadow"
  ],
  "area": "1582 m2"
}
```

Natura 2000 Tool

To assess potential conflicts with internationally designated conservation areas, a tool was developed to compare the relationship between specific basins and Natura 2000 sites, encompassing habitat- and bird-protected areas. The screening was important, as activities within or potentially affecting these areas are subject to stringent assessment requirements under the EU Habitats Directive and associated Danish legislation. The tool utilized the primary basin dataset along with the official Danish polygon layers representing designated Natura 2000 habitat- and bird-protected areas. Following basin identification, relevant Natura 2000 polygons within a 5 km radius were loaded, ensuring efficient computation.

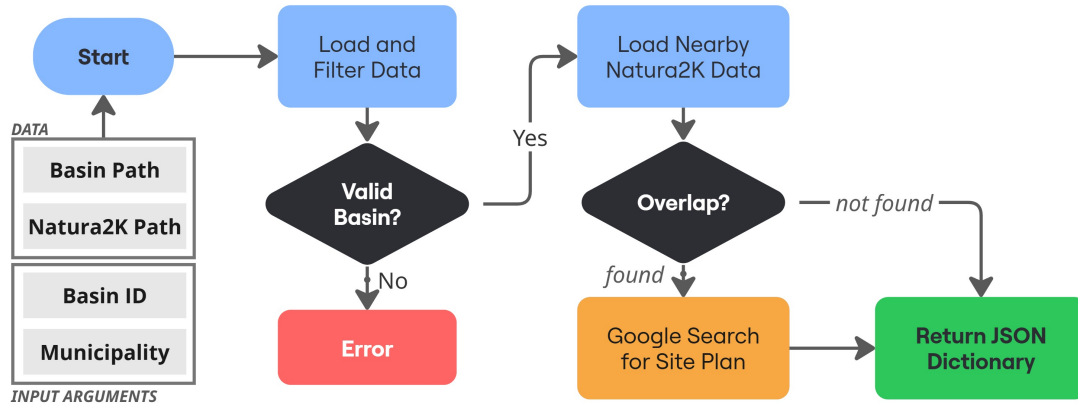


Figure 16: Workflow for checking basin relationship to Natura 2000 sites.

The tool generated a structured JSON output specifying the basin's Natura 2000 status as 'overlap', 'nearby', or 'none'. If overlap or proximity was detected, the output included the Natura 2000 area type(s) and official names. To facilitate quick access to additional contextual information, a Google Custom Search API was implemented to find the URL to the official management plan. As no centralized access point was available, the search used: the site name, a standard phrase, and a filetype filter for targeting PDFs. The query followed the format: "{site_name} Natura 2000 plan 2022 2027 filetype:pdf", returning the most relevant result among the retrieved documents. This setup enabled the agent to deliver accurate status, site details, and a direct link to planning documents. An example of the result is shown below.

Natura 2000 Result: Determined Protection Status & Found Site Documents (Runtime: 0.67 seconds)

```
"data": {
  "status": "overlap",
  "types": [
    "habitat area",
    "bird area"
  ],
  "name": "Vestamager og havet syd for",
  "url": "https://edit.mst.dk/media/icddmblm/n143-natura-2000-plan-2022-27-\vestamager-og-havet-syd-for.pdf"
```

VP3 Streams Tool

To evaluate potential hydrological connections and associated regulatory implications, a tool was developed to assess the relationship between specific basins and VP3-classified streams. These VP3 streams are subject to specific environmental quality objectives, making it relevant to assess potential impacts from basin discharge. A default search distance of 200 m was used, and if a stream was found, the tool identified the larger coastal water catchment into which the stream drains, providing broader hydrological context. The analysis utilized the primary basin dataset, the official VP3 stream layer, and the coastal water catchment layer.

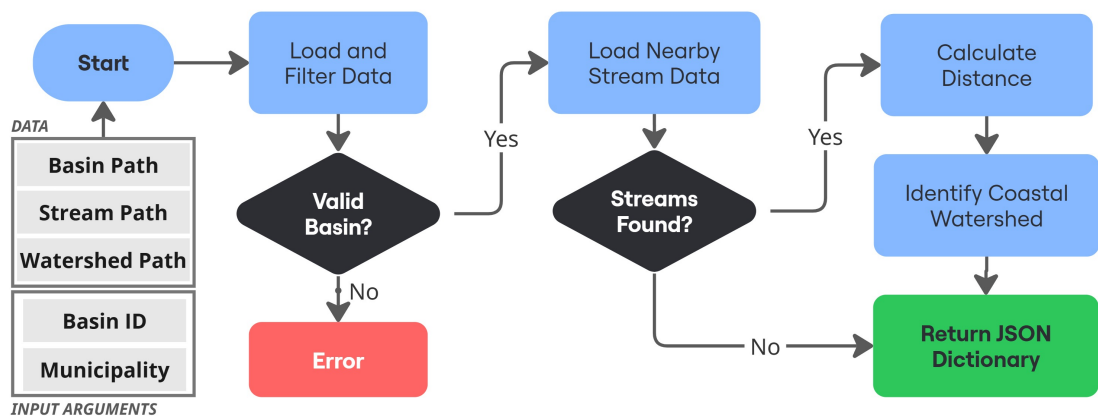


Figure 17: Workflow of determining basin relationship to streams.

The tool's findings were delivered as a structured JSON output, which indicated whether VP3 streams were found near the basin ('yes'/'no'). If a stream was found, the JSON included the distance to the closest segment along with detailed attributes for that segment from the VP3 dataset (name, VP3 category, type, ecological/chemical status, management objectives). If a stream was found, the coastal water catchment was also identified and provided. The output informed the agent about potential basin-stream interactions and the environmental context, particularly relevant due to stringent requirements for protecting streams from polluted discharges found in earlier dispensations. An example of the result is shown below.

Stream Result: Assessed Hydrological Connections & Catchment Context (Runtime: 0.18 seconds)

```
"data": {
  "stream_distance": "4 m",
  "stream_name": "Unknown",
  "stream_category": "Natural",
  "stream_type": "RW1",
  "stream_ecological_status": "Poor ecological status",
  "stream_ecological_goal": "Good ecological status",
  (shortened for brevity) ...
  "coastal_catchment_name": "Åbne vandomr. Gr. VI Øresund og Køge Bugt og\
  Østersøen (201)"
}
```

Annex IV Species Tool

To ensure consideration of protected species, this tool was developed to screen for species listed under Annex IV of the EU Habitats Directive, which mandates stringent habitat protection. Identifying the potential presence of these species near a basin was deemed important, as earlier dispensation cases indicated that confirmed occurrences often necessitate specific conditions for basin cleaning. The analysis utilized Vejdirektoratet's basin dataset and a national point dataset of recorded Annex IV species observations. A default search distance of 2000 m around the basin was employed for the proximity analysis, chosen as a general approximation of the potential dispersal range of species, although these might vary significantly between different species. The final workflow is shown in Figure 18.

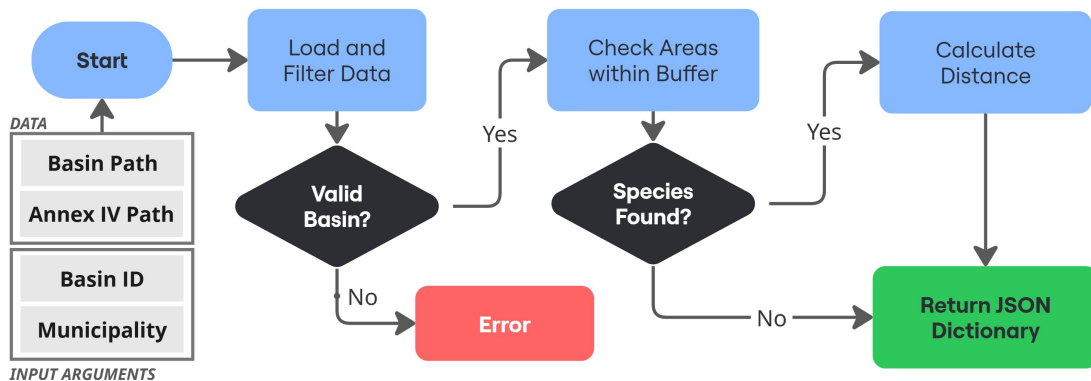


Figure 18: Workflow of the Annex IV Tool, outlining data processing and analysis steps.

The screening results were structured into a JSON dictionary for the agent. This output indicated whether Annex IV species observations were found within the search distance ('yes'/'no'). If observations were detected, the JSON provided a list of these, with each record containing the observation date, Danish species name, species group, and calculated distance to the basin. This list also served as an initial indication of whether a physical site visit might be required to assess potential impacts on species, ensuring alignment with Habitats Directive and Nature Protection Act requirements. An example of the result is shown below.

Annex IV Result: Screened for Nearby Protected Species Observations (Runtime: 0.21 seconds)

```
"data": {
  "search_distance": "2000 m",
  "species_found_within_distance": "yes",
  "observations": [
    {
      "observation_date": "2024-09-14",
      "species_name": "Sand lizard",
      "species_group": "Reptiles and amphibians",
      "distance": "124 m"
    },
    {
      "observation_date": "2024-09-14",
      "species_name": "Sand lizard",
      "species_group": "Reptiles and amphibians",
      "distance": "161 m",
    },
    (list continues)
```

4.5.3 Document Creation

To convert the technical data and spatial analysis results into accessible formats suitable for supporting municipal review and providing decision-making support, the final set of agent tools focused on document creation. The objective was to synthesize the outputs from the preceding information retrieval and geospatial screening tools into both visual representations and structured textual PDF reports. One tool was specifically developed to generate map visualizations, providing crucial spatial context. A second, more comprehensive tool was designed to produce a complete PDF report, integrating the map with LLM-generated summaries of all findings, offering a consolidated factual basis for the dispensation assessment along with the regulatory implications.

Generate Map

To provide important visual context for the assessed basin and its surroundings, a tool was developed to generate a map. The primary purpose was to create a clear cartographic representation showing the location and extent of the basin overlaid with key regulatory and environmental data layers analyzed by the geospatial screening tools, aiding interpretation and informing the dispensation decision process undertaken by municipal case handlers. The tool used the geospatial datasets from the screening tools, plotting them on the map extent. To supplement this, the most recently released official orthophoto, retrieved from Dataforsyningen's WMTS server via an API key, was chosen to visualize land cover, vegetation patterns, and terrain features around the basin.

Using Matplotlib for rendering, the tool plotted the vector layers over the orthophoto basemap. Specific styling choices were applied to ensure visual clarity and distinguishability between potentially overlapping

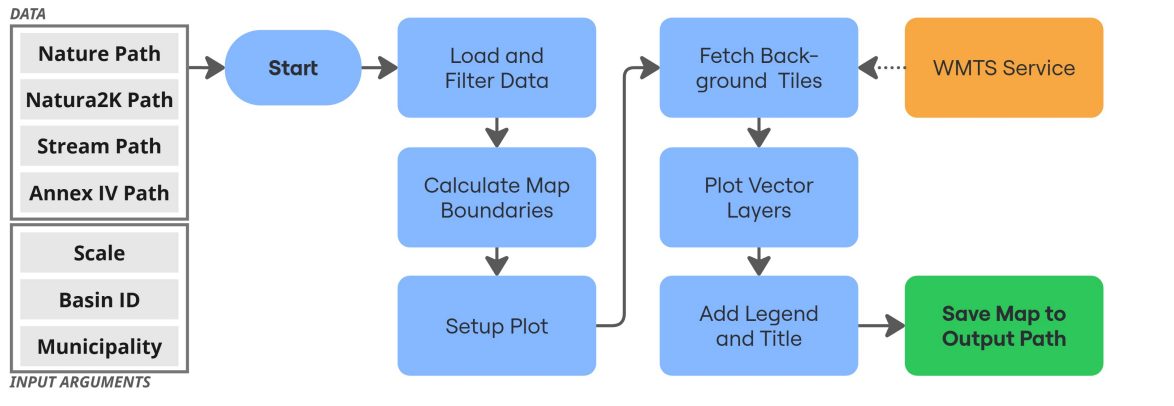


Figure 19: Simplified illustration of the processing steps for map creation.

layers, while still allowing the underlying orthophoto detail to be visible through the use of transparency. If the WMTS basemap service failed, the tool was designed to proceed with rendering the vector layers alone. A legend explaining the symbology for each layer and a title indicating the basin identifier, municipality, and map scale were included. The final map was generated and saved as a PNG file, a lightweight format suitable for web display and embedding within the PDF report. The tool returned the relative path to this generated PNG file to the agent executor.

PDF Report Generation

The final tool in the workflow was developed to combine the findings from previous tools and convey them in a structured and actionable PDF report. The report was designed as a decision-support output for municipal case handlers, offering a quick and factual overview to aid their assessment of §3 dispensation applications. It was set up to present regulatory zoning relationships and attribute information, and importantly, include a "Summary Insight" section to guide subsequent administrative steps. The internal workflow of the PDF generation tool, illustrated in Figure 20, sequentially executed the entire tool pipeline. The JSON outputs from the different tools were then aggregated into a combined JSON dictionary payload (see example in Appendix C.3), which was passed to an LLM - separate from the main agent orchestrator.

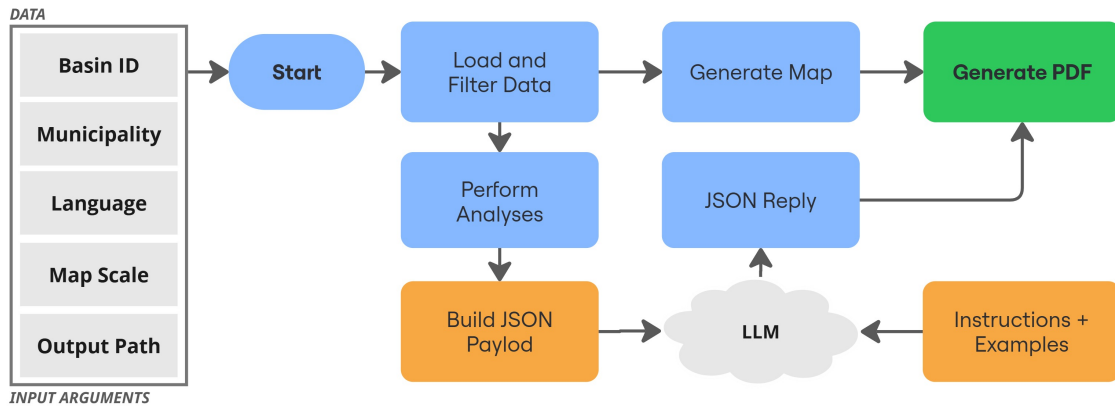


Figure 20: Simplified illustration of the model workflow for generating basin reports.

The LLM’s task of generating textual content was guided by the payload data alongside a detailed set of system instructions. These instructions carefully defined the LLM’s role as an assistant for municipal case handlers and its specific task with context while laying out explicit rules for output formatting, such as requiring valid JSON, fluent full sentences, and a professional tone. Importantly, they also included descriptive conditional logic through ”Guidelines for ’Summary Insights’,” which directed the LLM in interpreting various data combinations (e.g., basin protection status, maintenance history, Natura 2000 presence) to formulate actionable, regulatory-aware conclusions, including recommendations for site visits or specific dispensations. This was also supported by few-shot examples to clearly illustrate the desired output. Achieving reliable and consistent behavior from the LLM for this complex task necessitated a dedicated iterative refinement process across the system instructions

As an example of the challenges encountered, early instruction iterations led the LLM to generate non-JSON responses by adding explanatory comments or markdown formatting around the JSON. This caused errors, as the text failed to appear in the report. A similar issue arose with the main agent and tool descriptions, where the agent lazily abbreviated basin identifiers during tool invocation, leading to failed analyses. Both issues were resolved through iterative prompt refinements. Each iteration cycle involved generating reports, reviewing output, and adjusting prompts to be more explicit and unambiguous. This ultimately led to the instruction set described in Appendix C.2 and simplified in Figure 21. Feedback from municipal stakeholders was also crucial in aligning the LLM’s output with practical needs.

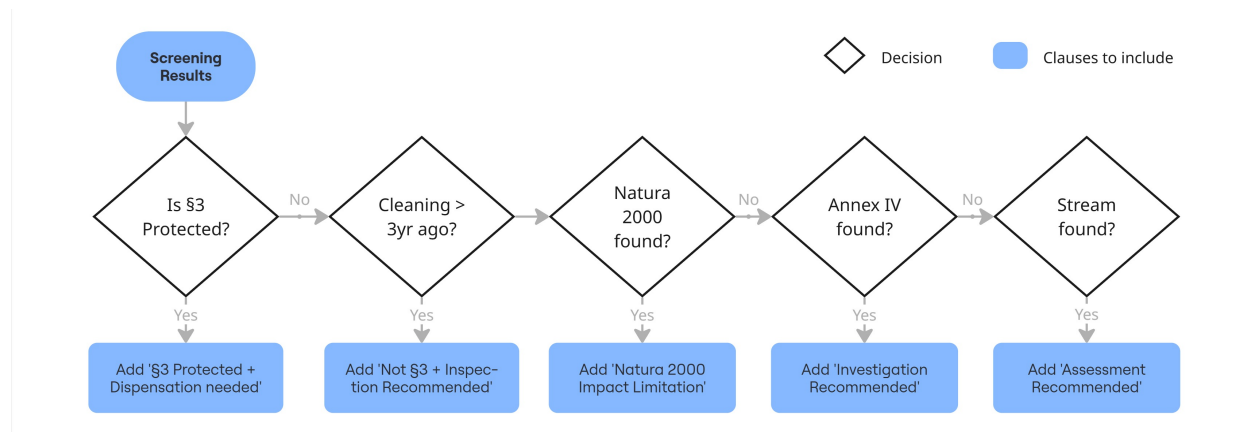


Figure 21: Visualized representation of conditional logic described in Appendix C.2.

For the summarization and interpretation logic, GPT-4.1 was selected due to its fast inference, low latency, cost-effectiveness, and appropriate reasoning capability for summarizing structured JSON dictionaries within the defined context. A temperature setting of 0.3 was configured to restrain creative deviations and ground the LLM’s responses in factual claims. The LLM was tasked with three primary objectives: **1.** generating concise summary results for each data category, **2.** concluding on the findings’ implications for dispensations and other approvals, and **3.** structuring its entire output as a single, precise JSON object. This JSON output, containing the LLM-generated titles and summaries for each report section, was then parsed by the ReportLab Python library. An example of a response from the LLM is shown on the next page. The total token count for the LLMs are shown in Table 5.

| Component | Reference | Token Count |
|---------------------------------|-----------------|-------------|
| <i>Chatbot Agent Context</i> | | |
| Agent system instructions | Instruction Set | 68 |
| Tool descriptions (all tools) | Appendix C.1 | 1,047 |
| <i>Report Generator Context</i> | | |
| LLM instructions (per report) | Appendix C.2 | 1,596 |
| Result payload (per report) | Appendix C.3 | 100–800 |

Table 5: Token usage breakdown across two distinct model contexts. The chatbot agent uses fixed agent and tool instructions, while the report generator receives its own instruction block along with dynamic, basin-specific data.

LLM Response Based on Payload

```

"Summary Insights": {
  "title": "Summary Insights",
  "summary": "The basin is registered as protected nature under §3, and a dispen..."
},
"Basin Information": {
  "title": "Basin Information",
  "summary": "The basin was last maintained on 3/1/1976. The general assessment is.."
},
"Protected Nature": {
  "title": "Protected Nature",
  "summary": "The basin is part of a larger natural area of at least 7331 m2, regi.."
},
"Annex IV Species": {
  "title": "Annex IV Species",
  "summary": "Annex IV species such as the Otter and the Great Diving Beetle have..."
},
"VP3 Streams": {
  "title": "VP3 Streams",
  "summary": "The basin is located 163 m from a natural RW2 stream with a high eco.."
}

```

The final PDF construction used ReportLab to assemble a standardized document. An internal helper function routed the JSON sections from the LLM to designated places within the report template. In total, the PDF contained a main title, a "Summary Insight" section with a programmatically constructed summarizing table, followed by the specific tool results presented sequentially, and the generated map on a subsequent page. ReportLab also facilitated the inclusion of design elements like logos and headers. Upon successful PDF creation, the tool returned the relative file path of the finished report to the agent executor, making the report available to the user.

4.6 Use of Generative AI in the Project

Generative AI platforms and models were central to the thesis project and served as digital assistants to support research and software development. Specific models were chosen for specific tasks based on their strengths. GPT-4o was used for most writing and research tasks due to its ability to handle most tasks appropriately. Grok 3 assisted in finding and discussing scientific studies due to it being trained in a vast amount of scientific literature. Gemini 2.5 and GPT-4o mini-high were used for model development due to their strong coding abilities. Finally, Claude 3.7 was used in visualization and figure creation due to its strong visual reasoning capabilities. The model selection for different tasks ensured outputs and sparring sessions yielded appropriate and useful support.

In the research and writing stages of the report, the *deep research* capabilities of Grok 3 and GPT-4o were used to explore the existing body of scientific literature relevant to NLP, AI agent architectures, and occasionally their relationship to land management. Deep research yielded tailored short-format reports based on literature with clear narratives for the tailored questions and objectives. This capability was often preceded by manual research as well as prior discussion with the LLMs in order to set the context. While extremely useful, the deep research capability often incorporated web-based articles, necessitating validation of sources, although the validation phase often lead to deeper insights gained from papers. This iterative research process can be seen in Figure 22.

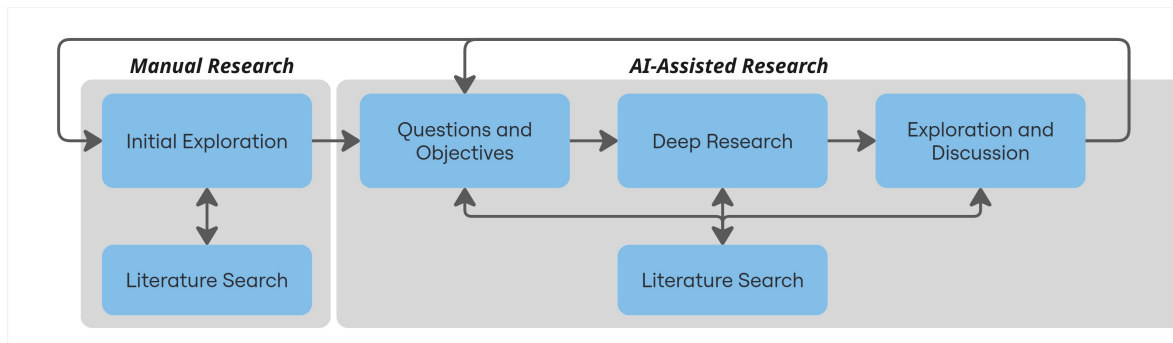


Figure 22: Illustration of the applied research process.

For the software development aspects of this project, models such as Gemini 2.5 Pro and GPT-4o mini-high were essential. These models not only understood agent architecture types but also helped develop them. Notably, they demonstrated a remarkable ability to produce functional scripts in a one-shot manner when provided with carefully crafted prompts that specified important inputs and outputs as well methodological steps and constraints. Beyond development, AI also supported the thesis writing process by providing feedback and drafting sections based on notes and outlines. Importantly, all AI-generated output was reviewed and validated for accuracy and appropriateness before being refined and included, ensuring that the work was critically analyzed and remained an authentic intellectual contribution.

5 Results: Model Assessment

This chapter presents the results from evaluating the developed geospatial AI-agent across its two primary functionalities, being interactive chatbot queries and the generation of structured basin reports. The evaluation examines how well the agent supports factual screening tasks relevant to §3 dispensations by testing its ability to interpret user intent, execute tool-based analysis, and produce clear, actionable outputs. The assessment is grounded in practical examples and focus on both interaction quality and the accuracy of regulatory-relevant findings. The following sections present these findings in two parts: first, the agent’s conversational performance, and second, the content and reliability of its generated reports.

5.1 Chatbot Interactions

The chatbot interactions are a core component of the model and its assessment, as they demonstrate the system’s ability to process user queries. To evaluate the chatbot’s performance, interactions are categorized into two types: simple and complex requests. Simple requests typically involve the agent performing one or two tool calls, with complex requests requiring the chatbot to perform a series of tool calls, which it must plan and execute effectively. By examining both types of interactions, the aim is to assess the chatbot’s versatility and reliability in handling diverse user queries. The performance is assessed based on criteria such as factual accuracy, adaptability to different queries, and the ability to provide a seamless user experience.

5.1.1 Simple Request

Simple requests are foundational interactions where the chatbot handles straightforward user queries. These types of interactions are crucial because they reflect common tasks that require the model to execute one or two tool calls only and are performed regularly. Since the model relies on using specific basin identifiers, most interactions in this workflow typically start by listing relevant basins based on the user’s query. An example of this interaction is shown below:

User: Can you make a list of basins in Hvidovre?

Chatbot: Sure, here are the basins located in Hvidovre:

- (1) 3-0 55/0307 Venstre
- (2) 3-0 58/0183 Højre
- (x) ...
- (10) 11-0 10/0424 Venstre

Let me know if you want information about any specific basin.

chat continues...

In the example, the agent returned with a list of basins, which were cross-checked and validated in the master list. It demonstrates the chatbot’s ability to handle simple requests by accurately retrieving a list of basins from a specified municipality. With the list of basins, the chatbot was prompted to randomly

select one basin from the list and present its registered attributes. The following example demonstrates this interaction:

User: Can you select a random one and tell me what we know about it from the registered data?

Chatbot: The selected basin is "3-0 55/0307 Venstre" located in Hvidovre. The registered information about it is as follows:

- Last cleaned: 1/2/2011
- Last cleaning of inlet/outlet: No data registered
- Last maintenance: 1/12/2018
- Last screening: 17-01-2018
- Type of basin: Lake
- General assessment: Functioning
- Access to the basin: Direct access
- Remark: Permanent deep water at the inlet = -30cm due to deep outlet. Water level in the basin rises 90cm above the outlet quota per 27/2. Sediment at the inlet 30-35cm. Close pump station for part of M3.
- Bank covered with bushes: 15%
- Bank covered with trees: 5%
- Bank covered with reeds: No data registered
- Bottom covered with bushes: No data registered
- Bottom covered with trees: No data registered
- Bottom covered with reeds: 25%
- Inlet: Ok
- Outlet: Ok

Let me know if you want to know more specific information or anything else.

chat continues...

The chatbot successfully handled both the list retrieval and the random basin selection, demonstrating its ability to perform basic queries effectively. However, the remark field within the basin data contains technical language, which may be challenging for non-expert users to interpret. Upon cross-checking with the source data, the remark accurately reflects the registered information. With the chatbot's primary function being to provide efficient data access rather than interpretation, conveying the remark as originally registered aligns with the model setup, although the appropriateness can be discussed.

5.1.2 Complex Request

Complex requests involve multi-step reasoning where the agent must chain several tool calls. These types of requests can be considered less routine, but nevertheless aims to provide insight into the data. These requests pushes the agent to plan a sequence of tool calls before executing them. One such request could be to ask which basins were registered as protected. The below interaction is a continuation of the example:

User: Which basins in Hvidovre are not registered as protected nature?

Chatbot: *System error – rate limit exceeded*

User: Can you check for three of them?

Chatbot: All three basins—3-0 55/0307 Venstre, 3-0 58/0183 Højre, 3-0 55/0400 Venstre—are registered as protected nature.

- 55/0307 Venstre – SØ, 1.868 m²
- 58/0183 Højre – Strandeng + SØ, 13.804 m²
- 55/0400 Venstre – SØ, 1.868 m²

The initial request to analyze all basins encountered a rate limit error, occurring as the agent performed an LLM call for every tool execution, sending the entire instruction set, tool descriptions, and chat history repeatedly. Since these combined elements contain around 3000 tokens, the sequence of LLM calls quickly exceeded the OpenAI API's rate limit of 10,000 tokens per minute. Despite this, the agent successfully responded when the scope of the query was reduced to three basins. The nature types are also conveyed in Danish, with the system omitting translation although instructed to respond in the user language. This might be due to the temperature setting (0.8), which potentially nudged the LLM to convey the data as registered.

The conveyed results of the basins were cross-validated within GIS and found to be accurate. The first and third basins were identified as spatially connected, which explains why they share the same information. However, an issue arose with the second basin, "50//0183 Højre." This basin, located adjacent to a highway as shown in Figure 23, was mistakenly marked as part of a protected nature area on the opposite side of the highway. This false positive resulted from the current geospatial processing method, which does not account for major physical barriers when determining spatial connectivity.



Figure 23: Map of basin 50/1083 Højre along with its relation to protected nature.

5.2 Basin Report Assessments

Following the assessment of direct chatbot interactions, this section evaluates the primary output of the agent’s analytical tools, being the basin reports, which are central to supporting municipalities and addressing the main research question. These reports offer a quick overview of relevant basin information while also providing summary insights into the next administrative steps. The evaluation focuses on both the accuracy of the spatial information conveyed and the clarity and quality of the summaries and takeaway formulations. To conduct the assessment, the 10 basins specified in Table 6 were selected through a semi-random process to ensure a representative range of situations. Two basins were run multiple times to assess consistency in the model’s output.

| Municipality | # | §3 Nature | Natura 2000 | Annex IV | VP3 Streams | Runtime (s) |
|-------------------------|----|------------|-------------|----------|-------------|-------------|
| Hvidovre | 1 | ✓ | | ✓ | | 7.62 |
| Ringsted | 2 | ✓ | | | ✓ | 7.37 |
| København | 3 | ✓(partial) | ✓ | ✓ | | 44.60 |
| Middelfart | 4 | | | | | 5.91 |
| Vejen | 5 | ✓ | ✓(partial) | ✓ | | 9.51 |
| Silkeborg | 6 | | ✓ | ✓ | | 14.94 |
| Kalundborg | 7 | ✓ | ✓ | ✓ | ✓ | 14.71 |
| <i>Repeated reports</i> | | | | | | |
| Horsens | 8 | ✓ | | ✓ | ✓ | 29.75 |
| | 9 | ✓ | | ✓ | ✓ | 30.31 |
| | 10 | ✓ | | ✓ | ✓ | 35.17 |
| Køge | 11 | | | ✓ | ✓ | 40.80 |
| | 12 | | | ✓ | ✓ | 22.20 |

Table 6: Overview of selected basins and their spatial relationships. Checks indicate the reported status. Green indicates factual correctness and validated PDF results; red indicates identified issues or inconsistencies. The report number links to the corresponding Appendix section.

5.2.1 Spatial Accuracy

The geospatial information presented in the generated reports originated directly from the deterministic outputs of the screening tools. Consequently, the spatial analyses were found to be logically true and consistently reflected accurate information based on the geospatial processing of the input datasets. However, the assessment of spatial accuracy extended beyond logical consistency, with the core consideration being whether the factually correct tool findings accurately describes on-the-ground situations and their regulatory implications without potential for misinterpretation.

A primary nuance identified concerns the interpretation of geometric overlaps. In the case of Basin 4 in København, the tool results described the basin as protected nature, triggered by a minor geometric overlap with an adjacent §3 area. However, the vast majority of the basin itself is not independently registered as protected, arguably making this a false positive. In contrast, the Basin 5 report in Vejen demonstrates a more nuanced situation, where a Natura 2000 site boundary crosses the middle of the lake, creating a partial overlap so significant that the Natura 2000 should rightfully be fully considered.

Similarly, the buffer-based logic inherent in certain tools for identifying connections to nearby protected nature and streams, while consistently applied, has limitations. The tools correctly reported features within the buffer distances, but their boolean-style output ("connection: yes/no") made it difficult to capture nuanced cases where connections might not be meaningful. In most of the generated reports, the results were factually correct, with the primary ambiguity being whether an area is genuinely connected to protected nature, a relatively minor issue.

The visual representation in the maps also had functional and technical limitations, as the default scale caused maps to fail in scaling adequately for Annex IV observations and, to some extent, for streams and protected nature. In some cases, issues with the WMTS service for retrieving the orthophoto also failed, resulting in missing tiles (Basin 9) and a missing map ((Basin 10). Despite the interpretative nuances, the underlying geospatial logic of the tools was found to be consistently correct across all 12 generated reports, with the primary challenge lying in ensuring the automated interpretation aligns perfectly with complex real-world ground conditions.

5.2.2 Descriptive Summaries

Another assessment aspect is to evaluate LLM performance in generating summaries of the attribute fields passed from the tool outputs. The primary goal of these summaries was to provide concise, descriptive, and professional overviews of the factual findings for each consideration. The task is to assess whether the summaries included all attribute information and any variations in the formulations. While the LLM correctly followed the instructed rules by responding with valid JSON structures essential for generating the reports successfully, several patterns and nuances in its summarization approach as guided by the instructions were observed.

A general observation concerns the relevance and scope of information included, particularly in the 'Basin Information' section. When assessing basins with a lot of attribute information, the LLM generally included most factual details. However, it had a tendency to highlight or include off-shoot remarks whose direct relevance to the §3 dispensation context may not immediately be apparent, such as operational notes and comments, which appear to be heavily influenced by the 'remark' attribute field from the source data. Examples of basin summaries include:

- For Basin 2 in Ringsted, the summary noted: *"The inlet and outlet are hidden and overgrown, suggesting that signage installation at these points would facilitate operations."*
- For Basin 5 in Vejen, it included: *"[the basin] requires action, as the outlet is submerged, and the gate needs replacement."*
- For Basin 7 in Kalundborg, the summary mentioned: *"the basin is likely to be decommissioned as part of project 11930. Observations from the last visit indicate a need for clearing to assess the basin's condition."*
- For Basin 9 in Horsens, the summary included: *"The basin is part of a VVM phase, and maintenance should only occur if absolutely necessary, as some basins may be removed or expanded."*

These examples illustrate how the LLM sometimes incorporated operational details or internal remarks that, while potentially informative, could distract from the primary focus of a §3 dispensation assessment or imply actions beyond the immediate scope. While the Horsens basin reports 8, 9, and 10 contains multiple off-shoot remarks stemming from a noted "VVM/EIA" phase, the inclusions of this, while irrelevant to the case, is generally consistent, though with minor phrasing variations. Variations of these remarks are observed for the other spatial summaries, where the LLM occasionally concludes by emphasizing the specific implication of the findings as the examples show below:

- For Basin 3 in København, the Annex IV summary urged *"to ensure compliance with their protection under the Nature Protection Act."*
- A similar more extreme phrasing for Annex IV appears for Basin 12 in Køge, where the LLM concluded to *"investigate their connection to the basin to ensure maintenance [as] not [to] violate protection under §29a of the Nature Protection Act."*
- For Basin 7 in Kalundborg, the stream summary concluded by saying that *"the impact of maintenance on this stream must be assessed."*

While these concrete examples were noted, most summaries were largely appropriate and descriptive. For protected nature, the LLM consistently reported area sizes, relevant nature types, and protected status or connections as per tool outputs. For Natura 2000 sites, it also reliably found and provided the correct site management plan URL, which the LLM then correctly following instructions by embedding the link. Multiple runs for the Horsens (Basin 8, 9, 10) and Køge (Basin 11, 12) showed negligible stylistic variations in these sections, indicating stable factual summarization when input data was straightforward and instructions were clear, though minor verbosity was occasionally noted.

The summaries of Annex IV Species showed more notable nuances. While the presence of species was correctly reported, the LLM often under-described the magnitude of sightings. For example, 21 observations for Basin 1 were summarized as "observed multiple times," and the quantity of 14 observations for Basin 3 was similarly not conveyed in the summary text. This lack of quantitative detail could diminish the perceived significance of species presence, and leaves the reader uninformed about their potential connection to the basin, which is compounded by the species' locations often not being visible on the default map scale.

A minor tendency for action-oriented or interpretive phrasing appeared in some descriptive summaries, which is likely due to the combined effect of general instructions to produce *"concise, professional text"* and field-specific prompts such as Annex IV's instruction to *"describe proximity and quantity while assuming presence is neither confirmed nor excluded"*. While the text remains relevant in the broader dispensation context, such phrasing represents slight scope creep and occasional verbosity. Still, repeated generations for the Horsens and Køge basins showed consistent factual output, suggesting overall stability despite these tendencies.

5.2.3 Adherence to Conditional Logic

The final and most critical aspect assessed is the 'Summary Insight' section, which synthesizes the attribute summaries into actionable conclusions. Its generation is guided by the LLM instructions specifying conditional logic described in Appendix C.2 and shown in Figure 21. The evaluation focused on the LLM's adherence to this logic through a combined quantitative and qualitative assessment. For each report, the generated insights were compared against the set of clauses expected based on the input data and rules. Results of this analysis are presented in Table 7.

| Report | Expected Clauses | Correctly Generated | Omitted Clauses | Incorrectly Added | Notes |
|--------------|------------------|---------------------|-----------------|-------------------|---------------------------|
| Basin 1 | 2 | 2 | 0 | 0 | - |
| Basin 2 | 2 | 2 | 0 | 0 | - |
| Basin 3 | 3 | 3 | 0 | 0 | - |
| Basin 4 | 1 | 1 | 0 | 0 | - |
| Basin 5 | 3 | 3 | 0 | 0 | - |
| Basin 6 | 3 | 3 | 0 | 0 | - |
| Basin 7 | 4 | 4 | 0 | 0 | - |
| Basin 8 | 3 | 3 | 0 | 0 | - |
| Basin 9 | 3 | 3 | 0 | 0 | - |
| Basin 10 | 3 | 2 | 1 | 0 | VP3 stream clause omitted |
| Basin 11 | 3 | 3 | 0 | 0 | - |
| Basin 12 | 3 | 3 | 0 | 0 | - |
| Total | 33 | 32 | 1 | 0 | Success Rate: 97% |

Table 7: LLM Adherence to Conditional Logic in Generating 'Summary Insight' Clauses Across 12 Basin Reports

Across the 12 reports, a total of 33 individual conditional insight clauses were expected based on the input data and the logic framework. The LLM successfully generated 32 of these, with only one omission occurring in the report for Basin 10, where the VP3 stream clause was not triggered despite the stream connection being correctly identified in the report. Importantly, no instances were found where the LLM included a clause not warranted by the logic. This yields a 97% success rate in the LLM correctly applying the core conditional logic for including intended summary statements.

Qualitatively, the 'Summary Insights' consistently identified when a §3 dispensation was needed, recommended site visits for unprotected basins with older maintenance dates, and questioned the need for action in recently maintained cases (e.g., Basin 11 and Basin 12 vs. Basin 4). Even when multiple regulatory flags were present (e.g., Basin 7; Basin 3; Basin 5), the LLM correctly integrated them into a single actionable paragraph.

While phrasing occasionally leaned toward verbosity, especially in reports with many regulatory flags (e.g., Basin 7; Basin 5), the factual logic was consistently applied. Minor stylistic variation across the repeated Horsens reports did not change the conveyed regulatory meaning. Overall, the LLM showed strong and reliable performance in applying conditional logic, supporting its use for generating structured, regulatory-aware conclusions in §3 dispensation screening.

6 Lessons Learned for Future Work

This chapter reflects on the key insights, challenges, and lessons learned throughout the design, development, and initial assessment of the geospatial AI-agent workflow for supporting the §3 basin dispensation application. It examines experiences related to system design, LLM integration, data handling, and validation. It concludes by discussing the limitation of the developed system and outlining directions for future work.

Limitations in Capturing Spatial Nuance

A lesson learned during development and assessment of the agent tools concerned the inherent limitations of the datasets in capturing real-world spatial nuance. While the deterministic tools accurately processed available data to identify overlaps or proximity, their outputs sometimes lacked the contextual depth needed to reflect on-the-ground situations. For instance, a geometric overlap or proximity between a basin and a protected area might not imply a relevant connection if infrastructure acts as a physical barrier. Likewise, proximity without considering surrounding urban development might misrepresent the actual environmental context. Other datasets could infuse the results with nuance, such as using road data to check for physical obstructions or local plans to assess land use conditions.

The identified limitation in capturing nuanced spatial context directly impacted the depth of assessment the AI agent and the report-generating LLM could provide. As a result, the 'Summary Insights' still require human validation of site conditions to confirm the conveyed statements. Future development could address this by enriching tool outputs with additional data sources, giving the LLM more nuanced input to reason upon. In this project, reliable results were less limited by the LLM's reasoning capacity than by the level of detail in its inputs—specifically, the prompt and instructions. This suggests that injecting more contextual data alongside well-crafted instructions could expand the capabilities of AI-driven decision support systems.

Instructional Design and LLM Interactions

In the context of the instruction and tool description iterations, the interplay between LLM instructions and the structure of the input data required careful consideration and refinement. For instance, when tool outputs indicated that a particular regulatory feature was not found near the basin, early iterations of the report-generating LLM, despite instructions to use a specific placeholder like *"Not found in conflict search"* would sometimes produce overly descriptive definitive statements, interpreting the lack of a finding as a conclusive absence relevant to the case. This highlighted a tendency of the LLM to fill in gaps or infer beyond the explicit data if the instructions were not sufficient in describing the task.

Further challenges in LLM interactions emerged in the agent's orchestration of tool calls. During development, GPT-4 occasionally simplified or abbreviated complex basin identifiers when passing them as arguments. While this was likely an attempt to reduce complexity, it led to invalid arguments and empty tool results. To address this, the input descriptions were made more stringent, explicitly requiring *"The EXACT basin identifier from the Match Basin Tool."* This highlighted that while LLMs can manage complex reasoning, enforcing technical precision requires clear constraints and validation mechanisms at the tool input level.

The structure of the data provided to the LLMs, primarily JSON dictionaries, was also learned to be an important factor. While each tool call generally returned structured data, the lesson points to the need for clear and unambiguous data representation in the input prompt. For tasks involving reasoning over multiple tool outputs across entities, including unique basin identifiers in each JSON snippet is essential. This, alongside understandable key naming with clear hierarchies within the JSON, is crucial for enabling an LLM to accurately process, associate, and synthesize information from different tool outputs without confusion.

Advancing Conflict Assessment Procedures

A key lesson from this project was the clear delineation of the AI agent's role in enhancing, rather than replicating, existing §3 dispensation screening workflows. Standard municipal practice involves an initial conflict assessment, often supported by GIS-based tools that generate factual reports detailing potential overlaps with regulatory layers, serving as a standardized checklist. While crucial for establishing a factual baseline, these existing reports typically present data without synthesized interpretation or tailored guidance on implications for the case. The development process, informed by the municipal decision-making and the need for site-specific nuance, confirmed that a fully autonomous decision-making agent remained impractical.

Consequently, the AI agent was designed to augment the conflict assessments. Its conversational interface provides easy access to specific tool-derived facts, while its core value lies in the generated PDF reports. These reports go beyond factual checklists by using an LLM to interpret the geospatial findings and highlight potential regulatory implications. The system supports the initial gathering of *case facts* and the inference of relevant *legal facts* from the data, but deliberately stops short of performing the final legal *subsumption*. This design ensures the AI acts as a technical assistant, streamlining the screening process and empowering case handlers to apply their expertise more effectively.

Stakeholder Feedback and Model Validation

Direct engagement with stakeholders and an iterative development phase proved central to refining the AI workflow and aligning it with municipal needs and regulatory understanding. For example, an early version of the PDF report incorrectly implied that the presence of Natura 2000 sites or Annex IV species would automatically require supplementary permits alongside a §3 dispensation. Feedback in interview 3 clarified that the key concern is avoiding negative impact on protected elements, leading to refinements to the system instructions and summary formulations. While stakeholders expressed a preference for simpler language, a deliberate choice was to retain explicit references to the statutory basis for research purposes, although future implementation might lean more toward accessibility.

The project's iterative approach to integrating domain expertise proved valuable. The initial semi-structured interviews (Interviews 1 & 2) helped map §3 dispensation processes and identify key information needs from both applicant and municipal perspectives. This understanding informed the development of the AI agent and its core functionalities. A follow-up session (Interview 3) clarified several points and provided targeted feedback. A lesson learned was that this stage could benefit from deeper stakeholder

involvement. Future projects could frame these sessions more as “co-design” workshops, where case workers help refine LLM instructions with prompt engineering and compare output formulations. This would better translate their domain knowledge into practical improvements in the agent’s reasoning and presentation style.

From a technical standpoint, the ability to verify the data on both sides of the report-generating LLM was a key aspect of development and preliminary validation. The JSON payload, aggregated from deterministic tool outputs, could be inspected before being passed to the LLM, and the structured JSON response generated in return could likewise be examined before PDF rendering. This “glass-box” setup, referring to the ability to inspect inputs and outputs of a specific system, enabled analysis of how the LLM interpreted instructions and generated results. This provided confidence in the internal consistency of the pipeline, even without detailed user validation, and demonstrated that the system could translate factual tool outputs into structured summaries and insights under the guidance of the instructions.

Architectural Enhancements for Scalability

A lesson from the model concerned the current agent architecture’s challenges with API rate limits when processing complex queries requiring multiple chained tool executions. As observed in the result section, requests to analyze several basins sequentially failed due to exceeding OpenAI API token limits (e.g., 10,000 input tokens per minute). While the agent’s baseline instructional context was around 1,200 tokens, the nature of the LangChain AgentExecutor, which resends accumulated chat history and intermediate agent steps with each LLM call, led to accumulation in token usage. This highlighted that while the current architecture is effective for simpler tasks, a more scalable approach is needed for the robust handling of complex queries.

To address this, a future architectural enhancement would involve decoupling the planning and execution phases of complex queries. This revised workflow would consist of: **(1)** An initial LLM call (*Planner*) to interpret the user’s request and generate a plan detailing the sequence of necessary tool calls and their arguments. **(2)** A deterministic execution engine to then run all planned tools in a batch, aggregating their results without intermediate LLM calls. **(3)** A final LLM call (*Responder*) to process the complete set of aggregated tool results and generate an answer for the user. This decoupled architecture can significantly reduce API calls, improve token efficiency, increase robustness against rate-limit errors, and enhance overall system reliability and performance for practical municipal use.

Beyond scalability, the architectural separation of the conversational agent from the dedicated, task-instructed LLM used for PDF report generation proved central in ensuring LLM-alignment with specific functionalities in the workflow. This design allowed one agent (LLM) to carry out user interaction and tool execution, with another LLM specifically instructed to operate under highly specific, domain-crafted instructions for generating the detailed report of single cases. For user experience, a further lesson pointed to the potential value of integrating an interactive map alongside the chatbot as current map visualization in the end of the workflow sometimes omits features due to default scaling. An interactive map offering auto-zoom and direct basin interaction would greatly improve efficiency and intuitiveness.

7 Conclusion

The upscaling of Vejdirektoratet’s basin cleaning efforts followed with an increase in §3 basin dispensation applications, represent a significant upcoming workload for Danish municipalities. This thesis investigated how geospatial analysis and NLP could address Vejdirektoratet’s goal of supporting municipalities in this task. Through an iterative methodology informed by stakeholder interviews and analysis of past cases, a geospatial AI-agent was designed and assessed, aimed at answering spatial queries and generating augmented conflict assessment reports to support municipal case workers by providing tailored, case-specific insights from the underlying spatial data.

To address the essential information requirements for municipal decision-making (**RQ1**), the research identified central legal statutes (such as NBL and HB), key procedural steps (including conflict assessment), the core geospatial datasets, and the role of site visits. Site visits were found to be central for assessing the determining the site-specific conditions not derivable from available data. This fundamental limitation in data appropriateness, as well as the case reliance on site-specific conditions, marked the limit of the model’s potential capabilities and resulted in the deliberate decision to derive as much information from the data as possible.

In exploring how NLP workflows could reliably extract and synthesize spatial data for case support (**RQ2**), the developed AI-agent demonstrated a viable proof-of-concept. The system was designed as a modular sequence of deterministic geospatial tools, each performing specific analyses (overlap, proximity, attribute retrieval) and integrating external web services for basemaps and Natura 2000 management plan lookups. These tools generated structured JSON outputs, which were then effectively received and responded upon by the system. For the augmenting of conflict assessment reports with coherent summaries and action-oriented insights, a task-instructed LLM utilized carefully crafted instructions to synthesize the findings.

The assessment of the AI workflow’s factual accuracy, reliability, and practical utility (**RQ3**) revealed strong performance. The spatial reasoning consistently conveyed verifiable results, with discrepancies primarily arising from the interpretation of nuanced spatial contexts rather than tool errors. The LLM responsible for generating summary insights for the basin reports showed a 97% success rate in correctly applying the conditional logic for deriving instructed implications, indicating high reliability in translating structured data into natural language and rule-based conclusions. However, as covered, the practical utility is inherently limited by the current datasets’ inability to capture site-specific conditions.

This project demonstrates the transformative potential of agent-driven architecture in case work, where LLM-powered agents facilitates user interaction with geospatial queries and document creation with separate task-specific LLMs are instructed to reason for the augmented conflict reports. The key contribution is showcasing how AI can augment existing conflict assessments by delivering tailored preliminary insights from spatial data, offering a more context-aware starting point for municipal case handlers. Future work should explore scalability and refine co-design methods with end-users through prompt engineering to further adapt these systems to practical municipal workflows.

References

- [1] Vejdirektoratet. *Personal communication*. Informal Communication. 2025.
- [2] European Parliament and Council of the European Union. *Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy*. Official Journal of the European Communities, L 327, 22 December 2000, pp. 1–72. 2000.
- [3] Miljøministeriet. *Vandområdeplanerne 2021–2027*. <https://edit.mst.dk/media/njvvhax/vandomraadeplanerne-2021-2027-22-9-2023.pdf>. ISBN 978-87-91824-01-2. 2023.
- [4] Miljø- og Ligestillingsministeriet. *Bekendtgørelse af lov om naturbeskyttelse (LBK nr. 927 af 28/06/2024)*. <https://www.retsinformation.dk/eli/lta/2024/927>. Danish Nature Protection Act. 2024.
- [5] Peter Høgstad Kelstrup and Jacob Adeel Madsen. *Interview regarding Applicant Process for §3 Dispensations*. Personal Communication (Teams Meeting). Interview conducted by Alexander Adiyasa. Mar. 2025.
- [6] Lene Kofoed et al. *Interview regarding Municipal Process for §3 Dispensations (HP25-mål om Bassiner)*. Personal Communication (Teams Meeting). Interview conducted by Alexander Adiyasa. Apr. 2025.
- [7] Miljø- og Ligestillingsministeriet. *Bekendtgørelse om udpegning og administration af internationale naturbeskyttelsesområder samt beskyttelse af visse arter (BEK nr. 1098 af 21/08/2023)*. <https://www.retsinformation.dk/eli/lta/2023/1098>. Danish Habitat Protection Act. 2023.
- [8] Miljø- og Ligestillingsministeriet. *Bekendtgørelse af lov om miljøbeskyttelse (LBK nr. 1093 af 11/10/2024)*. <https://www.retsinformation.dk/eli/lta/2024/1093>. Danish Environmental Protection Act. 2024.
- [9] Miljø- og Ligestillingsministeriet. *Bekendtgørelse af lov om forurennet jord (LBK nr. 282 af 27/03/2017)*. <https://www.retsinformation.dk/eli/lta/2017/282>. Danish Contaminated Soil Act. 2017.
- [10] Miljø- og Ligestillingsministeriet. *Bekendtgørelse om affald (BEK nr. 1749 af 30/12/2024)*. <https://www.retsinformation.dk/eli/lta/2024/1749>. Danish Waste Management Order. 2024.
- [11] Kulturministeriet. *Bekendtgørelse af museumsloven (LBK nr. 358 af 08/04/2014)*. <https://www.retsinformation.dk/eli/lta/2014/358>. Danish Museum Act. 2014.
- [12] Therese Monberg. *Slides from Lecture 2–3: Legal Fundamentals – On Thinking Legally and Legal Method*. Bachelor course, 3rd semester, Aalborg University Copenhagen. Unpublished teaching material. 2021.
- [13] Furesø Kommune, Center for By og Miljø. *Dispensation til oprensning af sø på matr.nr. 7000c, Ravnsholt Skov (Sags nr. 18/1369)*. Afgørelse/Dispensation. June 2018.
- [14] Fredericia Kommune, Natur & Miljø. *Dispensation til oprensning af beskyttet regnvandsbassin (Sagsnr. 18/8745)*. Afgørelse/Dispensation. Nov. 2018.
- [15] Greve Kommune, Teknik & Miljø. *Dispensation til oprensning af § 3 beskyttet regnvandsbassin (Sagsnr. 253-2018-35789)*. Afgørelse/Dispensation. July 2019.

- [16] Greve Kommune. *Dispensation fra naturbeskyttelseslovens § 3 til oprensning af regnvandsbassiner (Sagsnr. 22-003386)*. Afgørelse/Dispensation. Jan. 2023.
- [17] Gladsaxe Kommune, Klima Og Natur. *Dispensation fra naturbeskyttelsesloven til oprensning af regnvandsbassin (Hillerød motorvejen)*. Afgørelse/Dispensation. Sept. 2022.
- [18] Furesø Kommune, Natur- og Miljøafdelingen. *Dispensation til at oprensning af sø på matr.nr. 7000c, Ravnsholt Skov (Sags id: 22/15835)*. Afgørelse/Dispensation. Nov. 2022.
- [19] Esbjerg Kommune, Teknik & Miljø. *Dispensation fra naturbeskyttelsesloven til oprensning af Vejdirektoratets regnvandsbassiner (Sagsid 16/13277)*. Afgørelse/Dispensation. July 2019.
- [20] Fredericia Kommune, Natur og Miljø. *Dispensation til oprensning af beskyttet regnvandsbassin (Sagsnr. 22/10950)*. Afgørelse/Dispensation. Nov. 2022.
- [21] Brønderslev Kommune, Plan og Miljø. *Dispensation til Oprensning af regnvandsbassin 310,3H (Sags. nr. 06.03.00-P19-1-18)*. Afgørelse/Dispensation. Sept. 2018.
- [22] Esbjerg Kommune, Teknik & Miljø. *Dispensation efter naturbeskyttelsesloven til oprensning af et vejvandsbassin (Sagsid mig-a-2024-2173)*. Afgørelse/Dispensation. Feb. 2025.
- [23] Jonathan J. Webster and Chunyu Kit. “Tokenization as the Initial Phase in NLP”. In: *Proceedings of COLING*. Nantes, France, 1992. URL: <https://aclanthology.org/C92-4173/>.
- [24] Sabrina J. Mielke et al. “Between Words and Characters: A Brief History of Open-Vocabulary Modeling and Tokenization in NLP”. In: *arXiv* (2021). URL: <https://arxiv.org/abs/2112.10508>.
- [25] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. “GloVe: Global Vectors for Word Representation”. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2014, pp. 1532–1543. DOI: [10.3115/v1/D14-1162](https://doi.org/10.3115/v1/D14-1162). URL: <https://aclanthology.org/D14-1162>.
- [26] Rico Sennrich, Barry Haddow, and Alexandra Birch. “Neural Machine Translation of Rare Words with Subword Units”. In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*. 2016, pp. 1715–1725. DOI: [10.18653/v1/P16-1162](https://doi.org/10.18653/v1/P16-1162). URL: <https://arxiv.org/abs/1508.07909>.
- [27] Taku Kudo and John Richardson. “SentencePiece: A Simple and Language-Independent Subword Tokenizer and Detokenizer for Neural Text Processing”. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 2018, pp. 66–71. DOI: [10.18653/v1/D18-2012](https://doi.org/10.18653/v1/D18-2012). URL: <https://arxiv.org/abs/1808.06226>.
- [28] Encyclopædia Britannica. *Morphology (linguistics)*. <https://www.britannica.com/topic/morphology-linguistics>. Accessed: May 28, 2025.
- [29] Cagri Toraman et al. “Impact of Tokenization on Language Models: An Analysis for Turkish”. In: *arXiv* (2022). URL: <https://arxiv.org/abs/2204.08832>.
- [30] Anmol Nayak et al. “Domain Adaptation Challenges of BERT in Tokenization and Sub-Word Representations of Out-of-Vocabulary Words”. In: *Proceedings of the First Workshop on Insights from Negative Results in NLP*. Association for Computational Linguistics, 2020, pp. 1–5. URL: <https://doi.org/10.18653/v1/W19-4828>.

- [31] Ashish Vaswani et al. “Attention Is All You Need”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Curran Associates, Inc., 2017, pp. 5998–6008. URL: <https://arxiv.org/abs/1706.03762>.
- [32] Sepp Hochreiter and Jürgen Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (1997), pp. 1735–1780.
- [33] Alec Radford et al. *Learning Transferable Visual Models From Natural Language Supervision*. 2021. arXiv: 2103.00020 [cs.CV]. URL: <https://arxiv.org/abs/2103.00020>.
- [34] Jacob Devlin et al. “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT)*. Association for Computational Linguistics, 2019, pp. 4171–4186. URL: <https://arxiv.org/abs/1810.04805>.
- [35] Tom B. Brown et al. “Language Models are Few-Shot Learners”. In: *arXiv* (2020). URL: <https://arxiv.org/abs/2005.14165>.
- [36] Alec Radford et al. *Improving Language Understanding by Generative Pre-Training*. OpenAI technical report. 2018. URL: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.
- [37] Lei Huang et al. “A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions”. In: *ACM Transactions on Information Systems* 43.2 (Jan. 2025), pp. 1–55. ISSN: 1558-2868. DOI: 10.1145/3703155. URL: <http://dx.doi.org/10.1145/3703155>.
- [38] Artificial Analysis. *Artificial Analysis – Intelligence Research Tracker*. <https://artificialanalysis.ai/>. Accessed: May 28, 2025. 2024.
- [39] Long Phan et al. “Humanity’s Last Exam”. In: *arXiv* (2025). arXiv: 2501.14249 [cs.LG]. URL: <https://arxiv.org/abs/2501.14249>.
- [40] David Rein et al. “GPQA: A Graduate-Level Google-Proof Q&A Benchmark”. In: *arXiv* (2023). arXiv: 2311.12022 [cs.AI]. URL: <https://arxiv.org/abs/2311.12022>.
- [41] Naman Jain et al. *LiveCodeBench: Holistic and Contamination Free Evaluation of Large Language Models for Code*. 2024. arXiv: 2403.07974 [cs.SE]. URL: <https://arxiv.org/abs/2403.07974>.
- [42] Rishi Bommasani et al. “On the Opportunities and Risks of Foundation Models”. In: *arXiv preprint arXiv:2108.07258* (2021). URL: <https://arxiv.org/abs/2108.07258>.
- [43] Pranab Sahoo et al. *A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications*. 2025. arXiv: 2402.07927 [cs.AI]. URL: <https://arxiv.org/abs/2402.07927>.
- [44] Jason Wei et al. *Chain-of-Thought Prompting Elicits Reasoning in Large Language Models*. 2023. arXiv: 2201.11903 [cs.CL]. URL: <https://arxiv.org/abs/2201.11903>.
- [45] Long Ouyang et al. *Training language models to follow instructions with human feedback*. 2022. arXiv: 2203.02155 [cs.CL]. URL: <https://arxiv.org/abs/2203.02155>.
- [46] Eduardo Mosqueira-Rey et al. “Human-in-the-loop Machine Learning: A State of the Art”. In: *Artificial Intelligence Review* 56 (2023). Published online 17 Aug 2022, pp. 3005–3054. DOI: 10.1007/s10462-022-10246-w.

- [47] Zhi Jing, Yongye Su, and Yikun Han. “When Large Language Models Meet Vector Databases: A Survey”. In: *arXiv preprint arXiv:2402.01763* (2024). <https://arxiv.org/abs/2402.01763>.
- [48] Tomas Mikolov et al. “Efficient Estimation of Word Representations in Vector Space”. In: *Proceedings of Workshop at International Conference on Learning Representations (ICLR)*. 2013. URL: <https://arxiv.org/abs/1301.3781>.
- [49] Yichao Wu et al. *LoRA-SP: Streamlined Partial Parameter Adaptation for Resource-Efficient Fine-Tuning of Large Language Models*. 2024. arXiv: 2403.08822 [cs.LG]. URL: <https://arxiv.org/abs/2403.08822>.
- [50] Patrick Lewis et al. “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Curran Associates, Inc., 2021. URL: <https://arxiv.org/abs/2005.11401>.
- [51] Jeremy Howard and Sebastian Ruder. *Universal Language Model Fine-tuning for Text Classification*. 2018. arXiv: 1801.06146 [cs.CL]. URL: <https://arxiv.org/abs/1801.06146>.
- [52] Tianyu Gao, Xingcheng Yao, and Danqi Chen. “SimCSE: Simple Contrastive Learning of Sentence Embeddings”. In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Ed. by Marie-Francine Moens et al. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 6894–6910. DOI: 10.18653/v1/2021.emnlp-main.552. URL: <https://aclanthology.org/2021.emnlp-main.552/>.
- [53] OpenAI. *A Practical Guide to Building Agents*. <https://cdn.openai.com/business-guides-and-resources/a-practical-guide-to-building-agents.pdf>. Apr. 2025.
- [54] Lilian Weng. *LLM Powered Autonomous Agents*. <https://lilianweng.github.io/posts/2023-06-23-agent/>. Accessed: May 28, 2025. 2023.
- [55] Xinyi Li et al. “A survey on LLM-based multi-agent systems: workflow, infrastructure, and challenges”. In: *Vicinaearth* 1.9 (2024), pp. 1–43. DOI: 10.1007/s44336-024-00009-2. URL: <https://doi.org/10.1007/s44336-024-00009-2>.
- [56] Hailong Yang et al. “XAgents: A Framework for Interpretable Rule-Based Multi-Agents Cooperation”. In: *Proceedings of AAAI*. Association for the Advancement of Artificial Intelligence, 2024. URL: <https://arxiv.org/abs/2411.13932>.
- [57] Hema Kadia. *Rule-Based vs. LLM-Based AI Agents: A Side-by-Side Comparison*. TeckNexus. Accessed: March 8, 2025. 2024. URL: <https://tecknexus.com/rule-based-vs-llm-based-ai-agents-a-side-by-side-comparison/>.
- [58] Taicheng Guo et al. “Large Language Model based Multi-Agents: A Survey of Progress and Challenges”. In: *arXiv* (2024). arXiv: 2402.01680 [cs.CL]. URL: <https://arxiv.org/abs/2402.01680>.
- [59] Lee Boonstra. *Prompt Engineering*. Feb. 2025. URL: https://drive.google.com/file/d/1AbaBYbEa_EbPelsT40-vj64L-2IwUJHy/view%7D.

Appendices

A Case Document Examples

Notat

18.12.2018

Ejlskov A/S
Jens Olsens Vej 3
8200 Århus N
Danmark
www.ejlskov.com

Sag: 17061
X@ejlskov.com
Tel: XXXXXXXX

Klient: Vejdirektoratet**Projekt:** Bassinopretninger**Opgave:** Ansøgning til dispensation efter §3 i Naturbeskyttelsesloven**1. Indledning og formål**

Ejlskov A/S ansøger, på vegne af Vejdirektoratet, om dispensation efter §3 i Naturbeskyttelsesloven i forbindelse med oprensning af regnvandsbassinet:

Bassin 30-0 63/0270V: 7000e Nielstrup By, Ulse.

Regnvandsbassinet er beliggende inden for den overstående matrikel i Faxe Kommune.

Desuden ønskes et svar, på om der kræves tilladelse efter Miljøbeskyttelseslovens §19 til genplacering af sediment inden for vejmatrikel på areal omkring regnvandsbassinerne.

1.1 Baggrund

Vejdirektoratet ønsker som led i vedligeholdelse af deres tekniske afvandingsanlæg at oprense regnvandsvandsbassiner for aflejret sediment.

Der er ved den orienterende sedimentundersøgelse (se bilag 1) blevet konstateret en sedimentmægtighed på ca. 0,2 m. Der blev herfra udtaget sedimentprøver pr. 120 ton, i alt 1 sedimentprøve. Prøven herfra overskrider Miljøstyrelsens jordkvalitetskriterier og er på den baggrund kategoriseret som værende lettere forurenet.

Rapport for den miljøtekniske undersøgelse er vedlagt som bilag (bilag 1).

2. Oprensning og beskyttelsesforanstaltninger

Vejdirektoratet ønsker at sikre den fremtidige drift af regnvandsbassinet og har i den forbindelse udarbejdet forslag til oprensning og til fremtidigt vedligehold i form af en plejeplan. Dette er skitseret nedenfor.

Vejdirektoratets retningslinjer for oprensning af regnvandsbassiner omfatter, at:

- alt arbejde i og omkring bassinerne skal ske uden for fugle og padders yngletid, dvs. udføres i perioden fra 1. september til 1. marts.
- bassinerne skal oprenses til fast bund og ikke dybere.
- brinkerne må ikke gøres stejlere end de er i dag.
- det oprensede materiale genplaceres i videst mulige omfang indenfor vejmatriklen, såfremt der er kapacitet indenfor vejmatrikkel ved de respektive bassiner. Der kan være behov for indhentning af tilladelse efter Miljøbeskyttelseslovens §19.
- oprenset sediment opgraves og placeres direkte omkring bassinerne i en sikker afstand fra bassinernes kronekant og fra tilstødende recipienter for at undgå efterfølgende nedskyl eller udflydning. Afstanden afhænger af det aktuelle vandindhold.
- sedimenterne håndteres ved opgravning, så der så vidt muligt undgås op-hvirvling af sediment i vandet.
- bassinerne tømmes derfor først for vand i størst muligt omfang ved pumpning.
- der træffes foranstaltninger for at frafiltrere og minimere mængden af sediment i det oppumpede vand, som afledes via bassinernes eksisterende afløb.
- vegetation ryddes indledningsvist i et omfang der muliggør dels adgang med gravemaskine dels sikrer mulighed for at nå hele bassinets volumen.
- den endelig højde på det indbyggede sediment vil ikke overstige 0,5 meter over eksisterende terræn. Dette gælder også materialer fra knusning af vedplanter.
- der udføres ikke samtidig spuling af rør.
- der må ikke udsås græs, engblanding eller andet på arealet. Efter jordbearbejdning efterlades jorden uden vegetation til naturlig succession.
- bassinerne gøres mere lysåbne ved fjernelse af træer og krat, især mod syd og vest.
- opvækst af Pil og Rød-El anbefales fjernet ved oprykning med rode, da de skyder kraftigt efter beskæring.
- omtrent 1/10 af bredzonen skal efterlades urørt.
- der må ikke sprøjtes eller gødskes i søens umiddelbare nærhed (ca. 15 m)
- der må ikke laves øer i søen
- hvis der bliver fundet arkæologisk materiale ved bygge- eller anlægsarbejdet, skal arbejdet standses jævnfør § 27, stk. 2 i museumsloven. Fundet skal straks anmeldes til museet
- hvis der opdages forurening under arbejdets udførelse skal underrettes
- der gives besked til kommunens Natur- og Vandafdeling umiddelbart efter at projektet er gennemført.

Herudover er der opstillet retningslinjer for Vejdirektoratets fremtidige vedligeholdelse af bassiner. En plejeplan er skitseret som nedenstående:

- arealerne over vandspejlet vedligeholdes ved en simpel nedslåning af nye vækster hvert eller hvert andet år.

- foran ind- og udløb oprenses ca. hvert 5. år
- afklippet materiale fjernes fra arealerne.
- nedslåning af vegetationen omkring bassinerne over vandspejlet og oprensning foran ind- og udløb udføres uden dispensation fra naturbeskyttelseslovens § 3.
- større træer fældes i den udstrækning de er til gene for vedligeholdelsen af bassinet. Der er opmærksomhed på at der er restriktioner på fældningstidspunkter for større træer (redetræer).
- for at undgå skyggevirksomhed efterlades der normalt kun større planter nord for bassinet.
- større oprensninger under vandspejl udføres kun efter forudgående dispensation fra natur- og miljømyndighederne.

Af hensyn til dyrelivet i bassinet, vil den praktiske vedligeholdelse ske efter 1. september og før 1. marts.

3. Genplacering af sediment

Sedimentet fra bassinerne ønskes i videst muligt omfang udlagt omkring bassinerne, såfremt der er kapacitet indenfor vejmatrikel ved de respektive bassiner.

Idet der ved miljøteknisk undersøgelse af sedimentet er truffet indhold af miljøfremmede stoffer svarende til lettere forurenet jord, ønskes en tilkendegivelse af, hvorvidt der kræves tilladelse efter Miljøbeskyttelseslovens §19.

4. Bilag

Bilag 1 : Rapport - Orienterende sedimentprøver af regnvandsbassiner

Udarbejdet af
Ejlskov A/S

XXXXXXXX

Geolog, cand.scient.

Resultat af konflikt søgning

Konflikt søgning - miljø

Dato: 28. marts 2025

Sagsnummer: 1

Udarbejdet af: XXXXX

Bestilt af:

Der gøres opmærksom på, at konflikt søgningen er et øjebliksbillede og skal dermed regnes som vejledende. Konflikt søgningen kan bruges som del af grundlaget i vurderingen af hvorvidt projektet skal anmeldes til MKV-screening hos Trafikstyrelsen eller ikke (MKV-visitation). Hvis gældende lovgivning, registreringer eller vejprojektet ændrer sig fra det, som er brugt i denne konflikt søgning, kan det give anledning til at ændre på konklusionen og derfor skal visitationen foretages på ny.

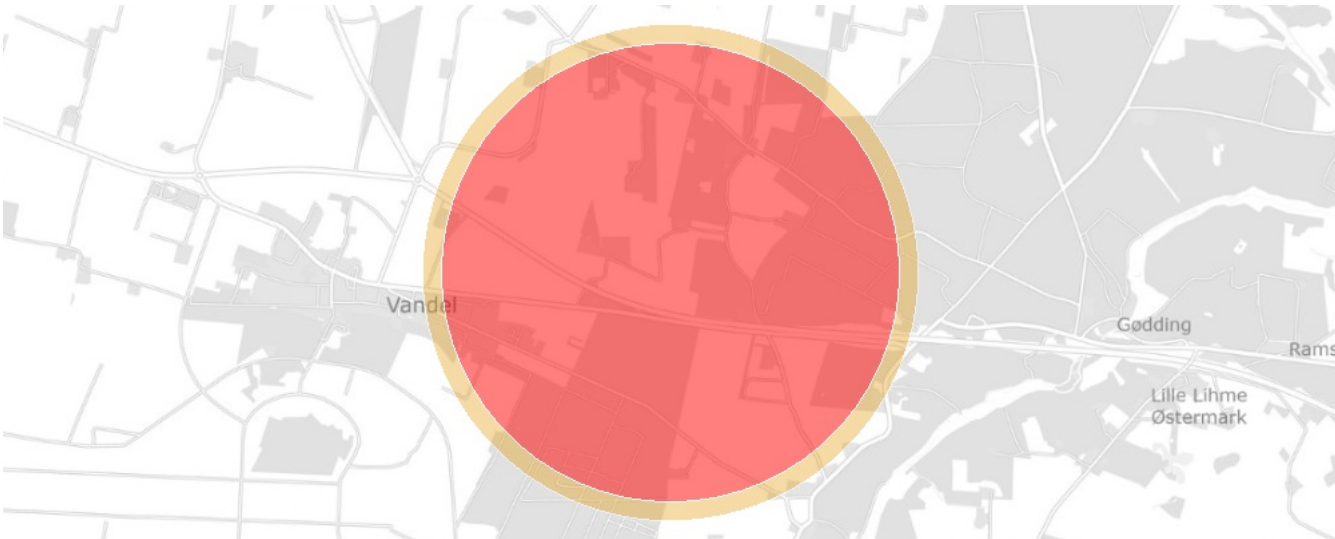
Konflikt søgningen tager udgangspunkt i eksisterende tilgængelig viden fra f.eks. Danmarks Arealinformation, Danmarks Miljøportal, PlanInfo, Danmarks Naturdata, MiljøGIS, Arter.dk, GEUS-databaser og kommunale webkort.

Dialog med kommune og andre relevante myndigheder er anbefalelsesværdig og supplerende materiale såsom "skrivebordsbesigtigelse" vha. COWI Multiviewer (service ala google streetview), som er tilgængelig på vejkort.vd.dk, kan med fordel bringes i spil for et mere komplet grundlag til MKV-visitationen.

Ved tvivlsspørgsmål kontakt da miljøteamet via miljøevma@vd.dk.

God fornøjelse!

Oversigtskort over udvalgt område



Analyserede temaer og fundne konflikter indenfor udvalgt område og buffer

| Tema | Konflikt | Buffer (100.0 m) |
|---|----------|------------------|
| *VP3 - Afgrænsning - Vandløb | | |
| *VP3 - Afgrænsning - Sø | | |
| *VP3 - Afgrænsning - Grundvand (dybde) | | |
| *VP3 - Afgrænsning - Grundvand (regional) | | |
| *VP3 - Afgrænsning - Grundvand (terremer) | | |
| *VP3 - Afgrænsning - Kystvande | | |
| Fuglekollision | X | X |
| Naturmæssigt særlig værdifuld skov på offentlige arealer (§25-skov) | X | X |
| Beskyttede naturtyper | X | X |
| Beskyttede sten- og jorddiger | X | X |
| Fredede områder | X | X |
| Fredede fortidsminder | X | X |
| Fortidsmindebeskyttelseslinje - polygon | X | X |
| Kirkebyggelinjer | X | X |
| Skovbyggelinjer | X | X |
| Overvågning af fugle flader | X | X |
| Overvågning af bilagsarter flader | X | X |
| Generelle artsfund flader | X | X |
| Værdifuld natur langs infrastruktur | X | X |
| Arter i kvadrater (2024) | X | X |
| Arter i kvadrater | X | X |
| Invasive arter | X | X |

| Tema | Konflikt | Buffer (100,0 m) |
|---|----------|------------------|
| Kommuneplan - oversigt - vedtaget | X | X |
| Lokalplan - vedtaget | X | X |
| Skovrejsningsområde - vedtaget | X | X |
| Fredskov - flade - gældende | X | X |
| Økologisk forbindelse | X | X |
| Bevaringsværdige landskaber - vedtaget | X | X |
| Naturbeskyttelsesområde - vedtaget | X | X |
| Drikkevandsinteresser, vedtaget | X | X |
| Følsomme indvindingsområder, vedtaget | X | X |
| Indsatsområder, vedtaget | X | X |
| Indvindingsoplande, alle | X | X |
| Boringer | X | X |
| Jordforurening - V1 kortlagt | X | X |
| Jordforurening - udgiet for kortlægning | X | X |
| Jordbundstyper | X | X |
| Habitatdirektivets bilagsarter | X | |
| Rødlistede arter | | X |
| Jordforurening - udgiet efter kortlægning | | X |
| Beskyttede vandløb | | |
| Strandbeskyttelseslinje - flade - gældende | | |
| Klitfredning - flade - gældende | | |
| Kystnærhedszone | | |
| Kulturarvsarealer | | |
| Fredede bygninger | | |
| Bygning fredningstatus - lav | | |
| Bygning fredningstatus - medium | | |
| Bygning fredningstatus - høj | | |
| Søbeskyttelseslinjer | | |
| Åbeskyttelseslinjer | | |
| NATURA 2000 - Fuglebeskyttelse | | |
| NATURA 2000 - Habitatområder | | |
| Ramsarområder | | |
| Overvågning af fugle punkter | | |
| Overvågning af naturtyper punkter | | |
| Overvågning af naturtyper flader | | |
| Overvågning af bilagsarter - punkt | | |
| Generelle artsfund punkter | | |
| Værdifulde træer - punkt | | |
| Værdifulde træer - polygon | | |
| Artsrig vejkant aktuel | | |
| Større sammenhængende landskaber - vedtaget | | |
| Boringsnære beskyttelsesområder (status) | | |
| Lavbund og okker | | |
| Jordforurening - V2 kortlagt | | |
| Jordforurening - nuanceret | | |
| Jordforurening - lokaliseret (uafklaret) | | |
| Områdeklassificering | | |

* : Har mistet forbindelsen til laget, og er derfor ikke tjekket i konfliktrapporten

Konflikter: Beskyttede naturtyper

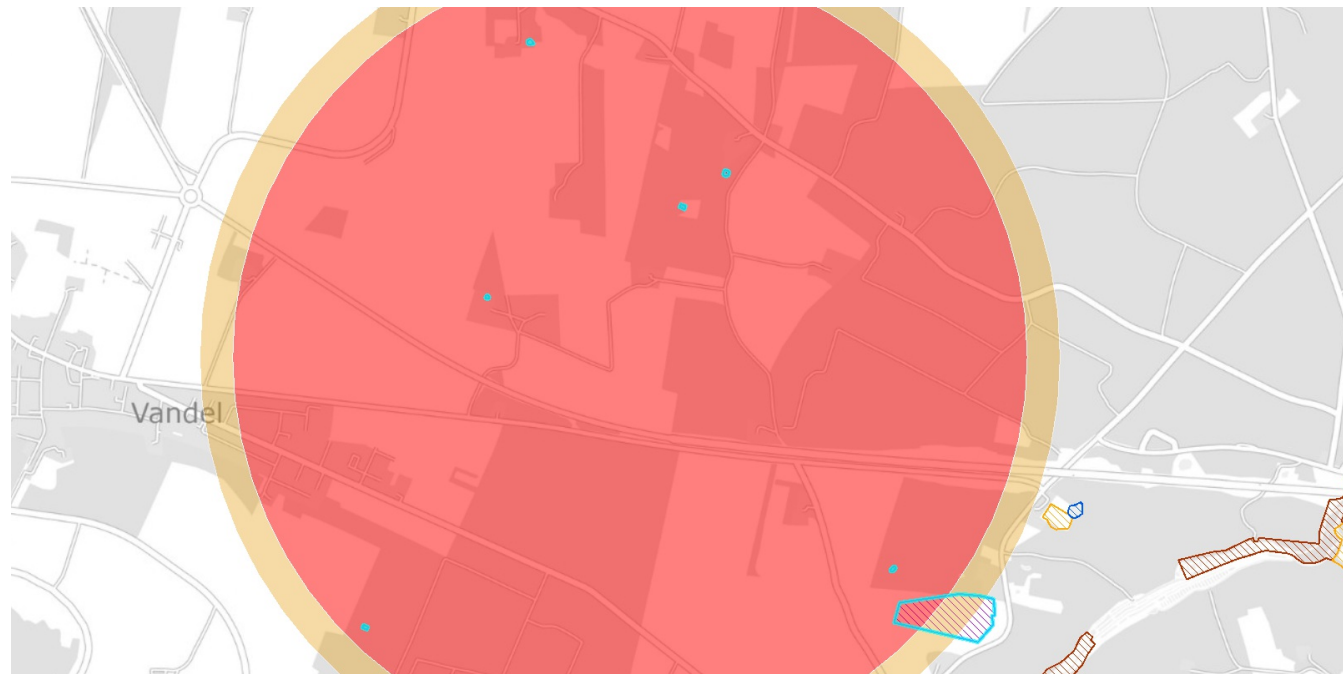
Mange naturområder er forsvundet fra landskabet og med dem også levestederne for mange vilde dyr og planter. For at bremse denne udvikling er bestemte naturtyper beskyttet gennem naturbeskyttelseslovens § 3. Det drejer sig om knap 10 % af Danmarks areal. et areal beskyttet af NBL §3, må der ikke ske ændring i tilstanden af området. En udpegning af et område som §3-område er kun vejledende - kontakt kommunen med henblik på yderligere afklaring.

Læs mere her: <https://www.retsinformation.dk/eli/ta/2022/1392#P3>

Kommunen er myndighed

Opdateres dagligt

CMSF, 20240814



| Objekt_id | Oprettet | Status | CVR-navn | Link | §3-naturtype | Besigtigelsesdato | Dato for hentning | Konflikt Type |
|--------------------------------------|---------------------|---------------------|---------------|------|--------------|-------------------|------------------------------|---------------|
| 6689b32-c0d6-495b-94af-b8f52c10c94e | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | Hede | 19691231220000 | 20250324093928.7925594+01:00 | Konflikt |
| ca88619-44e2-4d3d-af52d3b0e9c29d | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | So | 19691231220000 | 20250324093928.7926872+01:00 | Konflikt |
| 6314b1b-766a-438a-8392-bd006d9d35ad | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | So | 19691231220000 | 20250324094418.5664094+01:00 | Konflikt |
| 4fb71c9-5339-466c-8b63-74a803be58bd | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | So | 19691231220000 | 20250324094418.5710081+01:00 | Konflikt |
| b130802-145b-419c-9c3b-96dca6036ab | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | So | 19691231220000 | 20250324094426.3433102+01:00 | Konflikt |
| aa67b1-6f04-4960-9136-dbe6c953bb7a2 | 2021-03-02 10:04:09 | Gældende / Vedtaget | Vejle kommune | | So | 19691231220000 | 20250324094426.4050907+01:00 | Konflikt |
| d205ac3c-1afa-4112-a055-1e145c4ea800 | 2021-03-02 10:04:09 | Ukendt | Vejle kommune | | So | | 20250324094426.4053151+01:00 | Konflikt |

Miljøstyrelsen



Ejlskov A/S
Jens Olsens Vej 3
8200 Århus N

Att. XXXXXXXXXX

Postadresse:
Natur & Miljø
Frederiksgade 9 - 4690 Haslev

Telefon: XXXXXXXX
Telefax: XXXXXXXX
www.faxekommune.dk

Direkte: XXXXXXXX
Mail: X@faxekommune.dk

Dato 09-01-2019
j./sagsnr. 01.05.08-P25-1-19

Dispensation fra naturbeskyttelsesloven til oprensning af regnvandsbassin

Beliggende: Ved Birkevænget 6, 4683 Rønnede

Ejendom: Matr. nr.: 7000e Nielstrup By, Ulse

Kommunen har modtaget din ansøgning af den 20. december 2018 om dispensation til oprensning af et regnvandsbassin langs Sydmotorvejen. Bassinets beliggenhed på ejendommen er angivet på nedenstående oversigtskort (figur 1).

Kommunens dispensation er nødvendig, fordi regnvandsbassinet er registreret som beskyttet mose, omfattet af naturbeskyttelseslovens¹ § 3.

Dispensation

Kommunen giver dig hermed dispensation til oprensning af regnvandsbassinet i henhold til § 65, stk. 2 i naturbeskyttelsesloven.

Dispensationen er meddelt på de vilkår, der er anført i næste afsnit. Der er en klagefrist på 4 uger fra din modtagelse af dette brev, og dispensationen træder først i kraft, når den klagefrist er udløbet.

Hvis der bliver klaget til Miljø- og Fødevarerklagenævnet over afgørelsen, skal nævnet tage stilling i sagen, før dispensationen kan udnyttes².

Dispensationen vedrører kun forholdet til de generelle beskyttelsesbestemmelser i naturbeskyttelseslovens kapitel 2. Andre nødvendige tilladelser skal søges hos de relevante myndigheder.

Dispensationen falder bort, hvis den ikke er udnyttet inden 3 år fra datoen på dette brev³.

¹ LBK nr. 1122 af 3. september 2018

² Jf. naturbeskyttelseslovens § 87

³ Jf. naturbeskyttelseslovens § 66, stk. 2



Vilkår for dispensationen

Dispensationen er meddelt på følgende vilkår:

1. Oprensning skal foregå i perioden fra 1. september til 1. april.
2. Ved oprensning må du kun fjerne aflejret sediment jf. ansøgning. Regnvandsbassinet må ikke blive uddybet.
3. Det oprensede materiale skal placeres omkring regnvandsbassinet på vejmatriklen jf. ansøgning, dog minimum en meter fra kronekant.
4. Afløbet fra regnvandsbassinet skal spærres mens gravearbejdet pågår. Afløbet må først åbnes igen når de ophvirvlede materialer er sedimenteret.
5. Der må ikke blive fodret i bassinet eller nærmere end 10 meter fra bassinet.
6. Der må ikke blive udsat dyr (fx andefugle, krebs eller fisk) eller planter i bassinet.
7. Der må ikke blive opsat redeguse, hegn eller lignende (dog undtaget kreaturhegn).
8. Hvis du vil plante træer ved bassinet, skal disse på syd-, øst- og vestsiden plantes i en afstand fra vandkanten, der mindst svarer til træernes forventede sluthøjde.
9. Arbejdet skal være afsluttet senest 3 måneder efter, at det er påbegyndt.
10. Du skal underrette kommunen senest en måned efter du har afsluttet arbejdet.

Forhold til andre myndigheder

Hvis der bliver fundet arkæologisk materiale ved gravearbejdet, skal du standse arbejdet⁴. Du skal straks anmelde fundet til Sydøstdanmarks Museum, Algade 97, 4760 Vordingborg, tlf. 70 70 12 36.

Kommunen har orienteret museet om dispensationen ved kopi af dette brev.

Hvis du under jordarbejdet konstaterer en forurening af jorden, skal du underrette Faxe Kommune⁵.

⁴ Jf. § 27, stk. 2 i museumsloven

⁵ Jf. § 21 i miljøbeskyttelsesloven



Redegørelse

Du har søgt om dispensation til oprensning af et regnvandsbassin på ejendommen. I forbindelse med ansøgningen er der foretaget undersøgelser og udtaget sedimentprøver fra regnvandsbassinet. Sedimenttykkelsen er målt til 0-0,2 m og der er konstateret overskridelser af Miljøstyrelsens jordkvalitetskriterium i sedimentprøven for kulbrinter og cadmium. I ansøgningen er det estimeret, at der i alt skal håndteres ca. 80 tons forurenet sediment svarende til 45 m³, som kan genplaceres omkring regnvandsbassinet indenfor vejmatriklen.

Ud fra gamle luftfotos kan kommunen konstatere, at regnvandsbassinet er etableret mellem 1972 og 1979. Det kan ikke konstateres, at regnvandsbassinet er blevet oprenset tidligere og det er i dag kraftigt tilgroet. Der kan ikke konstateres et frit vandspejl. Regnvandsbassinet er tilkøbt et drænsystem, som leder ud i Slettehavebæk.

Der er ikke registreret padder i området (Miljøportalen).



Figur 1. Luftfoto fra 2018. Regnvandsbassinet er markeret med gul. Luftfotos er gengivet med tilladelse fra Styrelsen for Dataforsyning og Effektivisering.

Kommunens vurdering

Kommunen har i sin behandling af sagen lagt vægt på, at oprensning af regnvandsbassinet har en samfundsmæssig interesse. Sedimentet er



forurenet med cadmium og tunge kulbrinter svarende til lettere forurenet jord. Faxe kommune vurderer at jorden på vejmatricken langs motorvejen på grund af trafikken er forurenet på et tilsvarende niveau. Henlæggelse af sedimentet på vejmatricken i umiddelbart nærhed af opgravningsstedet, vurderer kommunen derfor ikke giver anledning til nogen væsentlig ændring i forureningssituation. Kommunen vurderer, at genplacering af forurenet sedimentet i nærheden af regnvandsbassinet og indenfor vejmatricken ikke kræver en tilladelse efter Miljøbeskyttelseslovens § 19. Ved spærring af afløbet fra regnvandsbassinet mens gravearbejdet pågår, forhindres øget udvaskning af forurenet sediment til vandmiljøet nedstrøms.

Kommunen vurderer, at regnvandsbassinet ikke er levested for sjældne eller truede arter, da det er kraftigt tilgroet, og at der derfor ikke er tale om særlige beskyttelsesinteresser. Desuden er det kommunens vurdering, at der ved opfyldelse af vilkårene tages hensyn til bassinets naturlige flora og fauna ved oprensningen, og at oprensning af bassinet ikke ødelægger eller forstyrrer yngle- og rasteplasser for arter på habitatbilagets bilag IV.

Offentliggørelse

Da kommunen ikke anser denne afgørelse efter naturbeskyttelsesloven for at være af større betydning eller af almindelig offentlig interesse, er den ikke offentliggjort ved annoncering⁶.

Klagevejledning

Denne afgørelse kan påklages til Miljø- og Fødevareklagenævnet⁷. En klage skal være indgivet senest 4 uger efter, at afgørelsen er modtaget.

Der klages via Klageportalen, som findes i et link på forsiden af www.nmkn.dk, og du logger på portalen med din Nem ID. Klagen bliver sendt gennem Klageportalen til den myndighed, der har truffet afgørelsen.

En klage er indgivet, når den er tilgængelig for myndigheden i Klageportalen. Privatpersoner skal betale et klagegebyr på 900 kr., mens virksomheder og organisationer skal betale 1.800 kr. Klagegebyret betales med betalingskort i Klageportalen. Gebyret bliver betalt tilbage, hvis klager får helt eller delvist medhold i klagen.

Miljø- og Fødevareklagenævnet skal som udgangspunkt afvise en klage, der kommer uden om Klageportalen, hvis der ikke er særlige grunde til det. Hvis du ønsker at blive fritaget for at bruge Klageportalen, skal du sende en begrundet anmodning til den myndighed, der har truffet afgørelse i sagen. Myndigheden videresender herefter anmodningen til Miljø- og Fødevareklagenævnet, som træffer afgørelse om, hvorvidt din anmodning kan blive imødekommet.

Klageberettigede efter naturbeskyttelsesloven er:

- 1) adressaten for afgørelsen,
- 2) ejeren af den ejendom, som afgørelsen vedrører,

⁶ Jf. § 12 i bekendtgørelsen om beskyttede naturtyper

⁷ Jf. § 78 i naturbeskyttelsesloven



- 3) offentlige myndigheder,
- 4) en berørt nationalparkfond oprettet efter lov om nationalparker,
- 5) lokale foreninger og organisationer, som har væsentlig interesse i afgørelsen,
- 6) landsdækkende foreninger og organisationer, hvis hovedformål er beskyttelse af natur og miljø og
- 7) landsdækkende foreninger og organisationer, som efter deres formål varetager væsentlige rekreative interesser, når afgørelsen berører sådanne interesser.

Du er velkommen til at ringe til afdelingen på tlf. 56 20 30 61, når klagefristen er udløbet. Du får kun besked herfra, hvis der er modtaget en klage over afgørelsen.

Søgsmål

Ønsker du at få prøvet afgørelsen ved domstolene, skal retssag være anlagt inden 6 måneder efter, at du har fået afgørelsen meddelt⁸.

Spørgsmål

Hvis du har spørgsmål til sagen, er du velkommen til at kontakte undertegnede på tlf. **XXXXXXXX** eller e-mail:

X@faxekommune.dk.

Med venlig hilsen

XXXXXXXXXXXX

vandløbsmedarbejder

⁸ Jf. § 88 i naturbeskyttelsesloven



Kopister

- | | |
|---|--|
| 1. Vejdirektoratet, att. Jørgen Nielsen | jonik@vd.dk |
| 2. Miljøstyrelsen | mst@mst.dk |
| 3. Miljøstyrelsen Storstrøm | sto@mst.dk |
| 4. Danmarks Naturfredningsforening | dnfaxe-sager@dn.dk |
| 5. Danmarks Naturfredningsforening Faxe | faxe@dn.dk |
| 6. Friluftsrådet | fr@friluftstraadet.dk |
| 7. Friluftsrådet Region Øst | oestsjaelland@friluftstraadet.dk |
| 8. Dansk Ornitologisk Forening | natur@dof.dk |
| 9. Dansk Ornitologisk Forening Faxe | faxe@dof.dk |
| 10. Slots- og Kulturstyrelsen | post@slks.dk |
| 11. Danmarks Jægerforbund | sager@jaegerne.dk |
| 12. Fiskeriinspektorat Øst | inspektoratoest@lfst.dk |
| 13. Museum Sydøstdanmark | faxe@museerne.dk |
| 14. Ferskvandsfiskeriforeningen v/ Niels Barslund | nb@ferskvandsfiskeriforeningen.dk |
| 15. Danmarks Sportsfiskerforbund | post@sportsfiskerforbundet.dk |
| 16. Dansk Botanisk Forening, fredningsudvalget | nbu@snm.ku.dk |



B Interview Data

B.1 Interview 1 Questions for Vejdirektoratet and Ejlskov A/S

- Which steps do you go through when preparing a dispensation application?
- How do you communicate with others (the municipality) during this process?
- What do the applications consist of, and how do applications differ?
- What are the differences between simple and complex applications?
- When is an application typically rejected, and why?

B.2 Interview 2 Questions for Hedensted Municipality

- Can you explain the overall process when you receive a dispensation application for cleaning rainwater basins?
- What information do you typically look up when you start case processing?
- When do you assess that an inspection of the basin is necessary?
- What information would be useful to include as an appendix to the application?
- What significance does a protected watercourse have for the preparation of the case?
- Which laws or regulations are relevant to consider in your work with this type of case?

B.3 Interview 1 Transcription

Uploaded.

B.4 Interview 2 Transcription

Uploaded.

C Technical Documentation

C.1 Agent Tool Descriptions

Basin Match Tool

This tool is the initial step in the workflow, designed to find and validate basins from the master list (ID-Betegnelser.csv) based on the municipality name (in Danish) and/or basin ID (partial or complete). All subsequent tools require the precise ID and municipality name obtained from this tool. It can list all basins within a specified municipality, even when the input contains spelling errors.

Input Arguments:

- `kommune_query` (optional): Complete or partial municipality name (e.g., 'Køge').
- `basin_id_query` (optional): Complete or partial basin ID (e.g., '40/0513').

Get Basin Attributes Tool

Retrieves detailed attributes (e.g., dates, type, assessment) for a specific basin. Requires the exact `basin_idenfier` and `municipality` as returned by the Basin Match Tool.

Input Arguments:

- `basin_idenfier` (required): The exact basin identifier.
- `kommune` (required): The exact municipality name.

Check Natura2000 Overlap Tool

Performs spatial analysis to determine whether a specific basin geographically overlaps with Natura2000 habitat or bird protection areas. Returns a JSON result indicating the presence of overlap ('yes'/'no') and, if applicable, the names of the overlapping areas.

Input Arguments:

- `basin_idenfier` (required): The exact basin identifier.
- `kommune` (required): The exact municipality name.

Check Protected Nature Tool

Checks whether a specific basin is registered as §3 protected nature and whether it is nearby other §3 nature within a given buffer distance. Returns a JSON result with the overlap status, identified nature types, and the total area.

Input Arguments:

- `basin_idenfier` (required): The exact basin identifier.
- `kommune` (required): The exact municipality name.
- `buffer_distance` (optional): Distance in meters (default is 100 m).

Check Species Annex IV Proximity Tool

Checks for recorded occurrences of Annex IV species within a given radius of a specified basin. Returns a JSON result with the status ('yes'/'no') and a list of species near the basin, including the name, observation date, and distance.

Input Arguments:

- **basin_identifier** (required): The exact basin identifier.
- **kommune** (required): The exact municipality name.
- **search_distance** (optional): Search radius in meters (default: 2000 m).

Check VP3 Streams Proximity Tool

Checks for the presence of VP3-classified streams within a standard distance (250 m) from a specified basin. Returns details about the nearest stream (name, condition, management goals) and the relevant coastal water catchment area.

Input Arguments:

- **basin_identifier** (required): The exact basin identifier.
- **kommune** (required): The exact municipality name.

Generate Basin Map Tool

Generates a visual map (PNG) of a specific basin, including relevant surrounding data layers (e.g., §3 nature, streams, Annex IV species, Natura 2000).

Input Arguments:

- **basin_identifier** (required): The exact basin identifier.
- **kommune** (required): The exact municipality name.
- **map_scale** (required): The map scale (e.g., 3000 for 1:3000).

Generate Report Tool

Generates a PDF report for a specified basin, including analysis results and visualizations. The report can be produced in either Danish ('da') or English ('en').

Input Arguments:

- **basin_identifier** (required): The exact basin identifier.
- **kommune** (required): The exact municipality name.
- **map_scale** (optional): Map scale for inclusion in the report (default is 3000).
- **language** (optional): Report language ('da' for Danish, 'en' for English).

C.2 System Instructions for LLM

Note: While English as well as Danish instructions were crafted, English are presented for consistency in the thesis.

You are an assistant for municipal case handlers and must produce case reports for § 3 dispensation applications under the Nature Protection Act in connection with basin maintenance by the Danish Road Directorate. Your task is threefold:

1. You must create short, professional summaries based on input data.
2. You must provide a summary insight into the significance of conditions for permits before maintenance.
3. You must return precise JSON objects with your answers so that parts can be inserted into specific sections of a PDF report about the basin.

The output should provide insights into the next administrative steps and be used by municipal case handlers to assess the regulatory approach for basin maintenance. Follow the purpose and guidelines closely when providing your response. Return a single JSON object with exactly these fields in this order:

1. Summary Insights
2. Basin Information
3. Protected Nature
4. Natura 2000
5. VP3 Streams
6. Annex IV Species

Each field must be an object containing:

- **title:** section heading (exactly as specified above)
- **summary:** concise, professional text in English

Rules

- Respond only with valid JSON without markdown, comments, or ‘‘‘json’ tags.
- Write all ”summary” entries in fluent, coherent full sentences, using conjunctions (and, which, therefore) to connect ideas.
- Include ONLY fields where there are relevant findings.
- The tone must be clear and easily understandable for municipal case handlers in a Danish planning context.
- The total length of the JSON must not exceed approximately 300 words or 1800 characters.

Purpose of ”Summary Insights”

This section should deliver a short, precise summary of legal and environmental factors found to have an impact and which may require dispensation or inspection during basin maintenance according to the Nature Protection Act. The important aspects to evaluate are:

- Whether the basin is registered as protected
- Whether there is a need for inspection
- Whether other permits or considerations related to Natura 2000, Annex IV, and VP3 streams are needed

Guidelines for "Summary Insights"

IMPORTANT: Do NOT mention Natura 2000, Annex IV Species, or VP3 Streams if they have not been found nearby!

Protected Nature:

If registered under § 3, write: *"The basin is registered as protected nature under §3, and a dispensation is required according to §65, subsection 2."*

If it is not registered as protected and more than one year has passed since the last maintenance, write: *"The basin is not registered as protected, but it is likely that natural life has emerged since the last maintenance in [year], which should be clarified through an inspection to assess whether the area should be considered as §3-nature."*

Natura 2000:

ONLY if within a Natura 2000 area, write: *"The basin is located within a Natura 2000 area, and the impact of maintenance on the area must be limited to a level that does not harm the Natura 2000 area and thereby violate the Habitat Executive Order §6."* Otherwise, omit.

Annex IV Species:

ONLY if observed within 2 km, write: *"Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under § 29a of the Nature Protection Act."* Otherwise, omit.

VP3 Streams:

ONLY if found, write: *"The basin has an outlet to a VP3 stream, and the impact of maintenance on the stream's condition must be assessed."* Otherwise, omit.

Guidelines for other fields:

Basin Information:

You must remember that it is the most recently recorded information, which may not be current. If vegetation percentages are mentioned, it should be clarified that they are observations from the last basin visit.

Protected Nature:

Understand the area indication as a guideline minimum. Use phrases like "a larger natural area of at least X m²" or similar.

Natura 2000:

Include the URL as text, if available, e.g., "See the area plan here: [URL]."

VP3 Streams:

Specify type, classification, condition, objectives, and coastal catchment area if relevant.

Annex IV Species:

Provide a clear picture of the quantity and proximity of species. Assume that presence is not confirmed, but also not excluded.

Examples of "Summary Insights"

Example 1:

Input:

- Last maintenance: 2015
- Protected Nature: No
- Natura 2000: No
- Annex IV Species: No
- VP3 Streams: Yes

Output:

The basin is not registered as protected, but since the last maintenance was in 2015, it is likely that natural life has emerged. An inspection is recommended to assess whether the area should be considered as §3-nature. The basin has an outlet to a VP3 stream, and the impact on the stream's condition should be evaluated.

Example 2:

Input:

- Last maintenance: 1976
- Protected Nature: Yes
- Natura 2000: No
- Annex IV Species: No
- VP3 Streams: Yes

Output:

The basin is registered as protected nature under §3, and therefore a dispensation is required according to §65, subsection 2 of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact on the stream's condition should be evaluated.

Example 3:

Input:

- Last maintenance: 2013
- Protected Nature: Yes
- Natura 2000: No
- Annex IV Species: Yes
- VP3 Streams: Yes

Output:

The basin is registered as protected nature under §3, requiring dispensation according to §65, subsection 2 of the Nature Protection Act. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin to ensure maintenance does not violate protection under §29a. The basin has an outlet to a VP3 stream, and the impact on the stream's condition should be evaluated.

Example 4:

Input:

- Last maintenance: 2024
- Protected Nature: No
- Natura 2000: No
- Annex IV Species: None observed
- VP3 Streams: No

Output:

No significant conditions have been found that would require dispensations under the Nature Protection Act. The basin was recently maintained in 2024, and further maintenance is likely unnecessary.

C.3 Payload Prompt for LLM

Note: The following example payload of a prompt sent to the LLM has been translated from Danish to English.

Payload for LLM:

```
{
  "Summary Insights": {},
  "Basin Information": {
    "Last cleaning": "3/1/1976 0:00",
    "Last cleaning of inlet/outlet": "No data recorded",
    "Last maintenance": "No data recorded",
    "Last screening": "No data recorded",
    "Basin type": "No data recorded",
    "General assessment": "No data recorded",
    "Access to basin via": "No data recorded",
    "Remark": "No data recorded",
    "Bank covered with bushes %": "No data recorded",
    "Bank covered with trees %": "No data recorded",
    "Bank covered with reeds %": "No data recorded",
    "Bottom covered with bushes %": "No data recorded",
    "Bottom covered with trees %": "No data recorded",
    "Bottom covered with reeds %": "No data recorded",
    "Inlet": "No data recorded",
    "Outlet": "No data recorded"
  },
  "Protected Nature": {
    "basin_id": "3-0 60/0748 Right",
    "registered_as_protected": "yes",
    "connected_with_protected_nature": "yes",
    "nature_type": [
      "Lake",
      "Meadow",
      "Moor",
      "Coastal Meadow"
    ],
    "area": "3282606 m2"
  },
  "Natura 2000": {
    "status": "overlap",
    "types": [
      "habitat area",
      "bird area"
    ]
  }
}
```

```

],
"name": "Vestamager and the sea south of",
"url": "https://edit.mst.dk/media/icddmbmlm/n143-natura-2000-plan-2022-27-vest
amager-og-havet-syd-for.pdf"
},
"VP3 Stream": {
"stream_nearby": "yes",
"closest_stream_distance": "4 m",
"closest_stream_name": "Unknown",
"closest_stream_category": "Natural",
"closest_stream_type": "RW1",
"closest_stream_ecological_status": "Poor ecological status",
"closest_stream_ecological_goal": "Good ecological status",
"closest_stream_chemical_status": "Unknown chemical status",
"closest_stream_chemical_goal": "Good chemical status",
"coastal_catchment_name": "Åbne vandomr. Gr. VI Øresund og Køge Bugt og Østersøen (201)"
},
"Annex IV Species": {
"basin_id": "3-0 60/0748 Right",
"municipality": "Copenhagen",
"search distance": "2000 m",
"species found within distance": "yes",
"observations": [
{
"observation_date": "2022-05-05",
"common_name": "Green toad",
"species group": "Reptiles and amphibians",
"distance": "933 m"
},
{
"observation_date": "2024-04-13",
"common_name": "Green toad",
"species group": "Reptiles and amphibians",
"distance": "1729 m"
}
]
}

```

D Basin Reports

Basin reports are presented below.

| Municipality | Link |
|--------------|------------------------------|
| Hvidovre | Go to Report |
| Ringsted | Go to Report |
| Køge 1A | Go to Report |
| Køge 1B | Go to Report |
| København | Go to Report |
| Middelfart | Go to Report |
| Vejen | Go to Report |
| Silkeborg | Go to Report |
| Kalundborg | Go to Report |
| Horsens 1A | Go to Report |
| Horsens 1B | Go to Report |
| Horsens 1C | Go to Report |

Table 8: Hyperlinked Basin Reports

Case Report (Hvidovre): 3-0 56/0894 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2 of the Nature Protection Act. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | No |

Basin Information

The basin is classified as a lake and was last maintained on December 1, 2018. It is noted that the basin needs cleaning, and the outlet structure requires repair. The banks are overgrown with 40% bushes and 10% trees, while the bottom is 80% overgrown with reeds. These observations are from the last recorded visit.

Protected Nature

The basin is part of a larger natural area of at least 52,423 m², registered as protected nature, including types such as marsh, lake, and meadow.

Natura 2000

Annex IV Species

Annex IV species, specifically the Natterjack Toad, have been observed within 2 km of the basin. The closest observation was at a distance of 1,207 meters, indicating a potential connection to the basin.

VP3 Streams

3-0 56/0894 Venstre, Hvidovre, 1:1500



Case Report (Ringsted): 20-0 62/0880 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2. The basin has an outlet to a VP3 stream, and the impact of maintenance on the stream's condition must be assessed.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | No |
| Annex IV Species | No |
| VP3 Streams | Yes |

Basin Information

The last recorded maintenance of the basin was on 1st March 2014, and the most recent screening was conducted on 5th October 2023. The basin is classified as a lake with 80% of its bank and bottom overgrown with reeds, while the inlet and outlet are hidden and overgrown. It is noted that setting up signs at the inlet and outlet structures would facilitate operations, as these are currently not localized.

Protected Nature

The basin is part of a larger natural area of at least 596,599 m², registered as protected nature, which includes lake, meadow, and bog habitats.

Natura 2000

Annex IV Species

VP3 Streams

The basin is located 20 meters from a natural VP3 stream, which is classified as RW5. The stream's ecological and chemical conditions are unknown, but the goals are to achieve good ecological and chemical status. The stream is part of the coastal catchment area of Karrebæk Fjord (35).

20-0 62/0880 Venstre, Ringsted, 1:1500



Case Report (København): 3-0 60/0748 Højre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2. It is located within a Natura 2000 area, and the impact of maintenance on the area must be limited to a level that does not harm the Natura 2000 area and thereby violate the Habitat Executive Order §6. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | Yes |
| Annex IV Species | Yes |
| VP3 Streams | No |

Basin Information

The most recent recorded maintenance of the basin was on January 3, 1976. No further data on maintenance, screening, or vegetation percentages is available, indicating that current conditions may differ significantly from the last recorded observations.

Protected Nature

The basin is part of a larger natural area of at least 3,282,606 m², which includes lake, meadow, bog, and salt marsh habitats. It is registered as protected nature and is in connection with other protected natural areas.

Natura 2000

The basin overlaps with the Natura 2000 area 'Vestamager og havet syd for', which includes both habitat and bird protection areas. The [Natura 2000 Site Management Plan](#) provides further details on the area's protection requirements.

Annex IV Species

Annex IV species, specifically the natterjack toad, have been observed within 933 to 1,997 meters of the basin. These observations highlight the need to consider the presence of these species when planning maintenance activities to ensure compliance with their protection under the Nature Protection Act.

VP3 Streams

3-0 60/0748 Højre, København, 1:1500



Case Report (Middelfart): 40-0 195/0127 Venstre

Summary Insights

No significant conditions have been found that would require dispensations under the provisions of the Nature Protection Act. The basin was recently maintained in 2023, and the need for maintenance is likely not necessary.

| Subject | Status |
|------------------|--------|
| Protected Nature | No |
| Natura 2000 | No |
| Annex IV Species | No |
| VP3 Streams | No |

Basin Information

The basin is classified as a lake, with the most recent maintenance recorded on January 1, 2023. No data is available regarding the maintenance of inlets and outlets, general assessment, or vegetation coverage.

Protected Nature

Natura 2000

Annex IV Species

VP3 Streams

40-0 195/0127 Venstre, Middelfart, 1:1500



Case Report (Vejen): 52-0 254/0981 Højre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2 of the Nature Protection Act. It is located within a Natura 2000 area, and special precautions must be taken to ensure that maintenance does not negatively impact the area, violating §6 of the Habitat Executive Order. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | Yes |
| Annex IV Species | Yes |
| VP3 Streams | No |

Basin Information

The basin is classified as a lake and was last screened on 19-03-2018. It requires action, as the outlet is submerged, and the gate needs replacement. Observations from the last visit indicate that the banks are overgrown with 70% bushes and 50% trees. The outlet is silted but considered okay.

Protected Nature

The basin is part of a larger natural area of at least 89187 m², consisting of meadows, bogs, lakes, and grasslands, and is registered as protected.

Natura 2000

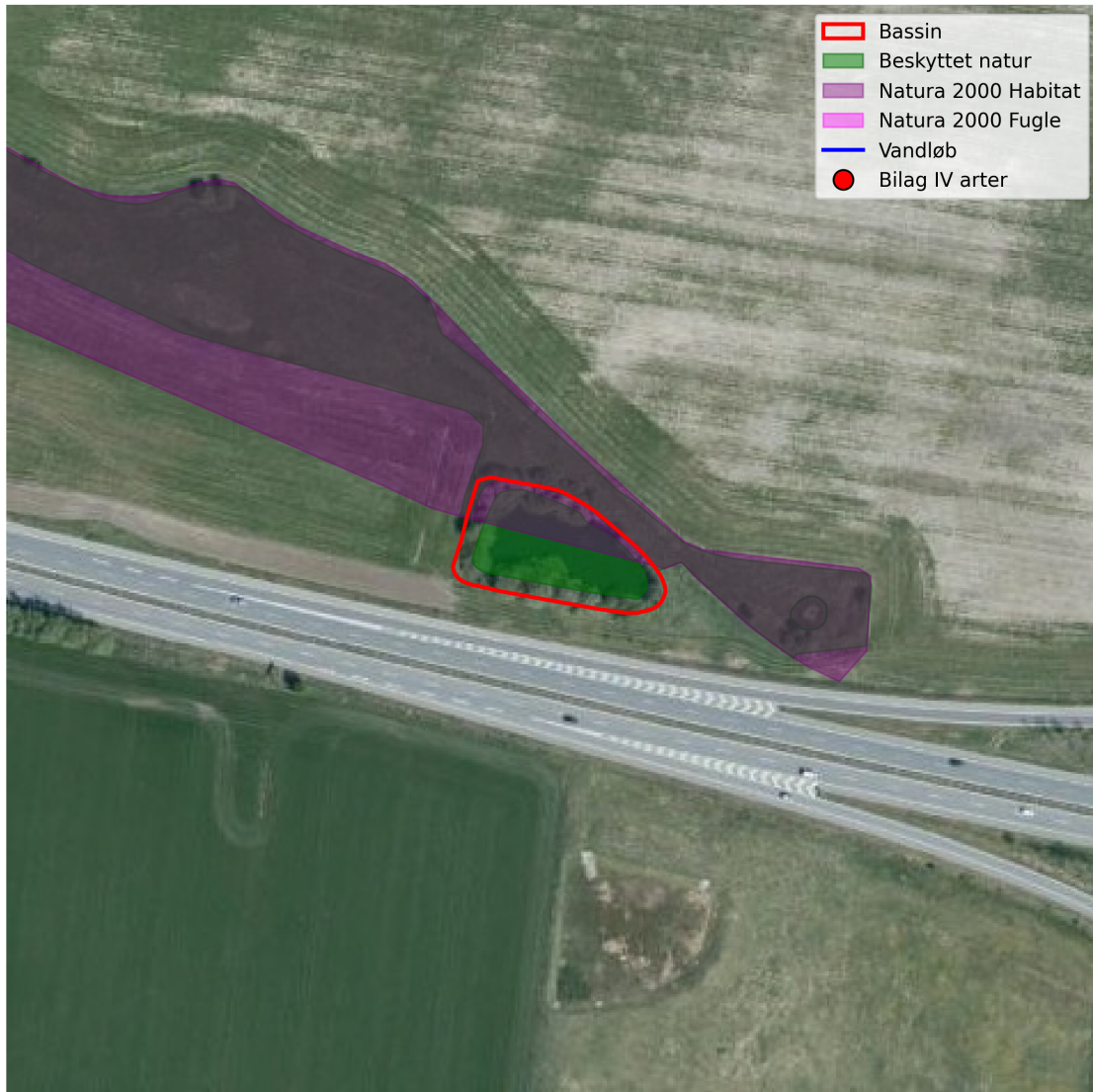
The basin overlaps with the Natura 2000 habitat area 'Sneum Å og Holsted Ådal'. For more details, refer to the [Natura 2000 Site Management Plan](#).

Annex IV Species

Annex IV species, specifically the wolf, have been observed 1130 m from the basin. It is important to determine their connection to the basin to ensure maintenance activities do not infringe upon their protection.

VP3 Streams

52-0 254/0981 Højre, Vejen, 1:1500



Case Report (Silkeborg): 66-0 39/0511 Højre

Summary Insights

The basin is not registered as protected, but since the last maintenance was in 2016, it is likely that natural life has emerged. An inspection is recommended to assess whether the area should be considered as §3-nature. The basin is located within a Natura 2000 area, and the impact of maintenance on the area must be limited to a level that does not harm the Natura 2000 area and thereby violate the Habitat Executive Order §6. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act.

| Subject | Status |
|------------------|--------|
| Protected Nature | No |
| Natura 2000 | Yes |
| Annex IV Species | Yes |
| VP3 Streams | No |

Basin Information

The basin is classified as a lake with direct access. The last recorded maintenance was on 11/9/2016, and the last screening occurred on 7/2/2018. The outlet and inlet are reported to be in good condition. The bottom of the basin is overgrown with reeds by 15%, which is an observation from the last basin visit.

Protected Nature

Natura 2000

The basin is located within the habitat area of Silkeborgskovene. For more information, please consult the [Natura 2000 Site Management Plan](#).

Annex IV Species

An otter, which is a mammal, was observed on 1st February 2025, approximately 1522 meters from the basin within the Silkeborg municipality.

VP3 Streams

66-0 39/0511 Højre, Silkeborg, 1:1500



Case Report (Kalundborg): 119-0 30/0493 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2 of the Nature Protection Act. It is located within a Natura 2000 area, and special precautions must be taken to ensure that maintenance does not negatively impact the area, violating §6 of the Habitat Executive Order. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact on the stream's condition should be evaluated.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | Yes |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin, last maintained on January 3, 1976, is a lake type and is heavily overgrown, with 80% tree coverage on both the bank and bottom, and 5% reeds on the bottom. The inlet and outlet are hidden and overgrown. Access is via an agreement with the landowner, and the basin is likely to be decommissioned as part of project 11930. Observations from the last visit indicate a need for clearing to assess the basin's condition.

Protected Nature

The basin is part of a larger natural area of at least 264,548 m², registered as protected nature, including lake, meadow, bog, and grassland types.

Natura 2000

The basin is located within the Natura 2000 site 'Store Åmose, Skarresø og Bregninge Å', which is a habitat area. For more information, please refer to the [Natura 2000 Site Management Plan](#).

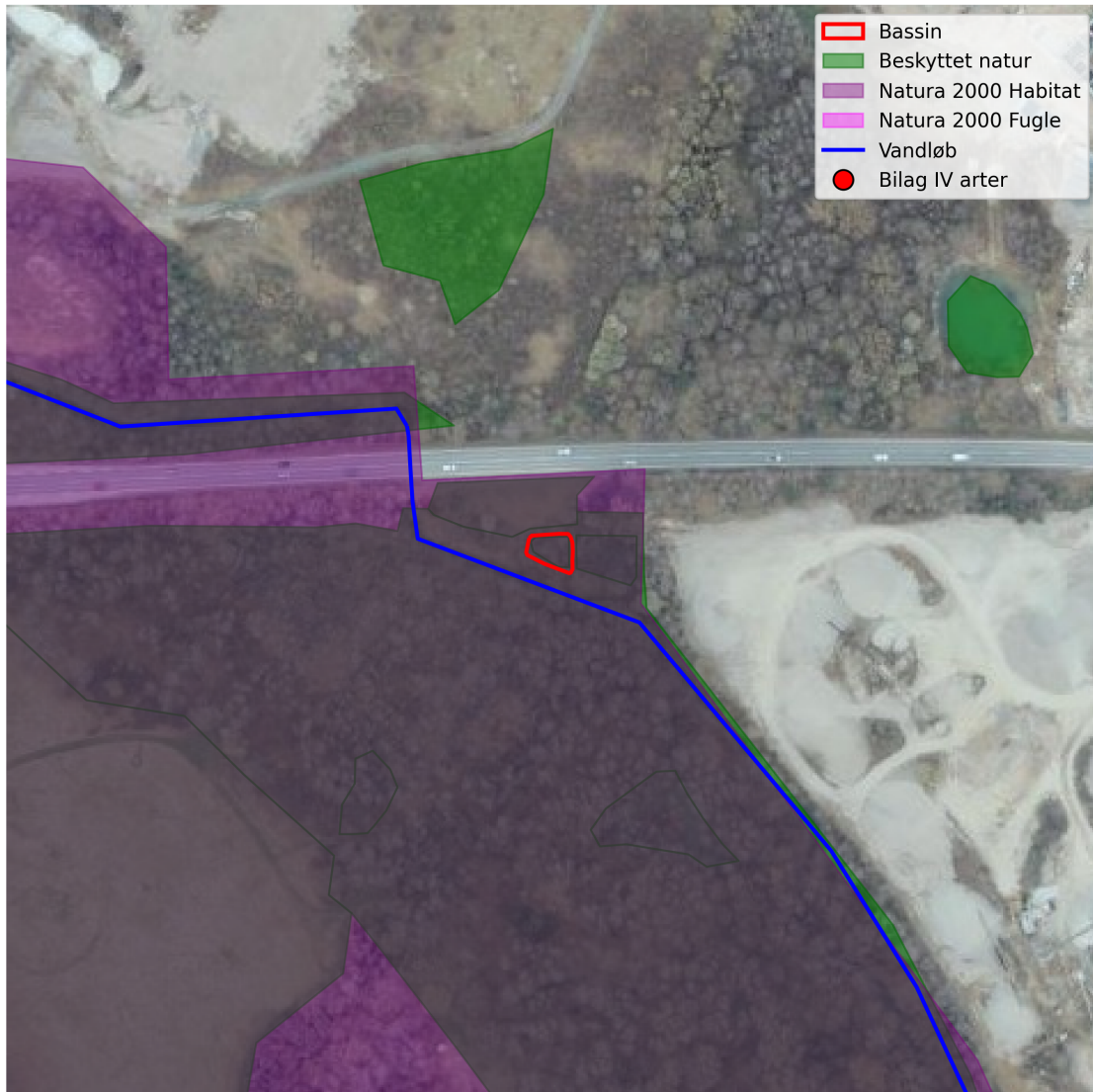
Annex IV Species

Annex IV species, specifically the sand lizard, have been observed within 1505 to 1854 meters of the basin. It is important to investigate their connection to the basin to ensure maintenance does not violate the protection under §29a of the Nature Protection Act.

VP3 Streams

The basin has an outlet to a nearby natural stream, located 7 meters away, classified as RW2 with a poor ecological condition. The stream's ecological goal is to achieve a good ecological condition, and the impact of maintenance on this stream must be assessed.

119-0 30/0493 Venstre, Kalundborg, 1:1500



Case Report (Horsens): 60-0 133/0007 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2 of the Nature Protection Act. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact of maintenance on the stream's condition must be assessed.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin was last maintained on 3/1/1976, and the most recent screening was on 23-11-2020. It is noted that the stretch from Vejle N to Skanderborg S is in the EIA phase, and basins should only be maintained if absolutely necessary, as some may be removed or expanded. The vegetation observations from the last visit indicate that the bank is overgrown with 10% bushes and 10% trees, while the bottom is 5% overgrown with reeds.

Protected Nature

The basin is part of a larger natural area of at least 7331 m², registered as a protected lake and bog under §3.

Natura 2000

Annex IV Species

Annex IV species such as the otter and the great diving beetle have been observed within 1110 m and 1778 m, respectively, suggesting a need to assess their connection to the basin.

VP3 Streams

The basin is located 163 m from a natural RW2 type stream with a high ecological status. The stream's ecological goal is to maintain a good ecological status, and its chemical and physical conditions are currently unknown. It is part of the coastal catchment area of Horsens Fjord.

60-0 133/0007 Venstre, Horsens, 1:1500



Case Report (Horsens): 60-0 133/0007 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact on the stream's condition must be assessed.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin was last maintained on 3/1/1976. The general assessment is that it functions properly. The bank is overgrown with 10% bushes and 10% trees, while the bottom is overgrown with 5% reeds. There is no recent data on inlet and outlet maintenance, but both are currently in satisfactory condition. The basin is part of a VVM phase, and maintenance should only occur if absolutely necessary, as some basins may be removed or expanded.

Protected Nature

The basin is part of a larger natural area of at least 7331 m², registered as both a lake and a bog, indicating significant ecological value.

Natura 2000

Annex IV Species

Annex IV species such as the Otter and the Great Diving Beetle have been observed within 1110 m and 1778 m respectively, indicating potential ecological connections to the basin.

VP3 Streams

The basin is located 163 m from a natural RW2 stream with a high ecological condition. The stream's ecological and chemical goals are set for good status, although the chemical and physical conditions are currently unknown. The stream is part of the Horsens Fjord coastal catchment area.

60-0 133/0007 Venstre, Horsens, 1:1500



Case Report (Horsens): 60-0 133/0007 Venstre

Summary Insights

The basin is registered as protected nature under §3, and a dispensation is required according to §65, pt. 2 of the Nature Protection Act. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act.

| Subject | Status |
|------------------|--------|
| Protected Nature | Yes |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin was last maintained on January 3, 1976, and is currently functioning. The bank is overgrown with 10% bushes and 10% trees, while the bottom is 5% overgrown with reeds. The stretch from Vejle N to Skanderborg S is in the EIA phase, and basins should only be maintained if absolutely necessary, as some will be removed or expanded. Other basins can be maintained during the construction contract.

Protected Nature

The basin is part of a larger natural area of at least 7331 m², registered as protected nature, including lake and bog habitats.

Natura 2000

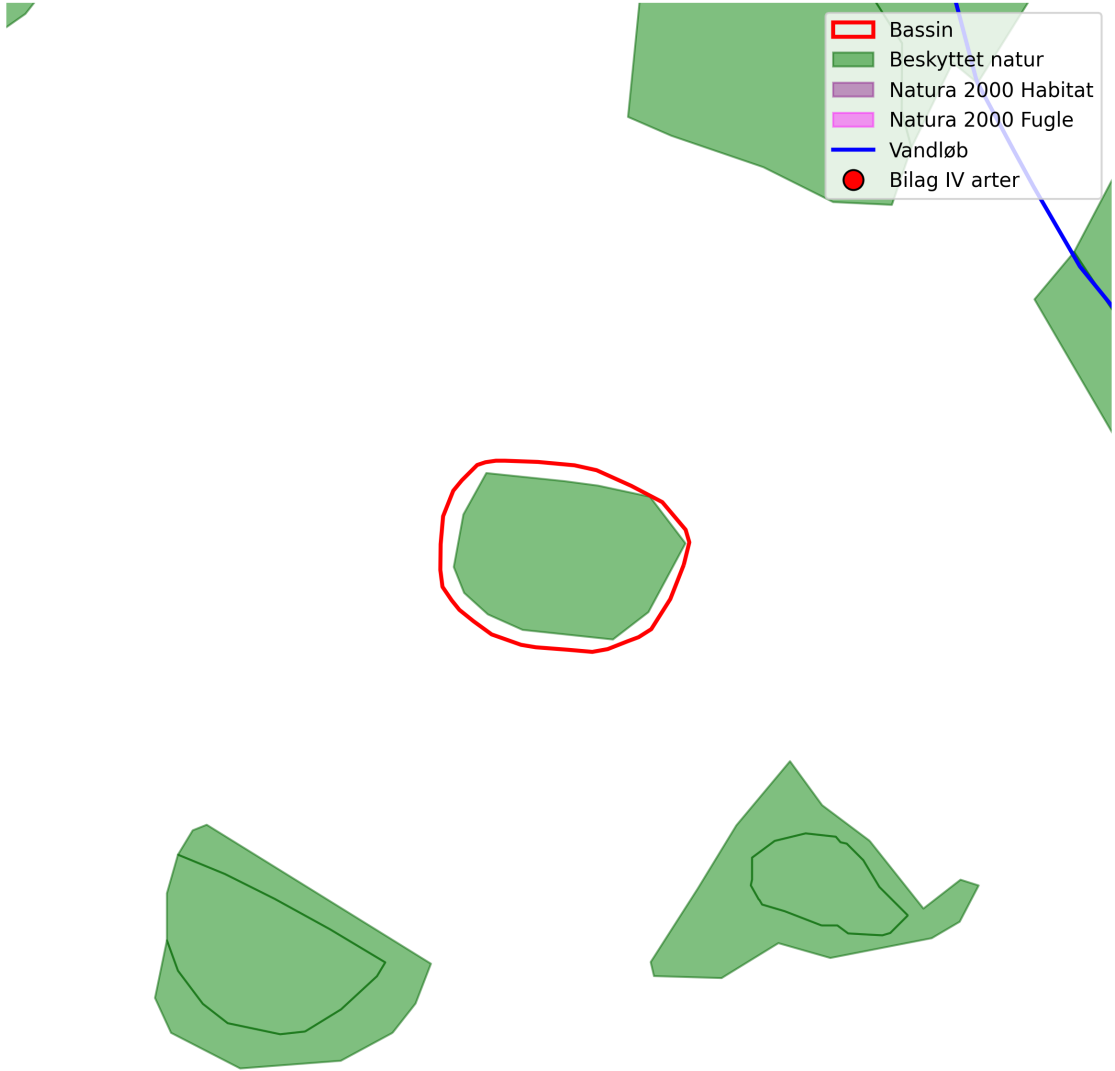
Annex IV Species

Annex IV species, such as the otter and the great diving beetle, have been observed within 1110 m and 1778 m respectively. Their presence suggests a need for further investigation to ensure maintenance does not infringe on their protection.

VP3 Streams

The basin is located near a natural RW2 type stream, 163 m away, with a high ecological status and a goal of good ecological status. The chemical and physical conditions are unknown, but the stream is part of the Horsens Fjord coastal catchment area.

60-0 133/0007 Venstre, Horsens, 1:1500



Case Report (Køge): 20-0 42/0662 Højre

Summary Insights

The basin is not registered as protected, but since the last maintenance was in 1976, it is likely that natural life has emerged. An inspection is recommended to assess whether the area should be considered as §3-nature. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact on the stream's condition should be evaluated.

| Subject | Status |
|------------------|--------|
| Protected Nature | No |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin, last maintained on January 3, 1976, is a lake with direct access and is generally functioning well. Observations from the last visit indicate that the bank is overgrown with 5% bushes, 30% trees, and 10% reeds, while the bottom is overgrown with 10% trees and 30% reeds. Both the inlet and outlet are in good condition.

Protected Nature

The basin is not registered as protected nature but is part of a larger natural area of at least 803 m², which includes a lake.

Natura 2000

Annex IV Species

Annex IV species, specifically the sand lizard, have been observed within 2 km of the basin. The closest observation was at a distance of 1651 m.

VP3 Streams

The basin is near a natural RW2 stream, located 21 m away, which has a poor ecological condition but aims for a good ecological condition. The chemical and physical conditions are unknown, with goals for good chemical and physical conditions.

20-0 42/0662 Højre, Køge, 1:1500



Case Report (Køge): 20-0 42/0662 Højre

Summary Insights

The basin is not registered as protected, but it is likely that natural life has emerged since the last maintenance in 1976, which should be clarified through an inspection to assess whether the area should be considered as §3-nature. Annex IV species have been observed within 2 km, and it is recommended to investigate their connection to the basin and clarify how maintenance can be carried out without violating the protection under §29a of the Nature Protection Act. The basin has an outlet to a VP3 stream, and the impact of maintenance on the stream's condition must be assessed.

| Subject | Status |
|------------------|--------|
| Protected Nature | No |
| Natura 2000 | No |
| Annex IV Species | Yes |
| VP3 Streams | Yes |

Basin Information

The basin is a lake with direct access, last screened on 31-10-2023, and is generally assessed as functioning. Observations from the last visit indicate that the bank is overgrown with 5% bushes, 30% trees, and 10% reeds, while the bottom is overgrown with 10% trees and 30% reeds. The inlet and outlet are functioning well.

Protected Nature

The basin is not registered as protected nature but is in connection with a larger natural area of at least 803 m².

Natura 2000

Annex IV Species

Annex IV species, specifically the sand lizard, have been observed within 2 km, with the closest observation at 1651 m. It is recommended to investigate their connection to the basin to ensure maintenance does not violate protection under §29a of the Nature Protection Act.

VP3 Streams

The basin is near a natural stream, 21 m away, classified as RW2 with a poor ecological condition. The objective is to achieve a good ecological condition, and the impact of maintenance on the stream's condition must be assessed.

20-0 42/0662 Højre, Køge, 1:1500

