Basic Climate Interaction Knowledge Base

Project KNOWING

Work package 1, Deliverable D1.4

Basic Climate Interaction Knowledge Base

Work package 1, Deliverable D1.4

Please refer to this report as follows:

Kaleta, P., Havlik, D., Parzer, V. (2024). Basic Climate Interaction Knowledge Base. Deliverable D1.4 of the Horizon Europe project KNOWING.

Project information	
Project name:	KNOWING – Framework for defining Climate Mitigation Pathways based on Understand- ing and integrated Assessment of Climate Impacts, Adaptation Strategies and Societal Transformation
Grant Agreement No.	101056841
Start date:	01/06/2022
Duration:	48 months
Coordinator:	Alexandra Millonig, Senior Scientist AIT Austrian Institute of Technology Giefinggasse 4, 1210 Vienna, Austria
	Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or CINEA. Neither the European Union nor the granting authority can be held responsible for them.
Deliverable details	
Description:	Deliverable description
Version:	Final
Dissemination level:	PU (Public)
Due date:	30/06/2024
Submission:	28/06/2024
Lead:	Kaleta, Patrick – Austrian Institute of Technology
Author(s):	Kaleta, P., Havlik, D., Parzer, V. (AIT), Austria

Legal Disclaimer

All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user, therefore, uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission and CINEA has no liability in respect of this document, which is merely representing the authors' view.

Users wishing to reuse material from this work that is attributed to a third party, such as tables, figures or images, are responsible for determining whether permission is needed for that reuse and for obtaining permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

Table of Contents

Content

Executive summary 1
1 KNOWING summary 2
2 Objectives of the Deliverable 3
3 Introduction 4
3.1 Benefits for KB application users 4
3.2 Benefits for tool/service developers 5
4 Relevant Datasets 6
4.1 Process for including datasets 6
4.2 Datasets to be added for next iteration of Knowledge Base 7
4.3 Discovering datasets of interest 7
5 System Design 9
5.1 Neo4j Database 10
5.2 Apollo with GraphQL 11
5.3 Frontend applications 13
6 Final Knowledge Base planning 16
6.1 Incorporating new data and relations 16
6.2 Incorporating the new user requirements 16
6.3 Application requirements 17
6.4 MAIA connection 17
7 Conclusions 18
References 19
Annex A: Data Summary 20
Annex B: Ethics Considerations 21

List of Figures

Figure 1 Current data schema for the initial version of the Knowledge base. 7
Figure 2: Architectural overview of the Knowledge Base 10
Figure 3: The property graph model 10
Figure 4: Cypher Query Example 11
Figure 5: Extract from the schema.graphql 11
Figure 6: GraphQL API – Example for CIC used by the Knowledge Base 12
Figure 7: GraphQL query, which retrieves the title and description for a CIC 12
Figure 8: Response of the GraphQL query 13
Figure 9: Table overview of elements in the Knowledge Base 14
Figure 10: Detail view showing connected elements for the climate-related risk "Heat on buildings" 14
Figure 11: Discovery view for different questions 15
Figure 12: Knowledge Graph based chatbot architecture 17

List of Acronyms

API	Application programming interface
CIC	Climate interaction context
DSS	Decision support service
FE	Front-End
GUI	Graphical User Interface
GDPR	General Data Protection Regulation
KB	(Climate Interaction) Knowledge Base
UI	User interface

Glossary

Angular (software)	TypeScript-based free and open-source single-page web application framework run on Node.js.
Apollo (soft- ware)	A suite of tools working with GraphQL, which includes the Apollo Server and Apollo client libraries. Apollo Server is a server-side library that helps building GraphQL APIs. On the other side Apollo Client is a client-side library that allows consuming those APIs from a front-end application.
Backend service	The backend refers to parts of a computer service that allow it to operate and that cannot be directly accessed by users. Back-end services typically contain the data and operating syntax of the application.
Climate im- pacts	The consequences of realized risks on natural and human systems, where risks result from the interactions of climate-related hazards (including extreme weather and climate events), exposure, and vulnerability. Impacts generally refer to effects on lives; livelihoods; health and well-being; ecosystems and species; economic, social and cultural assets; services (including ecosystem ser- vices); and infrastructure (based on IPCC, 2018)
Edge (graph data- base)	In the context of graph databases, the edges represent relationships between pairs of nodes. Edges are unidirectional and typed, thus each node stands for exactly one type of relationship between two nodes and the nodes can be linked by more than one edge type. Like nodes, edges can also contain various properties.
GRAND- stack	Full-stack framework for building data-centric web applications utilising the Neo4j graph database.
Graph data- base	A database that uses data graph structures for semantic queries with nodes, edges, and proper- ties to represent and store data. A key concept of the system is the graph. The graph relates the data items in the store to a collection of nodes and edges, the edges representing the relationships between the nodes.

GraphQL	GraphQL is an open-source data query and manipulation language for APIs and a query runtime engine. GraphQL enables declarative data fetching where a (software) client can specify exactly what data it needs from an API.
Neo4j	A fully featured, best-in-class open source graph database, that's also available in enterprise edition.
Node (graph database)	In the context of graph databases, the nodes (also called vertices) represent entities, for example concepts, events, places, and things. Nodes can have labels to define (classify) the node types and properties.
React (software)	JavaScript library that's used for building reactive websites. While not a framework, React does have a dedicated framework called Create React App that can be used to build web applications.

Executive summary

Deliverable D1.4 – *Basic Climate Interaction Knowledge Base* consists of the initial version of the Climate Interaction Knowledge Base (Knowledge Base or KB for short), which is being released simultaneously with this accompanying report, under https://knowing-climate.eu/knowledge-base/. The report explains the work performed in Task T1.4 – *Create Climate Interaction Knowledge Base* so far and summarizes the main tasks performed and decisions made during the past months.

The Knowledge Base is an open access database, which can be explored through a dedicated "search and matching" web GUI application (section 5.3), as well as through a powerful API for application developers and integrators (section 5.2).

The data basis of the KNOWING Knowledge Base primarily originates from WP1 – *Identify climate influences and interrelations*, WP2 – *Develop Modelling Framework* and WP3 – *Model CIC-specific Demonstrator Pathways*. Through the Knowledge Base, this data is made available to other work packages and potential users outside the project. Within the project, the main users of the Knowledge Base are the stakeholders and tools developed in WP4 – Establish communication & collaboration channels, most notably the KNOWING Decision Support System.

The central component of the Knowledge Base is the graph database that's interlinking various types of information into a knowledge web and allowing their use in applications. At the current development stage, the knowledge base makes the information that was generated in WP1 and WP2 deliverables D1.1 – *Literature Review Summary* (Jaroszweski et al., 2023a), D1.2 – *CIC-related Measures Inventory* (Jaroszweski et al., 2023b) and D2.1 – *Qualitative Interactive Model* (Zach et al., 2023) available for querying and presentation. This includes the hazards, sectors, risks, adaptation and mitigation responses, second order impacts thereof, and various types of relations between these elements.

As the KNOWING project progresses, additional data and requirements for the KB are expected to arise. Consequently, we concentrated on selecting a robust and flexible technology that will allow adjustments of data models and relationships, as well as introduction of new query possibilities, even at late stages of KB development. Graph databases offer this flexibility, allowing for the data model to be adapted or expanded with relative ease without the need for a complete overhaul of the already existing data schema. Unlike relational databases, where existing data tables need to be updated to accommodate new entities and relations, graph database allow us create new entities are created as new database nodes, as well as to add new relations between such entities ad new database "edges" without the need to redesign the whole database. This feature of graph databases will allow us to extend and refine the KB data model according to availability of the new data sets and new or changing requirements by the end-users and applications using the Knowledge Base in the background.

1 KNOWING summary

Climate change has been globally recognised as an existential threat requiring urgent action to avoid catastrophic consequences. Hence, the EU's Green Deal has been proposed "to make Europe the first climate neutral continent in the world". This includes not only the elimination of net emissions of greenhouse gases by 2050; this is to be achieved while decoupling economic growth from resource use and striving for a fair implementation, leaving no person and no place behind. This ambitious goal is additionally challenged by the need to adapt to unavoidable impacts.

According to the EU's Climate Adaptation Strategy (COM(2021) 82), "improving knowledge and managing uncertainty" is key for realising the vision of a climate neutral and climate-resilient Union, as "Climate change is having such a pervasive impact that our response to it must be systemic". Thus, there is an **urgent need for an integrated approach for enhanced understanding of the interaction, complementarity and trade-offs** between adaptation and mitigation measures, especially regarding the expected increase in regional mean temperature, changing precipitation pattern and soil moisture (IPCC AR6 WG I). Furthermore, this **understanding and knowledge needs to be provided to a broad audience to support local authorities** in EU countries for developing regional programmes.

KNOWING aims to develop a **modelling framework to help understand and quantify the interactions** between impacts and risks of climate change, mitigation pathways and adaptation strategies. The framework will be used to assess the **interrelations between public and private adaptation and mitigation strategies** in order to **identify mitigation pathways along optimised combinations of interventions** in different sectors (e.g. energy, mobility, land use, construction, agriculture). The framework will focus on **three main Climate Impact Contexts** (CICs): (1) Heat waves & health, (2) Soil fertility & agriculture, and (3) Flooding & infrastructure (including river and coastal flooding). It be applied **in four Demonstrator and five Follower Regions by involving authorities, stakeholders and citizens** to develop **enhanced activation and empowerment services, providing target-group-specific awareness, education and decision support tools** to improve the comprehensibility of complex interrelations and support strategic planning of combined adaptation and mitigation measures.

To achieve this goal, KNOWING will produce the following **key exploitable results (KERs)**:

- KER1 an **Impact Interaction Knowledge Base** comprising causal relations of climate and intervention impacts, rebound effects, coping strategies, etc. to inform Climate-ADAPT and IPCC Working Groups I, II & III
- KER2 an **Impact Interaction Model Framework** consisting of a system dynamics model, climate, and sector models for integrated assessment of impacts (direct and indirect) of climate change and countermeasures
- KER3 a Typology of transferable **Climate Mitigation Pathways** including optimised bundles of adaptation and mitigation measures for different typical Climate Impact Contexts (heat waves, soil fertility, flooding)
- KER4 **Climate Activation and Empowerment Services** addressing different target groups (citizens, businesses, authorities) to enhance climate literacy, provide playful trainings and support decision making

These results, developed with the support of an External Expert Advisory Board (EEAB) and a Stakeholder Reference Group (SRG), will **accelerate the transition to a climate-neutral and resilient society and economy** enabled through advanced climate science, mitigation and adaptation pathways and behavioural transformations.

This report is part of D1.4 deliverable of KNOWING WP1 "Identify climate influences and interrelations" and accompanies the initial version of the Climate Interaction Knowledge Base, which is being developed in Task T1.4. The Knowledge Base service will be made publicly available under https://knowing-climate.eu/knowledge-base/, simultaneously with the submission of the report.

2 Objectives of the Deliverable

This document is an accompanying report to the Knowledge Base service v1.0. It details the efforts made to design and develop a first iteration of the Climate Interaction Knowledge Base, which is being developed in task T1.4 to achieve KER1.

This report describes how the Climate Interaction Knowledge Base was designed, considering the various requirements from potential end-users and other KNOWING applications that might use the Knowledge Base as a source for data, and how it was implemented. This document will:

- Explain the purpose of the Knowledge Base (section 3).
- Introduce the data that the Knowledge Base is currently storing and planned to store in the future (section 4).
- Document the key design decisions and elements of the Knowledge Base v1.0, most notably the database design (section 5.1), the API allowing the third party applications to utilise the Knowledge Base (section 5.2), and the web front-end (GUI) application (section 5.3)
- Describe the envisioned features that will be implemented over the course of the coming months and published as a second (and final) version of the KNOWING Knowledge Base. Knowledge Base v2.0 is set to be released in the first quarter of 2026, as deliverable D1.5.

First iteration of the Knowledge Base incorporates the knowledge on **hazards, sectors, risks, adaptation, and mitigation responses, second order impacts thereof, as well as to various types of relations between these elements**. This knowledge can be accessed through an API (section 5.2), and through a dedicated "search and matching" web application that is presented in section 5.3.

Knowledge Base front end application can be reached online under https://knowing-climate.eu/knowledge-base/. In its first version, the primary objectives of this web GUI application are to visualize the data contained in the Knowledge Base, and the relations between the data entries, in a user-friendly frontend web-based application, and to demonstrate the types of "questions" the Knowledge Base can resolve.

One important role of this document will be to serve as a manual to end-users, on how to use the application's user interface, and to software developers explaining how data can be queried from the application's API.

3 Introduction

In accordance with the task definition for task T1.4, the goal of this deliverable is to **provide a first iteration of an open access database containing the knowledge gathered in the KNOWING project**. Based on the task T4.3 description, this "Climate Interaction Knowledge Base" (Knowledge Base for short) shall be accessible through a web interface and serve as a basis for the KNOWING Decision Support System (and possibly other applications). The principal purposes of the Climate Interaction Knowledge Base can thus be summarized as:

1) Make the knowledge generated in the project tangible for different user groups and provide insight on how this knowledge is interconnected through a flexible and user-friendly (web) graphical interface.

2) Make this knowledge accessible and usable to other tools that are developed as part of the KNOWING project and beyond, through an Application Programming Interface (API).

For the basic version of the Knowledge Base the primary data sources are WP1 and WP2 and the contained data is visualized in a separately developed "knowledge base search and matching" web-based application, which will offer users insight into the available data and how this data is interconnected in a knowledge graph.

3.1 Benefits for KB application users

Initial version of the Knowledge base GUI application (section 5.3) will be mainly used by the KNOWING project team, to foster the dialogue with potential Knowledge Base users and application developers.

Final version of the Knowledge Base will provide **tailored presentations of the underlying knowledge for various stakeholders and user groups** (as described in deliverable D4.1), from citizens and climate scientists to decision makers by providing each group of users only with the necessary and relevant information that they desire.

- **Scientific sector** will mostly benefit from having an application that can relate various database elements to the available data and models (domain models, SD models, or causal loop models) and view their respective parameters, meanings and if available also get details about the reliability and confidence of these models. Knowledge Base will transparently disclose its underlying data sources, making it easier for scientists and researchers to seek further information about relevant topics.

- **Public Services, Administration and Policy sector** will mainly use the Knowledge Base to get an overview of the available adaptation and mitigation measures, their interactions and how they affect different sectors. With this knowledge a decision maker can make a more informed choice, when first being presented and informed about other possible options. This can be especially powerful when used in combination with the KNOWING Decision Support System, which will utilise the Climate Interaction Knowledge Base to help the stakeholders understand how Climate Change mitigation and adaptation pathways can be applied in different regions and gain deeper insights into interdependencies between measures.

- **Industrial sector and Economy** stakeholders will mainly use the Knowledge Base indirectly, e.g. through KNOWING DSS, or through sector-specific planning applications utilising the KB to provide the application users with information related to risk, adaptation, mitigation, and relevant climate mitigation and adaptation pathways.

- **Finally, the Civil Society** stakeholders interested in learning about climate change and climate change adaptation will benefit from the Knowledge Base, since it allows them to discover the key elements of the climate adaptations (hazards, affected sectors, etc), view the complex interactions between different climate measures and their impacts on different sectors in an interactive web application, rather than having to read through scientific papers. The Knowledge Base gathers the collected datasets obtained by the literature research completed by other tasks in WP1 and visualizes the knowledge in an appropriate way to the users, also providing in addition descriptions to the displayed elements, giving users the necessary context for a better understanding.

3.2 Benefits for tool/service developers

The Knowledge Base aims to offer the data and relationships gathered and created in the KNOWING project not only to its users through the Knowledge Base frontend, but also to other applications within the KNOWING project and beyond it. This is achieved by means of a powerful API, which allows other tools to easily query and retrieve relevant data. This enables a wide range of different applications to benefit from the available datasets on climate change and climate adaptation and their many, complex interconnections.

One of the significant benefits for other application developers is that they don’t need to learn or work directly with the unique query syntax of the chosen graph database. Instead, the API component acts as an intermediary, translating complex graph database queries into JSON, a well-known standard for file formats. This abstraction allows developers to effortlessly access the data stored in the Knowledge Base. By receiving data in JSON format, which is widely used and easily parsed, developers can integrate this information into their applications more efficiently. This streamlines the integration process, making the Knowledge Base a more attractive data source for other applications, as it simplifies data integration.

In this project, where multiple software tools are being developed, offering the Knowledge Base as a centralized data source can provide significant advantages. By relying on the Knowledge Base instead of each application maintaining its own database, data duplication can be avoided, ensuring that the same information is not redundantly stored in multiple locations. This centralized approach also prevents data inconsistencies, where separate databases might contain conflicting versions of the same dataset. Additionally, using a shared, well-maintained data source like the Knowledge Base can streamline data management and simplify updates and maintenance in the future. It also facilitates better data integrity and consistency across all applications, enhancing the overall reliability and coherence of the project's outputs.

However, it shall be noted that the usage of the Knowledge Base by other tools developed as part of the KNOWING project remains voluntary. It should be decided by each developer team whether to incorporate the provided data from the Knowledge Base and to what extend their applications shall rely on that data. This is done to prevent other tool developments to stall due to the currently active development of the Knowledge Base.

4 Relevant Datasets

4.1 Process for including datasets

The initial version of the Knowledge Base contains data primarily gathered by work efforts performed in WP1. As the basis served both deliverables D1.1 (Literature Review Summary) and D1.2 (CIC-related Measures Inventory). The relevant datasets from these deliverables were compiled in an Excel file and enriched with additional information from D2.1 about the models that will be utilized in that work package. This Excel file is then imported into the graph database utilized by the Knowledge Base after being first pre-processed by several scripts, whose task is to structure and prepare the datasets into a machine-readable form. This ensures for a successful transfer of information into the database. Results from deliverable D2.2 were not yet included in the first iteration of the Knowledge Base, since these results are being finalized simultaneously with the release of the initial version of the Knowledge Base. These results will however be included into the Knowledge Base in the coming months.

For its initial release, the Knowledge Base contains the following entities in its database:

- Climate Impact Contexts (CICs)
- Sectors
- Hazards
- Mitigation responses
- Adaptation responses
- Responses of sectors
- Response opportunities
- Response risks
- Climate-related risks
- Second-order impacts

All these elements relate to one another, as illustrated in Figure 1. This schema highlights all the node types and the edges connecting them, showcasing the complex web of relationships within the database.

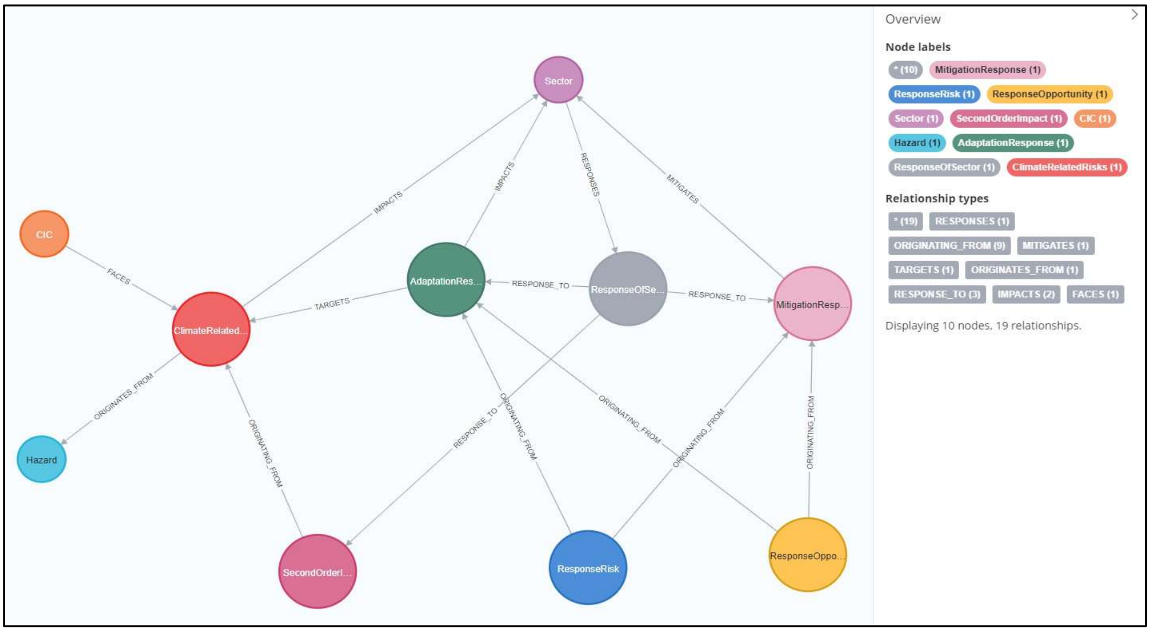


Figure 1 Current data schema for the initial version of the Knowledge base.

4.2 Datasets to be added for next iteration of Knowledge Base

As mentioned in the previous section, additional results of WP2 and WP3 will be integrated into the Knowledge Base in the coming months. The focus will shift from hazards, sectors, impacts and responses towards models, pathways, and transferability of KNOWING results, in anticipation of the KB being used by KNOWING Decision Support Service (DSS), to indicate how the WP3 pathways can be optimally applied to different regions and how selection of alternative adaptations in a pathway affects the regional resilience.

Only after the characteristics of the models and the pathways are defined and understood, we can begin altering and enhancing the existing data schema of the database. Additionally, once we receive initial feedback from potential end-users, we can begin incorporating other datasets from various work packages. These may include information about different regions, geo-referenced data, KPIs for validating mitigation and adaptation responses, and other relevant datasets. This iterative approach ensures that the Knowledge Base evolves to meet user needs and accommodates a wide range of data, thereby enhancing its utility and comprehensiveness. It will also become necessary to discuss with other tool developers in the KNOWING project, what kind of information their tools might require from the Knowledge Base, which might lead to the inclusion of datasets not yet foreseen for the Knowledge Base.

4.3 Discovering datasets of interest

To help find potentially interesting datasets to be integrated into the Knowledge Base, two Excel sheets have been created as part of task T1.4 and circulated among the other work packages. One Excel file is targeted at finding and describing potentially interesting datasets that have not yet been identified by task T1.4. Each project partner can describe datasets of interest by providing the following information:

- Short description of the data in this dataset
- The relevant KNOWING deliverable (for datasets produced by the project)

- A responsible contact person for task T1.4 to reach out to in case of questions
- Justification for why this dataset could be relevant for the Knowledge Base
- Information on where the dataset can be obtained and in what formats it is available

In parallel, a second Excel file was provided to the project partners as well, which focuses on the questions users would like the Knowledge Base to be able to answer. Since the Knowledge Base is also envisioned to be used as a source of information to other tools developed in the KNOWING project, the pool of potential questions can also be extended to those additional tools as well. For describing potential questions, project partners can fill out the following information:

- Contact information (name, partner, or region) for further discussion with T1.4
- The desired question
- Justification for why or for who this question is relevant
- Optionally, the application(s) in which this question should be answered (Knowledge Base, DSS, other)
- Optionally, what data is necessary for answering this question and where can it be found

Both Excel files will remain open for project partners to fill out in the upcoming months, while the development of the Knowledge Base will continue. The goal is to incrementally include new datasets into the database and provide more questions in the FE application for users to select from. Another goal of these efforts is to identify and highlight missing datasets that would be required to answer questions relevant for the regions and the users of the Knowledge Base. Such a lack of required information will be documented and further examined in deliverable D1.5.

5 System Design

Key requirements on the Knowledge Base can be summarised as follows:

- Flexible database that is capable of handling complex data with complex interconnections.
- Possibility to easily add new data types and relations and evolve the existing data models at a late development phase.
- Powerful API that allows the application developers to utilise the full search and matching capability of the database without having to deal with its inner structure.
- Modular “search and matching” web application that can be easily integrated in different web sites and adopted to specific needs of the different user groups.

Clearly, the KB is a data-centric application, and the choice of the database is at the core of the Knowledge Base system design. Initially, the use of **relational databases, vector databases and graph databases** has been considered by the development team.

While relational databases can store relationships, they navigate them with expensive JOIN operations or cross-lookups. So, it turns out that “relational” databases handle relationships poorly, especially if only a part of them is known at the initial design stage. In a graph database, relationships are stored natively alongside the data elements (the nodes) in a much more flexible format than in the relational databases. Everything about the system is optimized for traversing through data paths quickly, even millions of connections per second depending on the available hardware used. On the other hand, the vector databases can perform contextually-aware searches within large collections of poorly structured data, and discover implicit relations in such data sets, whereas the KB data is already highly structured and interconnected.

Consequently, we decided to **build the Knowledge Base on a graph database**, for following reasons:

- Graph solutions are focused on highly connected data with an intrinsic need for relationship analysis
- Graph databases are optimized to retrieve data and they offer the power of graph analytics that can help discover hidden patterns in the data that would be unseen with rational databases.
- Graph databases are best suited for data models that are inconsistent and demand frequent changes like:
    - Additional attributes will be added at some point
    - Not all entities will have the attributes in the table
    - The attribute types are not strictly defined

Neo4j implementation thereof was chosen for following reasons: Neo4j Community edition is a fully featured, best-in-class **graph database** that uses the GPL v3 license. It is actively maintained, and available in the “enterprise edition”, which assures that commercial support and advanced scaling and availability features will be available if required by the sustainability and exploitation plans.

In addition to the database, Neo4j also provides a very powerful **GraphQL API library**, which maps GraphQL neatly onto a graph database. Using a graph database backend for a GraphQL Service has performance advantages as graph databases are optimized for traversing paths in a graph.

Overall architectural design of the KNOWING Knowledge Base is shown in Figure 2. It is based on the **GRANDstack**, a full-stack framework for building applications. The components of the GRANDstack are: **GraphQL, React, Apollo and the Neo4j database**. In this project **Angular** is used instead of React for building the frontend.

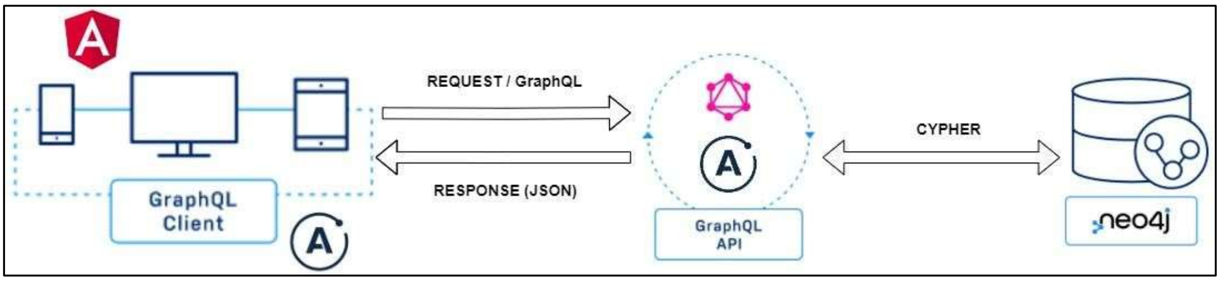


Figure 2: Architectural overview of the Knowledge Base

- **GraphQL** is an **API query language**, modelling application data as a graph. This means that there's one-to-one mapping between the data model presented in the API layer to the database all the way in the backend.
- The **Neo4j GraphQL library** is a Node.js library and works with any JavaScript GraphQL implementation such as GraphQL.js and Apollo Server. It is designed to easily build a GraphQL API backed for a Neo4j database.
- **Apollo Server** is used to serve the GraphQL endpoint. Apollo provides tools, both on the server and the client for working with GraphQL.

5.1 Neo4j Database

As mentioned in the previous section, Neo4j is a graph database. A graph database uses nodes, edges, and properties to represent and store data (see Figure 3). Beyond the core graph, Neo4j also provides ACID transactions, cluster support and runtime failover.

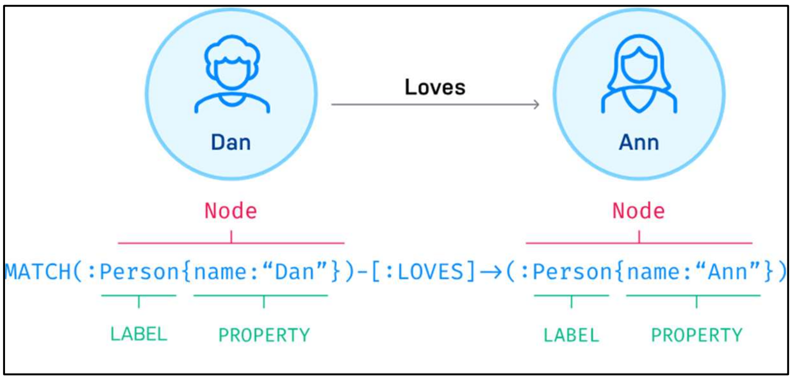


Figure 3: The property graph model

In graph databases, the data objects are represented by **nodes**. Each node can have an unlimited number or type of relationships without sacrificing performance. Relationships between nodes are represented by **edges**. Edges have a start node, an end node, a type, and a direction.

Nodes can be tagged with labels and can hold any number of key-value pairs (properties). Furthermore, their labels may also attach metadata (such as index or constraint information).

Like nodes, edges can also have additional properties. Moreover, the type of relations between the nodes (one to one, one to many, many to many) doesn't need to be explicitly defined. These relationships are the main concept of graph databases, and the database is optimised for fast querying of relationships.

Graph databases mostly use Cypher query language since it is widely adopted, fully specified, and is an open query language optimized for graph databases. With Cypher expressive and efficient queries can be constructed to do any kind of create, read, update or delete operations on the graph. The expressive nature of Cypher and the underlying mechanisms for storing and retrieving data allow us to use Neo4j to build a recommendation engine as the backbone of the Knowledge Base.

For comparison, Cypher is like mapping English language sentence structure to patterns in a graph. Mostly nouns are nodes in the graph, while verbs act as the edges in the graph representing relationships, and adjectives and adverbs are the properties of those nodes and edges (see Figure 4).

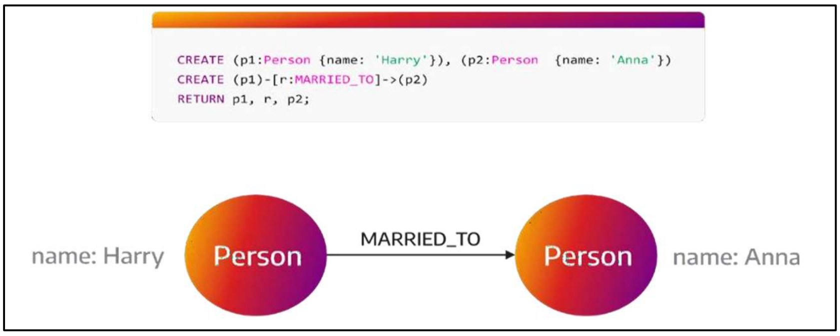


Figure 4: Cypher Query Example

5.2 Apollo with GraphQL

Apollo is a suite of tools working with GraphQL, which includes the Apollo Server and Apollo Client libraries. Apollo Server is a server-side library that helps building GraphQL APIs. On the other side Apollo Client is a client-side library that allows consuming those APIs from a front-end application. The Apollo Client instance is pointing to the GraphQL API endpoint to consume the services provided by that API.

GraphQL is an open-source API query language and a server runtime for building APIs. One of GraphQL's most powerful capabilities enables the API designer to express the entire data domain as a graph using nodes and relationships. So, API clients can traverse the graph to find the relevant data, as illustrated in figures 5 to 8.

```
Climate Impact Contexts
"""
type CIC {
    "Internal Unique Code"
    code: String!
    """Name of the CIC"""
    title: String!
    """Detailed Description"""
    description: String!
    """Which climate-related risks faces this CIC"""
    facesClimateRelatedRisks: [ClimateRelatedRisks!]! @relationship(type: "FACES", direction: OUT)
}
```

Figure 5: Extract from the schema.graphql

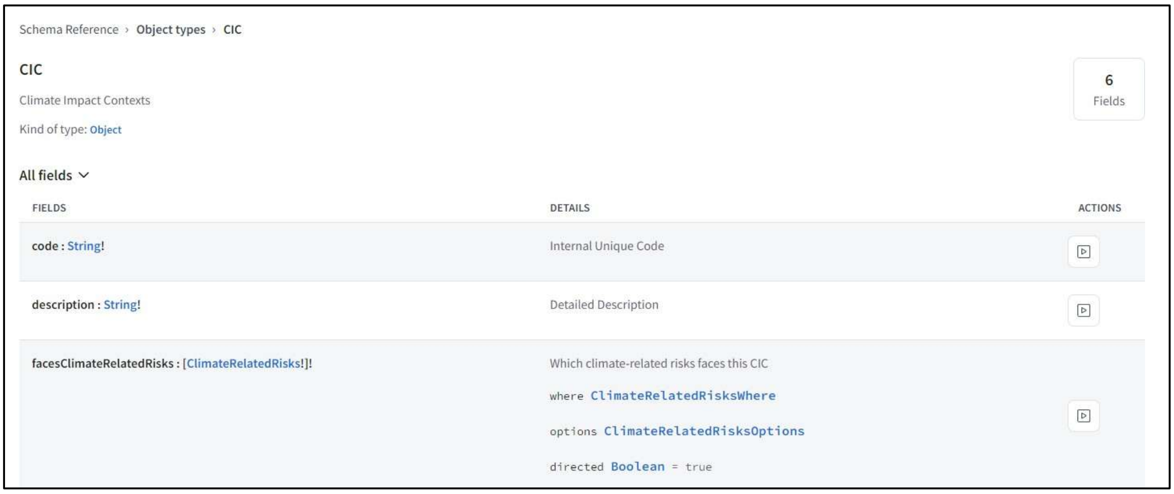


Figure 6: GraphQL API – Example for CIC used by the Knowledge Base

In GraphQL the schema (see Figure 5) becomes the specification for the API and is used to mock an API (see Figure 6). From type definitions the query types are defined, which are entry points for the API. The schema is defined with the Schema Definition Language. A GraphQL query specifies a called selection set, which can for instance contain only some of the overall available attributes of the requested element. See Figure 7 for reference, which return only the title and description attributes for all available CICs.

```
query CIC {
  cics {
    title
    description
  }
}
```

Figure 7: GraphQL query, which retrieves the title and description for a CIC

The response matches the fields in the selection set and the query (see Figure 8). Only the data of the fields requested is sent back as a response to the caller. This reduces the network traffic and the need for the frontend application to iterate over all defined attributes of an element, since only the requested information is sent over the internet and received by the application.

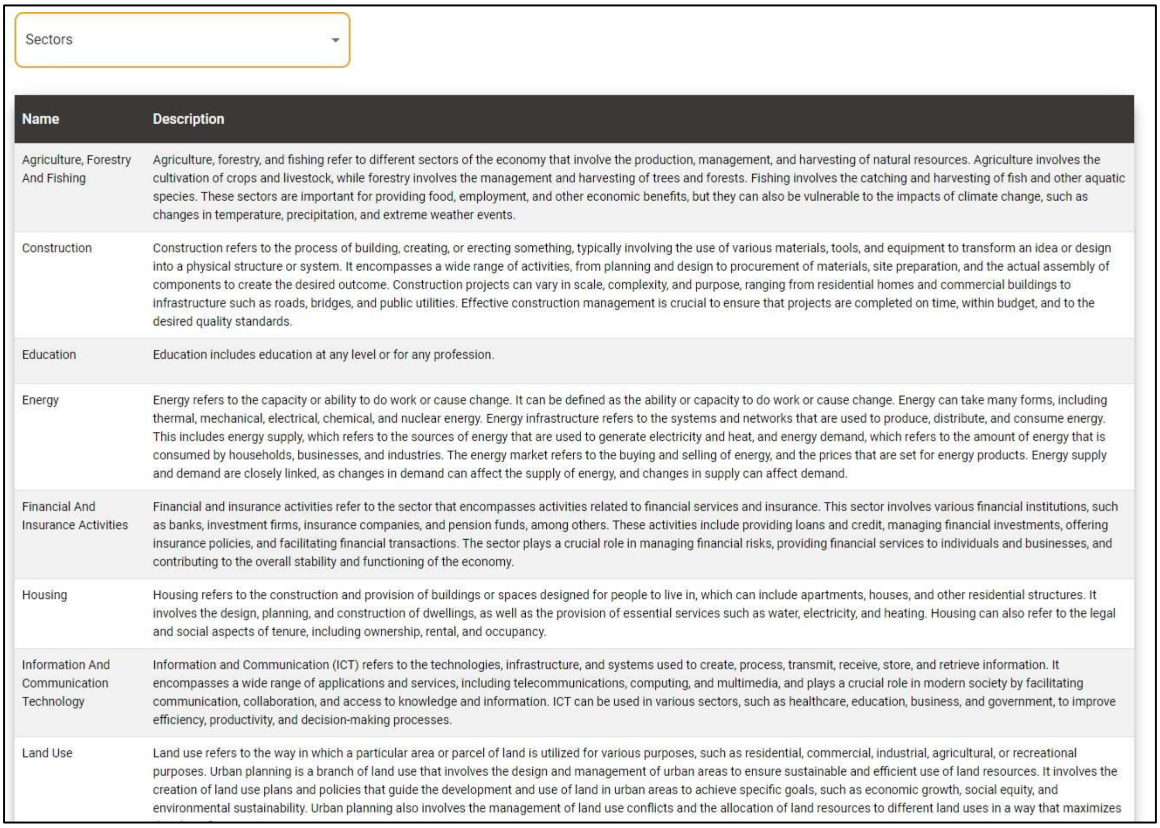
Figure 8: Response of the GraphQL query

5.3 Frontend applications

The KB frontend application for the initial version of the Knowledge Base is realized with Angular, a framework that allows for the creation of flexible and dynamic user interfaces. Angular was chosen because it offers a variety of libraries and packages for third-party software, one such package being the Apollo-Client enabling a fast and efficient data exchange with the Knowledge Base backend components.

The frontend's primary purpose is to visualize the contents of the database that resides in the backend of the Knowledge Base. The data can be presented to the user in three different modes:

On the one hand, users can **view and filter the database's content in the form of a table** as can be seen in Figure 9, allowing the users to quickly get an overview of the contained elements and their key attributes. On top of the table the users can **select the topic** that they are interested in, which will result in the table to be updated, displaying only the chosen elements. In addition, it is possible to use the **full-text search** box next to the selected items to further limit the search results.


Sectors

Name	Description
Agriculture, Forestry And Fishing	Agriculture, forestry, and fishing refer to different sectors of the economy that involve the production, management, and harvesting of natural resources. Agriculture involves the cultivation of crops and livestock, while forestry involves the management and harvesting of trees and forests. Fishing involves the catching and harvesting of fish and other aquatic species. These sectors are important for providing food, employment, and other economic benefits, but they can also be vulnerable to the impacts of climate change, such as changes in temperature, precipitation, and extreme weather events.
Construction	Construction refers to the process of building, creating, or erecting something, typically involving the use of various materials, tools, and equipment to transform an idea or design into a physical structure or system. It encompasses a wide range of activities, from planning and design to procurement of materials, site preparation, and the actual assembly of components to create the desired outcome. Construction projects can vary in scale, complexity, and purpose, ranging from residential homes and commercial buildings to infrastructure such as roads, bridges, and public utilities. Effective construction management is crucial to ensure that projects are completed on time, within budget, and to the desired quality standards.
Education	Education includes education at any level or for any profession.
Energy	Energy refers to the capacity or ability to do work or cause change. It can be defined as the ability or capacity to do work or cause change. Energy can take many forms, including thermal, mechanical, electrical, chemical, and nuclear energy. Energy infrastructure refers to the systems and networks that are used to produce, distribute, and consume energy. This includes energy supply, which refers to the sources of energy that are used to generate electricity and heat, and energy demand, which refers to the amount of energy that is consumed by households, businesses, and industries. The energy market refers to the buying and selling of energy, and the prices that are set for energy products. Energy supply and demand are closely linked, as changes in demand can affect the supply of energy, and changes in supply can affect demand.
Financial And Insurance Activities	Financial and insurance activities refer to the sector that encompasses activities related to financial services and insurance. This sector involves various financial institutions, such as banks, investment firms, insurance companies, and pension funds, among others. These activities include providing loans and credit, managing financial investments, offering insurance policies, and facilitating financial transactions. The sector plays a crucial role in managing financial risks, providing financial services to individuals and businesses, and contributing to the overall stability and functioning of the economy.
Housing	Housing refers to the construction and provision of buildings or spaces designed for people to live in, which can include apartments, houses, and other residential structures. It involves the design, planning, and construction of dwellings, as well as the provision of essential services such as water, electricity, and heating. Housing can also refer to the legal and social aspects of tenure, including ownership, rental, and occupancy.
Information And Communication Technology	Information and Communication (ICT) refers to the technologies, infrastructure, and systems used to create, process, transmit, receive, store, and retrieve information. It encompasses a wide range of applications and services, including telecommunications, computing, and multimedia, and plays a crucial role in modern society by facilitating communication, collaboration, and access to knowledge and information. ICT can be used in various sectors, such as healthcare, education, business, and government, to improve efficiency, productivity, and decision-making processes.
Land Use	Land use refers to the way in which a particular area or parcel of land is utilized for various purposes, such as residential, commercial, industrial, agricultural, or recreational purposes. Urban planning is a branch of land use that involves the design and management of urban areas to ensure sustainable and efficient use of land resources. It involves the creation of land use plans and policies that guide the development and use of land in urban areas to achieve specific goals, such as economic growth, social equity, and environmental sustainability. Urban planning also involves the management of land use conflicts and the allocation of land resources to different land uses in a way that maximizes

Figure 9: Table overview of elements in the Knowledge Base

Each element in the table can also be expanded and viewed in a dedicated **element detail view** (Figure 10), which lists all the available attributes of the given element and shows all other directly connected elements for the selected item. These connected elements form the “neighbourhood” of the previously selected element and can again all be individually clicked upon, which will open the detail views for those elements. This allows the users to **explore the knowledge base organically**, by following the chain of “neighbour” links.

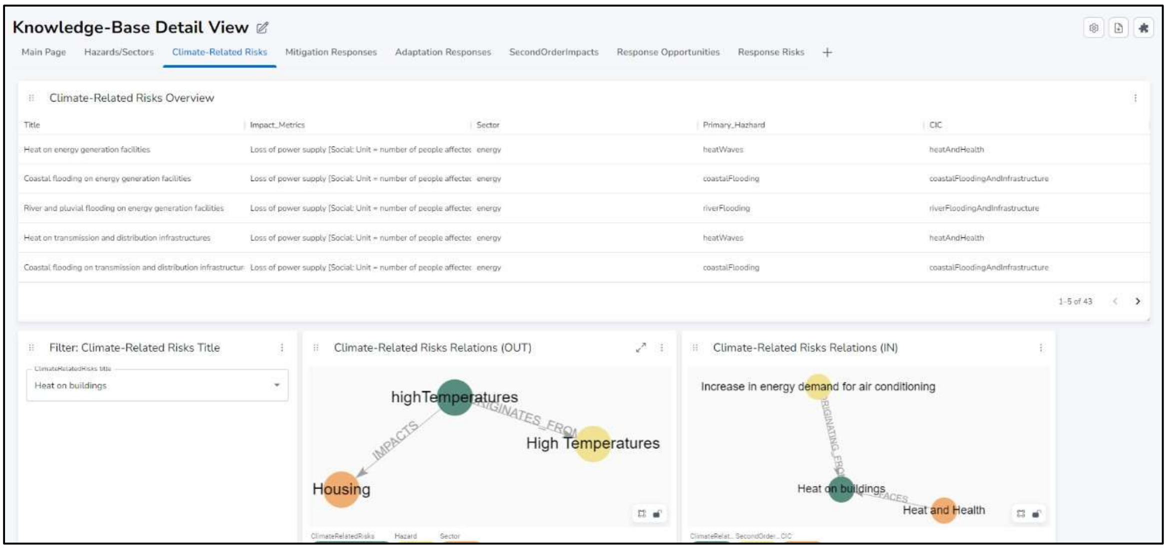
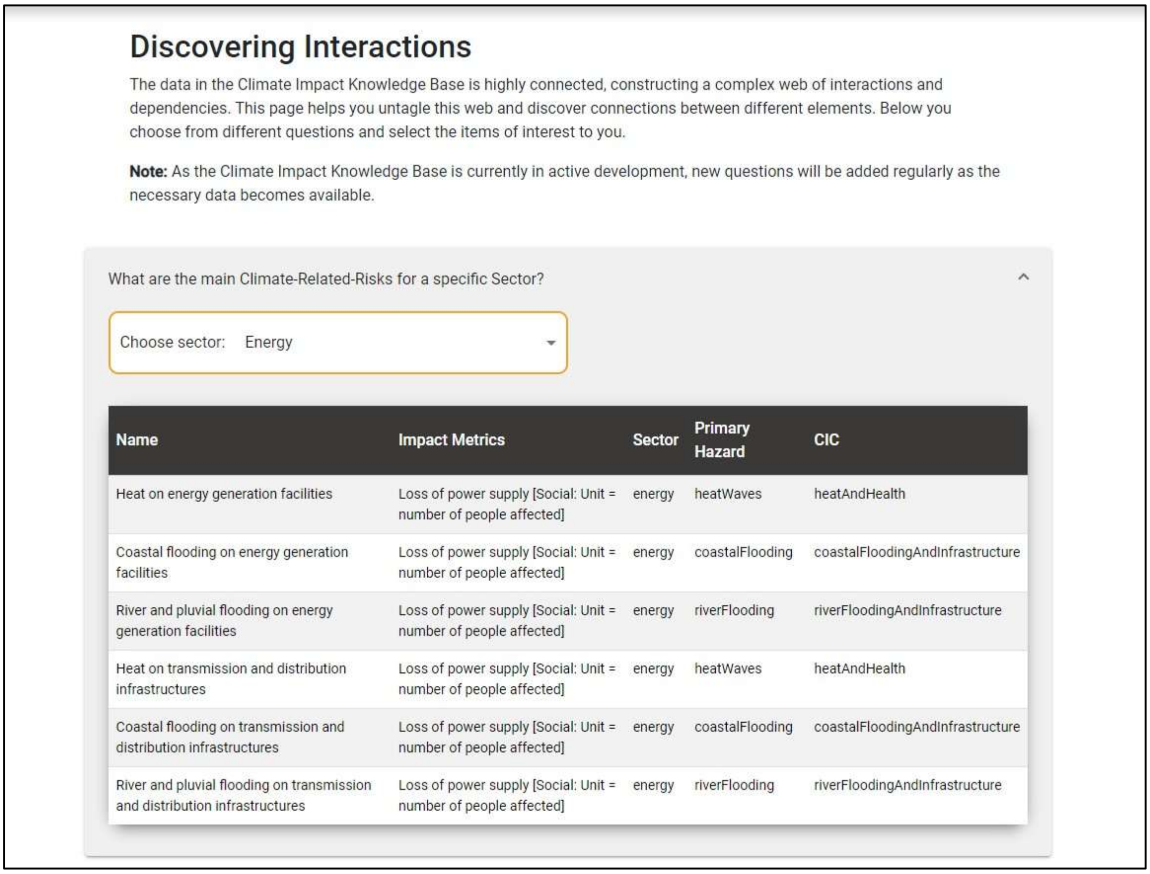


Figure 10: Detail view showing connected elements for the climate-related risk “Heat on buildings”

An alternative approach for exploring the Knowledge Base is by **browsing through a list of pre-defined questions (**Figure 11**).** When a specific question is selected, the UI displays the elements related to the specified question in a tabular form. In this way the complex connections between different elements will be placed in the centre of attention and users will be aided in understanding the interactions between, for example, measures and their impact on different sectors. Pre-defined questions demonstrate how the Knowledge Base can be embedded into applications and will be used to facilitate discussions with potential end-users and application developers. The number of available questions will be incrementally increased based on user suggestions and on the features of new datasets that will be added to the Knowledge Base in upcoming months (see section 4.3).

This view also allows the user to expand the chosen elements in the detail view (Figure 10).



Discovering Interactions

The data in the Climate Impact Knowledge Base is highly connected, constructing a complex web of interactions and dependencies. This page helps you untagle this web and discover connections between different elements. Below you choose from different questions and select the items of interest to you.

Note: As the Climate Impact Knowledge Base is currently in active development, new questions will be added regularly as the necessary data becomes available.

What are the main Climate-Related-Risks for a specific Sector?

Choose sector: Energy

Name	Impact Metrics	Sector	Primary Hazard	CIC
Heat on energy generation facilities	Loss of power supply [Social: Unit = number of people affected]	energy	heatWaves	heatAndHealth
Coastal flooding on energy generation facilities	Loss of power supply [Social: Unit = number of people affected]	energy	coastalFlooding	coastalFloodingAndInfrastructure
River and pluvial flooding on energy generation facilities	Loss of power supply [Social: Unit = number of people affected]	energy	riverFlooding	riverFloodingAndInfrastructure
Heat on transmission and distribution infrastructures	Loss of power supply [Social: Unit = number of people affected]	energy	heatWaves	heatAndHealth
Coastal flooding on transmission and distribution infrastructures	Loss of power supply [Social: Unit = number of people affected]	energy	coastalFlooding	coastalFloodingAndInfrastructure
River and pluvial flooding on transmission and distribution infrastructures	Loss of power supply [Social: Unit = number of people affected]	energy	riverFlooding	riverFloodingAndInfrastructure

Figure 11: Discovery view for different questions

6 Final Knowledge Base planning

Knowledge Base development will be actively continued by the work efforts put into task T1.4 and culminate in the release of the final version of the Knowledge Base in deliverable D1.5.

In general, we intend to continuously integrate more results (data, relations) from other KNOWING WPs, improve the overall user experience for the end-users and further refine the visual appearance of the KB web GUI application, so that it fully adheres to the general theming guides provided for the KNOWING project. Further, we plan to continue collecting from potential end-users their questions that they would like the Knowledge Base, DSS and potentially other applications to answer by utilising the Knowledge Base as a backend service.

6.1 Incorporating new data and relations

As indicated in section 4, the initial KB implementation mainly reflects the results of WP1 and WP2 that were available in time to be included in the KB, most notably the data and relations that are pertinent to deliverables D1.2 and D2.1. Additional information resulting from the work of WP2 and WP3 will be added as it becomes available, such as:

- Climate Mitigation and Adaptation Pathways: region-specific timelines of adaptation and mitigation measures
- Region-specific indicators (e.g. climate, sectors, socioeconomic): set of indicators – an initial list of indicators was already provided in deliverable D3.2 - allowing us to establish a similarity metrics between an arbitrary European region and the KNOWING case study regions.
- Semi-quantitative indicators related to price and complexity of different adaptation and mitigation methods, their effects, and side-effects.
- Data, models, and modelling parameters that were used to calculate actual impacts of hazards, adaptations, and mitigation in the project.

Concrete choice of the data and relations to be included in the final KNOWING Knowledge Base will depend both on the data availability, user's requirements, and the needs of the application developers.

6.2 Incorporating the new user requirements

The future tool development, improvement of current features and implementation of new features will be also driven by repeated consultations with different stakeholders and potential end-users. Such meetings will be undertaken as part of the upcoming meetings planned in tasks T4.1 and T4.3, where the tool developers will come together with representatives of the different regions to discuss the plans and requirements for the tools developed by WP4.

The initial version of the Climate Impact Knowledge Base will serve as the basis for further discussions with users, allowing them to test out the current capabilities and give feedback on possible improvements and missing features. We envision this to be an iterative process of multiple rounds of discussions and presentations of the Knowledge Base, coupled with an agile development process to react to the given feedback and steer further development in the desired directions.

One of the key requirements that needs to be incorporated in the final KB web GUI application is providing different application profiles for different user classes. This may include the use of different language for describing the KB elements and relations, different choice of information to be shown and different visualisation of the data and relationships, depending on the needs of different user types (e.g. decision makers, stakeholders, scientists, citizens).

Since KB is designed to be used in read-only mode, we do not anticipate the need for a fully-fledged user management, where certain users would have the rights to alter existing or add new data also to avoid the entry of potentially conflicting or misleading information. As a positive side effect, this also means that the KB service will not need to manage any GDPR-relevant user data.

6.3 Application requirements

As indicated in sections 3.2 and 5.2, KB backend is designed as a flexible knowledge source that can be utilised by other KNOWING applications and by third party applications beyond KNOWING. Most notably, the KNOWING DSS will make extensive use of the KB as a source of information related to similarity of regions, and transferability of the pathways, adaptation, and mitigation measures.

Key technical requirement for using the KB by third party applications is some form of service authentication and authorisation, to avoid accidental disclosure of sensitive information (if any is incorporated in the KB in the future) and minimise the risk of denial-of-service attacks.

6.4 MAIA connection

In parallel with the KNOWING project, the MAIA coordination and support action project (Horizon Europe - Grant Agreement 101056935) is developing a set of tools simplifying the use of generative AI models in the context of climate adaptation and mitigation.

KNOWING KB can profit from these services, mainly by using them to **generate property descriptions for inclusion in the KB database**. Other possible use cases include:

- KNOWING-aware AI Chatbot providing an alternative access path to the knowledge stored in the KB service (Figure 12)
- **Automated linking of relevant (new) articles to various KB elements**, e.g., to relate the KB outputs with new developments after the project end.
- **Semi-automated discovery of new elements for the KB**

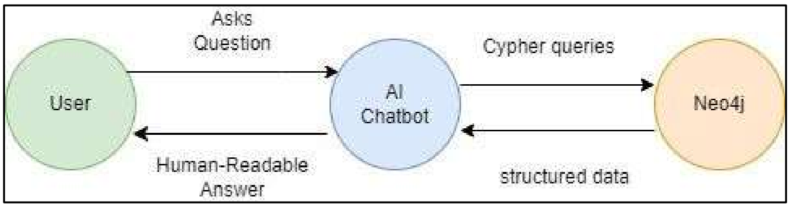


Figure 12: Knowledge Graph based chatbot architecture

In cooperation with the MAIA team, we intend to test the usability of the MAIA generative AI services in the KB context and report the findings in D1.5 deliverable. It is, however, important to understand that MAIA AI services are a work in progress, and that the generative AI applications currently aren't fit to be directly used by intended end-users of the KB (Havlik & Pias, 2024). They are good at producing text and even multimedia materials based on user requests, but their responses are often incomplete or even completely wrong. This is an inherent limitation of the generative AI technology that currently cannot be completely suppressed by any technological means. The main use case for utilisation of generative AI in KNOWING is therefore very conservative:

1) Make the KNOWING knowledge available to the AI
2) Ask the AI to produce textual descriptions of various KB elements (e.g. hazards, sectors, impacts, adaptation and mitigation options, and relations between them.) in a style that is appropriate for different KB user types.

3) Let the KNOWING experts inspect and if necessary, improve the descriptions before integrating them in the Knowledge Base.

Our working hypothesis is that this will allow us to generate the necessary texts with less efforts than what would be needed for the experts to write them without AI help.

Other use cases will be discussed and tested to some extent. This may reflect in our exploitation and sustainability planning, but we do not currently intend to integrate the results of these tests in the KB service, nor to make a “KNOWING chatbot” available to the public.

7 Conclusions

This deliverable accompanies the first iteration of an open access “Climate Interaction Knowledge Base” containing the knowledge gathered in the KNOWING project. It explains the reasons for KB development, expectations on the KB, and introduces the KB system design.

The deliverable also anticipates further KB development, both as a standalone web GUI application and as a service to be used by other applications, most notably by the KNOWING DSS. Some of the forward looking statements in chapter 6 indicate our current understanding of the KNOWING DSS specifications and may need to be altered to accommodate the final DSS concept, which is still in development and will be published as deliverable D4.6 *KNOWING Decision Support Service (DSS) Concept*, which is scheduled for September 2024.

Final results of the Knowledge Base development will be published as deliverable D1.5 *Final Climate Interaction Knowledge Base*, which is scheduled for March 2026. The Climate Impact Knowledge Base is one of the key exploitable results of the project, and realising its full potential will only be possible in teaming up with other actors from the climate research cluster.

References

Havlik, D., & Pias, M. (2024). Common errors in Generative AI systems used for knowledge extraction in the climate action domain. *arXiv preprint arXiv:2402.00830*.

IPCC. (2018). Annex I: Glossary. In IPCC (Ed.), *Global Warming of 1.5°C* (pp. 541–562). Cambridge University Press. https://doi.org/10.1017/9781009157940.008

Jaroszweski, D., 2023a. Literature Review Summary. Deliverable D1.1 of the Horizon Europe project KNOWING.

Jaroszweski, D., 2023b. CIC-related Measures Inventory. Deliverable D1.2 of the Horizon Europe project KNOWING.

Zach, M. et al (2023). Qualitative Interactive Model. Deliverable D2.1 of the Horizon Europe project KNOWING (Accompanying Report).

Annex A: Data Summary

Data accessibility:

This deliverable mainly consolidates the data outputs of D1.2 and D2.1 deliverables and republishes the results in the form of a graph database. All of the input data, as well as the Knowledge Base database graph are considered Open Data and will be published either through Zenodo or as the part of the deliverables (or both).

However, neither the D1.2/D2.1 deliverables, nor the underlying data have been made public by consortium at a time this deliverable was written. Since the KB database graph cannot be published as Open Data before the underlying input data is officially published, its publication on Zenodo will be delayed until this requirement is met.

Data interoperability:

Input data used in preparing this deliverable is available in the form of excel sheets. The KB knowledge graph, which is a main data output of this deliverable is available as Neo4j database dump.

Data reuse:

KB knowledge graph is primarily useful as a starting point for making a local copy of the Knowledge Base or extending it with additional data.

Compared to input data from D2.1 and D2.2 deliverables, the main value added by D1.4 consists of additional quality checks to insure the internal consistency of the data (nodes) and cross-references between data objects (edges).

Knowledge base data dump can be easily imported in Neo4j databases.

Security and Ethics:

The work performed for this Deliverable and the data used and produced is not considered sensitive in terms of ethics or security.

Table 1: Data used in preparation of KNOWING Deliverable D1.4

Dataset name	Format	Size	Owner & re-use conditions	Potential utility within & outside of KNOWING	Unique ID[1]
KNOWING D1.2	PDF	~1MB	Open Data	D1.2 CIC-related measures Inventory is a main source of data for the first version of the Knowledge base	CORDIS
KNOWING D2.1	PDF	~4MB	Open Data	D2.1 builds upon the D1.2 and expands upon the relations between hazards, sectors, risks, and positive and negative effects of various types of responses.	CORDIS

Input data that was used in preparing this deliverable is integral part of deliverables D1.2 and D2.1 and will be made available on CORDIS, under https://cordis.europa.eu/project/id/101056841/results after the mid-term project review. Qualitative SD models resulting from D2.1 are also available on Zenodo, under DOI 10.5281/zenodo.10210763 and 10.5281/zenodo.10160177.

[1] Unique ID in the tables is typically an URI or DOI, e.g. a Zenodo DOI for KNOWING output.

Table 2: Data produced in preparation of KNOWING Deliverable D1.4

Dataset name	Format	Size	Owner & re-use conditions	Potential utility outside of KNOWING	Unique ID
Knowledge Base	tbd.	?	tbd.	This data is mainly useful as a starting point for extending the KB, or as a way to implement a local version of the Knowledge Base.	N/A

KB knowledge graph has not been made public at the time of preparing this deliverable, for two main reasons. First, the underlying data sets from D1.2 and D2.1 haven't been made public yet, and second, no decision has been made how to publish this data and under which license. In principle, the KB data set is almost equivalent to union of the D1.2 and D2.1 data sets at the time of writing this deliverable, but this will change in the future as more data is added to the KB. The data itself will either be published as a set of CSV files that are used to populate the Knowledge Base, or as a Neo4J database dump file, which can be directly imported into a neo4J database.

Annex B: Ethics Considerations

No relevant ethical areas for KNOWING were touched in preparation of this deliverable.