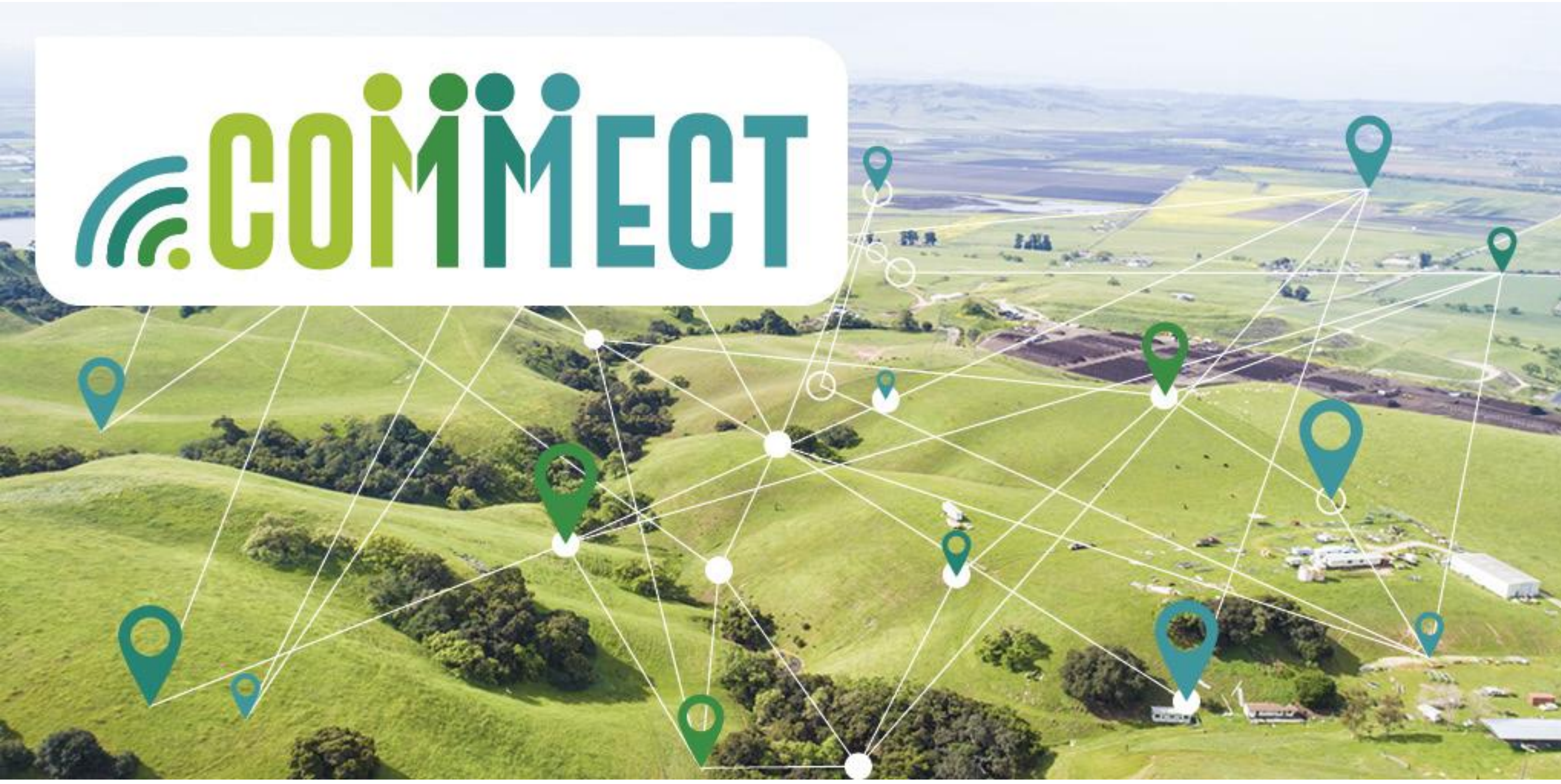
Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solutions



Deliverable 3.3
Decision-making Support Tool

February 2024

PUBLIC

This project has received funding from the European Union's Horizon Europe Research and Innovation
Programme under grant agreement no. 101060881.

COMMECT

Bridging the digital divide and addressing the need of Rural Communities with Cost-effective and Environmental-Friendly Connectivity Solutions

Grant agreement No. 101060881

Decision-Making Support Tool version 1

WP3 Impact Assessment Framework and Business Models

Deliverable ID	**D3.3**
Deliverable Name	**Decision-Making Support Tool version 1**
Lead Partner	DNET
Contributors	VITECH

PUBLIC

COMMECT Project Abstract

In recent years, the importance and need for broadband and high-speed connectivity has been steadily increasing, and even accelerated by the Covid-19 pandemic, driving the process towards a more connected society. However, acceleration has largely taken place in urban communities. In Europe, 13% of people are still without broadband access, and this mainly affects the most rural and remote areas (European Commission, Directorate-General of Communications Networks, Content & Technology, 2022). Those areas are the most challenging to address since they are the least commercially attractive. COMMECT aims to **bridge the digital divide**, by providing quality, reliable, and secure access for all in rural and remote areas. The **goal of extending broadband connectivity in rural and remote areas** will be achieved by *integrating Non-Terrestrial Networks with terrestrial cellular XG networks, and low-cost Internet of Things (IoT), complemented by Artificial Intelligence, Edge, and Network Automation.*

A participatory approach with the end-users and ICT experts working together on development challenges will be the key **for the digitalization of the sector**. To ensure the rich exchange of best-practice and technical knowledge among the actors of the agri-forest value chain, COMMECT has set up **five Living Labs across and outside Europe**, *where end-user "pains" and (connectivity) "gains" will be discussed in detail from different perspectives.*

COMMECT aims to contribute to a balanced territorial development of the EU's rural areas and their communities by making smart agriculture and forest services accessible to all. COMMECT will facilitate that by developing a **decision-making support tool** able to advise on the optimal connectivity solution, according to technical, socio-economic, and environmental considerations. This tool, incorporating collaborative business models, will be a *key enabler for jobs, business, and investment in rural areas, as well as for improving the quality of life in areas such as healthcare, education, e-government, among others.*

Document Information

Document Name	Decision-Making Support Tool version 1
Version No.	0.3
Due Date Annex I	29.02.2024
Report Date	28.02.2024
Number of pages	35
Lead Authors	S. Krco (DNET),
Other Authors	D. Nedic (DNET), D. Pradas (VITECH)
Dissemination Level	Public

Document History

Version	Date	Description	Authors	Reviewed by
0.0	06.01.2024	Document Creation Draft Table of Contents	S. Krco (DNET)	D. Pradas (VITECH) M.R. Palattella (LIST)
0.1	05.02.2024	First integrated draft	S. Krco (DNET), D. Nedic (DNET), D. Pradas (VITECH)	M. Lopez Lechuga (AAU) K. McDonnell (HWEI) R. Gilsing (TNO)
0.2	16.02.2024.	Updated after internal review	S. Krco (DNET), D. Nedic (DNET), D. Pradas (VITECH)	M.R. Palattella (LIST)
0.3	28.02.2024	Final version	S. Krco (DNET), N. Mišić (DNET), D. Pradas (VITECH)	M.R. Palattella (LIST)
0.4	10.07.2024	Revision after RP1. Updated Executive Summary, and examples in the Annex.	S. Krco (DNET), D. Pradas (VITECH)	M.R. Palattella (LIST)

Document Approval

Version	Name	Role in the Project	Beneficiary	Date
0.3	M. R. Palattella	Project Coordinator	LIST	28.02.2024
0.4	M. R. Palattella	Project Coordinator	LIST	19.07.2024

Executive Summary

The COMMECT project aims at empowering rural communities by using digital services supported by sustainable communication solutions. A variety of branches of rural area production and nature preservation are considered in the project, such as wine growing, olive growing, arable crop growing, forestry, preservation of natural environment and livestock transport. This report describes the design and implementation of a Decision-making Support Tool (DST) that will support users (such as farmers, forest operators, municipalities, etc.) to select and implement digital technologies in their farms and in the region, in general.

The tool is leveraging Generative AI/LLM (Artificial Intelligence/ Large Language Model) as the underlying technology, which over the last two years has witnessed an extremely rapid development, creating the foundation for the development of very flexible tools. Accordingly, this technology is used in favour of traditional approaches based on wizards and multiple-choice options for the generation of decision-making support. In addition, instead of using public knowledge to provide answers, the COMMECT DST relies on information and knowledge provided by the project (deliverables) and sources selected and curated by the project partners to generate responses. An enterprise-grade environment is used for development and provision of the service to ensure high reliability and security. Web and mobile interfaces are provided for interaction with the tool, including access via the WhatsApp messaging application.

This document provides an overview of the DST architecture design together with a description of the first release (v0.2) of the DST software (SW) tool. The rationale for selecting Microsoft Azure and OpenAI LLM as the development and production environment is provided. The initial validation process is described, together with the overall validation process, planned for the next phase of the project.

The next releases of the DST will focus on building up the private knowledge base (with data collected from LLs, end users and COMMECT partners) to incorporate the insights generated through the COMMECT efforts. In addition, it will focus on validating the accuracy of the responses, together with the work on prompt engineering and creating a conversational tool capable of driving discussion with users to collect all relevant information for generating customized support responses, based on specific needs, and specific context – all in all to support users in rural areas in terms of decision making on (the realization of) connectivity solutions.

The goal of the initial DST version was to create the technical foundation for implementing COMMECT DST, i.e., to evaluate different approaches in the rapidly evolving domain, including different LLMs, and to make the initial functionality to the consortium for better understanding the taken approach and to initiate collection and curation of data to create the COMMECT knowledge base.

In other words, the goal of implementing the initial DST version was to create a working tool with all required technological blocks, not on the quality of answers. The examples of the interaction with DST given in the document are provided to show that the tool works from the technological perspective, not to showcase the quality of the responses. This is in line with the project implementation plans.

How close to replicating the service provided by agriculture consultants we will be able to get, will depend on the quality of the information in the COMMECT knowledge base as well as on the capabilities of the underlying large language models and other tools used to implement the DST.

The current version of the DST is made available at: https://decision-support-tool.dunavnet.eu/

Table of Contents

List of Figures

List of Tables

Glossary of Terms

Term	Description
AGRI	Agriculture, Forestry, Rural Areas Domain
AGRI-KB	Agriculture, Forestry, Rural Areas Domain - Knowledge Base
AI	Artificial Intelligence
AIDR	Artificial Intelligence for Disaster Response
API	Application Programming Interface
DB	Data Base
DL	Deep Learning
DST	Decision-making Support Tool
GAI	Generative Artificial Intelligence
GDELT	Global Database of Events, Language, and Tone
GPT	Generative Pre-trained Transformer
HTTP	Hypertext Transfer Protocol
ICP	Intelligent Connectivity Platform
ICT	Information and Communication Technology
KPI	Key Performance Indicator
LL	Living Lab
LLM	Large Language Model
ML	Machine Learning
NLG	Natural Language Generation
NLP	Natural Language Processing
RAG	Retrieval Augmented Generation
SDK	Software Development Kit

1. Introduction

The COMMECT project aims at bridging the existing digital divide experienced by rural communities. Following a participatory approach, the project advocates for a close collaboration between the Information and Communication Technology (ICT) experts and the end-users (as illustrated in Figure 1). This will provide all the technical experts involved in the project with a better understanding of the end-users' businesses, their needs, and the benefits they expect from digitalization.

Five Living Labs (LLs) have been setup by the project across and outside Europe to carry out this initiative. The proposed connectivity solutions integrated in current or future ICT solutions will not only be evaluated from a technical perspective. In fact, the project will also investigate how those solutions can boost socio-economic growth of rural communities. Additionally, the environmental impact of the connectivity solutions will also be studied, providing the end-users with insights of how digitalization can support them in becoming more resilient and sustainable.

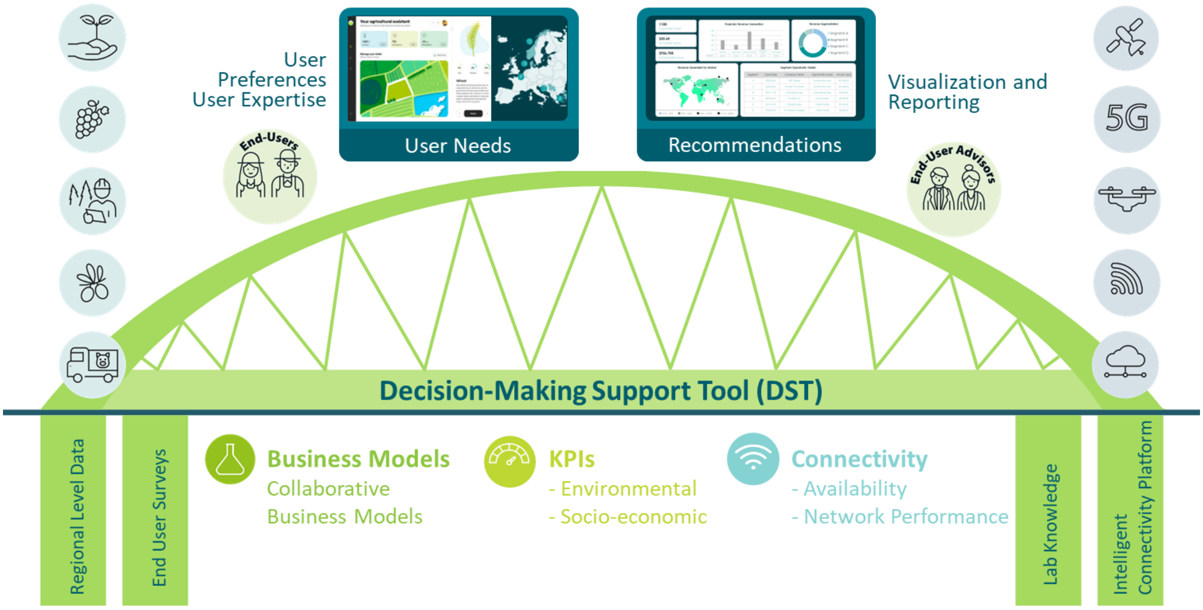


Figure 1. Overview of the COMMECT methodology.

The objective of the deliverable D3.3 - Decision-Making Support Tool version 1, output of WP3 - Impact Assessment Framework and Business Models, is to describe the initial implementation of the COMMECT DST made available to the COMMECT consortium (v0.2). The feedback collected from the project partners, based on the activities carried in each LL, will be used as an input into the implementation of the upcoming releases of the tool. The upcoming releases will be published quarterly to enable continuous testing and feedback collection.

At the time of writing the COMMECT proposal, Generative Artificial Intelligence (GAI) and Large Language Model (LLM)-based solutions did not exist or were in their infancy and were not considered for implementation of the tool. However, over the last two years we witnessed extremely rapid development of these technologies, which created foundations for development of very flexible tools with potential to scale beyond the project lifetime and create new exploitation opportunities. With that in mind, the COMMECT consortium based the implementation of the COMMECT DST on these novel technologies instead of using a more traditional approach based on wizards and series of questions with multiple choice answers.

As detailed in deliverable *D1.3 - COMMECT solution architecture version 1* [1], Figure 1 shows the DST as COMMECT's main mechanism to "bridge" the digital divide between COMMECT's connectivity solutions and the rural communities needs and use cases where choice is limited. DST can be seen as an advanced assistant, based on AI/LLM and exploiting powerful decision models, specifically trained in the domain of rural connectivity. This AI understands and generates text relevant to agriculture, forestry, rural areas, and connectivity challenges. It is therefore well-versed in the specific jargon, data, and scenarios pertinent to these fields, enabling it to provide more accurate, context-aware suggestions, analyses, and solutions tailored to the unique needs of rural connectivity projects. The domain-specific jargon or knowledge is stored in the AGRI-KB (a knowledge base of agri-forestry domain) which is then itself the basis of a LLM.The main pillar to build the bridge in the user side are the user needs and regional data that can be exploited. The COMMECT knowledge data base and the Intelligent Connectivity Platform (ICP) are the main pillar of the bridge in the advisors' side. The DST will allow to identify and propose relevant business models, suitable connectivity solutions, technical solutions, market/sector information, etc. while considering environmental and socio-economic effects or KPI's. The conceptual architecture of Figure 2 shows the main components of the DST, including the chat-based user interface for the end-users as the front-end, the analysis of the questions from users and their embedding based on LLM, the semantic search performed in the different Data Bases (DBs) looking for the required information, and finally the generation of human-like text to create answers. On the left side, we observe also how AI/LLM extract all the COMMECT Knowledge (relevant information from large datasets, documents, etc.) to create the structured DBs.

This interaction between the DST and the ICP is very important for the full DST implementation, and both are being developed in parallel within COMMECT. The ICP implementation, which is also based on LLM solutions, is detailed in deliverable *D2.3 - Intent-driven Service Orchestration version 1* [2]. As observed in the conceptual architecture of Figure 2, once integrated, the DST will obtain a list of connectivity solutions from the ICP (lower green block). The role of the ICP is thus to facilitate the interaction between the DST and the available connectivity options, acting as a mediator and provider of connectivity candidates for solving user questions or requests. The choice of solutions will be based on collected data from all COMMECT knowledge, including LLs, stakeholders, connectivity platforms, etc. The list of connectivity solutions will include their corresponding costs, features, and chances of success/failure.

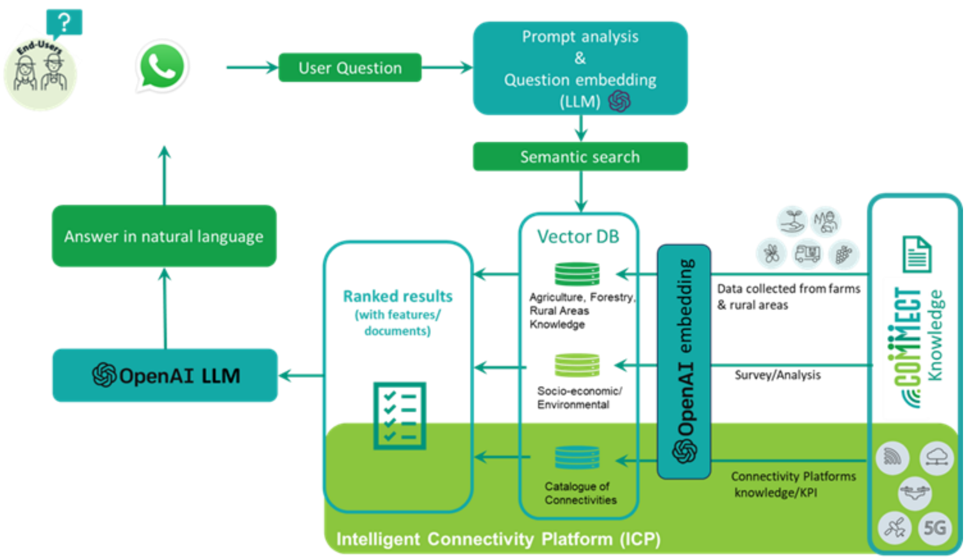


Figure 2. Conceptual architecture showing the integration of ICP in the DST.

To that end, the document starts with a literature review and state-of-the-art in generative AI and LLM, and AI-based decision support tools (Section 2). Section 3 provides the description of the overall methodology for requirements specification. Section 4 addresses the technical aspects of the initial and current DST implementations (releases v0.1 and v0.2 respectively), while Section 5 outlines the methodology for testing the tool. In Section 6, an overview of the current status of the tool is provided together with an outline of the plans for the second release. Finally, Section 7 summarizes this deliverable.

2. Background and State-of-the-Art

2.1 Decision support

In the business world, decision-making is often a high-stakes task reserved for authorized decision-makers [3][4]. These are typically key stakeholders like Chief Executive Officers (CEOs), managers, or project and team leaders. Still, it is impossible to guarantee the success of a business decision, regardless of who was responsible for it.

Making critical decisions can be particularly challenging when it involves data and analytics, as is now the case for most businesses. In today's data-driven landscape, companies are often dealing with vast amounts of information that can be difficult to interpret and utilize. Their challenge lies in assessing this data and using it to make better decisions that can impact the success of the business.

However, this much information covering a wide range of operational areas is not only overwhelming but also prone to errors and inaccuracies. Bad decision-making can be very risky for business.

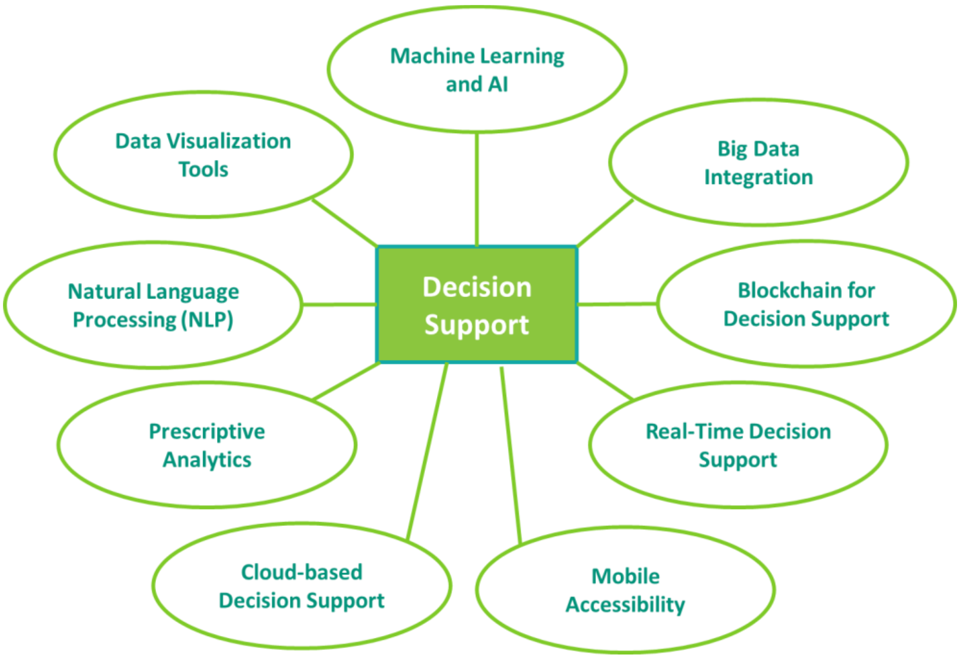


Figure 3. Features of Decision Support.

Decision support tools continue to evolve [5] with advancements in technology, and even if the decision-making process can use different technologies depending on the industry and its purpose, all the tools integrate some (and sometimes many) of the features listed below and show in Figure 3:

- *Machine Learning (ML) and AI Integration:* Decision support tools increasingly leverage ML and AI algorithms. These algorithms analyze vast amounts of data to provide insights and predictions, aiding in the decision-making processes. The prediction process of these algorithms aims at identifying the likelihood of future outcomes based on historical data. This helps decision-makers anticipate trends and make informed choices.

- *Data Visualization Tools:* Advanced data visualization tools help transform complex datasets into easily understandable visual representations. Interactive dashboards, charts, and graphs enable decision-makers to quickly grasp key insights and trends.
- *Natural Language Processing (NLP):* NLP technologies are employed to enable machines to understand and respond to human language. Decision support tools may use NLP to process unstructured data, such as text documents or customer feedback, providing valuable information for decision-making.
- *Natural Language Generation (NLG):* NLG enables analytics tools to offer easily understandable responses and generate written reports in a specific language, automating routine analysis to save time and money.
- *Prescriptive Analytics:* Going beyond predictive analytics, prescriptive analytics suggests actions to optimize decision outcomes. These tools not only forecast what is likely to happen but also recommend actions to achieve desired results.
- *Cloud-Based Decision Support:* Cloud computing has become integral to decision support tools, allowing for scalable and flexible solutions. Cloud-based platforms facilitate easy access to data, collaboration among team members, and the integration of various data sources.
- *Big Data Integration:* Decision support tools often integrate with big data platforms, enabling organizations to process and analyze large volumes of structured and unstructured data. This integration enhances the accuracy and depth of insights available for decision-making.
- *Blockchain for Decision Support:* In certain industries, such as supply chain management, blockchain technology is being used to enhance decision support. It provides a secure and transparent way to track and verify transactions, improving data integrity and trust.
- *Real-Time Decision Support:* The need for real-time decision-making has led to the development of tools that can process and analyze data in real-time. This is particularly crucial in dynamic environments where timely decisions are essential.
- *Mobile Accessibility:* Decision support tools are increasingly designed with mobile accessibility in mind. This allows decision-makers to access critical information and insights on the go, fostering agility in decision-making processes.

Though all the features are important in the process of decision-making, this section will focus on the impact of using ML (and AI) and the NLP.

2.2 Relevance of generative AI and LLM in decision-making

Concepts like ML, AI, Deep Learning (DL), and Data Science have always been a little confusing for the non-technical community, and something similar is happening between NLP and LLM. NLP is the study of human language by AI. It enables computers to comprehend, decipher, and produce text. LLMs are AI models, such as Generative Pre-trained Transformer (GPT)-3, that can produce, manipulate, and comprehend text in a way that is human-like, improving NLP skills. LLMs are thus a subset of NLP. As shown in Figure 4, they are not the same, but LLMs are a crucial component of NLP as they are an advancement of NLP capabilities.

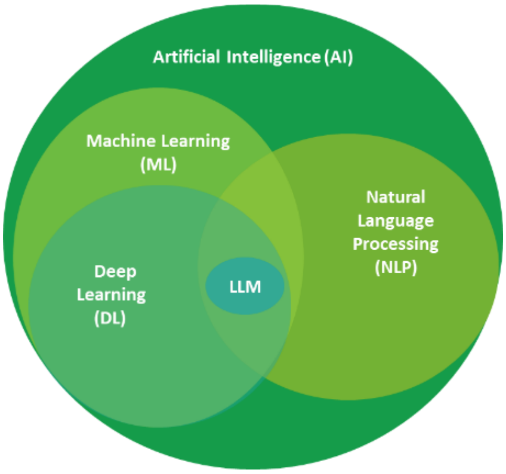


Figure 4. LLM, NLP and AI relationship.

LLMs like GPT have demonstrated significant relevance in decision support tools across various domains. Indeed, they can be fine-tuned for specific domains or industries, making them adaptable to the needs of decision support tools in various sectors.

LLMs excel in understanding and generating human-like text. This capability is crucial for decision support tools that involve interacting with users, understanding their queries, and providing informative responses. LLMs can process and generate text, documents and responses in a way that makes interactions more intuitive for users and save time for users.

LLMs can be used for extracting relevant information from large datasets, documents, or articles. This is particularly valuable for decision support tools that require summarization of complex information, enabling users to quickly grasp key insights without having to go through extensive data. Their language querying capabilities allow users to interact with decision support systems using plain language queries, making it more accessible for individuals who may not have a technical background. Not only that, but they can enable the development of conversational interfaces that can understand user queries, provide relevant information, and engage in dynamic interactions.

LLMs can be used to monitor news, social media, or other text sources in real-time. This is beneficial for decision support tools that require up-to-date information and insights from external sources.

Finally, LLMs can be integrated with other AI technologies, such as ML models, NLP, and recommendation systems. This integration enhances the overall capabilities of decision support tools by combining different AI approaches.

2.3 State-of-the-art of Decision Support Tools using AI

AI and Machine Learning have become an increasingly popular tool in recent years, given their ability to automatically detect patterns in data and make predictions about future events. This can be extremely useful for making decisions in a wide range of domains. AI can provide better information given that it can sift through extremely large amounts of data, and thus provide better information to decision makers. AI is able to automate the decision-making process. It is simply not possible for human beings to make optimal decisions all of the time. This is especially true in industries where the data is constantly changing. In these cases, machine learning algorithms can be used to automatically make decisions as trends change and evolve. AI improves accuracy by identifying patterns in data that humans may not be able to see, it can drastically improve the accuracy of its predictions. It can also create models that simulate different decision scenarios and help identify the best course of action. And as new data becomes available, it can be used to constantly update and refine decision models.

There are many different ways end-users or stakeholders can use decision support software to their advantage, to evaluate specific operations, to manage inventory, to optimise sales, etc. The fields where decision support is used are numerous, we discuss some examples below.

Examining the LLs in COMMECT, we see that decision support could contribute towards the following domains and use cases: disaster response (related to Norway LL), education and network communication (related to all LLs), livestock transport (related to Denmark LL) and agriculture (related to Luxembourg, Türkiye, and Serbia, LLs).

In education, universities use decision support tools to know how many students they currently have enrolled. This helps them predict how many students will register for particular courses or whether the student population is sufficient to meet the university's costs.

Other examples of decision support tools are 1) DreamBox [1], which uses AI to provide personalized learning experiences for students, adapting the curriculum based on individual progress and learning styles, 2) Cognii [7], an AI-powered virtual tutor that assists students in understanding and mastering complex subjects through natural language processing. Indeed, the DST and COMMECT will aim at educating rural and remote communities (and stakeholders), often isolated, due to lack of Internet access, on digitalization, connectivity, business opportunities, socio economic and environmental impact, etc.

In the field of humanitarian aid and disaster response, we can mention:

- Artificial Intelligence for Disaster Response (AIDR) [8]: Uses machine learning to analyse social media data during disasters, helping humanitarian organizations identify critical information and respond more effectively.
- Global Database of Events, Language, and Tone (GDELT) [9]: Analyses news articles and reports worldwide to provide real-time insights into global events, aiding decision-makers in disaster response and crisis management.

In the scope of network communications, manufacturers develop their own decision support tools, e.g.:

- Nokia AVA Cognitive Services Platform [10]: Uses AI to automate network management tasks, optimize network performance, and predict potential issues in telecommunications networks.
- Ericsson Operations Engine: Leverages AI for predictive analytics and automation in managing telecommunications networks, improving efficiency and service quality.

In agriculture, decision support tools using AI play a crucial role in optimizing crop management, resource allocation and overall farm productivity. In forestry, they can help not only with wildfire risk prevention and timber harvesting optimization, but also to improve forest health.

Table 1 shows examples of AI-based decision support tools used in agriculture and forestry. It should be highlighted that COMMECT DST goes beyond the usual decision-making skills, e.g., the ones shown in Table 1, and it also advises on the most adapted connectivity solutions to be used and the business models that can be implemented, while taking into account the ecological and economic impact of the choices and decisions.

Table 1. Examples of decision support tools using AI.

Decision support tools	Objective	How is AI used?
FarmWise [11]	*Agriculture* Weed detection and precision farming	Computer vision and machine learning algorithms analyse images to identify and selectively eliminate weeds, reducing the need for herbicides and optimizing crop yields.
Blue River Technology [12]	*Agriculture* Precision agriculture and crop optimization	Utilizes computer vision and machine learning to identify and apply precise herbicide sprays, reducing the overall use of chemicals while targeting weeds.
AgroSight [13]	*Agriculture* Disease detection in crops	Employs image recognition and machine learning to identify diseases and pests in crops based on images captured by drones or smartphones, aiding in early detection and treatment.
Motorleaf [14]	*Agriculture* Precision agriculture and crop optimization	AI-driven agri-tech company that has created an automated indoor farming operating system. Using AI and advanced machine learning algorithms, their system monitors and predicts environmental conditions to adjust the farm settings automatically for optimal plant growth.
Gamaya [15]	*Agriculture* Precision agriculture and crop monitoring	Integrates hyperspectral imaging and AI algorithms to provide insights into crop health, identify nutrient deficiencies, and optimize the use of fertilizers.
NCX [16]	*Forestry* Forest Inventory and Management	It utilizes machine learning algorithms to analyse satellite imagery, creating detailed forest inventory maps. This tool assists in decision-making related to timber harvesting, carbon sequestration, and overall forest health.
IBM PAIRS Geoscope [17][18]	*Forestry* Wildfire risk assessment, vegetation analysis	Used by *The Weather Vegetation Management – Predict* solution. It integrates AI and machine learning with satellite data to analyse historical and real-time data. It is used for applications such as predicting wildfire risks, vegetation and tree monitoring (average height, buffer zones and threats).

The specific application of AI in decision support tools for livestock transport might be a niche area with limited examples. However, there is potential for AI to play a role in optimizing logistics, ensuring animal welfare, and improving the overall efficiency of livestock transport.

2.4 State-of-the-art of Decision Support Tools using LLMs

Decision support tools using LLM are currently being used and applied to different domains. Some examples are shown in Table 2.

Table 2. Examples of decision support tools using LLM.

Decision support tools using LLM	Description
OpenAI's GPT-3 [19]	OpenAI has an Application Programming Interface (API) for GPT-3, and developers can use it to integrate the model into various applications.
GitHub Copilot [20]	GitHub Copilot, built on OpenAI's Codex, is designed to assist developers in coding.
Copy.ai [21]	Copy.ai is a platform that uses language models to assist with content generation.
ROSS Intelligence [22]	ROSS Intelligence is a legal research tool that utilizes AI, including LLMs.
DeepL Translator [23]	DeepL is known for its language translation services using LLMs.

The use of LLMs in decision support tools for domains like agriculture, transport, etc. is indeed being discussed by the scientific community, but tools are not yet available. An example of NLP-based tool is shown in the AgriBERT approach [7] for matching food description and nutrition data.

The International Food Policy Research Institute (IFPRI) recently discussed the impact of AI/LLM in agricultural extension services, which often face significant challenges, including weak institutional capacity, inadequate reach, and limited access to up-to-date scientific knowledge [25]. Language is also a significant barrier. Advisories provided only in the world's predominant languages cannot easily be understood in many communities speaking regional or local languages and dialects. A recent study published in Nature Food [26] explores the potential of generative artificial intelligence technologies (chatbots powered LLM) as a human-machine interface to facilitate agricultural extension services. Using ChatGPT as an exemplar LLM-based application, they particularly focused on two aspects where LLMs can be transformative:

- *Making science simple*: Farmers or extension agents are not typically trained to find scientific knowledge, and even if they do search for it, they may not have time and background to read through lengthy, technical reports. LLMs can be tasked to find the most relevant knowledge products, analyse them, and distil key points, which can be presented in an accessible (and translated) language that is easier to understand and apply in practice.
- *Making advisories personal*: Most agricultural extension services provide a standardized package of advisories pre-developed for crop and agricultural zones. However, those may not apply to many of the unique and complex challenges smallholders face. With larger training datasets and scalable computational capacities, LLMs can be used to assess farmers' specific challenges and generate data-driven, personalized recommendations.

2.5 The capabilities of LLMs

Informed decision-making remains one of the key factors for successful and efficient execution of business processes across all domains. While data analysis and expert opinions traditionally guided our choices, the landscape is undergoing a transformative shift. Generative AI, particularly LLMs, are emerging as powerful tools, leveraging the power of language to create foundations for a new creative data-driven era [27]. LLMs like GPT-3 and PaLM boast immense neural networks trained on huge amounts of text and code to:

- Grasp complex questions: by understanding the nuances of natural language, LLMs can delve into problems with layered contexts and multiple facets.
- Analyze massive datasets: sifting through mountains of data, they extract key insights, trends, and relationships that might elude traditional methods.
- Generate diverse perspectives: LLMs can simulate scenarios, explore alternatives, and even formulate counterarguments, offering a spectrum of possibilities for consideration.
- Synthesize knowledge: they can bridge gaps between disparate information sources, drawing insightful connections and forging novel pathways to solutions.

These capabilities translate into powerful decision support tools, reshaping how we navigate complex choices, for example:

- Empowers strategic decision-making with ChatGPT [28]: analyzes complex data, generates hypothetical scenarios, and predicts future trends to help users navigate complex choices.
- Scenario planning with NarrativeAI [29]: this platform leverages NLP to generate hypothetical narratives based on users' decisions, allowing them to visualize potential consequences and identify vulnerabilities in their plans.
- Gemini [30], AI Assistant, developed by Google AI: the tool allows exploration of complex data through interactive questioning, uncovering hidden patterns, and generating informative summaries, aiding in the analysis and interpretation of critical information.
- Perplexity [31] supercharges research & decision-making integrates real-time web search with AI language models, offering deeper insights, factual context, and diverse perspectives for tackling complex choices.

LLMs, however, are not simply data crunching machines, but can offer support in the creativity domain:

- Brainstorming and ideation: LLMs can generate a multitude of fresh ideas, sparking innovation and helping users break free from conventional thinking.
- Personalized Recommendations: LLMs can tailor solutions to specific context and needs, considering unique goals, resources, and risk tolerance.
- Human-Centered Communication: By generating clear, concise, and persuasive messaging, LLMs can facilitate effective communication with stakeholders, boosting buy-in and collaboration.

While transformative, LLMs are not without limitations. Mitigating data bias, where unfair or prejudiced information in the data can color their recommendations, remains a pressing challenge. Additionally, their internal workings can be less transparent, making it difficult to fully understand how they arrive at their outputs. Despite these challenges, the future of decision-making with generative AI and LLMs is bright. As research advances, LLM capabilities will undoubtedly expand, improving their accuracy, explainability, and ethical robustness. Moreover, their integration with other AI technologies like computer vision and robotics will create even more comprehensive decision support systems, blurring the lines between human and machine intelligence.

Ultimately, LLMs are not here to replace human judgment. Rather, they offer a powerful new toolset to augment our decision-making capabilities, leading to more informed, creative, and ultimately, better choices. As we learn to harness the power of these "language wizards," we have an opportunity to unlock a new era of intelligent decision-making.

3. Requirements

This section outlines the methodology employed to define and prioritize requirements for the COMMECT Decision-making Support Tool (DST). The overarching goal is to create a DST that bridges the gap between end-user needs and available connectivity solutions, providing users with tailored guidance and informed recommendations, taking into account the socio-economic and environment related impacts. By facilitating optimal decision-making for diverse user groups (farmers, municipalities, community members, and companies), the DST empowers informed choices for both initial investments and ongoing operations. The table below presents a list of the main requirements for the DST. These requirements have been derived using the following activities:

- Experience of COMMECT partners working with farmers: insights gained through direct interaction with the target user group.
- Survey analysis: findings from a survey conducted to capture user needs and preferences.
- LL leader interviews: valuable perspectives gleaned from discussions with representatives of LLs involved in the project.
- Workshop outcomes: key takeaways from a COMMECT internal workshop organized during a plenary project meeting in Oslo (October 2023) helped specifically to address DST development. During this workshop, the expected features of the COMMECT DST were discussed with the representatives of each COMMECT LL. The ask was not to limit themselves to the project scope, i.e., is a specific feature implementable technically or implementable in the project scope, but to take perspective of their respective stakeholders and define the requirements assuming no limits in resources required to implement.

These diverse sources of information have ensured that the DST requirements are grounded in real-world needs and priorities. Evaluating the effectiveness, efficiency, impact, adoption, technical performance, and return on investment of the DST platform will be crucial for LLs organizations, and communities.

Table 3. Requirements for development of DST.

ID	Description	Expected outcomes	Priority
R100	The DST must ensure transparency when generating responses by clearly identifying the sources of information used when requested.	Transparent and accountable information source will empower users to assess the credibility and reliability of the presented data and recommendations. This transparency builds user trust and confidence in the tool's capabilities and enhances its overall effectiveness.	High
R110	The DST must provide accurate answers.	Accuracy is of paramount importance. The DST has to be designed in such a way to maximize the accuracy and minimize the so-called “hallucinations” (i.e. confident sounding incorrect answers).	High

R120	The DST must provide a clear process for updating the private knowledge base used to generate responses.	The context in which the DST works evolves with the technological, business, environmental, societal and legislative changes, requiring the private knowledge base to evolve as well. The process for updating the knowledge base has to be described to allow changes by non-experts.	High
R130	The DST must be designed as an open and extendable tool that can be integrated with the 3rd party tools to enhance functionality.	This requirement aims to make the DST versatile and more powerful by integrating with third-party tools. Users can benefit from an extended ecosystem, ensuring that the DST remains adaptable and relevant in the domain of evolving connectivity solutions and technologies.	High
R140	The DST must provide a multilingual interface.	Enabling users to use the DST in their own language will increase usability of the tool from the user's perspective. This contributes to a positive user experience.	High
R150	The DST must ensure compliance with relevant ethical considerations.	By meeting this requirement, the DST becomes a tool that not only provides valuable recommendations but does so in a manner that respects ethical principles. This alignment will enhance user trust, satisfaction, and the overall positive impact of the DST on decision-making processes.	High
R160	Incorporate a feedback loop to collects insights from users.	Gather qualitative feedback from users and stakeholders to assess their satisfaction with the DST responses. Use this feedback to continuously improve the DST's usability and functionality.	Medium
R170	The DST should be aware of relevant legal requirements in the rural region where it is used.	Ensuring the responses comply with local and national regulations.	High
R180	The DST should provide location specific recommendations.	Tailored advice for connectivity solutions based on user's location.	Medium
R190	Cost-sensitive analysis	The DST should have the capability to analyse and suggest solutions within specified budget constraints (based on inputs of users).	High
R200	The DST must be able to generate responses related to connectivity options.	Generated responses will not only list available connectivity options in the region but will also rank them based on the user needs and constraints (scenario, financials, environment).	High
R210	The DST should have knowledge of the region that can have impact on the choice of communication options and digital services.	The DST should have access to maps of the region incorporating standard information as well as information about natural heritage and environmental protection maps, to ensure that this information is considered for biodiversity preservation and/or ecological impact. This	Low

		includes visualizing the footprint of infrastructure and its effects on the local ecosystem.	
R300	The DST must be able to generate responses related to potential and available business models based on the information provided by the user.	As various business models are possible and used on the market, the DST will provide support to users regarding the design of business models for connectivity solutions and to understand the benefits and costs of certain design decisions and the impact on the users and environment based on the provided information.	High
R400	The DST must be able to generate responses related to characteristics of rural areas such as agriculture, viticulture, and forestry.	By providing accurate and reliable information about the rural areas based on the information provided by users (farmers, winegrowers, forest owners), factoring in the local and global market conditions, climate, and weather, as well as user's preferences, the DST becomes an invaluable tool for rural stakeholders, empowering users to make informed decisions that maximize economic returns while considering the broader market dynamics and sustainable operations in rural areas.	High
R410	The DST should be able to generate responses related to servicing and maintenance of agricultural machines.	The DST will become a comprehensive tool for farmers and agricultural stakeholders, offering guidance on routine. The result is improved reliability, reduced downtime, and enhanced overall efficiency in agricultural operations.	Low
R420	The DST should be able to generate responses related to servicing and maintenance of viticultural machines and devices.	The DST will become a comprehensive tool for winegrowers and viticultural value chain stakeholders through providing advice on the use, maintenance and operation of devices and machines used in viticulture. This will help to improve the effectiveness and efficiency of viticultural operations	Low
R430	The DST should be able to generate responses related to servicing and maintenance of forestry-related machines and devices	Analogously to R410 and R420, the DST will provide advice on equipment, machines and devices used in forest management. This will help to improve the effectiveness and efficiency of forest management operations.	Low
R500	The DST should be able to generate responses related to the environmental impact of selecting different operations in agriculture, viticulture, and forestry	While sustainable production and reduction of CO2 are becoming key points that needs to be considered when planning and execution agricultural, viticultural and forest operations, understanding the impact of different choices regarding environment is not readily available. The role of the DST is to help bridging that gap.	Medium

R600	DST should provide a mechanism for evaluating and comparing the performance of various solutions.	In addition to providing recommendations on the most suitable solution based on the user's inputs, DST should be able to compare solutions based on their characteristics and the user's needs.	Medium
R900	The DST should be able to drive conversations with users to elicit all relevant information before generating responses to the original query.	DST users are not experts in defining questions. Hence, the DST should, based on the initial query, seek for clarifications and additional information to be in the position to generate useful answer tailored to the user and her needs.	High
R910	The DST should incorporate a speech/audio interface to enhance user interaction.	Increased DST accessibility and inclusivity, providing a seamless and natural means for users to engage with the tool.	Low
R920	The DST should be able to generate images and video/audio clips as responses.	The ability to provide responses in formats other than text will increase the richness of responses and the ability to present generated responses in ways that are easier to understand by users, contributing to more effective decision-making processes.	Low
R930	The DST should allow users to customize the depth of responses, offering options for both brief and verbose information based on individual preferences and needs.	This customization enhances user satisfaction and engagement, contributing to a more effective decision-making experience.	Low

The requirements listed in the table above are defined from the long-term perspective of various stakeholders, i.e. what functionalities they would like a DST to offer, without taking into account the feasibility of implementing these requirements in the scope of the COMMECT project. The goal was to have a rather comprehensive list of features of a DST that can be utilized for expansion of the COMMECT DST towards a full-fledged commercial DST offering (i.e., beyond the scope of the COMMECT project and aiming at the future exploitation of the DST and its results).

Task 3.4, responsible for implementation of DST in the project, will prioritize implementation and validation of the core features, given the available resources (work effort, available time and the domain specific knowledge that will be available/generated by the project).

4. COMMECT DST implementation description

The starting point for implementation of the DST is the COMMECT architecture described in the COMMECT deliverable D1.3. Following up on that, the initial release of the COMMECT DST leveraged generative AI/LLM and the Retrieval-Augmented Generation (RAG) [32][33] approach for enhanced language understanding and response generation.

We considered the following two approaches for implementation of the COMMECT DST: 1) fine tuning a LLM model and 2) using RAG. We chose RAG for several reasons:

- Training a new model requires considerable time and resources.
- Having in mind that the DST is envisaged as a tool for a specific region, training a model for each region would make it prohibitively costly. Further, it is expected that new relevant knowledge will be provided on regular bases, which would require additional model tunings.
- The DST should provide summarized information from the available information pool, which is one of the main application scenarios for the RAG-based solutions.

Instead of using public knowledge to provide answers, the COMMECT DST relies on information and knowledge provided by the project (deliverables) and sources selected and curated by the project partners to generate responses. Enterprise-grade environment is used for development and provision of the service to ensure high reliability and security. Web and mobile interface are provided for interaction with the tool, including access via WhatsApp messaging application.

The remainder of the section provides descriptions of two implementations of the COMMECT DST done so far:

- DST v0.1: this was the very first implementation based on the tools available at the moment, used for internal testing of the concept and the approach. This version was shown to consortium during Plenary COMMECT meeting in Oslo in October 2023.
- DST v0.2: this version is based on an enterprise-grade platform availing of the most recent frameworks. This version is made available to the COMMECT consortium together with the publication of D3.3 (February 2024).

The roadmap for subsequent DST releases during 2024 is provided in the Table 4 below.

Table 4: Roadmap for subsequent DST releases

Release #	Date	Features
0.3	Jun 2024	Ability to use different LLMs. Updated version of LLM used for testing.
0.4	Sep 2024	Conversational style initial version LLM updates. Initial feedback addressed.
0.5	Dec 2024	Conversational style updated. LLM updates. Integration with ICP (see Figure 2)

4.1 DST v0.1 implementation

The initial DST implementation utilized the LangChain library [34] to orchestrate interactions with the OpenAI GPT-3.5 model [35] FAISS vector database [36], supporting file formats like .txt, .pdf, and .doc. This approach offered advantages like control over data chunking and conversation history storage within the API logic. For interaction with users, a text-based interface accessible via web browsers was implemented first, followed by integration of WhatsApp messaging app as an even more convenient approach.

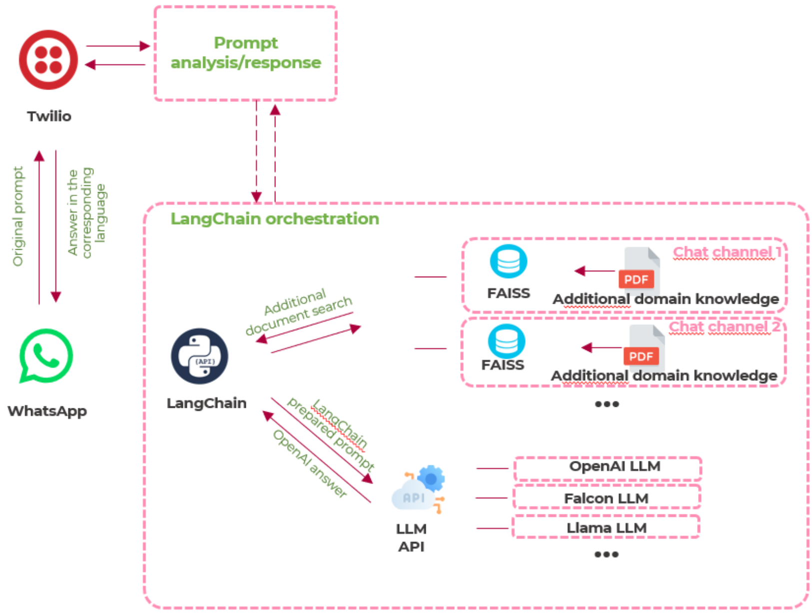


Figure 5. First architecture of Decision-making Support Tool.

4.2 DST v0.2 implementation

The rapid evolution of the AI domain, including LLM-related technologies, results in new tools and approaches becoming available almost daily. In parallel with the initial testing of DST 0.1, we were looking into other architectural choices such as Microsoft Azure AI Search vector database and search functionalities. The Microsoft Azure platform's capabilities, particularly the ability to select specific "skillsets", yielded superior results. By tailoring the skillset within Azure AI Search, we achieved a higher quality search across a larger document set (440 documents) compared to the initial implementation. For that reason, the second version of the DST tool was implemented using a different architecture, presented in Figure 6.

The core part of DST 0.2 contains two AppServices [37], an Azure OpenAI instance [38], Azure AI Search [39] instance and Azure Storage [40].

The AppServices host two API services. The first API is linked to a relational database to log interactions between the end user and the LLM. This logging is especially important for communications through the WhatsApp application, as it helps providing context to the queries due to the API's stateless design. This API is developed using C# ASP.NET Core. The second API packaged as a Docker container [41] and uses PromptFlow [42] logic created with ML Studio [43]. It takes a user's question, converts it into an embedding [44], then searches a knowledge base for relevant information. After finding the relevant pieces of information, it sends everything to LLM to generate the complete response. The API then takes the LLM's answer and sends it back to the user as the final response.

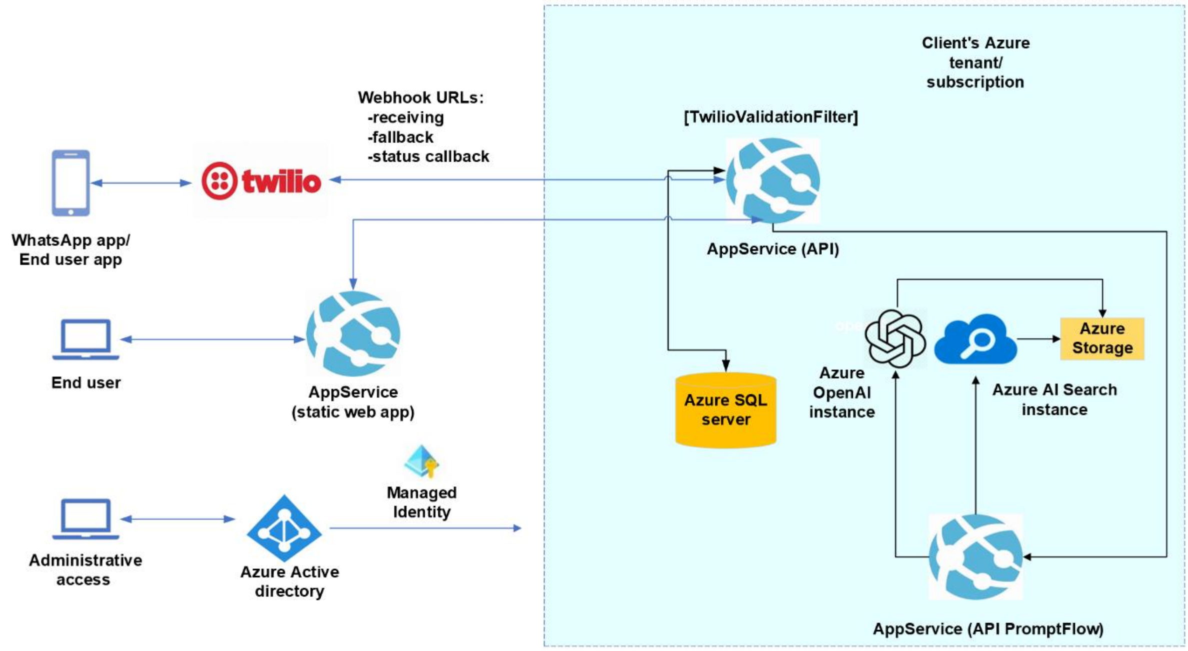


Figure 6. Current architecture of Decision-making Support Tool.

As in DST v0.1, WhatsApp and web browsers are available as user interfaces. The **WhatsApp** service connects with the first API described above. This connection is made possible through a Twilio service [45], which utilises the Twilio SDK (Software Development Kit) for integration. The API is protected by a feature called "TwilioValidationFilter", which only allows access from the Twilio service. This ensures that the interactions are secure and controlled.

The **Web Application** is designed as a static application built using React framework [46]. It connects to the previously mentioned API using the Hypertext Transfer Protocol (HTTP), with Axios [47] used for sending and receiving data. An Azure SQL Server Database (DB) is linked to this API and serves as the storage for user questions, their answers, and additional logs. Importantly, the database uses the phone number from which a question came as a unique identifier. This allows the system to provide responses that are relevant to the user's previous interactions, making the conversation more personalized and context aware.

Azure Cognitive Search is a cloud-based search service provided by Microsoft Azure that offers a rich set of capabilities for building sophisticated search experiences in web and mobile applications. It is built on top of the scalable infrastructure of Microsoft Azure, allowing us to quickly implement full-text search functionalities without the need for managing infrastructure. We utilised its primary workloads, indexing and querying. Indexing is an intake process that loads content into search and makes it searchable. Querying can happen once an index is populated with searchable content when client sends query requests to a search service and handles responses.

Azure AI Search Service is used for creating vector databases and search data. **Azure AI Search Index** (Vector DB) stores information obtained from documents uploaded to Azure Storage and translated into the form of embeddings. So, Vector DB is database that stores data as high-dimensional vectors, which are mathematical representations of features or attributes. A vector has a certain number of dimensions, which can range from tens to thousands, depending on the complexity and granularity of the data. The vectors are made by applying a transformation or embedding function to the raw data. In this case, vectors are created using an embedding model, text-embedding-ada-002. Figure 7 represents how vectors are created from text using embedding model. This service acts as the knowledge base's search engine, efficiently indexing and querying documents to retrieve the most relevant information for the LLM's response generation. **Azure OpenAI** is a cloud-based platform that has enabled us to build and deploy AI models quickly and easily.

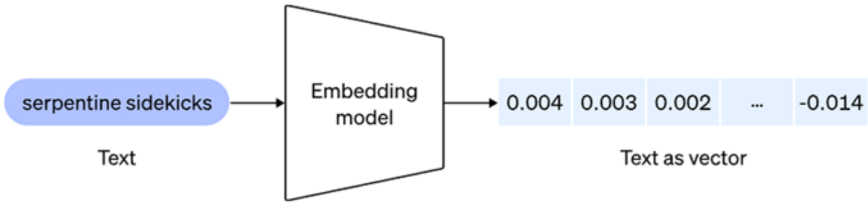


Figure 7. Creating vectors from test.

The main advantage of a vector database is that it allows fast and accurate similarity search and retrieval of data based on their vector distance or similarity. This implies that rather than relying on traditional database search methods, which look for exact matches or follow specific rules, we can utilize a vector database. A vector database allows us to search for data that is closely related in meaning or context, not just in exact wording or predefined parameters. This ability to find data based on its underlying significance is a key reason why we chose to use a vector database for this project. Key advantages of the described system lie in its scalability, adaptability, efficient knowledge retrieval, secure communication, and contextual understanding. Its cloud-based architecture allows for easy scaling to meet user demand and integrates smoothly with new technologies as they emerge. With the help of Azure AI Search and a vector database, it quickly and accurately finds the necessary information within the knowledge base, which enables the Language Learning Model to provide well-informed responses. Secure communication channels, such as Twilio, protect data privacy and authenticate users, building a foundation of trust and transparency. Moreover, the system's ability to store conversation logs and user profiles in the Azure SQL Server database enhances its understanding of the context of ongoing interactions. This capability ensures that the support offered is both personalized and relevant to the user's needs.

4.3 Explanation of the data flow and user interaction

The steps that represent the data flow and interaction with the user are listed below.

- Users can submit questions via WhatsApp or WebApp.
- These questions are forwarded to the AppService API.
- The AppService API sends the questions to the PromptFlow API.
- The PromptFlow API process is depicted in Figure 8:
  - A query vector is created from the user's question using a transformation model.
  - This query vector is used to perform a search in the database.
  - Azure AI generates a ranked list of search results.
  - These results, along with a prompt, are sent to the LLM.
  - The prompt guides the LLM to generate a relevant response.
- The response from the LLM is returned to the AppService.
- The AppService API logs the questions and responses in a database and deliver response back to the user.

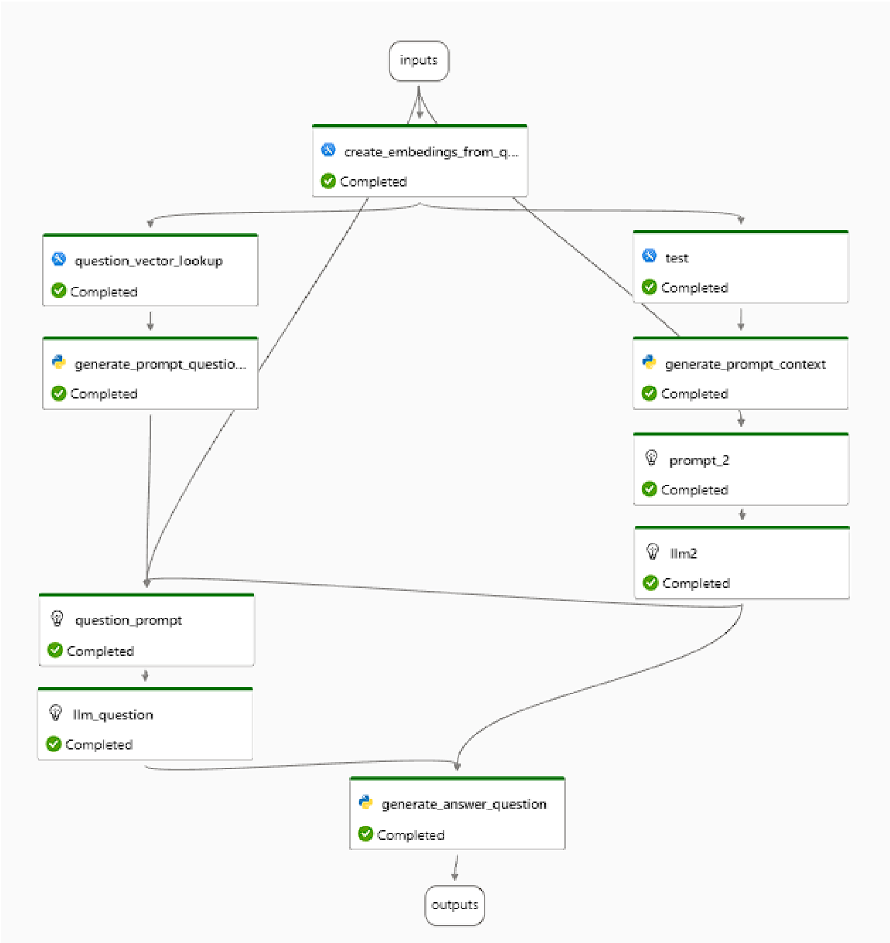


Figure 8. Architecture of API PromptFlow.

4.4 Private knowledge base

The DST is designed to utilize a specialized knowledge database compiled from documents related to the COMMECT project, i.e. AGRI-KB as defined in the COMMECT Architecture (Deliverable D1.3). These documents come in various formats, including .pdf, .docx, and .csv. To prepare this database, it was essential to convert the documents into a format that makes the data both understandable and accessible. At the foundation of document preparation for RAG solution is the logical organization of content. This includes employing hierarchical headings, bullet points, and clearly demarcated sections to delineate topics and subtopics. Such a structure aids the RAG model in efficiently locating and retrieving specific information, like an index in a scientific manual allows for quick reference. Every piece of information included should be relevant and presented in a clear, straightforward manner. Redundancy and overly complex language are to be avoided, as they can obscure the core information and reduce the document's utility for both retrieval and generation processes. Moreover, the integrity of the document is crucial. It should only contain information that is accurate, current, and verified, with all facts and data sourced from reputable references. This commitment to quality and reliability ensures that the RAG model has access to dependable data, significantly enhancing the relevance and accuracy of its generated content. Therefore, our task was to process the collected documents following the described principles.

To construct the database, we employed Azure Indexer. This tool is capable of extracting text from the documents and then populating a search index by mapping the extracted data from the documents directly to corresponding fields within the search index. Whenever there is a need to add new data or modify existing entries, it is necessary to rebuild the index. This rebuilding process is efficient, as it only incorporates the new or updated data.

DST v0.2 is using GPT-3.5-Turbo model. It is a state-of-the-art language model developed by OpenAI. GPT-3 is a variant of the transformer architecture and is known for its remarkable ability to generate human-like text based on the input it receives. It was trained on a massive amount of text data and can perform a wide range of natural language processing tasks, such as text generation, translation, summarization, question answering, and more. It has a lot of hyperparameters that can be used to configure and guide performance of the model in terms of the generated output. For our implementation, we used "*temperature*" and "*max_tokens*" parameters.

"*Temperature*" is a hyperparameter that controls the randomness of the model's output when generating text. Temperature affects the level of diversity in the generated text. When the temperature is set to a low value (close to 0), the model produces more deterministic and focused output. On the other hand, when the temperature is set to a high value (1 or above), the model becomes more creative and generates text that is more random and diverse. For COMMECT DST 0.2 the "temperature" parameter is set to a low value because we want answers based on data from database, i.e., limiting the creativity of the model when generating responses. The current architecture involves two calls to the LLM. For the first portion of the DST response, a temperature of 0 is utilized to prioritize retrieving the most accurate information possible from the knowledge base. The second part of the response is formulated to prompt further dialogue with the user; here, a temperature of 0.5 is employed to encourage greater creativity.

The "*max_tokens*" hyperparameter is used to limit the length of generated text or the number of tokens in the output. Tokens can be individual words, punctuation marks, or sub-word units, depending on the tokenization used by the model. By setting this parameter to a specific value, we control the length of the response generated by the LLM. This can be useful to ensure that the generated text remains within a certain character or token limit. By testing different options, we concluded that the optimal value for the max_tokens parameter in our solution is 2000 tokens. Azure OpenAI uses specialized tokenizers for text input in models GPT-3.5 Turbo based on Byte-Pair Encoding (BPE). These tokenizers efficiently manage a vast vocabulary by breaking down text into smaller units, enabling accurate text processing in various languages and styles.

The DST is deployed and can be accessed at the following URL:

https://decision-support-tool.dunavnet.eu/.

The interface and exemplary use of the tool is illustrated in the Annex I.

5. Testing and validation plan

LLMs integrated with specialized knowledge bases hold immense potential for decision support across various domains. However, ensuring their effectiveness from the end-user perspective necessitates a rigorous testing and evaluation process. In this section, we outline the COMMECT DST evaluation strategy focusing on user-centric test cases and addressing both objective and subjective aspects of its functionality. Validation will be done when the COMMECT knowledge base stabilizes, i.e., with the release of COMMECT DST v1.0.

The evaluation will commence by defining a range of tasks aligned with the intended functionalities of the decision support tool. These tasks will directly cater to the end-user's, primarily farmers, typical decision-making needs. For each task, tailored metrics will be developed to measure the LLM's effectiveness from the end-user's perspective. Such metrics may include:

- **Accuracy and Relevance**: Does the tool recommend relevant options aligned with the user's query and context? Are the presented options factually accurate and aligned with the private knowledge base?
- **Completeness**: Does the tool provide a comprehensive overview of potential solutions, considering various aspects and trade-offs?
- **Clarity and Explainability**: Are the tool's explanations and justifications for recommendations clear, understandable, and tailored to the user's level of expertise?
- **Actionability**: Does the tool provide concrete next steps and resources to facilitate informed decision-making?
- **User Satisfaction**: Do users find the tool helpful, easy to use, and reliable in assisting their decision-making process?

The following are examples of user-centric test cases that will be considered when creating test plans for evaluation of COMMECT DST:

- Scenario 1: A farmer seeking to invest in sustainable irrigation solutions receives recommendations tailored to their land size, crop type, and budget, along with clear explanations of potential benefits and drawbacks for each option.
- Scenario 2: A municipal official tasked with selecting a renewable energy provider to support farming activities in the municipality receives a comprehensive analysis of different options, considering factors like cost, environmental impact, and compatibility with existing infrastructure.

While quantitative metrics offer valuable insights, human judgment played by user groups representing the intended target audience remains crucial for comprehending the subjective aspects of the tool's performance through methods like:

- Usability testing: Observing users interact with the tool and gathering feedback on its ease of use, interface design, and navigation.
- Comparative evaluation: Comparing user experiences with the LLM-powered tool to alternative decision support methods or no support at all.
- In-depth interviews: Probing user perceptions of the tool's helpfulness, reliability, and impact on their decision-making confidence.

Similar to testing LLM chat interfaces, several challenges require careful consideration:

- **Data Bias and Fairness**: The private knowledge base and training data should be thoroughly evaluated for potential biases that could skew recommendations and disadvantage certain user groups.
- **Explainability and Interpretability**: Users should be able to understand the rationale behind the tool's recommendations, fostering trust and enabling more informed decision-making.

- **Evolving User Needs and Context**: The tool should be adaptable to accommodate changing user needs, preferences, and evolving real-world contexts.

The evaluation plan will be based on several best practices:

- **Iterative Testing and Refinement**: Continuously testing and refining the tool based on user feedback, focusing on areas with the most significant impact on decision-making effectiveness.
- **Transparency and Trust Building**: Providing clear explanations of the tool's limitations and capabilities, along with mechanisms for user feedback and improvement.
- **Integration with User Workflows**: Seamlessly integrating the tool into existing user workflows and decision-making processes.

By diligently implementing this testing and evaluation plan with user-centric test cases, we aim to gain a comprehensive understanding of the LLM-powered decision support tool's effectiveness from the end-user's perspective. This knowledge will guide its further development and refinement, ultimately contributing to more informed, efficient, and user-centric decision-making across various domains.

6. Status of implementation and the next steps

The digital age demands agility and expertise even in rural regions, where access to information and specialized knowledge can be limited. To that end, we designed and implemented the first version of the COMMECT DST. The tool is built on generative AI/LLM technology. As the AI domain is rapidly changing as we speak, the implementation also had to be very dynamic and adaptable to the evolution of the technology resulting in the change of some of the functional blocks between the v0.1 and v0.2 of the DST releases. While the current version of the DST, COMMECT DST v0.2, is stable (for the moment) and performs well, (during the writing of the deliverable), we expect further changes in the choice of LLMs and adaptations of DST to avail of the most recent advancements.

First, technical implementation was tested using a set of documents not related to rural regions to create the DST's private knowledge base. The main goal of this testing exercise was to verify that information flow is working correctly, that interaction between the solution components is implemented correctly, etc. The accuracy of the generated responses was not considered in this phase.

Once we were confident that the solution worked technically, we focused on testing the accuracy of the content. For that purpose, we used documentation from a known solution (agroNET [48]) combined with COMMECT documents like COMMECT DoA (Description of Action) as these are domains in which the COMMECT DST development team has expertise and can verify the accuracy of the responses. In addition, we also leveraged discussions and interactions with COMMECT partners during the Plenary meeting in Oslo, as well as interactions with potential stakeholders as part of the Synergy Days to further test this. In this phase, the DST was configured to cite the sources used when generating responses to ease their validation. This activity helped us not only to validate accuracy of the generated responses, but also to better understand how to structure documents in the knowledge base to improve the accuracy of the responses which will be of great help going forward.

The process of collecting and curating documents for the DST has started as well. The requirements regarding the documents, supported types, etc. were shared with the COMMECT partners. A shared folder for collecting documents was created and the initial documents have been already uploaded. In the coming period, as the documents are gathered, the DST knowledge base will be periodically updated and integrated into the tool.

In addition to following development of the technology, updating the tool accordingly and creating the knowledge base, the work on prompt engineering and implementing more intelligence into the tool to enable the DST to drive conversations with end users will continue. By implementing this conversational intelligence, we expect the tool to be able to collect all required information from end users (location of the fields, crops, machinery, etc.) through a dialogue that is resembling situations when farmers ask consultants for help. This approach, we believe, will make the tool even more usable and appreciated by the users.

Finally, as detailed before, the ICP will be integrated to the DST in future releases allowing to interact by means of an open API. More details of these implementations will be provided in future deliverables, i.e., version 2 of this deliverable (i.e. *D3.5*) and version 2 of deliverable *D2.3 (i.e. D2.6).*

7. Conclusion

In this deliverable, we have presented an overview of the architecture design for the COMMECT DST, as well as a description of the first release (v0.2). The DST is intended to support rural stakeholders in decision making on the (deployment, realization or use) of connectivity solutions. The tool is leveraging Generative AI/LLM as the underlying technology to support the generation of responses to questions posed. A list of requirements is given for the DST which highlights what the DST is intended to support / be able to do for its final release (v1.0). Task 3.4 will prioritize implementation and validation of the core features, given the available resources (work effort, available time and the domain specific knowledge that will be available/generated by the project).

Feedback collected from the project partners, based on the activities carried in each LL, will be used as an input into the implementation of the upcoming releases of the tool – this will include continuously updating the private knowledge base as well as testing and validation of the DST regarding its use and generated outputs (for which we have indicated important metrics to consider for doing so). A provisional roadmap in terms of future releases is presented as part of this deliverable.

References

[1] Deliverable D1.3, “COMMECT solution architecture version 1”, Horizon Europe COMMECT project, February 2024.

[2] Deliverable D2.3, “Intent-driven Service Orchestration version 1”, Horizon Europe COMMECT project, February 2024.

[3] *Nickolay Rudnichenko and Vladimir Vychuzhanin and Igor Petrov and Denis Shibaev, “Decision support system for the machine learning methods selection in big data mining”, International Workshop on Computer Modeling and Intelligent Systems 2020.*

[4] Using AI for Intelligent Decision-Making, https://www.mad.co/en/insights/using-ai-for-intelligent-decision-making, December 2023.

[5] Unleashing the Power of AI in Decision Support Systems, https://medium.com/ubuntu-ai/unleashing-the-power-of-ai-in-decision-support-systems-cb95c7594177, December 2023.

[6] Dreambox Learning, https://www.dreambox.com/, January 2024.

[7] Cognii, https://www.cognii.com/, January 2024.

[8] Artificial Intelligence for Disaster Response (AIDR), https://aidr.qcri.org/, January 2024.

[9] Global Database of Events, Language, and Tone (GDELT), https://www.gdeltproject.org/, January 2024.

[10] “Nokia AVA Energy Efficiency Harness AI for more sustainable energy use”, NOKIA white paper.

[11] FarmWise, https://farmwise.io/; January 2024.

[12] Blue River Technology, https://bluerivertechnology.com/ January 2024.

[13] Agrosight, https://golden.com/wiki/Agrosight-XKEZ98Z, January 2024.

[14] Motorleaf: ”The Farming AI Which Helps Boost Modern Indoor Farm”, https://en.reset.org/motorleaf-farming-ai-which-helps-boost-modern-indoor-farms-05222018/, January 2024.

[15] Gamaya, https://gamaya.com/, January 2024.

[16] NCX, https://ncx.com/, January 2024.

[17] Vegetation Management – Predict, https://www.ibm.com/downloads/cas/ZOYKGYWJ, February 2024.

[18] Lu, S., Hamann, H.F. “IBM PAIRS: Scalable Big Geospatial-Temporal Data and Analytics As-a-Service", Handbook of Big Geospatial Data, Springer, 2021.

[19] OpenAI’s GPT-3, https://beta.openai.com/signup/, December 2023.

[20] Github Copilot, https://github.com/features/copilot, December 2023.

[21] Copy.ai, https://www.copy.ai/, December 2023.

[22] ROSS Intelligence, https://www.rossintelligence.com/, December 2023.

[23] DeepL Translator, https://www.deepl.com/en/translator, December 2023.

[24] Saed Rezayi, Zhengliang Liu, Zihao Wu, Chandra Dhakal, Bao Ge, Chen Zhen, Tianming Liu, Sheng Li “AgriBERT: Knowledge-Infused Agricultural Language Models for Matching Food”, Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-22).

[25] IFPRI Blog, “Can we trust AI to generate agricultural extension advisories?”, https://www.ifpri.org/blog/can-we-trust-ai-generate-agricultural-extension-advisories, December 2023.

[26] Tzachor, Asaf; Devare, Medha; Richards, Catherine; Pypers, Pieter; Ghosh, Aniruddha; Koo, Jawoo; Johal, S.; and King, Brian. 2023. “Large language models and agricultural extension services”, Nature Food 4: 941-948.
[27] The Global Surge of Artificial Intelligence Generated Data: How GenAI and Large Language Models (LLMs) Are Reshaping the Global Information Landscape, https://www.linkedin.com/pulse/global-surge-artificial-intelligence-generated-data-how-lima/, LinkedIn, October 2023.
[28] ChatGPT, https://chat.openai.com/, December 2023.
[29] NarrativeAI, https://www.nrtvai.com/, December 2023.
[30] Gemini, https://gemini.google.com/, December 2023.
[31] Perplexity, https://www.perplexity.ai/, December 2023.
[32] What Is Retrieval-Augmented Generation aka RAG, https://blogs.nvidia.com/blog/what-is-retrieval-augmented-generation/, NVIDIA Blogs, November 2023.
[33] P. Lewis et. Al.: Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, arXiv:2005.11401, May 2020.
[34] Langchain library, https://www.langchain.com, December 2023.
[35] OpenAI GPT3.5 model, Models - OpenAI API, https://platform.openai.com/docs/models/gpt-3-5-turbo, December 2023.
[36] Faiss vector database documentation, https://faiss.ai/, December 2023.
[37] AppServices, https://azure.microsoft.com/en-us/products/app-service#overview, December 2023.
[38] Azure OpenAI, https://azure.microsoft.com/en-us/products/ai-services/openai-service, December 2023.
[39] Azure AI search, https://azure.microsoft.com/en-us/products/ai-services/ai-search, December 2023.
[40] Azure storage, https://learn.microsoft.com/en-us/azure/storage/common/storage-introduction, December 2023.
[41] Docker, https://www.docker.com/, December 2023.
[42] Get started in prompt flow - Azure Machine Learning | Microsoft Learn, https://learn.microsoft.com/en-us/azure/machine-learning/prompt-flow/get-started-prompt-flow?view=azureml-api-2, December 2023.
[43] ML studio, https://ml.azure.com/ , December 2023.
[44] Tutorial: Explore Azure OpenAI Service embeddings and document search, https://learn.microsoft.com/en-us/azure/ai-services/openai/tutorials/embeddings , December 2023.
[45] Twilio service, https://www.twilio.com/en-us, December 2023.
[46] React framework, https://react.dev/, December 2023.
[47] Axios, https://axios-http.com/, December 2023.
[48] agroNET - Digital Farming Management, https://agronet.solutions/, October 2023.

Annex I

In this Annex we describe the Decision-making Support Tool (DST), with its visual representation. The screenshots in Figure 9, Figure 10 and Figure 11 capture the prototype of user interface and the relevant (within scope of the COMMECT project) conversation flow, offering an overview of the DST's initial design and functionality. Figure 9 illustrates the interface and it also demonstrates how users can ask questions to the DST and receive input on (the realization or characteristics of) connectivity solutions (Figure 9 and Figure 10). It must be noticed that the DST also indicates the sources that have been leveraged to generate an answer to questions posed – as mentioned, this knowledge base is established through documentation created through COMMECT (deliverables, discussions, meetings) as well as interactions with relevant stakeholders in industries such as agriculture, viticulture and forestry. The knowledge base is still under development. Thus, whenever the DST cannot answer the question, notify the users that it does not have relevant knowledge to provide an accurate answer (see Figure 11).

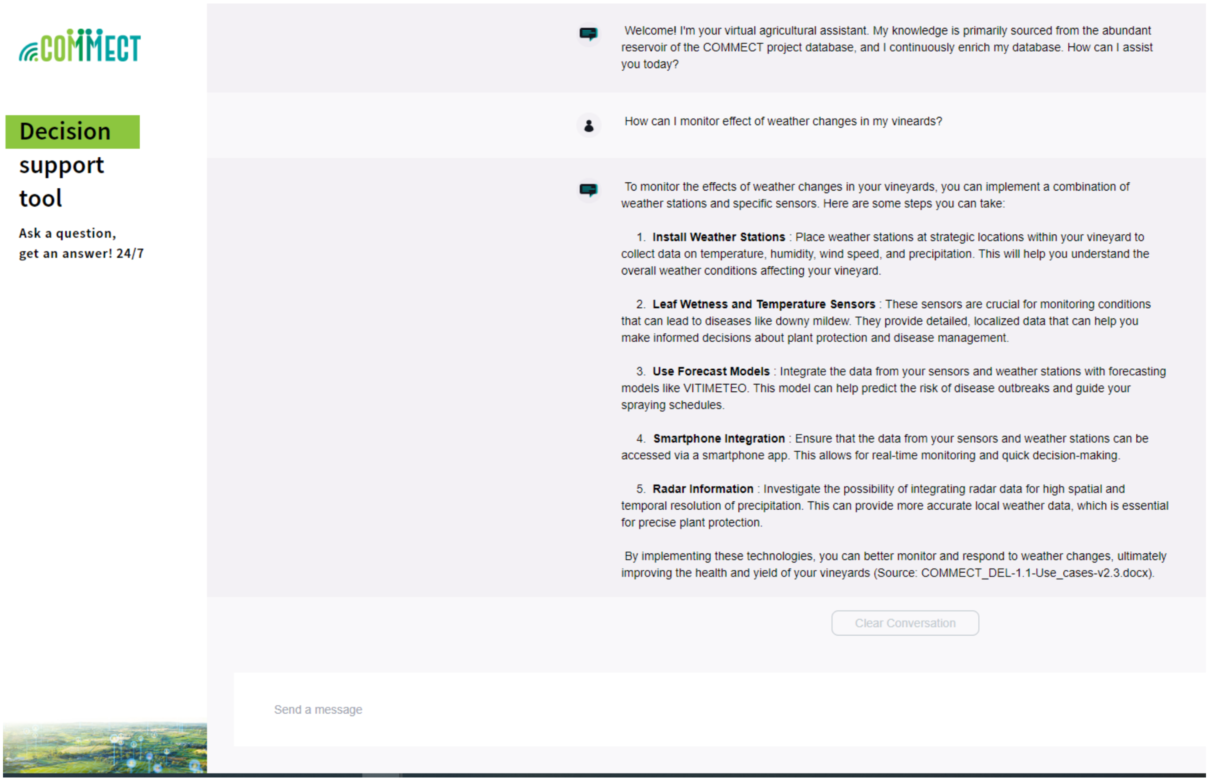


Figure 9 DST web page.

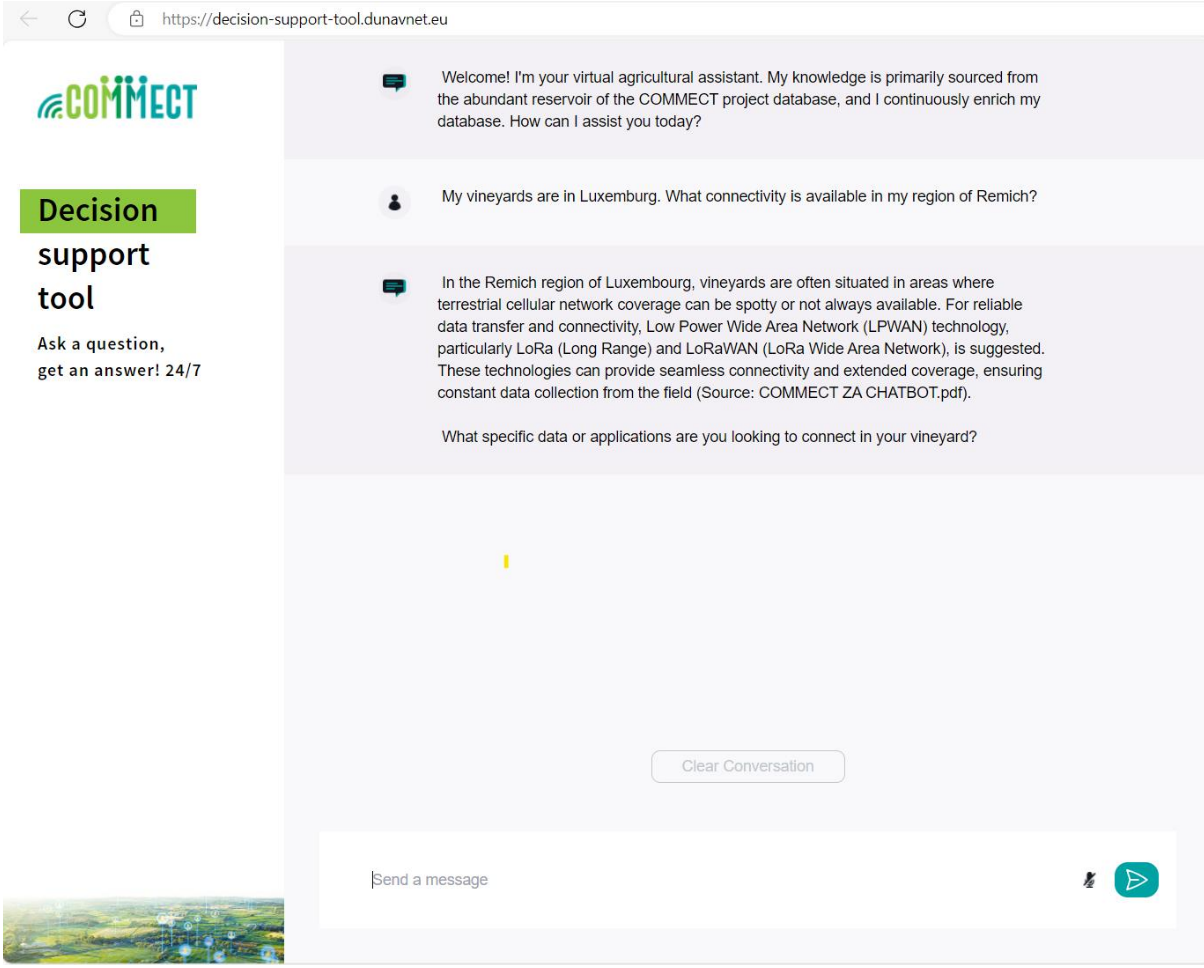


Figure 10 DST – user interface and an example of conversation.

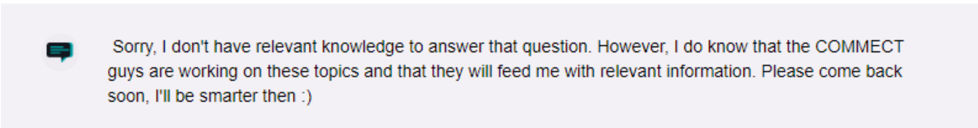


Figure 11 DST – user interface and an example of conversation (2).