DELIVERABLE D5.1 – IDT TOOLKIT MODULES - HUMAN IN THE LOOP - PHASE I

PROJECT ACRONYM:	DS2
PROJECT TITLE:	DataSpace, DataShare 2.0
GA NUMBER NO.	101135967
WEBSITE:	www.dataspace2.eu
DUE DATE OF DELIVERABLE:	30.6.2025
SUBMISSION DATE:	27.6.2025
LEAD BENEFICIARY:	ATC*
LEAD AUTHORS:	Elias Dakos (ATC) (Editor). Module/Additional activities content: Antonios Mpantis (ATC), Shmuel Bar (INTU), Gal Arnon (INTU), Carolyn Cole (VTT), Stefano Modafferi (UOS)
REVIEWERS:	Jens Schimmelpfennig (SWAG), Simon Sukič (MOMS)
TYPE:	OTHER
DISSEMINATION LEVEL:	PUBLIC

* ATC is WP lead, and should have been lead of this document but there was an anomaly in the DOA

DISCLAIMER

STATEMENT OF ORIGINALITY

This Deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

DOCUMENT HISTORY

VERSION	DATE	DESCRIPTION	NAME	ORG
V0.1	21-05-2025	Table of contents	E. Dakos	ATC
V0.2-0.5	06-06-2025	Modules and additional non-software activities added	A. Mpantis S. Bar, G. Arnon C. Cole S. Modafferi	ATC INTU VTT UOS
V0.6	10-06-2025	Final sections and editing	E. Dakos	ATC
V0.6.1	11-06-2025	Comments by Technical Manager	S. Campbell	ICE
V0.7	17-06-2025	Comments by TM Processed	E. Dakos	ATC
V0.8	24-06-2025	Comments by Reviewers	J. Schimmelpfennig S. Sukič	SWAG MOMS
V0.9	25-06-2025	Comments by Reviewers Processed	E. Dakos	ATC
V1.0	25-06-2025	Final Document	E. Dakos	ATC

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This document describes the work performed in DS2 Work Package 5 until M18 of the project. WP5 focuses on incorporating Human-in-the-loop mechanisms into the data lifecycle management process within the DS2 ecosystem. The work carried out is related to the tasks T5.1, T5.2 and T5.3 of the WP and the following objectives:

- O5: To produce a series of modules for the IDT of WP6 which deal with human-in-the loop functionality in terms of language, human-centric collaboration, and conversational support agents:
    - O5.1: Enable multi-cultural and multi-lingual NLP functionality with Cultural and Societal aspects for pervasive use across DS2 ecosystem.
    - O5.2: Support data consumer and data providers to assess the technical, business, operational and organisational capabilities of participating data spaces in the DS2 environment.
    - O5.3: Support data consumer and data providers to configure a commonly approved complex Data Life Cycle (DLC) in the DS2 environment.

Work Package 5 has addressed those objectives through the development of the following software modules:

- Culture Language Module (CLM), that derives from Task 5.1 and covers the ontological and Natural Language Process (NLP) aspects of DS2 and absorbs activity originally foreseen in WP6.
- Discovery, Assessment, Recommend, Configure module (DARC) which is the LLM Chatbot, and it is the consolidation of Tasks 5.2 and 5.3 into a single, unified architecture therefor one software module as well as into a unified implementation plan that provides a seamless stack of functionality.

Furthermore, Work Package 5 includes additional non-module software development activities which are part of Tasks 5.2 and 5.3 for the creation of an interface between DARC (LLM Chatbot) and the Knowledge Base of Sovereign Decision Support System (SDSS) that will serve as a provider of recommendations on security risks and mitigation for the Recommendation and Configuration steps of the DARC pipeline.

Work Package 5 also includes some additional non-software development activities on the development of a sociotechnical evaluation framework for DARC (LLM Chatbot). The framework entails technical evaluation recommendations alongside and informed by user needs and expectations, which are derived from engagement with future users through surveys and workshops. This work ensures that DARC is not only technically robust but also aligned with the diverse expectations and requirements of its intended user base. This specific activity will continue throughout the duration of Work Package 5 and will provide both pre- and post-assessment of the reliability, usability, accessibility and responsiveness rate of DARC (LLM Chatbot).

Together, all the Work Package 5 end products form a human-centric layer within the DS2 IDT Toolkit that enables intuitive, trustworthy interaction with the complex backend infrastructure, ensuring ethical and efficient configuration and deployment of data-sharing processes.

The formal WP deliverables are all software (Type: OTHER) which are not uploadable to EU Sygma, so the software, possibly source code, documentation, detailed progress tracking are PRIMARILY on the DS2 GitHub which will evolve and grow in content over time – see https://ds2-eu.github.io/documentation/. The online documentation includes module overview, architecture, components descriptions, example screen shots, information on installation and use and other matters. Within this document, which accompanies the formal software/information on the DS2 GitHub, an introduction to the WP is provided, followed by a brief description of the different modules developed. Then, there is a summary of the activity and progress for each module including links to the DS2 repository. In addition to this, some WPs have other non-module and/or non-software activities which are also reported where applicable. This is then followed by a description of KPI's, Primary Risks, and Primary Issues for the WP, followed by a conclusion section.

There are few significant risks and issues, and all software is on track or exceeding expectations.

1 INTRODUCTION

The WP5 suite of software delivers essential human-in-the-loop capabilities within the DS2 system. It provides tools that assist users to interact with the system through natural language, assess data space readiness, and guide configuration of complex data lifecycles.

This includes modules for the below which are documented in Section 3 :

- DS2 Culture and Language Module CLM
- DS2 Discovery, Assessment, Recommend and Configure Module DARC

In addition, additional activities are also documented in Section 4:

- Knowledge Base interface for Sovereignty Decision Support System

Some of the narrative texts in section 3 refer to other modules in this WP, or other WPs, so corresponding documents in other WPs may need to be referred to – namely, D3.2, D4.1, D5.1, D6.1. Similarly, modules may refer to additional activities, particularly [DASH, PLATFORM, and DS2CON] which appear later in section 4 of D6.1.

The technical partners involved are INTU (CLM), ATC (DARC) and UOS (Knowledge Base interface for SDSS) whilst VTT is responsible for the evaluation framework and the optimisation of LLM performance by engaging with end users and technical partners. All DS2 users have minor resources in the WP and participated intermittently/as necessary in meetings and/or email exchanges.

The formal WP deliverables are all software (Type: OTHER), so the software, possibly source code, documentation, and detailed progress tracking are primarily on the DS2 GitHub which will evolve and grow in content over time - https://ds2-eu.github.io/documentation/.

1.1 Document Structure

The document structure is as follows:

- Section Error! Reference source not found.: Introduction: Introduces the deliverable – this section
- Section 2: Module Description: High-level reminder of modules connected with the work package
- Section 2.3: Software Progress: Module, by module, progress
- Section Error! Reference source not found.: Additional Activities: Progress of any non-software activities
- Section 5: KPI, Risks, and Primary Issues: Main factors affecting progress and status
- Section 6: Conclusion: Conclusion and next steps

1.2 Glossary and Abbreviations

A definition of common terms related to DS2 as well as a list of abbreviations is available at https://www.dataspace2.eu/results/glossary.

1.3 External Annexes and Supporting Documents

External Documents:

- DS2 D2.2 Requirements, baselines, KPIs, Architecture & Specifications
- DS2 Communication, Dissemination, and Exploitation report (M18)

- Risk identification spreadsheet
- Questionnaire: “Needs and requirements for intelligent assistance in inter-organizational data exchange”
- Survey results: “Needs and requirements for intelligent assistance in inter-organizational data exchange”

2 MODULE DESCRIPTION AND OVERALL STATUS

2.1 Modules

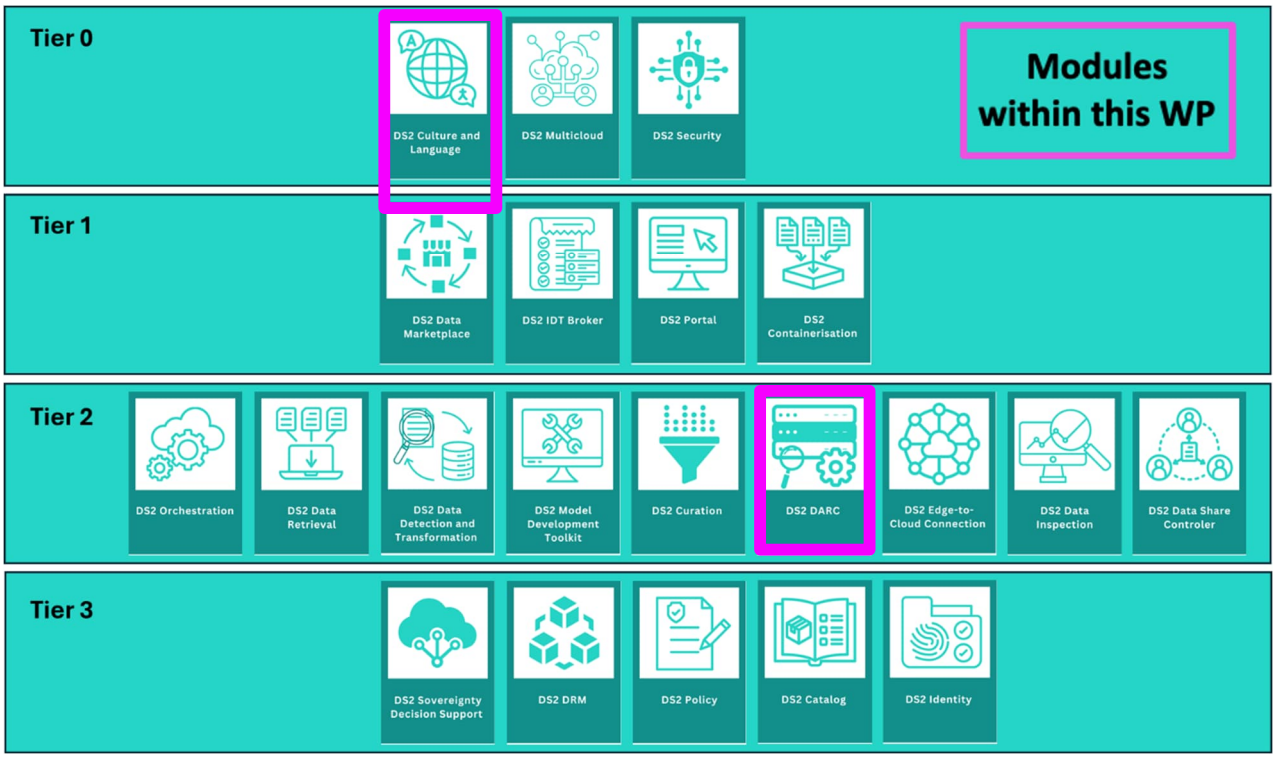


Figure 1: WP5 Modules and Tiers

This WP has the following modules:

Module	Purpose
Tier 0	
CLM 	The Culture and Language Module (CLM) aims to give a data consumer a better understanding of what data offers exist both within their own data space and in other data spaces, which may come from different sectors or countries and in different languages. This will increase the potential for use of a data offer. This is achieved through transformation of human language in shared information and queries into a rich, searchable hierarchical ontological description of the offered data set.

Module	Purpose
Tier 2	
DARC 	To inquire, discover and assess though the conversational UI of an AI-driven agent, the In-Dataspace and Inter-Dataspace data-oriented capabilities and limitations of DS2 Dataspaces as well as DS2 Modules (software prerequisites and/or recommendations) which will compose the "ideal use" scenarios/ paths that will fit the needs of the end-users. Then, to complement this by recommending to end-users, though the conversational UI of an AI-driven agent, the best DS2 Modules for the implementation of the selected "ideal use" scenario/ path, based on the outcomes of the T5.2 Module end-user interaction. Finally, to demonstrate the autoconfiguring, through APIs, a subset of Module based on these recommendations so that they can easily create their DS2 pipeline and start using and sharing data.

The module fit is as follows:

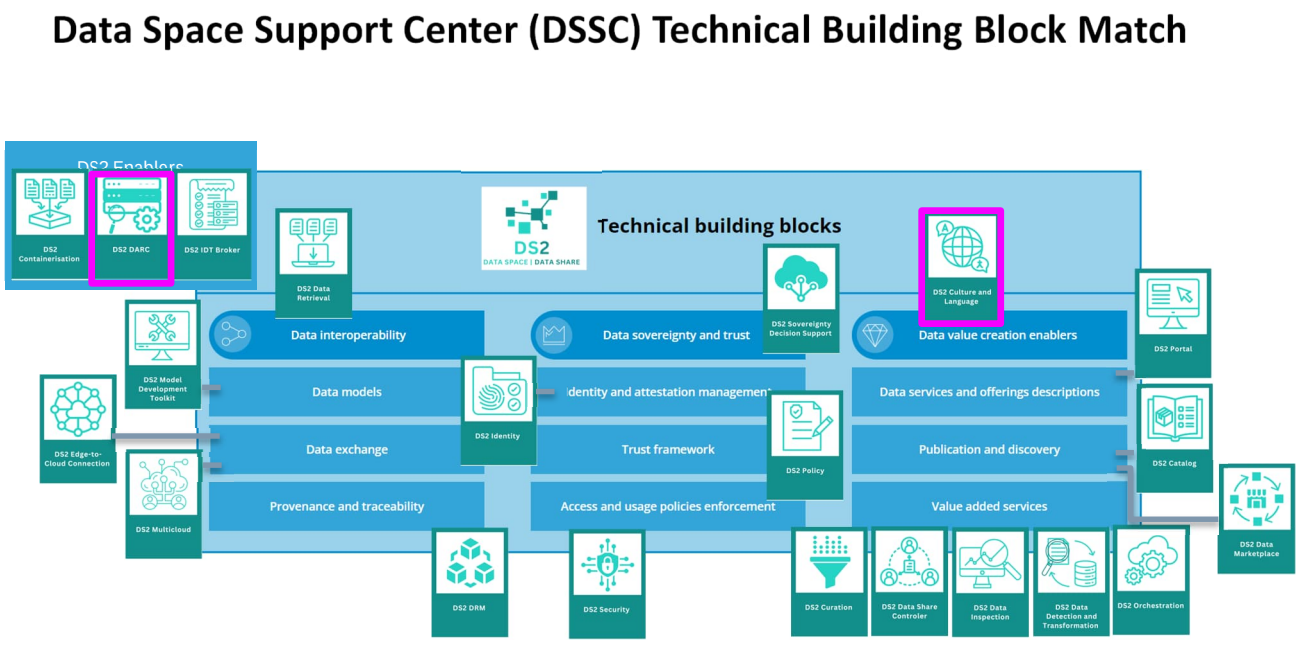


Figure 2: Module fit

2.2 Status

In terms of the percentage tracking / status columns, these are a high-level best-estimates given by the developers based on the detailed per-module tracking activity for features express in Section 3 and according to the methodology identified in Annex A of the D6.1 deliverable.

Module	Task	Partners	License Type	Software Status @ M18	Estimated Completion Month	How to Install "vs 0.1" at M18	How to Configure "vs 0.1"	Marketing Video	PORTAL Helm Char	Marketplace Entry
CLM	T5.1	INTU	Commercial/ Other	50%	M24: 70% - Sufficient for 1st Validation M30: 100%	On GitHub	By education session	M30	M24	M30
DARC	T5.2 T5.3	ATC	Commercial/ Other	50%	M24: 75% - Sufficient for 1st Validation M30: 100%	On GitHub	By education session	M30	M24	M30

2.3 Major Changes/Deviations impacting this WP

- INTU resources from WP6 consolidated in WP5: INTU has resources in WP6, but these have been consolidated in WP5 since the tasks were overlapping, so from a management perspective, and bearing in mind WP6 is overloaded with modules, this was seen as the most appropriate approach.
- ATC software module development results from Task 5.2 and Task 5.3 consolidated in one module (end product): The DOA envisioned this activity divided into two separate tasks, however, upon further technical investigation, it became appropriate to deliver them as part of the same technical architecture and thus merge them.

3 SOFTWARE PROGRESS

The software, documentation and progress for the modules developed is located at the DS2 GitHub repository accessible by the links below. These links give the current module documentation and at the top further links to the software and module progress. At this interim state of the project the modules are still under development, and documentation is limited to essential information. The framework for documentation and progress monitoring is identified in Annex A of WP6 equivalent deliverable "DS2 D6.1 - FEDERATED IDT PLATFORM - PHASE I". Progress is monitored through two-weekly sprints with detailed highlights. In terms of the documentation, it will be improved with How-Tos, API definitions, etc. overtime.

Module	Link
CLM	https://ds2-eu.github.io/documentation/modules/CLM/
DARC	https://ds2-eu.github.io/documentation/modules/DARC/https://ds2-eu.github.io/modules/DARC/

The rest of this major section identifies for each module:

- A reminder of the architectural diagram.
- Summary table of primary feature progress based on the functionalities and status of completion at M18 and which were defined for each module in the D2.2 Architecture annexes. Progress was measured via bi-weekly highlight/sprints/meetings. Highlight 1 represented 2025-02-04 and highlight 9 is the last two weeks of June 2025. Post review, the feature lists will be updated/enhanced. Also indicated in the table (on the right) is an expectation of when each feature will be completed (100%) at milestones M18, 24, 30, and 36 noting that to ensure a suitable validation plan all software should

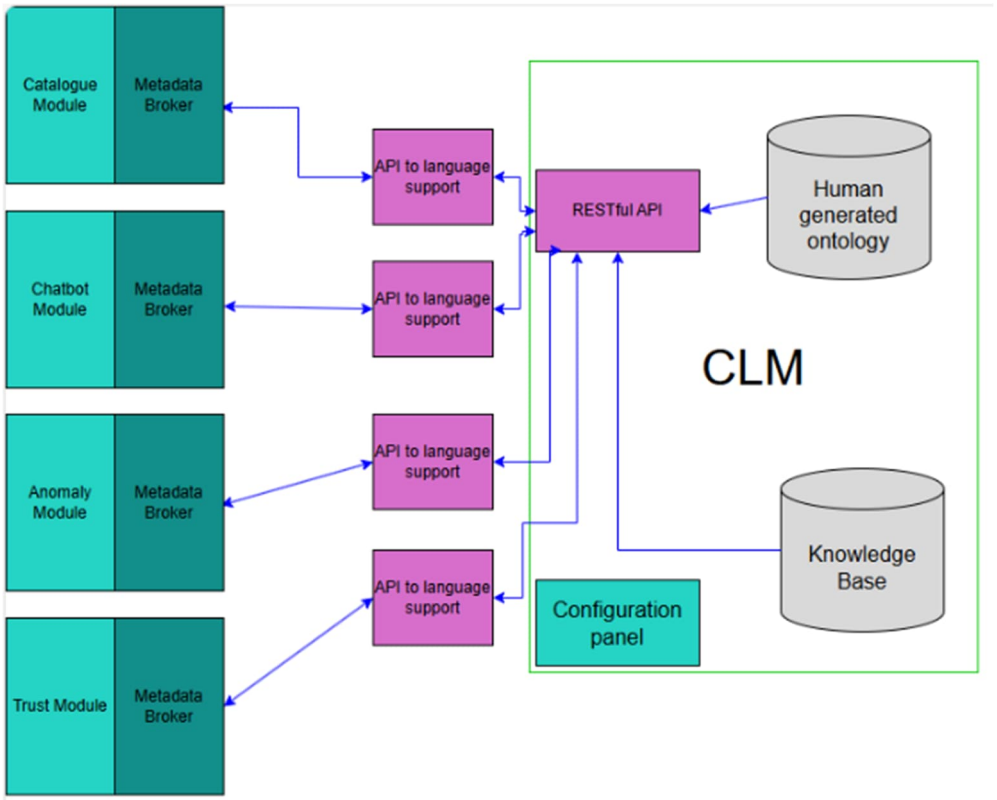
be completed by M30. Since at M18 some feature implementations will still be in-progress, then the percentages may be less than 100%. Details and highlight narratives are at the above links.

- Activity progress of the partner(s) involved in terms of features implemented noting that feature descriptions are at the above links.
- A use case validation plan, with an example of the PORTAL below. "Default Install" (BLUE) means that the module is either used or installed by IDT by default, whereas "Specific to Case" (GREEN) implies its use is specific to a particular use case, for example, UC1.2 – these cases were described in D2.2 and D7.1. A BLUE with no use case mention implies that the module will be needed/used whereas if a case is listed it means it will be needed/used AND that formal validation will take place in that pilot.

	KEY:	**Default install, UC = Validation**	**Specific To Case UC = Validation**
Module ID	**Precision Agriculture**	**Green Deal**	**City Scape**
PORTAL		UC2.2	UC1.1, UC1.2

3.1 Module: CLM - Culture and Language Module

3.1.1 Architecture Diagram



3.1.2 Sample Interface

The CLM module operates without a dedicated User Interface, offering its functionality exclusively through a well-defined API.

3.1.3 Primary Feature Progress in %

Functionality	0	1	2	3	4	5	6	7	8	9	--	M18	M24	M30	M36
Romanian NLP	60	60	60	60	60	60	60	60	60	100	-	100	100	100	100
Romanian Tokenizer	100	100	100	100	100	100	100	100	100	100	-	100	100	100	100
Romanian Morphology	80	80	80	80	80	80	80	80	80	100	-	100	100	100	100
Romanian POS models	60	60	60	60	60	60	60	60	60	100	-	100	100	100	100
Annotation of texts in Romanian for SNER	30	30	30	100	100	100	100	100	100	100	-	100	100	100	100
Integration and Testing (Romanian)	20	20	20	20	20	20	20	20	20	20	-	100	100	100	100
Slovenian NLP	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Slovenian Tokenizer	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Slovenian Morphology	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Slovenian POS models	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Annotation of texts in Slovenian for SNER	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Integration and Testing (Slovenian)	5	5	5	5	5	5	5	5	5	5	-	15	25	100	100
Greek NLP	50	50	50	50	50	50	50	50	50	50	-	60	70	100	100
Greek Tokenizer	25	25	25	25	25	25	25	25	25	25	-	35	45	100	100
Greek Morphology	80	80	80	80	80	80	80	80	80	80	-	90	100	100	100
Greek POS models	15	15	15	15	15	15	15	15	15	15	-	25	35	100	100
Annotation of texts in Greek for SNER	45	45	45	45	45	45	45	45	45	45	-	55	65	100	100
Integration and Testing (Greek)	10	10	10	10	10	10	10	10	10	10	-	20	30	100	100
Ontology Generation per domain	25	25	25	25	25	25	25	25	25	25	-	35	45	100	100
Integration into the Catalogue	10	10	10	10	10	10	10	10	10	10	-	20	30	100	100
Integration into the Chatbot	35	35	35	35	35	35	35	35	35	35	-	45	55	100	100
API	80	80	80	80	80	80	80	80	80	80	-	90	100	100	100
Automated Ontology R&D	25	25	25	25	25	25	25	25	25	25	-	35	45	100	100
Registration Process	25	25	25	25	25	25	25	25	25	25	-	35	45	100	100

3.1.4 Activity

In the initial phases of the project, INTU focused on building the foundational components necessary for seamless system integration. A key achievement was the development and iterative refinement of the INTU API, which now serves as a robust and flexible interface for accessing and processing data across various system components. The API reached 90% completion by the latest project iteration and is central to enabling advanced NLP capabilities, user-driven data access, and domain-specific functionality.

The API improvements allow for seamless interaction between system elements and structured knowledge sources, including ontologies and domain datasets. This interface facilitates data retrieval, supports semantic enrichment, and ensures accurate, context-aware responses for downstream applications. In the process, INTU engaged with other modules to explore ways to use the CLM (the module that INTU is tasked with) in additional modules that were not initially planned for use with the CLM. These include vetting of upload of catalogue descriptions, registration, and use of metadata, after ontologisation, to narrow the search options and to expedite matches.

The development of the Registration Process enables structured onboarding of users and use-case definitions. This system allows for the collection of user information, the specification of domain interests, and the configuration of system behaviour based on user needs. The Registration Process supports dynamic adaptation of the NLP components and system-wide optimization. Significant progress was also made in developing language support for Romanian and Greek.

The NLP pipelines for these languages have been extended with tokenization, morphology analysis, and part-of-speech models, with Romanian components reaching up to 80% completion and Greek at varying levels between 15% and 80%. Efforts for Slovenian language support are ongoing, with foundational components established and further work planned. INTU has engaged with the Use Case partners for building of domain-specific ontologies and lexicons in Romanian and Greek. All of these also contain English lexicons (assuming English as a "lingua franca").

INTU has provided the partner in Cluj-Napoca with an interface for viewing the analysis of its documents. In addition, all the Use Case partners have received templates and training on building a basis for their ontologies. These will be then integrated and loaded in the system. Collaborative efforts across the project have laid a strong foundation for future developments, including further API enhancements, optimized data flows, and advanced NLP model tuning.

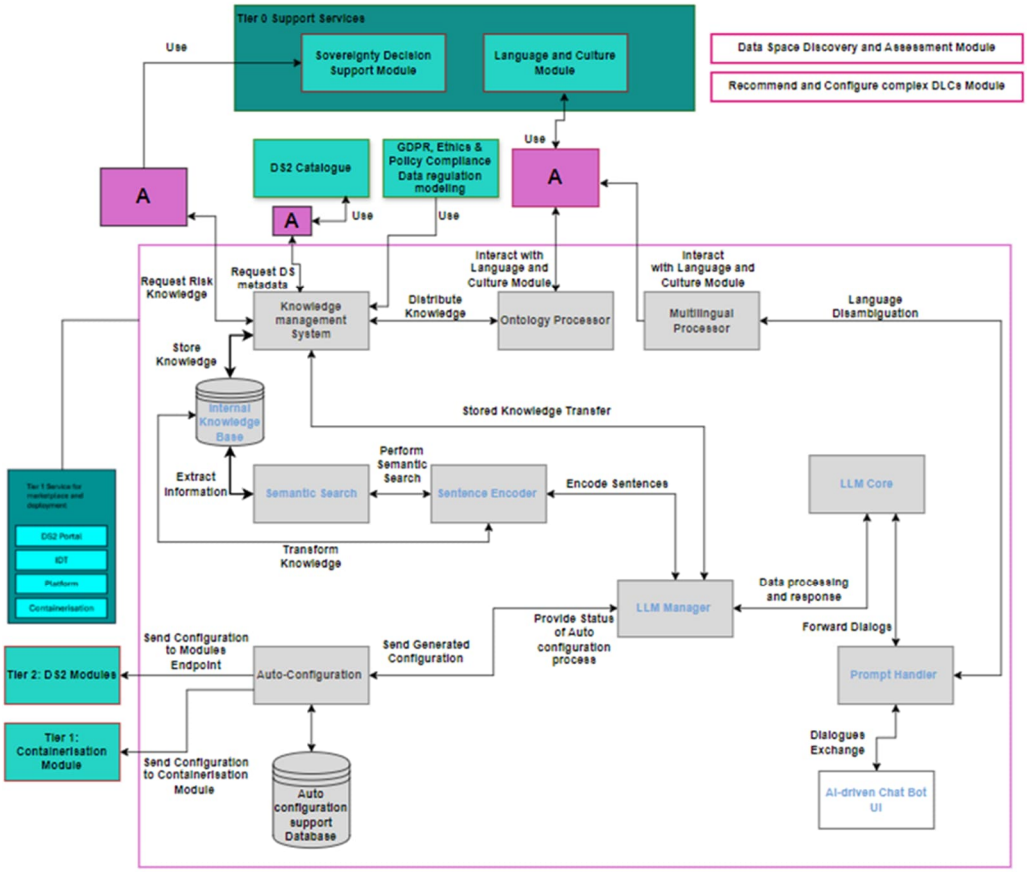
Ongoing work focuses on strengthening the API's capabilities, improving integration with domain-specific knowledge, and supporting diverse application scenarios.

3.1.5 Use Case Validation Plan

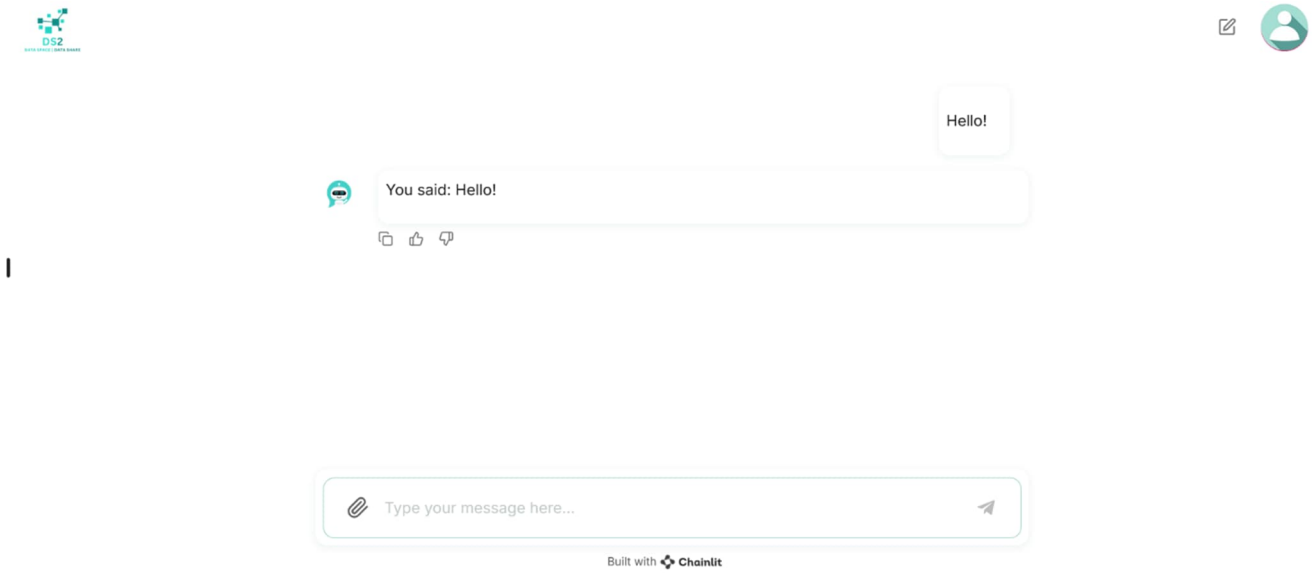
	KEY:	**Default install, UC = Validation**	**Specific To Case UC = Validation**
Module ID	**Precision Agriculture**	**Green Deal**	**City Scape**
CLM		UC2.2	

3.2 Module: DARC – Discovery, Assessment, Recommend and Configure Module

3.2.1 Architecture Diagram



3.2.2 Sample Interface



3.2.3 Primary Feature Progress in %

Functionality	0	1	2	3	4	5	6	7	8	9	--	M18	M24	M30	M36
Knowledge Management System	70	75	75	85	90	95	95	95	95	95	-	95	100	100	100
Internal Knowledge Base	50	50	50	70	80	85	85	85	87	90	-	90	90	100	100
Semantic Search	45	75	80	85	90	95	100	100	100	100	-	100	100	100	100
Sentence Encoder	45	75	80	85	85	90	100	100	100	100	-	100	100	100	100
Prompt Handler	70	70	85	85	90	95	95	95	95	95	-	100	100	100	100
LLM Manager I	50	60	60	60	60	75	85	85	85	90	-	90	100	100	100
LLM Manager II	50	50	50	50	50	50	60	75	85	85	-	90	95	100	100
LLM I	10	15	20	25	30	35	40	55	60	60	-	60	70	100	100
LLM II	10	10	10	10	15	20	35	40	45	45	-	45	65	100	100
LLM III	20	20	20	20	20	20	35	40	45	45	-	45	60	100	100
Ontology Processor	0	5	5	5	10	15	25	25	25	30	-	30	100	100	100
Multilignual Support	0	5	5	5	5	5	15	25	25	30	-	30	100	100	100
Auto-configuration I	0	0	0	0	0	0	0	0	0	5	-	5	50	100	100
Auto-configuration II	0	0	0	0	0	0	0	0	0	5	-	5	50	100	100
Auto-configuration III	0	0	0	0	0	0	0	0	0	5	-	5	50	100	100

3.2.4 Activity

In the initial stages of DARC (LLM Chatbot) backend software development, the foundational schema for the vector store was established and tested, integrating it with both the Internal Knowledge Base and the broader Knowledge Management System, both components being the core of DARC's backend. Early efforts also included uploading and processing descriptive data from the DS2 project (descriptions of modules inside DS2) to validate these foundational components.

Parallel to this, focused research began on enhancing the AI-assistant's performance through Semantic Search. Several new methods were introduced for the Sentence Encoder subcomponent, and experimental testing was carried out to refine results. A similar research-oriented approach was applied to the Prompt Handler, where basic prompting techniques were tested and more advanced strategies explored, such as creating personas based on the user's input query, such as business, technical and generic.

To support broader system integration, initial use-case scenarios were defined using structured questionnaires, with scenario-based optimization in mind. Basic fine-tuning of the LLM core was also conducted, laying the groundwork for more advanced tuning iterations later on.

Advanced Prompt Handling techniques like the "bag of concepts" model, leveraging INTU's CLM API, were introduced to improve semantic understanding and response generation. Ontology support was added through the Ontology Processor, which feeds structured knowledge into the Internal Knowledge Base for more context-aware prompting.

Multiple rounds of fine-tuning further improved LLM outputs, particularly for discovery and assessment tasks. The LLM Manager also saw performance gains through optimizations in routing logic. At the same time, the Knowledge Management System and preprocessing workflows were refined to ensure better handling of both structured and unstructured data.

Regarding the DARC (LLM Chatbot) UI/ front end, an initial version of GUI was designed and deployed to better align with project requirements and user needs, creating a more intuitive and coherent experience across components and following the overall look and feel and branding of DS2 project.

Ongoing work focuses on strengthening diverse application scenarios, work on the configuration and auto-configuration of DS2 modules LLM functionality, offer personalised engagement to the users, improve LLM reliability and consistency, improve GUI for advanced usability and accessibility as well as integration with DS2 PORTAL.

3.2.5 Use Case Validation Plan

	KEY:	**Default install, UC = Validation**	**Specific To Case UC = Validation**
Module ID	**Precision Agriculture**	**Green Deal**	**City Scape**
DARC	UC3.1, UC3.2	UC2.1, UC2.2	UC1.1, UC1.2

4 ADDITIONAL ACTIVITIES

4.1 Knowledge Base interface of Sovereign Decision Support System

The purpose of the Knowledge Base interface of Sovereign Decision Support System (SDSS) is to receive from DARC (LLM Chatbot) input regarding the generic description of a deployment scenario/ configuration of DS2 modules, and then to provide output in the form of recommendations/ advice on potential risks so that users can be conscious of them should they decide to process with the proposed configuration. Therefore, the Knowledge Base interface exposes a REST API, making specific data from SDSS accessible to DARC through this API.

More specifically the Knowledge Base interface is a wrapper for the Spyderisk (UOS Background) Knowledge Base and thus it is meant to provide general advice for deployment scenarios and to interact with DARC.

Furthermore, to calculate threats and risk Spyderisk Knowledge Base requires a system model representation. Considering the stage in which DARC will be used (before deploying DS2 modules) is not feasible to assume that the Chatbot will have all the necessary information. For this reason, only basic information is collected as input and then it is mapped on existing standard system representations that are then used to calculate the threats and risk (Figure 3).

Figure 3 :Steps of SDSS output creation for risks and threats based on DARC input (user input)

4.1.1 Interface with External Systems

//input

```
{
  "deployment_type": [
    "cloud",
    "local"
  ],
  "data_type": {
    "classification": "sensitive",
    "categories": [
      "personal",
      "health"
    ]
  },
  "existing_mitigations": [
    "data_encryption",
    "role_based_access"
  ],
  "user_priorities": [
    "confidentiality",
    "availability",
    "integrity",
    "..."
  ],
```

//output

```
  "threat_ranking": [
    {
      "threat": "Data Breach",
      "priority": 1,
      "risk_factors": {
        "likelihood": "high",
        "impact": "high"
      }
    },
    {
      "threat": "Unauthorized Access",
      "priority": 2,
      "risk_factors": {
        "likelihood": "medium",
        "impact": "high"
      }
    }
  ],
  "recommended_controls": [
    {
      "control":
"multi_factor_authentication",
      "coverage": [
        "unauthorized_access",
        "credential_theft"
      ]
    },
    {
      "control":
"network_segmentation",
      "coverage": [
        "....",
        "...."
```

	] }] }

Figure 4: Sample of SDSS interface with external systems

4.1.2 Activity

In this phase, the work was mainly focussed on designing the solution to create the wrapper interface for the Knowledge Base and then to extract relevant information.

4.2 DARC sociotechnical evaluation framework

VTT made progress on the development of a sociotechnical evaluation framework for the DARC conversational agent. The framework entails technical evaluation recommendations alongside and informed by user needs and expectations, which are derived from engagement with future users through surveys and workshops. This work ensures that DARC is not only technically robust but also aligned with the diverse expectations and requirements of its intended user base.

The work in this reporting period has included a literature review focused on the evaluation of conversational agents, including both technical performance metrics (e.g., for retrieval-augmented generation (RAG) systems) and user experience dimensions (e.g., usability, trust, and satisfaction). This forms the theoretical foundation for our sociotechnical evaluation and this review was also used to directly inform the content choices and design of our first user survey, which was run in Q2 2025.

To capture user needs and requirements, a survey was designed and deployed ("Needs and requirements for intelligent assistance in inter-organizational data exchange"). The survey targeted potential users of the DARC module, including DS2 participants, and representatives from industry, research, public administration, and IT. The survey was circulated to DS2 participants, sister projects, and broader professional networks via LinkedIn. The survey questions captured user expectations regarding functionality, trustworthiness, communication style, and other critical features of the DARC conversational agent. There were 21 responses to the survey and the response data will be taken forward in three main work streams: to inform DARC system development, inform evaluation design, and guide the development of user interaction simulations for testing and evaluation of the agent.

Following the survey, DS2 organized and facilitated a workshop involving DARC developers and DS2 use case members to discuss user needs and expectations in more depth, and how these needs can be met. This session enabled shared understanding of user requirements, validation of survey findings, and co-development of the most important features and functionalities the agent expects to perform.

The next steps of this work will include delivering tailored evaluation recommendations e.g. Human-in-the-Loop interaction preferences based on user background (i.e. technical, business) to the DARC module developers based on the analysis of survey and workshop data. VTT will also develop AI user personas, derived from the survey and workshop findings, to simulate realistic DARC interactions with various user profiles. When the agent is ready for interactive testing, a second survey will be designed and run, asking users to interact with and evaluate the agent based on the needs and requirements identified from the first survey.

The AI user personas will be used to simulate interactions with the DARC (LLM Chatbot) and respond to this second survey based on those interactions. This will enable the project to supplement and bolster the data and insights from the second survey. Finally, the process and findings will be compiled and reported in a scientific publication.

Links:

- Questionnaire: "Needs and requirements for intelligent assistance in inter-organizational data exchange".
- Survey Results: "Needs and requirements for intelligent assistance in inter-organizational data exchange".

5 KPI, RISKS, AND PRIMARY ISSUES

KPI Status:

KPI ID	Description	Status
KPI 5.1	Multi-cultural / multi-lingual user trials undertaken with different nationalities using DEC SUS4 to evaluate usability and acceptance. Respondents from all groups confirm that the prototype enables them to share, find and use data.	On track: Ontology developed, ongoing annotation of Romanian, Greek, Slovenian languages for use cases, Lexicons under development as well.
KPI 5.2	Acceptance of the Conversational Agent by end users with respect to functionality and usability.	On track: First workshop and end user needs and requirements/ pre-evaluation survey conducted, Results incorporated into Conversational Agent's development.
KPI 5.3	Number of users involved in DS2.	Achieved: 21 end users; Validation method: end user needs and requirements/ pre-evaluation survey conducted.
KPI 5.4	Understanding capabilities of complex DLC.	On track: Research and development activities as part of T5.2 and T5.3.
KPI 5.5	Acceptance of the Conversational Agent by end users with respect to functionality and usability.	On track: First workshop and end user needs and requirements/ pre-evaluation survey conducted, Results incorporated into Conversational Agent's development.
KPI 5.6	Number of users involved in DS2.	Achieved: 21 end users; Validation method: end user needs and requirements/ pre-evaluation survey conducted.
KPI 5.7	Understanding capabilities of complex DLCs.	On track: Research and development activities as part of T5.2 and T5.3.

Primary Risks (Those in DS2 Risk Register whose likelihood/impact are both > Low):

- 'No case study on same domain across languages or data types' Related to CLM module. None of the case studies have similar issues in which DS2 can demonstrate that information in one case corresponds to information in another space but in different languages. DS2 created a virtual data space data to demonstrate the cross-linguistic element.
- 'Privacy or other constraints in providing data'. Related to CLM module. Still waiting for training data from two use cases. Not extremely relevant before M18, will be addressed later in the project.
- 'The chatbot has to be built on the basis of a search mechanism and only then implementation of the NLP'. Related to CLM and DARC module interfacing. The CLM now includes its own process for fine-tuning the queries as they arrive from the chatbot.
- 'Failing to understand the human-centric aspects of complex DLCs is very critical for the successful implementation of the Conversational Agent'. Related to DARC module. The project has not reached a state where this risk is relevant.
- 'Many of the queries can be dealt within simple ontological wizards - this may turn out too late' Related to DARC module. The project has not reached a state where this risk is relevant.

- ‘Trying to achieve near-perfect status will introduce biases in the training of the Conversational Agent; not be useful in multiple contexts’ Related to DARC module. Risk is not present at current period. Risk can be mitigated with fine-tune model.
- ‘Limited initial input for the module configuration and auto-configuration functionality for DARC’. Related to DARC module. The project has not reached a state where this risk is relevant.
- ‘Limited technical and domain-specific lexicon/ vocabulary for better organisation/ optimisation of dataspace information for CLM’. Related to CLM module. Being addressed, information being collected from use cases.

Primary Issues:

- No primary issues to report.

6 CONCLUSION

This deliverable presents an overview of the progress achieved in WP5 during the first 18 months of the project. It focuses on the development of software modules that enable multilingual interaction, ontology-driven processing, and intelligent guidance for data space interoperability and DS2 module configuration. Full technical details and source code are maintained in the DS2 GitHub repository, as referenced in Section 3.

The deliverable also includes a summary of the WP5 KPIs and a risk table covering potential issues identified during the initial development phase. All modules are progressing according to schedule, with some benefiting from extended functionality to enhance their impact within the DS2 ecosystem. No critical risks have been encountered, and baseline implementations for most components are already functional.

WP5 follows the Agile methodology adopted across WP3–6, operating on a two-week sprint cycle with regular coordination meetings and reporting milestones. This process is only documented in the annex of WP6 (equivalent to this deliverable) but applies to WP3, 4, 5, and 6.

Looking ahead, the focus will be on enhancing the capabilities of the CLM and DARC modules as well as finalising the interface between SDSS and DARC.

This includes improving the precision of semantic processing, expanding support for domain-specific concepts, and advancing auto-configuration logic to better assist end users within the DS2 framework as well as continuing working on the GUI of DARC. Ongoing collaboration with users and other module owners, combined with iterative testing, will ensure these modules deliver effective and reliable support across the DS2 environment.

The DS2 GitHub repository will continue to serve as the central resource for software updates and supporting documentation.

In summary, WP5 is on track, and its outputs are expected to meet and, in several aspects, exceed its objectives.