Digital twins and Ontology for Robot Assisted Decommissioning Operations

DORADO

Innovation Action (IA)

This project has received funding from the Nuclear Research and Training (HORIZON-EURATOM-2023-NRT-01), project #101165990. The content of this website reflects only the author's view. The European Commission is not responsible for any use that may be made of the information it contains.

Start date: 1 Sep 2024

Duration: 36 months

Website: https://www.dorado-project.eu

Deliverable: D3.1

System Design and Data Plan

Authors: Petri Honkamaa (VTT), Dusan Daniska (WAI), Omar Zahra (IFE), Abdenour Benkrid (IFE), Leny Turmel (Spix), Mahmoud Abdelrahman (SCK CEN), Per-Arne Jørgensen (IFE), Paul Kemppi (VTT), Antti Räty (VTT), Joseph Ridao (KIT), Francois Patrice (ASNR), Dag Fjeld Edvardsen (Catenda)

Project information

Project acronym	DORADO
Project title:	Digital twins and Ontology for Robot assisted Decommissioning Operations
Start date:	1 Sep 2024
Duration	36 months
Website:	www.dorado-project.eu
Project EU officer:	Seifallah BEN HADJ HASSINE

Document information

Document title	System Design and Data Plan
Author(s)	Petri Honkamaa (VTT), Dusan Daniska (WAI), Omar Zahra (IFE), Abdenour Benkrid (IFE), Leny Turmel (Spix), Mahmoud Abdelrahman (SCK CEN), Per-Arne Jørgensen (IFE), Paul Kemppi (VTT), Antti Räty (VTT), Joseph Ridao (KIT), Francois Patrice (ASNR), Dag Fjeld Edvardsen (Catenda)
Number of pages	27
Document type	Report
Work package	WP3
Issued by	2025-05-31
Date of completion	2025-06-27
Dissemination level	Public

Summary

This deliverable describes the system design and data plan which is created in co-operation with the consortium members to suit each of their needs. The architecture consists of core technical components and modules, that will be used when preparing for the demonstrations in WP4.

Developing the modules requires access to a dataset that represents the final data used in the demonstrations. The presented data plan describes the needed data, and how we plan to collect it both in this project and through co-operation with other projects.

Approval

Date	**Approved by**
v1.0, 2025-06-26	Albert Nardi

Executive summary

The present document is the Data Management Plan (DMP) for the DORADO project, funded by the EURATOM Programme under Grant Agreement No 101165990. The DMP aims to ensure smooth implementation and professional data management in the framework of DORADO throughout its duration.

This document describes the system design and data plan for the DORADO project. We represent a generic modular architecture, where various data sources can be processed with a wide range of applications and services/APIs. The end users can access and manage the provided functionality through a variety of user interface options. In addition, the data flow between specific modules is detailed further using several technology examples. We list also the datasets to be collected on this project and the planned timetables. Datasets to be achieved through co-operation with other projects are listed similarly.

Table of Contents

Abbreviations and acronyms

Acronym	Description
ALARA	As Low As Reasonably Achievable
API	Application Programming Interface
BIM	Building Information Modelling
CAD	Computer-aided Design
D&D	Decommissioning and Dismantling
DTXS	Digital Twin Exchange Standard
EU	European Union
ISDC	International Structure for Decommissioning Costing of Nuclear Installations
LiIDAR	Light Detection and Ranging
LLM	Large Language Models
REST	Representational State Transfer
ROS	Robot Operating System
SLAM	Simultaneous Localization and Mapping
SSC	Systems, Structures and Components
STRIDE	Spoofing, Tampering, Repudiation, Information disclosure, Denial of service, Elevation of privilege
TLS	Terrestial Laser Scanning
WP	Work Package
YAML	Human-readable data serialization language

Figures and Tables

Figures

Tables

1 Introduction

1.1 Purpose

The purpose of this deliverable is to define system design and data plan to ensure efficient, secure, and scalable interoperability between technological components developed/used by the consortium members. Given the wide variety of data sources, technologies, and use cases, the system architecture must be modular and flexible to ensure both current and potential future needs.

The development datasets, representing the final data to be used in demonstrations, need to be specified and collected. This includes planning the specifics of each dataset, timetables and how datasets are acquired either within the DORADO project or through co-operation with other related projects.

1.2 Relation to other activities in the project

The focus of WP3 is the design and implementation of the core technical components for the system. The modules needed are determined based on the formats and applications identified in WP2, and they are developed to enable demonstration of the technologies in WP4. For more details, see also chapter 2.5 Handover and support for use case demonstrators.

In addition to implementing the technical components, the identification of necessary data to develop the work will be undertaken. This process will involve close collaboration with other work packages to ensure that the common data plan can be applied consistently across all the relevant WPs. Data collection is also one of the key areas of collaboration with other related projects. See more details about ongoing co-operation efforts in the next chapter.

1.3 Co-operation with other projects

The DORADO consortium aims to have open co-operation with other related projects. So far, two main EURATOM projects have been recognized in this regard: PLEIADES (PLatform based on Emerging and Interoperable Applications for enhanced Decommissioning processes) and XS-Ability. Already finished, the PLEIADES project was carried out between late 2020 and 2023 by a consortium composed by, among other partners, 7 participants in DORADO: VTT, IFE, KIT, WAI, iUS, IRSN, and CATENDA . The DORADO consortium has been discussing, for example, the possibility of using the data sets that were collected when the project was ongoing. XS-Ability is an ongoing project whose consortium is partially composed by former PLEIADES partners and/or current DORADO partners (CEA, VTT, IFE, and SCK-CEN). Unofficial meetings are regularly hold between the projects to discuss the current development status and potential co-operation possibilities.

The DORADO consortium is also allowed to utilize the 3D data from the FiR1 decommissioning project research reactor. It was created prior to reactor decommissioning for VTT's internal purposes. The 3D data of this model will be completed by re-scanning the facility after decommissioning in May 2025.

In addition, a set of existing and potential case studies from SCK CEN are available for integration in DORADO. These include historical datasets and future data collection opportunities at BR1, BR2,

and BR3, as well as a contaminated hot cell facility. Together, they will provide relevant use cases for testing digital twin components such as change detection, dose estimation, predictive planning, and mission execution.

Project	Data	Status
FiR1 decommissioning project	FIR1 model collected prior to decommissioning (point cloud, 3D model)	Collected (to be supplemented with a re-scan in May 2025)
PLEIADES	Halden research reactor model	Collected
SCK CEN (ANUBIS, DORADO, XS-Ability)	BR3: temporary waste storage facility: historical photogrammetry, dose rate survey, inventory data BR1: new case within XS-Ability project BR2: potential new use case involving temporary storage during transportation or maintenance operations Hot Cell: dismantling scenario with existing photogrammetry data	BR3: Collected (data through the ANUBIS project) BR1: Need to be agreed with XS-Ability BR2: Will be collected in Q4-2025 Hot Cell: Point cloud data available, dose rate mapping and spectrometry to be collected
XS-Ability	Robot data together LIDAR and RGBD camera. Also visual SLAM in addition to LIDAR SLAM. Multi-robot data.	Robot data will be collected in June-Sep Multi-robot data will be collected Autumn 2025

Table 1 Data sets from other projects to be used in Dorado

In addition to the previously mentioned projects, contacts with R&D projects CLEANDEM (Cyber physicaL Equipment for unmAnned Nuclear DEcommissioning Measurements) and IAEA DEDICATE (Deployment of Innovative Digital Technologies for Efficient Decommissioning of Nuclear Facilities) have been established. The DORADO consortium is actively working on establishing new co-operation possibilities by participating in conferences and interacting with the research community. For more details, see the deliverable D5.1 Dissemination, exploitation and communication plan.

To complement the data available, simulations of real-world scenarios are necessary to validate the integrated platform's capabilities before physical deployment. The DORADO project will utilize some radiation simulation tools that integrate easily with robotics simulation tools. On this direction, the **University of Manchester's radiation simulator** in Gazebo will be utilized. This simulator calculates expected radiation measurements at a sensor based on the inverse square rule (1/r^2) for intensity drop-off, accounting for attenuation through objects and sensor collimation. It can simulate various environments, including those built from CAD files (e.g., reactor rooms), and allows for the mapping of radiation fields using virtual robots.

Additionally, the **VRdose simulator from IFE** will be employed. VRdose is a real-time software tool that models and characterizes nuclear environments, enabling the planning of activity sequences and optimization of radiation protection. It supports real-time calculation of shielding effects, doses, and contributions from different isotopes, offering 3D radiation visualization.

Both simulators are designed to connect with ROS (Robot Operating System), facilitating easier integration and testing with the platform's robotic and data management components. These two simulation tools will allow for validation of the system's ability to handle diverse radiological scenarios.

2 System design

2.1 System architecture

The DORADO system architecture is an extensible, modular structure based on ontology, clearly specified interfaces and standardized conventions between modules to ensure compatibility and flexibility. It is designed to support various data types/sources, accommodate different applications and integrate various user interface types. Many of the technological components will provide well-documented interfaces to enable easy data flow between components. This makes it also possible to extend or even change individual implementations without disturbing the functionality of the whole system.

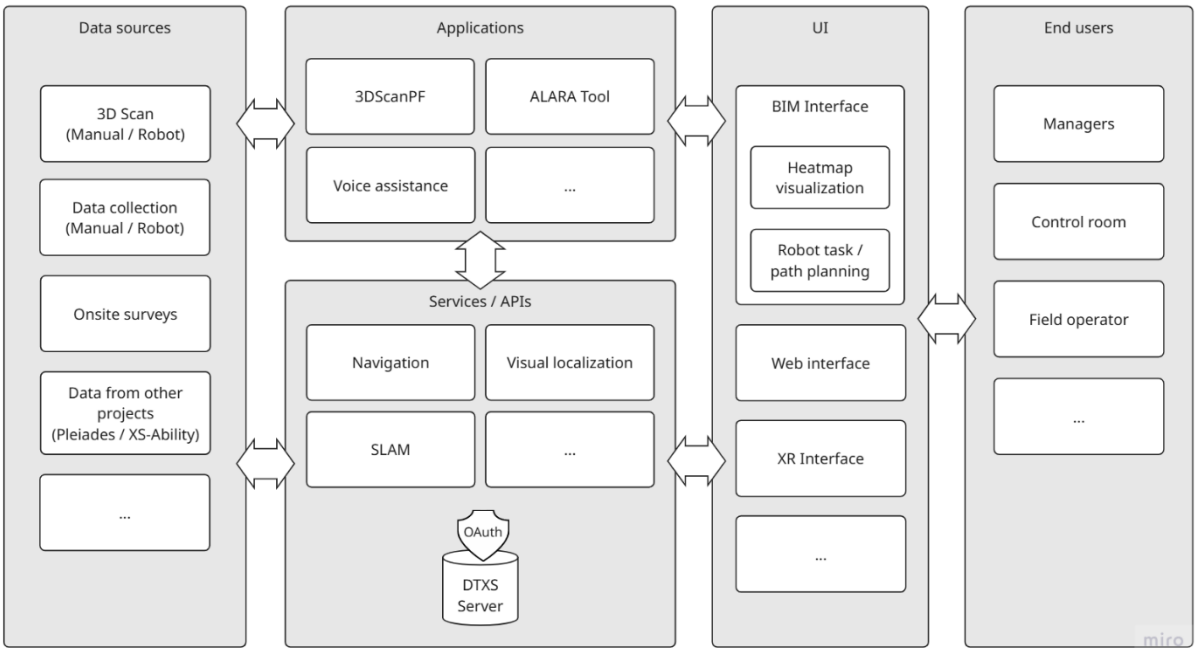

Figure 1. The Dorado system architecture

The Figure 1 shows a generic system architecture, where we divide the technical components (modules) into a few main categories: Data sources, Applications, Services/APIs, and User

interfaces. In addition, the end users can access and manage these through various User Interface possibilities. This diagram gives a basic overview of the whole system. In practice, each separate use case needs only part of the available modules. In the following chapter (2.2 Technologies) we present more detailed data flow diagrams for a selected set of technologies one at a time.

The central piece for this functionality is the server providing functionality to exchange data between all connected modules (DTXS Server). This data exchange is provided via a REST API developed in the DORADO project and called “Digital Twin Exchange Standard”, or “DTXS”. Access to this API is secured by several security principles on applicational layer and on architectural layer. These layers are described in chapter 2.4 Data security. Along with the server, the appropriate client to access the server resources via the DTXS API will be developed.

2.2 Technologies

Here we present some example methods and technologies in more detail, while describing also some real-world use cases where the presented approaches are applicable. This gives a good overview of what kind of use we are foreseeing for the proposed solution.

The diagrams in some of the following subchapters (e.g. Figure 2) are closely related to the system architecture diagram (Figure 1). Both illustration types refer to the same technical components. In the system architecture we show the generic module categories, and we display how the whole system is organized. Instead, in the following chapters we will focus on the data flow between the modules in more detail, while describing only one specific use case at a time.

2.2.1 Sensor data mapping with temporal dimension

In DORADO we want to automate data collection and preprocessing as much as possible. One of the key aspects for this is to align various sensor data and point cloud data to a common coordinate system as automatically as possible.

Here, we describe how visual localization in relation to a reference 3D point cloud can be used to align a mobile device (or a robot) into a global coordinate system. This way any measurements made using those devices can also be stored in the database using directly the correct coordinates.

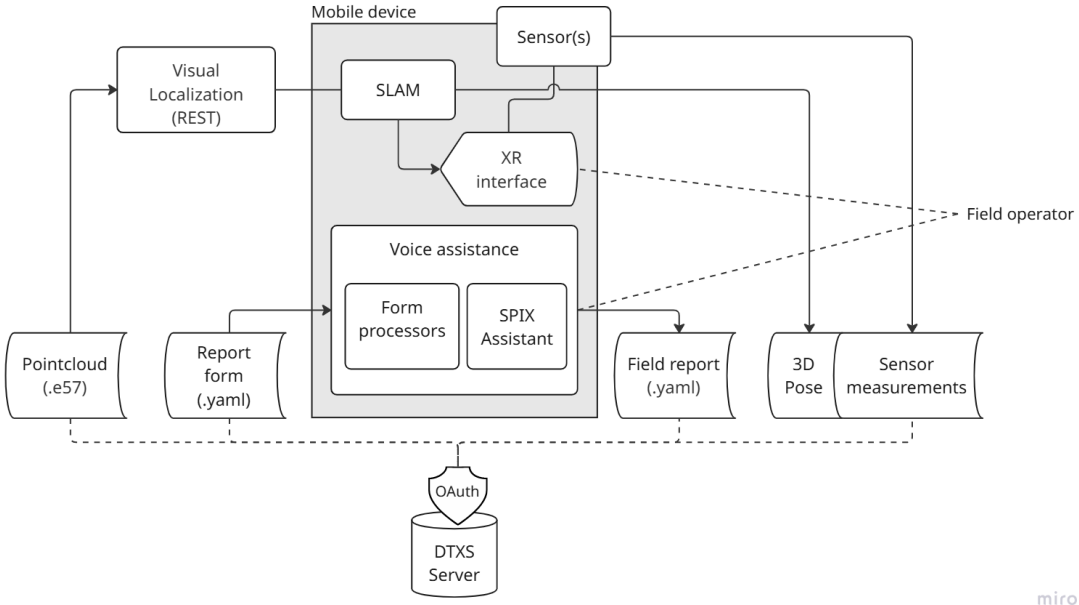

Figure 2. Field inspection using a mobile device will use voice assistance and automated 3D tracking to ensure seamless interaction.

The sensor data collection and mapping using both the handheld mobile device (Figure 2), and the robot (Figure 6), use very similar approaches. The platform itself is tracked in local coordinates with SLAM, using any combination of onboard sensors including cameras, depth cameras, IMUs, or even LiDARs. Most new handheld devices (phones/tablets) have software platforms like ARCore or ARKit implementing the SLAM and providing API for accessing it. Similarly, robot platforms (e.g. ROS) provide various SLAM implementations.

With SLAM, the mobile platform is often tracked using ad hoc local coordinates, which might also be different in every independent session. Therefore, while collecting data, the local coordinate system needs to be properly aligned with the common global coordinate system. This can be done manually, but to ensure up to date data uploads, the process should be automated as much as possible.

For this purpose, an automatic method for aligning local coordinate systems is being implemented, using photos taken from the said mobile platform. These photos are uploaded to a visual localization server along with the estimated camera pose in the local coordinates system. As feedback, the server responds with an estimated transformation between the two coordinate systems: how the

local coordinates need to be transformed and perfectly overlaid with the global coordinates. In theory, one query is enough for making the alignment, but the process should be repeated multiple times to minimize the estimation error. Furthermore, if the local SLAM generates a local 3D point cloud map of the environment, this can be used to finetune the transformation between the coordinate systems even further (e.g. using ICP). However, this would require more data transfer between the client and the server.

The visual localization server of DORADO uses a novel approach where the uploaded photo is matched directly with the LiDAR point cloud of the environment. In decommissioning projects, the more frequently the LiDAR scans are updated, the better the current situation is always reflected. This allows to use these latest LiDAR scans also for positioning. Visual localization is implemented as a service with a REST interface. This approach enables different implementation possibilities. One potential alternative could be to recognize fiducial markers (with known coordinates) in the photos and use those for alignment purposes.

In addition, it is needed to calibrate in advance the relative coordinates and potential timestamp differences between the various components included in the mobile rig. These include, the base coordinate of the platform (handheld device/robot), the relative coordinate of the camera, and the relative coordinates of additional sensors from which the data is collected.

2.2.2 Mission planning

Support for basic mission planning will be created. A mission is a set of tasks that will be carried out by one or more humans or robots. An example of a task can be to do radiological measurements on a specific surface or to scan a specific area in 3D. The mission can be done in sequence or in parallel depending on the resources available and if there are benefits or negatives with parallel execution. For instance, if a set of robots were available and the capabilities exist, they could scan a large area more quickly than a single robot.

There is a need to have some tool support for planning the mission based on the existing requirements. Notice that there is a difference between high-level mission planning (which is the main topic here) and low-level mission planning. The latter is taking the high-level mission planning as input and then using that and other environmental data and robot capabilities to technically plan the mission for the robot(s), including doing pre-mission optimization of travel routes, as well as interacting directly with ROS2 (Robot Operating System v2). High level mission planning is also input for human workers when they will do the scanning and not a robot.

For the user, the flow will be that he/she is presented with the existing schedule. From here, a task can be selected or a new one can be created. One part of this task is the mission plan, which will be created or edited in the mission planning tool. A task in the schedule can have 0 or 1 mission plans.

The user interface will show a BIM-representation of the relevant area of the facility, typically one floor. Most likely this will be presented in 2D (extracted from the 3D-model). What will be shown is an outline of the BIM-objects at a certain height. Objects that will typically be shown are walls, doors, pipes etc. In addition, there will be support for showing information about radiation zones (either from direct input or extracted from a special BIM model). The GUI will allow the user to add start-

points, waypoints, scan points (or -areas), changepoints, endpoints, and position points. When the actual mission is carried out, the results from this high-level planning are inputs to the low-level planning for the robot. The actual situation and optimization in the low-level planning as well as dynamic decisions during the actual mission will decide how this will actually be solved.

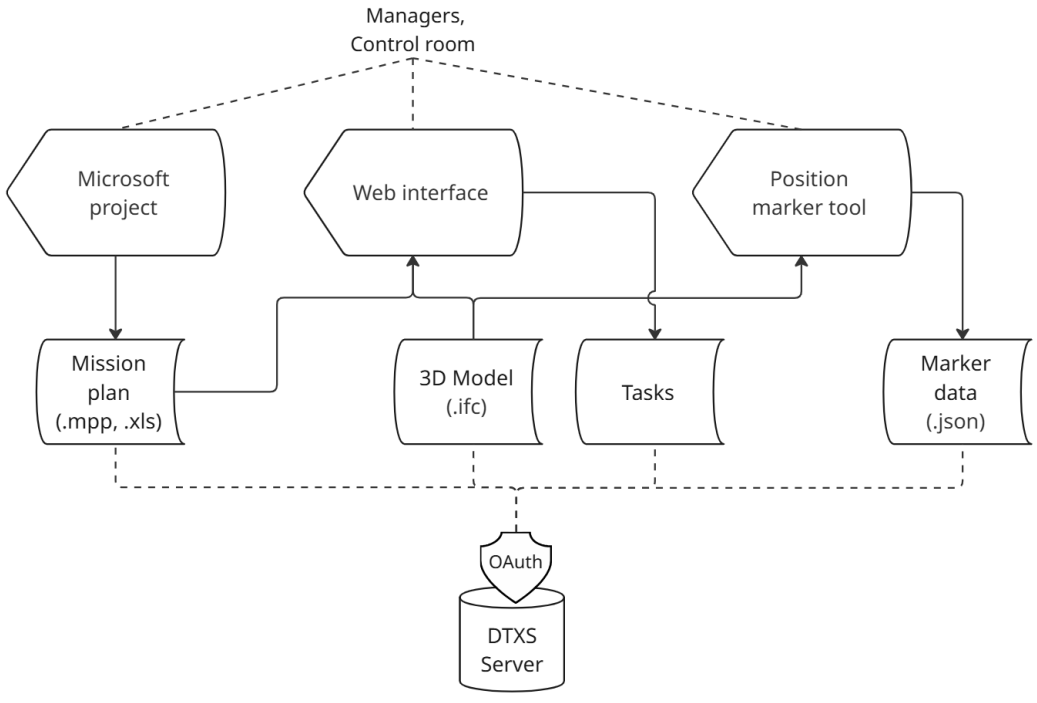

Figure 3. Mission planning

In the execution of the mission, printed position labels with exact position markers will be used to enhance/adjust/train the robot/human positioning. These markers are what is meant by the position points in the previous paragraph. If AR (Augmented Reality) is used, these markers can be used for that. A marker-position tool (that keeps track of printed markers and their positions) will be available either as a separate tool or alternatively as a part of the mission planning tool.

After the user has created a high-level scan mission, this can be exported as Json and/or related to the task in the schedule. This high-level mission can be assigned to a person (a robot operator typically).

It will be investigated if it is practical to create a real time visualization (or playback) of an actual robot mission as it is happening. This can be very useful for continuous learning of optimal mission planning/execution, and for an automatic documentation of the work that has been done in a mission.

Technically, the user interface will most likely be using HTML5, but a GUI built on top of “Pygame” (for a local client on the user’s machine) will also be considered. Data exchange with the backend is important; ideally, the inputs for this tool (BIM, schedule) and outputs (Json) should be stored on the

backend. It will be investigated if the whole tool should be deployed on the backend physically or if it makes more sense to have it as a satellite.

Note that this work has been refocused from the original topic suggested in the application "Environmental comparison against BIM" into "Mission planning". This is because we recognized that the latter topic was seen important in relation to task planning, and the original topic had too much overlap with the technology 2.2.3 Point-cloud and 3D model change detection.

2.2.3 Point-cloud and 3D model change detection

Change detection is very useful for comparing changes that have been made during two or more different periods of time. In the case of nuclear facilities that were built several decades ago, several changes may have been made without having a digital inventory. As they contain industrial technologies in buildings, nuclear facilities are especially complex, so the demolition of the buildings is not the same in all cases. In the AEC (Architecture, Engineering, and Construction) sector, change detection plays a more and more significant role in documenting construction progress. However, in the demolition and dismantling tasks, change detection is relatively new. Change detection in demolition has traditionally been performed with the help of measurements and visual inspections. In recent years, there has been considerable development of measurement devices, which has led to an increase in digital information being produced. Comparison between point clouds, between digital 3D models, or between point clouds and 3D models is possible (e.g. identification of the geometrical deviations, creation of heatmaps according to a certain tolerance established, etc.). However, some of these tasks are difficult to fully automate. Especially difficult is automating the process of converting raw point cloud data into 3D models (geometries) and use them into the BIM-methodology – like IFC extension, which uses semantic labeling of various objects. Therefore, focus will be on semi-automatic approach where in addition to some automated steps, the work is manually assisted using a combination of existing software like Cloud compare, SkethUp Pro, or Autodesk Revit among others. The process will be documented, and a set of best practices to manage the challenges will be collected

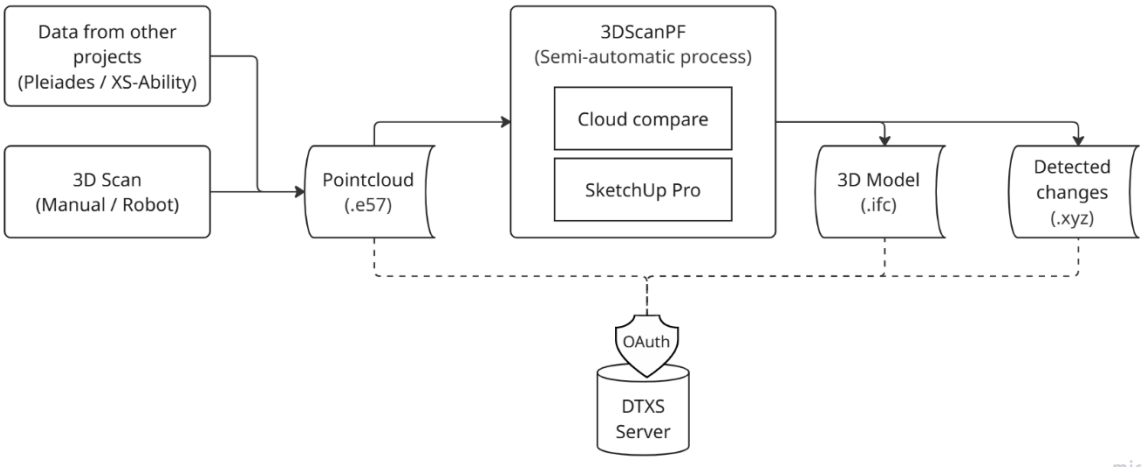

Figure 4. 3D scans can be transferred with a well-defined semi-automatic process into IFC models. Here we can also detect changes compared to previous LIDAR scans.

The input point cloud data will correspond to different periods of time in a decommissioning project. They are bing converted into IFC models and changes between time periods are being detected. The resulting IFC models and change heat maps are being stored in the DTXS Server, and methods to view change heatmaps (and other heatmaps) overlaid on the 3D model will be provided

2.2.4 Dynamic modelling of digital twins for ALARA dose estimation and planning

The DORADO platform incorporates dynamic digital twin modelling to support dose estimation and planning in accordance with the ALARA (As Low As Reasonably Achievable) principle during dismantling activities. This functionality enables continuous representation and simulation of the radiological environment as the facility geometry and source term distribution evolve over time.

The ALARA tool (Figure 6) receives input from multiple sources, including spectrometry measurements (e.g. .spe), dose rate maps (.xyz), and spatial models of the facility (.e57, .ifc). These datasets are processed by the RADTwin module, which is under development. RADTwin combines radiation transport simulations with geometric data to model the current radiation field and incorporates a predictive machine learning component that aims to estimate the evolution of source terms and dose distributions in response to dismantling progress and environmental changes.

The tool supports both the interpolation of measured data and predictive dose mapping under dynamic conditions. This includes forward-modelling of changes due to shielding removal, structural modifications, and updated source configurations based on dismantling sequences. The outputs consist of updated spatial dose rate distributions and structured reports (.pdf, .xls, .json), which can be visualized as 3D heatmaps to support operational planning, briefings, and monitoring.

All data exchanges are managed via the DTXS server, with secure data handling and user authentication provided through OAuth-based protocols. The overall objective is to establish a consistent and extensible framework for evaluating radiation exposure under evolving site conditions and to support informed planning decisions within radiological protection and decommissioning workflows.

This modelling approach is intended to enhance adaptability to change, improve predictive capacity, and contribute to dose optimization throughout the dismantling process.

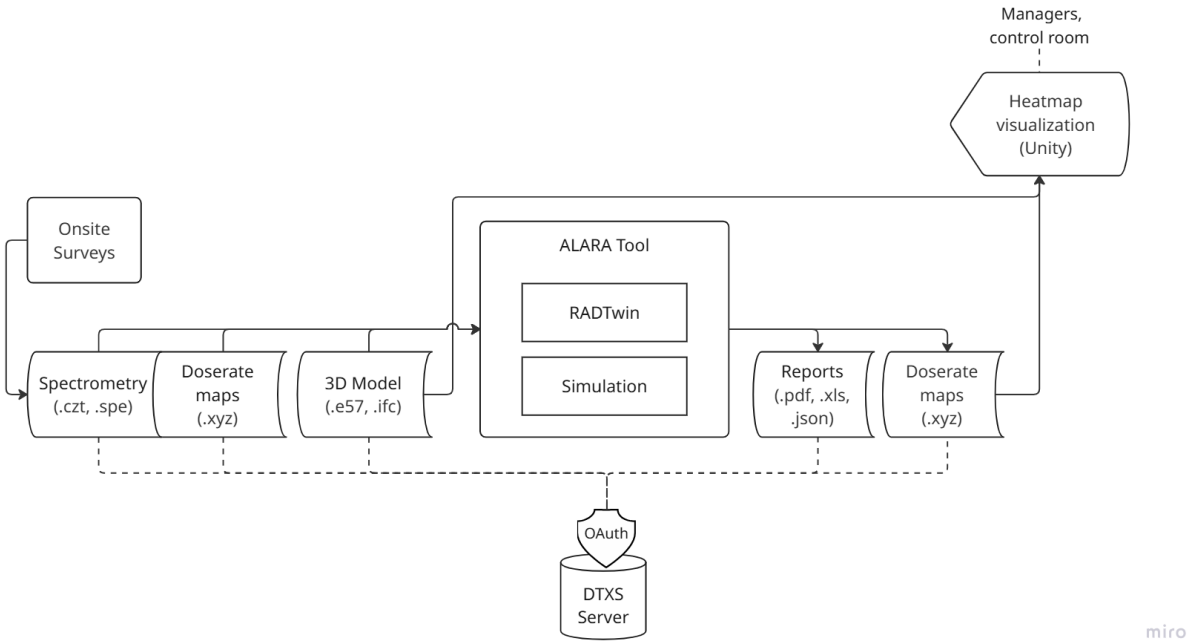

Figure 5. ALARA Tool allows filling in the gaps in surveyed dose rate maps as well as further simulation. The results can be collected into various reports as well as visualized.

2.2.5 Server-based integration with IFC file format and extended data queries

The integration of all technologies used in the project will be achieved by a central server and the REST API provided by this server. The central server will contain features to manage, store and provide data or other resources necessary for other technologies to operate in a collaborative manner. The principles of this architecture are illustrated in the diagram below.

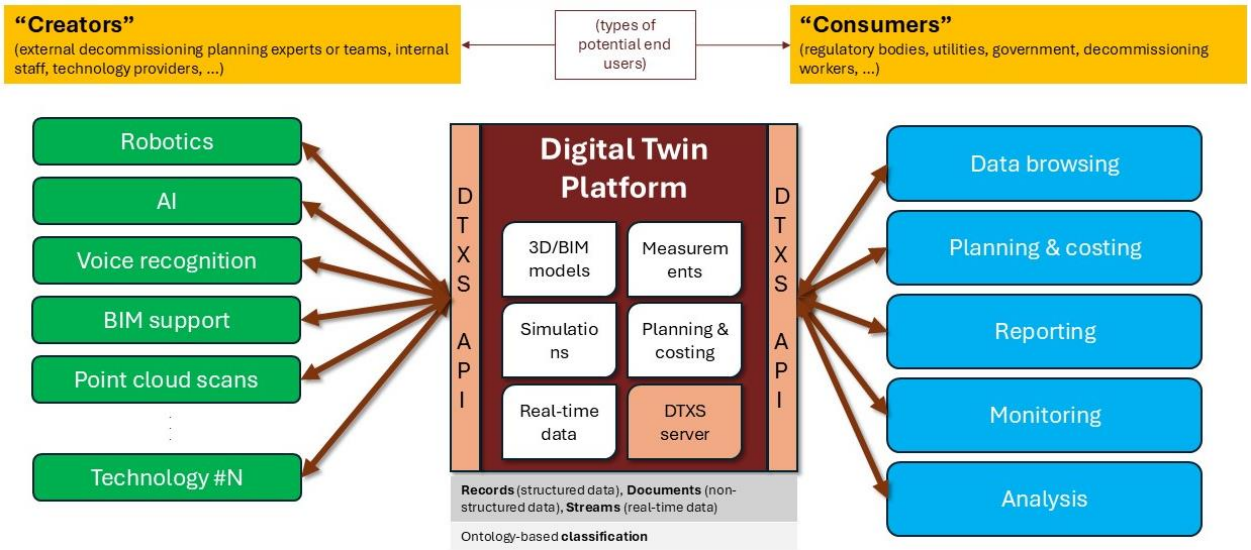

Figure: Principles of the server-based integration

Furthermore, one of the data types used in this project are files (documents) containing the 3D model of the facility as the basis for the visualization or planning activities. These models will be stored as IFC files. This file format contains a large set of structured information related to the 3D geometry and this information can be very useful for the actual decommissioning planning activities. To make this information more easily available to the user, extracting this information from the IFC file will be automatically performed. Extracted data will be stored as so-called 'records' for further querying.

2.2.6 Mission control and robot route optimization

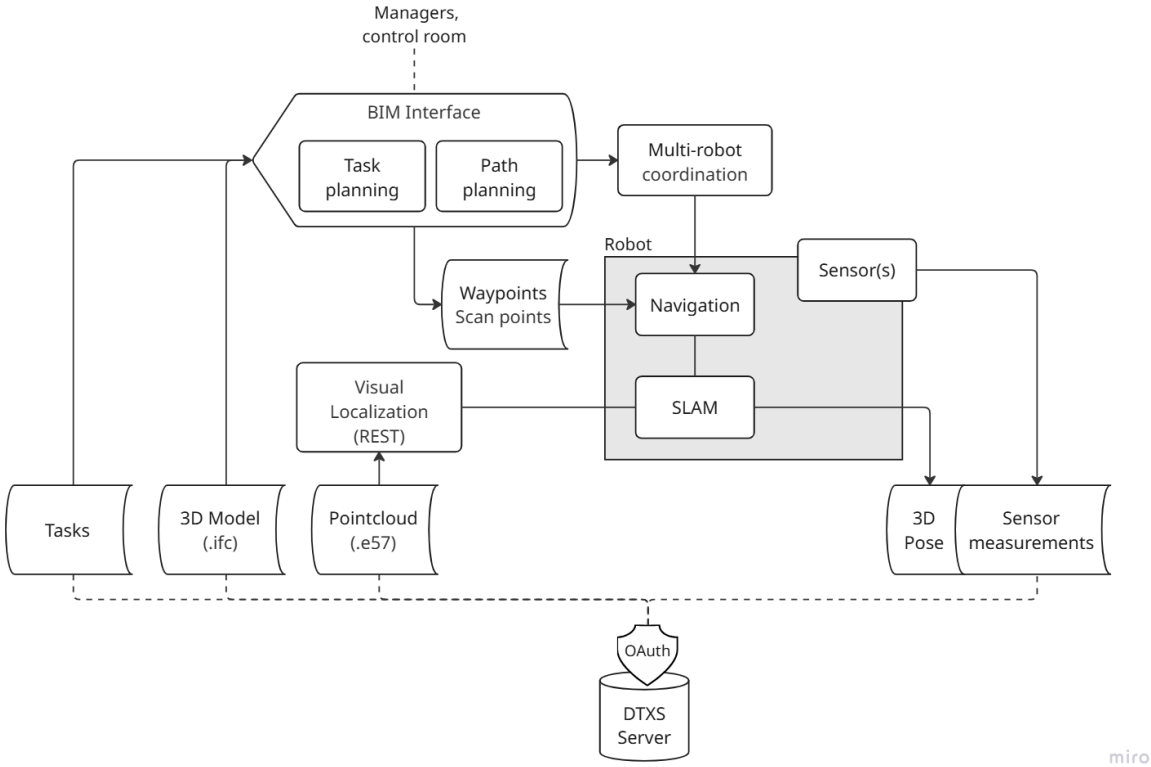

Figure 6. Mission control and robot route optimization

Robot mission planning and execution will be optimized using advanced algorithms for multi-robot coordination and task allocation (e.g., heuristic search algorithms and reinforcement learning). These algorithms will consider factors such as task deadlines, resource availability, and safety constraints to create an optimal sequence of tasks/missions. Robot capabilities, including mobility, sensor types, and payload capacity, significantly impact mission planning. Robots need to be assigned tasks that align with their capabilities to ensure successful mission execution. Moreover, route optimization takes into account the spatial and radiation-related constraints within the nuclear decommissioning environment. Advanced algorithms, such as A* (A-star) and Dijkstra's algorithms, will be employed to calculate the most efficient routes for robots, minimizing travel time, resource consumption and optimizing radiation exposure (based on the defined criteria). Moreover, the digital twin module will enable analyzing and validating the radiation data collected. If needed, additional scan points will be generated to improve the quality of the radiation maps generated.

Given the frequent change in the map of the facility during decommissioning, SLAM solution based multi-robot system will be employed to update the map while navigating through the facility. For precise localization of the robotic agents, the SLAM as well as the techniques introduced in subsection 2.2.1 will be employed. To easily command robots for (semi-) autonomous operation, the BIM-based interface will allow monitoring the navigation goals and updates (as detailed in subsection 2.2.2). While the robots operate through ROS2 middleware, communication will be established with DTXS server to make sure all the gathered data is shared with the other modules whenever needed.

2.2.7 Human to system smart voice assistant interface

The voice assistance is included in the mobile device used by field operators (See Figure 2).

To collect structured data, the assistant uses forms. These forms contain all the fields that the operator must fill out live in the field using voice input.

The structure of the form can vary in complexity, including sequences, repetitions, fields with sections, or optional fields. The possible types of input are:

- Measurements;
- Conformities (yes/no questions);
- Single or multiple choices from lists;
- References (or serial numbers);
- Free-speech comments (unstructured).

The format of these forms is YAML.

The output data of the voice assistant is the completed form in YAML format.

2.2.7.1 Components

DTXS Client

Form templates and filled forms are stored on the DTXS Server.

The role of the DTXS Client is to manage these exchanges.

This component requires a network connection to transfer forms between the voice assistant embedded on mobile devices and the DTXS Server.

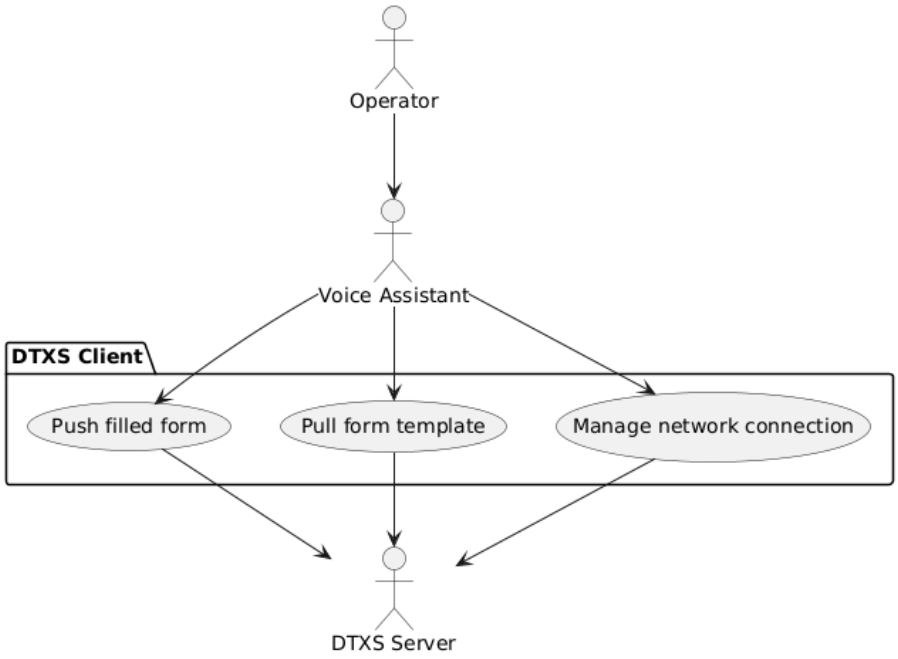

Figure 7. DTXS Client

Forms Processor

The Forms Processor component is responsible for interpreting the YAML form and processing it step-by-step, according to the operator's instructions.

This component works offline, without a network connection.

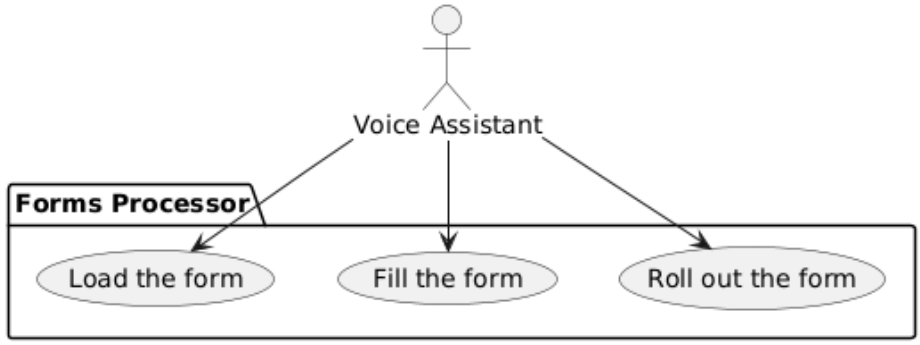

Figure 8. Forms processor

Spix Assistant

The Spix Assistant component is the conversational agent of the system.

It includes an audio management module to capture audio from the microphone and to send audio to the speakers.

It integrates a speech recognition module to convert audio into text.

It integrates a speech synthesis module to convert text into audio.

It also includes an NLU (Natural Language Understanding), a dialogue management, and an NLG (Natural Language Generation) modules. All these modules are used to understand the operator's intentions, to extract information (measurements for example) and to react by giving feedback (GUI, sounds or speech synthesis).

This component works offline, without a network connection.

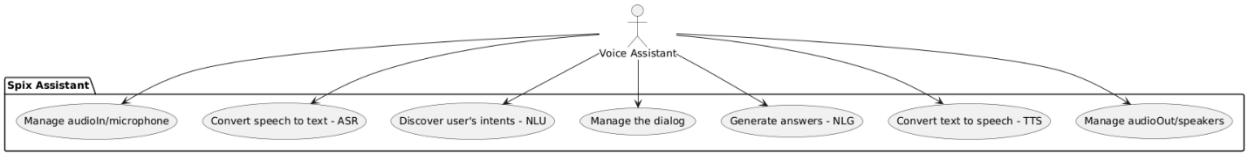

Figure 9. Spix assistant

2.2.8 Standardization using the common ontology

The formalized ontology approach is reflected especially at the borders of different ontologies like voice interface domain, the IFC domain and the decommissioning domain. The ontologies ensure the flexibility for the future without losing backward compatibility. Making software packages to use explicit ontologies in their structures and processes improves the communication across the participants and their different industrial and regulatory backgrounds. This approach provides a good basis for further development.

The use of ontologies is one of the core principles in the software development work of the DORADO project. The schemas to be used in the DTXS Server are based on this, and the server is used on most inter-module communication as described above. Furthermore, the voice-assistance interface (2.2.7 Human to system smart voice assistant interface) relies heavily on well-defined ontology and terminology.

The role of AI in Dorado is detailed in D2.1 State-of-the-art and business value analysis report. In relation to ontology, AI can be a very valuable tool. For example, Large Language Models (LLM) can be used to help make better and smarter interfaces for the user in voice-assistance interfaces. With better voice interaction capabilities, the operator can free his/her hands for the actual work. This improves the efficiency and security of the work. Also, the quality of the data reported will improve.

2.3 Software interfaces

Technology modules and their software will be mainly communicating through a managed central server and its REST API. This API will enable transferring of structured information (so-called ‘records’), files (3D models, point clouds, text documents, JSON files, ...) or potentially real-time streams as well. In certain cases, other APIs might be used as well.

The tables below describe file formats planned to be used, potential use for structured information and REST interfaces used in the software architecture.

File format	Purpose / Utilization
.ifc	3D models
.e57	Point clouds
.xyz	Heatmap of spatial values, e.g. for dose rate maps
.czt, .spe	Spectrometry data

Table: List of file formats planned to be used in DORADO architecture

Ontology class	Purpose / Utilization
Assets.Tangibles.Parts	Definition of objects in the 3D model, e.g. the SSC inventory of the plant.
Actors.Persons	Information about staff performing specified tasks.
Actors.Robots	Information about robots performing specified tasks.
Tasks	Mission planning, definition of tasks (jobs).
Costs.ISDC	Cost estimated for a specific task presented in ISDC format.
Safety.Risks.Register	List of risks to be assessed during risk management.
Assets.Intangibles.Measurements.CartesianMeasurements	Results of performed measurements.

Table: Illustration of ontology classes planned to be used in DORADO architecture

REST API	Purpose / Utilization
DTXS	Main data exchange interface for structured records, documents or real-time streams.
VTT Visual Positioning	Localize camera images relative to colored point cloud scanned from the environment using Terrestrial Laser Scanning (TLS)

Table: REST API interfaces planned to be used in the DORADO architecture

2.4 Data security

Data managed by the DORADO platform is based on various formats and levels of confidentiality. Therefore, proper security layers should be implemented for all stages of manipulation with the data uploading, downloading, storage and backup.

This security should be achieved by:

- Securing the transfer layers by using SSL and HTTPS communication;
- Securing user access by providing different levels of permissions;
- Securing access to the data by user authentication mechanisms, such as OAuth2.0.

In this project, the opensource identity access management tool called Keycloak will be used as the authentication layer. This tool also provides features to configure permissions for each user. Then, the DTXS server developed during the project, will be integrated with the Keycloak service and use it to retrieve the information and permission levels for the authenticated user. The HTTPS transfer layer will be implemented using self-signed SSL certificates.

2.5 Handover and support for use case demonstrators

The development made in WP3 should match well the requirements specified in WP2. Also, one of the key aspects is that the handover of the results for WP4 is made efficiently, so that it best supports the use case demonstrator work. This will involve close collaboration with the consortium members.

To enforce the needed collaboration, WP3 includes a specific task for the handover and case demonstrator support. This task culminates into a workshop between technology developers, demonstrator developers, and appropriate end-users. The data and memos collected from this workshop are released as a public deliverable D3.3 Technology handover and support.

3 Data plan

3.1 Data management

The data management approach, including data storage, data security and ethical aspects are specified in the deliverable D1.2 Data Management plan. Further details about data formats and software interfaces are described earlier in the chapter 2.3 Software interfaces and even more specific information is included in the Annexes.

3.2 Datasets

Developing the modules and technologies described in the chapter 2 System design, requires access to datasets that represents the final data used in the demonstrations. Here, the content of the needed datasets is identified and the plans for collecting the data is described.

The chosen datasets are also related to some of the four selected use cases that are detailed in the deliverable D2.1 State-of-the-art and business value analysis report.

- Use Case 1: Intelligent Mission Planning and BIM Integration;
- Use Case 2: Precise Localization and Change Detection through Multi-View Scan Registration;

- Use Case 3: Ontology-Driven Reporting and AI-Enhanced Hazard Assessment;
- Use Case 4: Autonomous Waste Drum Monitoring.

In the chapters below, the new datasets to be collected in this project are described. However, this description is not complete; some further datasets and updates for the existing datasets may be needed later.

3.2.1 FIR1 dataset

This dataset is mainly related to Use Case 2: Precise Localization and Change Detection through Multi-View Scan Registration. This is the key resource related to the development of the technology described in chapter 2.2.1 Sensor data mapping with temporal dimension, but it will used also by other described technologies. The dataset will be collected at the FIR1 Research Reactor located in Espoo, Finland. The research reactor is now dismantled, and the data should provide a good basis for comparing the changes with the earlier dataset acquired in the PLEIADES project.

The facility is available for the project use only until June/July 2025, so all the needed datasets need to be collected before this deadline.

Data	Description
LiDAR (TLS)	Colored 3D point clouds using tripod laser scanners by Leica and/or Trimble.
LiDAR (Robot)	3D point cloud collected using Spot robot with a separate Ouster lidar with integrated IMU.
Dose rates (Robot)	Dose rates collected with dose rate meter Automess 6150AD attached to the Spot robot. Data is collected using an attached Arduino Mega microcontroller board.
Handheld RGB Camera + Thermal Camera	Handheld video data using a camera rig including RGB camera and Thermal camera. In addition to video files, we need also data for calibrating the used physical camera rig for registration purposes.

Table 2. Data to be collected

Date	Description
27.5.2025	LiDAR (TLS) LiDAR (Robot) Dose rates (Robot)
June 2025	Handheld RGB Camera + Thermal Camera
June/July 2025	(Repeat needed data collections)

Table 3. Planned timetable

3.2.2 BR1/BR2/BR3 dataset

This dataset is a combination of data from previous projects and data collected in collaboration with several ongoing projects. The collaboration partners and aspects are detailed in: Table 1 Data

sets from other projects to be used in Dorado. Below table details the parts that will be collected in this project.

Date	Description
Q4-2025	BR2: potential new use case involving temporary storage during transportation or maintenance operations
TBD	Hot Cell: dismantling scenario with existing photogrammetry data. Dose rate mapping and spectrometry to be collected

Table 4. Planned timetable

4 Annexes

Annex 1: DTXS REST API

This annex describes basic principles of the DTXS REST API interface used to exchange data between modules connected to the central server. Implementation of the features described is an ongoing process and will be active during the whole project.

Full documentation of the API is available online at https://docs.wai.blue/dtxs-digital-twin-data-exchange-standard/api/rest .

Introduction

There are several types of information managed by the DTXS server:

- **Records:** Records are used to manage static structured data/information. This means - all data that can be reasonably stored using JSON structures;
- **Documents:** Documents are used to manage static unstructured data/information like text documents, images, spreadsheets or any other kind of documents. Documents are stored in folders;
- **Streams:** Streams are used to manage real-time data/information. Streams are specific for having timestamps and for being able to be uploaded in a real-time manner.

Security

Each request to the server must be secured by previous authentication. This is implemented using widely recognized OAuth2.0 protocol with a bearer token. The token is provided by the OAuth2.0-compatible server tool. In DORADO project, the opensource Keycloak Identity and Access Management is used.

Ontology and data classification

One of the key DTXS server's features is management of diverse data structures from different technologies and used for different purposes. The server is meant to be the central piece for storing, manipulating and retrieving data used in digital-twin enabled environments.

Thus, a proper classification of all data stored on the server is critical. For this purpose, the principles of ontology semantics have been used to develop a list of classes and their appropriate data structures. Then, each record, document or stream is clearly classified. For example, record storing information about inventory information would get classified under 'Assets.Tangibles' class. Additionally, the flexibility of the server implementation allows utilization of any ontology for data classification, the server is not fixed only for digital twins for nuclear decommissioning.

Endpoints

Server's endpoints are grouped into 5 main categories: general purpose endpoints, endpoints for managing databases, endpoints for managing records, endpoints for managing documents and endpoints for managing streams. The following tables list and describe these endpoints.

General purpose endpoints:

Endpoint	Description
GET /server-info	Get information about DTXS server.
GET /ontologies	Get a list of used ontologies.
GET /classes	Get a list of available classes.

Endpoints for managing databases:

Endpoint	Description
POST /database/{$databaseName}	Create an empty database.
DELETE /database/{$databaseName}	Delete a database.
GET /databases	Get list of available databases.

Endpoints for managing records:

Endpoint	Description
POST /database/{$databaseName}/record	Create a record.
PUT /database/{$databaseName}/record/{$recordUid}	Update existing record.
DELETE /database/{$databaseName}/record/{$recordUid}	Delete a record.
GET /database/{$databaseName}/record/{$recordUid}	Get a record.
GET /database/{$databaseName}/record/{$recordUid}/history	Get history of record.
POST /database/{$databaseName}/records	Get a list of records based on a search query.

Endpoints for managing documents:

Endpoint	Description
POST /database/{$databaseName}/folder	Create a folder.
GET /database/{$databaseName}/folder/{$folderUid}	Get contents of a folder.
DELETE /database/{$databaseName}/folder/{$folderUid}	Delete a folder.
POST /database/{$databaseName}/folder/{$folderUid}/document	Upload new document.
PUT /database/{$databaseName}/folder/{$folderUid}/document/{$documentUid}	Update existing document.
GET /database/{$databaseName}/folder/{$folderUid}/document/{$documentUid}	Get info about document.
GET /database/{$databaseName}/folder/{$folderUid}/document/{$documentUid}/download	Download a document.
DELETE /database/{$databaseName}/folder/{$folderUid}/document/{$documentUid}	Delete a document.

POST /database/{$databaseName}/documents	List all documents in database.

Endpoints for managing real-time streams:

Endpoint	Description
POST /database/${databaseName}/stream/open	Open a stream.
POST /database/${databaseName}/stream/close	Close stream.

Annex 2: Visual Positioning REST API

This annex describes basic principles of the Visual Positioning REST API interface used to estimate location and orientation of a camera image. The input consists of the image and related metadata containing intrinsic information (focal length, principal point, and optional lens distortion coefficients). The reference data is a colored point cloud created with Terrestrial Laser Scanner (TLS).

Security

Each request to the server must be secured by previous authentication. This is implemented using widely recognized OAuth2.0 protocol with a bearer token. The token is provided by the OAuth2.0-compatible server tool. In DORADO project, the opensource Keycloak Identity and Access Management is used.

Alternatively, if the service is used only in LAN without external connections, the requests can be done without authentication.

Endpoints

Currently the server has only one endpoint listed below.

Endpoints:

Endpoint	Description
POST /cameraPose	Estimate camera pose.