



BEEYONDERS

Testing Plan

D2.5: Testing Plan **WP2, T2.5**

Authors: Jose Carlos Jimenez Fernandez (TECNALIA)

Contributors: ACC, All Tech. Developers (ITA, FAD, STA, IIT, CST, HIC, DTU, VTT, CAI, COB)



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101058548.

Technical references

PROJECT ACRONYM	BEEYONDERS
PROJECT TITLE	Breakthrough European tEchnologies Yielding cOnstruction sOvereignty, Diversity & Efficiency of ResourceS
PROJECT COORDINATOR	Antonio Alonso Cepeda antonio.alonso.cepeda@acciona.com
PROJECT DURATION	June 2022 – November 2025 (42 months)

DELIVERABLE NO.	D2.5
DISSEMINATION LEVEL*	PU
WORK PACKAGE	WP2 – Specifications & Testing
TASK	T2.5 – Testing Plan
LEAD BENEFICIARY	TECNALIA (TEC)
CONTRIBUTING BENEFICIARY/IES	All Tech. Developers (ITA, FAD, STA, IIT, CST, HIC, DTU, VTT, CAI, COB)
DUE DATE OF DELIVERABLE	31 May 2024
ACTUAL SUBMISSION DATE	31 May 2024

* PU = PUBLIC

PP = RESTRICTED TO OTHER PROGRAMME PARTICIPANTS (INCLUDING THE COMMISSION SERVICES)

RE = RESTRICTED TO A GROUP SPECIFIED BY THE CONSORTIUM (INCLUDING THE COMMISSION SERVICES)

CO = CONFIDENTIAL, ONLY FOR MEMBERS OF THE CONSORTIUM (INCLUDING THE COMMISSION SERVICES)

V	DATE	BENEFICIARY	DESCRIPTION
0.1	26/06/2023	Jose Carlos Jimenez (TEC)	Initial draft version with inputs from all collaborators (M15)
0.2	30/08/2023	Antonio Alonso (ACC)	Draft version review
0.3	30/08/2023	Fabrizio Federici (AISCAT)	Draft version review
1.0	31/08/2023	Jose Carlos Jimenez (TEC)	Final version (M15)
1.1	30/04/2024	Christian Di Natale (IIT)	Review for exoskeleton
1.2	16/05/2024	Jose Carlos Jimenez (TEC)	First draft version released with inputs from all collaborators (M24)
1.3	23/05/2024	Cristina Naso (STAM)	Review for smart wearables
1.4	23/05/2024	Qing Zhang (HIC)	Review for additive manufacturing
1.5	27/05/2024	Qing Zhang (HIC)	Review for additive manufacturing
1.6	29/05/2024	Natalia Lasarte (TEC)	Review of the draft version
1.7	31/05/2024	Goretti Arias (ACC)	Review of the draft version
2.0	31/05/2024	Antonio Alonso (ACC)	Review and final version released (M24)



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101058548.

Table of contents

1. Executive Summary.....	6
2. Introduction	7
3. Research & Development objectives	10
4. Methodology Approach for testing	20
5. Testing Procedure for each technology	22
Autonomous Ground Vehicles	23
Autonomous Wheel Loader	23
Autonomous Maintenance Vehicle	38
Autonomous Aerial Vehicles	50
Additive Manufacturing.....	65
Exoskeleton.....	76
Smart Wearables	83
Digital Twin	102
Human Robot Collaboration.....	116
6. Data Analysis	125
7. Conclusions	126
8. References.....	127

List of Tables

Table 1: List of milestones (extracted from Table 3.1d of proposal document)	10
Table 2: R&D objectives for BEEYONDERS technologies	11

List of Figures

Figure 1. Technology readiness levels	7
Figure 2. Design Science scheme.....	8
Figure 3. User-Centered Design scheme	8



[D2.5_Testing Plan_v2.0]	4
Figure 4. Drone applications in surveying tasks	13
Figure 5. Agent based model simulation for railways worksite	18
Figure 6. Testing Plan methodology scheme.....	21
Figure 7. Wheel loader model: HIDROMEK HMK 635 WL. Extracted from D4.1.	23

Acronym list

	DESCRIPTION
3DPC	3D Printing concrete
API	Application Programming Interface
AWL	Autonomous Wheel Loader
BIM	Building Information Modeling
BT	Breakthrouh technologies
CAD	Computer-Aided Design
DB	Database
DT	Digital Twin
EMS	Enviromental Monitoring Station
EXO	Exoskeleton
ECU	Engine control unit
FEM	Finite Element Model
GPS	Geopositioning system
GNSS	Global Navigation Satelite System
GDPR	General Data Protection Regulation
HD	High definition
HRC	Human Robot Collaboration
IMU	Inertial Measurement Unit
KPI	Key Performance Indicator
LiDAR	Light Detection And Ranging
LCA	Life Cycle Analysis
LCC	Life Cycle Costing
MQTT	Message Queuing Telemetry Transport



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101058548.

NIOSH	National Institute for Occupational Safety and Health
NOx	Nitrogen Oxides
OWL	Web Ontology Language
PM	Particulate Matter
PPE	Personal protective equipment
ROS	Robot Operating System
RPM	Revolutions Per Minute
RTK	Real Time Kinematic
RSSI	Received signal strength indicator
RGB	Red green blue
LTE	Long term evolution
SQL	Structured Query Language
SWD	Smart Wearable
TRL	Technology readiness level
UI	User Interface
UAV	Unmanned aerial vehicle
VOCs	Volatile Organic Compounds



1. Executive Summary

BEEYONDERS will develop key specific Breakthrough Technologies (BT) in the following domains: Autonomous vehicles and collaborative robotics, additive manufacturing, workers' safety & wellbeing monitoring devices, digital twin management platform & Artificial Intelligence. These BT will be demonstrated and validated in 6 use-case diverse construction sites in 4 different European countries. The participation of construction site workers will demonstrate the safety and adoptability of these BT.

This document, D2.5 Testing Plan (second version), is a deliverable of the BEEYONDERS project, which is funded by the European Union's Horizon Europe Programme (2021 – 2027) under Grant Agreement No 101058548. It gathers the work performed within the Task T2.5. Testing plan of the project, led by TECNALIA.

The deliverable: "D2.5. Testing Plan" has 2 versions: in M15 and M24. This deliverable is the final version from the initial Testing Plan was released in M15. This deliverable comprises the development of the methodology for the technologies to be followed, to assure that the Technology Readiness Level (TRL) medium TRL5 is well established prior to the TRL6 validation in WP7.

This methodology will be based in designing and assessment of trial tests per each Beeyonder's technology, in a laboratory environment, and validating the Digital Twin data integration at the end of this task.

Section 1 is this executive summary where the deliverable is explained. Section 2 is the introduction, and it refers to provide the context and need of this deliverable in terms of Beeyonder's project. Section 3 presents the R&D objectives in BEEYONDERS, and how these are related with defined milestones in the project. Section 4 presents the methodology approach has been designed for this testing plan, being the detailed testing execution to be followed by each technology in section 5. This methodology is based in designing and assessment of trial tests per each BEEYONDERS's technology, in a laboratory environment, and validating the Digital Twin data integration at the end of this task. Finally, section 6 presents the approach about how to interpret the results in terms of assuring that the correct developments for the technologies are ready to be tested at the use cases of BEEYONDERS.



2. Introduction

This deliverable D2.5 – Testing Plan has as main objective to be used as a technical guide to reach the TRLs objective of BEEYONDERS' technologies. The Technology Readiness Level (TRL) scale was introduced into the EU funded projects in 2014 as part of the Horizon 2020 framework program. The TRL scale was originally defined by NASA as a mean for measuring or indicating the maturity of a given technology, from a paper sketch to its entry into the market¹.

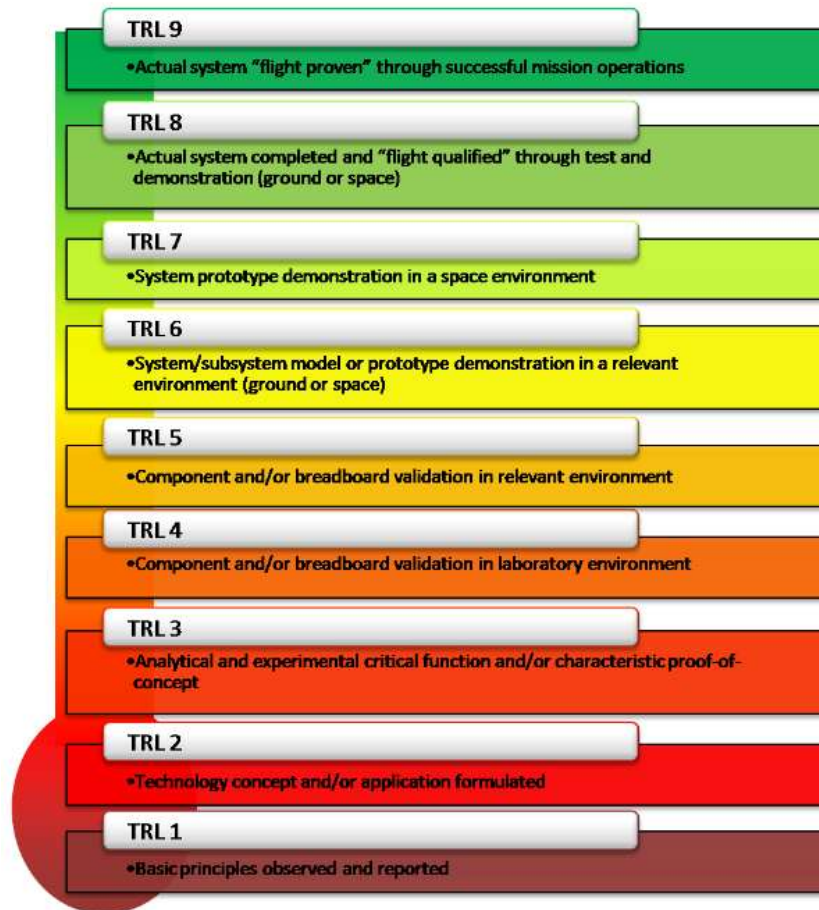


Figure 1. Technology readiness levels²

The task T2.5 – Testing Plan covers all BEEYONDERS breakthrough technologies (BT) that will evolve from TRL 4 to TRL 6 and adapted to their use in the different construction scenarios proposed in the project. At TRL 4, the validation of the technology is performed at the laboratory level. This includes the components, the Digital Twin as a common environment platform and their conceptual/lab scale integration as a whole system; then the BTs are validated in the context of a full integration, and their performance as a system, so at this point a laboratory prototype is available.

TRL 5 is a continuation of TRL 4, but the testing environment become as closer as possible to a realistic one, although still the testing environment is under a control mode. Reaching to this point, one can conclude that the new technology is feasible from a technological point of view.

¹ <https://horizoneuropencpportal.eu/sites/default/files/2022-12/trl-assessment-tool-guide-final.pdf>

² <https://www.nasa.gov/directorates/somd/space-communications-navigation-program/technology-readiness-levels/>



Finally, the BTs (components and common environment) and the system will be demonstrated in the relevant environment of real scale use cases reaching TRL 6 during the proposed use cases in BEEYONDERS in work package 7 – Use cases. The following scheme summarizes this evolution along the Work Plan of the project.

To achieve the objectives proposed of TRL evolution in BEEYONDERS, it is a requirement to establish a R&D methodology able to be adapted and suit each of the BTs of the project. The R&D methodology designed for BEEYONDERS is a continuous feedback loop between design, development, and validation-demonstration where design is, however, at the beginning of the whole process. This proposed methodology is a combination of the following statements:

- **Design Science approach** will be used as Scientific coordination methodology. This research methodology incorporates the principles, practices, and procedures necessary to carry out this type of research and is focused on the following aspects: To be consistent with previous literature, to provide a model and methodology with which to conduct the research, and to provide a system for presenting and evaluating the research data.³



Figure 2. Design Science scheme

- The **User-Centered Design approach** is the most widely used development process in human's wearable devices. This user-driven design process focuses attention on the usability goals, user characteristics, environment, and tasks at each stage of the development.^{4 5} In BEEYONDERS, this approach is mainly applied in the WP5 technologies (smart wearables, exoskeleton) as are devices to be worn by human workers, so the relation with the work packages of the project is as stated in the following scheme:

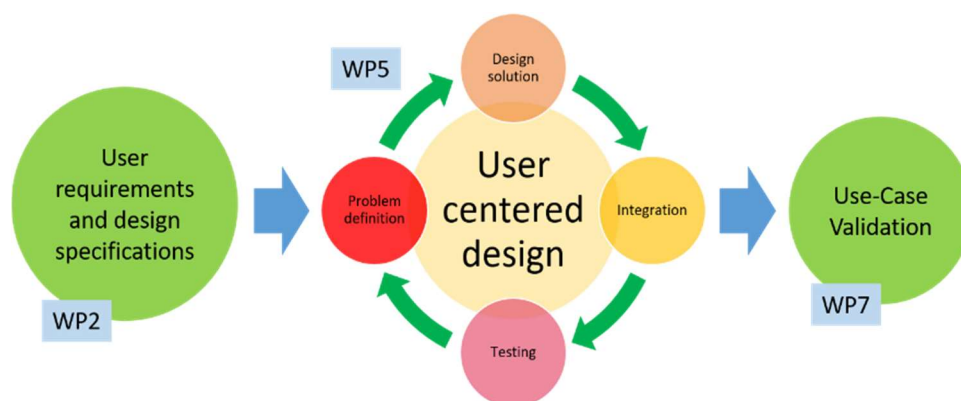


Figure 3. User-Centered Design scheme

³ Dresch, A., Lacerda, D.P., Antunes, J.A.V. (2015). Design Science Research. In: Design Science Research. Springer, Cham. https://doi.org/10.1007/978-3-319-07374-3_4

⁴ Standardization, I.O.f., Ergonomics of Human-system Interaction: Part 210: Human-centred Design for Interactive Systems. 2010: ISO.

⁵ Shore, L., et al., Technology acceptance and user-centred design of assistive exoskeletons for older adults: A commentary. 2018. 7(1): p. 3



From these figures, it is clearly highlighted the importance of the aspects of “testing” and “validation”, as per any prototype development proposed in the project. This aspect clearly outlines the importance and the need of providing a Testing Plan, creating a good mechanism to ensure the objectives for these are achieved during the project timeline.

Design is understood as a qualitative and quantitative set of principles that will be breakdown in requirements, objectives and, finally, Key Performance Indicators (KPIs) referring to Human-centric principles, automation of the construction site, technology integration (from other industries and in BEEYONDERS) in construction, resource efficiency, reduction of waste, reduction of embodied CO2 emissions, safety, wellbeing, and prevention, economic performance and competitiveness with non-EU competing solutions.

The KPIs have been set at the beginning of the project (WP2: Specifications & testing, D2.3 LCA/LCC KPIs ⁶, D2.4 Workers KPIs⁷) to guide technological and non-technological performance targets for development of the individual solutions and for their integration as a whole and to blueprint the test and validation program to iteratively validate the results at the different stages and TRLs of the project. Design is set at the beginning of the project, but, while the main principles and KPIs will remain untouched, these will be revisited, nurtured, and enhanced with new relevant information, to keep pace with the progress of reality, which will affect all parameters as they evolve in time (e.g., change in external requirements and conditions of any type: markets, legal, technological, etc.). The market and technology watch performed in WP8 (Communication, Standardization, Dissemination and Exploitation) will be key to support this evolution.

Development starts at TRL 4 with WP3: Digital Twin & Diagnostics, WP4: Autonomous vehicles and collaborative robotics, WP5: Worker-centric technologies to improve health, safety and well-being, and WP6: Additive manufacturing, and takes the solutions and their integration to TRL5, evolving to prototypes, under the guiding KPIs, and with the feedback of testing at the level of TRLs 4 (lab) and 5 (relevant DT-integration environment).

The prototypes and the whole BEEYONDERS integrated solution as applied to the different use cases in WP7 (Case Studies) are, at the final stage, reaching TRL6, further fine-tuned and adjusted with the inputs from the validation campaign in the case studies. As a mean of assuring the successful achievement of target TRL6, the WPs time distribution (Gantt) has been designed in order to permit an overlap between the beginning of the technical developments with the testing stage in the lab scale, and starting the case studies approximately at the midterm of the technological work packages in order to allow a further iteration of tuning these technologies while these are initially deployed on the case studies.

⁶ Stipanovic, I., Skaric Palic, S., Kucinic L. et al (2023) D2.3: LCA/LCC KPIs, BEEYONDERS project, <https://beeyonders.eu/>

⁷ Eva Martinez, E., Sestrem, J., Folco, G. et al (2023) D2.4: Workers KPIs, BEEYONDERS project, <https://beeyonders.eu/>



3. Research & Development objectives

The formulation of clear and well-defined research objectives serves as a pivotal compass that guides the course of BEEYONDERS project as a Research and Innovation Action (RIA). A Research and Innovation Action (RIA) primarily consists of activities aiming to establish new knowledge and/or to explore the feasibility of a new or improved technology, product, process, service, method, tool, or solution⁸. The previously mentioned objectives outline the specific goals and ambitions of the research & development proposed, providing a roadmap for the research design, methodology, and eventual findings.

The process of defining research objectives involves a meticulous balance between overarching aspirations and practical feasibility. These objectives delineate the boundaries within which the research will unfold, shedding light on the expected outcomes and potential contributions to the field. By articulating the intended outcomes and outputs, researchers are empowered to maintain focus, allocate resources judiciously, and assess the success of their endeavours objectively.

This section delves into the detailed Research & Development objectives for each of the BEEYONDERS' BTs and highlights their role in shaping the trajectory of these during the project's lifetime. Thus, and serving as the main guideline and objective, during the proposal stage of BEEYONDERS a set of milestones were set. These are presented in the table below extracted from the proposal document:

Table 1: List of milestones (extracted from Table 3.1d of proposal document)

Milestone number	Milestone name	Related WPs	Due date	Means of verification
MS1	Project requirements	2	M6	Consortium has agreed the documents that specify the general and technical requirements.
MS2	Tested initial prototypes	3-6	M18	Existing initial prototypes of wheel loader machine, drones, exoskeleton, smart wearables, and positioning software.
MS3	Prototypes validation at TRL5	2-6	M24	Registers from test validations prove that BEEYONDERS's prototypes have reached the requirements.
MS4	Demonstrations in case studies are set up	3-7	M30	Prototypes deployed on the case studies locations.
MS5	Demonstrations of the case studies are completed and TRL6 achieved	7	M40	Registers from case studies demonstrations prove that BEEYONDERS's prototypes have reached the requirements.
MS6	Exploitation strategy is defined	8	M42	Exploitation strategy of the partners is available and next steps to reach TRL9 are defined.

According to the presented milestones, it can be understood as the Research and Development is organized in two main segments. The initial segment will occur within WP2 (MS1 – Project requirements) defining the initial requirements and continues with the definition of this deliverable “Testing Plan” for the solutions and their integration along with the detailed planning and configuration of the tests needed to evaluate the performance of the solutions at two stages of development:

- Lab scale validation of the individual solutions and of their conceptual integration at TRLs 4 (related with MS2 – Tested initial prototypes through the testing of the initial prototypes of the BTs).

⁸ https://cordis.europa.eu/programme/id/H2020_ECSEL-2017-2/es



- Validation of the BT solutions and their integration in relevant environment (TRL 5), this is, at a scale where the actual prototypes are working together in an integrated manner leaving the lab to work in test facilities and real interconnection at a level prior to field case studies.

The final segment of validation will follow the same main scheme in WP7, beginning with the planning of the validation campaigns in each of the case studies (MS4 – Demonstrations in case studies are set up) and performing the validation at real scale to reach TRL6 (MS5 – Demonstrations of the case studies are completed and TRL6 achieved). The market and technology watch performed in WP8 during the full lifetime of the project (Communication, Standardization, Dissemination and Exploitation) will be key to support this evolution.

Once outlined the global R&D objectives of BEEYONDERS, the next step is to analyze in detail per BT the specific objectives are meant to be validated during the use cases demonstration. These objectives have been defined at the initial stage of the project, and are in line according to the BEEYONDERS' TRLs, available resources and current state of the art.

The objectives are collected in the following table:

Table 2: R&D objectives for BEEYONDERS technologies

AUTONOMOUS GROUND VEHICLES

When most of the people think in connected and autonomous vehicles (CAV) or machines, what automatically comes to mind are leading car manufacturers out in the race of building autonomous vehicles. Many other industries are incorporating autonomous technologies (Artificial Intelligence methods, high resolution sensors, path planning) with the ambition of improve efficiency, productivity, and safety. The construction sector has not been stranger and automation in construction have been a hot topic during the last two decades⁹, but as processes in the construction sites are digitalized, teleoperated or fully autonomous machinery are becoming more important.

The project will contribute with the development and test of a complete software stack for tele-operated and autonomous machines for construction sites and maintenance tasks that will implements, in addition to the advanced driving functions (path planning, obstacle avoidance, etc) a task execution module that will be designed to deploy cutting-edge algorithms based on AI specifically designed for constructions and maintenance machine's tasks like bucket loading, dumping or visual inspection.

According to multiple reports over the last year, the top trends shaping the autonomous construction equipment market include Artificial Intelligence (AI), Internet of Things (IoT) and High-Resolution Imaging Sensors to boost productivity (equating manual and autonomous driving), improve the safety (by reducing workers risks) and reducing CO2 emissions (path planning optimization and learning systems). Although autonomous driving technology can be used on construction sites, only partial sub-tasks could be achieved because:

- Construction sites are very hazard environments where perception and control systems must be designed and trained to be robust.
- Environments change and evolve continuously so maintaining HD or statics maps is not an option.
- Some specific tasks like controlling the bucket for loading and dumping are needed (material and pile characterization problems, throttle, and brake smart control to avoid wheel slip, fuel consumption and maintenance cost).

⁹ A. Hemami, F. Hassani, An overview of autonomous loading of bulk material, Proceedings of the 26th ISARC, Austin TX, U.S.A. 2009, pp. 405–411.



- High level tasks are performed by multiple machines in collaboration, so a high-level decision-making system is needed (Digital Twin). Concretely, wheel loaders, excavators, and dumpers have been an active area of research because they are very versatile and make quick and easy earth moving.

In summary, wheel loaders machines perform three main tasks during one cycle (short or large): 1) loading 2) navigating and 3) dumping while excavators only need tasks 1) and 3) plus some type of spinning. In most construction site scenarios, machines collaborate with other machines and humans to perform high-level tasks, and dumpers play an important role moving the load on long paths. Many works have been done in the main challenges described¹⁰ and specifically on bucket filling using neural network and reinforcement learning^{11,12}. One of the latest systems proposed is an Autonomous System Excavator (AES), which operate the arm of one excavator in different scenarios for handling materials¹³, using a 3D (LiDAR, cameras) perception system to sense the environment, task and motion planning and control.

In order improve efficiency, productivity, and safety in construction machines (performing high level task in a collaborative way) the following advancements to the state-of-the-art are necessary and will be developed in BEEYONDERS:

1. Develop the base of a full, machine agnostic scalable software stack for teleoperate/autonomous construction machines in hazards (indoor & outdoor) environments able to handle advanced driving functions, teleoperated features (avoiding collision system) and specific task execution (bucket loading or dumping) by adding plugins ad-hoc for the desired task.
2. Enable mechanisms for create, maintain, and share accurate 3D maps in real-time that encode infrastructure information, with multisession capabilities and completeness in the long term.
3. Improve the cognitive capabilities of the machines designing and developing a collaborative multi-sensor perception pipeline for enhanced shared situational awareness in construction sites, for heavy machines or aerial robots.
4. Designing a machine agnostic module based on reinforcement learning technics to teach machines task execution with the data capture in teleoperation mode.
5. Robotic platform with autonomous functionalities will be developed and demonstrated in road maintenance tasks, as:
 - a. Measure retro-reflectivity of road horizontal signals.
 - b. Link with Digital Twin for data exchange.
 - c. Coordinate with other elements (aerial autonomous vehicles) for full coverage.

AUTONOMOUS AERIAL VEHICLES

Traditionally, inspection tasks in construction sites have been performed manually, constituting intense, costly, and sometimes dangerous work. In the past few years, the use of aerial robots for this kind of tasks is on the rise, being employed for the inspection of roads, railway, wind turbines, mines, or bridges. However, most of these operations are still carried out manually, requiring an experienced safety pilot, and the inspections are usually local, i.e., within visual line of sight of the operator. Furthermore, the data

¹⁰ S. Dadhich, U. Bodin, U. Andersson, Key challenges in automation of earth-moving machines, Automation in Construction, Volume 68, 2016, Pages 212-222, ISSN 0926-5805,

¹¹ Dadhich, Siddharth & Sandin, Fredrik & Bodin, Ulf & Andersson, Ulf & Martinsson, Torbjorn. (2020). Adaptation of a wheel loader automatic bucket filling neural network using reinforcement learning. 1-9. 10.1109/IJCNN48605.2020.9206849.

¹² Backman, Sofi & Lindmark, Daniel & Bodin, Kenneth & Servin, Martin & Mörk, Joakim & Löfgren, Håkan. (2021). Continuous control of an underground loader using deep reinforcement learning.

¹³ L. Zhang, J. Zhao, P. Long, L. Wang, L. Qian, F. Lu, X. Song, D. Manocha, An autonomous excavator system for material loading tasks. Sci. Robot. 6, eabc3164 (2021).



obtained is typically not incorporated into a digital twin of the infrastructure and, as a result, it cannot be coupled with the rest of the information obtained by other systems, allowing only isolated analysis.

Autonomous navigation is one of the most critical functions for achieving an automated inspection system using robotic technologies, especially in situations where Global Navigation Satellite System (GNSS) signals are degraded or unavailable, such as in the proximity of buildings or indoor environments. The maturity of Simultaneous Localization and Mapping (SLAM) techniques during the last decades is a major factor in achieving consumer-grade deployment of autonomous navigation for ground robots. However, in the case of aerial robots, which move and operate in 3D environments, these methods imply greater computational complexity. Apart from the spatial dimensionality, aerial robots are exposed to additional challenges as their onboard capabilities are limited in terms of payload, computational resources, and power constraints.

In large-scale and long-term operations, wrong data associations are often unavoidable. Current SLAM systems heavily rely on robust estimators and geometric verification to mitigate this effect. For map persistence, loop closure and global localization techniques are vital. However, the presence of distinct and observable visual or structural features continues to have a significant impact on the perception algorithms. For vision-based approaches, viewpoint changes, poorly textured areas and lighting conditions often pose major challenges. Other localization systems use LIDAR as its main sensor since it does not depend on external light conditions. While their application in autonomous driving is on the rise, their use in aerial robots is still limited because they require more available payload and more powerful onboard processing capabilities.

BEEYONDERS will develop advanced aerial robotic technologies that will provide a new tool for the inspection of varied civil infrastructure, for both indoor and outdoor scenarios. In particular, BEEYONDERS will focus on the development of a GPS-free autonomous navigation system for the inspection of tunnels, and a long-range system for road and building inspection system, which will drastically reduce the total inspection times. In addition, the developed systems in BEEYONDERS will integrate the information acquired by the aerial robots into an object-oriented model for feeding information into a digital twin of the infrastructure.



Figure 4. Drone applications in surveying tasks

ADDITIVE MANUFACTURING

3-D Concrete Printing (3DCP) is changing the way concrete/mortar are used. Layer-by-layer deposition of concrete managed digitally removes the need of formwork, offer more design options to enable the construction of optimized structures with the right material, and improve construction safety reducing construction cost by involving less labour workers but with limited information relating to the structural performance of the 3D printed material and components.



This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101058548.

Additive manufacturing focuses on leveraging 3D Concrete Printing (3DCP) technology for marine caisson design & manufacturing, to boost the efficiency of the structure and its environmental impact, optimizing both the design and manufacturing of both the caisson itself but also of artificial reefs to be introduced on caisson's surfaces, to promote bio-marine life development. 3DCP robotic device will manufacture a prototype of caisson and a 3D printed reef attached to one of the vertical parament.

For the construction of caisson, a 3D printable material will be developed. To boost the bio-receptive function of the caisson, a second material and its related applying method will be developed. An extensive IoT deployment during the manufacturing of the printing material will be carried on, where AI models would provide the best mixing proportions of these kind of materials and knowledge on their mechanical properties. Thus, a rheological-temperature in-line monitoring and control both for feeding the optimal fresh concrete for marine infrastructure and marine life and its integration to the caisson will be studied.

The material development starts by sourcing raw materials that are available near case study demonstration cite and their characterization. Once the specifications (such as maximum aggregate size, strength level, durability class, bioreceptivity, etc) are clarified, lab trials will be carried out to for optimizing the recipes. By going through lab trials, for 3D printable material, its fresh properties will be optimized according to the 3D printing process and its long-term properties will be optimized to achieve the requirement. The feasibility and adequacy of different printing processes, such as extrusion of large layers including inclusion of reinforcements or printing the structure without reinforcement by working on fiber reinforced material or by applying post-tension, will be assessed in detail considering the final caisson and its operation environment. The functional requirements for the caisson material are:

- Adapted fresh properties development to ensure the pumping and printing process so and the integrity of printed element during construction phase, measured by yield strength.
- Adapted hardened properties to assure long term integrity of the caisson structure, measured by compressive and flexural strengths at age of 1 day, 7 days and 28 days, durability indicators such as water porosity and coefficient diffusion of chloride ion.

For bio-receptive material, in addition to its own properties, the interaction with 3D printable material and how the impact of apply method will also be accessed and optimized. Once succeeded lab trial, the 3D printable material will go through printing trials through which minor adjustments are possible to fit it with the printing process. Before the demonstration, industrial trial will be carried out so that the performance materials can be validated from factory production, with possible some final adjustments on the recipes. The functional requirements for the bio-receptive material are:

- Adapted fresh properties and applying method to assure its application on the surface of caisson.
- Bio-receptivity adapted to the marine environment of interest.

EXOSKELETON

In construction, the back and shoulders are the body area most affected by MSD (about 64% and 24% respectively)¹⁴. To accomplish many extremely physical tasks coordinated actions of the back and arm muscles are necessary.

¹⁴ Schneider, E., S. Copsey, and X. Irastorza, *Osh [occupational safety and health] in figures: Work-related musculoskeletal disorders in the EU-facts and figures*. 2010: Office for Official Publications of the European Communities.



Occupational exoskeletons have been developed with the goal of assisting workers during manual material handling operations, where they can improve safety when the risk of injury becomes significant. In the past few years, exoskeletons have started to become a commercial product^{15, 16}. These pioneering systems are technically simple, offering partial passive assistance (relying for their assistance on springs and elastic belts), at a relatively low cost¹⁷.

Unfortunately, these devices are not well suited to the unstructured working environment typically found in the construction sector, which is also characterized by a diverse range of working tasks and routines. In fact, current commercially available exoskeletons, which have mainly been developed for the automotive sector, are actuated passively, and are usually designed to assist one unique body segment at a time. Few examples use a modular approach.

On the other hand, active exoskeletons are being researched that offer the capability to introduce, on demand, additional energy supplied from an external source (e.g., batteries). These exoskeletons are suitable for dynamic tasks, but they are also heavier and more complex¹⁸. Moreover, because of the unstructured working environment (e.g. rough, uneven and broken ground) and the non-repetitive nature of tasks that are common in the construction sector, there is a need to develop exoskeletons that are both smart and powerful. Exoskeletons for the construction sector should be actuated by active actuators (e.g. brushless motors) and controlled in such a way that the modulation of the assistance is based on bioinspired sensor fusion and intelligence of specific cognitive control algorithms^{19, 20, 21}.

The key innovation in the exoskeleton technology that BEEYONDERS will develop, is the adaptation, customization and enhancement of an active exoskeleton that supports manual material handling and repetitive tasks and aligning these developments and refinements to the very specialized demands that arise in the construction industry.

It will be developed a lightweight active exoskeleton, which will be versatile, modular, and capable of assisting the most fatigue inducing operations found in the construction sector (e.g. fetching of material, static and dynamic operations in uncomfortable/tiring postures, etc.). From a technical point of view, cutting-edge technologies including sensors, actuators, and smart concepts will lead to a versatile device suitable for multiple purposes.

When considering the current state of the art in occupational exoskeletons²², and the highly interactive action of the back and arms while accomplishing manual material handling^{23, 24}, it is noticed that current occupational exoskeletons assist only one unique body segment at a time. Additionally, they are designed

¹⁵ LAEVO, <http://en.laevo.nl/home-en/>.

¹⁶ SUITX, <https://www.suitx.com/>.

¹⁷ Sugar, T., et al., *Hip exoskeleton market-review of lift assist wearables*. 2018.

¹⁸ Toxiri, S., et al., *Back-support exoskeletons for occupational use: An overview of technological advances and trends*. 2019: p. 1-13.

¹⁹ Di Natali, C., et al., *Design and evaluation of a soft assistive lower limb exoskeleton*. *Robotica*, 2019. 37(12): p. 2014-2034.

²⁰ Poliero, T., et al. *Enhancing Back-Support Exoskeleton Versatility based on Human Activity Recognition*. in *2019 Wearable Robotics Association Conference (WearRAcon)*. 2019. IEEE.

²¹ Poliero, T., et al. *Assessment of an On-board Classifier for Activity Recognition on an Active Back-Support Exoskeleton*. in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. 2019. IEEE.

²² Nussbaum, M.A., et al., *An Introduction to the Special Issue on Occupational Exoskeletons*. 2019, Taylor & Francis.

²³ Zhu, Z., A. Dutta, and F.J.A.i.C. Dai, *Exoskeletons for manual material handling—A review and implication for construction applications*. 2021. 122: p. 103493.

²⁴ Yoo, I.-g. and W.-g.J.W. Yoo, *Effects of different lifting and lowering heights on upper arm, shoulder and back muscle activity during a manual material handling task*. 2016. 53(1): p. 175-179.



to address very specific single goal working tasks. However, manual material handling activities require a coordinated effort of the arms and the back. Therefore, BEEYONDERS aims at filling this gap by developing an innovative technology that will support a large variety of working tasks commonly encountered into the construction sector.

Although the development of such a system is very challenging, integrating and developing onboard CPU, motors, wearable sensors, and control strategies, will provide a great level of flexibility and adaptability to the specific user needs.

A final point worth noting is that an active device that includes electric motors, sensors, CPU and batteries will not only provide assistance to workers while carrying out their daily work but will also improve safety through increased/improved connectivity and data sharing with a Digital Twin. Safety will be enhanced through the reduction of the risk of injury due to the assistance provided by the exoskeleton.

SMART WEARABLES

Construction site forms a challenging working environment from the point of view of ergonomics and risk of injuries. Among construction workers musculoskeletal disorders are common²⁵, ²⁶ due to difficult postures and repetitive movements²⁷. Different sensors have been used for detecting various construction site hazards for workers, the main types are accelerometer, gyroscope, ECG/EKG, infrared and RFID sensors²⁸. To gain more information of the physical activities and postures of the worker, data obtained from inertial motion unit (IMU) sensors has been collected and analysed by also utilizing machine learning models²⁹, ³⁰. Embedding these sensors in the already available work clothes provide an easy way to gather precise information concerning postures and activities directly from the workers. This smart work clothes no longer require preinstalled external sensors on the working site in specific locations and can support the detection of certain type of movements better than simple activity trackers worn on wrists. Wearable technologies enable a new data-based approach to improve the health and safety of workers. However, gaining information about postures and physical activities is not enough to tackle the challenges related to ergonomic and risky behaviour with the aim of mitigating these issues. The information has to be utilized in different levels of the organisation, for example, by combining the detected postures with models of physical strain available in the literature. This will provide more insight about the well-being and the working conditions of workers. According to prior research, constructions workers are positive on using wearable sensors for this purpose³¹.

In this project, the objective is to develop a smart wearable sensor platform that can be integrated into work clothes that will support the identification of risky behaviour in construction site. This smart wearable will be able to identify the physical activities of the worker and evaluate the related ergonomic risks by

²⁵ Schneider, S. P. (2001). Musculoskeletal injuries in construction: a review of the literature. *Applied occupational and environmental hygiene*, 16(11), 1056-1064.

²⁶ Holmström, E.B. (1992). Musculoskeletal disorders in construction workers related to physical, psychosocial, and individual factors, *Acta Orthopaedica Scandinavica*, 63:sup247, 55-55, DOI: 10.3109/17453679209155037

²⁷ CPWR – The Center for Construction Research and Training. (2018). *The Construction Chart Book - The U.S. Construction Industry and Its Workers*, 6th edition.

²⁸ Awolusi, I., Marks, E., & Hallowell, M. (2018). Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automation in construction*, 85, 96-106.

²⁹ Mäkela S-M, Lämsä A, Keränen J, Liikka J, Ronkainen J, Peltola J, Häikiö J Bordallo, Lopez M, Introducing VTT-Conlot: A realistic dataset for activity recognition of construction workers using IMU devices *IEEE Access*

³⁰ Mäkela S-M, Lämsä A, Keränen J, Liikka J, Ronkainen J, Peltola J, Häikiö J, Bordallo, Lopez M, VTT-Conlot: A realistic dataset for activity recognition of construction workers using inertial motion units

³¹ Häikiö, J.; Kallio, J.; Mäkelä, S.M.; Keränen, J. IoT-based safety monitoring from the perspective of construction site workers. *International Journal of Occupational and Environmental Safety* 2020, 4, 1–14.



estimating the muscular strain, by identifying any hazardous movement, and also by detecting “man down” situations. The technical challenges of this platform lie in the communication infrastructure of the working environment (e.g., poor quality or non-existing connections between wearable devices and other systems) and diversity of the physical activities of the workers.

During the development of the smart wearables, great attention will be paid to the investigation and data gathering of the key features for identifying ergonomic risk, worker’s fatigue, environmental risk conditions, etc. This information will be further sent to the digital twin to benefit use cases in BEEYONDERS. The research will focus on the assessment of the ergonomic and environmental risks experienced by workers when performing work activities at construction sites, by integrating also data collected by environmental and noise sensors. The research will concern, also, the evaluation on the usability of and user experience with the proposed device.

Technological components will be developed and tested in laboratory environment based on the project requirements. Some of the technical components’ development include machine learning which needs a data collection for training the specific component.

A final approach for this device is focused on the safety of the workers allocated into the road construction case, as in this stage the workers will be executing their tasks in the same area together with the autonomous wheel loader proposed in the project. Construction sites are often filled with hazards, which can hardly be prevented in the construction site layout design. In addition, there are many inevitable interactions between workers and heavy machinery. Construction workers are constantly exposed to hazards and close-call incidents, i.e., incidents that might have evolved into accidents. In order to enhance workers’ safety at the construction site, researchers have proposed tracking the workers’ real-time location for intervening in and analysing close-call incidents in time at the construction site. Various technologies and designs can be used to build a locating system, such as Ultra-wideband, Bluetooth Low Energy, and RTK-GNSS.

The objective is to demonstrate that smart wearables can detect proximity of workers with the construction machinery, so raise an alert in real time to the workers or even automatically stop the wheel loader in case of a dangerous proximity situation. With the input of real-time location data and a predefined safety layout or model of the construction, detect if there are close-call incidences, e.g., workers are within the preidentified hazard zones or the protective envelope of the heavy machinery. Based on close-call detections, further functions are implemented, such as risk assessment on close-call incidences, mitigation measures suggestions, and personalized feedback.

HUMAN – ROBOT COLLABORATION

Agent-based models (ABM) are a class of computational models aimed to simulate the actions and interactions of autonomous agents (both individual and collective, such as organizations and groups) to evaluate their effects on the system as a whole³². ABM combines elements of game theory³³, complex systems, emerging behavior, computational sociology, multi-agent systems. ABMs allow accurate studies on the behavior and actions of individuals but can also be applied to model (with different levels of accuracy) and simulate the interactions among human agents³⁴ and between the human agent and other agents. In the framework of the BEEYONDERS project, ABM will be applied to simulate and inspect the

³² Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems. In Proceedings of the National Academy of Science 99, 3 (2002), 7280--7287.]]

³³ Axelrod R. Axelrod R., (1997) The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration (Princeton Univ. Press, Princeton, NJ)

³⁴ MACY, Michael W.; WILLER, Robert. From factors to actors: Computational sociology and agent-based modeling. Annual review of sociology, 2002, 28.1: 143-166.



coexistence of human workers and robotized autonomous machineries in the worksite, to allow the analysis of the risks related to the work operation, also allowing the extension of the project assessment and validation to a higher number of possible scenarios to guide the selection of the best possible layout.

An agent-based model will be developed as support tool for the activities in the case studies where coexistence and cooperation between machines and human is planned. The system aims to identify possible risks and conflicts by simulating both human operators and machines tasks in terms of multiple parameters such as motion in the worksite or sequence of sub-tasks. Agent-based model will be fed with the algorithm and behaviour patterns developed for determining the tasks and the movements of the robotic machines, also considering the perception system on board of each machine for collision avoidance. An accurate description of the activities will complete the framework of information necessary for the development of the agent-based model.



Figure 5. Agent based model simulation for railways worksite

The agent-based model developed in BEEYONDERS allows a dual analysis: the first concerns the risks related to the work operation, while the second focuses on the operation planning and its efficiency. Concerning the risks associated to the work environment, particular attention will be paid to the workspaces of each agent involved in the inspection activities, to identify and determine possible conflicts between operations and movements of different agents. This analysis will allow planning several working layouts reducing the level of risk for the operator as well as maximizing the operation efficiency by optimizing the planning. The efficiency of the operation planning is relevant when considering operation timing and related cost. Therefore, BEEYONDERS will carry on the identification of a subset of working layouts that minimize the risk levels and maximize the activities efficiency.

This approach represents a novelty in collaborative robotics for the workplace, with previous research mainly focusing on domestic robots for healthcare³⁵ and proposes a complementary high-level planning of collaborative activities, where human behavior is at the center of the analysis.

DIGITAL TWIN

One of the main objectives of BEEYONDERS project is to demonstrate that new disruptive technologies can be applied in the specific context of a construction site. These new technologies will bring to the

³⁵ CARLOTTA, Belviso; RAINER, Hasenauer; BECHTOLD, Ulrike. Socially Assistive Robots Diffusion in Elderly Care A Pre-Adoption Study by Agent-Based Modeling. *Journal of Strategic Innovation and Sustainability*, 2018, 13.5: 58-75.



construction site additional machines and equipment, some of which will be automatic or autonomous. This means that a very high-definition knowledge of the site will be needed to plan and organize the work while ensuring the highest security to workers. Therefore, a fully integrated numerical environment must be specified and implemented: A Digital Construction Platform.

The results obtained in BEEYONDERS will be centralised and assessed in the Digital Twin environment. A fully integrated Digital Twin must be specified and implemented and stored in a dedicated database. The initial Digital Twin model is built, through specific data importer (static data importation layer), from the BIM models provided by the conception phases of the project, and from initial planification data. This model is connected to a set of “Input services”, that will produce the specific indicators based on the KPIs defined in WP2. All the “active” parts of the construction site (workers, equipment, machines, robots, drones) will provide, through sensors, a big amount of information (position, orientation, cloud points, distances, ...). All this information will be plugged to the Digital Twin and treated by dedicated “Output services”.

The core objective of the Digital Twin platform in the BEEYONDERS project is to conceptualize, implement and demonstrate breakthrough technologies in knowledge-driven construction work process, in which decision makers are supported by, for example, wearable sensors, tracking technologies, worker-assistance robotic devices, and autonomous construction vehicles.

The use of Digital Twins for runtime data gathering and representation linked to a digital counterpart of the reality for further data analytics, prediction and visualization has become highly relevant in many industries. In the AEC/FM domain, however a digital model is still restricted to the stage of a static design BIM model. Its application area remains limited to building design and planning work tasks instead of taking advantage of the rich data that is available in construction operations. The three main ambitions of the project, from which all others are derived, can be summarized as:

- Reach beyond the state-of-the-art applications of traditional data acquisition and analytics (e.g., brute force approaches) which use all monitored data streams for a specific purpose, towards a broader vision which spans several knowledge-driven monitoring technologies and several domains (i.e. levels of safety, efficiency, automation).
- Provide as an initial result the full situational awareness as expressed in alternative planning and performed construction site status modeling (i.e., yielding a high-level holistic knowledge model of the construction process, seamlessly integrated with runtime data to products and equipment states. It will be achieved using property-graph representation techniques. The outcome is previously not discovered knowledge that decision makers can access in the digital twin.
- Generate Artificial Intelligence techniques that drive learning from sensed data and knowledge extraction when aligning it with process information, predicts future problems or risks (i.e., related to safety, health, or well-being issues of workforce) and suggests alternative outcomes. KPIs that are initially derived by users from historical values will be continuously updated (i.e., advancing from lagging to leading safety indicators) and will allow detailed evaluations, e.g. measuring performance change quantitatively instead of using only human observation and subsequent subjective judgement. Benchmarking current vs. previous performances has the potential to influence rules and even regulations of decision-making processes.

The BEEYONDERS concept derived from these three points ultimately makes decision making over the entire construction operation more efficient, while projects will become safer, more productive, and stay on schedule or within the planned budget.



4. Methodology Approach for testing

In this section is described the methodology approach to accomplish the testing plan presented in this document for each BT, where the results and conclusions are presented in this final version of this deliverable (M24).

Developing a comprehensive testing plan is essential to ensure the quality and reliability of the technologies proposed to be developed in BEEYONDERS, and these steps are proposed to be followed by technologies developers. The testing methodology covers the following aspects:

- **Revision of technical requirements:** Before starting any testing, ensure a clear understanding of the project's requirements. This involves reading the documentation, discussing with stakeholders, and clarifying any ambiguities. To perform this point, the following documents of BEEYONDERS are related:
 - D2.1: System Requirements: General objectives of BEEYONDERS.
 - D2.2: Technical Requirements: Technical objectives of BEEYONDERS.
 - D2.3: LCA/LCC KPIs: Economic & environmental KPIs with workflows of use cases.
 - D2.4: Workers KPIs: Social and workers KPIs with workflows of use cases.
 - D3.1: Digital Twin architecture: Proposed architecture for Digital Twin related to BTs.
- **Define the research objectives:** The objectives to be achieved by the technologies to be demonstrated during the use cases of BEEYONDERS.
- **Determine testing methods:** These can include experiments, surveys, interviews, observations, or data analysis, considering the nature of the research and the data is required to collect.
- **Execute the testing:** Conduct the testing according to the designed protocol, collecting data from the participants using the specified methods and tools. The testing is mainly divided in the following stages:
 - Data analysis: Once the testing is complete, analyse the collected data using appropriate statistical or qualitative analysis techniques. Use the analysis to answer your research questions or test your hypotheses.
 - Iterations: Based on the testing results, refine the test cases and testing approach. This could involve running additional tests, modifying existing ones, or focusing on specific areas.
 - Validate iterations: Once defects are fixed, retest the affected areas to ensure that the issues have been resolved without introducing new problems.
 - Final verification: Conduct a final round of testing to ensure that all identified defects have been resolved and that the software / device meets the specified requirements.
 - Reporting: Interpret the results of your data analysis and draw conclusions based on your findings. Summarize the outcomes, implications, and limitations of the research. Prepare a comprehensive report or presentation of the testing results.
- **Document Results:** Preparing a testing report to be included in the related deliverable per each technology, that summarizes the testing process, results, and any relevant metrics. This will provide stakeholders with insights into the quality of the software / device. The testing results will be presented in each deliverable related to the technologies developed.
- **Acceptance from partners:** Get formal approval from stakeholders and partners, indicating that testing is complete, and the software / device is ready for release and deploy into the use cases of BEEYONDERS according to the project objectives.

These steps describe the methodology to be followed by each BT ensuring the TRL5 is reached. As a summary, below it is presented a scheme of this testing methodology:



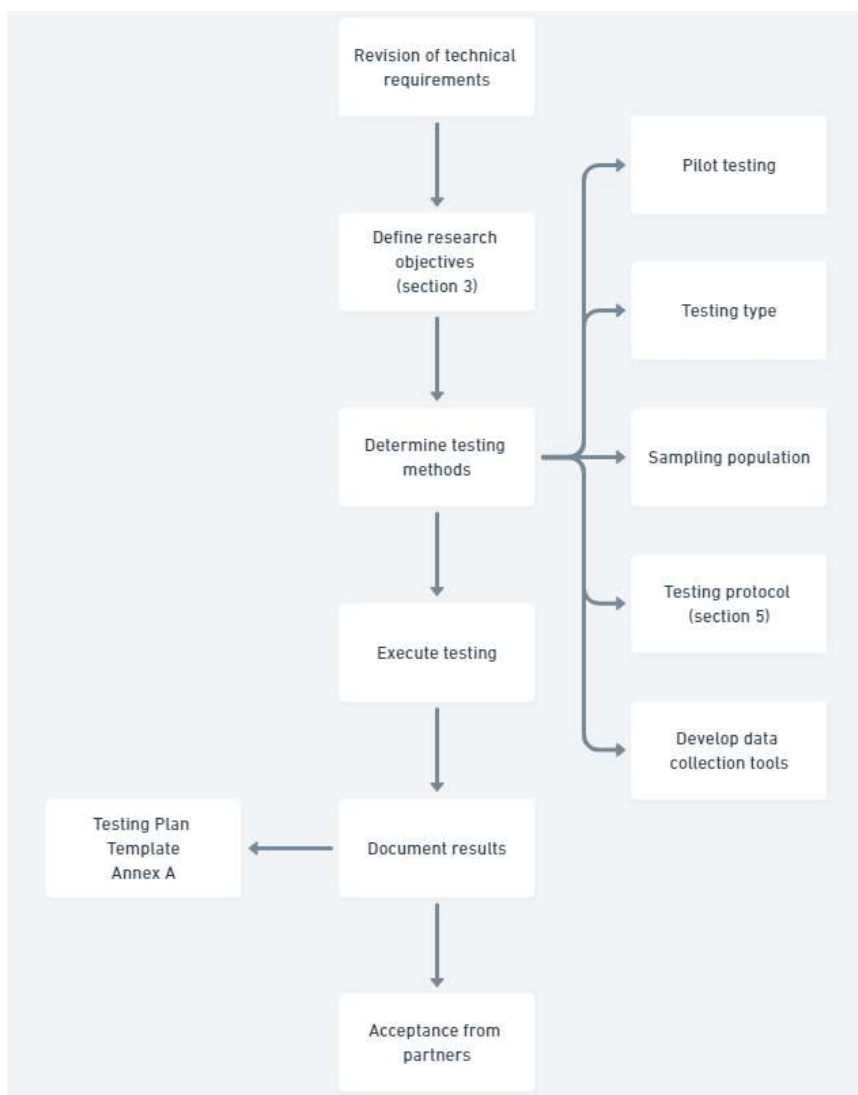


Figure 6. Testing Plan methodology scheme



5. Testing Procedure for each technology

A testing procedure is developed to provide a structured approach to evaluate the performance and capabilities of the new technologies are object of development in BEEYONDERS. Whether it's software applications, hardware devices, or complex systems, a rigorous testing process serves as a validation test for their viability and effectiveness.

The testing procedure serves various critical purposes, such as identifying defects, optimizing performance, validating functionalities, and mitigating potential risks. It entails a systematic series of activities designed to assess the technology's alignment with its intended purpose and user expectations, specially aligned with their objectives to be achieved during the use cases proposed in BEEYONDERS's WP7 – Use cases. Through meticulous testing, developers gain insights into a technology's strengths, weaknesses, and areas for improvement to reach the mentioned TRL in section 3 of this document.

The work carried out in this final version comprises the collection of the inputs from the technological partners of BEEYONDERS to gather, assess and validate the definition of the testing procedures will be executed regarding each technology, focused on the final objective in the related use cases will be meant to achieve.

The detailed testing procedures for each technology are presented in the following subchapters.



Autonomous Ground Vehicles

In BEEYONDERS, two types of autonomous ground vehicles are going to be developed: A wheel loader machine (to be deployed in tunnel use case and road use case), and a maintenance vehicle (to be deployed in the road maintenance use case).

Autonomous Wheel Loader

The testing plan for the autonomous wheel loader involves systematic testing and evaluation to ensure safe and efficient operation, and to demonstrate the capabilities to perform the prototype objectives in the proposed use cases to be deployed. The related task T4.6: “Autonomous and connected wheel loader machine” will develop an autonomous and tele-operated wheel loader system (AWL) for earth-moving tasks transforming a wheel loader machine in an autonomous vehicle able to perform advanced driving tasks in short or large cycles (tunnel and road construction) and specific earth-moving tasks (loading and dumping).

In the deliverable D4.1 – Modelling and data link interface³⁶, section 3.3, can be found the details about the machine description and the implication of this autonomous wheel loader in BEEYONDERS use cases. As a summary, the machine prototype is presented below:

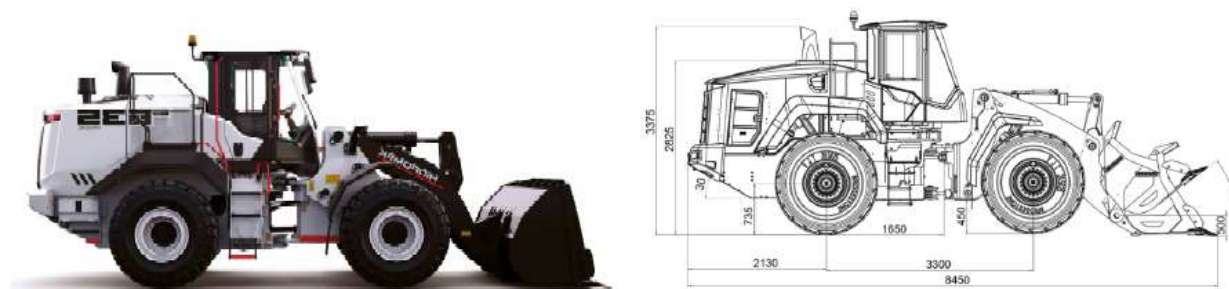


Figure 7. Wheel loader model: HIDROMEK HMK 635 WL. Extracted from D4.1.

³⁶ Nicolas P., Javier H., Marco M., Ivan V. et al. D4.1: Modelling and Digital Twin Data Link, (2022), BEEYONDERS project, <https://beeyonders.eu/>

The testing plan for the autonomous wheel loader involves systematic testing and evaluation to ensure safe and efficient operation, and to demonstrate the capabilities to perform the prototype objectives in the use cases of Road Construction and Tunnel Construction being the testing types to be executed and presented as follows:

1. Functional testing:
 - Verify that all autonomous functions of the wheel loader are working correctly, including navigation, obstacle detection and avoidance, load management, and communication with other systems (Digital Twin, Smart Wearables) or vehicles (Autonomous Aerial Vehicles).
 - Test the functionality of various modes, such as autonomous loading, unloading, and transport and the semi-autonomous mode.
2. Safety systems testing:
 - Evaluate the effectiveness of safety systems, such as emergency stop, collision avoidance, and object detection.
 - Conduct tests to ensure the wheel loader can detect and react appropriately to unexpected obstacles or hazards in its path.
 - Establish the data flow of current RTK location of the wheel loader to Digital Twin, and test in concordance with the location received from the smart wearables the distance calculation time lapse is safe enough.
3. Navigation and path planning:
 - Test the accuracy and reliability of the navigation system, including GPS or other positioning technologies (RTK) used by the autonomous wheel loader.
 - Assess the ability of the wheel loader to plan and follow optimal paths to reach designated locations or complete specific tasks.
 - Assess the optimized routes received from Agent Based Modelling simulations and integrate into the wheel loader's path planning.
 - Assess the manoeuvrability of the wheel loader inside the tunnel designed for the use case, as for executing the unloading path of debris it is a key element to design the operational steps of the machine.
4. Load management and efficiency:
 - Evaluate the autonomous wheel loader's ability to accurately and efficiently pick up, transport, and deposit debris / soil.
 - Test the load sensing capabilities to ensure proper load management and avoidance of overloading or instability.
5. Communication and coordination:
 - Test the communication and coordination with the drones and any other vehicle / device related to the task to be accomplished.
 - Evaluate the effectiveness of information exchange and task allocation between autonomous vehicles or human operators, and the data integration into Digital Twin database.

6. Environmental adaptation:
 - Assess the performance of the autonomous wheel loader in various environmental conditions, such as different types of terrain, lighting conditions, or weather (e.g., rain, snow).
 - Assess the performance of the autonomous wheel loader in the dusty environment proposed for the tunnel use case.
 - Test the wheel loader's ability to adapt to changes in the environment, such as adjusting navigation paths or behaviour in response to varying conditions.
7. Emergency procedures and Fail-Safe features:
 - Develop and test emergency procedures for critical situations, such as loss of control, system failure, or unexpected events.
 - Evaluate the effectiveness of fail-safe features and autonomous responses to emergency scenarios.
 - Evaluate the effectiveness of security stop if the distance to any human workers (wearing smart wearables or not) becomes dangerous.
8. Data logging and analysis:
 - Keep a record of all test sessions, including data logs, performance metrics, and any anomalies or issues encountered.
 - Analyze the collected data to identify areas for improvement, assess system performance, and validate the effectiveness of autonomous functions.
9. Safety measures and protocols:
 - Prioritize safety during all testing procedures and adhere to relevant safety guidelines and regulations.
 - Conduct risk assessments and implement safety measures such as physical barriers, warning systems, or designated safety zones when conducting tests in shared or public spaces. It is crucial to ensure that proper safety protocols and risk management strategies are followed throughout the testing process.
 - Engage with relevant stakeholders, including regulatory authorities, to ensure compliance with applicable regulations and guidelines for autonomous vehicles (this point is related to WP8 – Standardization).
10. Iterative testing and validation:
 - Continuously refine the testing plan based on the results and feedback obtained during each testing session.
 - Collaborate with stakeholders and consortium partners, including operators, engineers, and safety experts, to review and validate the testing plan and results.

Detailed Testing Plan

The detailed proposed testing plan is described below:

FUNCTIONAL TESTING		
DESCRIPTION		
Objectives: Ensure all autonomous functions operate correctly to perform tasks efficiently.		
Means of verification: Video recordings (simulations and tests), system logs showing successful task completion.	Methods: • Check multi-agent perception capabilities under controlled scenarios of progressive difficulties. • Evaluate HD multi-session semantic mapping algorithms with sensors for good and low visibility conditions. • Program the wheel loader to navigate predefined paths with obstacles (with and without GPS). • Simulate loading and unloading tasks and check learning capabilities.	Resources required: Virtual environment, controlled testing environments, obstacles, AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Conduct tests in simulation and in controlled environments resembling tunnel and road construction sites. Perception: - Indoor test: use of two perception stations (Lidar+camera) in indoor environment fusing data in real time: identification of people and velocity estimation. Test in laboratory with people moving around the stations with blind spots for the perception systems. - Preliminary outdoor test: use of two ground mobile platform with perception capabilities. Replication of indoor test (people and platforms moving). - Low visibility test: comparison of perception with RGB camera+Lidar vs Thermal camera+Radar in outdoor (variable climate conditions/dust/fog) - Virtual simulation of AWL-drone fused perception. Simulation of environment with objects and people. - TRL5 test: AWL-drone moving in teleoperation in a controlled area fusing information from moving people. Mapping: - Algorithm test vs data sets in the literature for basic functionality evaluation. - Recording of data sets with Lidar and Radar and comparison of results of algorithm application to each data set (urban environment).		

- Long-term multisession semantic mapping check. Use of a perception subsystem in a vehicle and fusion of information from different recording sessions in the same dynamic environment (urban environment).
- Virtual simulation of construction dynamic environment multi-session semantic mapping. Scenario replicating a road construction environment in low scale. Identification of piles of rubble/debris, machines, people and changing volumes in the environment (possibility to be added to the long-term map by exterior action).

Navigation:

- Trajectory generation in a known map. Use of a map from a controlled testing area (previously generated) including some non-transitable areas. Define different (A, B) points to check the trajectory definition according to the available information. At least one of the scenarios must involve the use of reverse gear.
- Navigation basic functionality virtual test. Replicate the testing area in a simulation environment and check navigation algorithms moving from A to B.
- Navigation obstacle avoidance virtual test. Include a static object in the basic simulation scenario in the theoretical trajectory and check avoidance strategy.
- Navigation response to person detection virtual test. Include a moving person with different trajectories and check the safety stop according to the defined distance thresholds.
- Navigation basic functionality TRL5 test. Replicate the virtual testing area in a controlled environment and check navigation algorithms and AWL controllers moving from A to B.
- Navigation obstacle avoidance TRL5 test. Include a static object in the testing area in the theoretical trajectory and check avoidance strategy.

Navigation response to person detection TRL5 test. Include a moving person during machine navigation with different trajectories and check the safety stop according to the defined distance thresholds.

Load/unload tasks:

- Loading task virtual testing. In a simulation environment with different debris piles (size and granulometry) check the algorithm is able to perform the loading task.
- Unloading task virtual testing. In a simulation environment including a virtual dumper in different relative positions to the AWL check the algorithm is able to perform the unloading task.
- Loading and unloading task TRL5 testing. Put a debris pile in the testing area and check the capability of the AWL to perform the loading task, navigate to a defined point and perform the unloading task.

Learning capabilities virtual test. Check the ability of the feedback loop and the learning algorithm to use operational data to improve the load algorithm in a simulation environment with different pile materials and tire-terrain behaviour.

SAFETY SYSTEMS TESTING		
DESCRIPTION		
Objectives: Assess the effectiveness of safety systems to prevent accidents.		
Means of verification: Video recordings of emergency stops, obstacle detection performance, and data flow analysis.	Methods: • Trigger emergency stop scenarios. • Introduce unexpected obstacles to evaluate collision avoidance and object detection. • Establish data flow between RTK location, Digital Twin, and smart wearables.	Resources required: Emergency stop mechanism, obstacles, RTK system, Digital Twin, smart wearables. AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Perform tests in simulated hazardous scenarios. Object detection testing. Ensure that the machine's sensors and detection algorithms can accurately identify and locate various types of obstacles, including stationary and moving objects. In a controlled testing area different obstacle, like rocks, construction materials, other machinery and moving people will be incorporated to check that they are correctly detected. Test will be carried out with the machine static and moving. Specific location systems will be used to assess the effective range and accuracy of the perception sensors. The metrics of the tests will be detection rate, false positives and reaction time. Collision avoidance. Verify that the machine's navigation and control systems can effectively avoid collisions after detecting obstacles. Introduction of an unexpected static object and check the effectiveness in recalculating the path and the ability to execute the needed manoeuvres (stopping, slowing down, rerouting). Simulate a person moving into the machine trajectory and check it stops in at the defined safety distance. The metric will be the success rate in the response according to the defined safety parameters and the response time in path recalculation. Stability Control. Ensure the machine remains stable under various operational conditions, performing tasks like lifting and moving in uneven terrain to prevent rollovers. Testing scenarios in the controlled area will be prepared including a variety of surfaces such as slopes, uneven ground, and loose materials. Navigation tests with different load weights will be performed. The metric evaluated in the tests will be the effectiveness of IMUs and other stability-related sensors in identify unbalance risk. Worker detection (wearables). – DT role is monitoring this safety distance: AWL sends location info to DT and receives a warning signal if distance to wearables is less than the defined threshold and check that the response is according to the specifications. The time needed in this communication process		

and the reaction time of the response of the AWL will be recorded.

Cybersecurity – Restricted access measures will be taken (not reported for security reasons)

NAVIGATION AND PATH PLANNING		
DESCRIPTION		
Objectives: Verify accurate navigation and path planning capabilities.		
Means of verification: GPS/RTK data analysis, video documentation of navigation paths, and simulation results.	Methods: - Analyse GPS/RTK data for accuracy. - Evaluate the wheel loader's ability to plan and follow optimal paths. - Integrate optimized routes from Agent Based Modelling simulations.	Resources required: GPS/RTK system, simulation software, Agent Based Modelling simulations. AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Test in simulated tunnel and road conditions. Preparation of the controlled and structured outdoor area with static obstacles (barriers and cones) and obtention of the reference map, with areas identified as non-transitable. Path planning tests: Benchmarking with literature navigation algorithms and datasets. Based on the map, define different sequences of points to travel around the area and check the path planning ability to define the route according to AWL driving characteristics. Navigation and trajectory execution tests: In the testing area define a sequence of points to travel. Execute the navigation with the basic AWL functionalities. Define a path imposed from an exterior system with varying degrees of complexity, including sharp turns and detours. Follow the required trajectory. In all scenarios (basic navigation and trajectory execution) record and analyse the accuracy of the path followed and the efficiency of the controls. Metrics: Comparison of the theoretical, GPS RTK and location algorithms (GPS denied) based trajectories. Time and fuel consumption to complete the path.		

LOAD MANAGEMENT AND EFFICIENCY		
DESCRIPTION		
Objectives: Ensure accurate and efficient load handling.		
Means of verification: Load sensor data, stability analysis reports, and video recordings.	Methods: • Monitor load sensing capabilities. • Assess stability during load handling.	Resources required: Load sensors, varying loads, stability analysis equipment. AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Perform load handling tasks with varying weights. Preparation of two piles of debris with different characteristics. Execution of autonomous loading tasks. Metrics for the test: evaluation of % of bucket use based on hydraulic pressure in the cylinders, volume of the debris in the bucket, loading time, tire slippage, average engine RPM, average fuel consumption.		

COMMUNICATION AND COORDINATION		
DESCRIPTION		
Objectives: Test communication with drones and other devices.		
Means of verification: Communication logs, task completion records, and database integration checks.	Methods: - Assess communication reliability and task allocation effectiveness. - Verify data integration into the Digital Twin database. - Evaluate communication with Autonomous Aerial Vehicles.	Resources required: Communication devices, drones, other vehicles/devices, Digital Twin database. AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Integrate communication protocols and coordinate tasks. In the controlled area, equipped with the necessary infrastructure to support wireless communications and a server platform where the digital twin replica can operate, deploy the AWL and drone. To establish baseline performance metrics for communications under optimal conditions a basic send/receive capability between the machine, the digital twin, and the drone will be tested over the network. The use of diagnostic tools to capture real-time data on communication performance and issues if foreseen to measure latency and data integrity. Additionally, the real-time data synchronization between the autonomous agents will be checked.		

ENVIRONMENTAL ADAPTATION		
DESCRIPTION		
Objectives: Evaluate performance in various environmental conditions.		
Means of verification: Performance evaluations, environmental condition simulation results.	Methods: - Assess performance in dusty tunnel environments. - Evaluate adaptability to changing environmental conditions. - Assess performance in different weather conditions: dry, rain...	Resources required: Testing environments with varied terrains and weather conditions. Perception kit (in the machine or as stand-alone)
Execution: Test in different terrains and weather conditions. Set up a perception station outdoor with a reference scenario (a static object and a moving person) under varying lighting conditions (dawn, daylight, dusk, and nighttime) and weather scenarios if available during the testing weeks (rain, fog). Dusty conditions will be simulated artificially. The main outputs of the perception will be compared system and differences evaluated.		

EMERGENCY PROCEDURES AND FAIL-SAFE FEATURES		
DESCRIPTION		
Objectives: Ensure effective response to critical situations.		
Means of verification: Emergency response logs, system behaviour analysis, and security stop checks.	Methods: • Trigger loss of control or system failure scenarios. • Evaluate fail-safe features and security stops. • Evaluate stop of the machine when a human is inside the near miss area.	Resources required: Emergency stop mechanism. AWL (machine, sensors, radar, lidar, cameras, industrial computer).
Execution: Simulate critical scenarios to test emergency procedures. - Remote emergency stop. Activate emergency stop buttons (remote and in the machine) and check the fail-safe response is activated. - Fail-safe response check (pending definition of fail-safe response). Measure response time and analyse machine reaction (mainly stopping distance) is according to the requirements. - Proximity sensors activation. Approach to the machine from different directions with the machine static and check that the fail-safe response is activated. Measure the distance to the machine. Move the machine close to an object and check the correct activation of the fail-safe response. Record and analyse the data to compare with the specifications. - Verification of ECU internal checking (to be defined with Hidromek). Provoke an internal failure and check that the ECU activates the fail-safe response. - Human Override (To be defined with Hidromek): Test the ability of a remote operator to take control in case the automated responses are considered insufficient.		

DATA LOGGING AND ANALYSIS		
DESCRIPTION		
Objectives: Maintain records and analyse collected data for performance assessment.		
Means of verification: Data logs, performance metric analysis reports, anomaly reports.	Methods: • Keep detailed records of test sessions. • Analyse performance metrics and anomalies.	Resources required: Data logging software, performance monitoring tools.
Execution: Log data during testing sessions and analyse collected data. • Equip the prototype wheel loader with sensors and data acquisition systems capable of capturing the related KPIs to the wheel loader, and assessment of the integration of the sensors with the vehicle's systems to gather real-time data during operations. • Ensure the sensors are accurate, reliable, and provide data in a format compatible with the data logging system. Quality of data analysis via data anomaly detection algorithms together with project global KPI "T12: Accuracy of collected data". • Evaluate the robustness data logging system onboard the wheel loader to capture sensor data continuously, ensuring the data logging system has sufficient storage capacity to handle large volumes of data generated during extended testing periods. • Establish a secure connection between the data logging system onboard the wheel loader and the digital twin database. Implement protocols for data transmission, ensuring data integrity and security during transfer. Verify compatibility between the data logging system's output format and the database requirements for seamless data integration.		

SAFETY MEASURES AND PROTOCOLS		
DESCRIPTION		
Objectives: Ensure adherence to safety guidelines and regulations.		
Means of verification: Safety protocol adherence checks, risk assessment reports, stakeholder feedback.	Methods: • Prioritize safety during testing procedures. • Engage with relevant stakeholders for compliance.	Resources required: Safety equipment, regulatory guidelines, stakeholder engagement.
Execution: Conduct risk assessments and implement safety measures. • Conduct a thorough risk assessment before the testing begins, identifying potential hazards related to the machine operations, environment, and human interactions. Specific warning lamp and/or buzzer during autonomous mode. • Develop specific safety measures based on the identified risks, including emergency response plans. Detecting people not participating in the testing or moving objects below 10 meters machine will stop. • Provide comprehensive safety training for all team members, focusing on the specific risks associated with the autonomous machine. • Hold a safety briefing before each testing session to review safety procedures, emergency contacts, and roles in case of an incident. • Determine the necessary PPE for the testing scenario, which may include helmets, gloves, eye protection, and high-visibility clothing. Ensure that all personnel always wear the required PPE during testing. • Clearly mark the testing area boundaries and ensure it is free from unnecessary equipment and objects. • Prepare or adapt the available emergency plan to effectively respond to emergencies. • Maintain records of all testing activities and incidents to improve future safety measures. • Review and feedback with stakeholders to continuously improve safety protocols based on testing experiences.		

ITERATIVE TESTING AND VALIDATION		
DESCRIPTION		
Objectives: Continuously improve system performance based on feedback.		
Means of verification: Testing iteration reports, stakeholder feedback, and performance improvement analysis.	Methods: • Review and refine the testing plan based on results and feedback. • Validate testing outcomes with stakeholders.	Resources required: Collaboration with stakeholders, testing iteration plans.
Execution: Collaborate with stakeholders for iterative testing. • Record results in all tests carried out in a test log which summarise the testing progress. • Explain reasons for software improvements and adaptation.		

Autonomous Maintenance Vehicle

As per the wheel loader, in the deliverable D4.1 – Modelling and data link interface, section 3.5, can be found the details about the machine description and the implication of this autonomous maintenance vehicle in BEEYONDERS use cases. The task T4.5: “Autonomous mobile robot for infrastructure maintenance” will focus on the development of a prototype of mobile robotic platform for the execution of multiple infrastructure maintenance operations, as defined in the associated use case. The prototype will incorporate suitable sensors, a pan-tilt unit and AI unit. This prototype will be used to implement and test the innovative software for this specific and complex scenario (developed in T4.2 & T4.3): autonomous navigation using 3D mapping; obstacle detection and avoidance; teleoperation for complex activities requiring low latency 4G LTE connectivity etc.

The testing plan for the autonomous inspection vehicle involves systematic testing and evaluation to ensure its safe and effective operation and the ability to perform the Road Maintenance use case objectives for this vehicle, and the following items have been identified:

1. Functional testing:
 - Verify that all autonomous functions of the inspection vehicle are working correctly, including navigation, obstacle detection and avoidance, sensor data collection, and analysis.
2. Sensor performance:
 - Evaluate the performance and accuracy of sensor used for inspection purposes (Retroreflectometer) or autonomous navigation required.
 - HW & SW integration of the retroreflectometer into the system.
 - Conduct tests to ensure that the sensors capture the necessary data with sufficient quality and resolution.
3. Navigation and path planning:
 - Test the accuracy and reliability of the navigation system used by the autonomous inspection vehicle.
 - Assess the vehicle's ability to plan and follow optimal paths to reach inspection locations and cover the required areas efficiently.
 - Validate the integration of the outdoor navigation suite from Tecnia Auto for basic road autonomous navigation.
4. Inspection task execution:
 - Evaluate the autonomous inspection vehicle's ability to perform the inspection task.
 - Validate the algorithms for lane detection and following of horizontal markings to correctly position the retroreflectometer when measuring horizontal signals.
 - Verify that the vehicle can accurately position itself in relation to the target objects or areas to be inspected.
 - Testing of the system operating independently in Tecnia's facilities or local available infrastructure (Bizkaia Connected Corridor, MUBIL, etc).

5. Data analysis and interpretation:
 - Develop algorithms or models to analyze and interpret the collected inspection data.
 - Test the effectiveness of the data analysis techniques and evaluate the accuracy of the inspection results or anomaly detection.
6. Communication and reporting:
 - Assess the autonomous inspection vehicle's ability to communicate inspection results or findings to relevant stakeholders or systems.
 - Test the integration of the system with the Digital Twin and integration with other systems (aerial autonomous vehicles according to Road Maintenance use case details).
7. Safety systems testing:
 - Evaluate the effectiveness of safety systems, such as emergency stop, collision avoidance, or fail-safe mechanisms.
 - Conduct tests to ensure the vehicle can detect and respond appropriately to unexpected obstacles or hazards during inspections.
8. Environmental adaptation:
 - Assess the performance of the autonomous inspection vehicle in various environmental conditions, such as different lighting conditions, weather (e.g., rain, snow), or varying surface conditions.
 - Test the vehicle's ability to adapt its inspection behaviour or parameters based on the environmental context.
9. Data logging and analysis:
 - Keep a record of all test sessions, including sensor data logs, inspection results, and any anomalies or issues encountered.
 - Analyze the collected data to assess the performance of the inspection vehicle, identify areas for improvement, and validate the effectiveness of inspection techniques.
10. Safety Measures and protocols:
 - Prioritize safety during all testing procedures and adhere to relevant safety guidelines and regulations.
 - Conduct risk assessments and implement safety measures, such as physical barriers, warning systems, or designated safety zones, when conducting tests in shared or public spaces and extrapolate to the future use case proposed.
 - Consider the specific requirements of the inspection tasks and engage with relevant stakeholders, such as domain experts or regulatory authorities, to ensure compliance with industry standards and regulations.

Detailed Testing Plan

The detailed proposed testing plan is presented below.

FUNCTIONAL TESTING		
DESCRIPTION		
Objectives: Ensure all autonomous functions operate correctly for effective inspection.		
Means of verification: Video recordings, system logs, successful task completion reports.	Methods: • Dynamic tests on the vehicle. • Program the vehicle to navigate predefined paths with obstacles. • Simulate sensor data collection and analysis.	Resources required: Controlled testing environment, obstacles, sensor suite.
Execution: Conduct tests in controlled environments resembling road conditions. • Prepare a controlled environment. • Check all the mounted sensors are connected. • Check the vehicle is with the battery fully charged. • Check the vehicle controller is fully charged. • Set the vehicle working conditions in manual control: ○ Simple or double Ackerman. ○ 2WD or 4WD. ○ Low or high speed. • Test if the vehicle works properly in manual control. • Autonomous navigation test: ○ Create a path where some obstacles are positioned. ○ Check if the vehicle can move along the path avoiding the obstacles. Collect all the required data during the test to check the data collection and analysis.		

SENSOR PERFORMANCE		
DESCRIPTION		
Objectives: Assess the accuracy and quality of sensor data.		
Means of verification: Sensor calibration reports, data quality analysis.	Methods: - Evaluate retroreflectometer performance and integration with vehicle systems. - Test sensor data collection in various lighting and weather conditions.	Resources required: Retroreflectometer, calibration tools, controlled testing environment.
Execution: Perform sensor calibration tests and controlled data collection. - Prepare a controlled environment. - Check the mounting of the retroreflectometer in the robot. - Check the mechanism to move the retroreflectometer in the robot. - Test the movement task to positioning of the sensor. - Test the full task of the retroreflectometer mounted in the vehicle. - Check the retroreflectometer in the following conditions: - In “simulated” rain - In hard light - In low light - Collect and assure the correctness of the generated data.		

NAVIGATION AND PATH PLANNING		
DESCRIPTION		
Objectives: Verify accurate navigation and path planning capabilities.		
Means of verification: GPS/RTK data analysis, navigation system logs.	Methods: - Analyse navigation system accuracy and reliability. - Evaluate path planning efficiency.	Resources required: GPS/RTK system, simulation software, test road environments.
Execution: Test navigation in simulated road environments. - Find a calibrated geodesic point. In case of these are not accessible, it will be required the creation of it using a RTK GPS base. - Check the accuracy of the vehicle GPS with the calibrated point. - Generate navigation plan based on recorded points: - Start recording a manual controlled trajectory. - Get navigation keypoints. - Send to the robot as new trajectory. - Compare the recorded trajectory with the followed. - Generate navigation plan selecting point directly in map. - Open maps and select the navigation keypoints. - Send to the robot as new trajectory. - Compare the followed trajectory with the keypoints.		

INSPECTION TASK EXECUTION		
DESCRIPTION		
Objectives: Validate the vehicle's ability to perform the inspection task accurately.		
Means of verification: Video documentation of inspection tasks, algorithm performance analysis.	Methods: - Assess lane detection and marking following algorithms. - Verify vehicle positioning accuracy during inspections.	Resources required: Real-world Road environments, lane marking samples, inspection targets.
Execution: Conduct tests in real-world and simulated road environments. - Move the robot to the inspection point. - Deploy the retroreflectometer to make the detection of the lane. - Get the measurements from the sensors: - Retroreflectometer - GPS position - Verify the data.		

DATA ANALYSIS AND INTERPRETATION		
DESCRIPTION		
Objectives: Develop and test algorithms for data analysis and interpretation.		
Means of verification: Algorithm performance metrics, anomaly detection reports.	Methods: • Develop algorithms for data interpretation. • Test algorithms on collected data for accuracy.	Resources required: Data analysis software, algorithm development tools.
Execution: Analyse collected inspection data with developed algorithms. • Prepare controlled environment: ○ Known GPS positions. ○ Known measurements sensors data. ○ Known expected results of the analysis. • Prepare invalid data to get the errors that can appear during the data recovery. • Test GPS data interpretation • Test sensor data interpretation		

COMMUNICATION AND REPORTING		
DESCRIPTION		
Objectives: Assess communication of inspection results to stakeholders.		
Means of verification: Communication logs, stakeholder feedback.	Methods: - Integrate inspection results with Digital Twin systems. - Test communication with drone.	Resources required: Digital Twin system, communication devices, drone.
Execution: Test communication systems integration. - Prepare controlled environment: - Known GPS positions. - Known measurements sensors data. - Known expected results of the analysis. - Test communication devices. - Test communication messages. Check correct visualization in the digital twin.		

SAFETY SYSTEMS TESTING		
DESCRIPTION		
Objectives: Ensure vehicle safety during inspections.		
Means of verification: Emergency response logs, collision avoidance system performance analysis.	Methods: • Test emergency stop and collision avoidance systems. • Evaluate response to unexpected obstacles.	Resources required: Emergency stop mechanism, obstacle simulation equipment.
Execution: Simulate hazardous scenarios and emergency situations. • Prepare controlled environment. • Emergency stop mechanism test: ○ Set the vehicle in the maximum working speed. ○ Press the emergency stop. ○ Check the correct and safe stop and the required distance. • Obstacle avoidance test: ○ Set an environment with avoidable obstacles. ○ Create a trajectory that goes through the obstacles area. ○ Check is the vehicle can go all the way without collide.		

ENVIRONMENTAL ADAPTATION		
DESCRIPTION		
Objectives: Assess vehicle performance in various environmental conditions.		
Means of verification: Performance evaluation reports, environmental condition simulation results.	Methods: - Evaluate performance in rain and varying light conditions. - Assess adaptability to changing surface conditions.	Resources required: Simulated environmental conditions, varied surface samples.
Execution: Test in different lighting, weather, and surface conditions. - Prepare controlled environment. - Rain conditions test: - Generate artificial rain situation. - Perform the tasks for what is prepared the robot. - Analyse the correctness of the recovered data. - Low light conditions test: - Generate low condition situation. - Perform the tasks for what is prepared the robot. - Analyse the correctness of the recovered data. - Surface conditions test: - Prepare multiple surface options environment. - Test the mobility of the robot and the difficulties to achieve the tasks. - Analyse the correctness of the dynamics of the robot to each surface.		

DATA LOGGING AND ANALYSIS		
DESCRIPTION		
Objectives: Maintain records and analyse collected data for performance assessment.		
Means of verification: Data logs, performance metric analysis reports, anomaly reports.	Methods: • Maintain detailed records of test sessions. • Analyse performance metrics and anomalies. • Assess integration of data logging with Digital Twin database.	Resources required: Data logging software, performance monitoring tools.
Execution: Log data during testing sessions and analyse collected data. • Prepare a data set that provides all the possible logging cases. • Test of input new data in logging. • Check if data input was made correctly. ○ Also, the fails messages if an incorrect data has been added • Check the status of the previously stored data: ○ If any new add modify it ○ The correctness of it in time • Test the data to replicate the logged data in digital twin: ○ Check accessibility ○ Check consistency		

SAFETY MEASURES AND PROTOCOLS		
DESCRIPTION		
Objectives: Ensure safety during testing procedures and compliance with regulations.		
Means of verification: Safety protocol adherence checks, risk assessment reports, stakeholder engagement feedback.	Methods: • Conduct risk assessments and implement safety protocols. • Adhere to relevant safety guidelines and regulations.	Resources required: Safety equipment, regulatory guidelines, stakeholder engagement sessions.
Execution: Implement safety measures and engage with stakeholders. • Generate first risk assessment list. • Compare the assessment with the safety guidelines and regulations. • Adapt the risk assessment list. • Based on the verification tests: ○ Check the status of the risk ○ Update the risk assessment priority • Re-evaluate the risk assessment. • Update the safety protocols.		

Autonomous Aerial Vehicles

Autonomous aerial vehicles (drones) are going to be deployed in three use cases in BEEYONDERS: Tunnel construction, Road construction, Building construction and Road Maintenance. As the requirements in these use cases are slightly different, two different drone models are going to be developed: DJI MATRICE 300 RTK and CATEC C750. For technical details about these models, please refer to D4.1, section 3.1, and section 3.2, where it also can be found their roles and objectives to be achieved in the different use cases.

According to the objectives and developments proposed for the drones to be demonstrated in the use cases, the testing plan covers the following aspects and validations:

1. Flight environment assessment (outdoor):
 - Assess the weather conditions, including wind speed, precipitation, and visibility, to determine if it is suitable for safe flight.
 - Identify and manage any airspace restrictions or regulations in the area where the testing will take place.
 - Choose an appropriate location for testing, considering factors such as open space, absence of obstacles, and minimal interference from other people or objects.
2. Flight environment assessment (indoor):
 - Assess the internal environment (tunnel and building) and visibility, to determine if it is suitable for safe flight.
 - Identify and manage any airspace restrictions or regulations in the area where the testing will take place.
 - Choose an appropriate location for testing, considering factors such as open space, absence of obstacles, and minimal interference from other people or objects.
 - Simulate the dusty environment proposed for the tunnel use case and assess the drone performance with this environment.
3. Flight performance and stability:
 - Conduct basic flight tests to assess the drone's stability, maneuverability, and responsiveness.
 - Test various flight modes and functions (e.g., altitude hold, auto-hover, return-to-home) to ensure they are working as expected.
 - Evaluate the drone's ability to maintain stability in different wind conditions.
4. GPS and navigation testing:
 - Assess the accuracy and reliability of the GPS system by conducting flights in different locations and monitoring the drone's positioning.
 - Test autonomous flight features, such as waypoint navigation to ensure accurate performance.

5. Non-GPS and navigation testing:
 - Assess the accuracy and reliability of the non-GPS system by conducting flights in different locations and monitoring the drone's positioning.
 - Test autonomous flight features, such as waypoint navigation to ensure accurate performance.
6. Battery and power management:
 - Evaluate the drone's flight time and battery performance under different conditions.
 - Test the low battery warning and auto-return-to-home features to ensure the drone can safely return to the landing location when the battery is low.
7. Payload and imaging system:
 - If the drone is equipped with a camera or other payload, test the functionality and image/video quality.
 - Assess the stability of the Lidar / photogrammetry camera gimbal or payload mount to ensure steady and accurate capture.
8. Emergency procedures:
 - Develop and test emergency procedures for scenarios such as loss of signal, motor failure, or in-flight emergencies.
 - Evaluate the effectiveness of fail-safe features, such as return-to-home or landing in case of a critical situation.
9. Documentation and data collection:
 - Keep a record of all flight tests, including flight logs, maintenance records, and any issues encountered during testing.
 - Collect data on flight parameters, such as altitude, speed, battery consumption, and GPS coordinates, for further analysis. Perform the same data collection for non-GPS environment flights.
 - Test accuracy of Lidar scanning and stability.
 - Test accuracy of photogrammetry scanning.
 - Validate the required REST - API to transfer live data to Digital Twin database (as location) and transferring the processed point clouds or photogrammetry in .pcd format.
 - Assessment of time required, and resources spent for the processing of the cloud points and photogrammetry before exchanging the final scanning to the Digital Twin.
10. Safety measures:
 - Follow local regulations and safety guidelines during all testing procedures.
 - Ensure proper safety equipment, such as propeller guards or protective barriers, is used when necessary.

- Have a spotter or observer, if necessary, according to use case locations requirements, present during flights to assist with situational awareness and safety.

11. Iterative testing and evaluation:

- Continuously refine the testing plan based on the results and feedback obtained during each testing session.
- Document any improvements or modifications made to the drone or testing procedures.

Detailed Testing Plan

According to the objectives and developments proposed for the drones to be demonstrated in the use cases, the testing plan covers the following aspects and validations:

FLIGHT ENVIRONMENT ASSESSMENT (OUTDOOR)		
DESCRIPTION		
Objectives: Ensure safe flight in outdoor environments.		
Means of verification: Weather reports, airspace maps, testing location assessment reports.	Methods: • Check airspace restrictions and regulations. • Select appropriate locations for testing, considering space and obstacles. • Assess weather conditions, including wind speed and visibility. • Simulate environment for road use cases.	Resources required: Weather monitoring tools, airspace regulation databases, simulation equipment.
Execution: To carry out the outdoor flights and to ensure a satisfactory result, the following steps shall be carried out: • The first step to obtaining a Flight Authorization for the ‘specific’ category in a different country from the one the operator is registered at, is applying for the permit in its State of Registration. To do that, it is necessary to find a very similar scenario which shares the following characteristics with the target scenario: ○ Ground risk and mitigations to be applied. ○ Airspace category and air risk. ○ Adjacent area categorisation. Once the Operational authorization has been obtained, the operator shall follow the procedures established in the desirable State to obtain a cross-border authorization. During this step, the actual location will be shared, and the main task will be adapting the permits to the local conditions. • Select a safe and obstacle-free flying area. In addition, comply with the restrictions of the flight category for which permissions have been obtained.		

- Perform an observation of the historical weather history of the selected area. In addition, on the days before the experiments, as well as on the day itself, assess the weather in the selected area as to whether the weather is favorable for carrying out the experiments.

Ensure that good coverage of the different radio links used can be guaranteed in the flight area: the safety pilot's radio link, which is the most important one, and the link between the ground computer and the drone.

FLIGHT ENVIRONMENT ASSESSMENT (INDOOR)		
DESCRIPTION		
Objectives: Ensure safe flight in indoor environments.		
Means of verification: indoor maps, testing location assessment reports.	Methods: • Monitor environment conditions (e.g., visibility). • Check owner safety restrictions and regulations. • Select appropriate locations for testing, considering space and obstacles. • Simulate dusty environment for tunnel use case. • Evaluate cooperation between drone and Autonomous Wheel Loader in the tunnel use case.	Resources required: indoor safety databases, simulation equipment.
Execution: to carry out the indoor flights and to ensure a satisfactory result, the following steps shall be carried out: • Comply with and ensure compliance with all restrictions and guidelines given by the owner of the space where the experiments will be carried out. Such as fencing and signposting to cut off the flight zone or notifying all persons in the vicinity of the experiments to be carried out. • Assess that there are no small objects or obstacles that could affect the flight indoors, as well as objects that could be moved by the generated airflows by the aircraft (such as papers or light objects). Conduct of flights in manual mode by the safety pilot in the flight area in such a way as to assess that the surrounding conditions do not affect the flight of the aircraft itself or its behavior or stability. • Ensure that good coverage of the different radio links used can be guaranteed in the flight area: the safety pilot's radio link, which is the most important one, and the link between the ground computer and the drone. • Ensure stable and continuous communication between the drone and the wheel loader during the whole process of cooperation in the case of use.		

FLIGHT PERFORMANCE AND STABILITY		
DESCRIPTION		
Objectives: Assess drone's stability, manoeuvrability, and responsiveness.		
Means of verification: Flight test reports, stability analysis.	Methods: • Test various flight modes and functions. • Evaluate drone's stability in different environment conditions (visibility).	Resources required: Drone, flight controller, environment measurement tools.
Execution: To assess the stability of the aircraft in the different flight environments, tests shall be carried out to evaluate the angle limits that the aircraft takes in that environment, in particular, certain angles in the roll and pitch axes may not be exceeded to ensure safe and stable flight. In addition, the maximum speeds of the aircraft to be used should be considered: • DJI M300 (Road Construction Use Case): ○ Max Speed: 15 m/s. ○ Max Permissible Wind Speed: 12 m/s. ○ Max Pitch Angle: 30°. • CATEC 750 (Road Maintenance Use Case): ○ Max Speed: 17 m/s. ○ Max Permissible Wind Speed: 10 m/s. ○ Max Pitch Angle: 25°. • CADRIN (Tunnel Construction & Building Construction Use Cases) ○ Max Speed: 5 m/s. ○ Max Pitch Angle: 25°. ○ It should be noted that to ensure safe flight in environments with reduced visibility conditions, the drone will have an artificial lighting system on board.		

GNSS AND NAVIGATION TESTING		
DESCRIPTION		
Objectives: Assess accuracy and reliability of GNSS system.		
Means of verification: GNSS data analysis, navigation performance assessment.	Methods: Test autonomous flight features like waypoint navigation.	Resources required: GNSS module, flight controller, test locations.
Execution: The accuracy and reliability of the aircraft used for outdoor use cases is directly related to the GNSS receiver. In the case of the DJI M300, it has an on-board GNSS RTK receiver. The accuracy of this receptor is: - Hovering Accuracy: $\pm 0.1\text{m}$ vertical and horizontal. - RTK Positioning Accuracy: $1\text{ cm}+1\text{ ppm}$ (Horizontal); $1.5\text{ cm} + 1\text{ ppm}$ (Vertical). Where ppm is part per million. In the case of CATEC 750, it has the HERE3 GNSS receiver on-board. The accuracy of this receptor is: - GNSS Positioning accuracy: 3D FIX: $\pm 2.5\text{m}$. To assess the performance of the outdoor navigation, a GNSS route shall be defined and followed in flight. Relative trajectory tracking errors shall be assessed.		

NON-GNSS AND NAVIGATION TESTING		
DESCRIPTION		
Objectives: Assess accuracy and reliability of non-GNSS navigation system.		
Means of verification: Positioning data analysis, navigation performance assessment.	Methods: Test autonomous flight features in non-GNSS environments.	Resources required: Non-GNSS navigation sensors, flight controller, test locations and site map.
Execution: Evaluate non-GNSS navigation of the aerial vehicle. The positioning system for GNSS-Denied environments has been tested on CATEC's indoor testbed. This testbed is equipped with a motion capture system with millimeter accuracy operating at a high frequency (around 100Hz). For the evaluation of the indoor localization algorithm, the positioning system described above shall be used. Mainly, this evaluation will be done through two metrics in the trajectory: Absolute Pose Error (APE), that measures the error in position and orientation along the entire trajectory, and Relative Pose Error (RPE), that measures the error in position and orientation at each instant of the trajectory. In addition to that, the main sensor of the GNSS-Denied positioning system is the Ouster OS0 whose accuracy is around +/- 3cm.		

BATTERY AND POWER MANAGEMENT		
DESCRIPTION		
Objectives: Evaluate flight time and battery performance.		
Means of verification: Battery usage data analysis, emergency feature testing.	Methods: • Test low battery warning and auto-return-to-home features. • Test battery duration for different fly modes and conditions.	Resources required: Battery packs, flight controller, battery management system.
The battery and flight time is a function of different external and internal factors, so even if the datasheet reflects some values, these may differ. Some of the factors that can affect this are: payload weight, wind speed, ambient temperature, etc. It should be noted that in the case of ambient temperature, it greatly affects the performance of drone batteries. If temperatures are low, the flight time will be much shorter. According to the datasheet of each aircraft and from the default battery set and payload configuration, the flight time of each aircraft is as follows: • DJI M300: ~30 min. • CATEC 750: ~15 min. • CADRIN: ~ 8 min. With regard to low battery warnings or safety mode, the default safety modes associated with each of the on-board autopilots have been configured. These safety modes can be return to home or auto landing, depending on the use case.		

PAYLOAD AND IMAGING SYSTEM		
DESCRIPTION		
Objectives: Assess functionality and image quality of payload.		
Means of verification: Image quality assessment, payload stability analysis.	Methods: - Evaluate stability of gimbal or payload mount. - Evaluate image quality obtained while flying. - Evaluate method of measuring retroreflectivity of road signals with drone and vegetation detection.	Resources required: Camera or other payload, gimbal, imaging software.
All on-board cameras will be stabilized through a 3-axis gimbal. This will ensure the stability of the camera and the quality of the images. Each aircraft is equipped with the following gimbals for image stabilization: - DJI M300: Zenmuse H20T (with RGB camera). - CATEC 750: Gremsy T3 V3(tentative) (with LiDAR and RGB camera). - CADRIN: Gremsy G-Hadron (with thermal camera). Regarding the method of measuring the reflectivity of traffic signals with a drone and vegetation, the following setup of on-board sensors has been selected: a solid-state lidar and an RGB camera. - The Blickfeld lidar has a built-in field that measures the relative retroreflectivity of the objects that the laser collides with. This will allow the relative retroreflectivity of traffic signals to be evaluated. - The on-board RGB camera has been selected so that with the resolution of the sensor and the flight altitude, traffic signs and encroaching vegetation can be clearly seen.		

EMERGENCY PROCEDURES		
DESCRIPTION		
Objectives: Develop and test emergency procedures.		
Means of verification: Emergency response testing reports, fail-safe feature performance analysis.	Methods: • Test fail-safe features like return-to-home or safe emergency landing. • Test response of the drone against unexpected obstacles.	Resources required: Simulation equipment, emergency response protocol documentation.
When applying for flight permissions to conduct experiments in outdoor use cases, it is necessary to carry out and submit a procedure for different types of emergencies, which is detailed in the documentation provided to the regulatory authorities. Some of these procedures are for: Battery Failure, Communication Loss, Control loss of the aircraft, In-flight power loss, In-flight fire, among others. When experiments are carried out indoors, a safety analysis is also made of the risks and/or emergencies that may exist and the mitigation measures to be taken, coordinated with the owner of the space. In addition to this, drone autopilots themselves have internally defined procedures for different low-level errors, such as problems with magnetometer or similar.		

DOCUMENTATION AND DATA COLLECTION		
DESCRIPTION		
Objectives: Maintain records of flight tests and data collection.		
Means of verification: Flight logs, maintenance records, data analysis reports.	Methods: • Collect flight parameters data for analysis. • Evaluate communication with DT database.	Resources required: Documentation templates, flight data recorder.
There shall be two types of data logging systems on aircraft: On the one hand the logging system of the autopilot itself and on the other hand a logging system of the onboard software. • Autopilot Logging: It shall manage all low-level aircraft information such as commands sent to the engines, raw data from the Inertial Measurement Unit (IMU), etc. and it will be stored in the autopilot itself. • Onboard Software Logging: All relevant flight information, such as position estimation, captured images, etc. will be stored in the on-board computer. In addition to this, all flights will be documented in flight plan documents with all necessary information such as expected behaviour, errors, etc. To evaluate the communication with the DT database, a server mock-up has been created that implements the messages defined in document D4.1.		

SAFETY MEASURES AND PROTOCOLS		
DESCRIPTION		
Objectives: Ensure compliance with regulations and safety guidelines.		
Means of verification: Compliance documentation, safety equipment checklist.	Methods: • Ensure proper safety equipment during flights.	Resources required: Safety equipment, regulatory guidelines.
As a safety protocol, it is important to make sure that you have a work plan that is known to everyone involved. In addition to this, ensure the correct use of Individual Protection Equipment (PPE), carry out safety checks prior to work, protect the environment, etc. In addition, select a safe and unobstructed flying area, where the restrictions of the category of flight in which the flight permit has been obtained are fulfilled.		

ITERATIVE TESTING AND EVALUATION		
DESCRIPTION		
Objectives: Continuously refine testing plan based on feedback.		
Means of verification: Testing plan updates, modification documentation.	Methods: Review testing results and feedback.	Resources required: Testing team collaboration tools, documentation system.
All developed functionalities have been tested following an iterative process. First, they are defined and created in simulation environments. After that, and once the correct functioning has been assessed, they are evaluated in CATEC's indoor testbed. Subsequently, when it has been evaluated that the performance is as expected, flight tests are attempted in environments similar to the final use case and representative. At each of these steps, the algorithms are evaluated and refined and, if necessary, the process goes back to the previous step to fix certain functionalities. Finally, a series of experiments will be carried out in the real use case environment, in order to evaluate whether the proposed system architecture is valid. After these experiments, a report of errors and improvements will be generated, and the process will return to the beginning of this to incorporate them into the system.		

Additive Manufacturing

Under the objective of test and demonstrate new, environmental-friendly, additive manufacturing solutions, this task focuses on developing 3D printable caisson with bio-diversity promotive function. We first establish a pre-caisson of 3D printable caisson with biodiversity promotive feature. And then we establish its manufacturing procedure by concrete 3D printing. On material perspective, a 3D printable material will be developed. To boost the bio-receptive function of the caisson, a second material and its related applying method will be developed. The performance of the material will be accessed by lab testing and on lab fabricated samples and on in-situ 3D printed elements. A demonstrator will be built to validate the structural performance and biodiversity promotive feature of 3D printed caisson. LCC and LCA analysis will be done, which benchmarks the cost and environmental impact of 3D printable caisson with traditional reference caisson.

The testing plan for additive manufacturing involves a systematic approach to evaluate the quality, performance, and reliability of printed parts. The proposed testing plan covers the following aspects:

1. Material performance Testing:
 - Evaluate the mechanical properties, such as tensile strength, elasticity, and hardness, of the materials used in additive manufacturing.
 - Evaluate fresh properties of fiber reinforced 3DPC.
 - Evaluate the deformational properties of fiber reinforced 3DPC, elastic modulus, shrinkage, and creep.
2. Material processing performance testing (to evaluate the compatibility of the material with 3D printing process):
 - Pumping test for pumpability of fiber reinforced 3DPC.
 - Application process validation for bio-receptive material.
3. In-situ performance of material:
 - 3D printing trial to evaluate the printability (yield strength of material from printing head) and buildability (stability of freshly printed structure).
 - Mechanical properties of printed elements from a structural monitoring plan established.
 - Bio-receptivity of printed element with second material applied on its surface, submerged in Spain, Acciona's testing center.
4. Dimensional accuracy:
 - Test the dimensional accuracy of printed parts by comparing them to the intended design specifications.
 - Assess the precision and repeatability of the 3D printing process in producing parts with the desired geometries.

5. Surface finish and texture:
 - Evaluate the surface finish and texture of printed parts to ensure they meet the required standards.
 - Conduct visual inspections and utilize surface metrology techniques, such as profilometers or 3D scanners, to measure roughness or texture characteristics.
6. Microstructural analysis:
 - Assess the presence of defects, such as voids, porosity, or inclusions, that could affect the integrity and performance of the printed components.
7. Environmental and chemical resistance:
 - Test the resistance of printed parts to various environmental conditions, such as temperature, humidity, or UV exposure.
 - Assess the chemical resistance of printed parts to specific substances or corrosive environments relevant to the intended application (maritime environment, both submerged in the sea water and not submerged).
8. Functional performance testing:
 - Evaluate the functional performance of printed parts by subjecting them to real-world operating conditions and functional tests. In fact, two aspects are to be evaluated: the structural and the bio receptive performance.
 - Test the compatibility and fit of printed parts within the overall system or assembly.
 - Evaluate the connection of the structural monitoring components to be installed on the caisson with the Digital Twin database.
9. Process optimization:
 - Conduct experiments to optimize the additive manufacturing process parameters, such as layer thickness, print speed, or temperature, to achieve the desired part quality and performance.
 - Monitor and document the effects of process parameter variations on the printed parts.
 - Establish an IoT framework during the manufacturing of the material to obtain the optimal composition for the printable material, testing the connection of the IoT data at the factory with the Digital Twin cloud database. Develop the framework to execute the optimisation algorithms.
10. Safety measures and protocols:
 - Follow relevant standards and guidelines specific to additive manufacturing, such as ASTM F42 for additive manufacturing technologies.
 - Ensure proper safety measures, such as propeller guards or protective barriers, is used when necessary, during the printing of the elements or the manufacturing of the material.

Detailed Testing Plan

The detailed testing plan cover the next aspects as below:

MATERIAL PERFORMANCE		
DESCRIPTION		
Objectives: Evaluate mechanical, deformational properties and durability of materials.		
Means of verification: - Process optimization reports, quality control data, video recordings and photos of the printing process. - Lab test reports.	Methods: - Perform compression and bending tests on lab cast samples. - Assess to fresh properties, such as workability and setting time on lab prepared material. - Evaluate deformational properties like elastic modulus, shrinkage, and creep on lab cast samples. - Validate lab method to characterize the bio-receptiveness of concrete, and apply it for evaluating bio-receptiveness of lab prepared bio-receptive material.	Resources required: Testing equipment (e.g., mechanical testing machine), material samples.
Execution: Conduct standardized tests on material samples. Materials development: Developed initial 3D printable structural concrete that meets structural requirements and bio-receptive concrete formulations capable of supporting marine life. Ongoing refinements based on 3D printing trial feedback.		

MATERIAL PROCESSING		
DESCRIPTION		
Objectives: Assess compatibility of materials with the 3D printing process		
Means of verification: Slump test, Pumping test results, application process validation reports.	Methods: - Determine the material production procedure (mixing procedure, batch size etc) of 3DPC. - Evaluate the pumpability of fibre-reinforced 3DPC. - Validate the application process for bio-receptive material.	Resources required: 3D printing facilities, experimental materials, sensor technology for real-time data capture. - Laboratory testing facilities, material scientists, a variety of raw material supplies.
Execution: Conduct 3D printing trials to evaluate pumpability and application process. - The pumpability of structural material has been validated through 3D printing trial in 2023 for the first version of the material. The minimum diameter of pumping hose was determined to be 75mm. The pumpability of the refined version of structural material were validated through 2nd trial organised in April 2024. - With progress on 3D printable caisson design, the application process for bio-receptive material will be 3D printing, too. We aim to handle two materials through one single printing process: a steel fibre reinforced 1K material for structural printing, the same material with bio-receptive agent additive at printing head (2K) for bio-receptive wall printing. - So far, the batch size, mixing procedure and the target mixing load has been optimized for the 2nd version of structural material.		

MATERIAL IN-SITU PERFORMANCE		
DESCRIPTION		
Objectives: Evaluate In-situ printability, buildability, mechanical properties, and bio-receptivity of materials.		
Means of verification: Printing trial reports, structural monitoring data, bio-receptivity test results.	Methods: • Conduct 3D printing trials to confirm the mixability, printability and buildability. • Cut samples from 3D printed elements and do mechanical test to quantify the impact of fibre orientation on flexural behaviour. • Access to the bio-receptiveness of sample cut from bio-wall through lab test.	Resources required: 3D printing equipment, structural monitoring sensors, real conditions testing centre, lab mechanical testing equipment.
Execution: Conduct 3D printing trials and testing on material produced and exposed to in-situ environment. • Completed 3D printing trials showing promising workability, mechanical performance test of printed elements undergoes. Design adjustments are being finalized to ensure successful submersion tests, with specific focus on integrating learnings from WP6 tasks.		

DIMENSIONAL ACCURACY		
DESCRIPTION		
Objectives: Ensure printed parts match intended design specifications.		
Means of verification: Dimensional measurement reports, comparison to design specifications.	Methods: • Measure dimensions of printed parts using precision tools. • Assess precision and repeatability of printing process.	Resources required: Precision measurement tools, design specifications.
Execution: Compare printed parts to design specifications. The dimensional accuracy of the printed elements will be evaluated according to the following techniques: • Material consistency: While printing, it must be ensured that the mortar mix is consistent and flows correctly through the printer nozzle. • Using digital calipers: For measuring small dimensions with high precision. • Employing laser scanners to the finished caisson: For creating a 3D representation of the printed object to compare against the CAD model. • Dimensional visual inspection: Measure the printed caisson at various points to check for adherence to design specifications. Measure lengths, widths, and heights at multiple points. • Tolerance analysis: Compare measured dimensions against the tolerances specified in the design.		

SURFACE FINISH AND TEXTURE		
DESCRIPTION		
Objectives: Evaluate surface finish and texture of printed parts.		
Means of verification: Inspection reports, surface metrology data.	Methods: • Visual inspection for surface quality. • Utilize surface metrology techniques for texture analysis.	Resources required: Visual inspection tools, surface metrology equipment.
The following methods are proposed to measure the correct surface finish and texture: • Surface roughness: Measure the surface finish to ensure it meets the required design. • Texture analysis: Evaluate the surface texture patterns, which might include periodicity, directionality, or random roughness. • Visual inspection: Conducting a thorough visual inspection under different lighting conditions to identify any irregularities, defects, or inconsistencies. • Tactile assessment: Feel the surface manually to detect any unexpected rough spots or smooth areas.		

ENVIRONMENTAL AND CHEMICAL RESISTANCE		
DESCRIPTION		
Objectives: Evaluate resistance of printed parts to environmental and chemical conditions.		
Methods: • Subject samples to environments with various temperature or/and relative humidity and do mechanical test after exposures. • Chloride penetration test.	Methods: • Subject samples to environments with various temperature or/and relative humidity and do mechanical test after exposures. • Chloride penetration test.	Methods: • Subject samples to environments with various temperature or/and relative humidity and do mechanical test after exposures. • Chloride penetration test.
Execution: Expose printed parts to various conditions. Environmental performance will be addressed through chloride penetration test and mechanical tests on samples cured in environments with different temperature/relative humidity.		

STRUCTURAL PERFORMANCE		
DESCRIPTION		
Objectives: Characterize the performance of printed structures.		
Means of verification: Test reports.	Methods: - Establish FEM model of 3D printable design and do stress analysis to ensure the integrity of structure. - Scale down 3D printable caisson design to representative and easy handling prototype, do structural analysis by FEM. 3D printing trial(s) to print out prototype. - Measure dimensions of printed parts using precision tools. - Integration of sensor on the prototype to monitor deformation of printed elements during printing, transport, and submersion. - Assess bio-receptivity of printed elements under real conditions.	Resources required: CAD software and structural design expert, concrete batching plant and 3D printing system, 3DPC material, deformation sensors, water-proof camera, image analysis, submerging sites.
- Structural deformation monitoring, bio-diversity development monitoring - Significant progress made on prototype design. The prototype consist of a 1: 4 scale down of 3D printable caisson design with simplified geometry and without bio-receptive walls, and several 3D printable independent bio-receptive walls. - Identified potential test sites and begun negotiations for permissions for submersion test and bio-diversity development monitoring.		

3D PRINTING PROCESS OPTIMIZATION		
DESCRIPTION		
Objectives: Adapt additive manufacturing process parameters to maritime caisson construction.		
Means of verification: Optimization reports, part quality assessment.	Methods: • Vary parameters such as layer thickness, print speed and check their impact on the properties of printed element. • Develop and integration of sensor solutions to improve the process efficiency.	Resources required: 3D printing equipment, process monitoring tools. Sensors and datalogger.
Execution: Conduct tests and trials to optimize process parameters. The relevant process parameters that will be evaluated are the following: • Testing different mortar compositions for marine conditions. • Optimize the mix for workability, setting time, and strength. • Print small-scale prototypes to validate design and material behaviour and monitoring of the printing process for consistency and address any issues in real-time. • Ensure proper curing of the printed structures to achieve maximum strength and durability. This will include controlled moisture and temperature conditions.		

SAFETY MEASURES AND PROTOCOLS		
DESCRIPTION		
Objectives: Ensure safety during additive manufacturing process.		
Means of verification: dedicated session on Health and Safety risk assessment in material development and manufacturing process reports.	Methods: • risk assessment before each 3D printing trial and demonstration. • Implement safety measures as necessary during printing. • Document learnings and give safety guidelines on 3D printable caisson construction.	Resources required: access to safety guide documents and papers, risk assessment tool and safety equipment.
Adhere to relevant standards and guidelines. Risk assessments was done before each 3D printing trial.		

Exoskeleton

The key innovation in the exoskeleton technology that BEEYONDERS will develop, is the adaptation, customization and enhancement of an active exoskeleton that supports manual material handling and repetitive tasks and aligning these developments and refinements to the very specialized demands that arise in the construction industry. Work-related Musculoskeletal Disorders (MSD) are an important cause of functional impairments and disability among civil construction workers with the lower back being the most overloaded body region and the shoulder as the second. A primary goal is to prevent this risk of injury by providing the worker with an assistive exoskeleton for lifting and handling heavy objects (up to 25kg), and therefore by reducing their perceived effort. IIT would provide the required technology, knowhow, and assistive strategies to ergonomically reduce the risk of injury.

Thus, the final goal is to enhance workers' working condition, health, and safety by providing a robotic support exoskeleton that will reduce the biomechanical loading during manual handling. The use of exoskeletons could also improve production efficiency by enhancing the workers performance, reducing the execution time, as well as by decreasing absenteeism and related costs. With the adoption of such device, both men, women, and aged workers could benefit by reducing working related difficulties. The use of this device not only reduce the workers fatigue but also mitigates the risks of developing work-related MSD.

This prototype will be demonstrated in the building use case. The testing plan proposed for the exoskeleton is designed according to the following assessment phases:

1. Basic Functionality, power, and safety demonstration (TRL4)
 - Basic Functionality characteristics:
 1. Verify that all components and subsystems of the exoskeleton are functioning properly.
 2. Test the range of motion, joint flexibility, and force/torque generation capabilities of the exoskeleton.
 3. Assess the responsiveness and accuracy of control mechanisms and user interface (if any).
 - Power and energy management:
 1. Assess the exoskeleton's power source and energy management system.
 2. Test the battery life, charging duration, and overall power efficiency of the exoskeleton.
 - Safety and emergency procedures:
 1. Evaluate the exoskeleton's safety features, such as emergency stop mechanisms or fail-safe systems.

The results of these are included in D5.1.³⁷
2. Main exoskeleton's features demonstration (TRL5):

³⁷ Christian Di Natali; Sergio Leggieri; Pinar Sencandan; Vasco Fanti; Paolo Leone. D5.1: Exoskeleton Prototype (2023). BEEYONDERS project, <https://beeyonders.eu/>

- Exoskeleton performances:
 1. Show the exoskeleton's ability to provide support for carrying heavy loads and reduce strain on the user.
 2. Show the exoskeleton's ability to support load carrying tasks.
- Mobility and manoeuvrability:
 1. Show the exoskeleton's use on various surfaces, slopes, or uneven surfaces.
 2. Show the user's ability to walk and climb stairs with the exoskeleton that could be necessary to go through during the building use case.
 3. Show that the impact of the exoskeleton on user mobility is minimum.
- 3. Onsite testing (TRL6)
 - Building workers related activities supported by exoskeleton:
 1. Testing on the Manual Material Handling (MMH) activities (back-support Exo) verifying the strain reduction on the user, by acquiring data also from users without exoskeleton to compare these data with those from workers wearing the exoskeleton (qualitative questionnaire).
 2. Testing on the upper limbs' activities (shoulder-support Exo) verifying the strain reduction on the user, applying the same considerations discussed in the previous point (qualitative questionnaire).
 3. Testing on combined activities (nice to have, it depends by the selected activities, please refer to dedicated Deliverable about onsite test) verifying the strain reduction on the user with same considerations of the points before.
 - User interface and control system:
 1. Evaluate the intuitiveness and usability of the exoskeleton's control interface (qualitative questionnaire).
 - User experience:
 1. Evaluate the proper fit and comfort for the user (qualitative questionnaire).
 2. Show adjustability and adaptability of the exoskeleton to different body sizes and shapes during use case demonstration.
 3. Evaluate the freedom of movement (qualitative questionnaire).
 - Training
 1. Propose a familiarization procedure with the exoskeleton before use.
 2. Evaluate the learning curve and user adaptation to the exoskeleton's functionalities (nice to have, it depends by the duration of the test, literature shows a week of test to be able to gather consistent data).
- 4. Data collection and analysis:
 - Implement data logging mechanisms to capture relevant performance and user data as trunk angle, , number of liftings, and the machine status.
 - Analyse the collected data to assess the exoskeleton's performance, identify areas for improvement, and validate its effectiveness in meeting predefined objectives.
- 5. Data streaming:
 - Test connection of exoskeleton gateway to Digital Twin database in terms of data streaming.

Detailed Testing Plan

Below can be found the proposed detailed testing plan for exoskeleton:

FUNCTIONAL, POWER AND SAFETY TESTING		
DESCRIPTION		
Objectives: Ensure the proper functioning, power management, and safety of the exoskeleton.		
Means of verification: Test and documentation on list of features mentioned above.	Methods: Functional tests and simulations on system characteristics and features.	Required resources: Exoskeleton prototype, testing equipment.
Execution: • Verify components and subsystems functionality. • Test range of motion, joint flexibility, and force generation. • Assess the responsiveness and accuracy of control mechanisms and user interface (if any). • Assess power source, battery life, and safety features. The results concerning the functionality of components and subsystems (e.g., user interface), the range of motion, force generation, the responsiveness and accuracy of control have already been presented in deliverable D5.1. Concerning Safety features, IIT has tested and verified the safety of main electronics components with the main goal of being compliant to EN ISO 13849-1 and 2. IIT has evaluated fault simulations carried out on the board components, achieving the related results. Test execution and methods: All the possible faults of the components of the main electronics board of the IIT exoskeleton are simulated one at a time. During the test, the detectability of the fault, card behavior, and status of the signal output/function of the card are recorded. Furthermore, where the fault may lead to an overload on a component, the protection temperature declared by the manufacturer on the datasheet is considered. Considering the calculations and considerations carried out, it is possible to declare that all the circuits the tested board achieves, based on what is reported in point 4.5 of the EN ISO 13849- 1, a PL equal to "C." Concerning Power consumption, the results will be included in Deliverable 5.2.		

LAB TESTING		
DESCRIPTION		
Objectives: Evaluate exoskeleton performance and mobility in controlled environments.		
Means of verification: testing and analysis of performance indexes of following activities: • Lifting data. • Load-carrying capacity data. • mobility performance evaluation.	Methods: testing of Lifting tasks, Load carrying, and mobility assessment on various surfaces. Tests were accomplished in laboratory with 5 to 10 subjects carrying out the target activities.	Required resources: Lab space, equipment for assessment (e.g., full body motion capture system, surface EMGs, metabolic consumption mask).
Execution: • Test exoskeleton performance in providing support when lifting heavy loads, reducing strain on the user. • Test load-carrying capabilities and effectiveness. • Assess user mobility on different surfaces, e.g., stairs, slopes, uneven terrains. Results have been published in the following papers: Evaluation of lifting task: V. Fanti¹, S. Leggieri, T. Poliero, M. Sposito, D. G. Caldwell and C. Di Natali, Multi-Exoskeletons Performance Evaluation: integrated muscle energy indices to determine the quality and quantity of assistance. Evaluation of carrying task: Vasco Fanti, Sergio Leggieri, Darwin G. Caldwell, Christian Di Natali. Proprioceptive-Based Control Strategy to Assist Walking and Carrying Tasks in Back-Support Exoskeletons. The results concerning the mobility performances of exoskeletons will be included in the following deliverable release.		

ON-SITE TESTING		
DESCRIPTION		
Objectives: Validate exoskeleton performance in real-world construction scenarios.		
Means of verification: User feedback (questionnaires), task completion time, ergonomic assessment results.	Methods: Field trials, ergonomic evaluations, user feedback collection, test of familiarization procedure (nice to have).	Required resources: Construction site access, construction materials, user training materials, questionnaires.
Execution: • Test exoskeleton support for manual material handling and upper limb activities. • Assess user interface, comfort, and adaptability. • Provide user training and evaluate user adaptation. Results will be shared in the following release of deliverables concerning the exoskeleton in the worksite. A tentative experimental protocol for the Hauling materials (lifting, carrying, lowering) consist in: 1. Test description: workers lift, carry, and lower objects/materials of 10-20kg from point A to B (20m). 2. Duration of the test: 1 week, 1h with EXO 1h w/o EXO. 3. Number of subjects: 10 workers. 4. Repetitions per subject: 2 (morning, afternoon). 5. Data gathering questionnaires based.		

DATA COLLECTION AND ANALYSIS		
DESCRIPTION		
Objectives: Capture and analyse performance and user data to assess exoskeleton effectiveness.		
Means of verification: Data analysis reports, performance metrics.	Methods: Data logging during trials, data analysis post-trials.	Required resources: Data logging devices, analysis software.
Execution: • Implement data logging mechanisms to capture relevant performance and user data such as trunk angle, number of lifts and the machine status. • Analyse collected data to identify areas for improvement and validate performance. • Equip the prototype of exoskeleton with sensors and data acquisition systems capable of capturing the relevant data to the exoskeleton performance, through sensors to gather real-time data during operations. • Ensure the sensors are accurate, reliable, and provide data in a format compatible with the data logging system. Quality of data analysis via data anomaly detection algorithms together with project global KPI “T12: Accuracy of collected data”. • Evaluate the robustness data logging system onboard the exoskeleton to capture sensor data continuously, ensuring the data logging system has sufficient storage capacity to handle large volumes of data generated during extended testing periods. • Establish a secure connection between the data logging system onboard the exoskeleton and the digital twin database. Implement protocols for data transmission, ensuring data integrity and security during transfer. Verify compatibility between the data logging system's output format and the database requirements for seamless data integration.		

DATA STREAMING		
DESCRIPTION		
Objectives: Integrate exoskeleton data with Digital Twin database for real-time monitoring and visualization.		
Means of verification: Successful data transmission, visualization of exoskeleton performance.	Methods: Integration testing, dashboard development.	Resources required: Digital Twin infrastructure, dashboard development tools.
Execution: - Test connection of exoskeleton gateway to Digital Twin database. - Provide dashboards for visualization of post-processed information. The dashboard to visualize the data, is under the responsibility of CSTB, who is the responsible partner for the dashboards provided for BEEYONDERS. This testing will be done through STAM IoT wearable device, as described under the ‘connectivity and compatibility’ section of the wearables test plan.		

Smart Wearables

Smart wearables development has been proposed in BEEYONDERS and will focus on the manufacturing and implementation of a suite of wearable sensors that will be applied on workers clothes or PPEs to monitor the well-being of the workers through the different tasks executed in the road construction and building construction use cases. To evaluate any possible risk for the workers, environmental parameters will be measured and combined with physical parameters such as the workers' activity recognition through accelerometers, location on the worksite through GPS or ergonomics evaluation. An IoT framework will provide a continuous stream of data from the wearable sensors to the Digital Twin that will allow collecting relevant information to improve users' safety. AI algorithms will then analyse the collected data to extract patterns that could potentially indicate a high risk for injuries and inform the workers in a preventive action.

Concerning the technical details of the smart wearables prototypes, it is important to remark that there will be three different prototypes (partners that will develop the prototypes are IIT, VTT and STAM), covering different aspects each and with the following details up to now:

- **STAM prototype details:**

It will be deployed in two different use-cases: building construction (indoor) and road construction (outdoor).

Sensors worn by the workers:

- Localization devices:
 - RTK systems, used to localize the workers within the environment. It will be captured the latitude and longitude (degrees) associated to a timestamp. For road construction use case.
 - Bluetooth tag for building construction use-case.
- Accelerometer, used to detect possible 'man down' situations. For road and building construction use case.
- Noise sensor, used to measure the sound pressure level (dB SPL). For road and building construction use case.
- IoT module for edge-to-cloud communication.

Sensors positioned around the environment:

- Localization stations for building construction use-case. The developed system is positioned at each entrance door of each apartment and is able to detect at which time instant a specific operator has entered or exited a specific apartment.
- Environmental monitoring stations positioned in known and fixed locations which measure the following parameters:
 - PM [$\mu\text{g}/\text{m}^3$]
 - VOC [adimensional index]
 - Temperature [$^{\circ}\text{C}$]

- Humidity [%RH]
- NOx [adimensional index]

- **VTT prototype details:**

It will be deployed in building construction (indoor) use case.

VTT is developing a software which analyses IMUs data to detect fatigue patterns during workers' activities. The software will be executed within IIT Smart Suit processing unit, utilizing data coming from Smart Suit IMUs.

Transversal testing plan (STAM, IIT, VTT)

The testing plan proposed for the different prototypes of smart wearables involves a systematic approach to evaluate its performance, functionality, usability, and reliability. The different aspects covered are the following:

1. Functional testing:
 - Verify that all features and functionalities of the smart wearable are working correctly, such as physiological parameters monitoring, activity tracking, notification alerts, communication capabilities, etc.
 - Test different modes or settings, such as sleep mode, active mode, and GPS navigation, to ensure accurate and reliable performance.
 - Evaluate RTK positioning and accuracy of the worker wearing the smart wearables (for road case only).
2. Sensor performance:
 - Evaluate the accuracy and reliability of sensors used in the smart wearable, such as accelerometers, gyroscopes, or GPS.
3. Connectivity and compatibility:
 - Evaluate the connectivity with the other devices.
4. Battery life and power management:
 - Evaluate the smart wearable's battery life under different usage scenarios, such as continuous tracking or intermittent usage.
5. Durability and resistance:
 - Evaluate the wearability and comfort of the device during use.
 - Define limits of resistance to physical stress (e.g., pressure, temperature, drop distance, etc.) of the smart wearables from the components' specifications.
6. Health and safety:

- Ensure that the smart wearable complies with relevant health and safety regulations.
 - Evaluate the effectiveness of safety features, such as fall detection or emergency alerts according to proximity detection to wheel loader (for road case only).
7. Data collection and analysis:
 - Implement data logging mechanisms to capture relevant user and device data.
 - Analyze the collected data to assess the performance of the smart wearable, identify areas for improvement, and validate the accuracy of physiologic parameters or activity measurements.
 - As for the Autonomous Wheel loader device, establish the data flow of current RTK location of the smart wearable to Digital Twin, and test in concordance with the location received from the wheel loader the distance calculation between both, the time lapse is safe enough (for road case only).
 8. User experience and feedback (T5.5 related):
 - Conduct user testing to gather feedback on the overall user experience, including comfort, usability, and satisfaction.
 - Gather user feedback on specific features, functionalities, or user interface elements to identify areas for improvement.
 - Consider user feedback for iterative design updates or software enhancements.
 - Assess the intuitiveness and ease of use of the smart wearables for the workers, and integration with other wearables or devices.
 9. Regulatory compliance:
 - Ensure that the smart wearable complies with relevant regulatory standards and certifications, such as safety, electromagnetic compatibility, or data privacy regulations (GDPR proposed is pseudo-anonym).

Detailed Testing

FUNCTIONAL TESTING		
DESCRIPTION		
Objectives: Verify the correct functioning of all features and capabilities of the smart wearable.		
Means of verification: Successful operation of all features, accurate positioning data, consistent performance across scenarios.	Methods: Use case scenarios, feature testing, real-world simulations.	Resources Required: Testing environment, access to real-world construction sites.
Execution: • Test health monitoring, activity tracking, environmental parameters. • Assess RTK positioning accuracy for workers (road construction case). • Verify that the indoor localization system is functioning as expected. • Assess fatigue of workers evaluation methods. VTT PROTOTYPE: The initial fatigue evaluation methods on workers were field tested in Helsinki on an actual construction work site. The initial prototype was worn by three workers for three whole working days. The sensors didn't hinder the movements of the workers since they were small and light weight. The batteries of the sensors and data collection mobile phone lasted for the full working day (8 hours). Bluetooth operation of three sensors per worker and the data collection mobile phone worked as intended with minimal packet loss. STAM PROTOTYPE: <u>RTK system for outdoor localization:</u> The system provides a very high accuracy (around 1cm). To test the correct functioning of the system, an operational test has been conducted where the sensor was placed on a machine while it was moving. Sensor's data have been recorded and plotted in google maps. The recorded positions perfectly matched with the followed path.		

Environmental stations for air quality assessment:

One sensor which provides different data.

- PM1.0, PM2.5, PM4, PM10 [$\mu\text{g}/\text{m}^3$]
- Temperature [$^{\circ}\text{C}$]
- Humidity [%RH]
- VOC [adimensional index 0-500]
- NOx [adimensional index 0-500]

The plot of the sensor can be visualized in figure below:

DATE	TIME	PM1	PM2.5	PM4	PM10	RH	TEMP	VOC	NOX
20/02/2024	08:14:47	83.90	90.50	92.60	93.70	40.86	20.38	97.00	1.00
20/02/2024	08:15:47	83.30	88.40	89.20	89.60	40.28	20.62	107.00	1.00
20/02/2024	08:16:47	86.00	92.30	94.20	95.00	39.84	20.74	97.00	1.00
20/02/2024	08:17:47	83.70	90.80	93.30	94.50	40.25	20.46	98.00	1.00
20/02/2024	08:18:47	84.00	90.60	92.70	93.80	41.05	20.19	105.00	1.00
20/02/2024	08:19:47	82.60	88.70	90.40	91.20	41.10	20.26	95.00	1.00
20/02/2024	08:20:47	83.40	89.90	92.00	93.00	41.09	20.33	99.00	1.00
20/02/2024	08:21:47	86.00	92.60	94.70	95.70	40.92	20.36	96.00	1.00
20/02/2024	08:22:47	85.70	92.30	94.40	95.40	41.03	20.43	101.00	1.00
20/02/2024	08:23:47	84.60	90.20	91.40	92.00	40.79	20.54	104.00	1.00

Indoor localization system:

In figure below we reported a screen of the interface that we have developed for testing purposes. In the first status (associated to time instant 0), both operators are out of the apartment. After an entrance/exit operation is detected by the system, the status of both workers is updated.

Operator 1: outside apartment A [0]
Operator 2: outside apartment A [0]

Operator 1: inside apartment A [2024-05-16 15:36:14]
Operator 2: outside apartment A [0]

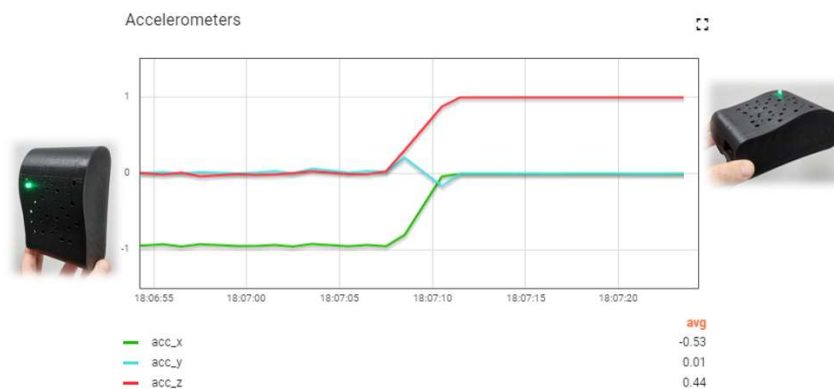
Operator 1: inside apartment A [2024-05-16 15:36:14]
Operator 2: inside apartment A [2024-05-16 15:36:33]

Operator 1: outside apartment A [2024-05-16 15:36:58]
Operator 2: inside apartment A [2024-05-16 15:36:33]

Operator 1: outside apartment A [2024-05-16 15:36:58]
Operator 2: outside apartment A [2024-05-16 15:37:19]

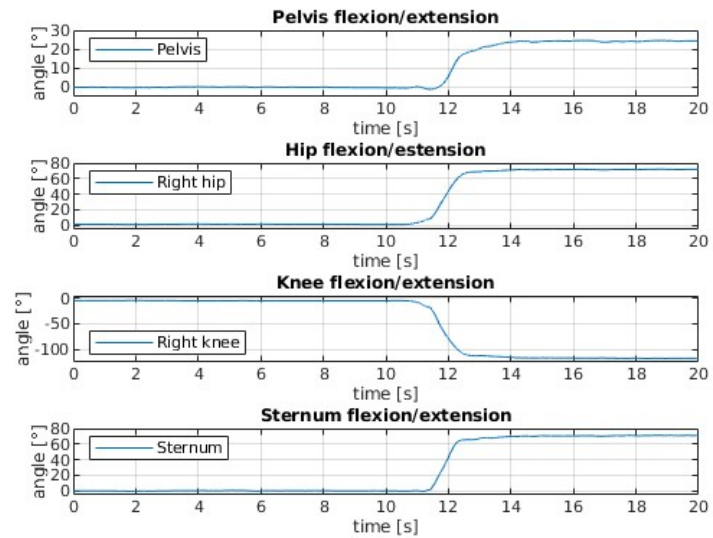
Accelerometer to detect man-down situations:

In figure below we can see that it is possible to determine when the case containing the wearable system is in horizontal position, which means that person wearing it is possibly laying down due to an accident.



IIT PROTOTYPE

The system is able to measure and save joint angles of the user and forearm muscle activity. The results show adherence with the movement carried out by the user.



Kinematic data collected by the Smart Suit during a squat. From the top to the bottom: the pelvis, the right hip, the right knee, and the sternum angles.



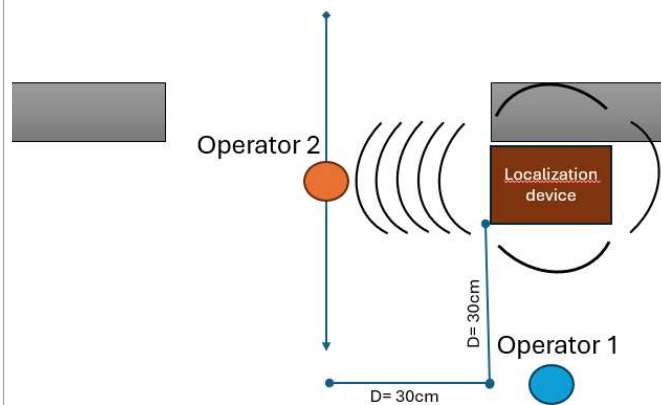
The kinematic data and the biomechanical model allow to reconstruct the subject motion.

SENSOR PERFORMANCE		
DESCRIPTION		
Objectives: Evaluate the accuracy and reliability of sensors used in the smart wearable.		
Means of verification: Consistent sensor readings across scenarios, alignment with reference data.	Methods: Sensor testing, data comparison with reference sensors.	Resources required: Sensor testing equipment, controlled environmental conditions.
Execution: Test sensor sensitivity and responsiveness in various scenarios (rest, exercise, environmental conditions). VTT PROTOTYPE: The initial sensors used in fatigue evaluation methods were pre-tested with laboratory tests to see that the readings are reliable. In artificial fatigue inducing laboratory tests: subjective Borg fatigue assessments, EMG, HR and HRV data values were cross-referenced with IMU sensors. Statistics, visualizations, and initial data analysis were implemented for the sensor values to do sanity checking, outlier values detection and overall assurance of correct sensor value operations. STAM PROTOTYPE: <u>RTK localization system:</u> To test the accuracy and reliability of the system, we have plotted the calculated GPS positions before and after the RTK corrections are activated. It can be noticed how the oscillating curve becomes a point after RTK corrections are used.		



Indoor localization system:

To test the system, we have positioned the localization device at the entrance of a room, and 2 Bluetooth tags were given to two persons, which are positioned as in figure below with respect to the localization device. In particular, operator 1 is standing still while operator 2 is moving along the line.



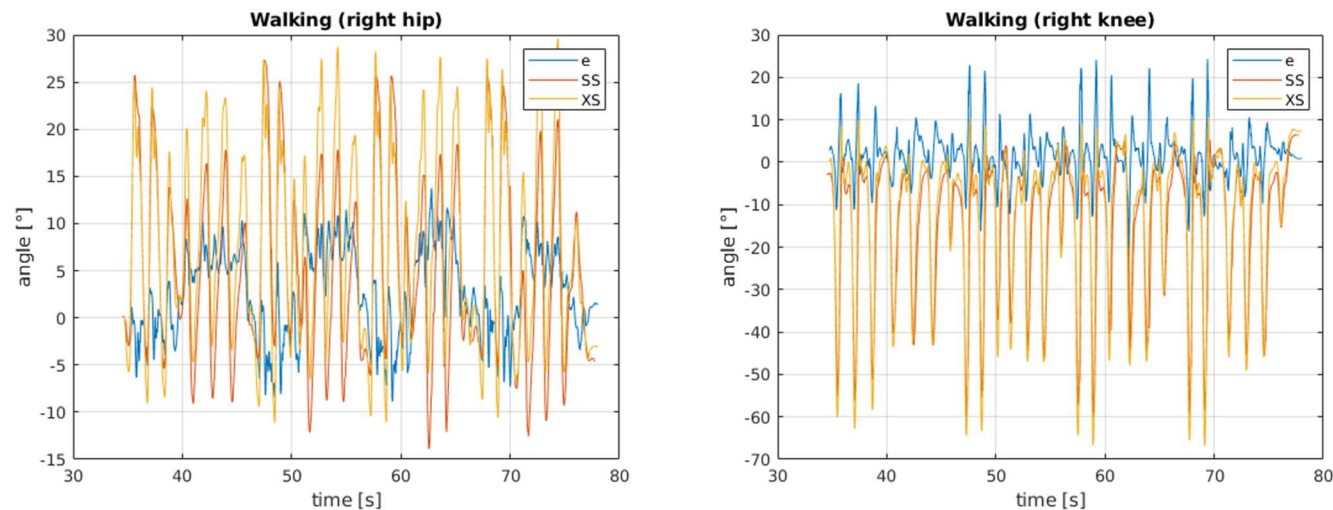
By plotting the RSSI of the two tags received by the localization device, we have verified that thanks to the shielding strategy implemented, the RSSI of the sensor which is not passing through the door is much lower compared to the other one. This allows us to discriminate between the two operators.

Environmental stations for air quality assessment:

We have performed some standard tests to check that each value registered by the environmental sensor is changing as expected:

- For temperature and humidity, the recorded value was compared with the output of other temperature and humidity sensors.
- To check if the VOC value was responding correctly, we have taken a big glass to create a close environment and a lighter. Then by pressing the button of the lighter without igniting the flame we can release gas and see how the VOC index increase while the NOx remains constant.
- By igniting the flame, we have observed changes also in NOx, PMs, temperature, and humidity.

IIT PROTOTYPE



The joint angles measured by the IIT prototype, in red, were compared with another standard motion capture system, in yellow, and the relative error, in blue, was computed. During walking, the root mean square error is 5.4 degrees at the hips and the root mean square error is 5.7 degrees at the knees.

CONNECTIVITY AND COMPATIBILITY		
DESCRIPTION		
Objectives: Ensure reliable wireless connectivity and compatibility with the Digital Twin architecture.		
Means of verification: Successful data synchronization, compatibility with Digital Twin architecture.	Methods: • Connectivity testing, compatibility checks. • Evaluate time lag for proximity alerts to wheel loader and accuracy.	Resources required: Digital Twin environment, connectivity testing tools.
Execution: • Test wireless connectivity (Bluetooth, Wi-Fi) for data synchronization. • Validate compatibility with proposed connection protocols (e.g., MQTT). VTT PROTOTYPE: The initial sensors used in fatigue evaluation methods were tested regarding data synchronization between different simultaneously used sensors, and connectivity tests were implemented to see that minimal packet loss would occur even with radio interference from other surrounding radio devices. STAM PROTOTYPE: The IoT device developed by STAM is equipped with a LTE CAT-1 module which is used to perform edge-to-cloud communication. The underlying data transfer protocol is MQTT. The IoT device is uniquely identified by a client_ID and publishes JSON encoded telemetry data to the MQTT broker hosted on the digital twin. The basic functionality of the IoT device is to transfer measurements from its onboard sensors, but it will also connect to the exoskeleton or smart suit to transfer their data stream, acting as a gateway. The IoT device firmware is implemented as an infinite loop, inside of which sensors are sampled and data is transmitted. Some preliminary tests were carried out, where STAM device was able to connect to the MQTT broker and publish some fictitious data. Some further tests will be conducted in the next months of the project to ensure that the IoT system is able to correctly interface with both IIT hardware and the DT platform (stability of the connection, data loss, update frequency, etc).		

BATTERY LIFE AND POWER MANAGEMENT		
DESCRIPTION		
Objectives: Evaluate the smart wearable's battery life and power management capabilities.		
Means of verification: Battery life meets expectations, effective power-saving features.	Methods: Battery life testing, power consumption monitoring.	Resources required: Battery testing equipment, power monitoring tools.
Execution: • Test battery life under different usage scenarios. • Assess power-saving modes and efficiency. VTT PROTOTYPE: The initial sensors used in fatigue evaluation methods and their battery life were tested to endure at least 8-9 hours of non-stop operation at 52Hz data frequency. This was also confirmed at actual field tests. STAM PROTOTYPE: The battery of STAM devices is designed to have a minimum autonomy of 8 hours/day. For some of them, such as the tag worn by the workers for the indoor localization, a power saving mode has been adopted, which only activates the tag when is moving. The most power consuming activity is related to the communication interface and exchange of data with IIT prototypes and the digital twin platform. Final tests will be carried out to test the battery life during collection and exporting of the data. IIT PROTOTYPE: The sensors' battery life is estimated to last for 350 minutes during nominal continuous usage. The system is powered by a 8750mAh of total capacity.		

DURABILITY AND RESISTANCE		
DESCRIPTION		
Objectives: Assess the smart wearable's durability and resistance to environmental factors.		
Means of verification: Device remains functional after exposure to environmental factors, positive user feedback on comfort.	Methods: Environmental testing, wearability assessments.	Resources required: Environmental testing facilities, user feedback mechanisms.
Execution: • Test resistance to water, dust, shock, and physical stress. • Evaluate wearability and comfort over extended use. For the three prototypes, a generic testing methodology is proposed: • Identify stress factors: Construction sites are subject to extreme conditions such as dust, moisture, temperature variations, mechanical shocks, and exposure to chemicals. • Usage scenarios: Consider the activities will be carried on by the prototypes, such as heavy lifting and exposure to environmental construction conditions and machinery. • Environmental usage: Ensure that the materials used in the construction of prototypes and on-board equipment are designed to be resistant to water and dust (e.g., IP67 or IP68), and can operate in high or low temperatures and high humidity. Evaluate exposure to oils, fuels, and other chemicals common in construction environments. • Field testing: Deploy prototypes in real construction environments for practical feedback. IIT PROTOTYPE: The system has been integrated into a working suit with windproof, waterproof, and tearproof clothing (EN 343:2003+A1:2007, class 2/2). The integration of IIT sensors does not affect these features.		

HEALTH AND SAFETY		
DESCRIPTION		
Objectives: Ensure compliance with health and safety regulations and evaluate the effectiveness of safety features.		
Means of verification: Compliance with regulations, effective operation of safety features.	Methods: Regulatory compliance checks, safety feature testing.	Resources required: Regulatory guidelines, safety testing equipment.
Execution: - Assess compliance with relevant regulations. - Test safety features like fall detection and emergency alerts. STAM PROTOTYPE: Data acquired by the sensors will be compared with some preselected threshold. If the value of the data exceeds the threshold, an alert will be generated in the user-interface so that the site manager can visualize it. Furthermore, an alert will be sent directly to the workers in case the distance between the run-time worker position and the run-time wheel loader position is below some threshold. In this case, an acoustic signal will be emitted by the buzzer worn by the worker. Since the distance is computed directly on the DT, the delay in sending the alert to the worker will be measured. VTT PROTOTYPE: In regards with compliance and certification: - Regulatory Standards: Ensure the components of the prototype complies with industry standards such as ANSI, ASTM, or relevant ISO standards for protective equipment. - Safety certifications: Obtain necessary certifications for electronic safety, such as CE or UL markings. IIT PROTOTYPE:		

The system integrates CE mark sensors in a commercial working suit (EN 343:2003+A1:2007, class 2/2). Thus, the IIT prototype and components comply with industry standards.

DATA COLLECTION AND ANALYSIS		
DESCRIPTION		
Objectives: Capture and analyse relevant user and device data for performance assessment and improvement.		
Means of verification: Actionable insights derived from data analysis, improved device performance.	Methods: Data logging, data analysis tools.	Resources required: Data logging tools, data analysis expertise.
Execution: • Implement data logging mechanisms. • Analyse collected data to assess performance and identify areas for improvement. • Equip the prototype with sensors and data acquisition systems capable of capturing the related KPIs to the Smart Wearables, and assessment of the integration of the sensors with the Digital Twin database to gather real-time data during operations. Verify compatibility between the data logging system's output format and the database requirements for seamless data integration. • Ensure the sensors are accurate, reliable, and provide data in a format compatible with the data logging system. Quality of data analysis via data anomaly detection algorithms together with project global KPI "T12: Accuracy of collected data". • Evaluate the robustness data logging system onboard the Smart Wearables to capture sensor data continuously, ensuring the data logging system has sufficient storage capacity to handle large volumes of data generated during extended testing periods.		

USER EXPERIENCE AND FEEDBACK		
DESCRIPTION		
Objectives: Gather user feedback to enhance the user experience and usability of the smart wearable.		
Means of verification: Positive user feedback, iterative design updates based on user input.	Methods: User testing, surveys, feedback collection.	Resources required: User testing facilities, feedback collection mechanisms.
Execution: • Conduct user testing sessions to gather feedback. • Gather feedback on comfort, usability, and satisfaction. Within WP5, Task 5.5: “Monitoring and assessing the impacts of breakthrough technologies on workers” will take care of gathering user feedback on the developed technologies. Indeed, a set of worker KPIs has been defined, and will be measured with interviews, questionnaires etc.		

REGULATORY COMPLIANCE		
DESCRIPTION		
Objectives: Ensure compliance with relevant regulatory standards and certifications.		
Means of verification: Compliance certifications, adherence to regulatory guidelines.	Methods: Compliance checks, regulatory review.	Resources required: Regulatory guidelines, compliance expertise.
Execution: • Review regulatory standards related to safety, electromagnetic compatibility, and data privacy. • Ensure adherence to GDPR regulations for data privacy. The prototype developed by STAM will not measure sensitive data apart from the position of workers. All recorded data will be fully anonymized and exclusively used for the project's purposes. Upon the project's completion, all data will be deleted. Additionally, all details about the technologies and their functionalities will be transparently shared with the workers, and the technologies will only be used with their explicit consent.		

Digital Twin

A Digital Twin is proposed to be used for monitoring, diagnostics, and prognostics to optimize asset performance and utilization of the technologies proposed in BEEYONDERS that will be deployed in the case studies. Complex prognostics modules based in Digital Twin will find the root cause of issues and improve productivity, planning performance, health & safety of workers and environmental impact.

At this stage of the project, Digital Twin is currently being developed. For technical details about the Digital Twin proposed, the deliverable D3.1³⁸ collects the fundamental aspects of BEEYONDERS digital twin architecture design. It examines the different components of the architecture, the challenges involved, and the best practices to ensure successful implementation.

At the end of this task, T2.5 (M24), the main goal for Digital Twin is to demonstrate that the different technologies of BEEYONDERS are connected to the database, streaming or off-line data can be visualized (so an interface is required) and BIM capabilities are available. Also, the architecture deployed for the Digital Twin will provide through the tasks T3.5 – Input services and T3.6 – Output services the intelligence of the system from the data collected, overlapping these two tasks the use cases demonstrations to allow an iteration and refine of the system.

To achieve the goals proposed, the following testing procedure is proposed:

1. Data integration and synchronization:
 - Verify that the digital twin accurately represents the physical system it is modelling in terms of point clouds obtained and BIM visualization.
 - Test the integration and synchronization of data from various sources, such as sensors, IoT devices, or historical data. Validate MQTT broker or any other proposed to be deployed.
2. Functional testing:
 - Test the functionality of specific features and interactions within the digital twin, such as data visualization both on live time (IoT) and static data (off-line data as captured cloud points, BIM models, KPIs results... data that refresh in time as to be considered static).
 - Test the functionality of any developed API to connect data to external services or servers, as Input/Output microservices or any other server required.
3. Performance and scalability:
 - Evaluate the performance and scalability of the digital twin when handling large datasets.

³⁸ A. Armijo., D. Zamora., J. Jimenez et al. D3.1: Digital Twin design and architecture (2022). BEEYONDERS project, <https://beeyonders.eu/>

- Test the response time and computational efficiency of the digital twin in real-time or near-real-time scenarios.
4. Validation and verification:
 - Compare the outputs of the digital twin against real-world observations and measurements.
 - Evaluate the effectiveness of the digital twin in providing real-time insights, recommendations, or optimization strategies.
 5. Security and data privacy.
 - Ensure that the digital twin implementation follows best practices for security and data privacy.
 6. Lifecycle management:
 - Develop testing procedures for updates, upgrades, or modifications to the digital twin throughout its lifecycle, as testing the compatibility and consistency of the digital twin during version changes or system upgrades.
 7. User experience and feedback (mainly interface where users will interact):
 - Gather user feedback on the usability, intuitiveness, and effectiveness of the digital twin interface.
 - Conduct user testing to validate the user experience and identify areas for improvement or additional functionalities.
 8. Documentation and training (same than point 7, focused on interface):
 - Develop comprehensive documentation for the digital twin, including user manuals, technical guides, and troubleshooting procedures.
 - Conduct training sessions or workshops to ensure users understand how to effectively use and interact with the digital twin.

Detailed Testing Plan

DATA INTEGRATION AND SYNCHRONIZATION		
DESCRIPTION		
Objectives: Ensure the accuracy and reliability of data representation and synchronization within the Digital Twin.		
Means of verification: Alignment between physical and digital representations, successful data integration tests.	Methods: Comparison of Digital Twin representation with physical system, data synchronization tests.	Resources required: Access to physical system data, testing environment for data synchronization.
Execution: • Test integration and synchronization of data from diverse sources (sensors, IoT devices, historical data). • Validate functionality of MQTT broker or other proposed data integration solutions. • MQTT connectivity test: Verify if the MQTT client can establish a successful connection with the MQTT broker. This includes testing different scenarios such as valid/invalid credentials, and handling of multiple simultaneous connections. • MQTT publish test: Test if the MQTT client can successfully publish a message to a topic on the broker. • MQTT subscribe test: Test if the MQTT client can subscribe to a topic and receive messages published to that topic. • MQTT broker stress test: Assess the broker's performance and stability under high load conditions. This involves simulating numerous clients, high message rates, and large message sizes. The broker's success is determined by its ability to handle the load without significant performance degradation or errors. • MQTT to InfluxDB data integrity test: Ensures that data received via MQTT is accurately stored in InfluxDB. This involves verifying the consistency and accuracy of the data stored in the database. The success of the test is determined by the match between the data retrieved from InfluxDB and the original data published by the MQTT client, with no missing, duplicated, or altered data. • Knowledge graph to InfluxDB integration test: Validates the consistency and accuracy of temporal data stored in InfluxDB against the contextual information in the Knowledge Graph, using the unique ID as the common link. The success of the test is determined by the match between the data associated with each unique ID in both systems, ensuring no discrepancies exist. • Knowledge graph to file storage integration test: Validates the consistency and accuracy of files stored in the repository against the contextual information in the Knowledge Graph, using the unique ID and Project paths as the common links. The success of the test is determined by the match between the data associated with each unique ID/path in both systems, ensuring no discrepancies exist.		

- Real-time synchronization test: Ensures that the system can accurately and promptly reflect changes in the physical system within the Digital Twin. This involves simulating real-time data generation, monitoring the synchronization process, and measuring the system's latency. The success of the test is determined by the system's ability to meet the defined real-time requirements, even under high load conditions. Additionally, the integrity of the data in the Digital Twin is verified after synchronization, ensuring no data is lost, duplicated, or altered during the real-time synchronization process.

FUNCTIONAL TESTING		
DESCRIPTION		
Objectives: Verify the functionality and performance of specific features and interactions within the Digital Twin.		
Means of verification: Successful execution of visualization features, API responses within expected parameters.	Methods: Use case scenarios, API testing, functional testing procedures.	Resources required: Access to visualization tools, testing environment for APIs.
Execution: Test visualization features for live and static data. Validate functionality of APIs connecting data to external services or servers. To evaluate the functional capabilities of the digital twin involves validating its accuracy, performance, and ability to replicate the physical counterpart, by comparing the digital twin's data and outputs with real-time data from the construction use case scenario. Check if it can accurately identify potential failures and assess the performance under changing conditions. At the end of the project, it will be engaged the use case leaders (acting as stakeholders) to provide feedback and insights, ensuring the digital twin meets operational requirements and provides actionable insights.		

PERFORMANCE AND SCALABILITY		
DESCRIPTION		
Objectives: Evaluate the performance and scalability of the Digital Twin under varying conditions and dataset sizes.		
Means of verification: Response time metrics, scalability testing results.	Methods: Load testing, stress testing, performance monitoring.	Resources required: Scalability testing infrastructure, performance monitoring tools (JMeter).
Execution: Test response time and computational efficiency with large datasets. Assess scalability under increasing data loads. Preliminary performance and scalability testing of the services show a good response from the Digital Twin architecture under different loads. However, full-scale testing must be carried out in a more real scenario, after all the use cases are deployed, using the JMeter tool for testing.		

VALIDATION AND VERIFICATION		
DESCRIPTION		
Objectives: Ensure that the Digital Twin outputs align with real-world observations and provide meaningful insights.		
Means of verification: Consistency between Digital Twin outputs and real-world observations.	Methods: Data comparison, validation against real-world scenarios.	Resources required: Access to real-world data, validation scenarios.
Execution: Compare Digital Twin outputs with real-world data. Evaluate the effectiveness of the Digital Twin in providing real-time insights and recommendations. The validation will be executed once the use cases are demonstrated, and the results will be analysed through the task 7.8 – Impact assessment of the technologies deployed.		

SECURITY AND DATA PRIVACY		
DESCRIPTION		
Objectives: Ensure the security and privacy of data within the Digital Twin ecosystem.		
Means of verification: Security audit reports, compliance certifications.	Methods: Security audits, compliance checks.	Resources required: Security an Analyzer tools (MQTTX, Wireshark, etc.), expertise in cybersecurity.
Execution: Implement best practices for data encryption, access control, and authentication. Conduct security audits and vulnerability assessments. The deployed Digital Twin architecture complies with legal requirements, industry best practices, and internal guidelines related to security and data privacy. The system implements different security measures: - Authentication mechanisms have also been implemented for accessing the system and internal services. - Server identity verified to clients through Certificates. - Encrypted communications (use of HTTPS, SSH, TLS), to ensure they cannot be intercepted, and sensitive information is not observed by third parties. Prowler security audit tool will be used periodically during the deployment identifying further vulnerabilities and reducing risks. 1. MQTT Security Test: Focuses on evaluating the broker's TLS/SSL encryption capabilities. It involves verifying the successful establishment of secure connections and the proper encryption and decryption of messages. The test's success is determined by the broker's ability to securely handle messages without any security breaches or errors. 1.1. Invalid Credential Test: Attempt to connect to the broker using invalid credentials. The broker should reject this connection. 1.2. Invalid Certificate Test: Attempt to connect to the broker using an invalid SSL certificate. The broker should reject this connection. 1.3. Data Breach Test: Attempt to intercept the data being sent from the client to the broker (a "man-in-the-middle" attack). With proper TLS/SSL encryption, this intercepted data should be unreadable.		

2. Web Service Security Test: Focuses on evaluating the security capabilities of the services. It involves verifying the successful establishment of secure connections and the proper encryption and decryption of messages. The test's success is determined by the service's ability to securely handle messages without any security breaches or errors.
 - 2.1. Invalid Credential Test: Attempt to connect to the services using invalid credentials. The server should reject this connection.
 - 2.2. Data Breach Test: Attempt to intercept the data being sent from the server to the client (a "man-in-the-middle" attack). With proper HTTPS encryption, this intercepted data should be unreadable.

LIFECYCLE MANAGEMENT		
DESCRIPTION		
Objectives: Develop procedures for managing updates, upgrades, and modifications to the Digital Twin over its lifecycle.		
Means of verification: Successful version updates, stable system performance post-upgrade.	Methods: Version control, regression testing, rollback testing.	Resources required: Version control tools, rollback procedures.
Execution: Test compatibility and consistency during version changes or system upgrades. Develop rollback procedures in case of issues with updates. This test plan details the procedures to be implemented to successfully pass the lifecycle management test, focusing on updates, upgrades, and modifications. 1. Version Control: Tracking Evolution Implementation: A robust version control system (e.g., Git) will be employed to meticulously track all changes to the DT's codebase, configuration files, and associated documentation. Procedure: • Every change, no matter how small, will be committed to the version control system with a clear description of the modification. • Branches will be used for developing new features or performing major upgrades, allowing for isolated development, and testing before merging changes into the main codebase. • Tags will be used to mark specific releases of the DT, providing clear reference points for rollback or analysis. Benefits: • Provides a complete history of changes, enabling easy identification of modifications and their impact. • Facilitates collaboration among developers, allowing them to work on different aspects of the DT concurrently. • Enables easy rollback to previous versions in case of issues with updates. 2. Regression Testing: Ensuring Stability		

LIFECYCLE MANAGEMENT
DESCRIPTION
Implementation: A comprehensive suite of regression tests will be developed to cover all critical functionalities of the DT. Procedure: • Before deploying any update or upgrade, the entire regression test suite will be executed to ensure that existing functionalities are not adversely affected by the changes. • Any failing tests will be investigated and addressed before the update is deployed. • New tests will be added to the suite as new functionalities are developed, ensuring continuous coverage. Benefits: • Provides confidence that updates and upgrades do not introduce new bugs or regressions. • Helps to maintain the stability and reliability of the DT over time. • Enables early detection of issues, reducing the risk of costly downtime or disruptions. 3. Rollback Procedures: Reversing Changes Implementation: Well-defined rollback procedures will be established for every update and upgrade. Procedure: • A detailed plan outlining the steps to revert the system to the previous stable version will be documented. • This plan will include procedures for restoring data, configuration files, and code to their previous states. • Regular rollback testing will be conducted to ensure that the procedures are effective and can be executed quickly in case of emergency. Benefits: • Enables rapid recovery from issues caused by updates or upgrades. • Minimizes downtime and disruptions to construction operations. • Provides a safety net for confidently deploying changes, knowing that there is a reliable mechanism for reversing them if needed. 4. Compatibility and Consistency Testing Implementation: Specific tests will be designed to verify compatibility and data consistency during version changes or system upgrades. Procedure: • Data migration procedures will be developed and tested to ensure that data from previous versions can be seamlessly transferred to the updated system.

LIFECYCLE MANAGEMENT
DESCRIPTION
• Compatibility tests will be conducted to ensure that the updated DT integrates smoothly with existing systems and data sources. • Performance testing will be performed after upgrades to ensure that the system continues to meet performance requirements. Benefits: • Ensures smooth transitions between versions, minimizing disruptions and data loss. • Maintains the integrity and consistency of data throughout the DT's lifecycle. • Prevents incompatibility issues that could hinder the system's functionality. 5. Documentation and Communication Implementation: All lifecycle management procedures, including version control, regression testing, and rollback plans, will be documented. Procedure: • Documentation will be kept up to date with every change and made readily accessible to all relevant stakeholders. • Communication channels will be established to inform stakeholders about upcoming updates, upgrades, and potential downtime. Benefits: • Ensures clarity and transparency in the management of the DT. • Facilitates effective communication and collaboration among stakeholders. • Promotes a proactive approach to managing changes, reducing the risk of surprises and disruptions.

USER EXPERIENCE AND FEEDBACK		
DESCRIPTION		
Objectives: Gather user feedback to improve the usability and effectiveness of the Digital Twin interface.		
Means of verification: Positive user feedback, improved interface usability.	Methods: User testing, surveys, feedback collection mechanisms.	Resources required: User testing environment, feedback collection tools.
Execution: Conduct user testing sessions to evaluate interface usability and intuitiveness. Gather feedback on user experience and identify areas for improvement.		

DOCUMENTATION AND TRAINING		
DESCRIPTION		
Objectives: Provide comprehensive documentation and training materials to support users in effectively utilizing the Digital Twin.		
Means of verification: User proficiency, reduced support requests.	Methods: Documentation creation, training sessions, knowledge transfer.	Resources required: Documentation tools, training materials, training environment.
Develop user manuals, technical guides, and troubleshooting procedures. Conduct training sessions or workshops to educate users on Digital Twin usage.		

Human Robot Collaboration

In BEEYONDERS, Human Robot Collaboration software techniques (HRC) will deal with the realization of an agent-based model environment to replicate the use cases scenarios and activities, to evaluate both the impact of the autonomous robot tasks on the human operators and from a safety and robot operation efficiency. Information will be collected from the end-users and from other technical activities (e.g.: robot workspace, autonomous path, tasks plan and cycle times).

From this information, the created model will be able to reproduce and simulate the different parallel activities occurring in the working area, providing each modelled agent with a precise behaviour pattern and schedule. An extensive simulation campaign will be performed, to study the interactions between the robot and the other agents involved in the working environment.

Testing plan proposed will evaluate the effectiveness and accuracy of the evaluations, optimisations, and simulations to be carried out by the dedicated software for HRC, and comprises the following aspects:

1. Model validation:
 - Verify that the model accurately represents the real-world system being simulated, specially focused on the use cases environments.
 - Validate the model's assumptions, inputs, and relationships against empirical data or expert knowledge.
 - Conduct sensitivity analysis to assess the model's response to changes in input parameters.
2. Functional testing:
 - Test the functionality of the model by running various scenarios and comparing the results with expected outcomes.
 - Verify that the model correctly implements the rules and behaviours of individual agents and their interactions.
3. Performance testing:
 - Evaluate the model's performance in terms of computational efficiency and resource utilization.
 - Test the model's scalability by increasing the number of agents or the complexity of the simulated system.
 - Assess the model's runtime, memory usage, and responsiveness to ensure it can handle large-scale simulations.
4. Sensitivity analysis:
 - Conduct sensitivity analysis to determine how changes in model parameters or initial conditions affect the model's behavior and outcomes.
 - Identify critical parameters or variables that significantly influence the model's results.
 - Test the robustness of the model by varying these critical parameters within plausible ranges.

5. Validation of the optimisation:
 - Compare the model's outputs with real-world data, validating the optimized coordinates obtained to assess its accuracy and validity with the coordinates system of the road construction cloud points from drone.
 - Compare the implementation of different optimization techniques providing all the pros and cons of each one.
6. Verification of Agent Behaviour:
 - Test the behavior and decision-making of individual agents within the model.
 - Validate that the agents correctly follow their assigned rules and interact with other agents according to predefined protocols.
 - Validate optimization routes (for road use case example) and the calculation process, as well test the inputs required for the simulation.
7. Documentation and reproducibility:
 - Document the model specifications, assumptions, and parameter values used in the simulations.
 - Provide clear instructions on how to reproduce the experiments and results.
 - Connection to Digital Twin data base test, in terms of receiving inputs required for the modelling as well data flow exporting results to Digital Twin.

Detailed Testing Plan

Testing plan proposed will evaluate the effectiveness and accuracy of the evaluations, optimisations, and simulations to be carried out by the dedicated software for HRC, and comprises the following aspects:

MODEL VALIDATION		
DESCRIPTION		
Objectives: Ensure that the model accurately represents the real-world system and validate its assumptions and inputs.		
Means of verification: Consistency between model inputs and real-world observations.	Methods: Empirical data comparison, expert consultation.	Resources required: Access to real-world data, expertise in the domain.
Execution: Compare the model's inputs with empirical data. The agent-based simulation inputs are retrieved from the DT platform. They consist in the geolocalized map of the environment generated by the autonomous drone and the GPS coordinates of workers and machines registered through a very precise localization system (RTK). Once the input data are downloaded from the DT and imported into Unity, they must be converted into the Unity reference system. This conversion has to be as precise as possible, to ensure that the simulated environment reliably represent the real one. To verify these aspects, we will select a portion of map from google maps and we will import it into Unity with a set of geolocated points. Then, we will define a trajectory in google maps within the selected map portion and save it in the form of GPS coordinates + timestamp. Such GPS coordinates will be imported into Unity as well. After the needed conversion between the global reference system and the Unity reference system, we will check that the trajectory looks the same in both reference systems. More information can be found in D4.7.		

FUNCTIONAL TESTING		
DESCRIPTION		
Objectives: Verify that the model functions correctly under various scenarios.		
Means of verification: Consistency between simulated and expected results, validation of agent behaviours.	Methods: Scenario-based testing, comparison with expected outcomes.	Resources required: Simulation environment, scenario design expertise.
Execution: • Run simulations for different scenarios and compare results with expected outcomes. • Validate correct implementation of agent rules and interactions. The purpose of the simulation is to reproduce the environment and the various agents tasks occurring in the road-construction use-case and calculate the optimal trajectory for the wheel loader. This trajectory is the result of a trade-off between finding the shortest path and avoid possible collision with workers. The trajectory was calculated implementing the A* algorithm and defining a custom cost-function, that includes a parameter called BALANCE. By modifying this parameter, we are able to influence the calculated trajectory depending on the needs. In particular, positive values of BALANCE prioritize finding the shortest path feasible, while negative values generate routes that traverse less congested areas. This behavior will be verified in the simulation. More information can be found in D4.7.		

SENSITIVITY ANALYSIS		
DESCRIPTION		
Objectives: Determine how changes in parameters affect model behaviour and outcomes.		
Means of verification: Understanding of parameter influence on model behavior, identification of critical parameters.	Methods: Parameter variation.	Resources required: Parameter range determination.
Execution: • Vary relevant parameters and observe changes in model outputs. • Identify critical parameters influencing model results. Apart from the BALANCE parameter described in the functional testing, there are other parameters that are internal to the simulation that influence the computation of the optimal trajectory. Such parameters define how the heatmap is generated, a grid-like structure that allows to record and visualize crowd flows in a specific environment. Once the heatmap is generated, it is integrated in the A* cost function and used to compute the optimal route of the wheel loader. The parameters which define how the heatmap is generated will be varied observing how the calculated trajectory is affected. More information can be found in D4.7.		

PERFORMANCE TESTING		
DESCRIPTION		
Objectives: Evaluate the model's computational efficiency and scalability.		
Means of verification: Meeting performance benchmarks, ability to handle larger simulations.	Methods: Performance monitoring, scalability testing.	Resources required: Computational resources, performance monitoring tools.
Execution: • Measure runtime, memory usage, and responsiveness of the model. • Test scalability by increasing the complexity of the simulated system. The most computational expensive process of the simulation is the computation of the heatmap, whose complexity depends on the values set for the customizable parameters. Three values have been taken into account for the tests: • Memory usage • CPU usage • Frames per second Different scenarios have been tested considering different number of workers and different values for the DISTANCE parameter of the heatmap: • 10 persons, heatmap distance =1 • 10 persons, heatmap distance =4 • 10 persons, heatmap distance =8 • 25 persons, heatmap distance =1 • 25 persons, heatmap distance =4 • 25 persons, heatmap distance =8 While factor value has been set to 5.		

The tests have been run on a HP ZBook Firefly 14 inch G9 Mobile Workstation PC, with a 12th Gen Intel® Core™ i71255U, 1700 MHZ, 10 core with 32 GB RAM, with no other relevant applications running at the same time.

Persons	heatmap dis	memory	cpu	fps
10	1	2370 MB	50%	2
10	4	2313 MB	45%	1.2
10	8	2282 MB	30%	0.9
25	1	2418 MB	55%	1.6
25	4	2337 MB	37%	1
25	8	22920 MB	30%	0.8

What we can derive from these results is that increasing the complexity of the calculation leads to an expected decrease in FPS, while OS-related values experience a parallel, unexpected, decrease. These reductions in the OS values may be related to internal logic of the OS process scheduling and memory allocation.

The most meaningful data is the FPS value, being it related to the amount of time required to calculate all the data that are required for each frame.

VALIDATION OF OPTIMIZATION		
DESCRIPTION		
Objectives: Validate the optimized route calculations against real-world data.		
Means of verification: Consistency between model outputs and real-world data, assessment of optimization techniques.	Methods: Data comparison, optimization technique assessment.	Resources required: Access to construction site data, optimization analysis tools.
Execution: • Compare model outputs (optimized coordinates) with actual construction site data. • Assess accuracy and validity of optimized routes. To calculate the optimized trajectory for the wheel loader, we mainly investigated 2 path planning algorithms: NavMesh and A*. We opted for the A* algorithm for its extensive customization capabilities. Indeed, in our scenario it's crucial to integrate an additional parameter into the optimization process, one that specifically take into account the paths taken by workers. Customizing the algorithm to accommodate this requirement would prove challenging with the NavMesh algorithm, given its near black-box nature, which makes customization difficult. Once the simulation will be integrated with the other beeyonders technologies and tested in the road-construction use-case, we will assess the accuracy of the generated routes comparing the routes covered by the wheel loader within unity with the ones covered by the wheel loader in the real scenario. If the conversion between the Unity reference system and the real one is accurate, then the two routes will coincide.		

DOCUMENTATION AND REPRODUCIBILITY		
DESCRIPTION		
Objectives: Document the model specifications and provide clear instructions for reproducing experiments.		
Means of verification: Clear and comprehensive documentation, successful reproduction of experiments.	Methods: Documentation creation, experiment replication.	Resources required: Documentation tools, experiment replication expertise.
Execution: - Document model assumptions, parameter values, and experimental setups. - Provide instructions for reproducing simulations and interpreting results. The simulation exchange information with the digital twin platform. Input data: - Map of the working site registered from the autonomous drone. It is provided in OBJ format and imported in unity through API request. - GPS coordinates of workers and wheel loader recorded trajectories. They are recorded using a RTK localization system and imported in unity through API request. Output data: - GPS coordinates of the optimized trajectory calculated for the wheel loader. They are uploaded to the DT through API request. To enhance the usability of the simulation, we've developed a Windows executable that streamlines several key steps of the process. This executable comes pre-loaded with the map of the environment and the recorded trajectories of workers. Users can conveniently set the main parameters for heatmap creation, as well as the balance parameter. Once the desired parameter values are inputted, users can initiate the simulation by clicking the "simulate" button. In real-time, the scenario unfolds, allowing to see workers moving along their defined trajectories, while the heatmap is dynamically generated. Upon completion of this process, the application transitions to another screen, where users can observe the wheel loader navigating along the optimized trajectory calculated based on the specified parameters.		

6. Data Analysis

Data analysis involves examining the collected data to achieve meaningful conclusions and answer research questions about each BEEYONDERS technology presented in section 5 of this document. The specific data analysis techniques will vary depending on the nature of the developed technology, the type of data collected, and the research objectives. The steps for data analysis comprises the following aspects:

- **Data cleaning and preparation:** cleaning and organizing the collected data. Removing any outliers, missing values, or errors. Checking for inconsistencies and ensure data uniformity and preparing the data in a format suitable for analysis.
- **Visualizations:** Providing visual representations of the data using graphs, charts, or plots. Visualizations help in understanding patterns, trends, and relationships within the data. Common types of visualizations include histograms, scatter plots, bar charts, and box plots.
- **Statistical analysis:** computing descriptive statistics to summarize and describe the data. Calculate measures such as mean, median, mode, standard deviation, and range. These statistics provide a basic understanding of the data distribution and central tendencies.
- **Qualitative analysis:** If the testing results includes qualitative data, such as interview transcripts or open-ended survey responses, conduct qualitative analysis. This can involve thematic analysis, content analysis, or other qualitative techniques to identify patterns, themes, or insights from the qualitative data. This aspect is more related to the human centred devices (exoskeleton or smart wearables).
- **Interpretation of findings:** Analysing the results of the data analysis and interpreting the findings in the context of each technology research objectives.
- **Validation and reliability:** Assessing the validity and reliability of the data analysis process, through considering factors such as the representativeness of the sample, the accuracy of the measurements, and the rigor of the analysis techniques used.
- **Reporting:** Presenting the data analysis results in a clear and concise manner including tables, figures, or visualizations to support the interpretation of findings with explanations and interpretations of the results.
- **Peer Review and discussion:** At the end, feedback from partners, advisors, or colleagues on your data analysis and findings is a good practice, participating in discussions to validating the interpretations and conclusions.

The selection of the appropriate data analysis techniques is based on the research objectives, type of data, and the specific hypotheses or technical developments being investigated. Reporting of the data analysis will be presented in each technology related deliverable.



7. Conclusions

In this final version of the deliverable D2.5, it has been successfully developed and documented an “ad hoc” testing methodology tailored to the unique requirements and future objectives of the diverse typologies of technologies developed in BEEYONDERS, including the autonomous wheel loader, drones, autonomous maintenance vehicle, digital twin, smart wearables, exoskeleton, additive manufacturing, and human-robot collaboration systems. This methodology has been designed to assess and validate these technologies, ensuring they meet the criteria necessary to advance from Technology Readiness Level (TRL) 4 to TRL 5, setting the stage for successful demonstrations and further development in the upcoming phases of the project.

This approach compiles detailed protocols for evaluating performance, durability, reliability, and integration capabilities, among others, within the context of their respective use cases. The results and conclusions obtained from this testing methodology will be critical in demonstrating the readiness and maturity of the technologies, paving the way for their future deployment and integration into real-world applications. While this document outlines the testing methodology, the subsequent results and findings will be comprehensively detailed in future deliverables, which will provide in-depth analysis and validation for each technology under the BEEYONDERS project.



8. References

- ¹ <https://horizoneuropencpportal.eu/sites/default/files/2022-12/trl-assessment-tool-guide-final.pdf>
- ² <https://www.nasa.gov/directorates/somd/space-communications-navigation-program/technology-readiness-levels/>
- ³ Dresch, A., Lacerda, D.P., Antunes, J.A.V. (2015). Design Science Research. In: Design Science Research. Springer, Cham. https://doi.org/10.1007/978-3-319-07374-3_4
- ⁴ Standardization, I.O.f., Ergonomics of Human-system Interaction: Part 210: Human-centred Design for Interactive Systems. 2010: ISO.
- ⁵ Shore, L., et al., Technology acceptance and user-centred design of assistive exoskeletons for older adults: A commentary. 2018. 7(1): p. 3
- ⁶ Stipanovic, I., Skaric Palic, S., Kucinic L. et al (2023) D2.3: LCA/LCC KPIs, BEEYONDERS project, <https://beeyonders.eu/>
- ⁷ Eva Martinez, E., Sestrem, J., Folco, G. et al (2023) D2.4: Workers KPIs, BEEYONDERS project, <https://beeyonders.eu/>
- ⁸ https://cordis.europa.eu/programme/id/H2020_ECSEL-2017-2/es
- ⁹ A. Hemami, F. Hassani, An overview of autonomous loading of bulk material, Proceedings of the 26th ISARC, Austin TX, U.S.A. 2009, pp. 405–411.
- ¹⁰ S. Dadhich, U. Bodin, U. Andersson, Key challenges in automation of earth-moving machines, Automation in Construction, Volume 68, 2016, Pages 212-222, ISSN 0926-5805,
- ¹¹ Dadhich, Siddharth & Sandin, Fredrik & Bodin, Ulf & Andersson, Ulf & Martinsson, Torbjorn. (2020). Adaptation of a wheel loader automatic bucket filling neural network using reinforcement learning. 1-9. 10.1109/IJCNN48605.2020.9206849.
- ¹² Backman, Sofi & Lindmark, Daniel & Bodin, Kenneth & Servin, Martin & Mörk, Joakim & Löfgren, Håkan. (2021). Continuous control of an underground loader using deep reinforcement learning.
- ¹³ L. Zhang, J. Zhao, P. Long, L. Wang, L. Qian, F. Lu, X. Song, D. Manocha, An autonomous excavator system for material loading tasks. Sci. Robot. 6, eabc3164 (2021).
- ¹⁴ Schneider, E., S. Copsey, and X. Irastorza, *Osh [occupational safety and health] in figures: Work-related musculoskeletal disorders in the EU-facts and figures*. 2010: Office for Official Publications of the European Communities.
- ¹⁵ LAEVO, <http://en.laevo.nl/home-en/>.
- ¹⁶ SUITX, <https://www.suitx.com/>.
- ¹⁷ Sugar, T., et al., *Hip exoskeleton market-review of lift assist wearables*. 2018.



- ¹⁸ Toxiri, S., et al., *Back-support exoskeletons for occupational use: An overview of technological advances and trends*. 2019: p. 1-13.
- ¹⁹ Di Natali, C., et al., *Design and evaluation of a soft assistive lower limb exoskeleton*. Robotica, 2019. 37(12): p. 2014-2034.
- ²⁰ Poliero, T., et al. *Enhancing Back-Support Exoskeleton Versatility based on Human Activity Recognition*. in *2019 Wearable Robotics Association Conference (WearRAcon)*. 2019. IEEE.
- ²¹ Poliero, T., et al. *Assessment of an On-board Classifier for Activity Recognition on an Active Back-Support Exoskeleton*. in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*. 2019. IEEE.
- ²² Nussbaum, M.A., et al., *An Introduction to the Special Issue on Occupational Exoskeletons*. 2019, Taylor & Francis.
- ²³ Zhu, Z., A. Dutta, and F.J.A.i.C. Dai, *Exoskeletons for manual material handling—A review and implication for construction applications*. 2021. 122: p. 103493.
- ²⁴ Yoo, I.-g. and W.-g.J.W. Yoo, *Effects of different lifting and lowering heights on upper arm, shoulder and back muscle activity during a manual material handling task*. 2016. 53(1): p. 175-179.
- ²⁵ Schneider, S. P. (2001). Musculoskeletal injuries in construction: a review of the literature. *Applied occupational and environmental hygiene*, 16(11), 1056-1064.
- ²⁶ Holmström, E.B. (1992). Musculoskeletal disorders in construction workers related to physical, psychosocial, and individual factors, *Acta Orthopaedica Scandinavica*, 63:sup247, 55-55, DOI: 10.3109/17453679209155037
- ²⁷ CPWR – The Center for Construction Research and Training. (2018). *The Construction Chart Book - The U.S. Construction Industry and Its Workers*, 6th edition.
- ²⁸ Awolusi, I., Marks, E., & Hallowell, M. (2018). Wearable technology for personalized construction safety monitoring and trending: Review of applicable devices. *Automation in construction*, 85, 96-106.
- ²⁹ Mäkela S-M, Lämsä A, Keränen J, Liikka J, Ronkainen J, Peltola J, Häikiö J Bordallo, Lopez M, *Introducing VTT-Conlot: A realistic dataset for activity recognition of construction workers using IMU devices* IEEE Access
- ³⁰ Mäkela S-M, Lämsä A, Keränen J, Liikka J, Ronkainen J, Peltola J, Häikiö J, Bordallo, Lopez M, *VTT-Conlot: A realistic dataset for activity recognition of construction workers using inertial motion units*
- ³¹ Häikiö, J.; Kallio, J.; Mäkelä, S.M.; Keränen, J. IoT-based safety monitoring from the perspective of construction site workers. *International Journal of Occupational and Environmental Safety* 2020, 4, 1–14.
- ³² Bonabeau, E. Agent-based modeling: Methods and techniques for simulating human systems. In *Proceedings of the National Academy of Science* 99, 3 (2002), 7280--7287.]]
- ³³ Axelrod R. Axelrod R., (1997) *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration* (Princeton Univ. Press, Princeton, NJ)
- ³⁴ MACY, Michael W.; WILLER, Robert. From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology*, 2002, 28.1: 143-166.



³⁵ CARLOTTA, Belviso; RAINER, Hasenauer; BECHTOLD, Ulrike. Socially Assistive Robots Diffusion in Elderly Care A Pre-Adoption Study by Agent-Based Modeling. *Journal of Strategic Innovation and Sustainability*, 2018, 13.5: 58-75.

³⁶ J. Jimenez et al (2023) D2.2: Technical Requirements (2022), BEEYONDERS project, <https://beeyonders.eu/>

³⁷ Nicolas P., Javier H., Marco M., Ivan V. et al. D4.1: Modelling and Digital Twin Data Link, (2022), BEEYONDERS project, <https://beeyonders.eu/>

³⁸ Christian Di Natali; Sergio Leggieri; Pinar Sencandan; Vasco Fanti; Paolo Leone. D5.1: Exoskeleton Prototype (2023). BEEYONDERS project, <https://beeyonders.eu/>

