DARKO

Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

H2020-ICT-2020-2 Grant agreement no: 101017274

DELIVERABLE 8.5

Final demonstration

Dissemination Level: PUBLIC

Due date: month 54 (June 2025)

Deliverable type: Report

Lead beneficiary: Technical University of Munich (TUM)

Contents

1 Introduction

This document reports on the results of the final public DARKO milestone demonstration on June 5, 2025. The demo coincided with the last stakeholder meeting with invited industry and academia representatives at the KI.Fabrik lab, located at the Deutsches Museum in Munich.

Building on the system architecture from the third milestone and demonstration (MS3, D8.4), this demonstration presented the improvements realized on the mobile platform used for the previous milestones, and on the second mobile platform equipped with the elastic arm introduced in MS3. The first platform was used to showcase a complete scenario, including risk-aware navigation and the safe motion unit, navigation, perception, manipulation (including throwing with and without a pneumatic tool), and for additional demonstrations of other components and features (e. g., human and object detection, gesture recognition, motion prediction, and human-aware planning). The second platform is used to demonstrate the throwing performance of the elastic arm.

2 Integration events

Before the demonstration and stakeholder meeting, and similarly to the previous stakeholder meeting in 2024, the consortium conducted an extended integration phase at the University of Pisa, and at the ARENA2036 and KI.Fabrik facilities in Stuttgart and Munich, respectively. The integration spanned over several weeks in April, May and June, and focused on deploying, integrating, and testing all features developed during the final project period.

The three integration events in Stuttgart, Pisa and Munich involved both on-site and online collaboration, marked by an intense and productive work environment. Team members put in extended hours, often working late into the night and starting early in the morning. Figure 1 provides a glimpse into key moments from these integration weeks.

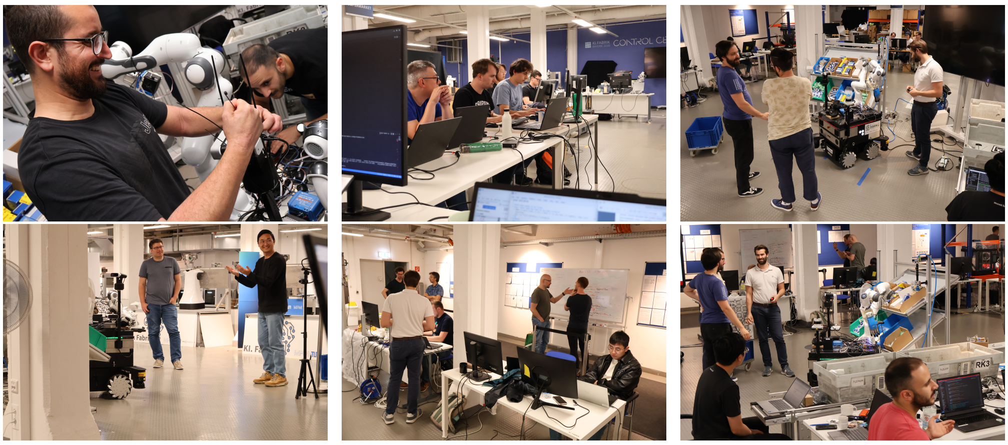


Figure 1: Snapshots captured during the integration session at KI.Fabrik, May–June 2025.

3 Stakeholder meeting

The milestone demonstration on June 5 2025 was performed as part of a stakeholder meeting with invited industry representatives. A video retracing this event is available via this link: https://www.youtube.com/watch?v=Q-M9PuG0DnE. Figure 2 shows some impressions from the meeting and demonstrations and Figure 3 details the agenda.

Figure 2: Demonstrations and presentations for the invited audience during the Milestone 4 stakeholder meeting.

The on-site audience included people from the Technical University of Munich (TUM), Munich Institute of Robotics and Machine Intelligence (MIRMI), Olive Robotics, KUKA, KION, Robco, Franka Robotics and Bosch Rexroth.

Similarly to the two previous milestone demonstrations, the meeting was divided in two sessions, each with presentations and live demonstrations. Additional information, such as the real-time visualization of the local map and cameras, was visible on dedicated screens.

After welcoming the participants, the event started in a seminar room with the first set of presentations: first, a project overview from the coordinator Achim Lilienthal, followed by a set of technical presentations from the work packages, including perception of objects, humans and gestures, mapping and localisation, human–robot spatial interaction, and motion planning. As for the previous milestone, the first set of presentations was concluded by a presentation of the KI.Fabrik by Achim Lilienthal.

We then proceeded to the demo area on the KI.Fabrik facilities in the Deutsches Museum for the first set of demonstrations, which consisted of showing the components related to online mapping, perception and recognition of human, objects and gestures, and risk- and human-aware navigation. These features, integrated in one of the mobile platform, are detailed in the following sections of the report.

The second presentation session was dedicated to the manipulation-related tasks, including the development and integration of the elastic manipulator and general-purpose gripper, planning and control for manipulation and throwing, perception for manipulation, and finally the risk representation and operations scheduling of WP7.

For the final set of live demonstration, located in the same demo area, we have separately presented the special features developed during the project, such as throwing objects with both the elastic arm and the elastic wrist. This demo session ended with a demonstration of all the features integrated into a dedicated scenario, which encompasses gesture detection to initiate the demo, navigation toward a shelf, grasping and throwing of an object in a box located on a conveyor. The contents of the demos are further described in Section 4.

After this demonstration, we concluded the stakeholder meeting with a joint standing lunch, socializing and discussing between the invited stakeholders and the consortium members.

Agenda of DARKO Stakeholder Meeting @ KI.Fabrik (5 June 2025)

09:30 – 09:40: Welcome (Large meeting room, first floor)

- DARKO in a Nutshell
  (Prof. Dr. Achim Lilienthal, TU München, Germany)

09:40 – 10:30: Presentations, Part I (Large meeting room, first floor)

- Perception I: Object and Human Detection, Gesture Recognition
  (Dr. Timm Linder, Robert Bosch GmbH, Germany)
- Mapping & Localization: Reconstruction, Introspection
  (Prof. Dr. Martin Magnusson, Örebro University, Sweden)
- Spatial Human-Robot Interaction I: MoDs, LHMP, Intent Communication, ARMoD
  (Yufei Zhu, Örebro University and Tim Schreiter, TU München, Germany)
- Motion Planning
  (Dr. Andrey Rudenko, Robert Bosch GmbH, Germany)
- KI.Fabrik in a Nutshell
  (Prof. Dr. Achim Lilienthal, TU München, Germany)

10:30 – 11:00 Demo, Part I (KI.Fabrik, basement)

- Object Perception, Human Perception, Gesture & Engagement Recognition
  (Dr. Narunas Vaskevicius, Robert Bosch GmbH, Germany)
- MoDs (Online Learning), MoD-LHMP, Intention Communication
  (Yufei Zhu, Örebro University, Sweden)
- Context-Aware MPC, Intent Communication: Human-Aware, Risk-Aware
  (Dr. Andrey Rudenko, Robert Bosch GmbH, Germany)

11:00 – 11:10 Coffee Break (Meeting rooms, first floor)

11:10 – 12:00 Presentations, Part II (Large meeting room, first floor)

- DARKO Platform
  (Dr. Valentin Le Mesle, TU München, Germany)
- Dynamic Manipulation
  (Prof. Dr. Paolo Salaris, University of Pisa, Italy & Yang Liu, EPFL, Switzerland)
- Perception II: Perception for Manipulation
  (Shih-Min Yang, Örebro University, Sweden & Prof. Dr. Martin Magnusson, Örebro University, Sweden)
- Spatial Human-Robot Interaction II
  (Luca Castri, University of Lincoln, UK)
- Orchestrator: Handling of Risks & Scheduling
  (Dr. Nicoló Mazzi, Spindox, Italy)

12:00 – 12:30 Demo, Part II (KI.Fabrik, basement)

- Throwing with Pneumatic Tool
  (Marco Baracca, University of Pisa, Italy & Giorgio Simonini, University of Pisa, Italy)
- Throwing with Elastic Arm
  (Dr. Valentin Le Mesle, TU München, Germany)
- Integrated DARKO Demo: Orchestrator, Navigation, Spatial Human-Robot-Interaction, Picking and Throwing with the SoftHand
  (Marco Baracca, University of Pisa, Italy & Dr. Nicoló Mazzi, Spindox, Italy)

12:30 - 14:00 Lunch (Meeting rooms, first floor)

Figure 3: Agenda of the stakeholder meeting on 5 June 2025.

4 Milestone demonstration

During a series of live demonstrations, we showcased the final implementation of core components for perception, planning and control for picking and throwing, human perception and scene understanding, human-aware mapping and human–robot spatial interaction, motion planning, task scheduling, as well as the mobile dynamic manipulation platform itself. The throwing capabilities of an elastic arm have also been presented on the second mobile platform.

The demonstrations took place in a designated area on the main floor of the KI.Fabrik facility. This space featured a custom-built setup with shelves containing boxed objects and a conveyor table with larger trays—mirroring the shelves used in DARKO's primary use case at BSH's warehouse. Similar to the BSH warehouse layout, and as in the milestone 2 and 3 demos (D8.3 and D8.4), the shelf and conveyor table were positioned close enough for efficient task execution but far enough apart that the robot could not reach both the picking (boxes) and placing (trays) points without either driving or throwing.

Similarly to the demonstrations presented in milestones 2 and 3, the audience stood near the demo area, enclosed by barrier tape, where they could observe the robot's actions directly while also viewing RViz visualizations on a large screen.

4.1 Dynamic Manipulation with the General Purpose Gripper

The DARKO robot's ability to manipulate objects involves picking them up from a shelf (see Figure 4) and throwing them to a desired location. Two separate platforms have been investigated and implemented, each showcasing a different approach. The first approach utilizes an anthropomorphic adaptive soft hand attached to a commercially available robotic arm through an elastic wrist. The second approach necessitated the development of a specialized elastic arm.

Figure 4: Dynamic manipulation with the general-purpose gripper

4.1.1 Picking and throwing objects from the shelf

In this section, we explore the part of the demo that showcases the picking and throwing actions performed during MS4. Similar to MS3, the demo presented a single cycle of the use-case scenario. The mobile base moved towards the shelf, stopping right in front of it. Subsequently, the system used its perception stack to acquire information about the location of the desired object. With this information, the manipulator picked up the object, and finally, the mobile base moved to its new position while the arm threw the grasped object towards the target box. To perform this demonstration, we integrated several components developed during the project. From the previous milestone, the rigid flange between the SoftHand 2 and the manipulator has been replaced by an elastic wrist (for more details, refer to deliverable D1.4). To enhance the predictability of the manipulator's arm movements for humans, as we did in previous milestones, we implemented a human-like motion planning algorithm that now includes an obstacle avoidance strategy. More information about the entire motion generation framework of the manipulator can be found in deliverable D4.2. The grasp pose for the arm was provided by VoteGrasp, a vision framework described in deliverable D2.2 and D2.3. This framework utilizes a learning-based algorithm capable of estimating feasible grasp poses for 2-finger grippers using as input directly the point cloud recorded by an RGB-D camera. To adapt the grasp pose for the anthropomorphic hand, we introduced an orientation offset to exploit the opposition between the thumb and the other fingers. For the throwing motion, we employed the

Figure 5: Single dynamic manipulation cycle of the use-case scenario

trajectory generation framework that leverages the elastic wrist, as described in deliverable D4.6. Figure 5 shows specific phases of the dynamic manipulation cycle.

4.1.2 Throwing with the pneumatic tool

In the MS3 demo, an updated version of the pneumatic tool was showcased, featuring a complete picking and throwing sequence. Two key innovations were introduced: first, the hand tool was integrated with a vision system based on AprilTag, enabling the robot to detect and interact with the tool autonomously. Second, the robot was programmed to execute the entire sequence. It utilized our human-like arm movement, as described in deliverable D4.2, to approach and pick up the pneumatic tool. Once the tool was acquired, the robot used it to pick an object whose position was assumed known from a box on the shelf. In the final step, the object was thrown into a target tray by inverting the pneumatic flow.

In MS4, the hand tool was also integrated with a vision algorithm capable of detecting stacked objects in a box and providing its configuration relative to the manipulator. Additionally, to compensate for the imprecision caused by the elastic wrist, a visual servoing algorithm was implemented to precisely position the sucker on top of the object.

Furthermore, we generalized the model for the pneumatic throwing mechanism to ensure broader applicability across various setups and compressor types.

4.1.3 Throwing with the elastic arm

As for the previous milestone, the throwing performances of the elastic arm (developed in the scope of T1.3 and detailed in D1.3) have been displayed using a second platform, depicted in Figure 6. The elastic arm, attached to one corner of the mobile platform, is equipped with the latest version of the Bi-Stiffness Actuator (BSA) at the final distal joint. While a magnetic gripper was used as an end effector for MS3, the manipulator is now equipped with a 2-finger gripper to demonstrate the possibility of throwing other kinds of objects. The integration of the arm on the mobile platform is the same as for MS3, and further details can be found in D1.5.

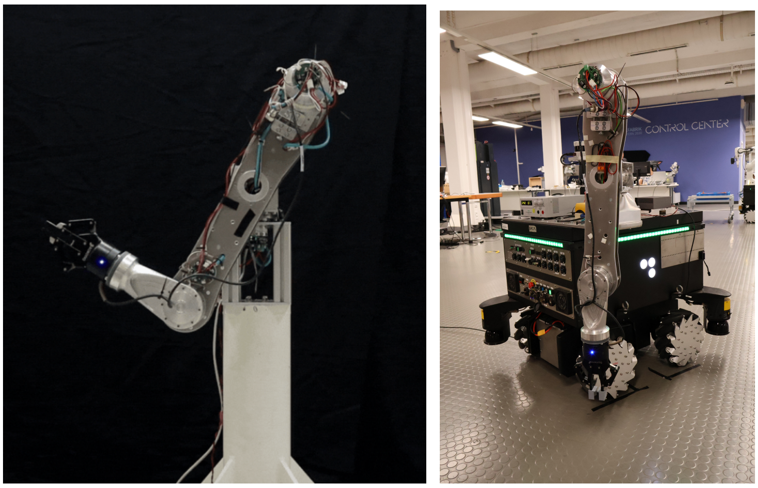


Figure 6: Snapshot of the developed 3-DoFs elastic arm equipped with a 2-finger gripper (left), and its integration on the second mobile platform (right).

Since MS4 is the final demonstration milestone of the DARKO project, it has been decided to emphasize on the elastic properties of the developed arm and its impact on throwing performances. As such, this demo aimed at throwing a small 3D printed sphere into two different boxes, as displayed in Figure 7. The first one, placed 2.5m away from the robot, shows the maximal distance that was reached during MS3 using a 'rigid' velocity-based controller developed by EPFL: the trajectory is optimized for the throwing, but the elastic properties of the BSA are not exploited. The second box is located 5m away from the platform, corresponding to the throwing distance when the elastic properties of the BSA are exploited. As for the previous milestone, the throwing motion is constrained to a fixed vertical plane to improve the safety around the robot and improve the reliability of throws. Finally, choosing 3D printed spheres as objects could simplify the implementation of the demo, and ensured that the location of the center of mass of the object and its inertia are precisely known.

During the demo session, the audience was invited to throw the small spheres into the different boxes. In addition to making the demonstration more interactive, this showed the audience that even if throwing seems to be an easy task for humans, performing it in a reliable manner can be challenging. Then, the spheres were manually placed in the gripper of the elastic arm, which threw them into the two boxes.

4.2 Human and Object Perception

The DARKO perception stack consisting of modules developed in WP2 (T2.1, T2.5) was demonstrated jointly with the "Causal Discovery" demo described in Section 4.3. A snaphsot of the demo as well as illustration of different outputs of the perception modules can be seen in Figure 8.

The demo showcased different perception capabilities from onboard RGB-D sensors including 3D human detection and tracking, articulated human pose estimation in 3D space, activity recognition and 3D object detection. The technical details of the demonstrated methods can be found in the deliverable D2.2. These technologies are necessary for various downstream applications and were utilized in other demos such as "Causal Discovery", "Risk- and Context-aware MPC", human-robot interaction, etc.

Beyond the stakeholder meeting, Bosch showcased intuitive task specification for

Figure 7: Throwing configuration during MS3 (left) and during MS4 (right). On the right picture, the box placed the closes to the robot corresponds to the throwing distance using only the rigid components of the BSA, while the second box indicates the longest throwing distance achieved by the elastic arm.

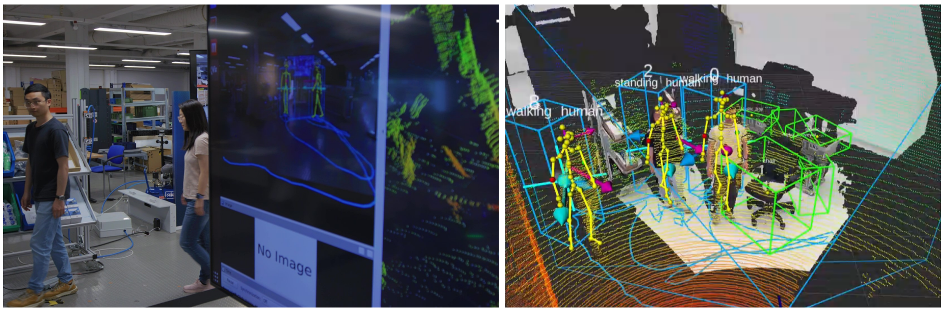


Figure 8: Human and object perception. The left subfigure shows impressions from the demo at KI.Fabrik during the stakeholder meeting. The right subfigure depicts example output of the DARKO 3D perception stack with tracked humans and their articulated poses (skeletons), detected objects and recognized activities.

a mobile robot within its robotics lab, utilizing DARKO's perception components. The demonstration leveraged the DARKO 3D object detector and a hand-held sensor suite (see D2.2 for more details) to semantically reconstruct the environment. This object-level semantic map was then used to specify navigation targets for a mobile ACTIVE Shuttle robot via pointing gestures (see Figure 9). DARKO's human perception stack played a key role in this interaction by detecting humans and their 3D articulated poses, which then were used to compute pointing direction and the target object.

The next milestone will focus on wide field of view perception based on fisheye and lidar sensors.

4.3 Causal Discovery

The "Causal Discovery" demo aimed to showcase the DARKO robot's ability to reconstruct a causal model of interactions between two people in a warehouse-like environment in real-time, directly onboard using data from its own sensors. Data acquisition, postprocessing, and causal discovery analysis were performed using the ROS-Causal framework (Figure 10), developed by WP5 as part of the DARKO project.

In the demo, the robot was positioned in a corner of the demonstration area, where it

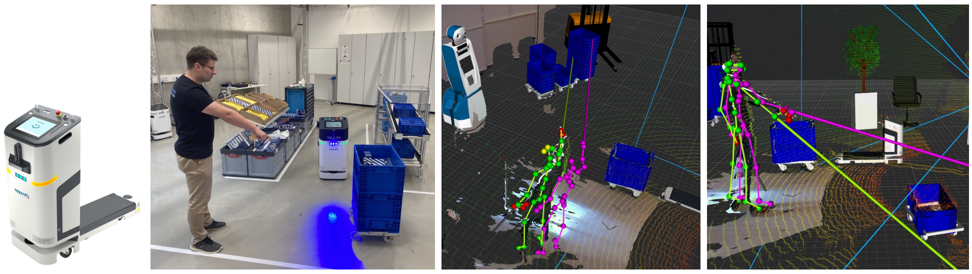


Figure 9: Intuitive robot control via pointing gestures demonstrated in Bosch robotics lab. From left to right: the ACTIVE Shuttle – an autonomous robot used to transport items e.g. in shopfloors of manufacturing plants; an example of pointing gesture; visualization of 3D pose estimation results by several methods from DARKO human perception stack, and the resulting semantic map with inferred pointing directions.

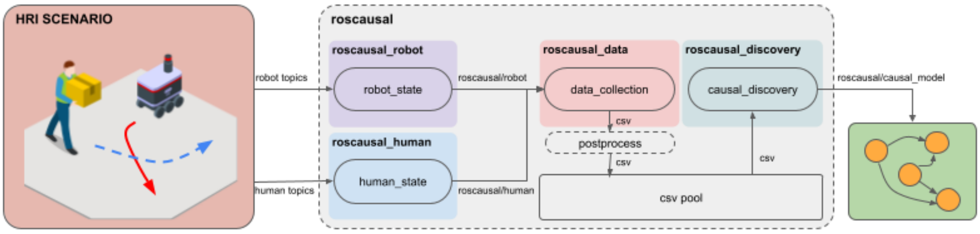


Figure 10: ROS-Causal pipeline: (i) data extraction from HRI scenarios; (ii) collection and post-processing of data to derive a high-level representation of the scenario; (iii) causal discovery conducted on the data, with the resulting causal model published on a dedicated ROS topic.

observed interactions between two people and their environment. Specifically, through ROS-Causal, the robot recorded data of human-human and human-target position interactions. It then postprocessed this data to derive the following three variables: v (human velocity); d_g (distance between the person and his/her target position) r (risk of collision with the other person). The robot then performed causal discovery analysis on these variables to reconstruct the causal model representing the relationships between them. (See Figure 11.)

4.4 Risk- and Context-aware Model Predictive Control

In the field of robotics, the ability to navigate autonomously in human-populated environments is a critical capability with numerous real-world applications, ranging from service robots in public spaces to collaborative robots in industrial settings. However, achieving safe and efficient navigation in such environments presents unique challenges, particularly in anticipating and adapting to human movements. As a result, there is a growing need for robot navigation systems to consider human motion and its predictions as integral components in their decision-making processes. Our team has made significant progress in developing a novel and efficient context-aware model predictive technique for generating robot trajectories. This technique takes into account various cues of humans, such as their short-term 2D predictions, body pose, and detected activity (e.g. standing, walking, sitting). By incorporating these cues into the trajectory generation process, our model predictive technique can better anticipate and adapt to human movements in real-time, leading to more efficient and safer robot navigation in human-populated environments. Furthermore, we have also integrated the notion of risk into our model predictive control.

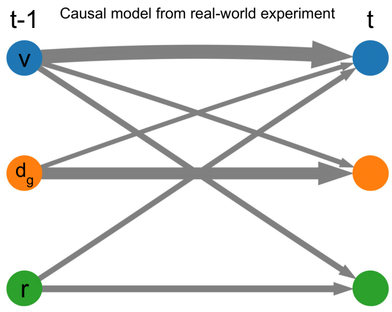


Figure 11: Causal model reconstructed by ROS-Causal using robot sensor data collected during the demonstration.

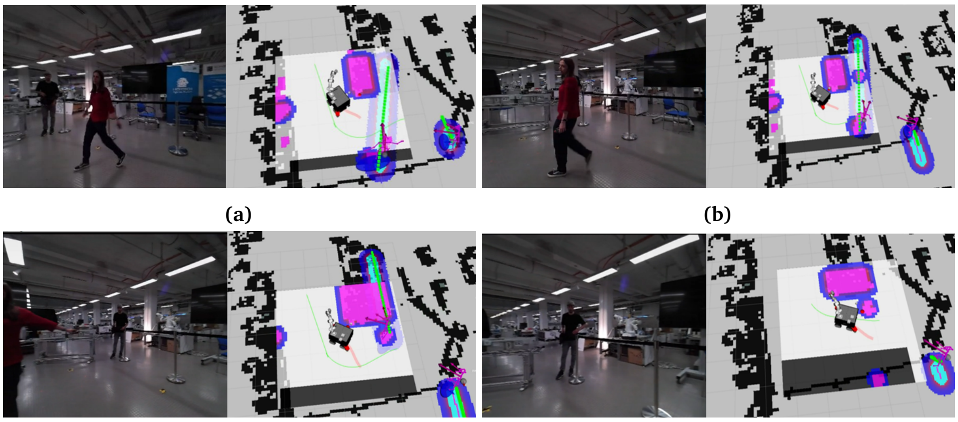


Figure 12: Rviz visualization and onboard camera image, while the robot moves in the KI Fabrik. The risk map is centered on the robot and includes detected objects, and static known risky areas (near the retractable belt stanchions).

This includes considering the risk of collisions with both static and dynamic obstacles in the environment. By incorporating risk assessment into the trajectory generation process, our technique can proactively avoid potential collisions and navigate the robot through complex and dynamic environments with greater safety and reliability. The culmination of our achievements was demonstrated in the milestone demo, where we showcased the successful integration of both the context-aware model predictive technique and the risk assessment component into the overall DARKO project navigation architecture. The demo highlighted the capabilities of our approach in enabling robots to navigate autonomously in human-populated environments while effectively considering and adapting to human movements and potential collision risks (see Figures 12 and 13).

4.5 Bidirectional communication of intent

The milestone demonstration also included an implementation of bidirectional communication of intent from T5.2. This module uses anthropomorphic intent communication by means of a Nao robot seated on top of the mobile DARKO platform (seen in Figure 15). The Nao robot acts as a "driver" and signals its intentions with human-like gestures like pointing to an object of interest or moving its head to make eye contact with a detected

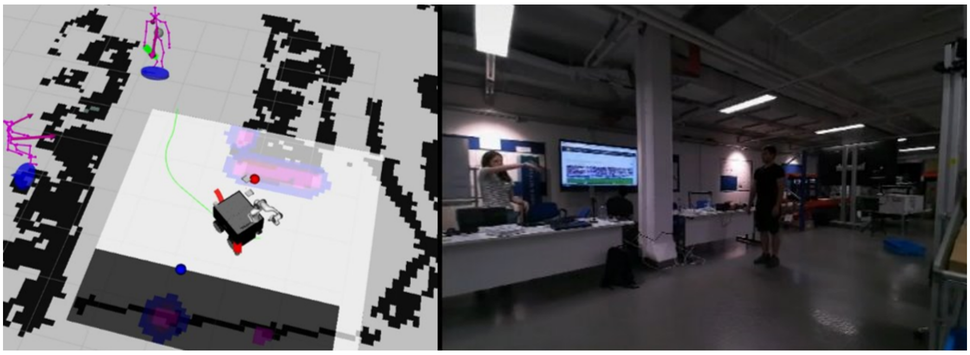


Figure 13: Focus on human skeletons detection and the constraints (blue and red markers) derived from the risk map.

person, as well as spoken language. The human-to-robot intent communication in this demo was implemented using a speech recognition interface, where a user could talk about objects in the view of the robot's camera and the Nao robot would point to the object in question and talk back about it.

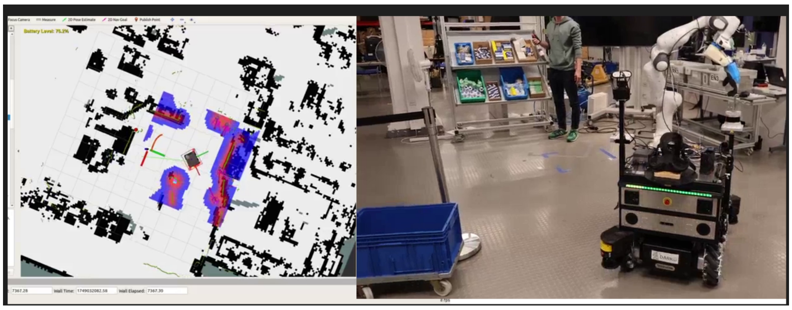
4.6 Orchestrator rescheduling

In this demo scenario, we demonstrate the risk-aware rescheduling capabilities of the Orchestrator. Its core objective is to assign missions that the robot can complete both successfully and safely. The robot is assigned a mission that specifies which objects to pick and the trays where they should be placed or thrown. In this scenario, the mission involves two objects (see Figure 14a):

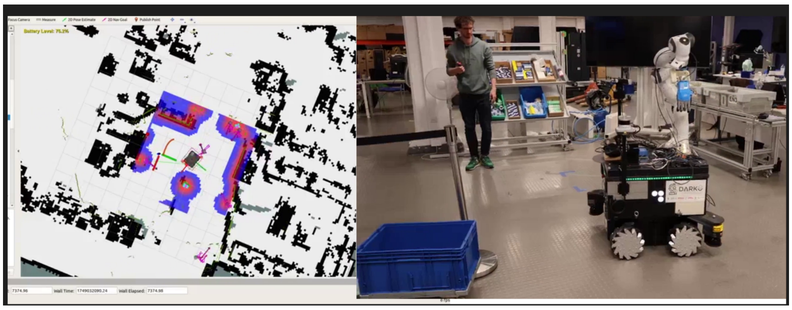
- *Object A*, located on the left side of the shelf
- *Object B*, located on the right side of the shelf

At the start of the mission, a human is standing in front of the picking location for *Object B*. Upon receiving a mission, the Orchestrator constructs a fully connected graph, where each node represents a location suitable for manipulation tasks—such as picking, placing, or throwing objects. Using historical success data, picking positions, locations of boxes and trays, risk maps, the Orchestrator estimates the probability of success for manipulation tasks.

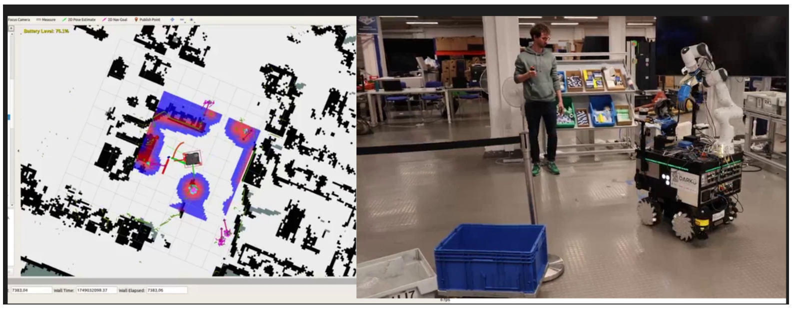
To plan safe and efficient navigation, the Orchestrator uses the risk maps to compute the optimal path for every possible movement in the graph. This includes estimating both travel time and associated risk. Initially, the Orchestrator assigns the robot to move toward *Object A* picking location (see Figure 14a), consistent with the safe state of the environment at mission start. Throughout the mission, the Orchestrator performs continuous risk assessment to verify that each current and upcoming action remains safe. If the estimated risk exceeds a predefined threshold, an alarm is triggered, and the Orchestrator halts the current task immediately. In the demo, the person later moves to block access to *Object A* picking location. The system detects this increased risk and recognizes that a safe approach is no longer possible. As the robot attempts to reach *Object A*, the person stands directly in its path. The Orchestrator reacts by stopping the robot (see Figure 14b) and replans the mission in real time. Finally, as shown in Figure 14c, the Orchestrator assigns a new task to pick *Object B*, which has now become safely accessible. The human than moves away from the working area and the Orchestator can safely guide the robot to mission completion.



(a)



(b)



(c)

Figure 14: Orchestrator demo. Figure a) shows the scene at the beginning of the mission, with a human standing in front of one of the picking location from the shelf. Figure b) shows the moment when the human moves from the right side to the left side of the shelf, blocking the picking area where the robot was planning to move and freeing the previously blocked one. Finally, Figure c) shows the robot re-planning towards the safest picking area.

4.7 Full Scenario

Figure 15: Impressions from the full scenario demonstration. DARKO robot approaching the shelf (left), a successful grasp of an item from a bin (middle), delivery and throw of the item into the target tray (right)

As the first event of the milestone 3 demonstration, we showed a complete scenario, orchestrated by the WP7 task scheduler, including localisation, navigation, perception, picking and throwing. See Figure 15. In this demo scenario, the robot first had to navigate to the shelf and position itself accurately in order to look inside the boxes on the shelf and reach the objects inside. We then regress grasp poses from the point cloud of the top-mounted RGB-D camera (T2.2) and grasp it with the Softhand 2 gripper (T4.1). After picking, the hand is held up for the camera to do in-hand perception, verifying that we have a good grasp for throwing. Finally, the robot drives a short distance to be within throwing distance of the conveyor with the target tray, and computes a throwing motion to throw the object into the target tray.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101017274