DARKO

Dynamic Agile Production Robots That Learn and Optimise Knowledge and Operations

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

DELIVERABLE 8.3
First demonstration

Dissemination Level: PUBLIC

Due date: month 29 (May 2023)

Deliverable type: Report

Lead beneficiary: Robert Bosch GmbH (Bosch)

Contents

1 Introduction

This document reports on the outcome of the public DARKO milestone demonstration on June 16, 2023. The demo coincided with the first stakeholder meeting with invited industry representatives, at ARENA2036 in Stuttgart.

Building on the system architecture from the first milestone and demonstration (MS1, D8.2), in this demonstration we have substituted several of the initial mock-up modules with real functionalities (e. g., grasp pose estimation, throwing, task scheduling), and consolidated and tested the integration of further components into the running system (e. g., risk-aware navigation and the safe motion unit). We have demonstrated a limited use-case replica consisting of navigation, perception, manipulation (including throwing); and showed separate demonstrations for further components (e. g., human and object detection, motion prediction, and human-aware planning). In particular, the use-case mission includes both motion and manipulation; we showed how certain induced risk factors can affect the execution of plans; we showed tracking and classification of 3D articulated poses of humans; and we showed integration of perception and manipulation for an initial demonstration of picking and throwing of objects relevant for the use case.

2 Integration events

Prior to the demonstration and stakeholder meeting, the consortium carried out an extended integration phase at the ARENA facility in Stuttgart through several weeks in April–June. The primary goal during this phase was the deployment and integration of all the new features that had been developed and refined in the preceding months. Spanning four weeks of on-site collaboration, this phase of the project was intense and highly productive – and characterised by extended working hours, including late nights and early mornings. Figure 1 shows some impressions from these integration sessions.

3 Stakeholder meeting

The milestone demonstration on June 16 was performed as part of a stakeholder meeting with invited industry representatives. Figure 2 shows some impressions from the meeting and demonstrations.

The on-site audience included people from ISG (Industrielle Steuerungstechnik GmbH), NAiSE, Franka Emika, Schaeffler, BSH Hausgeräte, Bosch Rexroth, Bosch Engineering, Stuttgart Media University, and University of Stuttgart. (The flyer sent out to our list of invitees is shown in fig. 3.)

The meeting was divided in two slots, each with presentations and on-screen demonstrations of pre-recorded activities plus live demonstrations.

After registration of the participants, the meeting started in a seminar room with a set of presentations: first, a project overview from the coordinator Achim Lilienthal, followed by a set of popular-science presentations from the work packages, including perception of objects and humans, mapping and localisation, human–robot spatial interaction, and motion planning.

We then proceeded to the demo area on the main ARENA2036 exhibition floor for the first set of demos. The contents of the demos are further described in section 4 but contained people detection (T2.5) and object-level semantic perception (T2.1), picking and throwing with a pneumatic tool developed by UNIPI (T4.4), prediction of human motion (T5.1, T3.5), and motion planning (T6.3, T6.4, T3.5). (We had also planned a

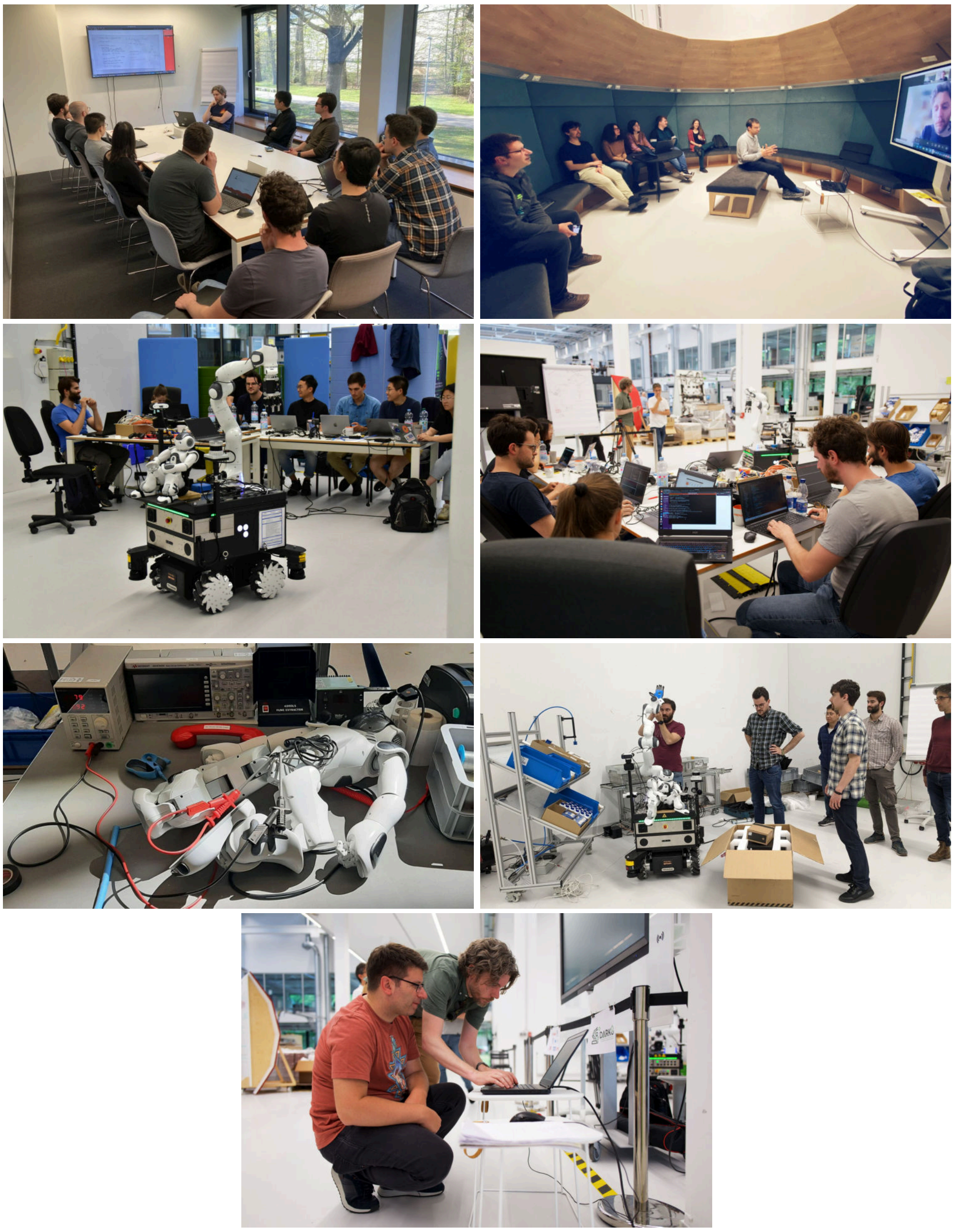


Figure 1: Snapshots from the integration sessions at ARENA2036, spring 2023.

Figure 2: Presentations and demonstrations for the audience invited to the milestone 2 stakeholder meeting.

Figure 3: Flyer emailed to the invited audience for the stakeholder meeting.

demo of human–robot intent communication from T5.2, but were not able to show it due to illness of the lead implementation responsible.)

The second presentation slot focused more on the manipulation related tasks: the development 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 demo slot, back in the main demo area, we showed a small scenario integrating several of the project components in a complete task: starting from a point outside the demo area, driving to the shelf, observing the objects on the shelf and picking one of them (using perception and manipulation components from WP2 and WP4), observing the in-hand grasp pose to ensure the object was picked, driving near to the target tray, and throwing the object into the tray – all of this controlled by a prototype task scheduler from WP7.

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

4 Milestone demonstration

We showcased, in a set of live demonstrations, the current 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 demo was performed in a dedicated area of the main showroom floor of ARENA2036, shown in fig. 4. In this area we have constructed a set of shelves with objects in boxes and a table with rolling pins (conveyor) with larger trays – similar to the shelves in DARKO's main use case at BSH's warehouse. As in the BSH warehouse, the shelf with boxes and the conveyor table with trays are located near to each other, but far enough so that the robot cannot reach both the picking point (boxes) and placing point (trays) without driving, or throwing.

The audience could stand next to the demo area, delimited by band barriers, and watch both the actions of the robot and at the same time Rviz visualisations on a large screen.

4.1 Manipulation with pneumatic tool

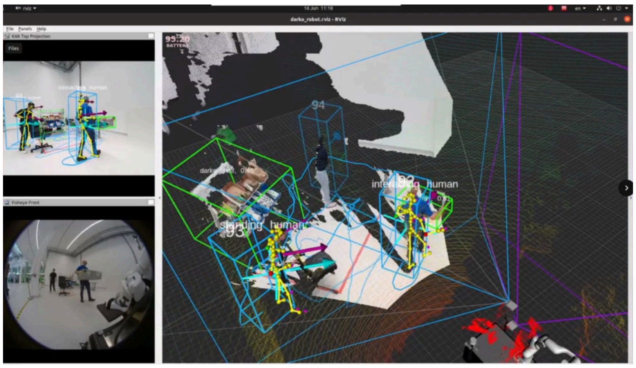
As the first demo item, we showed picking and throwing using a bespoke pneumatic tool developed by UNIPI, as shown in fig. 5. (See also D1.2.) In this mode of throwing, the pneumatic tool is mounted on the side of the shelf, and the robot is stationary. This demo was shown primarily to show the feasibility of this particular tool, and was not integrated with the main perception and planning components. The robot uses a *qb SoftHand* to pick the pneumatic tool and then uses the tool to pick an object from the box. After picking the object, it is "shot" into the target tray by inverting the flow of the pneumatic tool.

4.2 Human and object detection

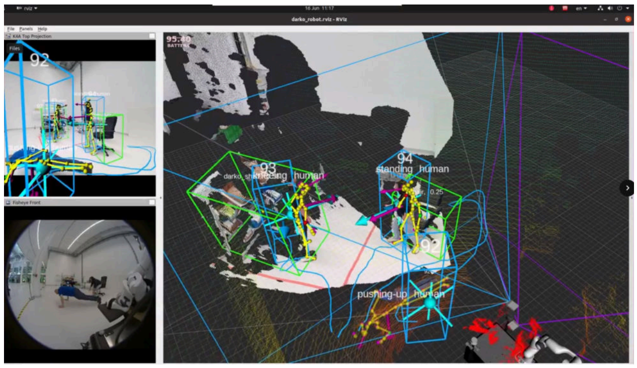
With the robot parked in a corner with a good view of the demo area, we demonstrated the object detector from T2.1 and the people perception module from T2.5, as shown in fig. 6. Objects are detected with an extended 9DOF RGB-D YOLO++ architecture (as further described in D2.2) which outputs 3D bounding boxes of objects of interest. Humans are detected with a similar RGB-D YOLO++ architecture, body poses for the

Figure 4: Demo area for the milestone 2 demonstration at the stakeholder meeting in June 2023. In front of the robot is the shelf with objects to be picked, and to the right is the table with trays where objects should be placed.

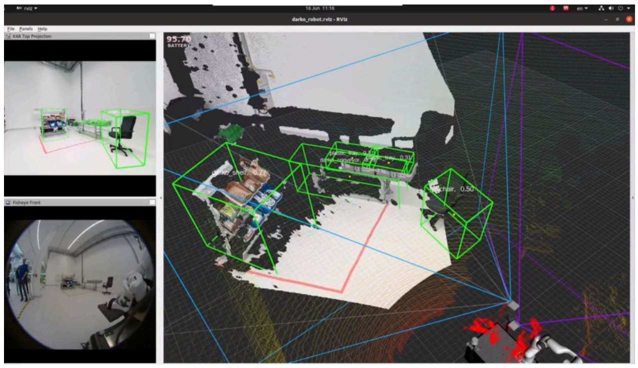
Figure 5: Demonstrating picking and throwing with a bespoke pneumatic tool.



(a) Estimating articulated human poses and tracking humans, including classification of human activities. Note the "standing" human vs the "interacting" human who is carrying a box.



(b) Estimating articulated human poses and tracking humans. Note the "pushing-up" human.



(c) Object-level semantics, detecting objects of interest from Kinect RGB-D data.

Figure 6: Human and object detection from the milestone 2 demonstration. In each figure, the large frame shows a perspective view of the multimodal sensor data and the labelled detections, the top-left frame shows the RGB data from the Azure Kinect camera, and the bottom-left frame shows the front-facing fish-eye camera.

detected humans are estimated with MeTRAbs and passed to a body pose classifier, while a Kalman filter-based tracker tracks the centroids of the detected persons. (These modules are further described in D2.2.)

4.3 Motion prediction

Prediction of human intents and their future motion is important because it enables a robot to take more informed decisions by anticipating behaviours and reacting in a timely and safe manner. Following the objectives of DARKO, this is especially important when striving for efficiency in time, where waiting for human actions to be completed could lead to unnecessary delays. The components for motion prediction aim to inform the motion planning and control pipeline of the robot to enable predictive planning and collision avoidance. Furthermore, predictions will enhance the perception system in periods of short occlusion to improve the safety aspects of the robot.

In this demo we have shown two methods for trajectory prediction running live on the robot: CliFF-LHMP and IOCMM (see fig. 7). Both methods are informed by the DARKO perception stack, which provides consistent tracking of position and velocity for

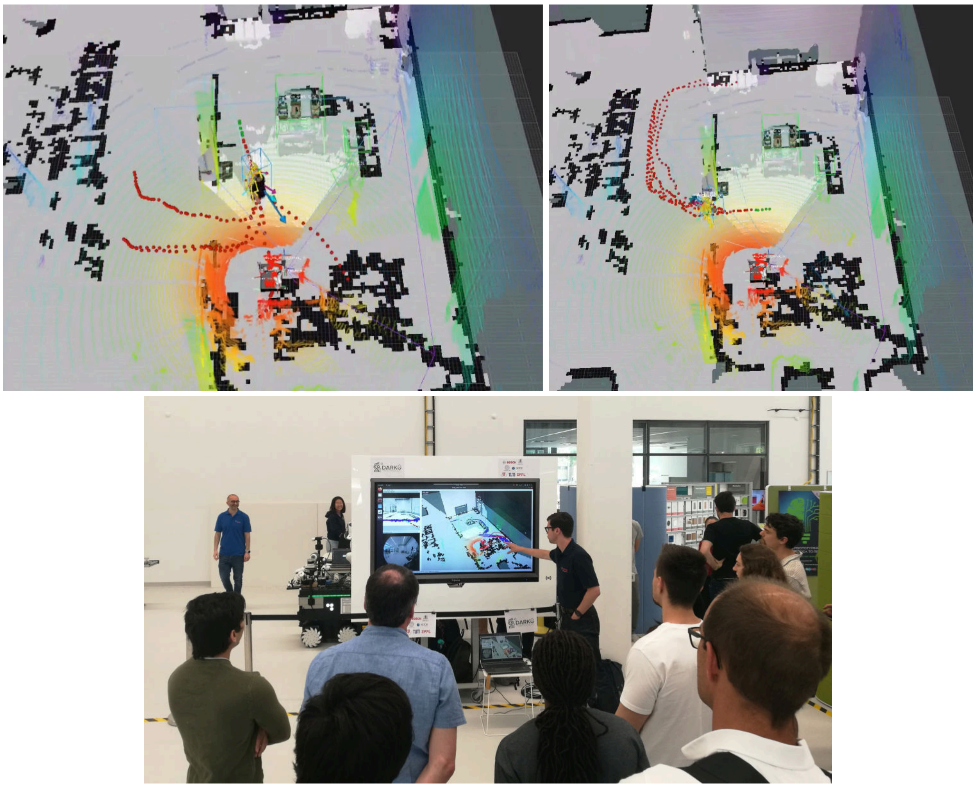


Figure 7: Human motion prediction shown live to the stakeholders.

each detected person. CLiFF-LHMP uses collected data on human motion in a specific environment to build a Map of Dynamics (MoD), and uses this MoD to predict the future states of every detected person. IOCMM, on the other hand, does not require collected data, but predicts optimal motion towards pre-defined "goal points" in the environment. The flow map in this case was created with data collected from the DARKO robot observing people walking in scenario-relevant patterns to/from the shelves set up in the demo area.

4.4 Motion planning

We also showed planning and execution of motions using the CLiFF-RRT* algorithm. This motion planning algorithm uses a 2D occupancy grid map as well as a CLiFF-map (as mentioned above) which encodes the usual flows (of people) in the environment. Flow-aware motion planners such as CLiFF-RRT* can take maps of dynamics into account to generate motions that align with the expected direction and speed of motion observed throughout the map – which we showed in these demonstration (fig. 8). We used the same maps as for motion prediction (above). (The current implementations of motion planning in DARKO are further described in D6.3.)

4.5 Safe motion

The current implementation of the generalised safe motion unit (gSMU) from WP6 was demonstrated, showing that it sets safe velocity limits when a person is detected near the robot – based on the reflected mass of the mobile platform and the injury analysis and

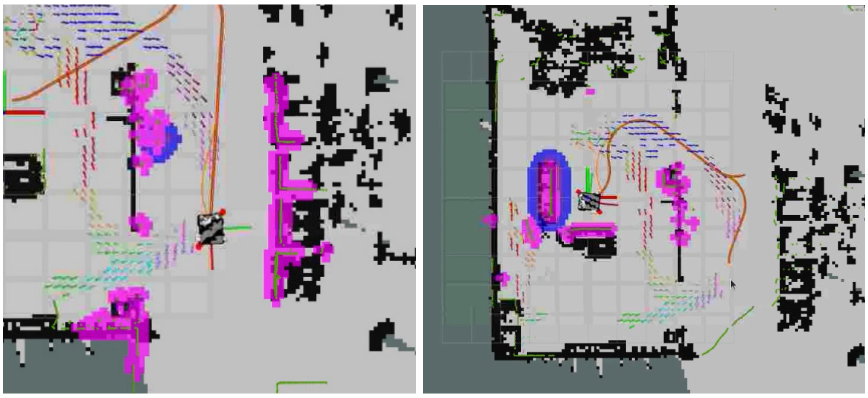


Figure 8: Example paths computed with CliFF-RRT* during the DARKO stakeholder meeting at ARENA2036. The CLiFF map of dynamics is shown here as a vector field with coloured arrows. *Left:* the computed path goes against the usual human movements. *Right:* the path adheres to the usual flow.

collision modelling that make up the gSMU – and resets to the maximally allowed velocity limit when the person is no longer present. The main objective of the gSMU is to reduce velocity to prevent human injury, but to do so in an *efficient* and *informed* way, so as to avoid unnecessarily conservative velocity reductions but still adhere to biomechanically safe limits.

4.6 Full scenario

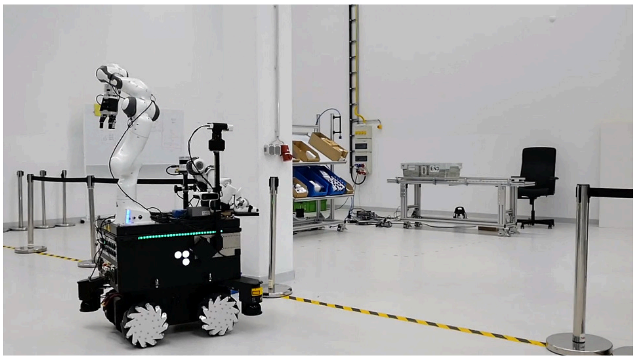
As the final event of the milestone 2 demonstration, we showed a complete scenario, orchestrated by the WP7 task scheduler, including localisation, navigation, perception, picking and throwing. See fig. 9.

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 (fig. 9a). We then regress grasp poses from the point cloud of the top-mounted RGB-D camera (fig. 10) We select the top-rated grasp pose and grasp it with the DH-3 gripper (fig. 9b). After picking, the hand is held up for the camera to do in-hand perception, verifying that we have a good grasp for throwing (fig. 9c). 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 (fig. 9d).

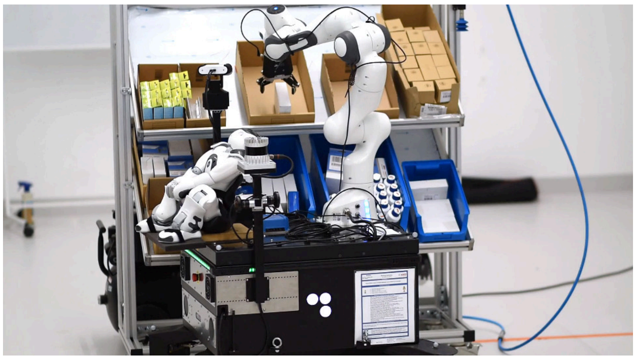
5 Automatica 2023

In addition to the milestone planned in the grant agreement and described above, we also got the opportunity to present at the Automatica trade fair in Munich, June 27–30 2023. DARKO's booth was part of the "munich I" AI.Society floor area at Automatica. Figure 11 shows some of the activities in and around our booth.

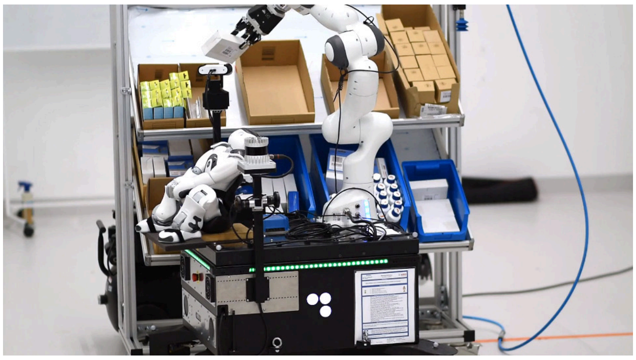
The robot platform used at ARENA was transported to Munich, together with the shelves and objects used in our demo scenario. A replica of the demo space was built in our booth, and we demonstrated several hundred picking and throwing cycles. We also demonstrated live people tracking and activity detection as well as human motion prediction from the robot's on-board sensors.



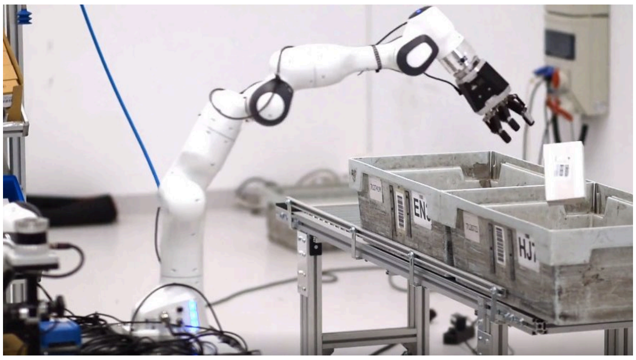
(a)



(b)



(c)



(d)

Figure 9: Final demo at the milestone 2 stakeholder meeting: move + pick + throw scenario.

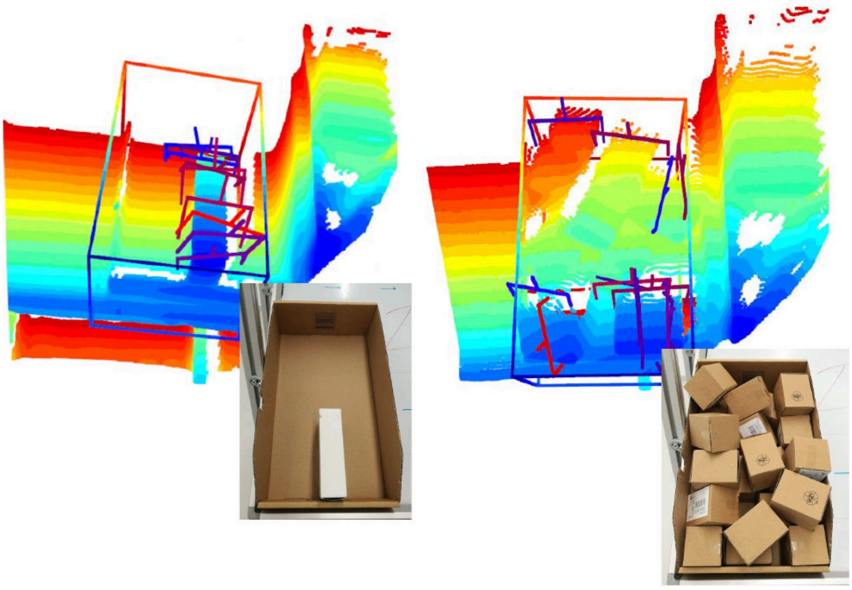


Figure 10: Generating grasp poses from a DARKO box with objects.

Figure 11: DARKO presenting at Automatica 2023.

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