**Physical AI — the path to a beautiful tomorrow.**



# PRAVDA REPLIKANTA

THE REPLICANT TRUTH
A NEWSPAPER ABOUT PHYSICAL ARTIFICIAL INTELLIGENCE

**BREUKELEN** (Netherlands)
Forecast · 2026-11-03 · CET

| | | |
|---|---|---|
| MORNING · 09:00 | 9° | Cloudy<br>Wind: NE 4.4 m/s · 11% |
| MIDDAY · 12:00 | 10.4° | Clear<br>Wind: NE 4.3 m/s · 6% |
| EVENING · 18:00 | 8.1° | Clear<br>Wind: NE 4.1 m/s · 0% |

Daily: 5.2–11.1 °C
Precipitation probability: 21%

*All we can do is learn to be human while the rain falls.*

**ISSUE № 6 · 2026-11-03** rev. 1 ·

Weather source: Open-Meteo · Retrieved: 2026-11-03 05:25 CET

## Robotics: data, contact and safety

Five studies examine how action representations, video, contact and risk checks alter robot performance.

**News**

# WorldBridge uses unlabelled task videos to reach 55.90% success

Imagine watching someone put objects on shelves without being told which movements to make. Learning from that video alone would be difficult for a robot too. A research team led by Jun Cen investigated whether videos could help robots perform tasks that were missing from their usual training examples.

The researchers used learning from video. One part of the system predicts what the next moments of an activity might look like. Another part chooses movements that could produce those pictures. This is not the same as copying a person directly. A robot has different hands, proportions and limits, so its own movements must still be calculated.

In a computer test of unfamiliar tasks, the starting system succeeded 25.07% of the time. Adding videos of those tasks, without recorded movement commands, increased success to 55.90%. Providing just one example with commands for each new task raised it to 70.70%. Those numbers describe a particular test collection, not every robot in every room.

WorldBridge combines separate video-prediction and action-generation experts, connected by structured attention and trained using joint flow matching. Its visual memory operates on two temporal scales. The authors distinguish video-only coverage of unfamiliar tasks from adding one action-labelled demonstration per task. The latter raises success to 70.70%, but requires a different supervision budget. Treating these two results as though they used identical training data would obscure the contribution of the action labels.

The experiment offers a reason to collect useful videos as well as demonstrations with movement commands. It does not mean a household robot can watch one clip and immediately cook dinner or assemble furniture. The researchers tested specific tasks under controlled conditions.

The reported 96.23% on the full RoboTwin 2.0 setting belongs to a different protocol from the unfamiliar-task experiment. Transfer to another gripper, camera arrangement or physical robot is not established by these figures. A reproducible comparison must specify exactly which task videos and action labels were available.

**Sources**

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

# 281 policies tested on real connector insertion

Putting a small network connector into its socket is harder than it looks. The plug must face the right way, line up and enter without slipping. A team led by Ammar Abbas compared ways to teach a five-fingered robot hand this task. They carried out roughly 1,700 attempts on real equipment, not just on a screen.

The team tested 281 versions of control software from seven groups. They changed what information the program received from cameras and how it described finger movements. This is called a data representation: the same task can be described using different pictures and numbers. Sometimes that choice mattered more than ordinary adjustments to the training process.

The researchers also examined 67 scores that could be calculated without performing the insertion. Only five,

related to how smooth and direct the predicted motion was, closely matched real success. Even those worked only within one group of programs. A movement that looks excellent in a computer test may still miss the actual socket.

For point-cloud policies, cropping the relevant region, using robot-base coordinates and normalizing features improved success by 70 to 100 percentage points in individual controlled comparisons. For a vision-language-action policy, absolute rather than relative actions improved insertion success by up to 70 points. These figures come from different configurations and cannot be added together or generalized to every architecture. Seven policy families were included, exposing how representation choices interact with model design.

The findings concern particular hands, connectors and test conditions. They do not establish the best program for every object. They do explain why a robot must be tested with the things it will really handle.

Only five of 67 offline measures reliably correlated with real insertion success, and only within one policy family. Correlations between minus 0.70 and minus 0.80 for predicted motion smoothness and directness do not turn these scores into substitutes for hardware trials. Selecting a controller from validation loss alone is particularly weakly supported here.

**Sources**

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

## Contact surfaces replace literal copying of finger joint angles

When you pick up a glass, bending your fingers is only part of the job. They must touch the glass in useful places or it may slip. A team led by Zhixuan Xu studied how to teach robot hands using the changing places where fingers contact an object. Their system includes pictures from cameras near the fingertips and information about contact surfaces.

One part records where human fingers touched an object. Another helps a robot choose contact points and wrist movements of its own. This matters because a human hand and a robot hand can be shaped quite differently. Simply copying the angles of human finger joints may not let a robot hold the same thing.

The researchers compared their approach with other methods on six combinations of robot hands and tasks. They report that average success in the final stage improved by at least 50 percentage points. A percentage point measures the difference between two success rates. For example, moving from 20% to 70% is an increase of 50 points.

The architecture separates demonstration capture from policy learning. Its demonstration interface combines shared contact geometry with colour fingertip views, aligning contact before retargeting the pose. The learning interface predicts relative contact-surface trajectories and then coordinates wrist and finger commands through a lower-level controller. The team has published software and hardware designs, allowing others to attempt experiments with different robotic hands. This is a contact-based representation rather than a claim that human and robot joint angles are interchangeable.

The tests used particular hands and tasks. They do not prove that a robot can now handle every cup or tool. But the experiment helps explain why similar-looking finger movements can produce very different results when the fingers touch an object differently.

The at-least-50-point gain concerns the final stage of six hand-task combinations, not every complete manipulation sequence. Different materials, calibration errors and initial approach motions require further tests. The quantity preserved is the pattern of physical contact, not the visual resemblance of finger poses.

**Sources**

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

## TRACE reduces person-tracking collisions with selective intervention

Imagine a robot following one person down a busy corridor. Someone wearing similar clothes crosses in front of it, and the robot may confuse the two people. A team led by Mingyang Lyu studied how improving a robot's ability to keep following someone affects the safety of its movements.

Training the software to keep track of its person made it better at following. But in some tests it also increased collisions. The researchers therefore added a separate danger check. It looks at clues about the target person's location and how confidently the robot can tell that person apart from others.

When the risk is high, the check shortens the next forward movement. It does not change the robot's turning command. In the experiments, more than 98% of steps remained unchanged. In a difficult tracking test, the collision rate fell by 3.96 percentage points compared with the version without this added check.

The authors fine-tuned only the planner head of a compact vision-language-action model while freezing the backbone. A separate risk predictor processes 84 cues, including candidate targets, identity margins, estimated range and bearing, and the proposed command. Above a calibrated risk threshold, it shortens the next translational waypoint without changing the turning component. This targets dangerous moments instead of applying a blanket speed reduction or replacing the learned controller with a wholly different planner.

That does not make the robot safe in every crowded shopping centre. The researchers compared programs under particular test conditions and did not claim protection from every possible mistake. Interestingly, simply slowing the robot all the time did not produce the same improvement.

A concurrent evaluation found a 3.77-point collision reduction versus 0.29 points in matched null pairs. The approach is not a formal safety guarantee beyond the studied tracking scenarios. Constant slowing without target-state information did not reproduce the benefit, distinguishing the proposed filter from a simple speed governor.

**Sources**

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

## Hourglass shares navigation across robots with different dynamics

A small wheeled robot and a larger robot with a different steering system cannot move through a narrow corridor in exactly the same way. The command to turn right produces different motions. A team led by Zhidong Wang proposed teaching one program to plan paths for several kinds of robots while still respecting those differences.

The method uses a world model: software that predicts how the scene around a robot may change after it moves. But pictures alone are not enough. The researchers also describe how each particular robot accelerates and turns. The planner can therefore consider what the machine can actually do, not merely which space looks empty.

In computer experiments, the researchers used seven robots representing five kinds of movement. The shared program succeeded 3.8 to 10.6 percentage points more often than some comparison methods. They then drove a real robot manually for three minutes to learn how it moved, before testing navigation in an unfamiliar indoor space.

The method couples a shared world model with a differentiable reduced-order model of each robot's motion through the change in camera viewpoint. The world predictor learns from multiple platforms, while kinematic parameters are identified separately. During imagined rollouts, a proposed command first passes through the robot template and then the visual world predictor. This makes the planner sensitive to turning and speed limits without training an entirely separate policy for each platform.

That real-world test matters, but it does not show that the same program works for every vehicle or on staircases. The study specifically discusses movement on flat ground and particular robot designs.

In the physical experiment, the researchers identified a new robot's parameters from three minutes of teleoperated driving and then tested unfamiliar indoor navigation. The flat-ground assumption does not cover stairs, steep slopes or contact-heavy manoeuvres. The numerical baseline gains were measured in simulation; the physical test supports a narrower transfer claim.

**Sources**

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

**Analysis**

## Representation choices change outcomes

**Performance depends not only on model training but also on coordinate frames, contact data and execution constraints.**

Results from different protocols cannot be directly compared without checking hardware and available data.

**Sources**

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

## Issue glossary

**Key terms in this issue.**

01 **WorldBridge**

A model linking predicted video with robot actions.

02 **RoboTwin 2.0**

A simulated manipulation benchmark for comparing methods.

03 **action representation**

The format used to express robot commands and coordinates.

04 **point cloud**

A collection of three-dimensional points representing surfaces.

05 **Contact Imitation Interface**

A skill-transfer method based on fingertip contact geometry.

06 **contact trajectory**

A sequence of positions of the touching region.

07 **TRACE**

A risk filter for person-following robot movements.

08 **risk estimation**

Computing a risk score from features of the situation.

09 **Hourglass**

A shared navigation architecture with robot-specific dynamics.

10 **world model**

A predictor of how the observed scene changes after an action.

11 **RJ45 connector**

A network connector used in the five-fingered hand insertion experiments.

**Sources**

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30