

# MolmoBot: An AI robot that learns from simulated data instead of the real world.

Ai2 introduces MolmoBot – an AI robot system trained entirely on virtual simulation data. This technology can reduce robot training costs and expand physical AI capabilities.

Virtual simulation data is becoming a key driver of the development of **physical AI** in enterprise environments, with notable projects such as **Ai2's MolmoBot** .

Previously, training hardware to interact with the real world relied on recordings of human-performed actions, a process that was both costly and time-consuming. Companies developing multi-purpose robotic systems often consider training with large amounts of real-world data as fundamental to building these AI agents.

**For example, the DROID** project collected approximately 76,000 remote sensing orbital trajectories from 13 different organizations, equivalent to about 350 hours of human labor. Meanwhile, **Google DeepMind's RT-1** model required 130,000 test runs collected over 17 months by operating technicians. This reliance on manual and proprietary data significantly increases research costs and concentrates technological capabilities in a few well-resourced industrial laboratories.

Ali Farhadi, CEO of **the Allen Institute for AI (Ai2)** , stated that the organization's goal is to build AI systems that advance science and expand humanity's exploratory capabilities. He believes robots can become a foundational scientific tool, enabling researchers to advance faster and ask new questions. To achieve this, AI systems that can be generalized in the real world are needed, along with tools that the global research community can collaboratively develop. Demonstrating the ability to transfer from simulated environments to reality is a crucial step in this direction.

The research team at Ai2 has proposed an economic model different from **MolmoBot** , a set of manipulating robots trained entirely on aggregated data. Instead of using humans to control the robots and collect data, the team created automated movement trajectories in a simulation system called **MolmoSpaces** .

The accompanying dataset, **MolmoBot-Data** , contains approximately 1.8 million expert-level manipulation trajectories. This dataset was created by combining the **MuJoCo** physics engine with 'domain randomization' techniques, which randomly change objects, camera angles, lighting, and dynamic elements to increase the diversity of the simulated environment.

Ranjay Krishna, PRIOR group leader at Ai2, says that most current methods try to bridge the gap between simulation and reality by adding data from the real world. However, his group is betting in the opposite direction: that gap can be narrowed by dramatically expanding the diversity of simulated environments, objects, and camera conditions. According to him, this advance shifts the focus of robotics from manual data collection to designing better virtual worlds, a problem that can be solved by technology.

To generate simulation data for physical AI, the team used 100 **Nvidia A100** GPUs . The system can generate approximately 1,024 experiments per GPU-hour, equivalent to over 130 hours of robot experience in just one hour of real-time.



Compared to real-world data collection, this method increases data throughput by nearly four times, thereby shortening development cycles and improving the return on investment for robotics projects.

The MolmoBot kit includes three different control policy types and was tested on two hardware platforms: **the Rainbow Robotics RB-Y1 mobile robot and the Franka FR3** tabletop robotic arm . The primary model uses the **Molmo2** vision-language platform , processing multiple RGB frames along with language instructions to determine the robot's actions.

For resource-constrained edge computing environments, the research team also offers **MolmoBot-SPOC** , a lightweight transformer model with fewer parameters. Additionally, there is **MolmoBot-Pi0** , which uses a **PaliGemma** architecture similar to Physical Intelligence's  $\pi_0$  model, allowing for direct performance comparisons.

In real-world testing, these models can transfer to real-world tasks without further refinement, even when working with objects or environments that were not present in the training data.

In the object picking and placing test, the main MolmoBot model achieved a success rate of 79.2%. This surpasses  **$\pi_0.5$**  , a model trained with a large amount of real-world data, which only achieved 39.2%. In mobile manipulation tasks, the robot was also able to successfully perform actions such as approaching, grasping doorknobs, and fully opening doors.

Offering a variety of architectures allows organizations to integrate powerful physical AI systems without being dependent on a single vendor or complex data collection infrastructure.

**The entire MolmoBot** ecosystem – including training data, data generation processes, and model architecture – is released as open source. This allows organizations to self-test, tune, and deploy physical AI systems at a controllable cost.

Ali Farhadi emphasized that for AI to truly drive science, progress cannot depend on closed data or isolated systems. Instead, a shared infrastructure is needed so that researchers worldwide can collaboratively build, experiment, and improve. According to him, this is the path for the continued development **of physical AI in the future.**

You finished reading the article "**MolmoBot: An AI robot that learns from simulated data instead of the real world.**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.