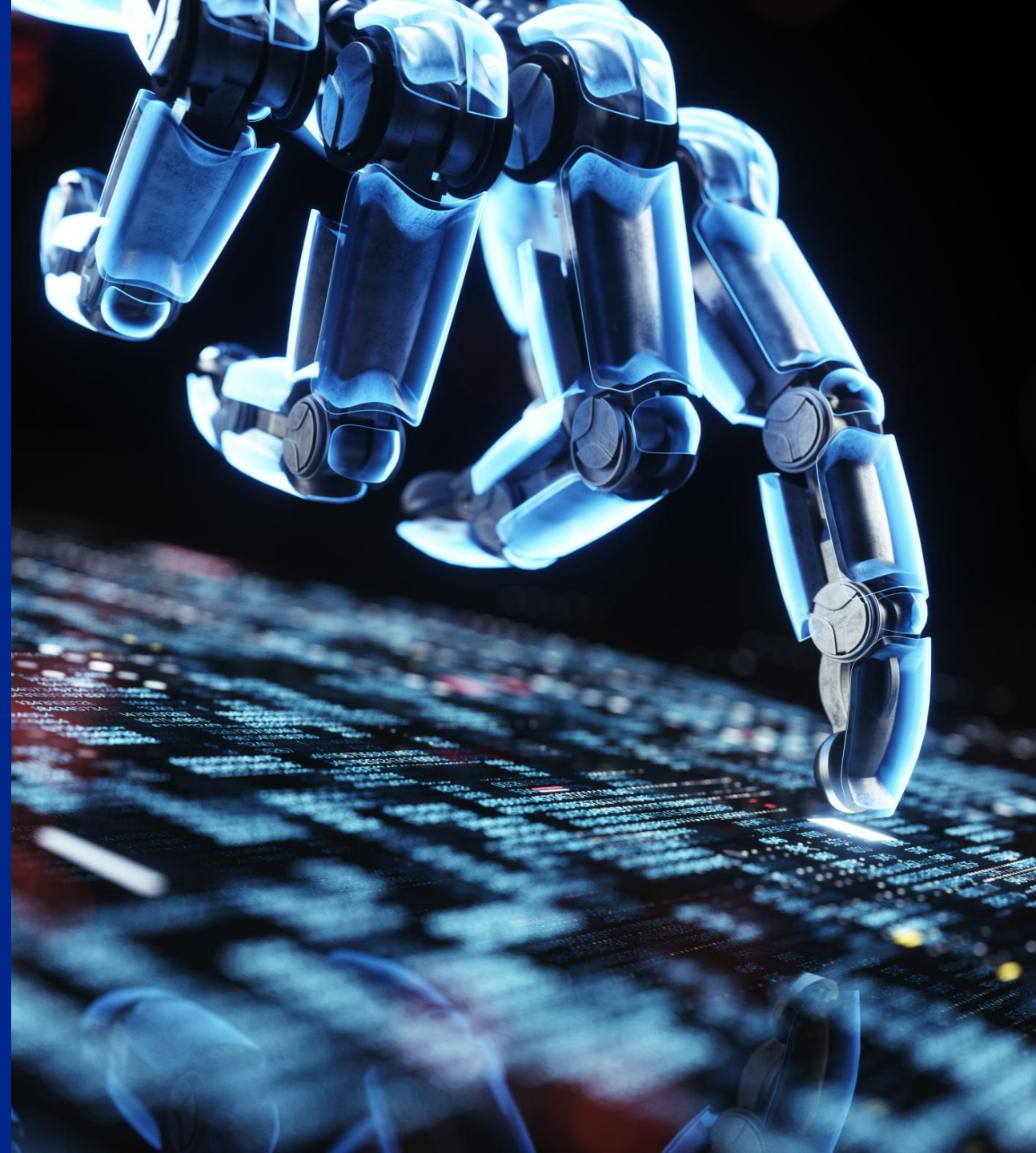


Synthetic Data Generation in Robot Learning

András Makány – R&D Associate

Antal Bejczy Center
for Intelligent Robotics

2026.03.24

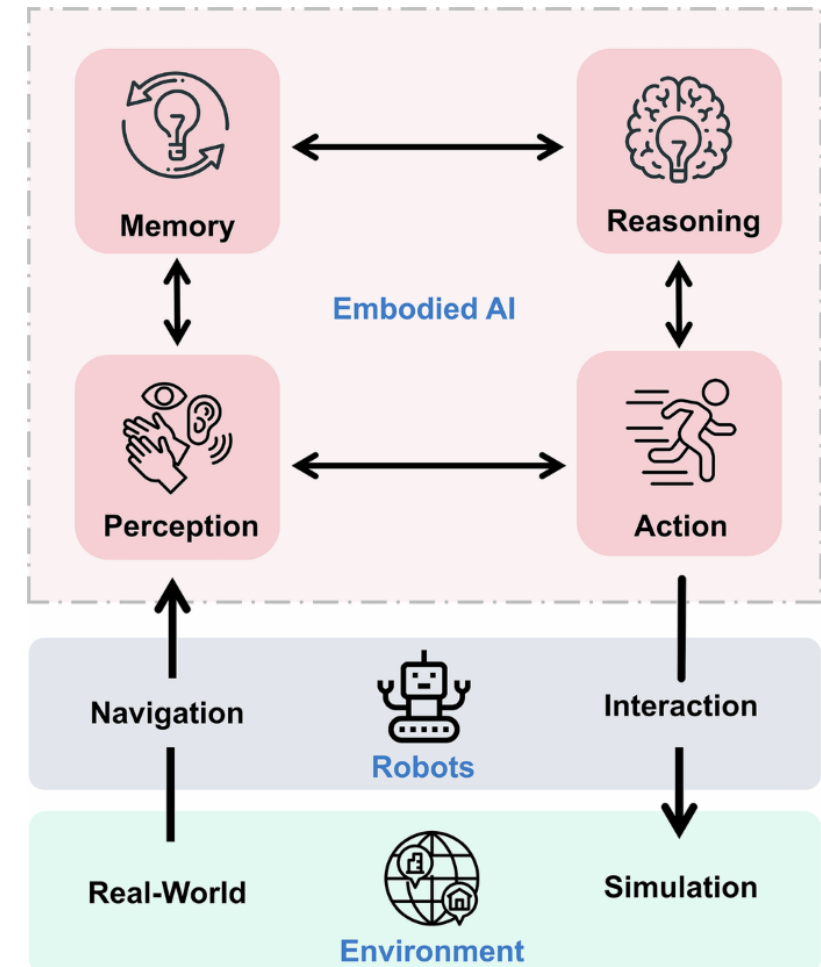


Current State of Embodied AI

Robotics control developed from simple program structures to sense-based control method, with **limited generalization**.

Embodied AI is the new wave of robot control, the integration of AI into robotic systems. These agents can **learn through** either **many tries**, or using **demonstrations** from experts.

Newest research focuses on **Vision-Language-Action** models. These multi-purpose agents can solve multiple tasks across many robot platforms.





The Data Collection

Method	Cost of Collection	Sim-to-Real Gap	Scalability & Automation
Real Data	High (Hardware & Time)	Physically accurate data	Low (Manual reset, limited scope)
Synthetic (Human)	Medium (Operator Time)	Present	Medium (Diverse environments)
Synthetic (Procedural)	Low (Initial definition)	Present	High (Automated randomization)

The **GUIDE** Pipeline



Generative

Automated generation of diverse, deterministic solutions and initial states.



Unified

Complete end-to-end integration mapping seamlessly with NVIDIA Isaac Sim and ROS 2.



Instruction-based

Built for minimal coding, enabling seamless robot task delegation through simple LLM prompts.



Demonstration

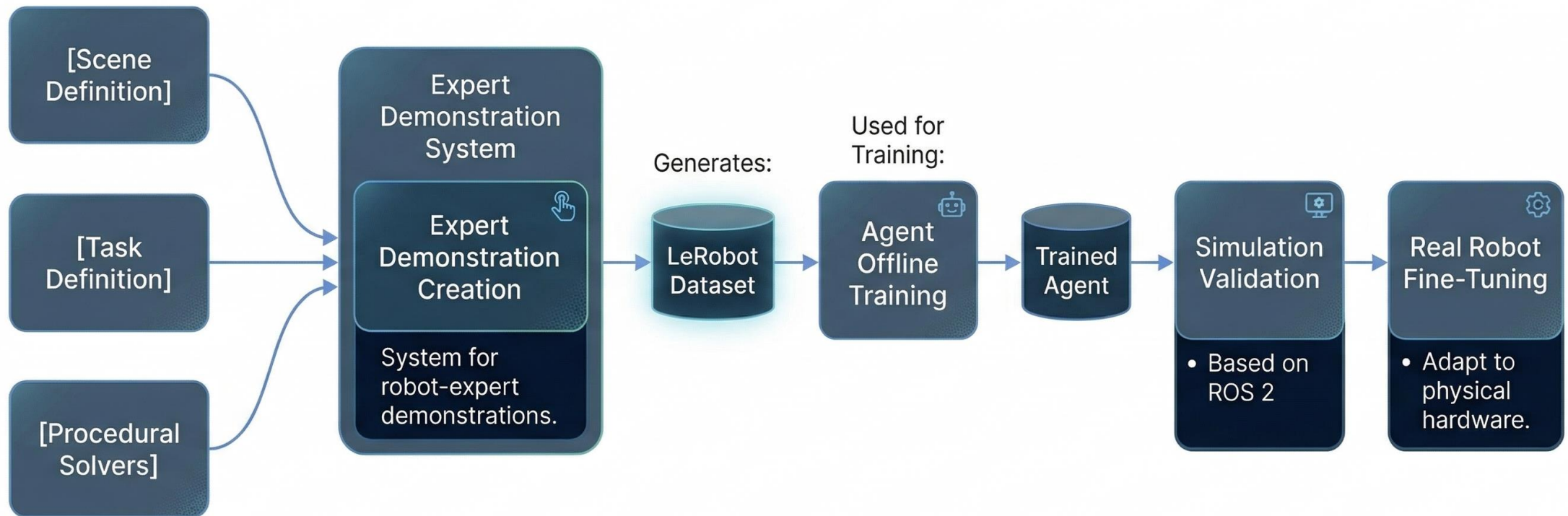
Extracts high-quality expert solver paths alongside multimodal images and robot states.



Environment

Built to deploy massive, domain-randomized scenes targeting precise edge cases.

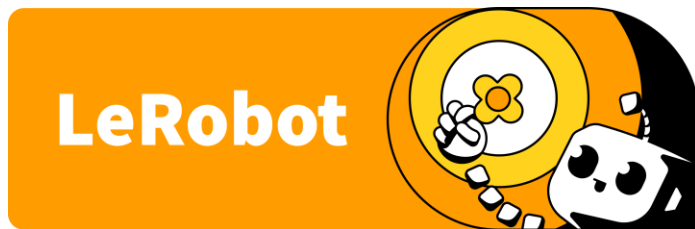
The **GUIDE** Pipeline



The **GUIDE** Pipeline

2 min

per demonstration



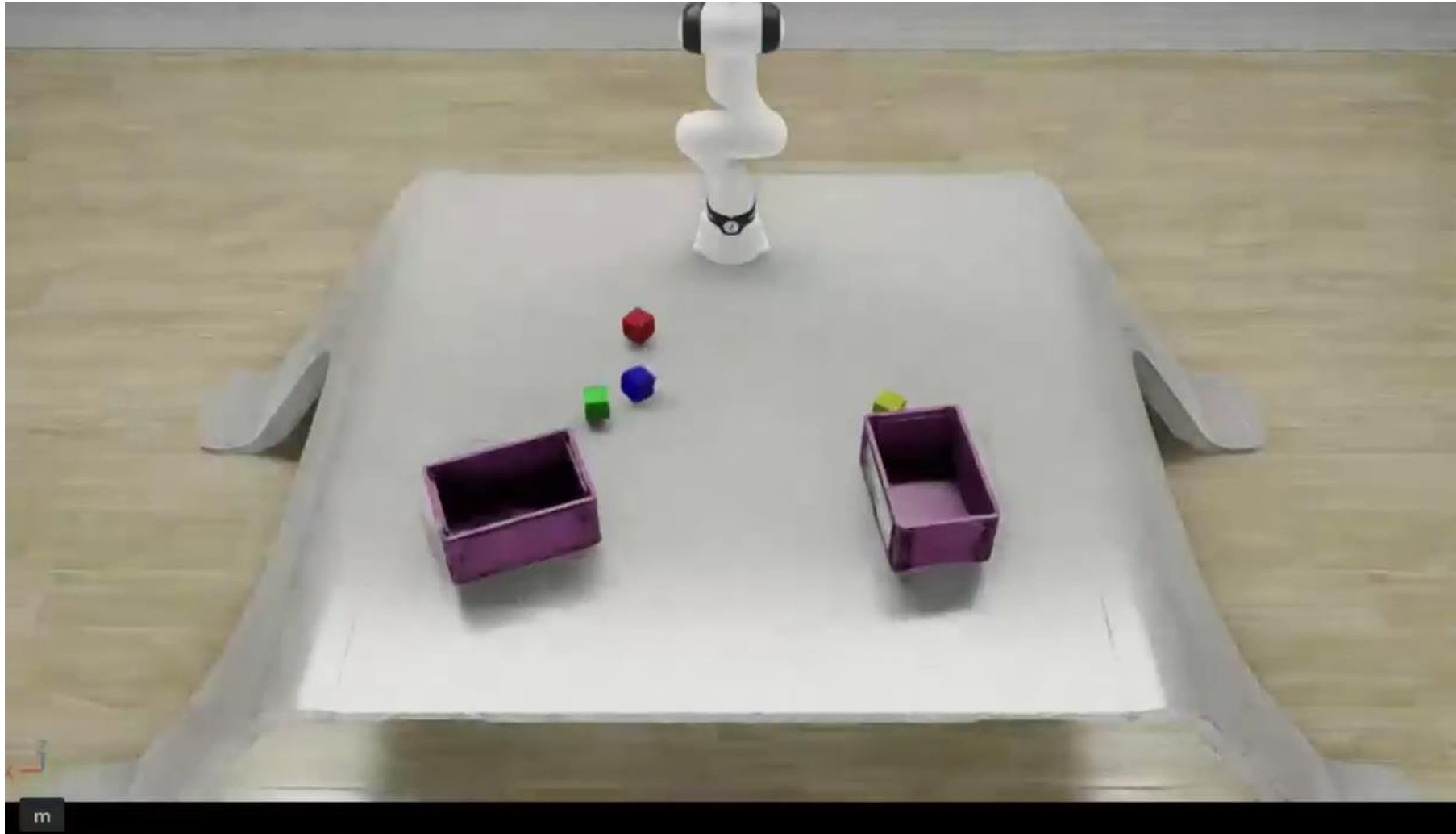
NVIDIA Isaac Sim + ROS 2

The pipeline solves a lab automation pick-and-place task and creates a dataset of **1000 demonstration** using a single simulated robot just under 34 hours.

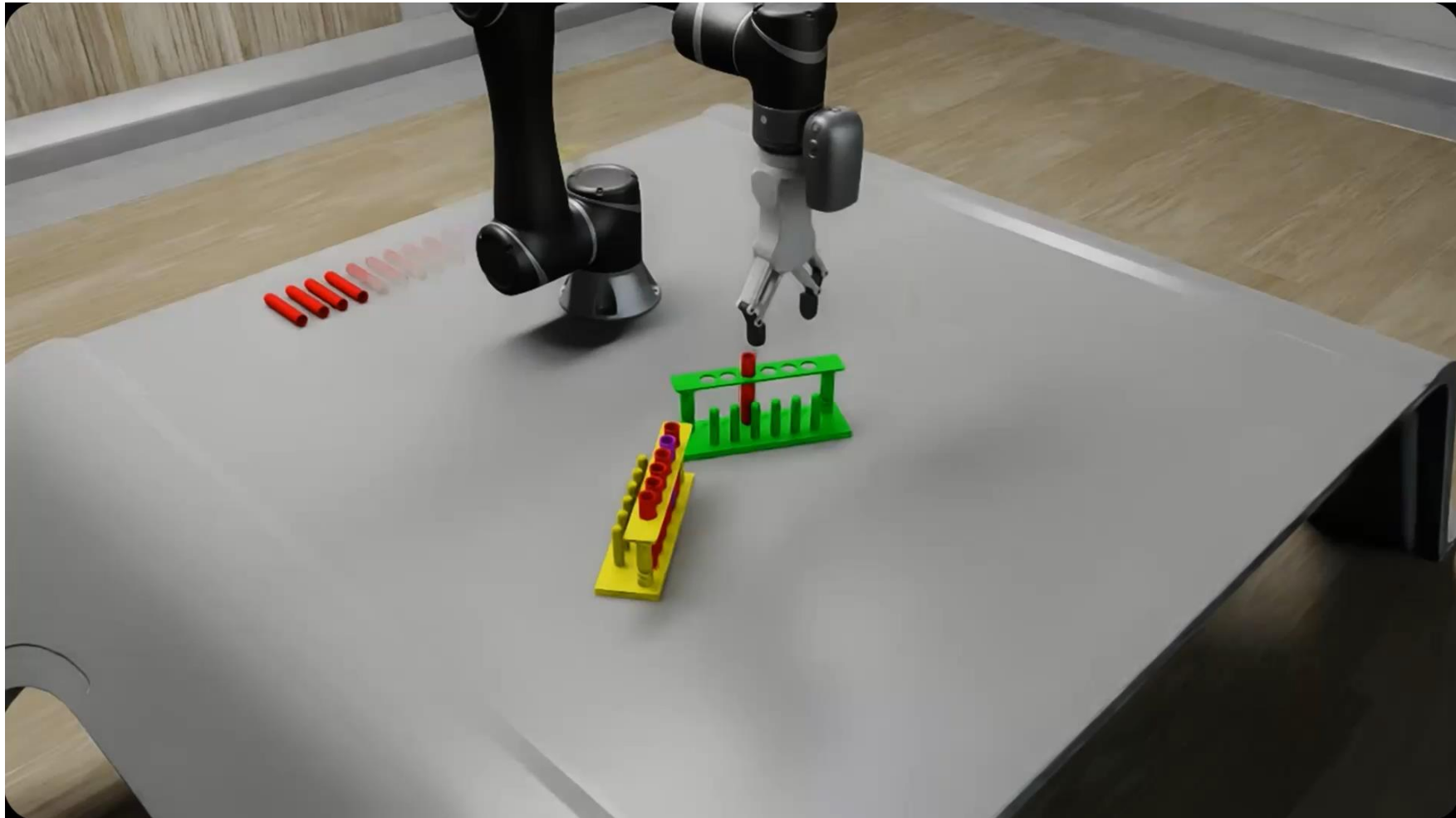
It seamlessly exports directly to learning-ready **LeRobot** dataset structure.

This massive data generation is perfectly tailored to feed modern sequence-models like **Action Chunking with Transformers (ACT)**.

Task Randomization






Automatic Dataset Generation



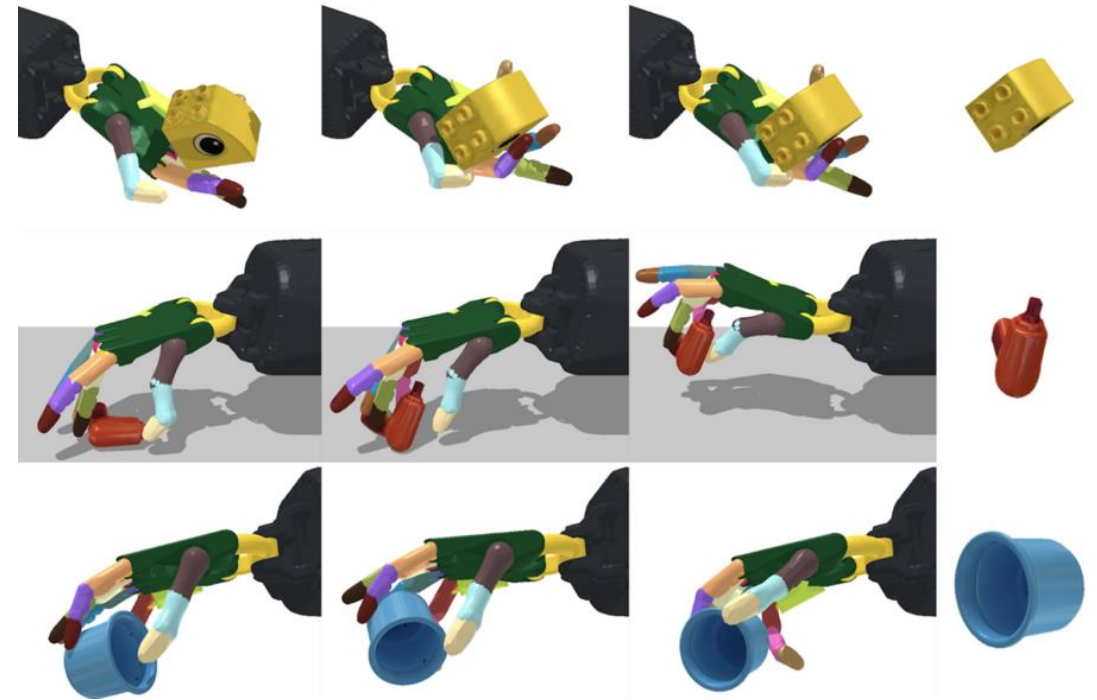
Limits: Distribution

Domain randomization can realistically provide immense **diversity** by altering task-relevant properties or camera feed variability:

-  Positions, sizes, and obstacle placements.
-  Mass and inertia (crucial for effort control).
-  Textures and lighting for vision models.

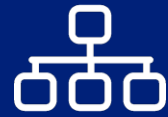
Possible tasks **grow exponentially**.

Coverage is bounded by the manual labor of texturing and the definition of parameters.



<https://www.eecs.mit.edu/dexterous-robotic-hands-manipulate-thousands-of-objects-with-ease/>

Limits: Sim-to-Real Gap



Sensing & Contact

The distance between simulated environments and physical reality is primarily driven by unsimulated nonlinear behaviors. Aligning sensing and contact dynamics remains a challenge.



Simulator Quality

While synthetic data acquisition is rapid and noiseless, the data quality is limited by the simulator's physical engine — particularly soft body and contact force simulations.

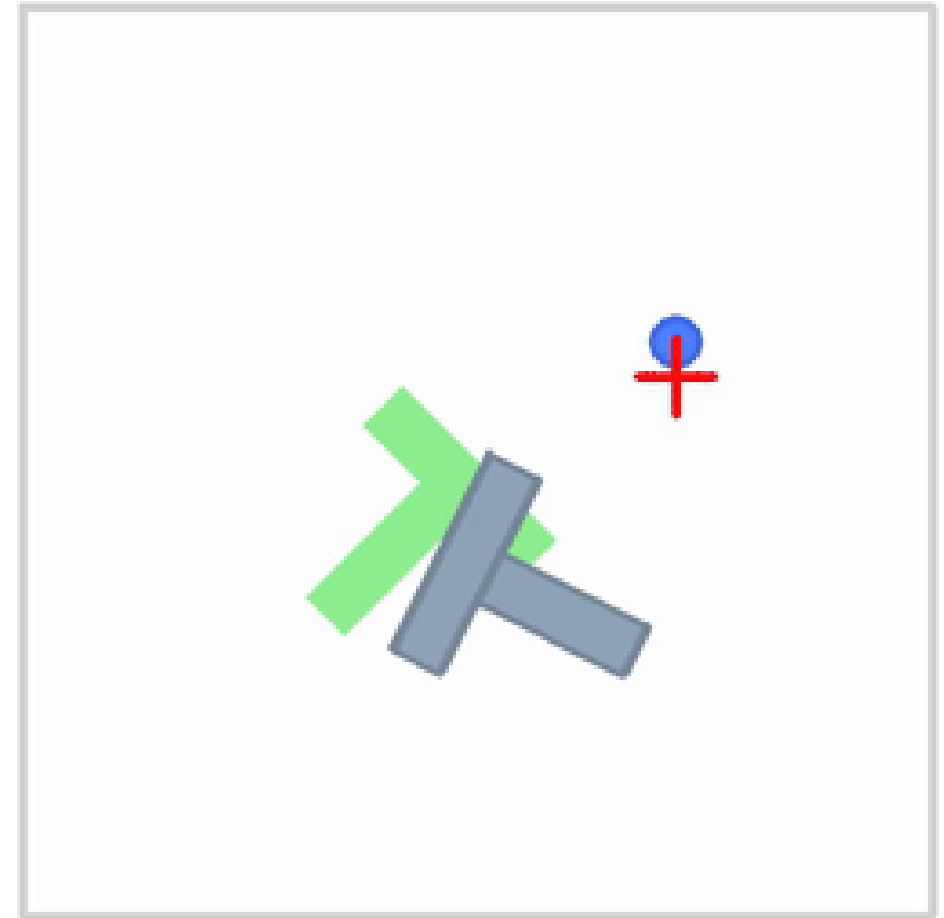
Authoring & Validation Burden

Using procedural solvers requires a **well-defined** solution pattern authored for every specific task.

While tasks like **bin picking** or **pick-and-place** are straightforward, scenarios requiring deep reasoning or complex relations (e.g. the PushT task) demand extensive specification effort.

Conclusion: Synthetic generation shines when the task **relies on precise environmental data** that can be substituted with simulation **ground-truth**.

Limits: Specification



<https://github.com/huggingface/gym-pusht>

Scalable Robot Roadmap



Curriculum Generation

Leveraging specialized LLM pipelines and WFMs to autonomously define and supervise comprehensive tasks directly from natural language prompts and example images.



Fine-Tuning

Combining large-scale domain randomization with learned residuals to aggressively narrow the lingering sim-to-real gap.



Evaluation

Implementing standardized assessment protocols for digital-twin-trained policies to guarantee unbiased, cross-embodiment scoring.

Questions?

Thank you for your
attention.

Project 2024-1.2.3-HU-RIZONT-00069 has been implemented with support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development, and Innovation Fund, financed under the 2024-1.2.3-HU-RIZONT funding scheme.