

WE DON'T MAKE ROBOTS. WE MAKE THEM INTELLIGENT.

The human-centric data backbone for Embodied AI and Humanoid Robotics.

Robots don't learn from code alone — they learn from experience. But raw data isn't enough: video shows what happened, not why it worked. Noitom Robotics is your dedicated Data Factory and World Compiler. Built on 13 years of motion capture expertise and spun off in 2025, we turn human physical experience into training-grade, multi-modal datasets — transferring human skill into robotic embodiments at a fraction of the cost of building it yourself.

Industrial Scale, Proven Economics



Established Global Expertise

Our technology is trusted by over 300 research institutions globally and proven in more than 500 commercial projects — today we work with close to 100 companies across the Physical AI field.



Industrial-Scale Production

Our Data Factory already produces high-precision embodied data on the order of 100,000+ hours per year — with structured acquisition pipelines engineered to scale by orders of magnitude as industry demand climbs past the million-hour mark.



Unmatched Facility Scale

Powered by our Data Factory: over 10,000 square meters of dedicated capture space engineered for mass data production — with facilities now online in Mountain View, Silicon Valley and Kuala Lumpur, Malaysia.



Validated, Not Just Collected

Every dataset is proven trainable through a closed validation loop — cross-embodiment retargeting, RL training in simulation, and sim-to-real deployment — with motion coverage benchmarked beyond leading datasets like AMASS and Motion-X++. You invest in results, not raw hours.

Commercial Value Summary

	Feature	Industry Standard	Noitom Robotics™
	Precision	Millimeter-level	Sub-millimeter (Optical + IMU)
	Compatibility	Platform-Locked	Cross-Ontology (Platform-Agnostic)
	Environment	Scripted / Lab-only	Unconstrained Stochastic "Wild" Data
	Signal Depth	Visual / Kinematic	Omni-modal (Tactile, Audio, Temp, RGB-D)
	Mapping	Manual / Rigid	Noitom Robotics Motion Retargeting (Any DoF)
	Output	Raw observations	Compiled, training-ready physical intelligence

Why Omni-Modal?



Finger-level truth

Millimeter-scale hand/finger kinematics plus pressure and contact signals — so models learn real grasp dynamics, including the signals that decide success or failure yet stay invisible in ordinary video.



Coverage across the full realism spectrum

Controlled “factory-grade” precision, large-space motion-with-vision, and truly natural “in-the-wild” behavior — so training data spans clean labels and messy real-world variance.



Built for scale, consistency, and deployment

Repeatable acquisition pipelines, standardized calibration/QA, and dataset structures designed for training workflows — validated end-to-end from capture through sim-to-real. Reliable data, not one-off demos.



Synchronized multi-modal ground truth

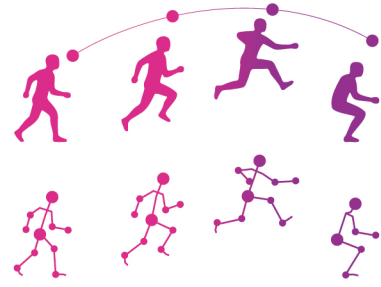
Motion + multi-view video (and additional signals where applicable) captured in tight time alignment, enabling strong visual grounding and cross-modal learning.

Inside the World Compiler

HiPHI-MOV: The Motion Prior

High Precision Human Interaction Motion with Object & Vision

HiPHI-MOV answers a fundamental question — how does the body move? It captures whole-body motion, interacted objects, and synchronized vision, engineered for robust locomotion and loco-manipulation policies.



- **Multimodal Deep Fusion:** Integrates 6-DOF full-body kinematics, optional hand-joint data, and synchronized RGB-D scenes to support integrated perception-control model training.
- **Co-registered Trajectories:** Ground-truth, full-body kinematic trajectories time-aligned with ego-centric and third-person visual observations in a single synchronized stream.
- **Benchmark-Proven Coverage (new — replaces “Scientific Interpretability”):** Motion primitives systematically span the human motion space, performed by hundreds of actors — covering motion regions that leading datasets such as AMASS, LAFAN1, and Motion-X++ miss.
- **Sim-to-Real Alignment:** Validated through a closed loop of cross-embodiment retargeting, RL training in simulation, and sim-to-real deployment — data that is proven trainable, not just collected.



HiPHI-OM: The Interaction Prior

High Precision Human Interaction Omni-Modality

HiPHI-OM captures what video can't see — contact, pressure, and precise hand-object geometry. Recorded in highly instrumented studios, it delivers ground-truth data for the signals that decide whether a task succeeds or fails.

- **Sub-millimeter Synchronization:** A robust pipeline fusing optical markers with inertial measurement units (IMUs) keeps tracking stable even through high-speed, high-acceleration motion.
- **Cross-Ontology Versatility:** Platform-agnostic by design — raw signals are decoupled from any single robot or semantic framework, maximizing cross-domain generalization and long-term value.
- **Omni-modal Capture:** Synchronized ego- and third-person RGB-D, hand-level tactile sensing, and high-precision spatial tracking, plus audio and temperature — with force data on the roadmap.
- **Hierarchical Structure:** Organized from micro-level movement primitives up to complete task sequences, supporting both low-level control and higher-level planning.

ITW: The Distribution Prior

In-The-Wild

ITW captures natural human behavior at life scale — with all the noise, clutter, and long-tail surprise no studio can script. It's the data that turns lab-optimized models into deployable ones.



- **Real-World Scene Generalization:** Captured across residential, hospitality, retail, and logistics settings (households, hotels, supermarkets), including changing lighting, dynamic obstacles, and unstructured layouts.
- **High-Efficiency Scaling:** Portable devices and large-scale parallel collection make acquisition dramatically more efficient than teleoperation, enabling rapid scaling of dataset volume.
- **Natural Behavior Capture:** Records unconstrained daily interactions and real work patterns rather than scripted lab actions — so robots learn motion that matches actual human operation.
- **Compiled Action Labels:** Models pretrained on HiPHI data reconstruct hand and wrist trajectories from ITW footage — turning raw real-world observation into labeled, training-ready data.
- **Compliance & Controllability:** A full toolchain for anonymization, compliance review, and post-processing, meeting R&D requirements for algorithm training.

Make the physical world learnable — with us.

Datasets, technical specifications, and the World Compiler research series at modalitynet.com



modalitynet.com
Explore the datasets