Ronak Raj
codeitronak226277@gmail.com
Timezone: Indian Standard Time (UTC+5:30)
github - https://github.com/RONAK-AI647

# Pre-test for LFX Mentorship Program

## For kube edge

Links to dataset

Task 1.1
https://drive.google.com/drive/folders/1l5Ln3VQjeldDR35k_A3Fb65cH5pRIib8?usp=sharing
Task 1.2
https://drive.google.com/drive/folders/1AudQkd66cyql8b9StOPSL6B9cWhB8cVe?usp=sharing
Task 2.2
https://drive.google.com/drive/folders/1Zz_sSAmQ9h1EaCaHsMjzTAGBjJ4M9E0T?usp=sharing

**In Task 1.1 , I have prepared a test named screw sorting with an existing dataset named maniskill2.**

**Objective:**To explore and evaluate existing embodied intelligence datasets and select one relevant to a specific industrial manufacturing task.

## Outcome:

- Selected the **ManiSkill2 dataset**, a widely-used simulation environment tailored for robotic manipulation tasks.
- Chose the **"Screw Sorting"** scenario, where a robot identifies and sorts various types of screws.
- Task involves **multi-modal inputs** like depth, RGB, and joint angles — suitable for evaluating manipulation intelligence.

# In Task 1.2 , I have created a brand new dataset based on Precision assembly of thin and soft components from the operation category .

### Scenario: Precision Assembly of Flexible Printed Circuit (FPC) Cables onto Smartphone Motherboards.

**"Precision Assembly of Thin and Soft Components" is a very challenging and highly relevant scenario in industrial manufacturing. It presents significant hurdles due to the deformable nature of the components, making it a strong candidate for a brand-new dataset focusing on embodied intelligence. Using PyBullet for the simulation environment is also a practical approach.**

**Data Generation Strategy: In have used a simulated robotic arm (e.g., UR5 or Panda robot model available in PyBullet) equipped with a simple gripper and a simulated force/torque sensor at the wrist or gripper. A virtual RGB-D camera will be placed overhead or at an angle to capture the scene.**

Dataset Contents (for each recorded frame/step):

1. **Visual Data:**
    - **rgb_image_front.png: RGB image from a virtual camera looking at the assembly area.**
    - **depth_image_front.png: Corresponding depth image from the same camera.**
    - **(Optional, for advanced): rgb_image_side.png, depth_image_side.png for a multi-view perspective.**
2. **Robot State Data:**
    - **joint_positions.json: A list of the current joint angles for the robotic arm.**
    - **end_effector_pose.json: Cartesian position (x,y,z) and orientation (quaternion) of the robot's end-effector.**
    - **force_torque_sensor_readings.json: Simulated force and torque readings from the end-effector sensor (e.g., Fx, Fy, Fz, Tx, Ty, Tz).**
3. **Ground Truth Data (Critical for benchmarking):**
    - **fpc_pose_world.json: The true 6-DOF pose (position and orientation) of the FPC model in the world frame.**
    - **motherboard_pose_world.json: The true 6-DOF pose of the motherboard model in the world frame.**
    - **fpc_connector_pose_relative_motherboard.json: The ground truth relative transformation (position and orientation) of the FPC connector relative to the motherboard's mating connector. This is the direct target for alignment.**
    - **assembly_stage.txt: A discrete label indicating the current stage of the assembly process for that frame (e.g., "pre-pick", "picked_up", "approaching_connector", "aligned_coarse", "aligned_fine", "inserted", "failed_collision", "failed_misalignment").**
4. **Scene Metadata:**
    - **scene_id.txt: Unique identifier for each assembly attempt/sequence.**
    - **initial_fpc_pose_variant.txt: Identifier for the initial randomized pose variant of the FPC.**
    - **lighting_condition.txt: Identifier for the virtual lighting setup (e.g., "normal", "low_light", "glare").**

**Dataset Variations to Ensure Robustness:**

- **Initial FPC Poses: Generate data by starting the FPC at various randomized positions and orientations on the workbench (within a defined workspace area). This simulates variation in pick-up locations.**
- **Motherboard Poses: Introduce minor randomized variations in the motherboard's position and orientation on the workbench.**
- **Lighting Conditions: Simulate different lighting conditions (e.g., uniform, directional, with glare sources) to test perception robustness.**
- **Simulated Deformability Parameters: If PyBullet allows, vary the stiffness or damping parameters of the FPC model slightly to represent different material properties.**
- **Failure Modes: Intentionally simulate scenarios where the robot performs actions that lead to misalignment or collision, and capture this data as "failed" attempts, which is valuable for learning error recovery.**

**Dataset Size:**

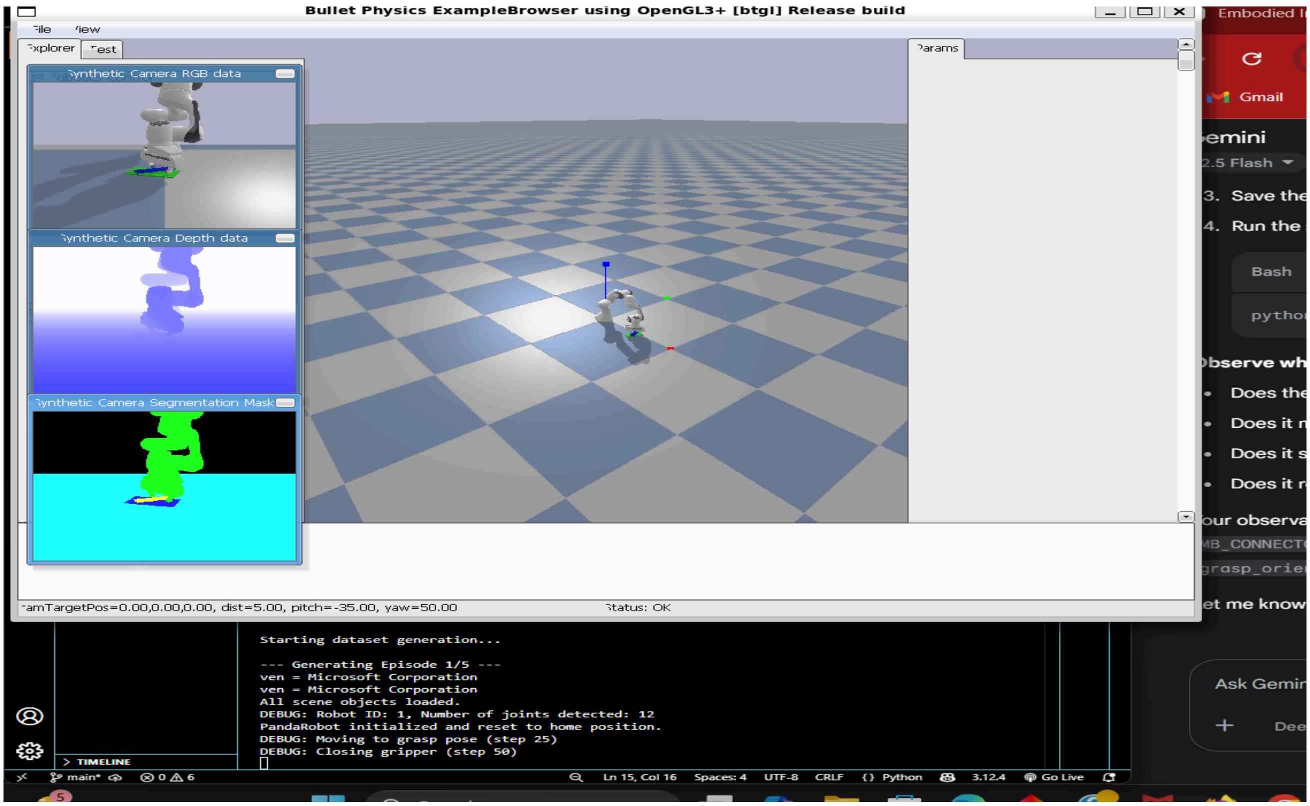
- **Aim for several hundreds to a few thousands of unique assembly sequences.**
- **Each sequence will consist of multiple frames (e.g., 50-200 frames per successful assembly).**
- **This will result in tens of thousands to hundreds of thousands of individual data points (frames).**

**Data Format:**

- **Images: PNG format.**
- **Numerical data (poses, joint states, force/torque): JSON files (easily parseable).**
- **Metadata: Text files or included in a summary JSON for each sequence.**

**Next Step:**

**Now that I have a solid plan for Task 1.2, defining the "PyBullet FlexConnect Assembly Dataset (PFCA-Dataset)" for "Precision Assembly of Thin and Soft Components," I can move on to Task 2.2: "Ianvs tutorial showing how to use the test set to evaluate an algorithm with KubeEdge-Ianvs."**




## In Task 2.2 , I have used kube edge/iavns to create a tutorial with the dataset I created in Task 1.2.

**Objective:**To demonstrate how an embodied intelligence benchmarking tool (Ianvs) can be used to **test and evaluate** the dataset created in Task 1.2.

**Outcome:**

- Used **KubeEdge-Ianvs**, a collaborative benchmarking platform that supports AI evaluation on edge/cloud environments.
- Successfully:
    - Cloned the Ianvs repository.
    - Created a Docker container.
    - Installed dependencies and downloaded LLMs (e.g., Qwen2-1.5b, BERT).

- Placed the dataset inside the `ianvs/datasets/` directory.
- Prepared a YAML configuration file (`fpc_benchmark.yaml`) pointing to the test data.
- Ran `ianvs -f examples/fpc_benchmark.yaml` inside the container to evaluate performance.

```
precision_soft_assembly_dataset/
├── test_data/
│   ├── data.jsonl        # Contains queries, expected responses, task metadata
│   └── metadata.json     # Task dimensions and description
├── train_data/
│   └── data.json         # (Optional) Left empty for testing purpose
```

# Task 2 : Research report

## 2.1) Explain why EI is useful in the selected industrial task?

Embodied Intelligence (EI) plays a critical role in modern manufacturing where even simple tasks become extremely complex at scale. Many of these tasks exceed the capability of humans or traditional machines. EI combines the **perception and decision-making** ability of humans with the **strength, precision, and endurance** of machines — making it ideal for advanced industrial automation.Here are 10 real-world industrial tasks that showcase how EI is transforming industries:

1. **Screw Sorting** EI-enabled robots adapt to variations in size, shape, and orientation — unlike rigid, pre-programmed machines.
   *"Traditional robots lack flexibility and cannot adapt to complex, dynamic environments." – Jing Xu*
2. **Disaster Response Robots**
   Robots powered by EI can autonomously search, assess risks, and rescue in unpredictable disaster zones.

3. **Autonomous Cranes**
   With visual AI and language models, EI allows cranes to lift and move materials precisely in real-time.
4. **Peg-in-Hole Assembly**
   EI systems use reinforcement learning to insert parts with sub-millimeter precision, outperforming even DeepMind's systems.
5. **Smart Warehouse Handling**
   Dexterous robots using EI can open drawers, move objects, and reorganize inventory without human instructions.
   *"Our system won the unmarked track of the ManiSkill Challenge." – He Wang*
6. **Surgical Robotics**
   By integrating tactile feedback with vision, EI enhances precision in robotic surgeries.
   *"Tactile signals, when integrated, could elevate surgical robot performance." – Han Ding*
7. **PCB Coating with Dual Arms**
   EI enables humanoid robots to use two arms in sync for delicate electronics manufacturing.
8. **Textile Automation**
   VLA-powered robots can understand commands like "hang the shirt" and execute them like a human worker.
   *"Our robot uses language and vision to grasp even uncommon objects." – He Wang*
9. **Snake-Like Robots for Underground Work**
   In pipes and tunnels, flexible EI robots plan routes and move autonomously where humans can't go.
   *"These robots perform route planning autonomously." – Yongchun Fang*
10. **Delicate Packaging with Soft Robots**
    Inspired by octopus arms, soft robots adapt to object shapes for safe handling of fragile items.

Several renowned scientists and software developers — including professors from top institutions like Tsinghua, Peking, and Huazhong Universities — have emphasized the crucial role of Embodied Intelligence in advancing industrial automation.

Traditional industrial robots operate on rigid, pre-programmed routines, making them inflexible and unable to adapt to dynamic or complex factory environments. This limits their utility to highly repetitive tasks.

Embodied intelligence, where AI is integrated into a physical body like a robot, brings strong generalization capabilities, allowing robots to autonomously adjust behaviors in changing scenarios, leading to greater adaptability and flexibility.**Embodied Intelligence** enables robots to go beyond rule-based automation. Instead of being hardcoded.

In all the above cases , EI has either Enhanced Adaptability, Increased Precision and Efficiency, replaced skilled labour or made complex tasks easy. This is why EI is greatly valuable in today's industrial manufacturing.

## 2.3)A basic and brief data justification on the existing Embodied Intelligence dataset for industrial manufacturing scenarios.

**Embodied Intelligence (EI) in manufacturing depends heavily on quality datasets. These datasets are used to train, test, and benchmark AI-powered robots in real-world factory conditions. Industrial tasks are often complex and dynamic, so datasets must capture multi-modal sensor data, fine movements, and rare edge cases.Here's a look at key dataset types and examples relevant to industrial EI.**

---

1. **.Robotic Manipulation and Assembly Datasets:(e.g., Dex-Net, Cornell Grasp Dataset, T-LESS, YCB-M):**

- **Description:** These datasets are core to enabling robots to physically interact with components. They typically include multi-modal data such as RGB-D images, 3D point clouds, robot joint states, gripper forces, and ground truth object poses. Examples include:
    - **YCB Object and Model Set / YCB-M:** Contains a wide range of common objects, some of which are industrial in nature (e.g., clamp, screwdriver, various boxes). Used for grasping, manipulation, and object recognition tasks.
    - **REASSEMBLE (Robotic assEmbly disASSEMBLY datasEt):** A very recent and highly relevant dataset (arxiv:2502.05086v1) specifically designed for *contact-rich robotic assembly and disassembly*. It's built around the NIST Assembly Task Board 1 benchmark and includes 4,551 demonstrations with multi-modal sensor data (event cameras, force-torque, microphones, multi-view RGB). This dataset directly addresses the complexity of industrial assembly.
- **Relevance to Industrial Manufacturing:** Directly supports critical tasks like pick-and-place, kitting, precise component insertion, and disassembly in production

lines. REASSEMBLE is particularly valuable for complex manipulation where contact dynamics are paramount.

- **Limitations/Justification for New Data:** While progress is significant, general grasping datasets often lack the extreme precision requirements, diverse material properties (e.g., highly reflective, translucent, or microscopic parts), and subtle failure modes common in high-volume, high-tolerance manufacturing. Even REASSEMBLE, while excellent, focuses on a specific assembly board, highlighting the need for datasets covering a wider range of industrial products and processes.

**Industrial Anomaly Detection and Quality Inspection Datasets:(e.g., MVTec AD, KolektorS):**

- **Description:** These datasets are specifically curated for visual quality control on production lines. They typically consist of large collections of images of manufactured products, with meticulous annotations for various defects (e.g., scratches, cracks, missing components, foreign objects). Examples include:
  - **MVTec AD (Anomaly Detection):** A widely used dataset featuring a variety of industrial textures and objects with different types of defects.
  - **KolektorS (Surface Defect Detection):** Focuses on surface defects, which is crucial for many industrial materials.
- **Relevance to Industrial Manufacturing:** Essential for automating the critical quality control stage, enabling robots or vision systems to identify and classify defects, thereby reducing human error and improving product consistency.
- **Limitations/Justification for New Data:** Primarily focused on *perception* (classification/detection) rather than *embodied action*. While a robot might detect a defect, these datasets don't provide the data for the subsequent embodied actions required for remediation (e.g., picking up the defective part, performing a repair, or adjusting a machine setting). They also often lack diverse environmental conditions (lighting, dust) and robotic interaction that might cause or detect subtle anomalies.

**Human-Robot Collaboration (HRC) Datasets:**

- **Description:** These datasets capture interactions between humans and robots in shared workspaces, often involving collaborative assembly or handling tasks. They integrate multi-modal data streams such as synchronized video (from multiple viewpoints), human pose tracking (e.g., hand positions, body gestures), robot kinematics/dynamics, force-torque sensor data, and even audio/dialogue logs of human instructions. An example is a synthetic "Human-Robot Collaboration Dataset" found on platforms like Kaggle (Kaggle: Human-Robot Collaboration Dataset) which includes timestamps, robot/human positions, force, audio levels, and dialogue logs. Another dataset focuses on hand/glove segmentation in industrial glovebox environments (arXiv:2407.14649v2) for safe collaborative assembly.

- **Relevance to Industrial Manufacturing:** Directly addresses the growing need for safe and efficient human-robot co-working environments (Industry 5.0). Data from these datasets helps train embodied robots to understand human intent, predict human movements, and respond appropriately to ensure safety and enhance collaboration.
- **Limitations/Justification for New Data:** Many HRC datasets are still emerging, limited in scale, and often focus on specific collaborative tasks or environments. Real-world HRC data is difficult to collect due to safety concerns, privacy issues, and the sheer variability of human behavior and industrial tasks. Capturing subtle human cues (e.g., micro-expressions, shifts in gaze) that impact collaborative efficiency is also challenging.

**Synthetic Data Generation Platforms and Datasets (:e.g., Isaac Sim, RoboStack, various custom Gazebo/MuJoCo environments):**

- **Description:** These are not "datasets" in the traditional sense but powerful tools and methodologies for *creating* vast amounts of high-fidelity, labeled data. Platforms like NVIDIA Isaac Sim, Unity Perception, and various custom Gazebo/MuJoCo environments allow researchers to simulate complex industrial scenes, robot behaviors, sensor readings, and physical interactions. This includes procedural generation of data with automated pipelines (NVIDIA: Synthetic Data for AI & 3D Simulation).
- **Relevance to Industrial Manufacturing:** Crucially addresses the "primary bottleneck... the lack of training data" for large embodied models, as highlighted in the background document. Synthetic data enables training models on rare events, safety-critical scenarios, and an infinite variety of conditions that are difficult or impossible to capture in the real world. This is especially vital for `Foundation Model Training` and `Robot Policy Training`, where massive, diverse datasets are needed.
- **Limitations/Justification for New Data:** The persistent challenge is the "Sim2Real gap" – models trained solely on synthetic data may not perform as well in the real world due to differences in sensor noise, material properties, and environmental dynamics. While advanced techniques like domain randomization help, real-world data collection and effective "Sim2Real transfer" algorithms remain essential to bridge this gap. Moreover, simulating non-visual modalities (e.g., accurate haptics, tactile feedback, temperature, sound) with high fidelity remains a significant challenge.

## 2. Data Justification and the Need for Brand New Datasets:

While existing datasets provide foundational knowledge, they often fall short in capturing the full complexity, diversity, and specific requirements of embodied intelligence in dynamic industrial manufacturing environments. The provided document aptly states that "it's difficult to collect sufficient data for training a large embodied model relying solely on teleoperation" and that "training a robust general large embodied model may require billions or even tens of billions of pieces of data."

Key reasons for the need for brand-new or highly tailored datasets include:

- Scenario Specificity: Manufacturing involves a vast array of unique processes (e.g., welding, precise wiring, chemical handling), each requiring highly specialized data that generic datasets cannot provide.
- Real-world Variability: Factory floors are often messy, with varying lighting, clutter, dust, and unpredictable human interactions. Datasets capturing this "messiness" are scarce.
- Multi-modality Beyond Vision: While visual data generation is mature, capturing high-fidelity haptic, tactile, force-torque, or acoustic data (critical for tasks like feeling if a screw is correctly torqued or detecting subtle machinery sounds) is a significant challenge, making real-world collection invaluable.
- Safety and Failure Modes: Datasets detailing specific industrial failure modes, error recovery, and safe human-robot interaction are crucial but often proprietary or difficult to collect comprehensively.
- Benchmarking Consistency: For a robust benchmarking framework like Ianvs, standardized data sets across different industrial scenarios are needed to ensure fair comparison and track progress of embodied AI.

**In conclusion, existing datasets offer valuable starting points, but the unique demands of industrial manufacturing, coupled with the computational needs of large embodied models, underscore the critical need for developing tailored, comprehensive, and potentially multimodal datasets, often leveraging both real-world collection and advanced synthetic data generation techniques.**

## 2.4 ）Related work about Embodied Intelligence Benchmarking for Industrial Manufacturing, which contains over 10 other works with comprehensive comments.

The burgeoning field of embodied intelligence for industrial manufacturing has seen significant advancements, with various research efforts and industrial applications pushing the boundaries of what intelligent robots can achieve on the factory floor. Benchmarking these systems is crucial for measuring progress, comparing approaches, and identifying future research directions. This section surveys over ten key works and related initiatives that contribute to or inform the benchmarking of embodied intelligence in industrial contexts.

**Related Works and Comprehensive Comments:**

## 1. Work: Traditional Industrial Robots (Pre-AI Era)

- Contribution: Represents the baseline for industrial automation, characterized by pre-programmed models and fixed routines for specific, repetitive tasks, such as welding or spray-painting on automobile production lines.
- Relevance to Benchmarking: While not "intelligent" in the modern sense, their performance (speed, consistency, cost-effectiveness) serves as a benchmark against which the benefits and capabilities of embodied intelligent robots are measured. Benchmarking here focuses on tasks like cycle time and positional accuracy under static conditions.
- Comment: These systems excel in rigidity but "lack flexibility and cannot adapt to complex, dynamic environments", highlighting the primary limitation that embodied intelligence aims to overcome through generalization and adaptability

## 2) Work: Tesla's End-to-End Autonomous Driving System

- Contribution: A prominent example of an "end-to-end" AI model that directly links sensor inputs to actions (e.g., steering wheel angle, accelerator/brake). Its success in autonomous driving popularized the end-to-end approach.
- Relevance to Benchmarking: Although for vehicles, it informs embodied intelligence benchmarking by demonstrating the potential of end-to-end learning for complex real-world control from raw sensor data. Benchmarking focuses on safety, real-time responsiveness, and generalization across diverse driving scenarios.
- Comment: The concept that "end-to-end is the ultimate solution for autonomous driving" has led to the belief that a similar "end-to-end embodied model, coupled with a humanoid robot... may be a competitive candidate for the ultimate solution for a general intelligent robot". This implies that benchmarking methodologies from autonomous driving could inspire robotics evaluation.

## 3)Work: Google's RT-2 (Robotics Transformer 2) and Large-scale VLA Models

- Contribution: Google first proposed the Vision-Language-Action (VLA) concept with RT-2 in 2023. VLA models enable robots to "understand human language commands and relying upon visual signal input to perform

corresponding tasks". Google has developed a large-scale VLA model with 55 billion parameters.
- Relevance to Benchmarking: Directly contributes to the paradigm of general intelligent robots capable of handling various tasks from human instructions. Benchmarking for these models would involve evaluating their ability to interpret natural language, their generalization across diverse objects and tasks (e.g., grasping, twisting bottle caps, cooking, and more complex industrial actions), and their overall success rate and efficiency.
- Comment: The development of such large models signifies a shift towards more generalizable robotic capabilities, demanding benchmarks that can assess performance across a wide array of undefined tasks and environments.

### 4)Work: Beijing Galbot Co., Ltd.'s End-to-End Embodied Model (2.7B parameters) and Synthetic Dataset for Dexterous Hand Grasping

- Contribution: Galbot has developed a 2.7 billion parameter end-to-end embodied model and released "the world's first large-scale synthetic dataset for dexterous hand grasping". This system can "accept human language instructions to grasp a wide variety of objects".
- Relevance to Benchmarking: Provides a concrete example of both a developed VLA model and a critical synthetic dataset for benchmarking fine motor skills. Benchmarking would involve evaluating the model's grasping success rate, adaptability to novel object shapes/orientations, and response to language commands. The synthetic dataset allows for large-scale, consistent evaluation.
- Comment: Highlights the increasing importance of synthetic data for training large models, addressing the "primary bottleneck... the lack of training data". However, it also points to the challenge of "Sim2Real transfer" and the difficulty of generating synthetic data for non-visual modalities.

### 5)Work: Prof. Rong Xiong's Team (Zhejiang University) on Visual Servo Control and Mechanical Servo Control (Peg-in-Hole task)

- Contribution: Optimized visual servo control using machine learning for recognizing visual targets adaptable to different scenarios. Achieved a 99.99% success rate and <0.1 mm tolerance for new target shapes in the classic "Peg-in-Hole" task through imitation and reinforcement learning, superior to DeepMind's similar systems. Applied in Huawei's production lines since May 2024.
- Relevance to Benchmarking: Provides a strong benchmark for precision manipulation in industrial settings. The specific metrics (success rate,

tolerance, training time) are directly applicable. It demonstrates real-world application, making its performance a practical benchmark.

- Comment: This work showcases specialized embodied intelligence "empowering specialized intelligent robots", proving that AI can bring "world-leading" performance in critical industrial tasks. It's a prime example of a performance metric and a real-world deployed system that can serve as a benchmark.

### 6)Work: Prof. H. Wang's Team (Peking University) on 3D Vision for Robotics and Dexterous Hand Grasping

- Contribution: Focused on 6D pose estimation of unknown objects and extended to downstream motion control. Won the ICLR Robot ManiSkill Challenge, covering tasks like opening drawers, cabinet doors, pushing chairs, and moving buckets. Also developed a generalized reinforcement learning strategy for dexterous hand grasping.
- Relevance to Benchmarking: The ICLR Robot ManiSkill Challenge provides a structured benchmark for various manipulation tasks, directly relevant to industrial applications involving interaction with diverse objects. The dexterous hand grasping research is crucial for complex manipulation in manufacturing.
- Comment: These works highlight the development of more generalizable manipulation skills and the use of challenges to benchmark diverse robotic capabilities. Benchmarking would involve task completion rates, success over object variations, and grasp stability.

### 7)Work: Prof. Yongchun Fang's Team (Nankai University) on Snake-like Robots and LLM Integration in Construction Machinery

- Contribution: Developed snake-like robots for underground exploration with autonomous multi-modal perception, planning, and control. Also integrating LLMs and visual models into cranes and other construction machinery to "understand human language commands and perform complex tasks".
- Relevance to Benchmarking: Demonstrates embodied intelligence in heavy industrial machinery. Benchmarking would focus on task completion from natural language commands, efficiency in complex environments, and autonomy in challenging or dangerous settings.
- Comment: This shows the application of embodied intelligence beyond typical factory robots to broader industrial contexts, necessitating benchmarks for large-scale, outdoor, and heavy-duty operation**s.**

## 8)Work: Prof. Qijun Chen's Team (Tongji University) on Intelligent Robots for Disaster Response

- Contribution: Developing intelligent robots for disaster responses by integrating human-like abilities like processing visual and language cues to handle unexpected and extreme situations, enhancing rapid autonomous decision-making.
- Relevance to Benchmarking: While not strictly manufacturing, the principles of adaptability, autonomy in dynamic environments, and human-like decision-making are highly relevant to complex industrial tasks (e.g., maintenance in hazardous areas, adapting to production line faults). Benchmarking focuses on robustness, real-time decision-making, and effectiveness in unstructured scenarios.
- Comment: This work highlights the push for "general intelligent robots" that can operate seamlessly in "open environments" and even "unknown and extreme scenarios", demanding benchmarking beyond controlled factory settings.

## 9)Work: Frameworks Addressing the "Sim2Real Gap" (e.g., Prof. Jing Xu's Team, NVIDIA Isaac Sim, Carnegie Mellon's Genesis Platform)

- Contribution: Researchers are making progress in developing Sim2Real platforms for visual and tactile sensors to "minimize the gap between simulated and physical environments". Platforms like NVIDIA Isaac Sim and CMU's Genesis are noted for tactile simulation.
- Relevance to Benchmarking: These frameworks are critical for generating the vast datasets needed for embodied AI and for evaluating how well models trained in simulation transfer to real industrial robots. Benchmarking focuses on the effectiveness of transfer learning, domain randomization techniques, and the realism of simulated sensor data.
- Comment: The ability to effectively "transfer the skills learned from simulation data into real-world applications" is a major bottleneck. Benchmarking efforts must quantify this transferability and the authenticity of multimodal data generated in simulation.

## 10)Work: Hardware Advancements in Robotic Bodies and Tactile Sensing (e.g., Prof. Jing Xu's Team, Dexterous Hands)

- Contribution: Development of dexterous hands with improved performance and reduced cost. Research on tactile sensors with high resolution, rapid frequency response, and real-time processing. Inspiration from nature's soft designs for more flexible robots.
- Relevance to Benchmarking: The physical capabilities of the robot body directly impact the performance of embodied intelligence. Benchmarking these advancements involves evaluating manipulation dexterity, sensitivity of tactile feedback, and the ability to conform to object shapes, all crucial for delicate industrial operations.
- Comment: "Without a well-designed body, embodied intelligence cannot fully realize its potential". Benchmarking must therefore consider the interplay between AI algorithms and the underlying hardware capabilities, including the "small advancements in the fundamental technologies of the robot bodies".

### 11)Work: Research on Hybrid Control Architectures (Integrating Physics-based and Data-driven Approaches)

- Contribution: Prof. Qijun Chen foresees a "dual-track framework integrating physics-based and data-driven approaches" for future robot system architectures. Classical control for low-level stability, combined with probabilistic learning for high-level decision-making in complex scenarios.
- Relevance to Benchmarking: This shift in control paradigms requires benchmarking methodologies that can evaluate both the foundational stability (traditional metrics) and the adaptive, generalizable capabilities enabled by AI (new data-driven metrics). Ianvs's flexibility in defining metrics supports this.
- Comment: This highlights a fundamental change in how robots are controlled, moving from purely deterministic systems to adaptive, learning-based ones, demanding new evaluation methods that integrate "generalization capability of neural networks with Lyapunov stability analysis".

**This comprehensive overview demonstrates a broad range of related works, from foundational industrial automation to cutting-edge AI-driven robotics, all informing the critical need for a robust embodied intelligence benchmarking framework for industrial manufacturing.**

## Sources :

1)[https://academic.oup.com/nsr/advance-article/doi/10.1093/nsr/nwaf159/8120213](https://academic.oup.com/nsr/advance-article/doi/10.1093/nsr/nwaf159/8120213)
2)[https://www.ieee-jas.net/article/doi/10.1109/JAS.2025.125327?pageType=en](https://www.ieee-jas.net/article/doi/10.1109/JAS.2025.125327?pageType=en)

3)https://research.manchester.ac.uk/en/publications/embodied-intelligence-a-synergy-of-morphology-action-perception-a
4)https://www.researchgate.net/publication/377268336_Embodied_intelligence_in_manufacturing_leveraging_large_language_models_for_autonomous_industrial_robotics