

ロボット基盤モデルへの取組み ー実データと波及ー

2026.3.23

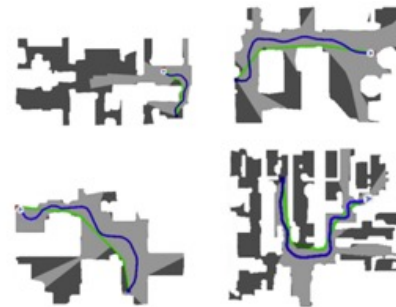
産業技術総合研究所 人工知能研究センター

実体知能研究チーム チーム長

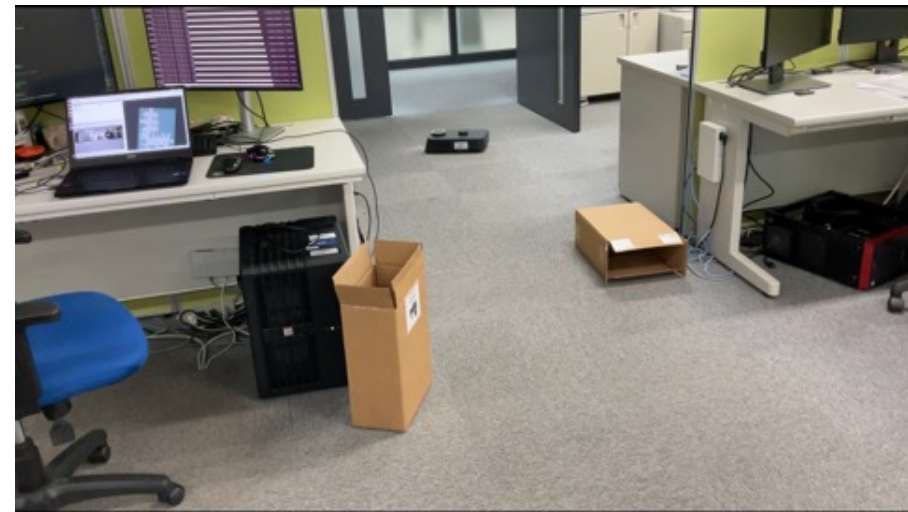
堂前 幸康

既存基盤モデルを実際に応用してみると：自律移動の場合

- Navigation基盤モデル NoMaDを屋内環境の3Dシミュレータでファインチューニング
- ダイナミックなシーン変化にもうまく対応

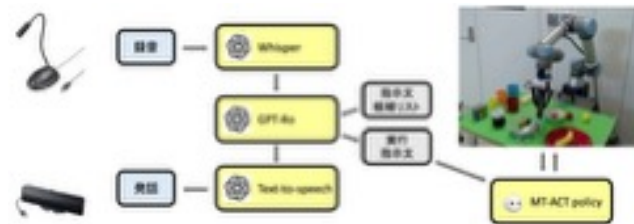


Yusuke
Yoshiyasu

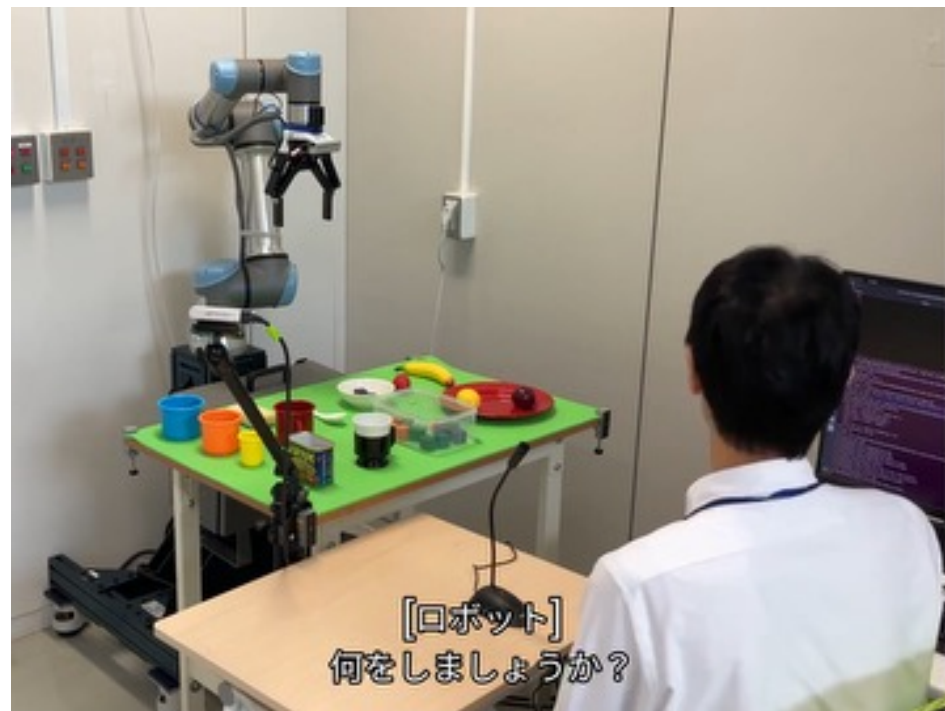


既存基盤モデルを実際に応用してみると：マニピュレーションの場合

- ロボット基盤モデル MT-ACT を1100エピソード（約六時間分）でファインチューニング
- 言語指示で動くシステム構築
- 視覚・言語・アクション，MMが滑らかに接続



Masaki Murooka



- ・ パフォーマンスの向上

様々な要因があるが、FMの本質は**データの質・量**。

→ 視覚、言語、音声などのモダリティと比較して、
実際に集める必要のあるアクションデータは未だ多くない。
また**オープンなデータ**も十分でない。

- ・ 普及に向けて

新しいモデルが多く、キャッチアップも容易でない。
開発者がトライしてみる環境が整っていない。

ロボットの“10万年データギャップ”問題

ScienceRobotics

Current Issue Archive About ▾ More ▾

HOME > SCIENCE ROBOTICS > VOL. 10, NO. 105 > GOOD OLD-FASHIONED ENGINEERING CAN CLOSE THE 100,000-YEAR “DATA...”

EDITORIAL | ARTIFICIAL INTELLIGENCE

f X 蝶 in 豆 聊 信

Good old-fashioned engineering can close the 100,000-year “data gap” in robotics



Ken Goldberg



インターネット規模の言語・画像データに比べ、**ロボットの実世界データは約10万年分不足**しており、この「データギャップ」が汎用ロボット開発を阻んでいる。

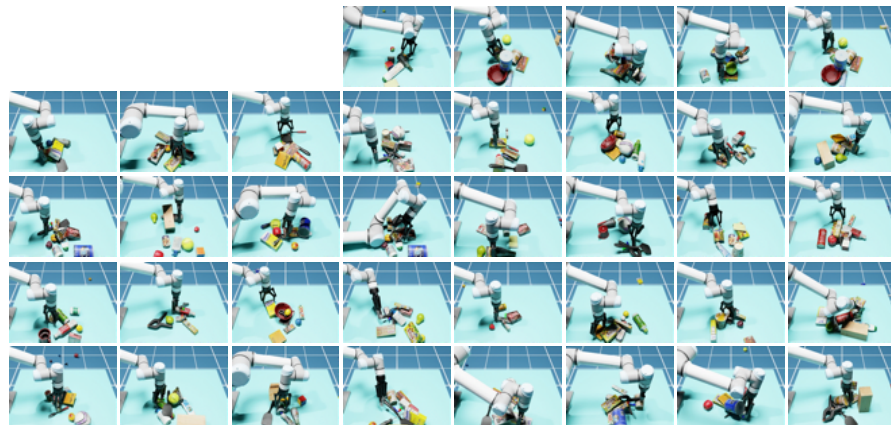
ロボットデータの収集方法例



実世界での人間の動作の**模倣**

- 正確で器用な作業
- ×収集の手間

実稼働ロボットデータを使う方法も



仮想環境（シミュレーション）データ
による学習

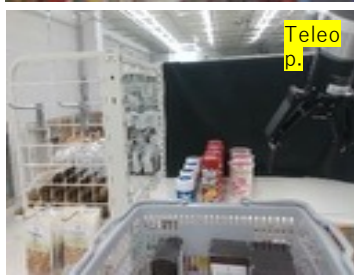
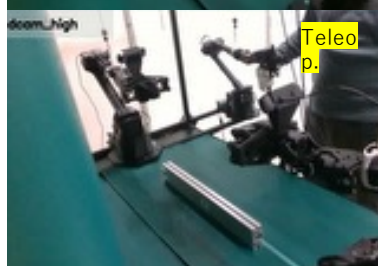
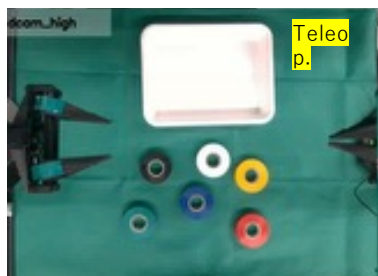
- 収集しやすい
- ×不正確な物理

WEB上の動画を使う方法も近い

AISTでも2023年より模倣学習データを収集開始



Tomohiro
Motoda

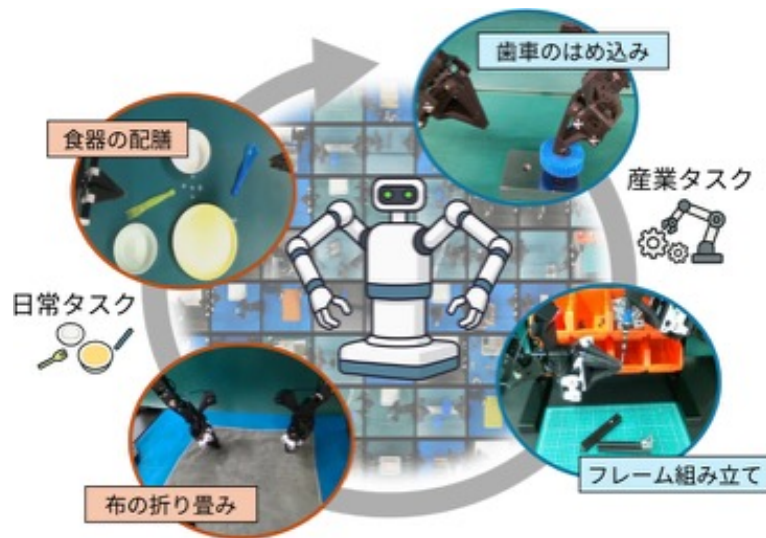
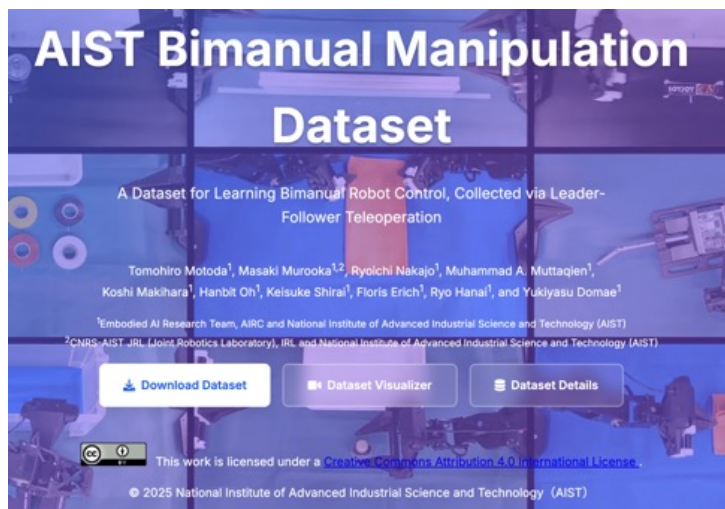




双腕ロボットAIの開発を支援するデータセットを無償公開

—両手を使うロボットAI開発の足がかりを提供—

- 器用な動作の実現を目指した双腕ロボットの作業データセット「AIST-Bimanual Manipulation」を無償公開
- 既に公開されているソフトウェアフレームワーク「RoboManipBaselines」との連携により、開発環境構築が容易となり、双腕ロボットAI開発の参入障壁が低減
- 製造業・物流・介護分野の人手不足解決と日本のロボットAI競争力向上に貢献



- 企業や研究者による利用多数
- 共同研究などでも活用中

シンプルな日常タスク、産業タスク、100タスク以上

AIST Bimanual Dataset : https://aistairc.github.io/aist_bimanip_site/

技術者がFMを試せるソフトウェアフレームワークを公開

- ・ データ収集からモデル訓練・推論実行までをサポート
- ・ 実機・シミュレーション両方をサポート
- ・ 環境・ロボット・モデルなど拡張可能



Masaki Murooka



- ・ GitHub over 300stars
- ・ CoRL 2025 workshopで発表
- ・ 企業や研究者による利用多数

RoboManipBaselines
<https://github.com/isri-aist/RoboManipBaselines>

FMの研究や応用が進んでいくためのフレームワークを提供



共同研究やコンサルティング契約で複数の国内企業との連携も実施中。

AISTのロボット基盤モデル研究

モデル構築の方法論

LLMに基づくプランニング
・対話的行動性

クロスモダリティ

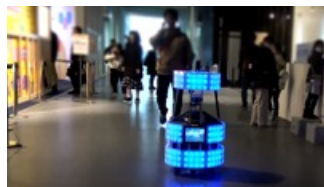
次世代
行動生成
モデル

ロボット
基盤モデル

一般家庭



商業施設、屋外



追加学習・
データセット

ドメイン適用



店舗



工場・物流



非現実 仮想環境 模倣
Unreal Sim Real

経験データ構築の方法論

モデル構築の原理原則理解から、
既存FMの応用研究、
さらに社会普及への活動を
一貫して実施。

データスケール化と
社会展開に向け
企業・大学・法人
などと連携