The Data Pyramid for Building Generalist Agents

Yuke Zhu

December 8, 2022
Building AI for Generalist Agents

- household robots
- virtual agents
- self-driving cars

Open-Ended Objectives
Massively Multitask
World Knowledge
Building AI for Generalist Agents

What data sources shall we use?
How can we best use them?
The Data Pyramid for Generalist Agents

Real-World Data
- Small scale and expensive to collect
- Ease of use for agent learning, direct transfer

Synthetic Data
- Unlimited simulated data (in theory)
- Content creation challenge, reality gap, computational burden

Web Data
- Massive scale and ever-growing
- Multimodal and unstructured
- Human-centered data

Common Crawl
- YouTube
- Reddit
- Wikipedia
The Data Pyramid for Generalist Agents

Real-World Data

Synthetic Data

Web Data

Leverage ALL data available!

Common Crawl
**MineDojo: Framework for Building Generalist Agents**

Open-ended Environment

Generalist Agent

Internet-scale Knowledge Base

**MineDojo**: Framework for Building Generalist Agents

Diverse open-ended environment

- Complex 3D world
- 1500+ programmatic tasks
  - Success conditions & reward functions are well-defined
  - Templated language prompts
- 1500+ creative tasks
  - No well-defined or easily automated success criterion
  - Reward function is very difficult to formulate

- Build a house in a cave
- Explore a desert temple
- Encircle llamas with fences
- Play fire ball with a ghast
**MineDojo**: Framework for Building Generalist Agents

Internet-scale knowledge base

MINE-DOJO Wiki

~7K multimodal pages:
texts, images, tables, diagrams
<table>
<thead>
<tr>
<th>Name</th>
<th>Ingredients</th>
<th>Crafting recipe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropper</td>
<td>Cobblestone + Redstone Dust</td>
<td><img src="image1" alt="Image" /></td>
</tr>
<tr>
<td>Daylight Detector</td>
<td>Glass + Nether Quartz + Any wood Slab</td>
<td><img src="image2" alt="Image" /></td>
</tr>
<tr>
<td>Piston</td>
<td>Any Planks + Cobblestone + Iron Ingot + Redstone Dust</td>
<td><img src="image3" alt="Image" /></td>
</tr>
<tr>
<td>Sticky Piston</td>
<td>Slimeball + Piston</td>
<td><img src="image4" alt="Image" /></td>
</tr>
<tr>
<td>Block of Redstone</td>
<td>Redstone Dust</td>
<td><img src="image5" alt="Image" /></td>
</tr>
<tr>
<td>Wood Door</td>
<td>Matching Planks</td>
<td><img src="image6" alt="Image" /></td>
</tr>
<tr>
<td>Golden Carrot</td>
<td>Gold Nugget + Carrot</td>
<td><img src="image7" alt="Image" /></td>
</tr>
<tr>
<td>Cake</td>
<td>Milk Bucket + Sugar + Egg + Wheat</td>
<td><img src="image8" alt="Image" /></td>
</tr>
<tr>
<td>Beetroot Soup</td>
<td>Beetroot + Bowl</td>
<td><img src="image9" alt="Image" /></td>
</tr>
<tr>
<td>Dried Kelp</td>
<td>Dried Kelp Block</td>
<td><img src="image10" alt="Image" /></td>
</tr>
<tr>
<td>Honey Bottle</td>
<td>Glass Bottle + Honey Block</td>
<td><img src="image11" alt="Image" /></td>
</tr>
<tr>
<td>Rabbit Stew</td>
<td>Cooked Rabbit + Carrot + Baked Potato + Any Mushroom + Bowl</td>
<td><img src="image12" alt="Image" /></td>
</tr>
<tr>
<td>Pickaxe</td>
<td>Any Planks or Iron Ingot or Gold Ingot or Diamond + Stick</td>
<td><img src="image13" alt="Image" /></td>
</tr>
<tr>
<td>Stone Pickaxe</td>
<td>Any stone-tier block + Stick</td>
<td><img src="image14" alt="Image" /></td>
</tr>
<tr>
<td>Stone Pickaxe</td>
<td>Cobblestone or Blackstone or Cobbled Deepslate + Stick</td>
<td><img src="image15" alt="Image" /></td>
</tr>
<tr>
<td>Pickaxe</td>
<td>Matching Damaged Pickaxes</td>
<td><img src="image16" alt="Image" /></td>
</tr>
<tr>
<td>Recovery Compass</td>
<td>Echo Shard + Compass</td>
<td><img src="image17" alt="Image" /></td>
</tr>
<tr>
<td>Spyglass</td>
<td>Amethyst Shard + Copper Ingot</td>
<td><img src="image18" alt="Image" /></td>
</tr>
</tbody>
</table>
340K posts
6.6M comments
560K images
My first time encasing an Ocean Monument. Help! What is the best way to sponge it!

With sponges

5x5 sections. Sponges only have so much power.

Sponges or a lot of dirt/sand

5x5 grid of sand squares. No joke but let’s you go down in a back and forth without fear of creating more source blocks

does anyone know why some of my wheat won't grow?(farming)

Gotta be the light add more torches

Light

It'll be your light level, as the wheat near the torch has grown. All you need to do is place more torches and you should see a difference.

light lv l too low
MineDojo: Framework for Building Generalist Agents

MineCLIP: scalable reward learning from YouTube videos

As I raised my axe in front of this pig, there’s only one thing you know is gonna happen… I’m also going to gather a little bit of stone from the side of this little hill here. … but I’m gonna go around gathering a little bit more wood from these trees.

[Xu et al. 2022; Radford et al. 2021]
MineDojo: Framework for Building Generalist Agents

Stack the history RGB frames

Observation

Action

MineCLIP
Correlation = 0.95

“Shear sheep to obtain wool”
MineDojo: Framework for Building Generalist Agents

- CLIP (OpenAI): Pre-trained CLIP model from OpenAI fails to generalize.
- Sparse Reward: Binary task-completion rewards lead to exploration challenges.
- Manual Reward: Dense manually curated rewards show oracle performances.
- Ours (Avg): Two variants of MineCLIP models learned from YouTube videos.
Towards Human-Level Generalist Minecraft Agents

Decorate winter wonderland

Recreate Hogwarts Castle

Simulate a tiny CPU

Fight Ender Dragon

mendojo.org
LID: Interactive Task Planning with Large Language Models

Trained on 300 billion tokens of text

GPT-3

Common Crawl

Learning objective: predict the next word

Application Domains:
text summarization, visual question answering, conversational AI, …

“Pre-Train Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022
Can we harness the commonsense knowledge in large language models for high-level task planning?

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022
LID: Interactive Task Planning with Large Language Models

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022
LID: Interactive Task Planning with Large Language Models

VirtualHome agent

current partial observation

history experiences

Inside(pancake, stove)

task goal

Tokenizer

Pre-trained Language Model

Policy Model

action

trained from scratch

pre-trained

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022
**LID: Interactive Task Planning with Large Language Models**

Significant improvement in combinatorial generalization to novel tasks

“Pre-Trained Language Models for Interactive Decision-Making.” Li, Shuang et al. NeurIPS 2022
**Input Prompt:**
How to “find material and craft a gold pickaxe” in Minecraft? Let’s think step by step.

**GPT-3:**
1) Find a place with a lot of trees;
2) Cut down the trees and gather the wood;
3) Find a place with a lot of stone;
4) Mine the stone and gather the cobblestone;
5) Find a place with a lot of iron;
6) Mine the iron and gather the iron ingots;
7) Find a place with a lot of gold;
8) Mine the gold and gather the gold ingots;
9) Craft a gold pickaxe.
VIMA: General Robot Manipulation with Multimodal Prompts

What if we can prompt a household robot to …

Input Prompt: Recite the first law of robotics

Output: A robot may not injure a human being

What if we can prompt a household robot to …

Bring me... This is a sponge. You can use it to clean the table.

Do this... Never enter this room!
VIMA: General Robot Manipulation with Multimodal Prompts

Visual Goal: Rearrangement
Rearrange objects to match this scene:

One-shot Demonstration
Follow this motion for

Novel Concept Grounding
This is a blinket
This is a wug
Put a wug into a blinket

Visual Constraint
Sweep all into without exceeding

VIMA

generalist robot agent for multi-task learning and zero-shot generalization

“VIMA: General Robot Manipulation with Multimodal Prompts.” Jiang, Yunfan et al. ArXiv 2022
VIMA: General Robot Manipulation with Multimodal Prompts

VIMA: Visuo-Motor Attention model

- Transformer encoder-decoder;
- Encode multimodal prompts with a frozen LM (Google T5);
- Decode robot arm actions given the prompt and interaction history.

“VIMA: General Robot Manipulation with Multimodal Prompts.” Jiang, Yunfan et al. ArXiv 2022
**VIMA:** General Robot Manipulation with **Multimodal Prompts**

- **Cross-attention** to condition history on prompt;
VIMA: General Robot Manipulation with Multimodal Prompts

- **Cross-attention** to condition history on prompt;

- Alternate **cross-attention** and **causal self-attention** to decode actions;
VIMA: General Robot Manipulation with Multimodal Prompts

- **Cross-attention** to condition history on prompt;

- Alternate **cross-attention** and **causal self-attention** to decode actions;

- **Objects** as tokens.
VIMA: General Robot Manipulation with **Multimodal Prompts**

Open-source large-scale simulation benchmark

- 6 task categories unified by the sequence model;
- 1000s of procedurally generated tasks, paired with multimodal prompts;
- Scripted oracles to generate expert demonstrations.
VIMA: General Robot Manipulation with Multimodal Prompts

Four-level evaluation protocol for systematic generalization

- systematically measuring the **zero-shot generalization capability** of trained agents.
- Each level *deviates more from the training distribution* and is strictly more challenging.
**VIMA:** General Robot Manipulation with Multimodal Prompts

Model scalability from 2M to 200M parameters

---

**Level 1**
Object Placement
- Put the block into the box

**Level 2**
Novel Combination
- Put the block into the container

**Level 3**
Novel Object
- Put the object into the container

**Level 4**
Novel Task
- Put all objects with the same texture as another object into it

---

**Model Scalability**

- **L1**
- **L2**
- **L3**
- **L4**

Success Rate (%) vs. Model Size (M)

- Ours
- Gato
- Flamingo
- DT
VIMA: General Robot Manipulation with Multimodal Prompts

Data scalability from 0.1% to full dataset

Level 1: Object Placement
Put the \[\text{object}\] into the \[\text{location}\]

Level 2: Novel Combination
Put the \[\text{object}\] into the \[\text{location}\]

Level 3: Novel Object
Put the \[\text{object}\] into the \[\text{location}\]

Level 4: Novel Task
Put all objects with the same texture as \[\text{object}\] into it

Graphs showing data scalability for levels L1 to L4, with success rate (%) on the y-axis and data size on the x-axis. The graphs compare success rates between different methods: Ours, Gato, Flamingo, and DT.
SAILOR: Data-Efficient Learning from Diverse Prior Data

Real-world data is expensive. Can we gather all prior data to accelerate learning of target tasks?

Prior Robotic Dataset: related tasks, task-agnostic data, etc.

large-scale prior robotic data collected in diverse settings

+ small amount of target task data
SAILOR: Data-Efficient Learning from Diverse Prior Data

Mapping past experiences into latent skills

Extracts reusable behaviors from rich, diverse data

How to learn skill representations that enable effective downstream learning?

How to learn robust policy given limited target task demonstrations?

"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022
SAILOR: Data-Efficient Learning from Diverse Prior Data

Learning Skill Representations

---

"Learning and Retrieval from Prior Data for Skill-based Imitation Learning." Nasiriany, Soroush et al. CoRL 2022
SAILOR: Data-Efficient Learning from Diverse Prior Data

Learning Skill Representations

Retrieval from Prior Data

Target Task Data

Aggregated Policy Learning Data

Retrieved Prior Data

“Learning and Retrieval from Prior Data for Skill-based Imitation Learning.” Nasiriany, Soroush et al. CoRL 2022
**SAILOR: Data-Efficient Learning from Diverse Prior Data**

Learning Skill Representations  
Retrieval from Prior Data  
Policy Learning

Target Task Data  
Aggregated Policy Learning Data

$z_0$ $o_0$  
$z_1$ $o_1$  
...  
$z_H$ $o_H$

Skill encoder
Policy

$L_2$  
L2 decoder  
$\hat{o}_{0:H-1}$

“Learning and Retrieval from Prior Data for Skill-based Imitation Learning.” Nasiriany, Soroush et al. CoRL 2022
SAILOR: Data-Efficient Learning from Diverse Prior Data

Prior Data
~24 hours play data across four environments

Target Tasks
30 human demonstrations collected in ~30 minutes

Setting Up

Cleaning Up

CALVIN: A benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks, Mees et al., IEEE-RAL 2022
**SAILOR: Data-Efficient Learning from Diverse Prior Data**

<table>
<thead>
<tr>
<th>Method</th>
<th>Task Success Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC-RNN (Mandlekar et al.)</td>
<td>41.0</td>
</tr>
<tr>
<td>BC-RNN (FT)</td>
<td>61.8</td>
</tr>
<tr>
<td>BC-RNN (R3M) (Nair et al.)</td>
<td>28.0</td>
</tr>
<tr>
<td>IQL (Kostrikov et al.)</td>
<td>14.7</td>
</tr>
<tr>
<td>IQL (UDS) (Yu et al.)</td>
<td>20.0</td>
</tr>
<tr>
<td>FIST (Hakhamaneshi et al.)</td>
<td>9.3</td>
</tr>
<tr>
<td>Ours</td>
<td>88.0</td>
</tr>
</tbody>
</table>
SAILOR: Data-Efficient Learning from Diverse Prior Data

- BC-RNN (Mandlekar et al.)
- BC-RNN (FT)
- BC-RNN (R3M) (Nair et al.)
- IQL (Kostrikov et al.)
- IQL (UDS) (Yu et al.)
- FIST (Hakhamaneshi et al.)
- Ours

Task Success Rate (%)

0 20 40 60 80 100

41.0 61.8 28.0 14.7 20.0 9.3 88.0
SAILOR: Data-Efficient Learning from Diverse Prior Data

- BC-RNN (Mandlekar et al.): 41.0%
- BC-RNN (FT): 61.8%
- BC-RNN (R3M) (Nair et al.): 28.0%
- IQL (Kostrikov et al.): 14.7%
- IQL (UDS) (Yu et al.): 20.0%
- FIST (Hakhamaneshi et al.): 9.3%
- Ours: 88.0%

Task Success Rate (%)

Bar chart comparing different methods for task success rate.
SAILOR: Data-Efficient Learning from Diverse Prior Data

<table>
<thead>
<tr>
<th>Method</th>
<th>Task Success Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BC-RNN (Mandlekar et al.)</td>
<td>41.0</td>
</tr>
<tr>
<td>BC-RNN (FT)</td>
<td>61.8</td>
</tr>
<tr>
<td>BC-RNN (R3M) (Nair et al.)</td>
<td>28.0</td>
</tr>
<tr>
<td>IQL (Kostrikov et al.)</td>
<td>14.7</td>
</tr>
<tr>
<td>IQL (UDS) (Yu et al.)</td>
<td>20.0</td>
</tr>
<tr>
<td>FIST (Hakhamaneshi et al.)</td>
<td>9.3</td>
</tr>
<tr>
<td>Ours</td>
<td>88.0</td>
</tr>
</tbody>
</table>
SAILOR: Data-Efficient Learning from Diverse Prior Data

Large-scale prior data collection in real world

• Diverse kitchen environment

• 8 food items, one pot, one pan

• Unstructured interaction data: 300,000 timesteps.

“Learning and Retrieval from Prior Data for Skill-based Imitation Learning.” Nasiriany, Soroush et al. CoRL 2022
SAILOR: Data-Efficient Learning from Diverse Prior Data

Effective imitation learning for vision-based manipulation

BC-RNN (FT): 46.7% success rate

Ours: 76.7% success rate

Vision-based policies trained on 30 demonstrations for the target task

ut-austin-rpl.github.io/sailor
VIOLA: Imitation Learning with Object-Centric Priors

VIOLA: Imitation Learning with Object-Centric Priors

Raw visual observation

General object proposals

Pre-trained region proposal network on large image datasets

VIOLA: Imitation Learning with Object-Centric Priors

Encoding object visual appearances and their spatial locations

VIOLA: Imitation Learning with Object-Centric Priors

Using transformers to select task-relevant objects and reason about their relations

VIOLA only needs 50 raw demonstrations through easy teleoperation and it learns to use the K-cup machine to make coffee without any annotations on raw data. VIOLA can reward a hard-working researcher with a cup of good coffee!
The Data Pyramid for Generalist Agents

Real-World Data

SAILOR & VIOLA [Nasiriany et al. & Zhu et al. CoRL 2022]: Data-efficient and robust imitation learning in the real world with diverse prior robot data

Synthetic Data

VIMA [Jiang et al. NeurIPS FMDM 2022]: Building generalist robotic agents with multimodal prompts in large-scale simulation

Web Data

MineDojo, LID [Fan et al. & Li et al. NeurIPS 2022]: Building generally capable agents with Internet-scale knowledge

Common Crawl

YouTube, Reddit
The Data Pyramid for Generalist Agents

- Real-World Data
- Synthetic Data
- Web Data

Massive, Multimodal, Multitask Model for Generalist Agents
Acknowledgement

Soroush Nasiriany
Yifeng Zhu
Shuang Li
Ajay Mandlekar

Linxi “Jim” Fan
Yunfan Jiang
Guanzhi Wang
Anima Anandkumar