

# PENG LI

Fudan University ◇ Shanghai, 200433, P.R.China

✉ [lip21@m.fudan.edu.cn](mailto:lip21@m.fudan.edu.cn) ☎ (+1) 312-241-6488 🏠 [artpli.github.io](https://artpli.github.io) 🔄 [artpli](https://artpli.github.io)

## EDUCATION

### Fudan University

*M.S., Computer Science*

FudanNLP Group, Adviser: [Prof. Xipeng Qiu](#)

Shanghai, China

2021.09 – 2024.06 (*expected*)

### East China Normal University

*B.Eng. Data Science*

NLP Group, Adviser: [Prof. Yuanbin Wu](#)

Shanghai, China

2016.09 – 2020.06

### Toyota Technological Institute at Chicago (TTIC)

*Visiting Student*

TTIC NLP & Robotics Lab, Advisers: [Prof. Hongyuan Mei](#) and [Prof. Matthew R. Walter](#)

Chicago, U.S.

2023.05 – *Present*

## RESEARCH INTERESTS

My research interest lies mainly in **foundation models for robot learning**. My research goal is to build general-purpose robots to autonomously interact with the real world and achieve self-evolution. To achieve this goal, I am (1) enhancing the perception, reasoning, planning, acting and learning capabilities of robots with existing CV/NLP foundation models [P2,P4]; (2) designing and pretraining robotics-specific foundation models [P3]; (3) building new evaluation benchmarks [P1].

## PREPRINTS AND PUBLICATIONS

\* indicates equal contribution.

### P1 MANGO: A Benchmark for Evaluating Mapping and Navigation Abilities of Large Language Models

Peng Ding\*, Jiading Fang\*, [Peng Li\\*](#), Kangrui Wang\*, Xiaochen Zhou\*, Mo Yu, Jing Li, Matthew Walter, Hongyuan Mei  
Submitted to *ICLR 2024*. ([OpenReview](#))

### P2 Statler: State-Maintaining Language Models for Embodied Reasoning

Takuma Yoneda\*, Jiading Fang\*, [Peng Li\\*](#), Huanyu Zhang\*, Tianchong Jiang, Shengjie Lin, Ben Picker, David Yunis, Hongyuan Mei, Matthew R. Walter.  
Submitted to *ICRA 2024*. ([ArXiv](#))

### P3 MOSS: Training Conversational Language Models from Synthetic Data

Tianxiang Sun, Xiaotian Zhang, Zhengfu He, [Peng Li](#), Qinyuan Cheng, Hang Yan, Xiangyang Liu, Yunfan Shao, Qiong Tang, Xingjian Zhao, Ke Chen, Yining Zheng, Zhejian Zhou, Ruixiao Li, Jun Zhan, Yunhua Zhou, Linyang Li, Xiaogui Yang, Lingling Wu, Zhangyue Yin, Xuanjing Huang, Xipeng Qiu

★ MOSS is the first ChatGPT-like LLM in China, the first plugin-augmented LLM in China, and fully open-sourced with 11.7k+ stars.  
In submission. ([GitHub](#))

### P4 CodeIE: Large Code Generation Models are Better Few-Shot Information Extractors

[Peng Li\\*](#), Tianxiang Sun\*, Qiong Tang, Hang Yan, Yuanbin Wu, Xuanjing Huang, Xipeng Qiu.  
Accepted by *ACL 2023*. ([ArXiv](#))

## RESEARCH EXPERIENCE

### Toyota Technological Institute at Chicago (TTIC)

May 2023 – Present

TTIC NLP & Robotics Lab, Advisers: [Prof. Hongyuan Mei](#) and [Prof. Matthew R. Walter](#)

Chicago, U.S.

#### Statler: State-Maintaining Language Models for Embodied Reasoning

- Utilized a second LLM to simultaneously manage the status information of objects while employing a first codewriting LLM to write the robot's policy code for decision-making tasks.
- The simulation and real-robot experimental results of three robotics tasks, namely “pick and place”, “block disinfection” and “weight reasoning”, indicate that its performance exceeds that of the Code-as-Policies method.

## MANGO: A Benchmark for Evaluating Mapping and Navigation Abilities of Large Language Models

- Constructed the first benchmark to assess the mapping and navigation abilities of LLMs. Extracted walkthroughs from 53 interactive text-games and designed two types of questions to test these two capabilities of LLMs
- Evaluated the effectiveness of several black-box and open-source LLMs including ChatGPT, Claude, LLaMA, RWKV, and analyzed the influencing factors.
- Experiments also indicated that good mapping and navigation capabilities can help LLMs play text-based games more effectively.

## Building World Models That Learn From Real-World Interactions

- Proposed and currently implementing a general reinforcement learning framework to enable LLMs to continuously enhance their reasoning and planning abilities through interaction with the real environment.

## **Shanghai AI Laboratory**

**May 2023 – Present**

*OpenLMLab, Advisers: Dr. Hang Yan and Dr. Kai Chen*

*Shanghai, China*

### Augmenting External Tools for 100B LLMs at Shanghai AI Lab

- Implemented and optimized the functionality to call external tools, such as search engines, calculators, equation solvers, and drawing tools of the LLM at Shanghai AI Lab (InternLM). Prepared the SFT data for utilizing these tools and fine-tuned 100B LLMs using these data.

### Exploring Improved Structures of Pre-training Corpora for Math-LLMs and Code-LLMs

- Currently engaged in extensive research to develop effective pre-training data strategies for training more powerful LLMs specifically designed for mathematical and coding tasks.

## **Fudan University**

**May 2021 – Present**

*NLP Group, Adviser: Prof. Xipeng Qiu*

*Shanghai, China*

### MOSS: China's First ChatGPT-like Large Language Model

- Participated in the design and implementation of the open-source tool-augmented 16B conversational language model “MOSS” as a core member.
- Helped prepare the training data for the first version of MOSS. The goal was to make MOSS helpful, honest and harmless (HHH).
- Augmented MOSS with external tools like the search engine, calculator, text2image tools etc. These tools enable MOSS to access up-to-date information, have better mathematical skills, have the ability to produce images and so on.
- Responsible for processing sensitive data in conversations with MOSS.

### CodeIE: Utilizing Code-LLMs for Better Few-Shot Information Extraction

- Proposed to utilize Code-LLMs like Codex with structured code-format prompts to perform structured few-shot information extraction tasks.
- Designed concrete code-format prompts for the few-shot NER and RE tasks. Then prompted OpenAI's *code-davinci-002* model with in-context demonstrations to perform these IE tasks.
- Experimental results on seven benchmarks show that our method consistently outperforms fine-tuning moderate-size pre-trained models specially designed for IE tasks (e.g., UIE) and prompting NL-LLMs (e.g., GPT-3) under few-shot settings.
- Further conducted a series of in-depth analyses to demonstrate the merits of leveraging Code-LLMs for IE tasks.

## **SERVICES**

### **Reviewer**

EMNLP

2021, 2022, 2023

ICRA

2024

### **Teaching Assistant**

Pattern Recognition and Machine Learning under Prof. Xipeng Qiu

2022-2023

Introduction to Artificial Intelligence under Prof. Xipeng Qiu

2022

Statistical Natural Language Processing under Prof. Yuanbin Wu

2021

## **SKILLS**

**Programming:** Python, C/C++, LaTeX; Scikit-learn, PyTorch

**Language:** Mandarin, English