

# Dongchan Shin

dcshin718@gmail.com | Github | Google Scholar

## EDUCATION

### The University of Hong Kong

Bachelor of Science (Major in Chemistry, minor in Computer Science)

Pok Fu Lam, Hong Kong

Sep 2018 - Dec 2024

## PUBLICATIONS

### AgentRewardBench: Evaluating Automatic Evaluations of Web Agent Trajectories

COLM 2025

Xing Han Lù, Amirhossein Kazemnejad, Nicholas Meade, Arkil Patel, **Dongchan Shin**, Alejandra Zambrano, Christopher J. Pal, Siva Reddy

### Spider2.0: Evaluating Language Models on Real-World Enterprise Text-to-SQL Workflows

ICLR 2025, *Oral*

Fangyu Lei\*, Jixuan Chen\*, Yuxiao Ye, Ruisheng Cao, **Dongchan Shin**, Hongjin Su, Zhaoqing Suo, Hongcheng Gao, Wenjing Hu, Pengcheng Yin, Victor Zhong, Caiming Xiong, Ruoxi Sun, Qian Liu, Sida I. Wang, Tao Yu

### OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments

NeurIPS 2024 Dataset & Benchmark Track

Tianbao Xie, Danyang Zhang, Jixuan Chen, Xiaochuan Li, Siheng Zhao, Ruisheng Cao, Toh Jing Hua, Zhoujun Cheng, **Dongchan Shin**, Fangyu Lei, Yitao Liu, Yiheng Xu, Shuyan Zhou, Silvio Savarese, Caiming Xiong, Victor Zhong, Tao Yu

## RESEARCH EXPERIENCE

### Vector Institute

Toronto, Ontario (Remote)

Research Intern, Advisor: Prof. Victor Zhong

Nov 2025 - Present

Topic: Automatic Structure Induction for Hybrid Knowledge Search

- Building a novel framework that can accurately and efficiently answer natural language questions over a large collection of unstructured documents

### Mila – Quebec AI Institute

Montreal, Quebec

Research Intern, Advisor: Prof. Siva Reddy

Mar 2025 – Jun 2025

Topic: Test-Time Adaptability of Web Agents

- Building a benchmark to evaluate how well web agents infer and adapt to implicit user preferences

### XLANG Lab, HKU

Pok Fu Lam, Hong Kong

Research Assistant, Advisor: Prof. Tao Yu

May 2023 – Dec 2024

Topic: Evaluating Language Agents

- Conducted evaluation of LLM/VLM agents in real computer environments and enterprise text-to-SQL workflows

## LANGUAGES & INTERESTS

**Languages:** Korean (Native), English (Fluent), Mandarin Chinese (Fluent)

**Interests:** Language Agents, Guitar, Baseball

## TECHNICAL SKILLS

**Language:** Python, SQL, L<sup>A</sup>T<sub>E</sub>X

**Library / Framework:** PyTorch, Langchain, transformers, vLLM

## SERVICES

**Reviewer:** ACL 2025 Demonstration Track

**Others:** Co-Founder of HKU Korean Coding Society