

# NARUTATSU (EDWARD) RI

wl2787@columbia.edu

---

EDUCATION	<b>Columbia University, School of Engineering and Applied Science</b> B.S. Candidate in Computer Science, GPA: 4.02/4.00. Egleston Scholar (Top 1% of undergraduate engineering applicants). Ezoe Memorial Foundation Academic Scholarship (Full-ride scholarship)	<b>Sep. 2020 – May 2024</b>
PUBLICATIONS	<p>[1] <i>Enhancing Few-shot Text-to-SQL Capabilities of Large Language Models: A Study on Prompt Design Strategies</i> Linyong Nan, Yilun Zhao, Weijin Zou, <b>Narutatsu Ri</b>, Jaesung Tae, Ellen Zhang, Arman Cohan, Dragomir Radev arXiv Preprint, 2023.</p> <p>[2] <i>IdEALS: Idiomatic Expressions for Advancement of Language Skills</i> <b>Narutatsu Ri</b>, Bill Sun, Sam Davidson, Zhou Yu arXiv Preprint, 2023.</p> <p>[3] <i>Contrastive Loss is All You Need to Recovery Analogies as Parallel Lines</i> <b>Narutatsu Ri</b>, Fei-Tzin Lee, Nakul Verma ACL 2023 Workshop (RepL4NLP)</p>	
RESEARCH EXPERIENCE	<b>Columbia University Department of Computer Science</b> <i>Undergraduate Researcher</i> , with Daniel Hsu <ul style="list-style-type: none"><li>Research on model selection capabilities of language models and transformer approximation theory.</li></ul>	<b>Jan. 2023 – Present</b>
	<b>Columbia University Department of Computer Science</b> <i>Independent Researcher</i> , with Nakul Verma <u>Recovering Analogies as Parallel Lines</u> <ul style="list-style-type: none"><li>Designed CWM (<i>Contrastive Word Model</i>), word embedding model with simple contrastive learning objective with analogy recovery performance competitive with popular word embedding models with 50× training efficiency.</li><li>Conducted theoretical analysis on CWM and proposed connection between analogy formation and word co-occurrence statistics.</li></ul> <u>Embeddability of Parallelograms in <math>\mathbb{R}^d</math> space</u> <ul style="list-style-type: none"><li>Research on metric embeddings to identify relation between embeddability of sets of points satisfying parallelogram conditions and dimensionality. Proved linear relationship between dimensionality and number of analogy conditions and empirically verified robustness.</li></ul> <u>Properties of Word Embeddings Optimized for Word Similarity Tasks</u> <ul style="list-style-type: none"><li>Developed t-SNE-based word embedding and empirically showed competitive performance on word similarity tasks.</li></ul> <u>Deep Embedded Clustering</u> <ul style="list-style-type: none"><li>Improved model introduced by paper <i>Unsupervised Deep Embedding for Clustering Analysis</i> by adding local structure preserving property.</li></ul>	<b>Jan. 2021 – Present</b>
	<b>Columbia University Natural Language Processing Lab</b> <i>Lead Researcher</i> , with Zhou Yu <ul style="list-style-type: none"><li>Research on word choice improvement and developing phrasal recommendation language models.</li></ul>	<b>Aug. 2022 – May 2023</b>
	<b>Columbia University Natural Language Processing Lab</b> <i>Undergraduate Researcher</i> , with Kathleen McKeown, Zhou Yu <ul style="list-style-type: none"><li>Research on faithfulness of large language model-generated explanations and simulatability.</li></ul>	<b>Sep. 2022 – May 2023</b>

**The University of Tokyo International Research Center for Neurointelligence**

**May 2022 – Sep. 2022**

*Visiting Researcher*, with Mingbo Cai

- Analyzed manifold structure in contextualized Transformer word embeddings.
- Developed model to predict and decode syntactic information of movie scene descriptions from raw fMRI brain activity.

**University of Electro-Communications iSYSLab**

**Oct. 2017 – Jun. 2019**

*Special Researcher*, with Takayuki Nagai

- Conducted research for 2.5 years on a dialogue system capable of utilizing only latent syntactic and semantic information for response generation.
- Presented work at the 33rd Annual Conference of the Japanese Society for Artificial Intelligence.

*A Dialogue System Implemented with Latent Parameters*

**Weida Li**, Chie Hieda, Takayuki Nagai

Journal of Proceedings of the Annual Conference of JSAI, 2019.

**WORK  
EXPERIENCE**

**Columbia University Department of Computer Science**

**Jul. 2022 – May 2023**

*Head Teaching Assistant, Machine Learning (COMS 4771)*

- Tested out of COMS 4771 (Graduate Machine Learning course) in sophomore year and serving as Head Teaching Assistant.
- Served as TA for Summer 2022, Fall 2022 (Head TA), Spring 2023 (Head TA) semesters.

**Project Thryving, DATAFLUCT, INC.**

**Jul. 2021 – Jul. 2022**

*Software Engineer*, Subcontractor

- Developed an interactive business intelligence providing dialogue system that analyzes trends and anomalies in corporate sales data.

**AWARDS &  
HONORS**

Dean's List

**2020 – 2023**

Tau Beta Pi candidate

**2022, 2023**

Egleston Scholar

**2020**

Ezoe Memorial Foundation Academic Scholarship

**2019**

(Selected as 1 out of 5 students among thousands of applicants)

University Robotics Competition (RoboCup Japan Open) Winner

**2018**

S-ISEF Finalist

**2018**

National Programming Contest Silver Medal

**2017**

**COURSEWORK**

**Graduate Coursework:**

- *Computer Science*: Machine Learning (Skipped), Unsupervised Learning, Advanced Algorithms, Natural Language Processing, Computational Learning Theory, Dialog Systems, Geometric Data Analysis, Robot Learning, Projects in Computer Science
- *Mathematics*: Probability Theory, Advanced Linear Algebra

**Undergraduate Coursework:**

- *Computer Science*: Data Structures, Fundamentals of Computer Systems
- *Mathematics*: Real Analysis, Ordinary Differential Equations