

# Asahi Ushio

Research Engineer, Google, London, UK  
asahi1992ushio@gmail.com  
<https://asahiushio.com/>

## RESEARCH AREA

Machine Learning, Natural Language Processing, Language Modelling, Audio Recognition/Generation.

## EDUCATION

10/2020 - 09/2023 **Ph.D.** in Computer Science & Information

**Cardiff University**, Cardiff, UK

Advisors: Jose Camacho Collados, Steven Schockaert

Area of study: Language model understanding, Question and answer generation,  
NLP for social media.

04/2016 - 03/2018 **M.E.** in Electrical Engineerings

**Keio University**, Kanagawa, Japan

Advisor: Masahiro Yukawa

Thesis: Projection-based Regularized Dual Averaging for Stochastic Optimization

Area of study: Stochastic optimization, Online learning, Sparsity-aware filtering

04/2011 - 03/2016 **B.E.** in Electrical Engineerings

**Keio University**, Kanagawa, Japan

Thesis:

(1) An Online Kernel Method Based on Orthogonal Projection Reflecting Multiple Tasks

Advisor: Masahiro Yukawa

Area of study: Multi-task learning, Adaptive learning, Kernel method

(2) Stock Return Prediction by Stochastic Model with Variance Risk Premium

Advisor: Teruo Nakatsuma

Area of study: Bayesian forecasting, Stochastic volatility, MCMC

## PUBLICATIONS

### Conference Proceedings

Dimosthenis Antypas, [Asahi Ushio](#), Francesco Barbieri, Jose Camacho-Collados, "Multilingual Topic Classification in X: Dataset and Analysis," in Proceedings of EMNLP main conference, November 2024.

Joanne Boisson, [Asahi Ushio](#), Hsuvas Borkakoty, Kiamehr Rezaee, Dimosthenis Antypas, Zara Siddique, Nina White, Jose Camacho-Collados, "How Are Metaphors Processed by Language Models? The Case of Analogies," in Proceedings of CoNLL main conference, November 2024.

[Asahi Ushio](#), Jose Camacho Collados, Steven Schockaert, "A RelEntLess Benchmark for Modelling Graded Relations between Named Entities," in Proceedings of EACL main conference, March 2024.

Dimosthenis Antypas, [Asahi Ushio](#), Francesco Barbieri, Leonardo Neves, Kiamehr Rezaee, Luis Espinosa-Anke, Jiaxin Pei, Jose Camacho-Collados, "SuperTweetEval: A Challenging, Unified and Heterogeneous Benchmark for Social Media NLP Research," in Proceedings of EMNLP findings, December 2023.

Asahi Ushio, Yi Zhou, Jose Camacho-Collados, "An Efficient Multilingual Language Model Compression through Vocabulary Trimming," in Proceedings of EMNLP findings, December 2023.

Asahi Ushio, Fernando Alva-Manchego, and Jose Camacho-Collados, "A Practical Toolkit for Multilingual Question and Answer Generation," in Proceedings of ACL system demonstrations, July 2023.

Asahi Ushio, Fernando Alva-Manchego, and Jose Camacho-Collados, "An Empirical Comparison of LM-based Question and Answer Generation Methods," in Proceedings of ACL findings, July 2023.

Asahi Ushio, Fernando Alva-Manchego, and Jose Camacho-Collados, "Generative Language Models for Paragraph-Level Question Generation: A Unified Benchmark and Evaluation," in Proceedings of EMNLP main conference, December 2022.

Jose Camacho-Collados, Kiamehr Rezaee, Talayeh Riahi, Asahi Ushio, Daniel Loureiro, Dimosthenis Antypas, Joanne Boisson, Luis Espinosa-Anke, Fangyu Liu, Eugenio Martinez-Camara, Gonzalo Medina, Thomas Buhrmann, Leonardo Neves, and Francesco Barbieri, "TweetNLP: Cutting-Edge Natural Language Processing for Social Media," in Proceedings of EMNLP system demonstrations, December 2022.

Asahi Ushio, Leonardo Neves, Vitor Silva, Francesco Barbieri, and Jose Camacho-Collados, "Named Entity Recognition in Twitter: A Dataset and Analysis on Short-Term Temporal Shifts," in Proceedings of AACL main conference, November 2022.

Dimosthenis Antypas\*, Asahi Ushio\*, Leonardo Neves, Vitor Silva, Francesco Barbieri, and Jose Camacho-Collados, "Twitter Topic Classification," in Proceedings of COLING main conference, October 2022 (\*Equal contribution).

Asahi Ushio, Steven Schockaert, and Jose Camacho-Collados, "Distilling Relation Embeddings from Pre-trained Language Models," in Proceedings of EMNLP main conference, November 2021.

Asahi Ushio, Federico Liberatore, and Jose Camacho-Collados, "Back to the Basics: A Quantitative Analysis of Statistical and Graph-Based Term Weighting Schemes for Keyword Extraction," in Proceedings of EMNLP main conference, November 2021.

Asahi Ushio, Luis Espinosa-Anke, Steven Schockaert, and Jose Camacho-Collados, "BERT is to NLP what AlexNet is to CV: Can Pre-Trained Language Models Identify Analogies?," in Proceedings of ACL main conference, August 2021.

Asahi Ushio and Jose Camacho-Collados, "T-NER: An All-Round Python Library for Transformer-based Named Entity Recognition," in Proceedings of EACL system demonstrations, April 2021.

Asahi Ushio, Masahiro Yukawa, "Projection-based Dual Averaging for Stochastic Sparse Optimization," in Proceedings of 42nd IEEE ICASSP, pp.2307–2311, New Orleans: USA, March 2017.

## Workshop Proceedings

Asahi Ushio and Jose Camacho Collados, "A Systematic Analysis on the Temporal Generalization of Language Models in Social Media," in Proceedings of 14th Workshop on Computational Approaches to Subjectivity, Sentiment & Social Media Analysis, August 2024.

Asahi Ushio, and Danushka Bollegala "Zero-shot Cross-Lingual Counterfactual Detection via Automatic Extraction and Prediction of Clue Phrases," in Proceedings of the 2nd Workshop on Multilingual Representation Learning, December 2022.

Aleksandra Edwards, Asahi Ushio, Jose Camacho-Collados, Helene de Ribaupierre, and Alun Preece "Guiding Generative Language Models for Data Augmentation in Few-Shot Text Classification," Proceedings of the 3rd Workshop on Data Science with Human in the Loop, December 2022.

## Journal Articles

Asahi Ushio and Masahiro Yukawa, "Projection-based Regularized Dual Averaging for Stochastic Optimization," IEEE Trans. Signal Processing, vol.67, no.10, pp.2720–2733, May 2019.

Ken McAlinn, Asahi Ushio, and Teruo Nakatsuma, "Volatility Forecasts Using Nonlinear Leverage Effects," Journal of Forecasting, 2019.

## Miscellaneous

Max F. Frenzel, Bogdan Teleaga, Asahi Ushio, "Latent Space Cartography: Generalised Metric-Inspired Measures and Measure-Based Transformations for Generative Models," arXiv:1902.02113, Feb 2019.

Yoshifumi Seki, Asahi Ushio, Takeshi Yoneda, Yutaka Matsuo, "Analyzing User Behavior in News App towards Detecting Click-Bait News," IEICE, Text Analysis symposium, Sep 2017 (Japanese only).

Asahi Ushio, Yoshinobu Takahashi, Seiya Satoh, Hiroshi Yamakawa, "A Research for Efficient Equivalence Structure Search by Local Distribution of Variables," IEICE, Information-Based Induction Sciences and Machine Learning, Nov 2016 (Japanese only).

Asahi Ushio, Ken McAlinn, Teruo Nakatsuma, "Nonlinear Leverage Effects in Asset Returns: Evidence from the U.S. and Japanese Stock Markets," 9th International Conference on Computational and Financial Econometrics, Dec, 2015.

Asahi Ushio, Masahiro Yukawa, "Multi-task Learning with Operator-valued Reproducing Kernels," IEICE, Signal Processing symposium, Nov, 2015.

Asahi Ushio, Ken McAlinn, Teruo Nakatsuma, "Estimate the Leverage Function of Stochastic Volatility Model by Particle Learning," The Japanese Association of Financial Econometrics and Engineering, JAFEE, 42th, Jan, 2014.

## EXPERIENCE

01/2025 - RESEARCH ENGINEER (Full-time)

**Google**, London, UK

· Research on multi-media generation.

01/2024 - 12/2024 APPLIED RESEARCH SCIENTIST (Full-time)

**Amazon**, Tokyo, Japan

· Research on information retrieval, language modeling, and product search.

· Develop production-level backend API for search index evaluation.

01/2024 - 12/2024 RESEARCH COLLABORATOR (part-time)

**Kotoba Technologies**, Tokyo, Japan

· Research on Japanese automatic speech recognition.

10/2023 - 12/2023 POSTDOCTORAL RESEARCH

**Cardiff University**, Cardiff, UK

· Research on NLP in social media.

06/2023 - 10/2023 RESEARCH SCIENTIST (Internship)

**Google**, Zurich, Switzerland

· Research on music captioning via large language model fine-tuning.

· Experiences:

◊ Fine-tune PaLM2 on a mixture of audio and text tasks with Pax.

◊ Implement custom datasets, tasks, and evaluation metrics on seqIO.

◊ Process audio dataset with audio representation models such as MuLan and w2v-BERT.

- Advisors: Andrea Agostinelli and Matej Kastelic (MusicLM/Muso)
- 09/2021 - 12/2021 RESEARCH SCIENTIST (Internship)  
**SNAP**, L.A, U.S.
  - Temporal-aware named-entity recognition in noisy SNS data.
- Advisors: Francesco Barbieri, Vitor Silva Sousa and Leonardo Neves
- 06/2021 - 09/2021 APPLIED RESEARCH SCIENTIST (Internship)  
**Amazon**, Tokyo, Japan
  - Research on multilingual counterfactual sentence detection on e-commerce.
  - Advisor: Danushka Bollegala
- 02/2018 - 10/2020 MACHINE LEARNING SCIENTIST (Full-time)  
**Cogent Labs**, Tokyo, Japan
  - Research on natural language understanding, information retrieval, and generative models.
  - Develop product-level APIs based on cutting-edge NLP models such as semantic search and entity recognition.
- 07/2017 - 11/2017 RESEARCHER (Internship)  
**Cogent Labs**, Tokyo, Japan
  - Research on brain computer interface and biomedical signal processing.

## TEACHING

- 10/2022 - 08/2023 TEACHING ASSISTANT  
**Cardiff University**, Computer Science and Informatics
  - Support modules of Fundamentals of Programming (CMT120), and Distributed and Cloud Computing (CMT202).
- 04/2016 - 09/2016 TEACHING ASSISTANT  
**Keio University**, Department of Electronics and Electrical Engineering
  - Support a module of electrics and electrical engineering experiment, a mandatory module for undergrad student in Department of Electronics, Electrical Engineering, Keio University.

## SCHOLARSHIP

- 10/2020 - 09/2023 PhD Scholarship 2020  
School of Computer Science & Informatics, Cardiff University  
(£15,285 per year & the international student tuition fees).
- 04/2016 - 03/2018 Postgrad Scholarship  
Fuji Seal Packaging Education and Scholarship Foundation  
(\$ 1,200 per year).

## ADDITIONAL SKILLS

LANGUAGE: Japanese, English, Chinese (Conversational)

SOFTWARE: Python (tensorflow, pytorch, sklearn, huggingface), Docker, Flutter for web, GCP, AWS, GitHub <https://github.com/asahi417>.