

# Yoshinari Fujinuma

## Office Address

Virtual

## Contact

fujinumay@gmail.com

<http://akkikiki.github.io>

Twitter: @akkikiki

## Education

University of Colorado Boulder, USA

Dec 2021

Ph.D. in Computer Science

Advisors

- Katharina Kann(2020-2021), Jordan Boyd-Graber(-2021), Michael J. Paul (-2019)

University of Tokyo, Japan

Sep 2014

M.S. in Computer Science; Advisor: Akiko Aizawa

International Christian University, Japan

Mar 2012

B.A. in Computer Science and Mathematics; Advisor: Grant Pogosyan

## Professional Experience

**Applied Scientist**, Amazon.com, Inc., USA

Sep 2021 - Present

**Part-Time Instructor**, Univ. of Colorado Boulder, USA

Jan 2021 - May 2021

- CSCI 4622 Machine Learning ([Github Repo](#))

**Teaching Assistant**, Univ. of Colorado Boulder, USA

Jan-Dec 2020, May-Jul 2021

- CSCI 2270 Data Structures

**Applied Scientist Intern**, Amazon.com, Inc., USA

May 2020 - Aug 2020

**Research Assistant**, Univ. of Colorado Boulder, USA

Aug 2016 - Dec 2019

**Applied Scientist Intern**, Amazon.com, Inc., USA

May 2018 - Aug 2018

**Software Development Engineer**, Amazon Japan K.K., Japan  
2016

Oct 2014 - Aug

**Software Engineer Intern**, Amazon Japan K.K., Japan

Nov 2013 - Feb 2014

**Part-time Engineer**, Atilika, Japan

Aug- Nov 2013, Apr- Sep 2014

**Software Engineer Intern**, Cookpad, Japan

July 2013 (one month)

## Publications

- Pietro Lesci, [Yoshinari Fujinuma](#), Momchil Hardalov, Chao Shang, Lluís Marquez: “[Diable: Efficient Dialogue State Tracking as Operations on Tables](#)”, ACL Findings (long) 2023
- [Yoshinari Fujinuma](#), Jordan Boyd-Graber, Katharina Kann: “[Match the Script, Adapt if Multilingual: Analyzing the Effect of Multilingual Pretraining on Cross-lingual Transferability](#)”, ACL (long) 2022
- [Yoshinari Fujinuma](#), Masato Hagiwara: “[Semi-Supervised Joint Estimation of Word and Document Readability](#)”, TextGraphs-15@NAACL (short) 2021
- Mozhi Zhang\*, [Yoshinari Fujinuma\\*](#), Michael J. Paul, Jordan Boyd-Graber: “[Why Overfitting Isn’t Always Bad: Retrofitting Cross-Lingual Word Embeddings to Dictionaries](#)”, ACL (short) 2020

- Mozhi Zhang, Yoshinari Fujinuma, Jordan Boyd-Graber: “[Exploiting Cross-Lingual Subword Similarities in Low-Resource Document Classification](#)”, AAAI 2020
- Yoshinari Fujinuma, Jordan Boyd-Graber, Michael J. Paul: “[A Resource-Free Evaluation Metric for Cross-Lingual Word Embeddings based on Graph Modularity](#)”, ACL (long) 2019
- Dasha Pruss, Yoshinari Fujinuma, Ashlynn R. Daughton, Michael J. Paul, Brad Arnot, Danielle Albers Szafir, Jordan Boyd-Graber: “[Zika discourse in the Americas: A multilingual topic analysis of Twitter](#)”, PLOS ONE 2019
- Yoshinari Fujinuma, Alvin Grissom II: “[Substring Frequency Features for Segmentation of Japanese Katakana Words with Unlabeled Corpora](#)”, IJCNLP (short) 2017
- Yoshinari Fujinuma, Hikaru Yokono, Pascual Martínez-Gómez, Akiko Aizawa: “[Distant-supervised Language Model for Detecting Emotional Upsurge on Twitter](#)”, PACLIC (long) 2015

\*denotes equal contribution

## **Selected Projects Analysis on Multilingual Pretraining 2020 - 2021**

- Investigating the effect of using different pretraining languages. ([Paper](#))

## **Intrinsic Evaluation Measure for Cross-Lingual Embeddings 2017 - 2019**

- Developed a graph-based intrinsic measure to evaluate the quality of cross-lingual word embeddings. ([Paper](#))

## **Finite State Transducer (FST) for Kuromoji 2015**

- Replaced a double-array trie to an FST for Kuromoji, a java-based Japanese tokenizer used in Lucene, Solr, and Elastic Search. ([Github Repo](#))

## **Academic Service**

- Area Chair:
  - 2023 ACL (Multilingualism and Cross-lingual NLP track)
- Program Committee/Reviewer:
  - 2023 AAAI, ACL Rolling Review (Feb., Apr.)
  - 2022 AAAI, ACL Rolling Review (Jan., Mar., Jun., Oct., Dec.), ACL, EMNLP, NAACL SRW, W-NUT, CoNLL
  - 2021 NAACL, ACL, EMNLP, ACL Rolling Review (Sept., Oct., Nov.), CoNLL, NAACL SRW, ACL SRW, W-NUT, Workshop on NLP for Indigenous Languages of the Americas (AmericasNLP)
  - 2019 Workshop on Noisy User-Generated Text (W-NUT)
- Secondary Reviewer:
  - 2019 ACL
  - 2017 EMNLP, WWW
- Student Volunteer: 2020 ACL

## **Academic Honors**

- Outstanding Reviewer for CoNLL 2021
- Departmental Travel Grant from CU Boulder (300 USD) 2017
- Graduate School Travel Grant from CU Boulder (500 USD) 2017
- Dean’s Fellowship from CU Boulder (1/2 of the yearly tuition and stipend) 2016
- Graduate School Travel Grant from Univ. of Tokyo (80,000 JPY) 2013
- Gödel Foundation Prize: Best Bachelor thesis in CS and Math (50,000 JPY) 2012
- Horie Takematsu and Koh Scholarship (2,000 USD) 2010
- Student Scholarship from ICU (1/3 of the yearly tuition) 2008-2012

**Computer and  
Language Skills**

Languages:

Software:

English:

Domain-specific:

Proficient: Python; Intermediate: C++, Java

PyTorch, Git, Vim, L<sup>A</sup>T<sub>E</sub>X, MySQL

TOEFL iBT 101 (2015)

machine learning, natural language processing