

Yoshinari Fujinuma

Office Address

Virtual

Contact

fujinumay@gmail.com

<http://akkikiki.github.io>

Twitter: @akkikiki

Education

University of Colorado Boulder, USA

Aug. 2016 - Present

Advisors

- Current: Katharina Kann, Jordan Boyd-Graber

- Past: Michael J. Paul

5th year Computer Science PhD student

University of Tokyo, Japan

Sep. 2014

M.S. in Information Science and Technology; Advisor: Akiko Aizawa

International Christian University, Japan

Mar. 2012

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

Professional Experience

Instructor, University of Colorado Boulder, USA

Jan. 2021 - Present

- For “CSCI 4622 Machine Learning” class

Teaching Assistant, University of Colorado Boulder, USA

Jan. 2020 - Dec. 2020

- For “CSCI 2270 Data Structures” class

Applied Scientist Intern, Amazon Web Services Inc, USA

May 2020 - Aug. 2020

Research Assistant, University of Colorado Boulder, USA

Aug. 2016 - Dec. 2019

Applied Scientist Intern, Amazon.com, USA

May 2018 - Aug. 2018

Software Engineer, Amazon/A9.com, Japan

Oct. 2014 - Aug. 2016

Software Engineer Intern, Amazon/A9.com, 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

- 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 paper), 2020
- Mozhi Zhang, Yoshinari Fujinuma, Jordan Boyd-Graber: “Exploiting Cross-Lingual Subword Similarities in Low-Resource Document Classification”, AAAI (long paper), 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 paper), 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 paper), 2017
- Yoshinari Fujinuma, Hikaru Yokono, Pascual Martínez-Gómez, Akiko Aizawa: “Distant-supervised Language Model for Detecting Emotional Upsurge on Twitter”, PACLIC (long paper), 2015

*denotes equal contribution

Current Projects	Cross-lingual Transfer from Multiple Languages	2020 - present
	<ul style="list-style-type: none"> • Investigating better cross-lingual transfer from multiple source languages on token-level using language embeddings. 	
	Joint Estimation of Word and Document Readability	2019 - present
	<ul style="list-style-type: none"> • Proposing to use graph convolutional network to exploit recursive relationship between word and document difficulties for readability assessment. 	
Selected Past Projects	Intrinsic Evaluation Measure for Cross-Lingual Embeddings	2017 - 2019
	<ul style="list-style-type: none"> • Developed a graph-based intrinsic measure to evaluate the quality of cross-lingual word embeddings. • Paper link: https://www.aclweb.org/anthology/P19-1489 	
	Finite State Transducer (FST) for Kuromoji	2015
	<ul style="list-style-type: none"> • Replaced a double-array trie to an FST to build a dictionary for Kuromoji, a java-based Japanese tokenizer used in Lucene, Solr, and Elastic Search. • Code available at https://github.com/atilika/fst 	
Academic Service	<ul style="list-style-type: none"> • Program Committee: <ul style="list-style-type: none"> – 2021 NAACL – 2020 EMNLP, AACL-IJCNLP SRW, ACL-SRW, W-NUT – 2019 Workshop on Noisy User-Generated Text (W-NUT) • Secondary Reviewer: <ul style="list-style-type: none"> – 2019 ACL – 2017 EMNLP, WWW 	
Academic Honors	• Travel Grant (CU Boulder)	2017
	• Dean’s Fellowship (CU Boulder)	2016
	• Best Bachelor thesis in CS and Math (Gödel Foundation Prize)	2012
Computer and Language Skills	<u>Languages:</u>	Proficient: Python; Intermediate: C++, Java
	<u>Software:</u>	PyTorch, Git, Vim, L ^A T _E X, MySQL
	<u>English:</u>	TOEFL iBT 101 (2015)
	<u>Domain-specific:</u>	machine learning, natural language processing