

Yoshinari Fujinuma

Office Address

Environmental Design 201
315 UCB, 1060 18th St., Boulder, CO 80309

Contact

fujinumay@gmail.com
<http://akkikiki.github.io>

Education

University of Colorado Boulder, USA Aug. 2016 - Present
Computer Science PhD candidate; Advisors: Michael J. Paul, Jordan Boyd-Graber;
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 Pogoyan

Professional Experience

Research Assistant, University of Colorado Boulder, USA Aug. 2016 - Present
• Proposed an intrinsic evaluation measure for cross-lingual embeddings which correlates with multiple cross-lingual downstream tasks.
Applied Scientist Intern, Amazon.com, USA May 2018 - Aug. 2018
• Built a cross-lingual slot tagger using cross-lingual embedding as the feature for LSTM-CRF with language-adversarial auxiliary function (gradient reversal).
Software Engineer, Amazon/A9.com, Japan Oct. 2014 - Aug. 2016
• Built a JA/ZH language detector for search queries with recall > 0.8.
• Built a semi-supervised CRF-based Japanese query label tagger with F1 = 0.91.
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

- Y Fujinuma, J Boyd-Graber, M Paul: “A Resource-Free Evaluation Metric for Cross-Lingual Word Embeddings based on Graph Modularity”, ACL 2019
- D Pruss, Y Fujinuma, A Daughton, M Paul, B Arnot, D Szafir, J Boyd-Graber: “Zika discourse in the Americas: A multilingual topic analysis of Twitter”, PLOS ONE 2019
- Y Fujinuma, A Grissom II: “Substring Frequency Features for Segmentation of Japanese Katakana Words with Unlabeled Corpora”, IJCNLP 2017
- Y Fujinuma, H Yokono, P Martínez-Gómez, A Aizawa: “Distant-supervised Language Model for Detecting Emotional Upsurge on Twitter”, PACLIC 2015

OSS Project

Finite State Transducer (FST) for Kuromoji 2015
• 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.
• Available at <https://github.com/atilika/fst>

Academic Honors

Dean’s Fellowship (CU Boulder) 2016
Best Bachelor thesis in CS and Math (Gödel Foundation Prize) 2012

Computer and Language Skills

Languages: Proficient: Python, Java
Software: PyTorch, Git, Vim, L^AT_EX, MySQL
English: TOEFL iBT 101 (2015)
Domain-specific: machine learning, natural language processing