

Naoki Otani

Language Technologies Institute
School of Computer Science
Carnegie Mellon University

Email: notani@cs.cmu.edu

EDUCATION

† *Indicates expected*

- 2017–2019 † Master of Science in Language Technologies, Carnegie Mellon University
Advisor: Dr. Eduard Hovy (2018-), Dr. Yiming Yang (2017-2018)
- 2015–2017 Master of Informatics, Kyoto University, Japan
Thesis: “Commonsense Knowledge Acquisition via a GWAP on a Dialogue System”
Advisor: Dr. Sadao Kurohashi
- 2011–2015 Bachelor of Engineering, Kyoto University, Japan
Thesis: “Quality Control of Hierarchical Crowdsourced Classification”
Advisor: Dr. Hisashi Kashima

EXPERIENCE

- Aug.–Oct. 2016 **Research Intern** *Microsoft Research Asia*, Beijing, China
Analyzed microblogs and developed a paraphrasing tool for chat-bots
- Apr.–Aug. 2016 **Teaching Assistant** *Kyoto University*, Kyoto, Japan
Taught C programming to undergraduate students and rated their assignments.
- Feb. 2016 **Research Intern (part time)** *Yahoo Japan Corporation*, Tokyo, Japan
–Feb. 2017 Developed a GWAP (game with a purpose) for acquiring knowledge. The game was published on a widely-used spoken dialogue system, *Yahoo! Voice Assist*.
- Sep. 2015 **Research Intern** *IBM Research - Tokyo*, Tokyo, Japan
Studied acronym disambiguation on biomedical papers using network structures and processed a massive network with millions of nodes.
- Aug. 2015 **Software Engineer** *Kyoto University*, Kyoto, Japan
Organized the Machine Learning Summer School (MLSS) [Predictive Model Challenge](#). Developed a competition platform and prepared datasets for the competition.
- Apr.–Aug. 2015 **Teaching Assistant** *Kyoto University*, Kyoto, Japan
Taught C programming to undergraduate students and rated their assignments.

RESEARCH PROJECTS

2018-	Finet-grained Sentiment Analysis based on Human Needs	<i>CMU</i>
	<ul style="list-style-type: none">• Developing a methodology to identify and explain sentiments in texts.	
2018-	Cross-lingual Word Embedding	<i>CMU</i>
	<ul style="list-style-type: none">• Studying how morphological difference impacts the quality of cross-lingual mapping of embeddings.• Conducted an error analysis of an unsupervised method [2].	
2017-2018	Low-resource Cross-lingual Event Type Detection	<i>CMU</i>
	<ul style="list-style-type: none">• Improved the performance of a neural network-based text classifier by cleansing training data and aggregating predictions.• Contributed to winning a first and second place in two language tasks at the third Low Resource Human Language Technology Evaluation (LoReHLT17).• Conducted a series of experiments to study how to rapidly develop a classifier using distant supervision. [4]	
2017	Translating Commonsense Knowledge Bases	<i>Kyoto University</i>
	<ul style="list-style-type: none">• Developed a method to translate a large English knowledge base into Japanese. [3]• Obtained 150,000 types of Japanese commonsense knowledge, which will be registered to ConceptNet.	
2015–2017	Commonsense Knowledge Acquisition on a Dialogue System	<i>Kyoto University</i>
	<ul style="list-style-type: none">• Developed a game plugin of a spoken dialogue system in C++ for acquiring commonsense knowledge in collaboration with Yahoo Japan Corporation.• Collected 200,000 types of commonsense knowledge from 100,000 players within a year. [7]	
2016	Paraphrasing for Chat-bots	<i>Microsoft Research Asia</i>
	<ul style="list-style-type: none">• Developed a rule-based method for translating casual expressions into formal expressions in Japanese.	
2014–2016	Crowdsourcing for Annotation and Evaluation	<i>Kyoto University</i>
	<ul style="list-style-type: none">• Utilized statistical models in item response theory and extended the model for crowdsourcing.• Proposed an unsupervised statistical model for aggregating many pairwise comparisons of machine translations. [5]• Produced a statistical model for quality control of crowdsourced classification. [1], [6]	

AWARDS

2017–2019	Funai Overseas Scholarship Full amount of tuition for two years. 9 out of 109 applicants were awarded.
2016	Award of Excellence, Internship program at Microsoft Research Asia
2016	Student Scholarship, EMNLP2016

SKILLS

Computing	Python, C/C++, Shell script, R
Language	Japanese (native), English (fluent), Chinese-Mandarin (conversant)

PUBLICATIONS

Journal

- [1] **Naoki Otani**, Yukino Baba, and Hisashi Kashima. 2016. Quality Control of Crowdsourced Classification Using Hierarchical Class Structures. *Expert Systems with Applications (ESWA)*, 58:155–63. [doi](#)

Conference

- [2] Ruo Chen Xu, Yiming Yang, **Naoki Otani**, and Yuexin Wu. 2018. Unsupervised Cross-lingual Transfer of Word Embedding Spaces. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, To Appear, Brussels, Belgium, November. Association for Computational Linguistics.
- [3] **Naoki Otani**, Hirokazu Kiyomaru, Daisuke Kawahara, and Sadao Kurohashi. 2018. Cross-lingual Knowledge Projection Using Machine Translation and Target-side Knowledge Base Completion. In *Proceedings of the 27th International Conference on Computational Linguistics (COLING)*, pages 1508–1520, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics. [PDF](#)
- [4] Aldrian Obaja Muis, **Naoki Otani**, Nidhi Vyas, Ruo Chen Xu, Yiming Yang, Teruko Mitamura and Eduard Hovy. 2018. Low-resource Cross-lingual Event Type Detection in Documents via Distant Supervision with Minimal Effort In *Proceedings of the 27th International Conference on Computational Linguistics (COLING)*, pages 70–82, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics. [PDF](#)
- [5] **Naoki Otani**, Toshiaki Nakazawa, Daisuke Kawahara, and Sadao Kurohashi. 2016. IRT-based Aggregation Model of Crowdsourced Pairwise Comparison for Evaluating Machine Translations. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 511–20, Austin, Texas, USA, November. Association for Computational Linguistics. [PDF](#)
- [6] **Naoki Otani**, Yukino Baba, and Hisashi Kashima. 2015. Quality control for crowdsourced hierarchical classification. In *Proceedings of 2015 IEEE International Conference on Data Mining (ICDM)* (short), pages 937–42, Atlantic City, New Jersey, USA, November. IEEE. [doi](#)

Workshop

- [7] **Naoki Otani**, Daisuke Kawahara, Sadao Kurohashi, Nobuhiro Kaji, and Manabu Sassano. 2016. Large-Scale Acquisition of Commonsense Knowledge via a Quiz Game on a Dialogue System. In *Proceedings of Open Knowledge Base and Question Answering (OKQBA) Workshop*, pages 11–20, Osaka, Japan, December. The COLING 2016 Organizing Committee. (Workshop at COLING). [PDF](#)

REFERENCES

Available on request.