Noriki Nishida, Ph.D.

RIKEN AIP, Nihonbashi 1-chome Mitsui Bldg., 15th floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan

noriki.nishida@riken.jp

https://norikinishida.github.io

https://www.github.com/norikinishida

Research Areas

I am a Research Scientist at RIKEN Center for Advanced Intelligence Project (AIP), Japan. Previously, I received my Ph.D. degree in Information Science and Technology from The University of Tokyo in 2020. I have been working on research and development in natural language processing, with a particular focus on knowledge acquisition, information extraction, discourse analysis, multimodal modeling, large language models, and their applications in domains such as medicine and healthcare. He is interested in automatically extracting and organizing useful knowledge from large document collections to facilitate knowledge discovery.

Employment History

Dec. 2023 – Present	**Research Scientist.** RIKEN AIP.
Apr. 2022 – Mar. 2025	**Part-Time Lecturer.** University of Tsukuba.
Jul. 2020 – Jun. 2021	**Visiting Researcher.** The University of Tokyo.
Apr. 2020 – Nov. 2023	**Postdoctoral Researcher.** RIKEN AIP.
Apr. 2018 – Mar. 2020	**Young Research Fellow (DC2).** The Japan Society for the Promotion of Science.
Apr. 2016 – Mar. 2020	**External Collaborator.** The PLU Group in AIRC.
Nov. 2014 – Aug. 2015	**Part-Time Software Engineer.** Logarhythm Inc.

Education

Mar. 2020 **Ph.D. of Information Science and Technology.**
Department of Creative Informatics, Graduate School of Information Science and Technology, The University of Tokyo.
Thesis title: *Unsupervised Induction of Natural Language Discourse Structure Based on Rhetorical Structure Theory.*
Advisor: Hideki Nakayama.

Mar. 2017 **Master's Degree in Information Science and Technology.**
Department of Creative Informatics, Graduate School of Information Science and Technology, The University of Tokyo.
Thesis title: *Unsupervised Learning of Syntactically Plausible Word Representations by Solving Word Ordering.*
Advisor: Hideki Nakayama.

Mar. 2015 **Bachelor's Degree in Engineering.**
Department of Information and Communication Engineering, Faculty of Engineering, The University of Tokyo.
Thesis title: *Hand Gesture Recognition Using Recurrent Convolutional Neural Networks.*
Advisor: Hitoshi Iba and Yoshihiko Hasegawa.

Teaching History

Oct. 2022 – Mar. 2025	**Data Science.** University of Tsukuba.
Apr. 2022 – Mar. 2025	**Information Literacy.** University of Tsukuba.
Oct. 2017 – Mar. 2018	**Data Science.** The University of Tokyo. (TA)
Oct. 2014 – Mar. 2015	**Basic Programming Exercise.** The University of Tokyo. (TA)

Research Publications

Journal Articles

1. Shibahara, T., Yamada, I., **Nishida**, **N.**, Teranishi, H., Kozaki, K., & Matsumoto, Y. (2024). Weakly Supervised NER using Thesaurus Hierarchical Structure. *Journal of Natural Langauge Processing*, *31*(3), 984–1014.

2. **Nishida**, **N.**, & Matsumoto, Y. (2022). Out-of-Domain Discourse Dependency Parsing via Bootstrapping: An Empirical Analysis on Its Effectiveness and Limitation. *Transactions of the Association for Computational Linguistics*, *10*, 127–144. Presented at ACL 2022. doi:10.1162/tacl_a_00451

3. **Nishida**, **N.**, & Nakayama, H. (2020). Unsupervised Discourse Constituency Parsing Using Viterbi EM. *Transactions of the Association for Computational Linguistics*, *8*, 215–230. Presented at ACL 2020. doi:10.1162/tacl_a_00312

4. Nakayama, H., & **Nishida**, **N.** (2017). Zero-Resource Machine Translation by Multimodal Encoder-Decoder Network with Multimedia Pivot. *Machine Translation*, *31*(1), 49–64. doi:10.1007/s10590-017-9197-z

Conference Proceedings (refereed)

1. Munne, R. F., **Nishida**, **N.**, Liu, S., Tokunaga, N., Yamagata, Y., Kozaki, K., & Matsumoto, Y. (2025). Zero-shot entailment learning for ontology-based biomedical annotation without explicit mentions. In *Proceedings of the 31st international conference on computational linguistics* (pp. 8148–8159). Retrieved from https://aclanthology.org/2025.coling-main.542/

2. **Nishida**, **N.**, Inoue, K., Nakayama, H., Bono, M., & Takanashi, K. (2025). Do multimodal large language models truly see what we point at? investigating indexical, iconic, and symbolic gesture comprehension. In *Proceedings of the 62nd annual meeting of the association for computational linguistics*. To Appear.

3. Chen, Y.-P., **Nishida**, **N.**, Nakayama, H., & Matsumoto, Y. (2024). Recent trends in personalized dialogue generation: A review of datasets, methodologies, and evaluations. In *Proceedings of the 2024 joint international conference on computational linguistics, language resources and evaluation (lrec-coling 2024)* (pp. 13650–13665). Retrieved from https://aclanthology.org/2024.lrec-main.1192/

4. El Khettari, O. *., Nishida, N. *., Liu, S., Munne, R. F., Yamagata, Y., Quiniou, S., … Matsumoto, Y. (2024). Mention-agnostic information extraction for ontological annotation of biomedical articles. In *Proceedings of the 23rd workshop on biomedical natural language processing* (pp. 457–473). Note: El Khettari and Nishida contributed equally to this work. doi:10.18653/v1/2024.bionlp-1.37

5. Wakamiya, S., Pereira, L. K., Reithel, L., Yeh, H., Han, P., Shimizu, S., ... Aramaki, E. (2023). NTCIR-17 MedNLP-SC Social Media Adverse Drug Event Detection: Subtask Overview. In *Proceedings of the 17th ntcir conference on evaluation of information access technologies (ntcir-17)*. Retrieved from https://research.nii.ac.jp/ntcir/workshop/OnlineProceedings17/pdf/ntcir/01-NTCIR17-OV-MEDNLP-WakamiyaS.pdf

6. Kamezawa, H., **Nishida**, **N.**, Shimizu, N., Miyazaki, T., & Nakayama, H. (2022). RNSum: A large-scale dataset for automatic release note generation via commit logs summarization. In S. Muresan, P. Nakov, & A. Villavicencio (Eds.), *Proceedings of the 60th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 8718–8735). doi:10.18653/v1/2022.acl-long.597

7. Takeuchi, J., **Nishida**, **N.**, & Nakayama, H. (2022). Neural networks in a product of hyperbolic spaces. In *Proceedings of the 2022 conference of the north american chapter of the association for computational linguistics: Human language technologies: Student research workshop* (pp. 211–221). doi:10.18653/v1/2022.naacl-srw.27

8. Kamezawa, H., **Nishida**, **N.**, Shimizu, N., Miyazaki, T., & Nakayama, H. (2020). A visually-grounded first-person dialogue dataset with verbal and non-verbal responses. In *Proceedings of the 2020 conference on empirical methods in natural language processing (emnlp)* (pp. 3299–3310). doi:10.18653/v1/2020.emnlp-main.267

9. **Nishida**, **N.**, & Nakayama, H. (2018). Coherence modeling improves implicit discourse relation recognition. In *Proceedings of the 19th annual SIGdial meeting on discourse and dialogue* (pp. 344–349). doi:10.18653/v1/W18-5040

10. **Nishida**, **N.**, & Nakayama, H. (2017). Word ordering as unsupervised learning towards syntactically plausible word representations. In *Proceedings of the eighth international joint conference on natural language processing (volume 1: Long papers)* (pp. 70–79). Retrieved from https://aclanthology.org/I17-1008/

11. Laokulrat, N., Phan, S., **Nishida**, **N.**, Shu, R., Ehara, Y., Okazaki, N., ... Nakayama, H. (2016). Generating video description using sequence-to-sequence model with temporal attention. In *Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers* (pp. 44–52). Retrieved from https://aclanthology.org/C16-1005/

12. **Nishida**, **N.**, & Nakayama, H. (2015). Multimodal Gesture Recognition Using Multi-Stream Recurrent Neural Network. In *Proceedings of the 7th pacific-rim symposium on image and video technology (PSIVT 2015)*. doi:10.1007/978-3-319-29451-3_54

Awards

Dec. 2020 **Outstanding Reviewer.** The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2022).

Mar. 2020 **Young Researcher Encouragement Award.** The Annual Meeting of the Association for Natural Language Processing.

Jul. 2017 **Annual Conference Award.** The Japan Society of Artificial Intelligence (JSAI).

Talks

Mar. 2025 **Choosing to Conduct Research at a National Research Institute in This Era.** The SOKENDAI Career Path Support Seminar.

Talks (continued)

Jun. 2023	**Standard Supervision vs. In-Context Learning in NLP.** RIKEN AIP.
May 2022	**Machine Learning for Knowledge Acquisition from Scholarly Articles.** The 2022 Annual Meeting of the Biometric Society of Japan.
Nov. 2018	**Towards Unsupervised Discourse Parsing.** Artificial Intelligence Research Center (AIRC), Japan.
Mar. 2016	**Deep Learning in Computer Vision.** Kansai Chapter of the Acoustic Society of Japan.
Sep. 2015	**Deep Learning in Video Recognition.** Prometech Simulation Conference 2025.

Research Grants

Jul. 2022 – Mar. 2023	**JST AIP Challenge Program.**
May. 2022 – Mar. 2025	**JSPS KAKENHI Grant-in-Aid for Transformative Research Areas (B)** (Co-Investigator).
Apr. 2021 – Mar. 2024	**JSPS KAKENHI Grant-in-Aid for Early-Career Scientists.**
Apr. 2018 – Mar. 2020	**JSPS KAKENHI Research Fellowship for Young Scientists (DC2).**

Academic Activities

Program Committee	EMNLP 2023 Publicity Chairs; SCIDOCA 2021,2022,2023
Journal Editor	Journal of Natural Language Processing (Apr. 2023 – Present)
Journal Reviewer	Language Resources and Evaluation; ACM Transactions of Asian and Low-Resource Language Information Processing; Journal of Natural Language Processing
Conference Reviewer	ACL Rolling Review; ACL; NAACL; EACL; EMNLP; COLING; AAAI; IJCAI.

Skills

Languages	Strong reading, writing and speaking competencies for English and Japanese.
Programming	Python, Java, C++, SQL, Linux, LaTeX, etc.