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

Profile

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.

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.

Teaching History (continued)

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

Preprint

1. Liu, S., **Nishida**, **N.**, Munne, R. F., Tokunaga, N., Yamagata, Y., Kozaki, K., & Matsumoto, Y. (2025). Ma-coir: Leveraging semantic search index and generative models for ontology-driven biomedical concept recognition. arXiv: 2505.12964 [cs.CL]. Retrieved from https://arxiv.org/abs/2505.12964

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

Refereed Conference Proceedings

1. **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 63rd Annual Meeting of the Association for Computational Linguistics (ACL 2025 Main Conference)*. (To Appear).
2. 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 (COLING 2025)* (pp. 8148–8159). Retrieved from https://aclanthology.org/2025.coling-main.542/
3. 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 (BioNLP 2024)* (pp. 457–473). (El Khettari and Nishida contributed equally to this work). doi:10.18653/v1/2024.bionlp-1.37
4. 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/

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 (ACL 2022)* (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 (NAACL-SRW 2022)* (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 2020)* (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 (SIGDIAL 2018)* (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 (IJCNLP 2017)* (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 (COLING 2016)* (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**, EMNLP 2022.
Mar. 2021	**Young Researcher Encouragement Award**, Shota Sugiura (co-authored), The Annual Meeting of the Association for Natural Language Processing.
Mar. 2020	**Young Researcher Encouragement Award**, The Annual Meeting of the Association for Natural Language Processing.
Sep. 2017	**Student Encouragement Award**, Yuki Kobayashi (co-authored), Japan Society for Software Science and Technology.
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.
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 2021–2025
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; COLM; AAAI; IJCAI.

Skills

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