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.

My research centers on natural language processing, with a particular focus on how knowledge can be extracted, organized, retrieved, and used to solve complex problems. Specifically, I develop methods for information extraction, knowledge base organization, information retrieval, and question answering, often targeting specialized domains such as biomedical science. Recently, my work has focused on expanding and refining the knowledge capabilities of large language models.

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.

Research Publications

Journal Articles

1 Takayoshi Shibahara, Ikuya Yamada, **Noriki Nishida**, Hiroki Teranishi, Kouji Kozaki, and Yuji Matsumoto. 2024. "Weakly Supervised NER using Thesaurus Hierarchical Structure". *Journal of Natural Langauge Processing*, 31(3), pp. 984–1014.
[paper]

2 **Noriki Nishida** and Yuji Matsumoto. 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, pp. 127–144.

(Presented at ACL 2022)
[paper] [code] [dataset]

3 **Noriki Nishida** and Hideki Nakayama. 2020. "Unsupervised Discourse Constituency Parsing Using Viterbi EM". *Transactions of the Association for Computational Linguistics*, 8, pp. 215–230.

(Presented at ACL 2020)
[paper] [code]

4 Hideki Nakayama and **Noriki Nishida**. 2017. "Zero-Resource Machine Translation by Multimodal Encoder-Decoder Network with Multimedia Pivot". *Machine Translation*, 31(1), pp. 49–64.
[paper]

Refereed Conference Proceedings

1 Yi-Pei Chen, **Noriki Nishida**, Hideki Nakayama, and Yuji Matsumoto. 2025. "Post Persona Alignment for Multi-Session Dialogue Generation"In. *Findings of the Association for Computational Linguistics: EMNLP 2025*.
[paper]

2 **Noriki Nishida**, Koji Inoue, Hideki Nakayama, Mayumi Bono, and Katsuya Takanashi. 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 (Volume 2: Short Papers)*, pp. 514–524.
[paper] [code]

3 Shanshan Liu, **Noriki Nishida**, Rumana Ferdous Munne, Narumi Tokunaga, Yuki Yamagata, Kouji Kozaki, and Yuji Matsumoto. 2025. "MA-COIR: Leveraging Semantic Search Index and Generative Models for Ontology-Driven Biomedical Concept Recognition"In. *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 4: Student Research Workshop)*, pp. 596–607.
[paper] [code]

4 Rumana Ferdous Munne, **Noriki Nishida**, Shanshan Liu, Narumi Tokunaga, Yuki Yamagata, Kouji Kozaki, and Yuji Matsumoto. 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.
[paper]

5 Oumaima El Khettari, **Noriki Nishida**, Shanshan Liu, Rumana Ferdous Munne, Yuki Yamagata, Solen Quiniou, Samuel Chaffron, and Yuji Matsumoto. 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)
[paper] [code] [dataset]

6 Yi-Pei Chen, **Noriki Nishida**, Hideki Nakayama, and Yuji Matsumoto. 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.
[paper]

7 Shoko Wakamiya, Lis Kanashiro Pereira, Lisa Reithel, Hui-Syuan Yeh, Peitao Han, Seiji Shimizu, Tomohiro Nishiyama, Gabriel Herman Bernardim Andrade, **Noriki Nishida**, Hiroki Teranishi, Narumi Tokunaga, Philippe Thomas, Roland Roller, Pierre Zweigenbaum, Yuji Matsumoto, Akiko Aizawa, Sebastian Möller, Cyril Grouin, Thomas Lavergne, Aurélie Névéol, Patrick Paroubek, Shuntaro Yada, and Eiji Aramaki. 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)*.
[paper]

8 Hisashi Kamezawa, **Noriki Nishida**, Nobuyuki Shimizu, Takashi Miyazaki, and Hideki Nakayama. 2022. "RNSum: A Large-Scale Dataset for Automatic Release Note Generation via Commit Logs Summarization"In. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (ACL 2022)*, pp. 8718–8735.
[paper] [dataset]

9 Jun Takeuchi, **Noriki Nishida**, and Hideki Nakayama. 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.
[paper]

10 Hisashi Kamezawa, **Noriki Nishida**, Nobuyuki Shimizu, Takashi Miyazaki, and Hideki Nakayama. 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.
[paper] [dataset]

11 **Noriki Nishida** and Hideki Nakayama. 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.
[paper]

12 **Noriki Nishida** and Hideki Nakayama. 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.
[paper] [code]

13 Natsuda Laokulrat, Sang Phan, **Noriki Nishida**, Raphael Shu, Yo Ehara, Naoaki Okazaki, Yusuke Miyao, and Hideki Nakayama. 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.
[paper]

14 **Noriki Nishida** and Hideki Nakayama. 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)*, pp. 682–694.
[paper]

Lectures

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

Talks

Mar. 2025	🔖	**"Choosing to Conduct Research at a National Research Institute in This Era,"** Invited Talk at The SOKENDAI Career Path Support Seminar.
Jun. 2023	🔖	**"Standard Supervision vs. In-Context Learning in NLP,"** Workhop Presentation at RIKEN AIP.
May 2022	🔖	**"Machine Learning for Knowledge Acquisition from Scholarly Articles,"** Invited Talk at The 2022 Annual Meeting of the Biometric Society of Japan.
Nov. 2018	🔖	**"Towards Unsupervised Discourse Parsing,"** Invited Talk at Artificial Intelligence Research Center (AIRC), Japan.
Mar. 2016	🔖	**"Deep Learning in Computer Vision,"** Invited Talk at Kansai Chapter of the Acoustic Society of Japan.
Sep. 2015	🔖	**"Deep Learning in Video Recognition,"** Invited Talk at Prometech Simulation Conference 2015.

Awards

Dec. 2022	🔖	**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).

Research Grants

Jul. 2022 – Mar. 2023	🔖	**"Application of Discourse Analysis for Knowledge Acquisition from Medical and Pharmacological Literature,"** JST AIP Challenge Program, Principal Investigator.
May. 2022 – Mar. 2025	🔖	**"Modality Crossing Based on Latent Structural Understanding in Multimodal Dialogue Translation,"** JSPS KAKENHI Grant-in-Aid for Transformative Research Areas (B), Co-Investigator.

Research Grants (continued)

Apr. 2021 – Mar. 2024   **"Discourse Structure Analysis in Biomedical Literature for Biomedical Knowledge Acquisition,"** JSPS KAKENHI Grant-in-Aid for Early-Career Scientists, Principal Investigator.

Apr. 2018 – Mar. 2020   **"Deep Learning Models for Natural Language Understanding Incorporating Discourse Structure and Knowledge,"** JSPS KAKENHI Research Fellowship for Young Scientists (DC2), Principal Investigator.

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, EMNLP, NAACL, EACL, COLING, COLM, AAAI, IJCAI.

Skills

Research   Natural Language Processing, Knowledge Acquisition, Information Extraction, Document Retrieval, Document Parsing, Knowledge Graph, Ontology, Multimodal, Machine Learning.

Programming   Python (NumPy, spaCy, PyTorch, HuggingFace, Transformers), C++, SQL, Neo4j, Docker, Git, Linux.

Languages   English (professional), Japanese (native).