

Jude Crener Junior Pierre

Ph.D. Candidate (ML & Medical NLP) | Social Computing Laboratory | NAIST, Japan

📍 Ikoma, Nara, Japan | ☎ +81 80 9756 3500 | 📩 judecrenerpierre02@gmail.com |

🌐 <https://jude-p24.github.io/portfolio/>

About Me

Ph.D. candidate at the Nara Institute of Science and Technology (NAIST) specializing in Natural Language Processing (NLP) and machine learning. My research focuses on building and evaluating scalable NLP workflows for medical text, with an emphasis on data preprocessing, data augmentation, and reliable model evaluation in low-resource settings. I am interested in applying these skills in production-oriented environments, working with real-world data and applied machine learning systems.

Education

- **Ph.D. in Information Engineering**, NAIST, Japan - Expected March 2027
 - Focus: Medical NLP and applied machine learning
- **M.Sc. in Information Engineering**, NAIST, Japan – 2024
 - Focus: Machine learning and natural language processing.
- **B.Sc. in Computer Science** (Minor in Economics), ESIH – 2018

Experience

- **Ph.D. in Information Engineering**, NAIST, Japan - Expected March 2027
 - Design and evaluate data-centric NLP pipelines for medical text, including word sense disambiguation, data preprocessing, annotation, and Named Entity Recognition (NER) data augmentation using large language models.
 - Build experimental workflows for training and evaluating ML models in low-resource settings, with a focus on data quality and reproducibility.
 - Collaborate with interdisciplinary research teams and document experimental results.
- **Communication Protocol Design Engineer** (R&D Internship) - Schneider Electric, Osaka, Jan 2023 - Mar 2023

- Designed and implemented an MQTT-based communication protocol for industrial HMI systems using C/C++
 - Worked on reliable data exchange between devices, focusing on performance and system-level constraints
- **Research Assistant** - NAIST, May 2022 - Dec 2022
 - Performed data cleaning, annotation, and quality checks for medical NLP datasets.
 - Assisted in preparing structured datasets for machine learning experiments and analysis.
 - NLP Models training and evaluation
- **Business Risk Analyst** - Digicel Group, 2018 – 2021
 - Conducted data analysis to identify and mitigate financial and operational risks.
 - Performed reconciliations and produced analytical reports to support decision-making.
- **Web Developer** (Freelance)- Digicel Group, 2017 – 2018
 - Developed web applications with frontend and backend components.
 - Built backend services using JavaScript (Node.js) and Python (Flask)-based frameworks

Research Interests

- Medical Natural Language Processing (NLP)
- Data-Centric Machine Learning and Data Augmentation
- Generative AI / Large Language Models (LLMs)
- Information Extraction (Named Entity Recognition, Relation Extraction, Knowledge Graph)
- Word Sense Disambiguation

Publications

- Jude Crener Junior Pierre, Gabriel Herman Bernardim Andrade, Shaowen Peng, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki: [Abbreviation Expansion using only a Single Word as Context](#), In Proceedings of the 20th World Congress on Medical and Health Informatics, MEDINFO 2025, Taipei, Taiwan.

Skills

- Programming: Python, JavaScript, C/C++, SQL
- Data Processing & Analysis: NumPy, Pandas

- Machine Learning: PyTorch, Keras, Scikit-learn
- Data Visualization: Matplotlib, Seaborn
- Backend & web: Node.js, Flask, HTML, CSS
- Research: Experimental design, evaluation, scientific writing

Languages

- French - Native
- English - Professional
- Japanese - Basic