

Albertus Andito

+44 7482 548897 | albertus.andito@gmail.com | linkedin.com/in/albertus-andito | Brighton, United Kingdom

Final-year PhD candidate in Informatics with experience in natural language processing (NLP), large language models (LLMs), machine learning (ML), and large-scale data systems. Skilled in research software development, building ML & data pipelines, and web applications. Seeking to apply this background in industry roles across AI & ML engineering, data science, or applied research.

WORK EXPERIENCE

PhD Researcher

Jan 2022 – Present

University of Sussex

Brighton, United Kingdom

- Conducting doctoral research in applications of NLP with a focus on schema matching and unsupervised assessment of LLM-generated web scraping configurations
- Developed a novel schema matching approach using self-supervised contrastive learning of sentence embeddings with **Hugging Face**; benchmarked against and reproduced multiple state-of-the-art competitors through large-scale experiments on HPC clusters (Slurm, multi-GPU)
- Introduced a new research task: unsupervised assessment of LLM-generated web scraping configurations, and designed embedding-based similarity measures to detect incorrect configurations
- Experimented with and evaluated LLMs (via **Pydantic AI**) in solving the assessment task, comparing their performance to embedding-based methods
- Fully funded doctoral project, supported by University of Sussex and CASM Technology

Research Software Engineer/Researcher (Part-Time)

Oct 2021 – Present

CASM Technology

Brighton, United Kingdom

- Consultancy working with non-profit organisations in using technologies to detect and understand online harms
- Developed and maintained daily social media data collection and analytics workflows using internal Method52 tool
- Built custom web scrapers to extract data from various news websites for downstream topic modelling and analysis
- Designed techniques to identify networks of related social media accounts given seed users, enhancing coverage of harmful activity detection
- Created interactive client dashboards (**Python**, **Streamlit**, **Plotly**) with metrics and visualisations for analysis and reporting
- Worked directly with clients to translate requirements into deployable, usable technical solutions

Doctoral Tutor (Part-Time)

Sep 2021 – Dec 2023

University of Sussex

Brighton, United Kingdom

- Delivered 2-4 lab classes per week for Databases and Advanced NLP modules, teaching both undergraduate and MSc students

Research Assistant

May 2021 – Aug 2021

Concept Analytics Lab, University of Sussex

Brighton, United Kingdom

- Applied NLP techniques to diary texts for the Mass Observation Visualisation project, enabling geographic visualisation of language use
- Developed a **Python** library and web API (**FastAPI**) to extract and query linguistics data (e.g. word and concept frequencies and probabilities, using **Spacy**), powering an interactive visualisation app

Student Demonstrator (Part-Time)

Oct 2020 – May 2021

University of Sussex

Brighton, United Kingdom

- Supported delivery of weekly labs for Introduction to Programming and Further Programming modules

Software Engineer Intern

Jul 2019 – Sep 2020

Oracle

Reading, United Kingdom

- Interned with the AI Apps team, contributing to frameworks that embedded machine learning to Oracle's enterprise products
- Built and maintained an example application and automated end-to-end testing suite, enabling application developers to adopt framework components more easily
- Implemented new features and maintained the framework using **Java**, **Spring Boot**, **Grafana**, and **Prometheus**

VOLUNTARY EXPERIENCE

Head of IT Department

Oct 2023 – Present

Doctrine UK (Doctoral Epistemic of Indonesian in the UK)

Remote, United Kingdom

- Lead a team of IT-focused PhD students in maintaining and expanding the organisation's digital infrastructure for 800+ members
- Introduced Kanban workflows (Trello) to improve scheduling and delivery of knowledge cluster events
- Directed migration of the organisation's website from WordPress to a static site generator (**Astro**), simplifying content structure and enabling richer interconnections between organisational content and entities
- Maintained and developed new features for the member portal web app (**MERN** stack)
- Supervised the development of a dynamic visualisation dashboard for membership census data

ICT Literacy Workshop Facilitator

Jun 2025 - Aug 2025

SEECAC (Southeast and East Asian Centre)

London, United Kingdom

- Facilitated biweekly workshops for 30 Filipino domestic workers, delivering training in essential ICT skills

Deputy Head of IT Department

Jan 2022 – Sep 2023

Doctrine UK (Doctoral Epistemic of Indonesian in the UK)

Remote, United Kingdom

- Developed a membership census portal web app used by 800+ members for registration and data updates
- Built and deployed an online election platform to support organisational voting processes
- Maintained and updated the organisation's WordPress website content

Full Stack Developer

Oct 2020 – Oct 2021

PPI UK (Indonesian Student Association in the UK)

Remote, United Kingdom

- Developed and deployed a web application to record and manage the database of Indonesian students in the UK
- Implemented a Single Sign-On (SSO) API integrating the database with the organisation's discussion forum (Vanilla Forum), streamlining member access

EDUCATION

University of Sussex

Brighton, United Kingdom

PhD in Informatics (Natural Language Processing)

Jan 2022 – Jan 2026 (Expected)

- Research on embedding similarity measures and large language models (LLMs) for schema matching and automated assessment of web scraping configurations

University of Sussex

Brighton, United Kingdom

BSc in Computer Science (with an industrial placement year)

Sep 2017 – May 2021

- Achieved First Class Honours (86%)
- Core modules: Software Engineering, NLP, Advanced NLP, Fundamentals of ML, Databases, Program Analysis
- Industrial placement: 1-year software engineering internship at Oracle, AI Applications team
- Final year project: Using dynamic knowledge graph for fake news early detection (Grade: 87%)

TECHNICAL SKILLS

Programming Languages: Python, Java, JavaScript/TypeScript, SQL

ML & NLP: HuggingFace Transformers, PyTorch, scikit-learn, Spacy, LLM (Pydantic AI)

Web & App Development: MongoDB, Express.js, React, Node.js, Streamlit, FastAPI, Astro

Infrastructure & Tools: Slurm/HPC, DigitalOcean, Git, GitHub Actions