

# Sotaro Takeshita

✉ [sotaro.takeshita@uni-mannheim.de](mailto:sotaro.takeshita@uni-mannheim.de) | 🏠 [sotaro.io/about](https://sotaro.io/about) | 📺 [sobamchan](#)

## Education

### University of Mannheim

PH.D. STUDENT IN NATURAL LANGUAGE PROCESSING

- Text summarization for scholarly documents.
- Supervisor: Prof. Dr. Simone Paolo Ponzetto

Mannheim, Germany

Sep. 2022 - Present

### The National University of Electro-Communication

M.S. IN COMPUTER SCIENCE

- Thesis: Multimodal models for commonsense reasoning.
- Supervisor: Prof. Dr. Yasuhiro Minami

Tokyo, Japan

Apr. 2018 - Mar. 2020

### The National University of Electro-Communication

B.S. IN COMPUTER SCIENCE

- Thesis: Unsupervised text summarization with reinforcement learning.
- Supervisor: Prof. Dr. Yasuhiro Minami

Tokyo, Japan

Apr. 2013 - Mar. 2018

## Experience

### VADIS (VARIABLE Detection, Interlinking and Summarization)

RESEARCHER

- A project funded by German Research Foundation (DFG)
- Development of a summarization system for social science research articles.

Germany

2022 - Present

### Build It Inc.

DATA ANALYST & ML ENGINEER

- Analysis of user behaviours using statistical and machine learning tools
- Development of recommendation SaaS

Tokyo, Japan

2015 - 2022

### The University of Campinas, Faculty of Electrical and Computer Engineering

RESEARCH INTERNSHIP

- Developed Neural Network-based Computer Vision system for self-driving wheelchairs.
- Advisor: Prof. Dr. Eric Rohmer

Campinas, Brazil

Nov. 2018

### NTT Media Intelligence Lab

RESEARCH INTERNSHIP

- Development of an NLP system to automatically generate funny stories in Japanese.
- Advisor: Dr. Ryuichiro Higashinaka

Kanagawa, Japan

Sep. 2018

### Recruit Holdings

DATA SCIENCE INTERNSHIP

- Development of an automatic scoring system for essays.
- Prized with a best intern award

Tokyo, Japan

Aug. 2017

### Sección de Estudios de Posgrado e Investigación de ESIME Culhuacan, Instituto Politecnico Nacional

RESEARCH INTERNSHIP

- Analysis on models of automatic face recognition models.
- Advisor: Prof. Dr. Mariko Nakano Miyatake

Mexico City, Mexico

Feb. 2016 - Apr. 2016

## Publications

### JOURNAL

#### Cross-lingual extreme summarization of scholarly documents

SOTARO TAKESHITA, TOMMASO GREEN, NIKLAS FRIEDRICH, KAI ECKERT AND SIMONE PAOLO PONZETTO

2023

International Journal on Digital Libraries (IJD)

## CONFERENCE

### GenGO: ACL Paper Explorer with Semantic Features

Bangkok, Thailand

SOTARO TAKESHITA, SIMONE PAOLO PONZETTO, KAI ECKERT

2024

The 62nd Annual Meeting of the Association for Computational Linguistics, System Demonstrations

### ROUGE-K: Do Your Summaries Have Keywords?

Mexico City, Mexico

SOTARO TAKESHITA, SIMONE PAOLO PONZETTO, KAI ECKERT

2024

The 13th Joint Conference on Lexical and Computational Semantics colocated at NAACL 2024, (Best Long Paper)

### ACLSum: A New Dataset for Aspect-based Summarization of Scientific Publications

Mexico City, Mexico

SOTARO TAKESHITA, TOMMASO GREEN, INES REINIG, KAI ECKERT AND SIMONE PAOLO PONZETTO

2024

In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)

### ZusammenQA: Data Augmentation with Specialized Models for Cross-lingual Open-retrieval Question Answering System

Seattle, Washington + Online

CHIA-CHIEN HUNG, TOMMASO GREEN, ROBERT LITSCHKO, TORNIKE TSERETELI, SOTARO TAKESHITA, MARCO

2022

BOMBIERI, GORAN GLAVAŠ, SIMONE PAOLO PONZETTO

Workshop on Multilingual Information Access (MIA) at NAACL

### X-SCITLDR: Cross-Lingual Extreme Summarization of Scholarly Documents

Cologne, Germany and Online

SOTARO TAKESHITA, TOMMASO GREEN, NIKLAS FRIEDRICH, KAI ECKERT AND SIMONE PAOLO PONZETTO

2022

ACM/IEEE Joint Conference on Digital Libraries (JCDL)

## Skills

---

**Programming** Python, SQL, Javascript

**Languages** Japanese (Native), English (Professional)