

# David Pohl

Phone: (+81) 70-8428-5072 | [de.david.pohl@gmail.com](mailto:de.david.pohl@gmail.com) | [linkedin.com/in/pohl-david](https://www.linkedin.com/in/pohl-david) | [github.com/david-pohl](https://github.com/david-pohl)

## EDUCATION

<b>Institute of Science Tokyo</b> — M.Sc. Computer Science — GPA: 97/100	Apr 2024 – Present
<ul style="list-style-type: none"><li>Research on tokenization in LLMs supervised by Prof. Naoaki Okazaki</li><li>Published two papers: LLM constrained generation, LLM efficient marginalization</li><li>Ongoing research on cross-tokenizer distillation</li></ul>	<i>Tokyo, JP</i>
<b>Heidelberg University</b> — B.Sc. Computer Science — GPA: 1.0/1.0 (with distinction)	Oct 2020 – Sep 2023
<ul style="list-style-type: none"><li>Research on word sense disambiguation supervised by Prof. Michael Gertz</li></ul>	<i>Heidelberg, DE</i>
<b>North Carolina State University</b> — Visiting Student — GPA: 4.0/4.0	<i>Raleigh, US / 2022</i>
<b>Schiller-Gymnasium Cologne</b> — Abitur (A-level) — GPA: 1.1/1.0 (valedictorian)	<i>Cologne, DE / 2020</i>

## EXPERIENCE

<b>Software Engineering Intern</b> — Woven by Toyota	Jun 2025 – Aug 2025
<ul style="list-style-type: none"><li>Built a Chrome extension that auto-checks and corrects software requirements using LLMs</li><li>Integrated support for checking against internal and industry standards (EARS, ISO, INCOSE)</li><li>Reduced revision time from minutes to seconds</li></ul>	<i>Tokyo, JP</i>
<b>Strategy Consulting Intern</b> — Boston Consulting Group	Aug 2024 – Oct 2024
<ul style="list-style-type: none"><li>Developed strategies for establishing a new risk division at a major German bank</li><li>Led the process redesign of tracking and implementing actions demanded by the European Central Bank</li><li>Consolidated a firm-wide document hierarchy for credit risk and compliance governance</li></ul>	<i>Frankfurt, DE</i>
<b>Research Intern</b> — University of Tokyo	Oct 2023 – Mar 2024
<ul style="list-style-type: none"><li>Led the proposal of a novel news recommendation technique for the Nikkei newspaper</li><li>Implemented and evaluated the method of training an estimator for the relevance decay of news</li></ul>	<i>Tokyo, JP</i>
<b>Software Engineering Intern</b> — Deutsche Telekom	Aug 2021 – Feb 2022
<ul style="list-style-type: none"><li>Automatized DevOps workflows by building pipelines with Ansible, Concourse, and Docker</li></ul>	<i>Dresden, DE</i>

## PUBLICATIONS

<b>Decoding-Free Sampling Strategies for LLM Marginalization</b>
<ul style="list-style-type: none"><li>Proposed a novel automata-based sampling method for efficient marginalization</li><li>Achieved consistent speedups of up to 36× over model-based sampling at similar accuracy</li><li><i>David Pohl</i>, Marco Cognetta, Junyoung Lee, Naoaki Okazaki. (under review, <a href="https://arxiv.org/abs/2510.20208">arxiv.org/abs/2510.20208</a>)</li></ul>
<b>Pitfalls, Subtleties, and Techniques in Automata-Based Subword-Level Constrained Generation</b>
<ul style="list-style-type: none"><li>Addressed key challenges in implementing constrained generation for subword-level LLMs using automata</li><li>Implemented and released a fast, reliable pipeline designed for direct use and easy extension</li><li>Marco Cognetta*, <i>David Pohl</i>*, Junyoung Lee, Naoaki Okazaki. TokShop at ICML 2025 (shared first authorship)</li></ul>
<b>Online DATEing: A Web Interface for Temporal Annotations</b>
<ul style="list-style-type: none"><li>Developed and deployed a web interface and API for easy access to temporal tagging models</li><li>Dennis Aumiller, Satya Almasian, <i>David Pohl</i>, Michael Gertz. SIGIR 2022</li></ul>

## PROJECTS

<b>Timothy - Your Personal LLM-Driven Shop Assistant</b>
<ul style="list-style-type: none"><li>Developed a full-stack application with a modern chat frontend for conversational assistance in e-commerce</li><li>Automated a pipeline of web scraping, data cleaning, and data storage</li></ul>
<b>Zero-Shot Word Sense Disambiguation using Word Embeddings</b>
<ul style="list-style-type: none"><li>Surveyed, designed, and evaluated embedding-based approaches to Word Sense Disambiguation (WSD)</li><li>Revealed that the inherent contextual understanding of BERT serves as a strong baseline in WSD</li></ul>

## SKILLS

<b>Prog. Languages</b>	Python, Kotlin, TypeScript, SQL
<b>Libraries</b>	PyTorch, NumPy, Pandas, Transformers, spaCy, NLTK, React.js, FastAPI
<b>Tools</b>	Docker, NX Monorepo, Git, Ansible, Concourse, OpenAI/Cohere API
<b>Languages</b>	English: Fluent (CEFR C2 / IELTS 8.5) — German: Native — Japanese: Basic

## AWARDS & HONORS

<b>Fellow of the German National Academic Foundation</b> ( <a href="https://studienstiftung.de/en/facts-and-figures">studienstiftung.de/en/facts-and-figures</a> )	2024 – Present
<b>DAAD Full Scholarship for an International Master's</b> ( <a href="https://daad.de/en/the-daad/what-we-do">daad.de/en/the-daad/what-we-do</a> )	2024 – Present
<b>B.Sc. Distinction “Excellent”</b> — Heidelberg University	2023
<b>U.S. Exchange Scholarship</b> — Baden-Wuerttemberg Foundation	2022
<b>Germany Scholarship</b> — Heidelberg University	2021 – 2023