

Will Roberts

Berlin, Germany · wildwilhelm@gmail.com · +49 176 4345 8016 · wkroberts.com · Open to remote

Software engineer with a career spanning Python backend systems, cloud infrastructure, and NLP/ML research — plus a PhD to back it up. Has shipped production systems, managed Kubernetes clusters, processed data at TB scale, and published at ACL and EMNLP. Equally at home designing a greenfield system, extending a complex platform, or diving into a hard algorithmic problem. Bilingual (English, German). Looking for a senior IC or staff engineering role in Berlin or remote.

EXPERIENCE

Software Engineer – Trust Lab, Berlin

Apr 2022 – present

- Designed and built a greenfield autonomous open-source intelligence (OSINT) system as sole architect and lead engineer, powered by generative AI with no human supervision; collected data on 2 million entities in 6 weeks; now in active development as a potential new product line
- Architected a TypeScript browser extension built around a custom domain-specific language for semi-automated structured scraping from walled-garden platforms; deployed across 6+ concurrent projects for 3+ years, covering 24 platforms (Facebook, Instagram, Twitter, LinkedIn, YouTube and others), 54 data schemas, 5 languages, and nearly 1 million structured records collected
- Led engineering as sole engineer on the Code of Practice second study (2024), sampling 18,385 social media posts to measure disinformation during the European elections; both studies were presented to the European Commission; the first study — which found Twitter to contain the most disinformation of all platforms surveyed — was covered by international media
- Took sole end-to-end ownership of two multi-year online monitoring contracts: a longitudinal harm-monitoring project for a government client, collecting 278K+ posts over 72 weeks and transitioning from manual to hybrid GenAI/human labeling; and an online monitoring product surfacing evidence of rental property misuse for a major accommodation platform, delivering 100–300 actionable leads monthly; hired and managed a contractor as one project scaled
- Core engineer on a full-stack internal data labeling platform that has processed approximately 1 million human reviews from 91 reviewers; extended the workflow orchestration layer to support majority voting schemes, disagreement detection and arbitration, and automated translation with language-matched agent routing; co-designed and owns the quality assurance infrastructure, comprising a stratified sampling pipeline and a metrics library for measuring labeling accuracy across all customers
- Served as technical lead on projects for several Western government agencies and on pilots with several medium and large social media platforms; conducted technical screening and system design interviews; mentored junior engineers

DevOps Engineer (freelance) – Realtainment, Berlin

Sept 2019 – Apr 2022

- Served as sole infrastructure architect for a €1M/month e-commerce platform: designed and executed a full migration from expensive dedicated rack servers to Kubernetes on DigitalOcean, reducing infrastructure costs by 10× and replacing manual incident resolution with automated, self-healing deployments
- Containerised a WordPress stack (NFS shared filesystem, MySQL cluster on K8s, later a managed database); subsequently migrated to a Vue.js storefront with Elasticsearch; deployed Redis, Prometheus, and Grafana via Helm; implemented GitLab CI/CD with automated builds and staging deployments
- Assisted with hiring and onboarding junior DevOps engineers; introduced cloud-native development practices across the team; built automated database backup tooling and Docker image lifecycle management

Software Architect (freelance) – Ventura Travel, Berlin

May – Dec 2018

- Designed a microservices migration architecture for a company intranet (internal tour package management and content platform), including a GCP and Kubernetes hosting strategy; ran knowledge-transfer sessions with the development team on cloud-native practices; implemented an authentication service in Node.js

Research Engineer – QTLearn, EU FP7 Project

Jun 2015 – Dec 2016

- Developed a word-embedding method to automatically collect idiomatic expressions across 8 European languages for use in machine translation research; co-authored the project report and presented results to the European Commission in person

University Lecturer – *Humboldt-Universität zu Berlin*

Apr 2012 – Apr 2018

- Taught 350+ hours of undergraduate courses in semantics and linguistics; assessed 500+ student exams and papers

Research Engineer – *DFKI (German Research Centre for AI), Saarbrücken*

Sept 2008 – May 2011

- Built Django web applications for sharing linguistic resources and managing machine translation infrastructure; developed bilingual terminology tools for the European Patent Office

Software Engineer – *Namzak Labs, Kingston, Canada*

Sept 2005 – Jul 2007

- Built a multithreaded C++ directory server and VoIP telephony client; ported the client from Symbian OS to Windows CE, working across resource-constrained embedded hardware

SKILLS

Languages	Python (contributor), TypeScript, C/C++, Shell, Go, Java, Perl
Backend	Django, Flask, PostgreSQL, MySQL, MongoDB, Redis, RabbitMQ, Elasticsearch, Snowflake, Nginx
Frontend	React, Angular, Next.js, Tailwind CSS
DevOps	Kubernetes, Docker, Helm, Terraform, GitLab CI/CD; AWS (ECS, EKS, Lambda, SQS, SNS, Step Functions, Aurora, Glue, Athena)
ML / NLP	scikit-learn, NLTK (contributor), SpaCy, word2vec, CoreNLP, neural networks; TB-scale data pipelines, NumPy, pandas, R; statistics and experimental design; OpenAI and Anthropic APIs
Practices	TDD, CI/CD, system design, technical documentation, high-performance and parallel computing; AI-assisted development (Cursor, Claude Code)

EDUCATION

PhD, Computational Semantics <i>Humboldt-Universität zu Berlin</i> <i>Thesis: Unsupervised lexical acquisition for German verbs</i>	2013 – 2022
MSc, Computational Linguistics <i>Universität des Saarlandes, Saarbrücken</i>	2007 – 2011
BSc (Hons), Cognitive Science <i>Queen's University, Kingston, Canada</i>	2000 – 2007

SELECTED PUBLICATIONS

Egg, Prepens, **Roberts**. Annotation and automatic classification of aspectual categories. *Proceedings of ACL 2019*, pp. 3335–3341.

Roberts and Egg. A comparison of selectional preference models for automatic verb classification. *Proceedings of EMNLP 2014*, pp. 511–522.

Full publication list available on request.