MICHAŁ ADAMCZYK

SOFTWARE DEVELOPER • RESEARCHER

code.madamczyk@gmail.com • https://github.com/mdm-code

Software developer with 4+ years of professional experience in business intelligence and natural language processing. Passionate to innovate. Eager to design and write effective software. Highly skilled in Python and Go coding, SQL, Bash automation and APIs.

EXPERIENCE

Software Developer / Researcher
February 2021 - Present

CLARIN-PL / University of Łódź (Remote / Łódź, Poland)

- Handled programming tasks and maintained RESTful APIs for natural language processing
- Developed and managed PostgreSQL database aggregating language data collected from hundreds of participants
- Worked on an annotation web platform for data collection and presentation
- Set up a production-ready Solr search engine for intent recognition and classification
- Trained and finetuned neural language models for downstream tasks like automatic punctuation and text classification
- Coauthored tasks for PolEval 2021 and 2022 (http://poleval.pl/)

Software Developer / Researcher
November 2019 - February 2021

VoiceLab (Remote, Norway / Gdańsk, Poland)

- Worked on shell scripting and automation for audio processing pipelines
- Written internal Python tooling package for named entity recognition
- Written a number of web scrapers gathering language material at scale
- Improved the statistical English language model used in the company with 10% increase in an extrinsic evaluation
- Developed vastly faster G2P converter for Polish
- Worked with most language modelling toolkits like SRILM and KenLM

PhD Fellow
January 2018 - December 2020

University of Stavanger (Stavanger, Norway)

- Delivered papers on international conferences and symposia
- Developed extensible data processing pipeline for historical language material
- Written up the whole monograph in LaTeX

Program Administrator
July 2016 - December 2017

TAPFIN/ManpowerGroup Solutions (Łódź, Poland)

Copywriter
June 2016 - August 2016

Softbox Group (Łódź, Poland, Remote)

English Language Instructor
Empik Language School (Łódź, Poland)

SKILLS

- Python
- Go
- SQL (PostgreSQL)
- Web frameworks, FastAPI, Flask (Python) and Gorilla, net/HTTP (Go)
- RESTful API, websockets, Transport Layer Protocols
- Solr
- Docker Containerization and IPC
- Bash
- Git
- UNIX/Linux Systems Architecture and Use
- Machine Learning
- Automatic Testing and TDD
- OOP, design patterns and data structures
- Scripting and Automation
- Microservices

EDUCATION

- PhD, Literacy Studies (Linguistics), University of Stavanger, 2018-ongoing
- MA, English (Linguistics), University of Łódź, 2012-2014
- BA, English (Linguistics), University of Łódź, 2009-2012

PROJECTS

- Opinionated snippet manager for Unix (https://github.com/mdm-code/gsnip)
- XDG base directory specification in Go (https://github.com/mdm-code/xdg)
- T5-based historical lemmatizer (https://github.com/mdm-code/manx)
- Zettelkasten as a single POSIX shell script (https://github.com/mdm-code/zet)
- Polish grapheme-to-phoneme converter (https://github.com/mdm-code/prg2p)

KEYWORDS

UNIX • Linux • BSD • Ubuntu • Docker • Containers • Docker-compose • Command Line Interface (CLI) • Terminal User Interface (TUI) • Secure Shell (SSH) • Networking • TCP/IP • Unix Domain Sockets (UDS) • Sockets • HTTP • REST • RPC • JSON-RPC • SQL • PostgreSQL • POSTGRES • Solr • TMUX • Screen • Vim • Vi • Neovim • Git • GitHub • GitLab • Python • Go (golang) • Bash • Zsh • Shell • POSIX • FastAPI • Flask • Pandas • Jupyter • Natural Language Processing (NLP) • Natural Language Understanding (NLU) • UML • YAML • JSON • Markdown • XML • LaTeX • Pandoc • TDD • DDD

LANGUAGES

- Polish (native)
- British English (native-like spoken and written)