

Michał Bień

✉ michal@mbien.pl 🌐 Glorf

EDUCATION

École polytechnique fédérale de Lausanne

Lausanne, CH

MSc in Digital Humanities

September 2020 -

- Studying DH with regard to ML & NLP applications (current GPA: 5.5/6)
- Some of this semester's courses: Applied Data Analysis, Machine Learning
- Extracurricular: ALPS 2021 (Advanced Language Processing School) organized by LIG Grenoble

Poznań University of Technology

Poznań, PL

BSc in Computing

October 2016 - February 2020

- Finished one of the best technical universities in Poland
- Final grade 5/5 (very good) with GPA 4.56/5
- Extracurricular: CLARIN workshops on NLP resources for polish language processing

PUBLICATIONS

RecipeNLG: A cooking recipes dataset for semi-structured text generation

INLG

M Bień, M Gilski, M Maciejewska, W Taisner, D Wiśniewski, A Ławrynowicz

2020

Big Data Analysis and Machine Learning at Scale with Oracle Cloud Infrastructure

CERN

M Bień, R Castellotti, L Canali

2019

PROFESSIONAL EXPERIENCE

EPFL dlab

Lausanne, CH

Student Research Assistant

January 2021 -

- Working as research assistant in Robert West's data science lab
- Participating in ACCOMOJI - an interdisciplinary NLP project - as data scientist

APPLICA.AI

Poznań, PL

Cloud Engineer

September 2019 - August 2020

- Participated in one of the largest commercial NLP research communities in Poland
- Worked closely with deep learning and data science teams to propose improvements of internal workflows

CERN

Geneva, CH

Openlab Summer Student

July 2019 - August 2019

- Participated in CERN Openlab Summer Student Programme as one of 40 selected participants
- Worked in CERN IT-DB-SAS team on deploying ML pipelines at scale with Spark and Kubernetes

Poznań Supercomputing and Networking Center - GÉANT

Poznań, PL

Backend/Cloud Developer

May 2018 - June 2019

- Participated in European research project
- Worked on Kubernetes deployments automation in distributed cluster system

TECHNOLOGIES

- **Languages:** Python, Bash, C
- **Big data:** Hadoop, Kafka, Spark
- **Data science:** NumPy, Pandas, Matplotlib, PyTorch, SkLearn
- **Cloud:** Helm, Kubernetes, Docker, Terraform, Ansible, AWS
- **Tools:** Git, VSCode, JIRA, JetBrains, Jupyter Notebooks
- **OS:** Expert level in Linux (Debian/Ubuntu, Fedora)

INVITED TALKS

- RecipeNLG: A cooking recipes dataset for semi-structured text generation** **Grenoble, FR**
Advanced Language Processing School 2021
January 2021
- o Attending PHD/Advanced MSc level NLP school
 - o Presenting the paper during the school's poster session
- Cooking recipes generator utilizing a deep learning-based language model** **Warsaw, PL**
IEEE Day
January 2021
- o Presenting the thesis in front of polish IEEE section
- RecipeNLG: A cooking recipes dataset for semi-structured text generation** **Dublin, IE**
International Conference on Natural Language Generation
December 2020
- o Selected to present the paper based on my bachelor thesis
- AI at CERN: Big Data and Machine Learning in the Cloud** **Warsaw, PL**
ML in PL 2019 Conference
November 2019
- o Presented contributed talk during main part of the conference
- Big Data Analysis and Machine Learning at Scale with Oracle Cloud Infrastructure** **Geneva, CH**
CMS Machine Learning Forum
August 2019
- o Presented my project in front of CERN committee dedicated to machine learning research
- Big Data Analysis and Machine Learning in the Cloud** **Geneva, CH**
Openlab Summer Student Programme Lightning Talks
August 2019
- o Presented my project in front of Summer Students and CERN IT department

ACHIEVEMENTS

- IEEE Thesis Contest** **1st place**
IEEE
November 2020
- o Won the nation-wide IEEE "Most innovative thesis of 2020" award
- CERN Hackfest** **3rd place**
CERN
June 2019
- o Created CERNLearn: an app that makes it easier for CERNies to learn from each other
- Data Ninja Competition** **4th place**
OLX Group
March 2019
- o Achieved 4th position, while competing with 58 other teams
 - o Learned multilabel text classification solutions (FastText) while preparing to competition

ADDITIONAL

- o **Research interests:**
 - natural language generation: data2text, text2data
 - evaluation of generative language models
 - big textual data analysis
 - computational social media
- o **Languages:** Fluent in English, intermediate in French
- o **Social:** Scout instructor
- o **Hobby:** hiking, climbing, literature