

Michał Bień

✉ michal@mbien.pl 🌐 Glorf

EDUCATION

École polytechnique fédérale de Lausanne

Lausanne, CH

MSc in Digital Humanities

September 2020 - July 2022

- Studied DH with regard to ML & NLP applications (GPA: 5.45/6)
- Thesis: "NLGRank: A discriminative evaluation for natural language generation metrics" with A. Bosselut
- Selected courses: Applied Data Analysis, Machine Learning, Deep 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

- Studied CS with regard to ML & NLP (GPA 4.56/5)
- Thesis: "Cooking recipe generator utilizing deep learning-based language model" with A. Ławrynowicz
- Published my thesis results on INLG2020 conference
- Extracurricular: CLARIN workshops on NLP resources for polish language processing

PUBLICATIONS

On the Context-Free Ambiguity of Emoji

ICWSM

J Czeżochowska, K Gligorić, M Peyrard, Y Mentha, M Bień, A Grutter, A Auer, A Xanthos, R West

2022

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

RESEARCH EXPERIENCE

Allegro.pl

Poznań, PL

Research Engineer - Machine Learning

September 2021 -

- Working on product recommendation in Allegro.pl research department
- Developing large-scale multimodal offer representation models

Epicer

Poznań, PL

Co-founder

December 2021 -

- Serving as a R&D head of a startup dedicated to NLG applications
- Creating scalable solution based on LLMs in Google Cloud

EPFL NLP

Lausanne, CH

Master Thesis Student

February 2022 - July 2022

- Worked on my master thesis in Antoine Bosselut NLP lab
- Analyzed novel methods for NLG evaluation

EPFL dlab

Lausanne, CH

Student Research Assistant

February 2021 - August 2021

- Worked as research assistant in Robert West's data science lab
- Participated in ACCOMOJI - an interdisciplinary NLP project

PUT NLP

Student lab member

Poznań, PL
February 2019 - December 2020

- Worked in Agnieszka Ławrynowicz NLP lab
- Prepared thesis and publication on semi-structured text generation

CERN

Openlab Summer Student

Geneva, CH
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

INVITED TALKS

Machine learning at Allegro

sMallTalk - Mall Group Tech Seminar

Prague, CZ
November 2022

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

Advanced Language Processing School 2021

Grenoble, FR
January 2021

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

International Conference on Natural Language Generation

Dublin, IE
December 2020

AI at CERN: Big Data and Machine Learning in the Cloud

ML in PL 2019 Conference

Warsaw, PL
November 2019

Big Data Analysis and ML at Scale with Oracle Cloud Infrastructure

CMS Machine Learning Forum

Geneva, CH
August 2019

Big Data Analysis and Machine Learning in the Cloud

Openlab Summer Student Programme Lightning Talks

Geneva, CH
August 2019

ACHIEVEMENTS

XLSR-Wav2Vec2 Fine-Tuning Week

Huggingface

1st place
March 2021

- Won first prize for my Polish Wav2Vec2 in model training hackaton organized by Huggingface

IEEE Thesis Contest

IEEE

1st place
November 2020

- Won the nation-wide IEEE "Most innovative thesis of 2020" award

ADDITIONAL

◦ Research interests:

natural language generation: data2text, text2data
evaluation of generative language models
big textual data analysis
computational social media

- **Languages:** Fluent in English, intermediate in French

- **Social:** Scout instructor

- **Hobby:** hiking, climbing, literature