

VILÉM ZOUHAR

vilem.zouhar@gmail.com

EDUCATION

Charles University in Prague 2017/18 - 2019/20

BSc. in Computer Science

Specialization in Computational linguistics

Thesis [Enabling Outbound Machine Translation](#)

Saarland University + Groningen University 2020/21 - exp. 2021/22

MSc. in Language Science and Technology

Double degree programme

Thesis TBD



SELECTED ACADEMIC PROJECTS AND PUBLICATIONS

[Google Scholar](#)

Providing Backtranslation Improves Users Confidence in MT, Not Quality

[NAACL 2021](#)

Vilém Zouhar, Michal Novák, Matúš Žilinc, Ondřej Bojar, Mateo Obregón, Robin L. Hill, Frédéric Blain, Marina Fomicheva, Lucia Specia, Lisa Yankovskaya

Sampling and Filtering of Neural Machine Translation Distillation Data

[NAACL SRW 2021](#)

Vilém Zouhar

Leveraging Neural Machine Translation for Word Alignment

[PBML 116](#)

Vilém Zouhar, Daria Pylypenko

WMT20 Document-Level Markable Error Exploration

[WMT20](#)

Vilém Zouhar, Tereza Vojtěchová, Ondřej Bojar

Extending Ptakopět for MT User Interaction Experiments

[PBML 115](#)

Vilém Zouhar, Michal Novák

Outbound Translation User Interface Ptakopet: A Pilot Study

[LREC 2020](#)

Vilém Zouhar, Ondřej Bojar

A Collection of Machine Learning Exercises

2018/2019

50 pages of ML tasks in R; full version available per request (used as [teaching material](#))

Paper on machine translation evaluation

in review

Survey of knowledge intensive NLP tasks with knowledge bases

in review

Paper on machine translation with eyetracking and EEG

in writing

Paper on dimension reduction for knowledge intensive NLP tasks

in writing

TECHNICAL KNOWLEDGE

LANGUAGE PROFICIENCY

Programming	Python, JS/TS, Rust, R, C/C++
Toolkits	PyTorch, TF (basics), NLTK, Marian NMT
Misc.	Linux, Matplotlib, LaTeX
	Always trying to learn more

Czech	Native
English	C1-C2
German	B2 (<i>in development</i>)
Polish	A1

WORK EXPERIENCE

[Spoken Language Systems](#) group (student research assistant)

2021-present

- University of Saarland (Germany)

- Language modelling with external source of information

[Statistical Natural Language Processing](#) class (tutor)

summer 2021

- University of Saarland (Germany)

- Prepared [class material](#)

Institute of Formal and Applied Linguistics (student research assistant)	2018-present
<ul style="list-style-type: none">- Charles Univeristy (Czech Republic)- Machine translation related projects- Bergamot project (in-browser MT)- Psycholinguistic project consultation- Miscellaneous research tasks	

Previo (intern software dev)	summer 2018
<ul style="list-style-type: none">- Development of multilayer CMS using JS, PHP, Zend and MySQL	

BIM Project (intern software dev)	summer 2017
<ul style="list-style-type: none">- Development of plugins for the ArchiCAD suite with C++/Boost and C#	

Web development	2015-2017
<ul style="list-style-type: none">- Participation in several commercial website projects using the PHP/JS/HTML/CSS stack- Brno bez vizuálního smogu, MrFox, Velab, Intimedcare	

EXTRA-CURRICULAR

Academic senate	2018-2020
<ul style="list-style-type: none">- Member of the Academic Senate at Charles University Faculty of Mathematics and Physics	

Game Jams	2015-2020
<ul style="list-style-type: none">- Various game jams, mostly Ludum Dare- Several games programmed and presented in limited time	

Misc. projects	present
<ul style="list-style-type: none">- Small projects either for convenience, or as a hobby; code hosted at GitHub	

Kasiopea	2017-2019
<ul style="list-style-type: none">- Organization of Kasiopea, an annual coding competition for talented highschool students	