

# Martin Fajčík

Researcher

Božetěchova 2  
Brno, 61200  
Czech Republic  
+420 54114 1421 (office L230.2)  
martin.fajcik@vut.cz  
mfajcik.github.io  
MFajcik  
mfajcik  
Google Scholar



## Education

- 2017–2024 **Ph.D., Brno University of Technology**, Brno, Czechia  
Ph.D. thesis on Factoid Question Answering and Fact-checking
- 2015–2017 **M.Sc. Intelligent Systems**, Brno University of Technology, Brno, Czechia  
M.Sc. thesis focusing on the stochastic optimization applied in automated processor verification
- 2012–2015 **B.Sc. Information Technology**, Brno University of Technology, Brno, Czechia  
B.Sc. thesis focusing on the development of static analysis and quick fix functionality for CodAL architecture description language

## Experience

- 2024–Present **Postdoctoral Researcher**, Brno University of Technology, Brno, Czechia  
Research interests: sequential architectures, LLM benchmarking, retrieval, fact-checking  
Projects: FactDeMice (PI), ELOQUENCE (co-solver)
- 2017–2024 **Research Assistant**, Brno University of Technology, Brno, Czechia  
Research assistant & Ph.D. candidate
- Nov **Research Assistant**, IDIAP Research Institute, Martigny, Switzerland
- 2021–Dec Supervisor: Petr Motlíček
- 2022 Research related to Criteria H2020 Project
- June **Java/Python Developer**, Codasip, Brno, Czechia
- 2015–Oct Development of Eclipse IDE plugins, CodAL language editor, python/bash scripts, and simple javascript front-ends based on doT.js. Design of miscellaneous Java APIs.
- 2017

## Publications

- 2025 Jarolím, A., Fajčík, M., & Makaiová, L. (2025). Can LLMs extract human-like fine-grained evidence for evidence-based fact-checking? *Proceedings of the 19th Workshop on Recent Advances in Slavonic Natural Language Processing (RASLAN 2025)*.
- 2025 Makaiová, L., Fajčík, M., & Jarolím, A. (2025). Examining the Metrics for Document-Level Claim Extraction in Czech and Slovak. *Proceedings of the 19th Workshop on Recent Advances in Slavonic Natural Language Processing (RASLAN 2025)*.
- 2025 Fajčík, M., Docekal, M., Dolezal, J., Ondřej, K., Beneš, K., Kapsa, J., Smrž, P., Polok, A., Hradis, M., Neverilová, Z., & others. (2025). Benczechmark: A Czech-centric multitask and multimetric benchmark for large language models with duel scoring mechanism. *Transactions of the Association for Computational Linguistics*, 13, 1068–1095.
- 2024 Stetina, J., Fajčík, M., Štefánik, M., & Hradis, M. (2024). A Comparative Study of Text Retrieval Models on DaReCzech. *Proceedings of the 18th Workshop on Recent Advances in Slavonic Natural Language Processing (RASLAN 2024)*, pp. 85–100.

- 2023 Fajčík, M., Motlicek, P., & Smrz, P. (2023). Claim-Dissector: An Interpretable Fact-Checking System with Joint Re-ranking and Veracity Prediction. *Findings of the Association for Computational Linguistics: ACL 2023*.
- 2022 Fajcik, M., Singh, M., Zuluaga-Gomez, J., Villatoro-Tello, E., Burdisso, S., Motlicek, P., & Smrz, P. (2022). IDIAPers @ Causal News Corpus 2022: Extracting Cause-Effect-Signal Triplets via Pre-trained Autoregressive Language Model. *Proceedings of the 5th Workshop on CASE* (pp. 70–78).
- 2021 Fajcik, M., Jon, J., & Smrz, P. (2021). Rethinking the Objectives of Extractive Question Answering. *Proceedings of the 3rd Workshop on Machine Reading for Question Answering* (pp. 14–27).
- 2021 Fajcik, M., Docekal, M., Ondrej, K. & Smrz, P. (2021). R2-D2: A Modular Baseline for Open-Domain Question Answering. *Findings of the Association for Computational Linguistics: EMNLP 2021*.
- 2021 Min, S., et al. (2021). NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned. *Proceedings of Machine Learning Research*. PMLR.
- 2020 Fajcik, M., Jon, J., Docekal, M. & Smrz, P. (2020). BUT-FIT at SemEval-2020 Task 5: Automatic Detection of Counterfactual Statements with Deep Pre-trained Language Representation Models. *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 437–444).
- 2019 Fajcik, M., Smrz, P. & Burget, L. (2019). BUT-FIT at SemEval-2019 Task 7: Determining the Rumour Stance with Pre-Trained Deep Bidirectional Transformers. *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 1097–1104).
- 2017 Fajcik, M., Smrz, P. & Zachariasova, M. (2017). Automation of Processor Verification Using Recurrent Neural Networks. *2017 18th International Workshop on MTV* (pp. 15–20). IEEE.

## Projects

- Jan 2025 – **FactDeMice**, TAČR bilingual CZ-TW grant TQ16000028, Principal Investigator  
Present Building the “FactDeMice Toolkit”: an AI-driven system for multilingual fact-checking in news comment sections based on automatic community notes generation—including automatic claim-extraction, evidence-based verification and translation (English, Czech, Mandarin). Aims to support journalists, analysts and organizations in combating misinformation.
- Jan 2024 – **ELOQUENCE (Multilingual Conversational AI)**, EU Horizon Europe grant 101135916  
Present Research & development of multilingual, bias-controlled conversational AI and voice/chat-bot technologies for safety-critical applications (e.g. emergency services, smart-home assistants). Development of explainable, knowledge-grounded and ethically aligned dialogue systems; hybrid LLM / speech-understanding methods; pilot implementations and validation across diverse EU languages.
- May 2023– **semANT – Semantic Document Exploration**, NAKI grant DH23P03OVV060  
July 2025 Improving the possibilities of searching in the full-text representation of digitized historical documents at the level of text meaning and the possibilities of natural navigation between thematically similar documents.
- Jan 2021–Apr 2024 **MASAPI – Multilingual assistant**, TAČR grant FW03010656  
Building the assistant that processes available information (local files, e-mails, indexed websites) in commonly used formats and supported languages, builds a knowledge base, replies to user's queries and looks up support for results in indexed documents.
- Nov 2021–Dec 2022 **Criteria**, EU H2020 grant 101021866  
Comprehensive data-driven Risk and Threat Assessment Methods for the Early and Reliable Identification, Validation, and Analysis of migration-related risks. Working on novel interpretable & efficient fact-checking system.
- Aug 2018–Jan 2020 **ROI Hunter**, Company contract  
Automatic product classification and detection of attributes from product descriptions. These are used to enhance product advertisements.

## Awards & Achievements

- 2025 Bronze Commemorative Medal of FIT BUT for excellent working results and performance of tasks at the institute
- 2024 Dean's Award for Ph.D. thesis
- 2022 Causal triplet extraction at CASE@EMNLP 2022 – **2nd place**
- 2020 EfficientQA competition at NeurIPS2020 (track <6GiB) – **3rd place**
- 2020 Counterfactual Reasoning at SemEval-2020 (task B) – **1st place**
- 2020 Counterfactual Reasoning at SemEval-2020 (task A) – **2nd place**
- 2019 RumourEval at SemEval-2019 – **2nd place**
- 2015 Deans Award for B.Sc. thesis

## Theses Supervision

### Master's Theses

- 2025 Lucia Makaiová – Monolingual Adaptation of Multilingual Language Models through Vocabulary Alignment
- 2025 Jakub Rysánek – Discovering Morphemes and the Structure of Emergent Language Across Tasks on Real-World Data
- 2024 Antonín Jarolím – Document Retrieval with Fine-grained Relevance Cues
- 2023 Jan Kočí – Automated Truth Discovery
- 2023 Daniel Kamenický – Designing a Multilingual Fact-Checking Dataset from Existing Question-Answering Data
- 2021 Michal Slávka – Multilingual Open-Domain Question Answering
- ### Bachelor's Theses
- 2024 Jakub Štetina – Efficient Multi-Vector Document Retrieval with Adaptive Representation Size
- 2021 Pavel Kocurek – Visual Question Answering
- 2020 Ksenia Bolshakova – Stylized Natural Language Generation in Dialogue Systems
- 2020 Jozef Karabelly – Non-Supervised Sentiment Analysis
- 2019 Alžbeta Štajerová – Automated Detection of Hate Speech and Offensive Language
- 2019 Tomáš Daniš – Machine Comprehension Using Commonsense Knowledge
- 2019 Jan Kočí – Recommender System for Web Articles

## Teaching

- 2023–Present Information Retrieval (PV211) at Masaryk University
- ### Master's Courses
- 2019–Present Natural Language Processing (ZPJa)
- 2019–2024 Artificial Intelligence and Machine Learning (SUI)
- ### Bachelor's Courses
- 2017–2021 Practical Aspects of Software Design (IVS)
- 2017–Present Scripting Languages (ISJ)
- ### Summer Schools
- 2019–Present NLP lectures at Summer Schools BISSIT

## Skills

**Communication** My love for speaking before audiences sprouted during summers while at the university (2013–2014), I worked at the Museum of Coins and Medals in Kremnica, Slovakia. I guided local and foreign tourist groups and helped with the organization of special events at the museum.

**Hobbies** Music (drums/harmonica), Scifi/Fantasy Reading, Amateur astronomer, Gardening, Hiking, Running, Swimming, Skateboarding, Skiing, Cycling, Volleyball, Snorkelling