Martin
Fajčík

NLP Researcher

—

Martin Fajčík
Božetěchova 2
Brno, 61200
Czech Republic

Telephone
+420 54114 1421 (office L230.2)

Email
ifajcik@fit.vut.cz

Web
https://www.fit.vut.cz/person/ifajcik/.en
https://mfajcik.github.io/

Google Scholar
https://scholar.google.com/citations?user=KUU5fUUAAAAJ&hl=en&oi=sra

GitHub
https://github.com/MFajcik

Linkedin
https://www.linkedin.com/in/mfajcik/

Education

Brno University of Technology / Ph.D. Candidate
Sept 2017 – , BRNO, CZECHIA
Advisor: Pavel Smrž

Ph.D. thesis focusing on Question Answering
Topics: Extractive QA, Open-Domain QA, Bias in Open-Domain QA, Document Retrieval

Brno University of Technology / M.Sc. Intelligent Systems
Sept 2015 – June 2017, BRNO, CZECHIA

M.Sc. thesis focusing on the stochastic optimization applied in automated processor verification

Brno University of Technology / B.Sc. Information Technology
Sept 2012 – June 2015, BRNO, CZECHIA

B.Sc. thesis focusing on the development of static analysis and quick fix functionality for CodAL architecture description language

Jobs

Brno University of Technology / Research Assistant
Sept 2017 – , BRNO, CZECHIA
Advisor: Pavel Smrž
Research assistant & Ph.D. candidate

IDIAP Research Institute / Research Assistant
Nov 2021 – Dec 2022, MARTIGNY, SWITZERLAND
Supervisor: Petr Motlíček
Research related to Criteria H2020 Project

Codasip / Java/Python Developer
June 2015 – Oct 2017 BRNO, CZECHIA

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.

Publications

Fajcik, M., Motlicek, P., & Smrz, P., 2021.**Claim-Dissector: An Interpretable Fact-Checking System with Joint Re-ranking and Veracity Prediction**. *In Findings of the Association for Computational Linguistics: ACL 2023 2023*

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**. In *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)* (pp. 70–78).

Burdisso, S., Zuluaga-Gomez, J., Villatoro-Tello, E., Fajcik, M., Singh, M., Smrz, P., & Motlicek, P., 2022. **IDIAPers@ Causal News Corpus 2022: Efficient Causal Relation Identification Through a Prompt-based Few-shot Approach**. In *Proceedings of the 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE)* (pp. 61–69).

Fajcik, M., Jon, J., & Smrz, P., 2021. **Rethinking the Objectives of Extractive Question Answering**. In *Proceedings of the 3rd Workshop on Machine Reading for Question Answering* (pp. 14–27).

Fajcik, M., Docekal, M., Ondrej, K. & Smrz, P., 2021. **R2-D2: A Modular Baseline for Open-Domain Question Answering**. *In Findings of the Association for Computational Linguistics: EMNLP 2021* 2021

Fajcik, M., Docekal, M., Ondrej, K. & Smrz, P., 2021. **Pruning the Index Contents for Memory Efficient Open-Domain QA**. *arXiv preprint arXiv:2102.10697*.

Min, S., Boyd-Graber, J., Alberti, C., Chen, D., Choi, E., Collins, M., Guu, K., Hajishirzi, H., Lee, K., Palomaki, J., Raffel, C., Roberts, A., Kwiatkowski, T., Lewis, P., Wu, Y., Küttler, H., Liu, L., Minervini, P., Stenetorp, P., Riedel, S., Yang, S., Seo, M., Izacard, G., Petroni, F., Hosseini, L., De Cao, N., Grave, E., Yamada, I., Shimaoka, S., Suzuki, M., Miyawaki, S., Sato, S., Takahashi, R., Suzuki, J., Fajcik, M., Docekal, M., Ondrej, K., Smrz, P., Cheng, H., Shen, Y., Liu, X., He, P., Chen, W., Gao, J., Oguz, B., Chen, X., Karpukhin, V., Peshterliev, S., Okhonko, D., Schlichtkrull, M., Gupta, S., Mehdad, Y. & Yih, W., 2021. **NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned**. In *Proceedings of Machine Learning Research. PMLR, 2021*.

Fajcik, M., Jon, J., Docekal, M. & Smrz, P., 2020, December. **BUT-FIT at SemEval-2020 Task 5: Automatic Detection of Counterfactual Statements with Deep Pre-trained Language Representation Models**. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 437–444).

Jon, J., Fajcik, M., Docekal, M. & Smrz, P., 2020, December. **BUT-FIT at SemEval-2020 Task 4: Multilingual Commonsense**. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 374–390).

Docekal, M., Fajcik, M., Jon, J. & Smrz, P., 2020, December. **JokeMeter at SemEval-2020 Task 7: Convolutional Humor**. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 843–851).

Fajcik, M., Smrz, P. & Burget, L., 2019, June. **BUT-FIT at SemEval-2019 Task 7: Determining the Rumour Stance with Pre-Trained Deep Bidirectional Transformers**. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 1097–1104).

Fajcik, M., Smrz, P. & Zachariasova, M., 2017, December. **Automation of Processor Verification Using Recurrent Neural Networks**. In *2017 18th International Workshop on Microprocessor and SOC Test and Verification (MTV)* (pp. 15–20). IEEE.

Projects

semANT (NAKI grant DH23P03OVV060) – Semantic Document Exploration
May 2023 - Dec 2027

We improve 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.

MASAPI (TAČR grant FW03010656) – Multilingual assistant for searching, analysing and processing information and decision-making support
Jan 2021 - Apr 2024

Building the assistant that processes available information (local files, e-mails, indexed websites) in commonly used formats (HTML, TXT, PDF, DOCX etc.) and supported languages, builds a knowledge base in these, replies to user's queries using this base and looks up support for results in the indexed documents.

Criteria (H2020 grant 101021866) – Comprehensive data-driven Risk and Threat Assessment Methods for the Early and Reliable Identification, Validation, and Analysis of migration-related risks
Nov 2021 - Dec 2022

Working on the novel interpretable & efficient fact-checking system. The project webpage is here.

ROI Hunter (company contract) – Automatic Extraction of Attributes from Product Feeds
Aug 2018 - Jan 2020

Automatic product classification and detection of attributes from the product description. These are used to enhance product advertisements.

Awards & Achievements

Causal triplet extraction at CASE@EMNLP 2022 – 2nd place 2022
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 2020
RumourEval at SemEval-2019 – 2nd place 2019
Deans Award for B.Sc. thesis 2015

Theses Supervision

Jan Kočí (M.Sc., 2023) **Automated Truth Discovery**
Daniel Kamenický (M.Sc., 2023) **Designing a Multilingual Fact-Checking Dataset from Existing Question-Answering Data**
Michal Slávka (M.Sc., 2021) **Multilingual Open-Domain Question Answering**
Pavel Kocurek (B.Sc., 2021) **Visual Question Answering**
Ksenia Bolshakova (B.Sc., 2020) **Stylized Natural Language Generation in Dialogue Systems**
Jozef Karabelly (B.Sc., 2020) **Non-Supervised Sentiment Analysis**
Alžbeta Štajerová (B.Sc., 2019) **Automated Detection of Hate Speech and Offensive Language**
Tomáš Daniš (B.Sc., 2019) **Machine Comprehension Using Commonsense Knowledge**
Jan Kočí (B.Sc., 2019) **Recommender System for Web Articles**

Teaching

M.Sc. Courses

Natural Language Processing (ZPJa) (2019 –)

Artificial Intelligence and Machine Learning (SUI) (2019 –)

B.Sc. Courses

Practical Aspects of Software Design (IVS) (2017 – 2021)

Scripting Languages (ISJ) (2017 –)

NLP lectures at Summer Schools BISSIT (2019 –)

Skills

Communication & Organization

During the summers at the university (2013–2014), I worked at the Museum of Coins and Medals in my hometown, Kremnica, Slovak Republic. I have guided lots of local and foreign tourist groups and helped with the organization of special events at the museum.

Hobbies

Music – Drums and harmonica

Activities – Hiking, Running, Skating, Cycling, Volleyball, Snorkling, ...