

# Ondrej Skopek

## PERSONAL DATA

---

EMAIL: [oskopek@oskopek.com](mailto:oskopek@oskopek.com) WEBSITE: [oskopek.com](http://oskopek.com)  
ADDRESS: Anna-Heer-Strasse 30 LINKEDIN: [linkedin.com/in/oskopek](https://www.linkedin.com/in/oskopek)  
8057 Zürich, Switzerland GITHUB: [github.com/oskopek](https://github.com/oskopek)

## WORK EXPERIENCE

---

- NOV 2019 – now | **Software Engineer** at GOOGLE, Zürich, Switzerland  
Developing new smart features for the [next-generation Google Assistant](#).
- FEB – JUL 2019 | **Teaching Assistant** at ETH ZÜRICH, Switzerland  
Student teaching assistant for the [Natural Language Understanding](#) course organized by the [Data Analytics Laboratory](#). Responsible for preparing and teaching NLP/Machine Learning tutorials for more than 200 Master's students.
- JUN – SEP 2018 | **Software Engineering Intern** at GOOGLE, Zürich, Switzerland  
Solving large-scale experimental Named Entity Recognition on an unlabeled enterprise dataset. Implemented a data conversion and processing pipeline, a state-of-the-art neural network model in TensorFlow with distributed training. Performed hyperparameter tuning and evaluation of the model.
- JUL – SEP 2017 | **Software Engineering Intern** at GOOGLE, Munich, Germany  
Ported and simplified the Voice Search feature on the New Tab Page of Desktop [Chrome](#) into Chromium's codebase, which helped enhance code quality and long-term maintenance. See [Chromium's repository](#) for all my contributions.
- JUL – SEP 2016 | **Software Engineering Intern** at MICROSOFT, Oslo, Norway  
Developed an engineering tool, which helped the team support upgrades of the Search module in SharePoint, in an effort to migrate to Continuous Delivery.
- JUL – SEP 2015 | **Associate Software Engineer (Intern)** at RED HAT, Brno, Czech Republic  
Added automatic statistical evaluation of [OptaPlanner's](#) Benchmarker results. Enables easier tuning of optimization algorithm parameters on practical combinatorial problems. See [OptaPlanner's repository](#) for all my contributions.

## EDUCATION

---

- 2017 – 2019 | Graduate Degree (MSc) in COMPUTER SCIENCE  
**Department of Computer Science, ETH Zürich**  
THESIS: [Mixed-curvature Variational Autoencoders](#)  
GPA (1 to 6, higher is better, 4 is passing): 5.52
- 2014 – 2017 | Undergraduate Degree (BSc) in COMPUTER SCIENCE  
**Faculty of Mathematics and Physics, Charles University, Prague**  
SPECIALIZATION: *General Computer Science*, focus: *Computational Linguistics*  
THESIS: [Planning for Transportation Problems](#) + [TransportEditor](#)  
GPA (1 to 4, lower is better, 3 is passing): 1.36

## SKILLS

---

Python (5 years)   Java (5 years)   C++ (2 years)   TensorFlow (2 years)   PyTorch (1 year)

## LANGUAGES

---

ENGLISH: Full professional proficiency   GERMAN: Basic working proficiency  
*TOEFL: 120/120 (3. 3. 2017)*   *High-school diploma (B2)*

## RESEARCH INTERESTS

---

Natural Language Understanding, Non-Euclidean Geometry, On-device learning, Generative Models, Reinforcement Learning, Medical Imaging, Computer Vision, CompSci Education

## PUBLICATIONS

---

1. **Mixed-curvature Variational Autoencoders.** Ondrej Skopek, Octavian-Eugen Ganea, Gary Bécigneul. INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS, 2020. [Paper](#).
2. **Adversarial Augmentation for Enhancing Classification of Mammography Images.** Lukas Jendele\*, Ondrej Skopek\*, Anton S. Becker, Ender Konukoglu. ARXIV. 2019. [Preprint](#).
3. **Injecting and removing suspicious features in breast imaging with CycleGAN: A pilot study of automated adversarial attacks using neural networks on small images.** Anton S. Becker, Lukas Jendele\*, Ondrej Skopek\*, Nicole Berger, Soleen Ghaffoor, Magda Marcon, Ender Konukoglu. EUROPEAN JOURNAL OF RADIOLOGY 120. 2019. [Journal paper](#).
4. **TransportEditor — Creating and Visualising Transportation Problems and Plans.** Ondrej Skopek, Roman Barták. System Demonstrations track, INTERNATIONAL CONFERENCE ON AUTOMATED PLANNING AND SCHEDULING (ICAPS). 2017. [Workshop paper](#).

\* indicates equal contributions

## PROJECTS

---

- |                |  |
|----------------|--|
| FEB – DEC 2018 | <b>Generating cancerous features in mammograms using GANs</b><br>Generative model which learns to transform healthy mammography images into cancerous ones and back. Enables smarter class balancing in small datasets. <a href="#">RSNA Press Release</a> . <a href="#">Journal paper</a> + <a href="#">Preprint</a> . <a href="#">GitHub</a> .<br>Worked in a small team supervised by <a href="#">Prof. Ender Konukoglu</a> (ETH Zürich). |
| FEB – JUN 2018 | <b>Story Cloze Task</b> ( <a href="#">Natural Language Understanding</a> , ETH Zürich)<br>Choosing the correct short story ending sentence out of two candidate sentences ( <a href="#">Story Cloze Task</a> ). Training data only contains the correct ending sentences. Achieved close to state of the art results at the time. <a href="#">Paper</a> . <a href="#">GitHub</a> .   |
| FEB – JUN 2018 | <b>Tweet Sentiment Analysis</b> ( <a href="#">Computational Intelligence Lab</a> , ETH Zürich)<br>Sentiment analysis of a large dataset of tweets using weakly supervised learning. The dataset was scraped and labeled automatically based on the presence of positive or negative emoji. We perform an extensive study of different models and provide a comparison. <a href="#">Paper</a> . <a href="#">GitHub</a> .                      |
| FEB – MAY 2018 | <b>Eye Gaze Estimation</b> ( <a href="#">Machine Perception</a> , ETH Zürich)<br>Estimating the 3D gaze angle of where a person is looking from single RGB images from a standard laptop webcam. <a href="#">Paper</a> .   |
| SEP – DEC 2017 | <b>Disease Stage Classification</b> ( <a href="#">Advanced Machine Learning</a> , ETH Zürich)<br>Three smaller projects as part of the Advanced ML course: age regression and degenerative disease stage classification from 3D brain MRI scans; and disease stage classification from electrocardiogram time-series data. <a href="#">GitHub</a> .  |
| SEP – DEC 2017 | <b>MapReduce projects</b> ( <a href="#">Data Mining</a> , ETH Zürich)<br>Four smaller projects implemented in a MapReduce-like environment: locality-sensitive hashing, polynomial kernel SVMs for image classification, clustering using coresets and k-means++, and a recommender system based on the LinUCB algorithm (context-aware multi-armed bandits). <a href="#">GitHub</a> .   |
| NOV – DEC 2016 | <b>Native Language Identification</b> ( <a href="#">Deep Learning</a> , Charles University)<br>Identification of a TOEFL essay author's native language from the English essay text. Achieved $\approx 75\%$ accuracy in the 11 native language classification problem using an end-to-end convolutional neural network (TOEFL11 dataset).   |