

# JAN FILLIES

Krossener str. 27, 10245 Berlin, Germany  
(+49)172-7259957 ◊ jan@fillies.de



## EDUCATION

---

### **Freie Universität Berlin (FU Berlin)**

Doctoral Candidate in Computer Science

Focus: Advancing Hate Speech detection within the field of youth language

*March 2023 (Current)*

### **Technical University Munich (TUM)**

Master of Science in Information Systems (1.8)

Focus: Business Analytics & Optimization

*October 2017 - October 2019*

### **Technical University Munich (TUM)**

Bachelor of Science Information Systems

*October 2014 - October 2017*

### **Evangelisches Gymnasium zum Grauen Kloster**

High school degree: Abitur.

*August 2005 - July 2014*

## WORK EXPERIENCE

---

### **Institute für Angewandte Informatik, Berlin**

*Scientific Researcher - Full Time Since January 2023*

*März 2022 (Current)*

- Technical implementation and internal project responsibility for the European research project FAST LISA ([infai.org/fast-lisa/](http://infai.org/fast-lisa/)). I am developing tools to detect and analyze cases of hate speech with machine learning. The main goal is to develop a model for data collection, capacity building, and community engagement. In doing so, the scientific findings are groundwork for my PhD thesis at the Free University of Berlin. The focus of my research is on the challenges of biases in machine detection of hate speech in textual communication between youth.

### **3pc GmbH, Berlin**

*Data Scientist - Part Time*

*März 2022 (Momentan)*

*9 Month*

- The position is focused on the development and execution of exploratory analyses (Python, Jupyter Notebook) and initial prototypes (Python, Docker). This involved investigating specific customer problems in different sectors (e.g., health and culture) to determine their solvability with machine learning. This is always accompanied by data preparation and data analysis. Statistical methods (e.g. correlation, t-tests) for evaluation and assessment were used as well as a variety of AI methods (e.g. LDA, CNN, LSTM). My responsibilities also included various project management, workshop coordination, and client-centric work.

**3pc GmbH, Berlin***Software development for innovative technologies*

April 2020 (Current)

23 Month

- Programming of individual software solutions as well as the development of standardized products (Python, Docker, Kubernetes) based on current technologies in the field of artificial intelligence (AI), particularly machine learning focused on NLP (SpaCy, Gensim, Keras, TensorFlow). Development of intelligent processes for the machine-aided analysis and refinement of extensive multimedia data archives as well as for the generation of interactive user interfaces for storytelling. Further, implementing exploratory prototypes for quick functional analysis of problems regarding the technical implementation (research, analysis, specification) within research projects such as Qurator and Panqura. Project development, grant proposal drafting, and lead responsibility for internal production and external customer innovation projects.

**Telefónica NEXT, Munich***Masterand - Master's Thesis M.Sc.*

March 2019 (6 Month)

- Improving the localization of telecommunication users using machine learning. The master's thesis constructed three neural networks based on the data set of user specific telecommunication network data. The networks were trained to recognize patterns belonging to certain locations and through that improve the overall user location estimation. A FNN, a CNN and a MNN were used and evaluated. Part of the project was also the inclusion of third-party data. The used technology stack had a wide variety. Mainly, Zeppelin Notebooks on AWS ERM clusters were used in combination with Pyspark and Keras, but Spark, Python and Jupyter Notebooks were also applied. Overall the accuracy of localization was improved by about  $\sim 7.9\%$ .

**Telefónica NEXT, Munich***Data Scientist - Working Student*

September 2018 (6 Month)

- The main objective was supporting the analysts by creating reports based on mobile network data during the whole lifetime of a project from pre-sales to delivery. Data preparation (Python, R, Athena, S3) and cleaning (Python, R, Spark, QGIS), visual presentation (Tableau) and special zoning for public transport analyzes (QGIS) were part of the position. This position also included teamwork, project responsibilities and creative freedom.

**VMware Inc., Munich***Sales Support - Working Student*

January 2017 (12 Month)

- The focus of the position was the overall support of the German sales team. It was initially introduced as a technical insights position but by demand turned into general Account Manager support. It created a close-up view and hands on experience into the daily routine of a large high-tech sales organization. This was sided by the responsibility for small projects with much direct partner contact. Fast understanding, structured working, team spirit and good understanding of Salesforce were daily needed skills.

**NTT Data, Munich***Technical Consultant - Working Student*

January 2016 (6 Month)

- To systematically support the consultants by tasks as offer creation, contract setup and project realization were the key parts of this position. Close up contact to the daily business and deep dive into cases made it a long-lasting experience.

**TECHNICAL STRENGTHS**

Python, PySpark, R, Java, Ocaml

Kears, scikit-learn, SpaCy, Gensim

Docker, Continuous Integration, Kubernetes, AWS (EMR, S3, Athena), Google Firebase

SQL, REST, XML  
QGIS, Tableau, GIT, JIRA

## **LANGUAGES**

---

German, English, Latin, Ancient Greek

## **EXTRA-CURRICULAR**

---

Digital Human Rights, Social startups, Game Theory  
Stay abroad in Rumford, Maine, USA (August 2011 - July 2012)  
Sailing, Running, Skiing

---

Place, Date

---

Jan Fillies