

# Dr. Franziska Horn

[hey@franziskahorn.de](mailto:hey@franziskahorn.de) | [linkedin/franziska-horn](https://www.linkedin.com/in/franziska-horn) | [github/cod3licious](https://github.com/cod3licious) | [franziskahorn.de](https://franziskahorn.de)

A **technical leader** with a strong **product mindset** and a background in **data science** and **software engineering** with **11+ years of experience** building data & software products in both research and application contexts.

## WORK EXPERIENCE

**Freelance Software Architect & Tech Lead** | LEIPZIG (GERMANY) / REMOTE 12/2024 - present

- increase developer productivity and engagement through streamlined internal structures and processes

**Sabbatical** | LEIPZIG (GERMANY) 12/2023 - 11/2024

- worked on personal projects & learned new skills (frontend: Vue.js, mobile development: Kotlin incl. Compose Multiplatform)
- joined the Carbon13 Venture Builder Program (Berlin Cohort 3) with the intention of founding a climate tech startup (Sept-Nov '24)

**Head of Data & Solutions Engineering** | ALCEMY, BERLIN (REMOTE) 10/2022 - 11/2023

- hired and led a team of three data scientists and engineers, responsible for customer integration and data-related customer requests
- created team vision & strategy and prioritized projects & objectives together with the customer success, product, and sales teams
- designed & built web apps empowering the customer success team to conduct recurring customer data analyses independently
- analyzed, documented, and optimized internal processes, e.g., reduced onboarding time for new customers from 50 to 25 days

**Senior Customer-Facing Data Scientist** | ALCEMY, BERLIN (REMOTE) 02/2022 - 10/2022

- analyzed laboratory data from our customers (cement plants), e.g., to identify irregularities in their production processes
- reduced time spent on recurring analyses by implementing configurable, reusable report templates
- simplified data integration process for new customers by planning and implementing refactorings of core product components, minimized required code files per customer to 6 from 17

**Freelance Data Science Solutions Architect** | LEIPZIG (GERMANY) / REMOTE 10/2018 - 01/2022

- strategy & ideation workshops with department heads and product managers to identify potential AI use cases
- design, implementation, and evaluation of data science solutions tailored to my client's needs (using Python)
- 1:1 coaching sessions and multi-day trainings on how (and when) to use machine learning techniques in practice
- clients include BASF (consulting + coaching for ~2 years) and TRUMPF (regular trainings since 2019)

**Postdoc Visiting Scientist** | ML GROUP, TU BERLIN (REMOTE) 05/2020 - 11/2020

- developed continuously evolving word embeddings that account for meaning changes over time (published at ACL 2021)

**Data Science Consultant** | BASF, LUDWIGSHAFEN 09/2017 - 10/2018

- predictive maintenance project in collaboration with TU Berlin: designed, implemented, and evaluated time series analysis models to predict the degradation of catalysts in chemical plants (patented and published in Computers and Chemical Engineering)
- authored the open source Python library [autofeat](#) for automated feature engineering and selection with 500+ stars on GitHub

**Machine Learning Team Lead** | SPECTRM, BERLIN 07/2016 - 06/2017

- established the machine learning team and hired two ML engineers
- implemented a content recommendation API for newspaper articles to promote our clients' content (Python Flask App)
- developed a chatbot "AI" to respond to user messages automatically (using RiveScript)

**Data Scientist** | IDALAB, BERLIN 02/2014 - 06/2016

- advanced analytics consulting projects, ML algorithm development in Python, presentation of results, and project management
- clients included razorfish (NLP backend for automatic content classification) and outfittery (style prediction for curated shopping)

**Student Research Assistant** | ML GROUP, TU BERLIN 08/2012 - 09/2014

- research on text classification, unsupervised learning (word2vec embeddings, dimensionality reduction), and information extraction
- EEG data analysis at the Berlin Brain-Computer Interface Lab: developed and efficiently implemented new algorithms in MatLab

**Research Intern** | MIT (MASSACHUSETTS INSTITUTE OF TECHNOLOGY), CAMBRIDGE, MA 07/2011 - 10/2011

- at the McGovern Institute for Brain Research / Gabrieli Lab; analyzed fMRI data using Python (published in JAMA Psychiatry)

**Student Research Assistant** | FRAUNHOFER INSTITUTE FOR CHEMICAL TECHNOLOGY, PFINZTAL 07/2007 - 12/2009

- worked independently, responsible for collection of infrared spectroscopy data

# EDUCATION

## **Ph.D. (Dr. rer. nat.) Computer Science** | ML GROUP, TU (TECHNICAL UNIVERSITY) BERLIN 04/2015 - 04/2020

- in the machine learning group of Prof. Dr. Klaus-Robert Müller; funded by the Elsa Neumann scholarship from the universities of Berlin
- thesis: SIMILARITY ENCODER - A NEURAL NETWORK ARCHITECTURE FOR LEARNING SIMILARITY PRESERVING EMBEDDINGS: developed a novel NN architecture to map high dimensional data into a low dimensional embedding space, where arbitrary pairwise relations between the data points are preserved as the embedding vectors factorize a given target similarity matrix
- supervised bachelor and master students
- graduated magna cum laude

## **M.Sc. Computer Science** | TU BERLIN 10/2012 - 03/2015

- focus: intelligent systems (machine learning, big data) & computational neuroscience (at the BCCN)
- thesis: KNOWLEDGE EXTRACTION FROM COMPLEX BIOLOGICAL TEXTS: A MACHINE LEARNING APPROACH
- graduated top of my class (1.0 on a scale from 1 (best) to 5)

## **B.Sc. Cognitive Science** | UNIVERSITY OSNABRÜCK 10/2009 - 09/2012

- interdisciplinary study program including courses in neurobiology, computer science, psychology, artificial intelligence, mathematics, computational linguistics, neuroinformatics, and philosophy; taught in English
- thesis: COMPARING AND COMBINING MULTIPLE EEG FEATURES IN MOTOR IMAGERY BCI - A LARGE SCALE STUDY
- graduated with distinction (1.1 on a scale from 1 (best) to 5)

## **Abitur (secondary school)** | FICHTE-GYMNASIUM KARLSRUHE 09/2000 - 06/2009

- final mark: 1.5 (on a scale from 1 (best) to 5); 11th grade as a year abroad in Missouri (USA)

# SKILLS

## **Leadership**

I'm a “get sh\*t done (well)” person, who motivates empowered teams through a strong vision and clear priorities, while striving for operational excellence in an agile environment.

- hired and mentored other team members; gave constructive feedback in regular 1:1s and conducted performance reviews
- gathered requirements from external customers and aligned team objectives & KPIs with internal stakeholders
- managed the product backlog and held sprint planning meetings, while ensuring the tasks present growth opportunities for individuals
- facilitated workshops and architectural decision making processes across multiple teams

**Certificates:** Professional Scrum Master 1 & Facilitation Skills (Scrum.org)

## **Communication**

**Proficient in professional writing and public speaking, enthusiastic about structuring information, clarifying complex concepts with diagrams, data storytelling, and meticulous editing.**

- written two free online books: “A Practitioner’s Guide to Machine Learning” and “Clarity-Driven Development of Scientific Software”
- taught 50+ machine learning courses to various audiences, ranging from department heads to aspiring data scientists

**Tools:**  $\LaTeX$ , HTML/CSS, Affinity Designer, Whimsical, Quarto, AsciiDoc(tor), office applications

**Languages:** German (native), English (fluent), French (basics), Spanish (basics)

## **Software Engineering**

**Programming Languages & Frameworks:** Python (incl. Jupyter, streamlit, FastAPI, sklearn, pytorch, pandas, uv, ruff, mypy, pytest), SQL (mainly PostgreSQL, SQLite), bash, MatLab, R, Vue.js, Kotlin (incl. Compose Multiplatform)

**Tools:** git (GitLab & GitHub, incl. CI/CD pipelines), AWS (incl. S3, CloudWatch, SQS, SNS), Grafana, Argo Workflows, Sentry, Terraform, Docker, Kubernetes

# PUBLICATIONS

## Exploring Word Usage Change with Continuously Evolving Embeddings

Franziska Horn

In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 290–297, Online, August 2021. Association for Computational Linguistics (ACL).

## Forecasting Industrial Aging Processes with Machine Learning Methods

Mihail Bogojeski, Simeon Sauer, Franziska Horn, Klaus-Robert Müller

*Computers and Chemical Engineering*, 144:107123, 2021.

## The autofeat Python Library for Automatic Feature Engineering and Selection

Franziska Horn, Robert Pack, Michael Rieger

*ECML PKDD Workshops 2019*, Springer, Cham, 2020.

## Automating the search for a patent's prior art with a full text similarity search

Lea Helmers\*, Franziska Horn\*, Franziska Biegler, Tim Oppermann, Klaus-Robert Müller

*PLoS ONE*, 14(3):e0212103, 2019.

## Predicting Pairwise Relations with Neural Similarity Encoders

Franziska Horn, Klaus-Robert Müller

*Bulletin of the Polish Academy of Sciences: Technical Sciences*, 66(6):821-830, 2018.

## Context encoders as a simple but powerful extension of word2vec

Franziska Horn

In *Proceedings of the 2nd Workshop on Representation Learning for NLP*, pages 10–14, Vancouver, Canada, August 2017. ACL.

## “What is Relevant in a Text Document?”: An Interpretable Machine Learning Approach

Leila Arras, Franziska Horn, Gregoire Montavon, Klaus-Robert Müller, Wojciech Samek

*PLoS ONE*, 12(8):e0181142, 2017.

## Explaining Predictions of Non-Linear Classifiers in NLP

Leila Arras, Franziska Horn, Gregoire Montavon, Klaus-Robert Müller, Wojciech Samek

In *Proceedings of the 1st Workshop on Representation Learning for NLP*, pages 1–7, Berlin, Germany, August 2016. ACL.

## Robust Artifactual Independent Component Classification for BCI Practitioners

I. Winkler, S. Brandl, F. Horn, E. Waldburger, C. Allefeld, M. Tangermann

*Journal of Neural Engineering*, 11(3):035013, 2014.

## Predicting Treatment Response in Social Anxiety Disorder From Functional Magnetic Resonance Imaging

O. Doehrmann, S. S. Ghosh, F. E. Polli, G. O. Reynolds, F. Horn, A. Keshavan, ... & J. D. Gabrieli

*JAMA Psychiatry*, 70(1):87-97, 2013.

## Increasing the Spectral Signal-To-Noise Ratio of Common Spatial Patterns

Franziska Horn, Sven Dähne

*Proceedings of the Fifth International Brain-Computer Interface Meeting*, 2013.

## Combining Multiple EEG Features in Motor Imagery BCI

Franziska Horn, Johannes Höhne, Sven Dähne, Benjamin Blankertz

*BBCI Workshop - Advances in Neurotechnology, Berlin, Germany*, 2012.

# PREPRINTS

## The DALPHI annotation framework & how its pre-annotations can improve annotator efficiency

Robert Greinacher, Franziska Horn

*arXiv preprint arXiv:1808.05558*, 2018.

## Discovering topics in text datasets by visualizing relevant words

Franziska Horn, Leila Arras, Gregoire Montavon, Klaus-Robert Müller, Wojciech Samek

*arXiv preprint arXiv:1707.06100*, 2017.

## Exploring text datasets by visualizing relevant words

Franziska Horn, Leila Arras, Gregoire Montavon, Klaus-Robert Müller, Wojciech Samek

*arXiv preprint arXiv:1707.05261*, 2017.

## Interactive Exploration and Discovery of Scientific Publications with PubVis

Franziska Horn

*arXiv preprint arXiv:1706.08094*, 2017.