

Dr. rer. nat.

# Christian Staudt

*data scientist · machine learning engineer*

Hamburg, Germany

✉ [mail@clstaudt.me](mailto:mail@clstaudt.me)

<http://clstaudt.me>



data mining · machine learning · artificial intelligence · big data

I am a data scientist with 7 years of experience as a freelancer in diverse industries. My mission is enabling my clients and the teams I work with to implement data-driven innovation, making the connection between use cases, algorithms, and the appropriate tech stack. I am mainly focused on machine learning from prototype to deployment, but my work also regularly involves optimization and data mining.

---

## Focus

### Data-driven Innovation – end-to-end

To make data-driven innovation succeed, I work with my clients in all phases of a data science project – including business case discovery, structuring of requirements, selecting appropriate methods, algorithms and tech stacks, and finally deployment and evaluation in production.

### Full Stack Data Science

*Full stack* refers to the software engineer who is familiar with both the user-facing side and the backend infrastructure of software. Similarly, the full stack data scientist understands the analytics methods as well as user interaction and infrastructure needed to put it all into practice – a fitting description of my typical role in data science projects. I am especially interested in establishing software engineering best practices in the field of data science.

### Pragmatic AI

Artificial Intelligence is currently both overhyped and underrated. We have seen some remarkable progress in AI capabilities, yet many companies struggle to put intelligent systems to work on everyday use cases. Work with me to go beyond the hype and connect business cases to technology – because many valuable, pragmatic cases for self-learning systems are within reach for your enterprise.

---

## Education

- 2012-2016 **Karlsruhe Institute of Technology (KIT)**, *PhD candidate at the Institute of Theoretical Informatics/Parallel Computing Group*  
– **Doktor der Naturwissenschaften** ( $\approx$  PhD) *magna cum laude* in Computer Science, awarded in June 2016 for the algorithm engineering-focused dissertation *Algorithms and Software for the Analysis of Large Complex Networks*
- 2005-2012 **Karlsruhe Institute of Technology (KIT)**, *computer science studies*  
– subjects: algorithm engineering, software engineering, compiler construction, parallel programming, advanced object-oriented programming, physics, sociology  
– **Diplom** ( $\approx$  Master's degree)

---

## Career

- 2016 - **Freelance Data Scientist**  
– contract roles in Data Science - see section Projects & Roles
- 2015 **Visiting Researcher**, *University of Illinois, Chicago / Clemson University*  
– 2 month research visit funded by a grant from the Karlsruhe House of Young Scientists
- 2012 - 2016 **Researcher**, *at the Institute of Theoretical Informatics/Parallel Computing Group (Prof. Henning Meyerhenke), KIT*  
– research focus: algorithm engineering for the analysis of large complex networks
- 2007 - 2010 **Student Research Assistant**, *at the Institute of Theoretical Informatics, research group Algorithmics I (Prof. Dorothea Wagner), KIT*  
– assistance in algorithmics research and scientific software development

---

## Freelance Projects & Roles

- Jan 2023 - **Artificial Intelligence Advisor**, *Limebit*  
**machine learning** **deep learning** **artificial intelligence**  
– project development for AI applications: medical data mining & AI, AI for renewable energy control
- Jun 2021 - Jul 2022 **Data Scientist / Machine Learning Engineer**, *Flaschenpost*  
**simulation** **optimization** **machine learning** **Python** **operations research**  
**pandas** **scikit-learn** **TensorFlow** **PyTorch** **Docker** **streamlit** **Scrum**  
– the planning and steering data science team improved the operations and supply chains, through simulation, forecasting and optimization  
– project lead for innovative solutions based on machine learning and optimization
- Oct 2020 - Jul 2022 **Data Scientist / Data Engineer**, *Henkel*  
**machine learning** **regression** **data engineering** **data quality**  
**Python** **scikit-learn** **pandas** **MLflow**  
– machine learning for sustainability reporting & product life cycle assessment  
– model and data pipeline engineering

Jun 2020 - **Innovation Consultant / Technical Writer, Andre Peters & Robin Peters Technovationen**

artificial intelligence machine learning software engineering

– securing funding of innovative technology projects (software & AI) for small and medium-sized enterprises

Apr - Oct 2020 **AI Consultant/Architect, Chami Consulting / SysDICE GmbH**

artificial intelligence NLP conversational AI graph databases

TensorFlow Rasa SpaCy neo4j

– enabling the adoption of model-based systems engineering with artificial intelligence  
– consulting, solution design and software architecture for the MBSE tool SysDICE

Sep 2017 - **Data Scientist, Point 8 (point-8.de)**

machine learning time series deep learning predictive maintenance simulation

Python scikit-learn TensorFlow

– **research and development** on machine learning in industrial applications, e.g. time series analytics, time series forecasting and anomaly detection  
– for **Deutsche Bahn**: analysis of graph-, geo- and time series- data from rail traffic; applying simulation to predictive analytics  
– for **client in mechanical engineering sector**: data mining and predictive maintenance on sensor data from IoT applications

Nov 2018 - **Data Scientist, denkbares GmbH (denkbares.com) for Fahrzeugwerk KRONE**  
Jun 2019 (www.krone-trailer.com), analytics application

Python big data dashboards analytics application DevOps plot.ly Dash

– design, development and deployment of a data analytics application for insight into customer demand and production process

Feb 2018 - **Data Scientist, PTV Group (ptvgroup.com), machine learning on traffic data**

May 2018

Python Spark big data geodata time series machine learning

MLOps

– applying machine learning to time series and geodata from road traffic analysis  
– from classification model engineering to a deployment strategy in a cloud environment.

Oct 2017 - **Data Scientist, Deutsche Telekom, multiple data science projects**

Apr 2018

Python machine learning predictive analytics process mining big data

Spark

– supporting Deutsche Telekom's think tank dedicated to improving customer service through advanced analytics  
– data mining, reporting and data visualization  
– machine learning for predictive analytics and root cause analysis  
– process mining

Mar 2017 - **Data Scientist, Boehringer Ingelheim**, data-driven measurement of marketing efforts  
Jun 2017  
**Python** **big data**  
– developing a prototype for data-driven measurement of global marketing campaign performance across channels

---

## Training & Teaching

Oct 2022 - **Artificial Intelligence Instructor, ExperTeach**  
**machine learning** **deep learning** **artificial intelligence** **TensorFlow** **PyTorch**  
– custom trainings on artificial intelligence and their industry use cases

Feb - Apr 2021 **Artificial Intelligence Instructor, Wild Code School**  
**AI** **machine learning** **deep learning** **reinforcement learning**  
**Python** **TensorFlow**  
– authorship, teaching & coaching for the artificial intelligence curriculum - computer vision, anomaly detection, sequence models, time series forecasting, reinforcement learning

Apr 2018 - **Data Science Instructor, Point 8 for clients including ExperTeach, Innogy, Bayerischer Rundfunk**  
Dec 2021  
**Python** **big data** **statistics** **machine learning** **deep learning** **AWS**  
– concept and development of a modular data science curriculum for industry trainings  
– taught over 25 courses from beginner to advanced level

## Workshops

2023 **The AI Dojo**, <https://www.ai-dojo.dev>  
2023 **Software Patterns, BMW**  
2022 **Latest Developments in Artificial Intelligence, ExperTeach / T-Systems**  
2021 **Big Data Analysis - End to End, Point 8 / ExperTeach**  
2020 **Deep Learning with TensorFlow / PyTorch, Point 8 / ExperTeach, x3**  
2020 **Big Data Analysis with PySpark, Point 8 / Heise iX Trainings**  
2020 **Software Engineering, Hochschule Fresenius**  
2018 - 2020 **Big Data Analysis with Python and PySpark, Point 8 / ExperTeach, x9**  
2020 **Big Data Analysis with PySpark, at data2day conference**  
2020 **Machine Learning for Time Series Forecasting, at ML Essentials Heidelberg**  
2019 **Data Analysis and Machine Learning with Python, Point 8 / ExperTeach, x6**  
2019 **Advanced Machine Learning, Point 8 / ExperTeach, at Telekom Technical Security Summit, x3**  
2018 **Machine Learning with PySpark, Point 8 / ExperTeach, x2**

---

## Open Source

2017 - 2023 **Author / Instructor, Data Science Learning Paths, Point 8**  
**Python** **data science** **machine learning**  
– Practical data science courses - from basic to intermediate

2021 - 2022 **Lead Developer / Open Source Maintainer**, Tuttle, Prototype Fund

Python

app development

GitHub

- The Prototype Fund supports prototyping innovative open source projects.
- Tuttle, an free software project, is building an application for the time, finance and paperwork management needs of freelancers

2012 - 2016 **Principal Researcher / Developer & Maintainer**, *NetworkKit*, Karlsruhe Institute of Technology

Python

C++

parallel computing

algorithm engineering

graph algorithms

- *NetworkKit* is an open-source toolkit for large-scale network analysis.

---

## Speaking

- 2019 **Machine Learning for Time Series Forecasting**, at *data2day conference Ludwigshafen*
- 2019 **Scaling Data Analysis and Machine Learning with PySpark**, at *buildingIoT conference Cologne*
- 2017 **PyCon.DE Karlsruhe**, *The Python Ecosystem for Data Science: A Guided Tour*
- 2017 **PyData Warsaw**, *The Python Ecosystem for Data Science: A Guided Tour*

---

## Volunteering

- 2021 **Program Committee Member**, *PyData Global*, pydata.org, The conference provides a forum for the international community of users and developers of open source data analysis tools
- 2020 - 2021 **Co-Organizer**, *PyData Hamburg*, pydata.org, Bi-monthly meetups with talks on innovation with open source data tools
- 2020 - **Podcast Producer / Editor**, *Scientists for Future Podcast*, s4f-podcast.de, conversations with researchers about a sustainable future

---

## Bibliography

**academic publications** [DBLP](#)

---

## Languages

**German** first language

**English** bilingual proficiency

**Italian** limited working proficiency

**Spanish** limited working proficiency