

Dr. rer. nat.

Christian L. Staudt

Data Scientist

22309 Hamburg
✉ mail@clstaudt.me
<http://clstaudt.me>



I am a computer scientist specializing on algorithms for data analysis by training, and a generalist by nature. My skill set combines a strong research background with experience in software architecture and development, especially in solutions for the analysis of big data. I offer consulting and development services in the area of data science, where I accompany the client on the entire way from data to decisions.

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* awarded in June 2016 for the 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) awarded 2012

Career

- 2016 - **Independent 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
– developed and maintained *NetworKit* (<http://networkit.itl.kit.edu>), an open-source software package for high-performance network analysis
– advised students for 5 Bachelor's theses and 1 Master's thesis
– peer review for leading conferences and journals
– taught as exercise instructor for undergraduate course on algorithms
- 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

Projects & Roles

Nov 2018 - **Data Scientist**, *denkbare GmbH (denkbare.com)* for *Fahrzeugwerk KRONE*
Apr 2019 (*www.krone-trailer.com*), reference contact: Dr. Joachim Baumeister, Managing Director / Michael Lehmann, Head of Product Management.

Python big data dashboards analytics application

ground-up design, implementation and deployment of a data analytics application for insight into customer demand and production process

Dec 2018 - **Software/Algorithm Engineer**, *Point 8 (point-8.de)* for *Deutsche Bahn*, reference
Mar 2019 contact: Dr. Tobias Brambach, Co-Founder.

Python time series simulation performance tuning

traffic data analysis, simulation and forecasting

Sep 2018 - **R&D Data Scientist**, *Point 8 (point-8.de)*, reference contact: Dr. Florian Kruse, Co-Founder and Managing Director.

Python machine learning time series deep learning TensorFlow

data science R&D: advanced machine learning & time series analytics

Apr 2018 - **Data Science Instructor**, *ExperTeach (experteach.eu)* & *Point 8 (point-8.de)*, reference contact: Dr. Christophe Cauet, Co-Founder.

Python Spark statistics machine learning

creating and conducting a series of data science training workshops: Introduction to Data Analysis with Python, Data Science 101 Big Data Analysis with PySpark, Machine Learning with PySpark

February 2018 - **Data Scientist**, *PTV Group (ptvgroup.com)*, machine learning on geodata, reference
May 2018 contact: Axel Gußmann, Director.

Python Spark big data geodata time series machine learning MLOps

applying machine learning to large geographical datasets - from model engineering to a deployment strategy in a cloud environment.

December 2017 **Data Scientist**, *Point 8 (point-8.de)* for *Festool (festool.com)*, IoT data mining, reference contact: Dr. Tobias Brambach, Co-Founder.

Python time series IoT

Sensor data analysis for IoT applications

October 2017 **Data Scientist**, *Deutsche Telekom*, from data exploration to predictive and explanatory
- April 2018 models, reference contact: Deniz Demirci, Vice President.

Python machine learning predictive analytics process mining big data

Spark

I supported Deutsche Telekom's think tank dedicated to improving customer service through advanced analytics.

March 2017 - **Data Scientist**, *Boehringer Ingelheim*, data-driven measurement of marketing efforts,
June 2017 reference contact: Dr. Philipp Diesinger, Global Chief Data Scientist.

Python Adobe Data Workbench big data

I developed a prototype for data-driven measurement of global marketing campaign performance across channels, relying on an elaborate big data model.

October 2016 **Consultant/Software Architect**, *LAVRIO.solutions for 100 Worte (100worte.de)*, design and implementation of sentiment analysis software, reference contact: Daniel Spitzer, founder.

Python

NLP

The core technology of this startup enables real-time analysis of natural language to derive psychological traits and indicators. I consulted on a scalable software architecture, efficient algorithms and software engineering best practices.

2012 - 2016 **Algorithm Engineer/Software Architect**, *Karlsruhe Institute of Technology*, developing NetworKit, an open-source software package for high-performance network analysis, reference contact: Prof. Henning Meyerhenke, Head of Research Group.

C++

OpenMP

Python

parallel computing

algorithms

During my PhD research I initiated the NetworKit open source project and contributed as algorithm engineer, software architect, lead developer and maintainer. NetworKit provides a tool suite of state-of-the-art high-performance algorithms and data structures for complex network analysis and machine learning on graphs.

Tools of the Trade

Python

matplotlib

scikit-learn

pandas

numpy

scipy

seaborn

Jupyter

Spark

SQL

TensorFlow

KNIME

Java

C++

OpenMP

git

Linux

shell

macOS

Speaker

2017 **PyData Warsaw**, *The Python Ecosystem for Data Science: A Guided Tour*.

2017 **PyCon.DE Karlsruhe**, *The Python Ecosystem for Data Science: A Guided Tour*.

For academic conference presentations, see section Conference Publications

Trainer

Technical trainings for industry clients

2019 **Machine Learning on Time Series**.

2019 **Machine Learning with PySpark**.

2019 **Scaling Data Analysis and Machine Learning with PySpark**, at *buildingIoT conference Cologne*.

2019 **(Big) Data Analysis with Python and PySpark**, x3.

2018 **Introduction to Data Analysis with Python**.

2018 **(Big) Data Analysis with Python and PySpark**.

2018 **Machine Learning with PySpark**.

Volunteering

2016 - 2017 **Co-Organizer**, *Karlsruhe Machine Learning University Group*, <http://ml-ka.de>.

2012 - 2013 **Founding Member**, *Karlsruhe Python User Group*.

2012 - **Proud Member**, *Chaos Computer Club e.V.*

Core Skills

research

computer science

data science

big data

algorithms

parallel computing

I am a computer scientist by training. During my PhD research on graph algorithms for network science – the subfield of data science concerned with complex network data – I have developed an in-depth understanding of algorithmics and data analysis, as well as the methods of science and engineering. Together with my collaborators I have published research in leading journals and presented at international conferences.

software development

software engineering

software architecture

programming

I have extensive software development practice in multiple languages (Python, C++, Java), with a focus on data analysis software, efficient algorithms and parallel programming. I work with code both at the hands-on and the conceptual level, aiming for well-thought-out architectural decisions and optimized designs.

consulting

agile project management

technical communication

presentation

I enjoy translating between technical knowledge and real-world application domains, which helps me to accompany projects from idea to implementation. Through publishing and presenting my research I have become proficient at communicating technical subjects in a clear way. While leading an open-source project I have gathered practical experience with agile methods for software project management which are useful in building and coordinating productive teams.

Languages

German first language

English bilingual proficiency

Italian limited working proficiency

Spanish elementary proficiency

Bibliography

Journal Articles

- 2017 ***Generating Realistic Scaled Complex Networks***, *Applied Network Science*, with M. Hamann, A. Gutfraind, I. Safro, H. Meyerhenke.
- 2016 ***NetworKit: A Tool Suite for Large-scale Complex Network Analysis***, *Network Science*, with A. Sazonovs, H. Meyerhenke.
- 2016 ***Structure-preserving sparsification methods for social networks***, *Social Network Analysis and Mining*, with M. Hamann, G. Lindner, H. Meyerhenke, D. Wagner.
- 2016 ***An empirical comparison of Big Graph frameworks in the context of network analysis***, *Social Network Analysis and Mining*, with J. Koch, M. Vogel, H. Meyerhenke.
- 2015 ***Engineering Parallel Algorithms for Community Detection in Massive Networks***, *IEEE Transactions on Parallel and Distributed Systems*, with H. Meyerhenke.
- 2013 ***Dynamic Graph Clustering Combining Modularity and Smoothness***, *Journal of Experimental Algorithmics (JEA) Volume 18 Issue 1, April 2013*, with R. Görke, A. Schumm, P. Maillard, D. Wagner.

Conference Publications

- 2016 ***Generating Scaled Replicas of Real-World Complex Networks***, *Complex Networks 2016*, with M. Hamann, I. Safro, A. Gutfraind, H. Meyerhenke.
- 2015 ***Complex Network Analysis on Distributed Systems***, *Foundations and Applications of Big Data Analytics (FAB) 2015*, with J. Koch, M. Vogel, H. Meyerhenke.
- 2015 ***Structure-preserving Sparsification of Social Networks***, *Advances in Social Networks Analysis and Mining (ASONAM) 2015*, with G. Lindner, M. Hamann, H. Meyerhenke.
- 2014 ***Approximating Betweenness Centrality in Large Evolving Networks***, *ALENEX '15*, with E. Bergamini, H. Meyerhenke.
- 2014 ***Detecting Communities Around Seed Nodes in Complex Networks***, *IEEE BigData '14*, with H. Meyerhenke, Y. Marrakchi.
- 2013 ***Engineering High-Performance Community Detection Heuristics for Massive Graphs***, *International Conference on Parallel Processing (ICPP) 2013*, with H. Meyerhenke.
- 2012 ***An Efficient Generator for Clustered Dynamic Random Networks***, *1st Mediterranean Conference on Algorithms*, with A. Schumm, H. Meyerhenke, R. Görke, D. Wagner.
- 2012 ***Static and Dynamic Aspects of Scientific Collaboration Networks***, *International Conference on Advances in Social Network Analysis and Mining (ASONAM) 2012*, with A. Schumm, H. Meyerhenke, R. Görke, D. Wagner.
- 2010 ***Modularity-Driven Clustering of Dynamic Graphs***, *Symposium on Experimental Algorithms (SEA) 2010*, with R. Görke, P. Maillard, D. Wagner.