Caroline Viktoria Weis

Curriculum vitæ

✉ carolinevweis@gmail.com ⬤ github.com/cvweis
⬀ cvweis.github.io 🎓 Google Scholar

My research interests lie in the development of *personalized healthcare* through data analysis and machine learning on medical and biological data. In my PhD, I developed models predicting *antimicrobial resistance* from MALDI-TOF mass spectrometry data, through *kernel methods*, *domain adaptation* and *topological data analysis*.

Keywords: Machine learning, Antimicrobial Resistance Prediction, Personalized Medicine, Topological Data Analysis, Domain Adaptation, Kernel Methods

Skills

- Strong knowledge of `Python` for data analysis (`numpy`, `scipy`, `pandas`, `scikit-learn`) and experience with deep learning frameworks (`PyTorch`, `TensorFlow` and `Keras`).

- Working knowledge of `R` and `MATLAB`. Knowledge of `SQL`.

- Strong knowledge of data visualization tools in `Python`. Working knowledge of `ggplot2` in R, plus `Bokeh` and `Rshiny` for interactive visualization.

- Knowledge of digital typesetting language LATEXand of the `Git` revision control system.

- Strong writing and public speaking skills.

Work experience

2022–present	**AI/ML Engineer** for **Biomedical AI**, Switzerland *GSK.ai*
2017–2021	**Research assistant** in Machine Learning for Healthcare at ETH Zurich, Switzerland *Machine Learning and Computational Biology (MLCB) group*

Education

2017–2021	**Ph.D.** candidate in Machine Learning for Healthcare at ETH Zurich, Switzerland Thesis: *MALDI-TOF MS based clinical antimicrobial resistance prediction using machine learning* *Machine Learning and Computational Biology (MLCB) group* Adviser: Prof. Dr. Karsten Borgwardt
2014–2016	M.Sc. in Biotechnology at ETH Zurich, Switzerland, final grade **5.55** (very good)[1] Thesis: *Assessing the Potential of Feature-pairs in Predicting the Impact of Missense Variants* Advisers: Dr. L. Folkman, Dr. D. Grimm, Prof. Dr. Karsten Borgwardt
2010–2014	B.Sc. in Integrated Life Sciences at FAU Erlangen, Germany, final grade **1.6** (excellent)[2] Thesis: *Analysis of zinc-oxide particle growth under influence of triethylamine by using Small Angle X- Ray Scattering and UV/Vis Spectroscopy* Adviser: Prof. Dr. Tobias Unruh
2001–2010	*Abitur*[3], Alexander-von-Humboldt Gymnasium Schweinfurt[4], Germany, final grade **1.8** (very good)

---

[1] ETH Zurich's Grading System: https://ethz.ch/content/dam/ethz/special-interest/itet/department/Studies/Forms/20160112_Grading_System.pdf
[2] German Grading System
[3] General qualification for university entrance
[4] Secondary school

PUBLICATIONS

In the following list of publications, equal first-author contributions are indicated using a superscript 'dagger' symbol, i.e. †, while joint supervision is denoted by a 'double-dagger', i.e. ‡.

2022 **Caroline Weis**, Aline Cuenod, Bastian Rieck, Olivier Dubuis, Susanne Graf, Claudia Lang, Michael Oberle, Maximilian Brackmann, Kirstine K. Soegaard, Michael Osthoff, Karsten Borgwardt‡, Adrian Egli‡. *Direct antimicrobial resistance prediction from clinical MALDI-TOF mass spectra using machine learning*. Nature Medicine, January 2022.
https://doi.org/10.1038/s41591-021-01619-9

Caroline Weis†, Bastian Rieck†, Sebastian Balzer†, Aline Cuénod, Adrian Egli, and Karsten Borgwardt. *Improved MALDI-TOF MS based antimicrobial resistance prediction through hierarchical stratification*. bioRxiv, April 2022.
https://doi.org/10.1101/2022.04.13.488198

2020 Stefan Groha†, **Caroline Weis**†, Alexander Gusev, Bastian Rieck. *Topological Data Analysis of copy number alterations in cancer*. Learning Meaningful Representations of Life Workshop. Neural Information Processing Systems (NeurIPS) 2020
arXiv:2011.11070

Catherine R. Jutzeler†, Lucie Bourguignon†, **Caroline Weis**, Bobo Tong, Cyrus Wong, Bastian Rieck, Hans Pargger, Sarah Tschudin-Sutter, Adrian Egli, Karsten Borgwardt‡ and Matthias Walter‡. *Comorbidities, clinical signs and symptoms, laboratory findings, imaging features, treatment strategies, and outcomes in adult and pediatric patients with COVID-19: A systematic review and meta-analysis*. Travel Medicine and Infectious Disease, September 2020.
https://doi.org/10.1101/2020.05.20.20103804

Caroline Weis†, Max Horn†, Bastian Rieck†, Aline Cuenod, Adrian Egli, Karsten Borgwardt. *Topological and kernel-based microbial phenotype prediction from MALDI-TOF mass spectra*. OUP Bioinformatics, accepted at ISMB 2020.
https://doi.org/10.1093/bioinformatics/btaa429

Caroline Weis†, Catherine Jutzeler†, Karsten Borgwardt. *Machine learning for microbial identification and antimicrobial susceptibility testing on MALDI-TOF mass spectra: a systematic review*. Clinical Microbiology and Infection, March 2020.
https://doi.org/10.1016/j.cmi.2020.03.014

2017 Jamie R. Wallen, Hao Zhang, **Caroline Weis**, Weidong Cui, Brittni M. Foster, Chris M. W. Ho, Michal Hammel, John A. Tainer, Michael L. Gross, Tom Ellenberger. *Hybrid Methods Reveal Multiple Flexibly Linked DNA Polymerases within the Bacteriophage T7 Replisome*. Structure, 25. 157–166., 2017.
https://doi.org/10.1016/j.str.2016.11.019

Oliver Ratmann, Emma B. Hodcroft, Michael Pickles, Anne Cori, Matthew Hall, Samantha Lycett, Caroline Colijn, Bethany Dearlove, Xavier Didelot, Simon Frost, A.S. Md Mukarram Hossain, Jeffrey B. Joy, Michelle Kendall, Denise Kühnert, Gabriel E. Leventhal, Richard Liang, Giacomo Plazzotta, Art F.Y. Poon, David A. Rasmussen, Tanja Stadler, Erik Volz, **Caroline Weis**, Andrew J. Leigh Brown, Christophe Fraser, on behalf of the PANGEA-HIV Consortium. *Phylogenetic Tools for Generalized HIV-1 Epidemics: Findings from the PANGEA-HIV Methods Comparison*. Molecular Biology and Evolution, Volume 34, Issue 1, January 2017, Pages 185–203, 2017.
https://doi.org/10.1093/molbev/msw217

INVITED TALKS

2023 *AI assistants accelerating biological discovery*. Invited lecture. Data Science Applications in Industry. ETH Zurich 2023.

MALDI-TOF mass spectrometry based clinical antimicrobial resistance prediction using machine learning. Invited talk. FBS Paris. Summer School l'École de l'Inserm Liliane Bettencourt 2023.

2022 *Developing Good Machine Learning Practises for Clinical Prediction Models*. Invited talk. CAS ETH Modern Concepts in Clinical Research. ETH Zurich 2022.

2021   *Case Study: MALDI-TOF Mass Spectrometry for Antimicrobial Resistance Prediction.* Invited talk. Workshop on Geometrical and Topological Representation Learning. International Conference of Machine Learning (ICML) 2021.

Kernel-based microbial phenotype prediction from MALDI-TOF mass spectra. Invited talk. Luxembourg Institute of Science and Technology (LIST) symposium on MALDI-TOF mass spectrometry 2021.

2020   *Topological Data Analysis of Copy Number Alterations in cancer.* Abstract chosen for oral presentation. Learning Meaningful Representations of Life Workshop. Neural Information Processing Systems (NeurIPS) 2020.

Kernel-based antimicrobial resistance prediction from MALDI-TOF mass spectra. Abstract chosen for oral presentation. Machine Learning for Global Health Workshop. International Conference of Machine Learning (ICML) 2020.

Proceedings Presentation: Topological and kernel-based microbial phenotype prediction from MALDI-TOF mass spectra. Invited talk. Intelligent Systems for Molecular Biology (ISMB) 2020.

Posters and other publishing

2020   Mandana Samiei, **Caroline Weis**, Larissa Schiavo, Tatjana Chavdarova, Fariba Yousefi. *Convening during COVID-19: Lessons learnt from organizing virtual workshops in 2020.* Whitepaper.
arXiv:2012.01191

2019   **Caroline Weis**, Max Horn, Bastian Rieck, Karsten Borgwardt. *Sparse representations for MALDI-TOF based microbial classification.* Poster and peer-reviewed abstract. 14th Machine Learning in Computational Biology Meeting (MLCB).

Professional experience

10/2015–
03/2016
Industrial research intern at *Genedata AG*, Basel, Switzerland.

As a Machine Learning research intern in the Screener Business Unit I assessed different algorithms for the task of classifying screening images. Images depicting bacteria treated with new compounds of unknown effect were classified into effect category. I developed a pipeline in R that has the potential to speed-up screening for drugs with a desired effect. Classical approaches such as support vector machine algorithms and t-SNE clustering provided good results, and a Genedata business poster I prepared for SLAS conference 2016 in San Diego set a new record for requested poster downloads at Screener Business Unit. Additionally I implemented a GUI for active learning in MATLAB, allowing for images classified with low probability to be presented to a human expect to be classified.

10/2014–
03/2016
Research assistant at *Control Theory and Systems Biology group*, ETH Zurich, Basel, Switzerland.

As a wetlab research assistant I performed standard tasks – such as Minipreps, PCRs, gel electrophoresis etc. – to assist projects lead by Post-Doctoral researchers.

09/2013–
07/2014
Academic research intern at *Lawrence Berkeley National Laboratory*, Berkeley, USA.

I worked at the SIBYLS beamline of the Physical Biosciences Division. I performed Protein Crystallography and Small Angle X-Ray Scattering experiments and subsequent data analysis. I worked on project, which eventually led to a publication in *Structure* while also providing guidance on data analysis to SIBYLS beamline users.
Supervisor: Dr. Michal Hammel

04/2012–
07/2012
Summer research assistant at *Chair of Crystallography and Structural Physics*, FAU Erlangen, Germany.

In this summer internship I grew zinc oxide crystals, mounted them to a plate and aligned them using a goniometer.

Thesis supervision

2020   Sebastian Balzer. *Improved MALDI-TOF based antimicrobial phenotype prediction through incorporating phylogenetic structure*
Bachelor thesis, ETH Zurich

2019   Lucie Bourguignon. *Mortality prediction using self-reported health records and large scale genomic data*
Master thesis, ETH Zurich

Teaching experience

In each of the following courses, I have served as a teaching assistant. Duties are listed for each course individually.

2018 + 2020   Exercises *Data Mining II*, ETH Zurich
The course duties consisted of programming exercises in Python, creation and grading of bi-weekly exercises, as well as as the development and correction of exam questions.

2012 – 2013   Exercises *Mathematical modeling and statistics for scientists*, Department Mathematics, FAU Erlangen
This computer science tutorial consisted of modeling and statistical analyses programming exercises in R which I had to present and supervise. In addition I supervised and corrected the exam.

2012   Exercises *Structural physics*, Chair of Crystallography and Structural Physics, FAU Erlangen
The course duties consisted of presenting written exercises in crystallography and structural physics, as well as as the development and correction of exam questions.

2011 – 2012   Microscopy course *Biology for physicians*, Animal Physiology, Department of Biology, FAU Erlangen
In this microscopy course I facilitated laboratory exercises about plant biology and basic physiology for first-year physiology students.

Languages

German	native speaker
English	fluent
French	basic knowledge

Service to the community

03/2022– present   *Women in Machine Learning (WiML) board member*
Serving in the Events Committee.

05/2021– 07/2021   *'Computational Approaches to Mental Health' (CA2MH) workshop organizer*, ICML 2021 virtual conference
I co-organized the first installment of the CA2MH workshop at ICML 2021, including reviewing the workshop application draft, communicating with speakers and recruiting programme committee members.

10/2020– 12/2020   *'Topological Data Analysis and Beyond' workshop program committee member*, NeurIPS 2020 virtual conference
I reviewed several workshop paper contributions and shared my previous experience organizing a virtual workshop at ICML 2020 with the workshop organizers.

05/2020– 09/2020   *WiML un-workshop organizer*, ICML 2020 virtual conference
As the *Finance and Sponsorship Chair* I was solely responsible for the representation and communication with industry partners of the WiML organization.

02/2020– 05/2020   *Academic Jury member*, St. Gallen Symposium, St. Gallen, Switzerland
As Academic Jury member I evaluate essays on the topic of *Freedom revisited* and identify the top 100 contributions which receive an all-expenses-covered invitation to the 2020 St. Gallen Symposium. I will also participate in the three-day St. Gallen Symposium, to discuss the most urging problems threatening our freedom at this time and what actions should be taken.[5]

06/2019– 06/2020   *Founding member of peer mentoring group 'Women in Data Science'*
Along with two colleagues I successfully applied for a grant of 5'000 CHF provided by ETH Zurich's Fix-the-Leaky-Pipeline program. Among my tasks were to approach industry partners who serve as mentors for our peer groups, as well as leading peer meetings with fellow Data Science PhD students and Post-Doctoral researchers.

2018   *ETH representative*, ETH Zurich pavillion, World Economic Forum, Davos, Switzerland
I represented ETH Zurich at the universities pavillion in Davos during the 2018 World Economic Forum. Over the course of four days I made the research topic of 'significant pattern mining' accessible to visitors, which ranged from Switzerland's leading research representatives and honorary guests to several high-school groups.

02/2015– 02/2016   *Student and Academic Affairs Commissioner*, Biotechnology Student Association, ETH Zurich, Switzerland
Among my duties as Commissioner was preparing, holding and interpreting the lecture evaluations each semester. I was also the primary contact person for communication between student and professor concerning teaching

---

[5] symposium cancelled after evaluation round due to *SARS-CoV-2* pandemic

matters. I served as the student body contact person to admitted students before their arrival, organized welcome events and helped them settle in at our department at ETH Zurich.

02/2013–
08/2014

Board Member, Biotechnology Student Initiative Erlangen, Germany

At the Biotechnology Student Initiative I organized talks of professors and company representatives, soft skill workshops and company tours, e.g. at Siemens Healthcare in Erlangen.

REFERENCES

PROF. DR. KARSTEN BORGWARDT
Machine Learning and Computational Biology Lab
ETH Zurich
karsten.borgwardt@bsse.ethz.ch

DR. BASTIAN RIECK
Principal Investigator AIDOS Lab, Institute of AI for Health
Helmholtz Zentrum München
bastian.rieck@helmholtz-muenchen.de

PROF. DR. CATHERINE JUTZELER
Biomedical Data Science Lab
ETH Zurich
catherine.jutzeler@hest.ethz.ch

Further references and credentials are available on request.

Last updated on 29th June 2023.