Christian Dallago *November 11, 1992*

✉ christian.dallago.us **in** linkedin.com/in/sacdallago ⚬ github.com/sacdallago ☎ (843) 410-9775
Contact me in: English • Italiano • Deutsch • Español

Highlights

- Project lead of AI&Software solutions for predictive and analytical applications in biology and NLP
- 8+ years experience in software engineering, machine learning, Dev- and MLOps in bioinformatics research
- 7+ years project and team management in multi-stakeholder teams of 2 to 40, remote and on-site
- 5+ years research record in Computer Science, Biology, and Machine Learning with publications (20), scholarships (4), one individually awarded grant (100K EUR), lead organization of conference tracks (1), and invited talks (9). Projects featured on Wired, The Verge, Guardian and others

---

Experience

Senior Solution Architect, Healthcare and Life Science @ NVIDIA |Munich, DE   *Jul'22 — Present*
I helped customers build solutions around BioTransformers and other Machine Learning solutions for Biology.

Senior Data Scientist, Group Data Analytics @ Allianz SE |Munich, DE   *Apr'21 — Jun'22*
Owner of voice products to assist customers during their insurance journeys. I lead a diverse team of engineers and data scientists in building solutions leveraging AI & NLP.

Scientific Researcher, Rostlab @ Technical University of Munich |Munich, DE   *Feb '19 — Present*
I contributed over 15 peer-reviewed articles in two years in the fields of biology, artificial intelligence and computer science to uncover biological mechanisms and how to use AI to improve human health.

Software Engineering Researcher, Sanderlab @ Harvard Medical School |Boston, USA   *Jan'18 — Jan'19*
I lead the development of a scientific product to predict the 3 dimensional structures of proteins, and another product to formalize biological knowledge in computable form to consume via AI algorithms.

Scientific Collaborator, Rostlab @ Technical University of Munich |Munich, DE   *Oct '15 — Nov '17*
I assisted the chair of computational biology in teaching activities for bachelor and master students in computer and data science, with a focus on DevOps and MLOps.

Lead Engineer & Co-founder, Reebes (Reebes.Land) |Bolzano, IT   *Oct '14 — Jun '15*
I co-founded a software company which aimed at facilitating connections between local businesses and customers via a custom built online platform.

Scientific Collaborator, UniBZ (Free University of Bolzano/Bozen) |Bolzano, IT   *Jul '13 — Jul '14*
I contributed to the development of the first fully functional modular raspberry-pi cluster and to other creative uses of raspberry-pi's, for example to re-purpose broken electronics.

System Administrator, Zucchitours (Zucchitours S.N.C.) |Bolzano, IT   *Dec '10 — Dec '15*
I was responsible for the IT infrastructure of a travel agency. I implemented virtualization solutions to improve workflows and enable remote work.

To find out more about what keeps me busy, let's set up a call, or chat me up on LinkedIn.

Education

PhD in Informatics, TUM |Munich, DE   *Oct '17 — Present*
Major: Bioinformatics

MSc in Informatics, TUM |Munich, DE   *Apr '15 — Oct '17*
GPA: 3.92/4, Major: Bioinformatics, Minors: Software Engineering, Computer Vision

BSc in Computer Science and Engineering, UniBZ |Bolzano, IT   *Oct '11 — Oct '14*
GPA: 3.63/4, Major: Software Engineering

Visiting College of Charleston, SC, U.S.A., 2012 (6mo) • Universidad Pontificia de Salamanca, Campus de Madrid, Madrid, Spain, 2013 (6mo) • Nanyang Technological University, Singapore, 2016 (6mo) • Harvard Medical School, MA, U.S.A., 2018 (1y)

Selected Academic Work

Learned embeddings from deep learning to visualize and predict protein sets A python framework that leverages learned representations (embeddings) from trained protein language models (ELMo, BERT, and more) to provide computable representations for downstream protein predictions (e.g., subcellular location and secondary structure). The first framework of its kind: it quickly grew a community and dedicated users. It now features a hosted API and two front-ends, and is in use by large projects with over 3000 monthly users.

Role: Initiator, Owner & Lead Developer.
Languages and frameworks: Python (transformers, pytorch, flask, celery), JavaScript & HTML
Find it online: bioembeddings.com
Article/Proceeding: ISMB&LMRL&Current Protocols, 10.1002/cpz1.113

Modeling aspects of the language of life through transfer-learning protein sequences The first comprehensive attempt at learning protein sequence representations using contextually aware NLP techniques. This transformative piece led to companies like Facebook, Google and Salesforce following in the steps of our initial findings. Subsequent work (see ProtTrans in further publications) was backed by grants amounting to $4.5 Million in compute.

Role: Co-author & Developer
Languages and frameworks: Python (pytorch, AllenNLP), React
Find it online: embed.protein.properties
Article/Proceeding: BMC Bioinformatics, 10.1186/s12859-019-3220-8

Embeddings from deep learning transfer GO annotations beyond homology A first-principles approach to using protein-embeddings from protein LMs (see above) to annotate proteins with unknown Gene Ontology (GO) annotations based on proteins with known GO annotation. This piece of work highlighted the fundamental difference between purely sequence-based methods (like homology) and learned sequence models (like protein LMs), highlighting the larger information content of the latter in its most naive setting.

Role: Co-initiator, Co-author, & Developer.
Languages and frameworks: Python **Find it online**: embed.protein.properties
Article/Proceeding: Scientific Reports, 10.1038/s41598-020-80786-0

Other contributions (e.g., NeurIPS) on **SemanticScholar** and **GitHub**

Grants

Bundesministerium für Bildung und Forschung (BMBF) - 01IS17049 Grant for Junior Scientists, Software Campus 2.0 (TU München, 2020-2022; Total: 100.000 EUR)

Scholarships

Deutschlandstipendium & IBM Scholarship for extraordinary students, 2015/2016; 2016/2017; EUR 7.200
Procincia Autonoma di Bolzano Scholarship for post-graduate students, 2018; 2019l; EUR 18.000

Other academic contributions

Reviewer Nature Biotechnology, OUP Bioinformatics, PLoS Computational Biology, IEEE/ACM transactions on computational biology and bioinformatics
Organization Main organizer for the special session on Representation Learning in Biology@ISMB/ECCB2021

Teaching

Javascript technology Technical Univeristy of Munich - Fall & Spring 2016/2017
Web App Development with the Power of Node.js MOOC via edX & Technical University of Munich
Protein Prediction I & II for Computer Scientists Technical University of Munich - Spring & Fall 2019/2021

Media

'A Song of Ice and Data': students create death prediction software for Game of Thrones The Guarian
From Google Summer of Code to Game of Thrones on the Back of a JavaScript Dragon Google Open Source Blog
This Computer Algorithm Predicted Who Will Die Next on Game of Thrones TIME
Diese "Game of Thrones"-Figuren sind in Lebensgefahr WELT
3D analysis of SARS-CoV-2 reveals clues on virus tactics phys.org
Machine learning assisted structure analysis reveals SARS-CoV-2 virus tactics bionity