

# Emilio DORIGATTI

## Bioinformatics, Data Science, Machine Learning, Statistics

in linkedin.com/in/edorigatti  
@ emilio.dorigatti@mail.com

github.com/e-dorigatti  
e-dorigatti.github.io



### AT A GLANCE

<b>My Vision</b>	Create a fair, sustainable, and thriving society through my expertise in Data Science.
<b>My Values</b>	Reasonable, Reliable, Trustworthy, Competent, Autonomous, Open, Straightforward, Free.
<b>Technologies</b>	Python, R, SQL, PyTorch, TensorFlow, Keras, Scikit-Learn, Pandas, PyMC3, NLTK, OpenCV, Linux, Docker, Spark, AWS EC2/S3, SLURM, PostgreSQL, Elasticsearch, Redis, ...

Throughout my PhD I **worked with experts from multiple research fields**, including physics, biology, bioinformatics, statistics, mathematics, and computer science, **leading the development of novel data analysis methods**, and testing domain-specific hypotheses in order to **generate new insights**.

I also co-developed five university courses and seminars offered throughout the years, **supervising 18 Master's students** during four projects, four theses, and three seminars.

All of this, together with many presentations and my personal blog has taught me how to **communicate complex ideas in simple terms** as well as to understand and **collaborate with people who have widely different skills and capabilities**.

### SELECTED PROFESSIONAL EXPERIENCE

Now July 2019	<b>Ph.D. Student, DEPT. OF STATISTICS, LUDWIG MAXIMILIAN UNIVERSITÄT, Munich, Germany</b> My research focused on quantifying uncertainty in deep learning models, semi structured regression models, as well as self-supervised and positive-unlabeled learning. <ul style="list-style-type: none"><li>➤ <b>Organized</b> and <b>conducted</b> exercise sessions for a Master-level Deep Learning course (three years)</li><li>➤ <b>Supervised</b> 15 Master's students during thesis projects, group projects and seminars;</li><li>➤ <b>Interdisciplinary collaboration</b> with other researchers, both as a <b>team lead</b> and as a contributor;</li><li>➤ <b>Communicated</b> complex ideas to professionals and neophytes alike, in writing and speech.</li><li>➤ Additional courses: time management and agile project management, good scientific practice, presentation skills, scientific writing</li></ul>
Now July 2019	<b>Guest Scientist, INSTITUTE OF COMPUTATIONAL BIOLOGY, HELMHOLTZ ZENTRUM, Munich, Germany</b> My Ph.D. project was about developing new frameworks to aid <i>in silico</i> design of personalized vaccines for cancer. <ul style="list-style-type: none"><li>➤ Developed vaccine design frameworks based on <b>mixed-integer linear programming</b>;</li><li>➤ Used <b>deep learning</b> to improve the identification of neo-epitopes (proteasomal cleavage, MHC and TCR binding);</li><li>➤ Performed <b>single cell analyses</b> to quantify the response to vaccines.</li></ul>
April 2019 October 2018	<b>Digital Consultant, NETLIGHT CONSULTING GMBH, Munich, Germany</b> I consulted a Munich-based financial firm and helped them expand and improve their transaction processing systems. <ul style="list-style-type: none"><li>➤ Learned to navigate large, mature code-bases and quickly add value in an <b>agile environment</b>;</li><li>➤ Obtained the Payment Card Industry Data Security Standard (PCI DSS) Developer certification;</li><li>➤ Acted as a Quality Assurance engineer ensuring reliability in day-to-day operations of the systems.</li></ul>
June 2018 February 2018	<b>Master Thesis Student, RISE SICS, Stockholm, Sweden</b> I used machine learning to predict, informally, the relationship between the energy that comes off the earth surface and the consequent change in temperature and wind speed with altitude. This is important to make climate simulations more precise, and improve our understanding of Earth's climate. <ul style="list-style-type: none"><li>➤ Found a bug in PySpark and submitted a patch for it;</li><li>➤ Benchmarked several machine learning algorithms, ensuring a fair and unbiased evaluation;</li><li>➤ Applied and extended my knowledge of big data processing systems (<b>pySpark</b>).</li></ul>

June 2017 March 2017	<b>Business Analytics Consultant, BRIGHTCAPE B.V., Eindhoven, the Netherlands</b> I consulted a major Dutch bank and helped them develop a system to predict the next actions of their customers to improve staff allocation. <ul style="list-style-type: none"> <li>➤ Acted as technical advisor supporting the CTO in <b>meetings with prospective customers</b>;</li> <li>➤ Used predictive process mining for <b>customer journey analytics</b>;</li> <li>➤ Used time series analysis for <b>demand forecasting</b> and <b>capacity planning</b>.</li> </ul>
June 2016 February 2016	<b>Open Source Developer, WIKIDATA, Trento, Italy</b> I improved and extended a text mining framework I developed during the <b>Google Summer of Code 2015</b> , resulting in the addition of 4.5M new claims about 12M entities coming from more than 50 different web sources to the Wikidata knowledge base. <ul style="list-style-type: none"> <li>➤ Used <b>web scraping</b> to obtain textual documents with the required information;</li> <li>➤ Used <b>named entity recognition</b> and <b>machine learning</b> to parse the text and extract the new facts;</li> </ul>
June 2016 April 2014	<b>Software Developer, SPAZIODATI SRL, Trento, Italy</b> I worked on the data acquisition and processing systems underlying the company flagship product. <ul style="list-style-type: none"> <li>➤ Created decoupled and extensible APIs through a microservices architecture (<b>Python Flask</b>).</li> <li>➤ Offloaded heavy data processing jobs to a cloud infrastructure (<b>Spark</b> on <b>AWS EC2</b>);</li> <li>➤ Gained experience in developing big data systems and working in <b>agile teams</b> following <b>Scrum</b>.</li> </ul>

## SELECTED PUBLICATIONS

- Dorigatti E, Schubert B, Bischl B, Rügamer D, (2023) *Frequentist uncertainty quantification in semi-structured neural networks*, International conference on Artificial Intelligence and Statistics (**AISTATS**), PMLR
- Ziegler I, Ma B, Nie E, Bischl B, Rügamer D, Schubert B, Dorigatti E, (2022) *What cleaves? Is proteasomal cleavage prediction reaching a ceiling?*, **NeurIPS LMRL**
- Dorigatti E, Bischl B, Schubert B, (2022) *Improved proteasomal cleavage prediction with positive-unlabeled learning*, Machine Learning for Health (**ML4H**)
- Boniolo F\*, Dorigatti E\*, Ohnmacht A J\*, Saur D, Schubert B, Menden M P, (2021) *Artificial Intelligence in Early Drug Development enabling Precision Oncology*. **Expert Opinion on Drug Discovery**, 16:9, 991-1007
- Fritz C\*, Dorigatti E\*, Rügamer D (2021) *Combining Graph Neural Networks and Spatio-temporal Disease Models to Predict COVID-19 Cases in Germany*. **Nature Scientific Reports** 12 (1), 1-18
- Rezaei M, Dorigatti E, Rügamer D, Bischl B, (2022, *Joint Debaised Representation Learning and Imbalanced Data Clustering*, IEEE International Conference on Data Mining (**ICDM**) Deep Learning and Clustering (DLC) Workshop
- Dorigatti E, Schubert B (2020) *Joint epitope selection and spacer design for string-of-beads vaccines*. **Bioinformatics** 36, i643–i650. European Conference on Computational Biology (**ECCB**)
- Dorigatti E, Schubert B, (2020) *Graph-theoretical formulation of the generalized epitope-based vaccine design problem* (RD Kouyos, Ed.). **PLOS Computational Biology** 16, e1008237

## EDUCATION

2019 - Now	Ph.D. in <i>Statistics</i> focused on deep learning, uncertainty quantification and discrete optimization for the design of personalized cancer vaccines. Ludwig Maximilian Universität, Munich, Germany.
2016 - 2018	Master's Degree in <b>Data Science</b> with a specialization in <i>Distributed Systems and Data Mining for Really Big Data</i> , and a minor Degree in <b>Innovation &amp; Entrepreneurship</b> , European Institute of Innovation and Technology (EIT) Digital Master School, GPA 95% <ul style="list-style-type: none"> <li>➤ First Year: <i>Computer Science and Engineering</i>, Eindhoven Technical University, Eindhoven, the Netherlands, Final Grade "cum Laude"</li> <li>➤ Second Year: <i>Information and Communication Technology</i>, KTH Royal Institute of Technology, Stockholm, Sweden, Final Grade A</li> </ul>
2013 - 2016	Bachelor's Degree in <b>Computer Science</b> , University of Trento, Trento, Italy, GPA 95%, Final Grade 110/110 "cum Laude"
2008 - 2013	High School Diploma for <b>Informatics Industrial Engineer</b> , ITT Marconi Rovereto, Rovereto, Italy, Final Grade 100/100

## LANGUAGES

Italian ● ● ● ● ●

English ● ● ● ● ○

German ,  
Chinese ● ○ ○ ○ ○