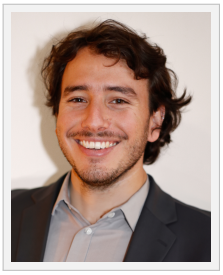
Santiago Manuel Castro Dau



General Information

Date of Birth	24th of May 1995
Nationalities	Mexico and United States of America
Residence	Zürich, Switzerland (student-B permit valid until 09/2024)
Lenguages	Spanish (native), English (fluent), German (B1)

Bio

- Computational Biologist with a Chemical Enginnering background.
- Extensive interdisciplinary research experience.
- Technical expertise in mathematical modeling, machine learning and software development.
- Paassionate about science and champion for curiosity-driven research.

Education

09/2020 – 06/2023 **Master of Science** in Computational Biology and Bioinformatics, ETH Zürich
Final average: 5.6. Master's thesis grade: 6.

08/2014 – 08/2019 **Bachelor of Engineering** in Chemical Engineering, Universidad Nacional Autónoma de México
Admission rate 2014: 10%. Graduated with honors, top 7% of cohort. Final average: 9.6/10.

02–07/2018 **Visiting Student**, University of Tokyo

Experience

08/2023–Now **Software Engineer at the Food Systems Biotechnology group at ETH Zürich.**

Duration	4 months at the moment of applying
Description	Developing microbiome-analysis software tools for Qiime2, a package collection with reproducibility at its core.
Supervisor	Michal Ziemski

12/2022 – 05/2023 **Master Thesis at the IBM Research in the Artificial Intelligence for Single Cell Research group.**

Duration	6 months for thesis. Ongoing involvement in the project.
Thesis Title	*Modeling the Tumor Microenvironment with Graph Concept Learning*
Description	Designed and implemented an innovative machine learning framework, enabling interpretable predictions of tumor tissue images through concept and geometric-deep learning.
Supervisors	Pushpak Pati, Marianna Rapsomaniki and Mark Robinson

03–06/2022 **Research Internship at the Cortical Computation group, Institute of Neuroinformatics at the University of Zürich, and ETH Zürich.**

Duration	3 months
Description	Modeled the learning process in biological neurons by representing the system with an artificial neural network and describing its dynamics through mixture model parameter inference.
Supervisor	Matthew Cook

06–09/2021 **Research Internship at the Computational Evolution group, D-BSSE, ETH Zürich.**

Duration	3 months
Description	Conducted a statistical analysis of COVID-19 epidemiological data to investigate the applicability of an *effective reproductive number estimator* to data from countries outside the EU.
Supervisor	Jana Huisman

☎ +41762011813 • ✉ sanntiago5@gmail.com • sann5.github.io
5antiago • Sann5

06/2019 – 06/2020 **Research Internship at the Computational Genomics group, National Institute of Genomic Medicine, Mexico City.**

Duration 12 months

Description Conducted a statistical analysis of copy number mutations in breast cancer to explore their impact on the structure of the transcriptional network.

Supervisor Enrique Hernández Lemus

Contributions to Repositories and Publications

2023 **Castro-Dau S.**, Martinelli A., Rapsomaniki M., Pait P. *Modeling the Tumor Microenvironment with Graph Concept Learning*. Paper in preparation.

2023 Currently working in a contribution to the PyTorch-Geometric package.

2022-2023 Contributed with pull requests to the packages q2-moshpit, q2-assembly, and ATHENA.

2020 Participated in the CTD^2 Pancancer Drug Activity DREAM challenge.

Special Achievements

2023 **Contributed Talk**

Title *Modeling the Tumor Microenvironment with Graph Concept Learning*, BC2

Selected for an oral presentation at Basel Computational Biology Conference 2023, Switzerland's main event in the domain of Computational Biology and one of the major events of its kind in Europe.

2021 **Winner of "ETH Week 2021: Health for Tomorrow"**

An interdisciplinary, innovative, problem-solving challenge, where we completed with more than 50 ETH students to come up with an attractive solution for a real-world health-related problem.

2020 **Recipient of "Jóvenes de Excelencia Citibanamex" scholarship**

Scholarship for promising graduates looking to continue their studies at a recognized international institution.

2019 **Contributed Talk**

Title *Effect of Structural Variants on Transcriptional Network Architectures*, INMEGEN

Selected for an oral presentation at Instituto Nacional de Medicina Genomica's 5th Student Encounter.

2018 **Recipient of "Movilidad Internacional" scholarship**

Scholarship to support my studies at the University of Tokyo, one of the world's top academic institutions.

Technical Skills

- Programming lenguages: Python, R, C++, SQL, Matlab, Bash, LaTeX.
- Fluency in popular libraries and command line tools, e.g. PyTorch, PyTorch-Geometric, scikit-learn, MLFlow, Snakemake, NetworkX, Qiime2, Tidyverse.
- Experience developing software adhering to various standard technologies and best practices including CI/CD tools.
- Proficient use of bash scripting, workflow management systems, and high-performance computing environments.
- Experience with a variety of machine learning approaches, e.g. bayesian, concept, reinforcement, geometric and active learning, vision transformers.
- Familiarity with proteomic, genomic, and transcriptomic datasets, both at standard resolution and spatially and single-cell resolved.