

# Geert-Jan Huizing

Phone : +33 6 70 86 38 09 — Email : gjhuizing@gmail.com

*PhD in machine learning & computational biology. Actively looking for a research scientist or postdoc position*

## Education

---

**Institut Pasteur & ENS PSL** *Ph.D. in computational biology* ..... **2020-2024**

Machine learning, optimal transport, single-cell multiomics, neural networks

Supervisors: Laura Cantini (CNRS, Institut Pasteur) and Gabriel Peyré (CNRS, ENS PSL)

**Institut Polytechnique de Paris** *M2 Data Science (masters' degree)* ..... **2019-2020**

Data science, machine learning, deep learning, mathematics, computer vision, signal processing

**Télécom Paris** *Diplôme d'ingénieur (masters' degree)* ..... **2017-2020**

Data science, machine learning, deep learning, mathematics, computer vision, signal processing

**Classe préparatoire Masséna, Nice** *MPSI-MP\** ..... **2015-2017**

Mathematics, physics, chemistry, computer science, engineering

## Research interests

---

My ongoing research is at the interface between **machine learning**, **optimal transport**, and **computational biology**. I have focused on the joint analysis of **single-cell multiomics** and **spatial transcriptomics**.

## Research and professional experience

---

**Single-cell multi-omics integration using Optimal Transport** *Ph.D. in computational biology* ..... **2020-2024**

CNRS, Institut Pasteur and ENS PSL, Paris, France

- Proposed a powerful framework to analyze single-cell data through Optimal Transport.
- Implemented machine learning models to jointly learn from different single-cell omics modalities.
- Developed open-source Python packages coded in PyTorch and JAX, easy to use for bioinformaticians.
- Presented my work in major international conferences and peer-reviewed in ISMB 2022 and ICML 2024.

**Optimal Transport for single-cell omics** *M.S. internship* ..... **May-Sept 2020**

CNRS and ENS PSL, Paris, France

- Proposed a novel way to analyze single-cell omics data using Optimal Transport distances between cells.
- Implemented my work using PyTorch and benchmarked it against traditional analysis pipelines.

**Data visualization at e-health unicorn Doctolib** *B.S. internship* ..... **Summer 2018**

Doctolib, Paris, France

- Using data visualization tool Periscope, SQL, and Python I created dashboards with team-specific KPIs.
- Worked with a product owner to visualize product adoption among doctors and hospitals and presented my work during a company 'tech talk'.

## Teaching experience

---

**Mathematics of data science** *Teaching assistant* ..... **2021-2022**

**Mathematics of data science** *Teaching assistant* ..... **2022-2023**

Teaching assistant for Gabriel Peyré's course at the ENS PSL. I made pen-and-paper exercises and Python practicals focused on the mathematics of signal processing. Themes include Shannon theory, Fourier theory, inverse problems, and wavelets.

## Selected talks

---

### Conferences and workshops

\*contributed, †invited

**CANUM**<sup>†</sup> *Learning cell fates landscapes from spatial transcriptomics with Fused Gromov-Wasserstein* ... Ile de Ré (France), **2024**

**ISMB/ECCB**\* *Paired single-cell multi-omics data integration with Mowgli* ..... Lyon (France), **2023**

**ECCB**\* *Optimal Transport improves cell-cell similarity inference in single-cell omics data* ..... Sitges (Spain), **2022**

**ICML**\* *Unsupervised ground metric learning using Wasserstein Singular Vectors* ..... Baltimore (USA), **2022**

**SMPGD**\* *Optimal Transport improves cell-cell similarity inference in single-cell omics data* ..... Online, **2022**

**Tri-Omics Summit**<sup>†</sup> *Defining cell-types using Optimal Transport* ..... London (UK), **2022**

**Journée Boris Ephrussi**\* *Paired single-cell multi-omics data integration with Mowgli* ..... Paris (France), **2022**

### Seminars

**Institut Curie single-cell seminar** *Paired single-cell multi-omics data integration with Mowgli* ..... Paris (France), **2024**

**scverse community meeting** *Paired single-cell multi-omics data integration with Mowgli* ..... Online, **2023**

**AI in Biology and Health** *Paired single-cell multi-omics data integration with Mowgli* ..... Institut Pasteur, Paris, **2023**

**MLSP** *Unsupervised ground metric learning using Wasserstein Singular Vectors* ..... ENS Lyon, **2023**

**DDisc** *Paired single-cell multi-omics data integration with Mowgli* ..... IMT, Toulouse, **2022**

**SingleStatOmics** *Optimal Transport improves cell-cell similarity inference in single-cell omics data* ..... Online, **2021**

## Publications

---

**GJ Huizing**, G Peyré, L Cantini, “Learning cell fate landscapes from spatial transcriptomics using Fused Gromov-Wasserstein”, *bioRxiv*, 2024.

**GJ Huizing**, I M Deutschmann, G Peyré, L Cantini, “Paired single-cell multi-omics data integration with Mowgli”, *Nature Communications*, 2023. **Contributed talk and best poster award** at ISMB/ECCB 2023, MLCSB session.

**GJ Huizing**, L Cantini, G Peyré, “Unsupervised ground metric learning using Wasserstein Singular Vectors”, *International Conference on Machine Learning*, 2022.

**GJ Huizing**, G Peyré, L Cantini, “Optimal transport improves cell-cell similarity inference in single-cell omics data”, *Bioinformatics*, 2022.

## Open-source contributions

---

### Hackathons

**OTT fest** *Hackathon at Apple Research in Paris to improve OTT-JAX* ..... **2023**

**OpenProblems Jamboree** *Hackathon on benchmarking tasks for single-cell genomics* ..... **2021**

### Open-source packages

#### STORIES

[github.com/cantinilab/stories](https://github.com/cantinilab/stories)

Learn spatially informed differentiation potentials using Optimal Transport. Built with JAX.

#### Mowgli

[github.com/cantinilab/Mowgli](https://github.com/cantinilab/Mowgli)

Perform single-cell multi-omics integration using Optimal Transport. Part of the scverse ecosystem.

#### wsingular

[github.com/CSDUlm/wsingular](https://github.com/CSDUlm/wsingular)

Compute different flavors of Wasserstein Singular Vectors, implemented efficiently using PyTorch and POT.

#### OT-scOmics

[github.com/cantinilab/OT-scOmics](https://github.com/cantinilab/OT-scOmics)

Compute Optimal Transport as a metric between cells in single-cell omics data. Leverages Pytorch and POT.

### Merged pull requests

**JKOnet** Fix call to deprecated JAX API [jknet/pull/5](https://github.com/jknet/pull/5)

**SciPy** Uniformize documentation of interpolator functions [scipy/pull/18413](https://github.com/scipy/scipy/pull/18413)

**OTT-JAX** Improve the plotting function ([ott-jax/ott/pull/280](https://github.com/ott-jax/ott/pull/280)), add a tutorial ([ott-jax/ott/pull/282](https://github.com/ott-jax/ott/pull/282)), and improve documentation ([ott-jax/ott/pull/288](https://github.com/ott-jax/ott/pull/288) and [ott-jax/ott/pull/270](https://github.com/ott-jax/ott/pull/270))

## Skills

---

**Mathematics** Matrix calculus, linear algebra, optimization, probability, statistics, variational inference, optimal transport.

**Machine Learning** Python environment, e.g. NumPy, Scikit-learn, PyTorch, JAX. Classical ML such as linear dimensionality reduction. Deep architectures, including VAEs and Transformers.

**Computational biology** Experience with analyzing single-cell omics, including gene expression, surface proteins, methylation, and chromatin accessibility. Clustering, visualization, integration, and gene set enrichment using R and Python.

**Software development** Git, Pytest, GitHub Actions, Readthedocs, Codecov, Poetry, PyPI, SLURM, Hydra, Weights & Biases.

**Scientific writing** Various Elsevier Researcher Academy certificates.  $\LaTeX$  is my go-to for writing and Inkscape for figures.

**Languages** French and Dutch (native), English (bilingual, *perfect score for TOEIC and TOEFL ITP*)

## Service

---

**Peer review** I sub-reviewed for ISMB 2022 and reviewed for ICML 2024.

**French Red Cross** I volunteer at the French Red Cross and hold a first aid certificate (PSC1).

## References

---

#### Laura Cantini

CNRS and Institut Pasteur, Paris, France  
PhD advisor  
[laura.cantini@pasteur.fr](mailto:laura.cantini@pasteur.fr)

#### Gabriel Peyré

CNRS and ENS PSL, Paris, France  
PhD advisor  
[gabriel.peyre@ens.fr](mailto:gabriel.peyre@ens.fr)