

# Stéphane Ghozzi

Data scientist with a training in theoretical physics, interested in analyzing, visualizing and interpreting data in close collaboration with data owners and users. Works in teams to apply machine learning methods and develop interactive dashboards.



[twitter.com/stephaneghozzi](https://twitter.com/stephaneghozzi) · [gitlab.com/stephaneghozzi](https://gitlab.com/stephaneghozzi) ·  
[linkedin.com/in/stephaneghozzi](https://linkedin.com/in/stephaneghozzi) · [scholar.google.com/citations?user=uGVLwREAAAAJ](https://scholar.google.com/citations?user=uGVLwREAAAAJ) ·  
[researchgate.net/profile/Stephane\\_Ghozzi](https://researchgate.net/profile/Stephane_Ghozzi)

---

## Experience

### Research assistant

#### Helmholtz Centre for Infection Research

June 1, 2020 – present · 6 mos

Brunswick (Braunschweig), Germany

[helmholtz-hzi.de/en/research/research-topics/bacterial-and-viral-pathogens/epidemiology/our-research/](https://helmholtz-hzi.de/en/research/research-topics/bacterial-and-viral-pathogens/epidemiology/our-research/)

#### Robert Koch Institute

April 15, 2016 – May 31, 2020 · 4 yrs 2 mos

Berlin, Germany

[rki.de/signale-project](https://rki.de/signale-project)

Machine learning, informatics, statistics and visualizations, Signale team, Department of Infectious-Disease Epidemiology.

##### Research and development:

- development of outbreak-detection algorithms and modelling of infection dynamics;
- performance evaluation and parameter optimization of the algorithms;
- natural language processing of online articles to support international infectious-disease surveillance;
- interactive visualizations of the results and other data for public-health professionals.

##### Management and organization:

- speaker for the team (four data scientists, two web developers, an average of two students);
- coordination of the international Topic Group Outbreaks under the Focus Group AI for Health of ITU and WHO: [itu.int/en/ITU-T/focusgroups/ai4h/Pages/tg.aspx](https://itu.int/en/ITU-T/focusgroups/ai4h/Pages/tg.aspx)
- supervision of master theses;
- organization of workshops and hackathons;
- writing of grant applications.

#### World Health Organization

May 1, 2019 – October 31, 2019 · 6 mos

Geneva, Switzerland

[who.int/eios](https://who.int/eios) · [who.int/emergencies/outbreak-toolkit](https://who.int/emergencies/outbreak-toolkit)

Machine learning and web-application development for epidemic intelligence and investigation of outbreaks of unknown origins.

In the Detection, Verification and Risk Assessment (DVA) and the Health Operations Monitoring and Data Collection (MDC) units of the Health Emergency Information and Risk Assessment (HIM) department within the WHO Health Emergencies (WHE) program of WHO.

**Visual artist      self employed**

March 1, 2012 – April 14, 2016 · 4 yrs 1 mo

Berlin, Germany

[stephaneghozzi.com/works](http://stephaneghozzi.com/works)

Drawing, generative animation, photography, video, computer animation.

Selected projects and collaborations:

- 2014: videos exhibited during the *backup* festival, E-Werk, Weimar;
- 2005: twelve illustrations for *Trace.project*, a compilation album of original electronic music;
- 2005: videography for the dance piece *Entre-Deux* by Mirjam Fruttiger, Paris and Rome (including one-week invitation at the Villa Médicis);
- 2004: videography on the documentary *Manchay Tiempo* by Florence Blum and María Pía Medina-Luna (four-weeks filming in Peru);
- 2002: drawings and photographs published in the magazine *R de réel*.

**Postdoctoral researcher      Institute for Theoretical Physics, Cologne University**

March 1, 2010 – February 29, 2012 · 2 yrs

Cologne, Germany

[www.thp.uni-koeln.de/~lassig](http://www.thp.uni-koeln.de/~lassig)

Statistical and mechanical models of biological evolution, in the Lässig group.

Mathematical model and analysis of bacterial growth and gene expression, interpretation of experimental results.

Signatures of selection in DNA sequences and comparison with population genetics models:

- long-term influenza evolution;
- transcription-binding-site motifs in yeast.

Co-evolution signatures in protein sequences.

**Teaching assistant      Cologne University**

September 1, 2010 – January 30, 2012 · 1 yr 5 mos

Cologne, Germany

Mathematics and statistical physics for Bachelor students.

**UPMC Sorbonne Universities**

October 1, 2005 – August 31, 2009 · 3 yrs 11 mos

Paris, France

Thermodynamics, optics and waves, mathematical methods for Bachelor students.

**Doctoral  
researcher**

**Laboratory of Statistical Physics, École normale supérieure**

September 1, 2005 – December 31, 2009 · 4 yrs 4 mos

Paris, France

[www.labos.upmc.fr/ljp/?article7](http://www.labos.upmc.fr/ljp/?article7) · [www.lps.ens.fr](http://www.lps.ens.fr)

Theoretical and experimental biophysics: Dynamics of gene regulatory networks, in the Chatenay group (now at Laboratoire Jean Perrin).

Won a 60 k€ grant to fund the experimental project (over 3 years, used to buy apparatus and consumables): program “Interface physique, biologie et chimie : soutien à la prise de risque 2007-2009” of the CNRS.

Time series of gene expression levels, via fluorescent microscopy, of the lysis-lysogeny decision network of the bacteriophage Lambda:

- molecular biology (extraction of viral genes, insertion of genes coding for fluorescent protein, modification of bacterial genomes);
- microbiology (bacterial and viral cultures);
- automation and fluorescent microscopy;
- image analysis.

Mathematical analysis of noise statistics of bacterial gene expression.

Computer simulations of the dynamics and evolution of gene regulatory networks.

**Volunteering  
Founding  
member**

**Celsius**

August 1, 2007 – August 31, 2009 · 2 yrs 1 mo

Paris, France

Celsius was a think tank with the goal of developing the European project.

- Elaboration of background documents on technical themes;
- preparation, organization of, and follow-up on, two-days meetings in Madrid, Brussels, and Paris, each with more than 30 participants;
- development of publication strategies, print and online.

**Internships**

**Laboratoire de l'Accélérateur Linéaire, Université Paris-Sud**

February 1, 2005 – March 31, 2005 · 2 mos

Orsay, France

Theoretical particle physics, in the Davier group: Computation of fundamental physical quantities for high energy physics from particle collision experiments data.

**DESY, Humboldt-Universität zu Berlin**

January 17, 2003 – August 31, 2003 · 7 mos

Zeuthen, Germany

Theoretical particle physics, in the Jegerlehner group: Analysis of experimental data with ad hoc and first-principle mathematical models of elementary particles.

**Laboratoire Kastler Brossel, École normale supérieure**

June 1, 2002 – August 31, 2002 · 3 mos

Paris, France

Experimental quantum physics, in the Grynberg group: Construction of an optical atom trap.

---

## Education

### Certifications    **Coursera**

January 1, 2016 – March 31, 2016 · 3 mos  
Practical Predictive Analytics: Models and Methods  
Machine Learning  
Data Manipulation at Scale: Systems and Algorithms

### Ph.D.                **École normale supérieure**

September 1, 2005 – December 31, 2009 · 4 yrs 4 mos  
Paris, France  
Theoretical and experimental biophysics: Dynamics of gene regulatory networks.  
See “Experience” above for details.

### B.Sc. & M.Sc.            **École normale supérieure**

September 1, 2001 – August 31, 2005 · 4 yrs  
Paris, France  
Theoretical and mathematical physics: specialization in particle physics and statistical physics.

Competitive fellowship.

Activities and societies: President of the Photography and Video-Making student associations in 2002 and 2003:

- presenting, defending, and managing budgets;
- initiation and support for members.

### M.A.                **École nationale supérieure des arts décoratifs**

September 1, 2003 – August 31, 2004 · 1 yr  
Paris, France  
CGI, 3d animation, post-production.

---

## Skills

**Programming**    R, Python, LaTeX, Mathematica, Matlab, Processing (Java), SQL, Perl, C++, Shell

**Tools**                Git, Team Foundation Server (agile management: Scrum, Kanban), Photoshop, Illustrator, After Effects, Microsoft Office, 3ds Max, Blender

**Languages**        French (native)  
German (professional)  
English (professional)  
Japanese (basics)

## Publications

- [1] A. Abbood, A. Ullrich, R. Busche, and S. Ghazzi, “Eventepi—a natural language processing framework for event-based surveillance,” *PLOS Computational Biology*, vol. 16, no. 11, pp. 1–16, 11 2020. [Online]. Available: <https://doi.org/10.1371/journal.pcbi.1008277>
- [2] T. Jombart\*, S. Ghazzi\*, D. Schumacher, Q. Leclerc, M. Jit, S. Flasche, F. Greaves, T. Ward, R. M. Eggo, E. Nightingale, S. Meakin, O. J. Brady, Centre for Mathematical Modelling of Infectious Diseases COVID-19 Working Group, G. Medley, M. Höhle, and J. Edmunds, “Real-time monitoring of COVID-19 dynamics using automated trend fitting and anomaly detection,” *medRxiv*, 2020. [Online]. Available: <https://doi.org/10.1101/2020.09.02.20186502>
- [3] O. Stojanović, J. Leugering, G. Pipa, S. Ghazzi, and A. Ullrich, “A Bayesian Monte Carlo approach for predicting the spread of infectious diseases,” *PLOS ONE*, vol. 14, no. 12, pp. 1–20, 12 2019. [Online]. Available: <https://doi.org/10.1371/journal.pone.0225838>
- [4] B. Zacher, A. Ullrich, and S. Ghazzi, “Supervised Learning for Automated Infectious-Disease-Outbreak Detection,” *Online Journal of Public Health Informatics*, vol. 11, no. 1, 2019. [Online]. Available: <https://doi.org/10.5210/ojphi.v11i1.9770>
- [5] F. Eckelmann, S. Ghazzi, and A. Ullrich, “Dashboards as strategy to integrate multiple data streams for real time surveillance,” *Online Journal of Public Health Informatics*, vol. 11, no. 1, 2019. [Online]. Available: <https://doi.org/10.5210/ojphi.v11i1.9701>
- [6] N. Sarma, A. Ullrich, H. Wilking, S. Ghazzi, A. K. Lindner, C. Weber, A. Holzer, A. Jansen, K. Stark, and S. Vygen-Bonnet, “Surveillance on speed: Being aware of infectious diseases in migrants mass accommodations - an easy and flexible toolkit for field application of syndromic surveillance, Germany, 2016 to 2017,” *Eurosurveillance*, vol. 23, no. 40, 2018. [Online]. Available: <https://doi.org/10.2807/1560-7917.ES.2018.23.40.1700430>
- [7] L. Perfeito, S. Ghazzi, J. Berg, K. Schnetz, and M. Lässig, “Nonlinear Fitness Landscape of a Molecular Pathway,” *PLOS Genetics*, vol. 7, no. 7, pp. 1–10, 07 2011. [Online]. Available: <https://doi.org/10.1371/journal.pgen.1002160>
- [8] S. Ghazzi, J. Wong Ng, D. Chatenay, and J. Robert, “Inference of plasmid-copy-number mean and noise from single-cell gene expression data,” *Phys. Rev. E*, vol. 82, p. 051916, Nov 2010. [Online]. Available: <https://doi.org/10.1103/PhysRevE.82.051916>
- [9] S. Ghazzi, “Expression Dynamics of a Genetic Regulatory Network: the Lysis/Lysogeny Decision of Bacteriophage Lambda,” Theses, Université Pierre et Marie Curie - Paris VI, Dec. 2009. [Online]. Available: <https://tel.archives-ouvertes.fr/tel-00515109>
- [10] S. Ghazzi and F. Jegerlehner, “Isospin violating effects in  $e^+e^-$  vs.  $\tau$  measurements of the pion form factor  $|F_\pi|^2(s)$ ,” *Physics Letters B*, vol. 583, no. 3, pp. 222 – 230, 2004. [Online]. Available: <https://doi.org/10.1016/j.physletb.2004.01.021>