

# Paul Scherer

Email: paul.scherer@cl.cam.ac.uk Github: paulmorio

Web: <https://www.cl.cam.ac.uk/~pms69/>

## Education

---

**University of Cambridge**  
**PhD Computer Science**

(Jan 2019 - Present)

My current research interests lie within graph representation learning and its applications in biomedical informatics. However, I carry a general interest in the design of inductive biases for learning algorithms, and machine learning applications in scientific research.

In addition to publishing research on biomedical informatics, I have authored several open source Python libraries for learning with graph structured data and active learning. I am supervised by Profs. Pietro Liò and Mateja Jamnik.

**University of Cambridge**  
**MPhil. Advanced Computer Science (Graduated with Distinction)**

(Oct 2017 - Nov 2018)

Undertook modules in machine learning and algorithms for data mining, biomedical information processing, affective computing, and natural language processing. A substantial focus of the degree was on critically assessing and replicating research papers as well as proposing original research.

Completed self-proposed dissertation on graph kernels supervised by Professor Pietro Liò

**University of Edinburgh**  
**BSc. Artificial Intelligence and Computer Science (Graduated with 1st Class Honours)**

(Sep 2012- Jun 2016)

Dissertation in developing clustering algorithms applicable to protein-protein interaction networks which incorporate biological knowledge encoded in ontologies.

## Selected Publications

---

Ramon Viñas Torné\*, **Paul Scherer\***, Nikola Simidjievski, Mateja Jamnik, Pietro Liò "Spatio-relational inductive biases in spatial cell-type deconvolution". (ICML2023 CompBio Workshop) 2023.

**Paul Scherer**, Pietro Liò, Mateja Jamnik "Distributed representations of graphs for drug pair scoring" (LoG2022). 2022. **Best reviewer award**.

**Paul Scherer**, Thomas Gaudelot, Alison Pouplin, Suraj M.S., Jyothish Soman, Lindsay Edwards, Jake Taylor-King. "PyRelationAL: A Library for Active Learning Research and Development." arXiv preprint 2022 (Under review).

Benedek Rozemberczki, **Paul Scherer**, Yixuan He, George Panagopoulos, Alexander Riedel, Maria Stefanoei, Oliver Kiss, Ferenc Beres, Guzmán López, Nicolas Collignon, and Rik Sarkar. 2021. "PyTorch Geometric Temporal: Spatiotemporal Signal Processing with Neural Machine Learning Models." (CIKM 2021), **Best resource paper award**

**Paul Scherer**, Pietro Liò "Learning distributed representations of graphs with Geo2DR" (ICML 2020 Workshop in Graph Representation Learning and Beyond), 2020

Nikola Simidjievski, Cristian Bodnar, Ifrah Tariq, **Paul Scherer**, Helena Andres-Terre, Zohreh Shams, Mateja Jamnik, Pietro Liò "Variational Autoencoders for Cancer Data Integration: Design Principles and Computational Practice" (Frontiers Genetics), 2019

## Select Open Source Projects

---

**PyRelationAL (2021-Present)** - Author/maintainer of modular open source active learning library for research and development.

**Pytorch Geometric Temporal (2020-Present)** - Part of the developing team for a dynamic extension of the Pytorch Geometric library for graph representation learning.

## Professional

---

### Research Intern / Research Consultant - Relation Therapeutics (Jan 2021 - Oct 2021)

Proposed and started the development of PyRelationAL a modular open source Python library for active learning. The package is now publicly available for research purposes.

### Part-time Research Assistant / Consultant - University of Cambridge (Jul 2018 - Present)

Developed an interactive ABM-CA platform for the Department of Land Economy. It is now used as a research and teaching tool for postgraduate students within the University. (2019-2022)

Investigated role privacy of multi-robot formations using generative adversarial networks (2018).

### Data Scientist - MRC Epidemiology Unit (Jul 2016 - Sep 2017)

Developed data harmonisation and analysis methods for an EU-FP7 funded epidemiological study using federated analysis techniques. This included onboarding 8 international institutions to participate under strict data governance rules. The pilot study has lead to a publication in the international journal for obstetrics and gynaecology BJOG.

### Informatics Supervisor - University of Edinburgh (Sep 2015 - Jun 2016)

Prepared and gave seminar like classes on topics covered in *Computation and Logic* and *Object Oriented Programming*. Wrote supplementary materials and enjoyed high attendance (>90%) in classes.

### Software Engineering Intern - Metasystems Asia Hong Kong Office (Jun 2015 - Aug 2015)

Implemented scrape and import tools for phpBB forums, e-mails, and SVN logs in PHP, Shell, and Python to transfer their entire internal knowledge-base off deprecated tools to Atlassian Confluence.

### Resident Assistant - University of Edinburgh (Aug 2014 - Jun 2016)

As a live-in RA the role focused on building relationships with residents and empathising with their needs, aspirations, and interests so that they can make the most out of University life.

Awarded the Edinburgh and Edinburgh Leadership Award for services to improve student well-being.

### Student Software Developer - Lightfields GmbH (Jun 2012 - Aug 2014)

Developed an in-house continuous integration suite and analysed different server/application monitoring tools to increase rate of development and maintenance.

### ADHS Voice Study Assistant - Charité Universitätsmedizin (Nov 2010 - Apr 2011)

Processed, analysed, and organized voice samples gathered from research subjects for the computational analysis and biometric diagnosis of ADHS (ADHD).

## Technical skills

---

**Python (Proficient):** Regularly use NumPy, SciPy, Pandas, Scikit-Learn, NetworkX, PyTorch, and comfortable using the Python ecosystem across operating systems.

**R (Proficient):** For data exploration, wrangling, visualisations, and quick observations.

**NetLogo (Familiar):** For agent based modelling and GIS simulation research.

**General:** Comfortable developing on Linux and Windows and modern coding practices.

Always happy to learn more! I have learned and used other languages (e.g. PHP, Java) on the job.

## Personal details

---

German national with full settled status in the UK.

Fluent in: English and German. Intermediate level: Japanese and Spanish.