

# Ana Pengelly, PhD

[GitHub](#) | [LinkedIn](#) | [anapengelly.info](mailto:anapengelly.info)

## PROFILE

Machine Learning (ML) researcher, passionate about using data for the greater good and transforming society for the better. My main goal is to use data and technology ethically in order to solve global socio-economic and health problems. I have broad international research experience and I have a unique combination of skills in Data Science/Machine Learning and Genetics/molecular biology. I am a proactive team player and enjoy collaborating and sharing my skills as well as learning from people with different skill-sets.

## EXPERIENCE

RESEARCH ASSOCIATE, COMPUTATIONAL EPIDEMIOLOGY, IMPERIAL COLLEGE, LONDON IN COLLABORATION WITH VISA, LONDON— 11/2020-PRESENT

Advanced statistical, time-series and deep-learning analysis with Visa transaction data to model COVID-19 spread in the UK. Explainable machine learning on high dimensionality health data (UK Biobank) to understand the impact of environmental exposures on cardiovascular diseases and construct health scores. **Tensorflow, Deep Learning (LSTM), ML experiment tracking with MLflow, advanced time-series statistical analysis (VARIMA, granger causality), XGboost, Shapley feature importance.**

STUDENT ANALYST, DATA SCIENCE LAB, VISA, LONDON — 05/2020-09/2020

Using Visa transaction data to model COVID-19 mortality in London boroughs. **Big data querying, data visualisation and dashboards and building advanced statistical models with Python.**

POSTDOCTORAL FELLOW, THE FRANCIS CRICK INSTITUTE, LONDON — 2016-2019

Functional analysis of transcriptional regulation by dFOXO during stress response. **Genetics, molecular biology, protein mass spectrometry, confocal microscopy.**

PHD RESEARCH, MAX PLANCK INSTITUTE OF BIOCHEMISTRY, MUNICH — 2010-2015

Functional Analysis of Histone Post-translational Modifications by the Polycomb group of Transcriptional Repressors. **Genetics, molecular biology, confocal microscopy, RNA-sequencing (from sample preparation to computational analysis on the Galaxy platform, using Python and R).**

FURTHER RESEARCH EXPERIENCE, EDINBURGH, SEATTLE, STOCKHOLM— 2007-2015

3 to 6 months internships at: Wellcome Trust Centre for Cell Biology, Edinburgh; Fred Hutchinson Cancer Research Center, Seattle; Karolinska Institute, Stockholm; University of Liverpool.

## EDUCATION

Imperial College, London — MSc Health Data Analytics and Machine Learning, 2020, **with distinction. Full scholarship from the School of Public Health.**

Max-Planck Institute of Biochemistry, Munich — PhD Molecular Biology and Genetics, 2015. **European Commission, Marie-Curie FPN7 fellowship.**

Université Paris Diderot, Paris 7 (now Université de Paris) — BSc & MRes Genetics, 2010. **Excellence-Major fellowship from the French Government.**

Institut Universitaire de Technologie, Saint-Etienne — Technical Diploma (DUT) Environmental Engineering, 2007. **Excellence-Major fellowship from the French Government.**

## SKILLS

Programming languages: Python, Unix, SQL, Git, R, basic HTML5 and CSS3. Database systems: mysql, postgresql, Hive, Hadoop, web-scraping with BeautifulSoup. Machine Learning, Deep Learning (Tensorflow, Pytorch), Computer Vision, Advanced Time-series Analysis, MLflow platform for ML experiment tracking. Python machine learning and statistics libraries: Pandas, Numpy, Scipy, Scikit-Learn, Statsmodels, Matplotlib, Plotnine. Statistics, Epidemiology, Transcriptomics, Proteomics, Molecular Biology, Confocal Microscopy, Biochemistry, Genetics. University teaching, student supervision, public speaking and article and report writing skills.

## SCIENTIFIC PUBLICATIONS

Stefana I.M., Driscoll P., Obata F., **Pengelly A. R.**, Bailey A., MacRae J.I., Gould A.P. Developmental diet regulates *Drosophila* lifespan via lipid autotoxins. *Nature Communications* 1384, (2017).

**Pengelly, A. R.**, Kalb, R., Finkl, K. & Müller, J. Transcriptional repression by PRC1 in the absence of H2A monoubiquitylation. *Genes & Development* 29, 1487-1492 (2015).

**Pengelly, A. R.**, Copur, Ö., Jäckle, H., Herzig, A. & Müller, J. A Histone Mutant Reproduces the Phenotype Caused by Loss of Histone-Modifying Factor Polycomb. *Science* 339, 698–699 (2013).

## LECTURES

I have given the following lectures to Imperial College, London, students as part of the MSc Health Data Analytics and Machine Learning:

- [Introduction to Molecular and Cellular Biology for computational scientists](#) as part of the Molecular Epidemiology.
- [Recurrent Neural Networks](#) as part of the Machine Learning module.