# KASIA KEDZIERSKA

For my PhD @ University of Oxford, I studied cancer of the uterus and chromatin organisation in disease progression. Over the summer, I joined the bioML team at the Microsoft Research New England as an intern where I worked on Foundation Models in single-cell biology. Previously, I worked with NLP methods and knowledge graphs during my internship @ Novo Nordisk Research Centre in Oxford. Using Transfer Learning and YOLO model I built a framework to identify sea pens from a video footage of an ocean floor @ Turing Data Study Group.

With extensive domain knowledge in **Computational Biology** complemented by my experience in **Data Science** & **Machine Learning** I am eager to drive advancements at the crossroads of these disciplines.

## □ SELECTED WORK EXPERIENCE

Intern @ Microsoft Research New England, Cambridge, Massachusetts, USA

During the summer I investigated the potential of the Foundation Models in the space of single cell biology. I was mentored by Alex Lu, Ava Amini, and Lorin Crawford.

Intern @ Novo Nordisk Research Centre Oxford, Oxford, United Kingdom

I worked with NLP and knowledge graphs to screen biomedical articles to identify and prioritise therapeutic targets. To increase the impact of the analysis and increase reach I built and deployed an interactive dashboard (using R Shiny) to allow colleagues within the company to investigate and visualise the results directly in real-time.

DPhil Researcher @ Wellcome Centre for Human Genetics, Big Data Institute,
University of Oxford, UK

In my PhD project I looked at how chromatin organization influences disease initiation and progression in uterine cancer using multimodal data. I was also working on building and refining ML models of cancer evolution, specifically identyfing evolutionary trajectories in the cancer of the uterus.

Visiting Graduate Student @ Ratan group, University of Virginia, USA

I carried out the research for my Master thesis Analysis of the mutational burden
across gene sets in cancer in which I modeled somatic mutations background
distribution using germline variation. I also developed SONiCS - algorithm to genotype
Short Tandem Repeats (STRs) using dense forward simulations of the polymerase
chain reaction (PCR).

### **EDUCATION**

present

2018

2018

2015

2022

2019

DPhil in Genomic Medicine and Statistics @ Nuffield Department of Medicine,
 Brasenose College

PhD fully funded by the Wellcome Trust Four-year PhD Studentships in Science

M. Sc. Eng., Biotechnology @ Warsaw University of Technology

Master thesis *Analysis of the mutational burden across gene sets in cancer* awarded the title of **The Best Master Thesis in Bioinformatics** defended in 2018.

# SELECTED AWARDS AND HONORS

2023 JXTX + CSHL 2023 Biology of Genomes Scholarship @ JXTX Foundation, Cold Spring Harbor Laboratory

Awarded to outstanding graduate students in genomics and data sciences.

Graduate Prize in the 'Outstanding work outside degree' category @ Nuffield Department of Medicine, University of Oxford

Each year Nuffield Department of Medicine, based on nominations, awards selected PhD students based on their performance within and outside of their degree.

Senior Hulme Scholarship @ Brasenose College, University of Oxford

Senior Hulme Scholarship is awarded by Brasenose College, University of Oxford to DPhil students whose academic performance is deemed to be exceptional.

## SUMMER SCHOOLS & HACKATHONS

Sea pen identification from video footage challenge @ Turing Data Study Group @ The Alan Turing Institute, London, UK

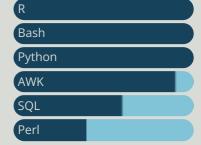
 Machine Learning Summer School @ Imperial College London, University College London, London, UK View this Resume online

kasia.codes/resume/

#### CONTACT

- ✓ kasia@well.ox.ac.uk
- **y** kzkedzierska
- github.com/kzkedzierska
- kasia.codes
- **in** kzkedzierska

#### **CODING SKILLS**



#### Python:

PyTorch, keras/Tensorflow, scverse/scanpy, scikitlearn, NumPy, Pandas, Seaborn, matplotlib and more

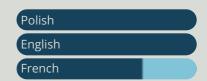
#### R:

tidyverse, data.table, Shiny, plotly, tidymodels, caret and others

# High Performance / Cloud Computing:

SLURM, Sun Grid Engine (SGE), Microsoft Azure, Azure Blob Storage, AWS Cloud Storage and more

#### **LANGUAGES**



CV source code. Made with pagedown, based on nstrayer/cv.

# ■ SELECTED PUBLICATIONS

Full list of publication is available through my Google Scholar profile scholar.google.com/citations?user=Yv6poTwAAAAJ.

- Assessing the limits of zero-shot foundation models in single-cell biology @ bioRxiv
   K. Z. Kedzierska, L. Crawford, A. P. Amini, A. X. Lu
- Data Study Group Final Report: CEFAS Automated identification of sea pens using OpenCV and machine learning @ Zenodo

  In alphabetical order: M. Asthana, R. Blackwell, S. Davis, A. Downie, I. Forsyth, K. Kedzierska, R. Mestre, Z. Reza, I. Ribeiro

In alphabetical order: M. Asthana, R. Blackwell, S. Davis, A. Downie, J. Forsyth, **K. Kedzierska**, R. Mestre, Z. Reza, J. Ribeiro, P. Palola, Y. Said

Functional analysis reveals driver cooperativity and novel mechanisms in endometrial carcinogenesis @ EMBO Molecular Medicine

M. Brown, A. Leon, **K. Kedzierska**, C. Moore, H. L. Belnoue-Davis, S. Flach, J. P. Lydon, F. J. DeMayo, A. Lewis, T. Bosse, I. Tomlinson, D. N. Church

Prognostic integrated image-based immune and molecular profiling in early-stage Endometrial Cancer @ Cancer Immunology Research

N. Horeweg, M. de Bruyn, R. A. Nout, E. Stelloo, **K. Kedzierska**, A. León-Castillo, A. Plat, K. D. Mertz, M. Osse, I. M. Jürgenliemk-Schulz, L. C.H.W. Lutgens, J. J. Jobsen, E. M. van der Steen-Banasik, V. T. Smit, C. L. Creutzberg, T. Bosse, H. W. Nijman, V. H. Koelzer and D. N. Church

Dynamics of cardiomyocyte transcriptome and chromatin landscape demarcates key events of heart development @ Genome Research

M. Pawlak, **K. Z. Kedzierska**, M. Migdal, K. A. Nahia, J. A. Ramilowski, L. Bugajski, K. Hashimoto, A. Marconi, K. Piwocka, P. Carninci and C. L. Winata

- 2018 SONICS: PCR stutter noise correction in genome-scale microsatellites @ Bioinformatics K. Z. Kedzierska, L. Gerber, D. Cagnazzi, M. Krützen, A. Ratan, L. Kistler
  - SELECTED CONFERENCE PRESENTATIONS
- Systematic characterisation of chromatin modifiers in endometrial cancer, Poster @ European Association for Cancer Research 2022 Congress, Seville, Spain
- Analysis of the mutational burden across gene sets in cancer, Invited talk @ Polish Bioinformatics Society Symposium, Cracow, Poland
  - ♣■ SELECTED TEACHING EXPERIENCE
- Data visualization in bioinformatics hackathon mentor @ Online hackathon NGSprint, Discord

  I led the hackathon in data viusalisation with emphasis on computational biology. Teaching materials are available at github.com/kzkedzierska/NGSprint\_data\_viz.
- Online tutorials: Python for Data Science and Introduction to Python @ NGSeminars, YouTube

  I led two Python tutorials: Introduction to Python kasia.codes/talk/intro\_to\_python/ and Python for Data Science kasia.codes/talk/py4ds/.
- Introduction to R @ Wellcome Centre for Human Genetics, Oxford, UK
   8 week course in Introduction to R, Data Manipulation, Data Visualisation and RNA-seq data analysis.
- Introduction to Managing Code with Git @ Wellcome Centre for Human Genetics, Oxford, UK

  I led a 2-hour introduction to working with Git. Materials, including slides and exercises are available at kasia.codes/talk/into\_to\_git/.
  - **€** SELECTED GRANTS
- Visegrad Grant to organise NGSchool2022 @ Visegrad Fund

  32,190 EUR awarded towards organising affordable training and conference focusing on ML application in

  Computational Biology. During this project I managed an international team of volunteers and led the organisation of summer school, conference, online seminars and hackathon.

## NON-PROFIT WORK

President @ NGSchool Society

The goal of the Society is to promote and support science, with emphasis on computational biology.