KASIA KEDZIERSKA

I am a Computational Biologist and Machine Learning Researcher dedicated to advancing methods development and application within biology and healthcare.

Currently, I lead an R&D project focused on **biological Foundation Models** (FMs). During my internship with the **bioML team** at Microsoft Research New England, I evaluated FMs for single-cell biology applications. Previously, at Novo Nordisk Research Centre in Oxford, I used NLP techniques and knowledge graphs to enhance biomedical research. Additionally, with the Turing Data Study Group, I developed a framework using **Transfer Learning** and the **YOLO** model to identify sea pens from ocean floor video footage.

With my extensive expertise in Computational Biology, coupled with a robust background in Data Science and Machine Learning, I am enthusiastic about driving innovations at the intersection of these disciplines.

🖥 RESEARCH EXPERIENCE

present | 2024

Deep Learning Researcher @ DeepLife, Remote

Leading an R&D project on Foundation Models in biology, including **mentoring** and supervising a **Master's student**.

2023

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.

2021

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.

2024 | 2018

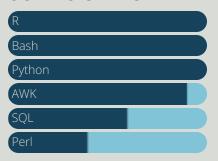**DPhil Researcher @ Wellcome Centre for Human Genetics, Big Data Institute, University of Oxford, UK**

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.

CONTACT

- ✉ me@kasia.codes
- 🐦 kzkedzierska
- 🐙 github.com/kzkedzierska
- 🔗 kasia.codes
- in kzkedzierska

CODING SKILLS

R

Bash

Python

AWK

SQL

Perl

Python:
PyTorch, keras/Tensorflow, scverse/scanpy, scikit-learn, NumPy, Pandas, Seaborn, matplotlib and more

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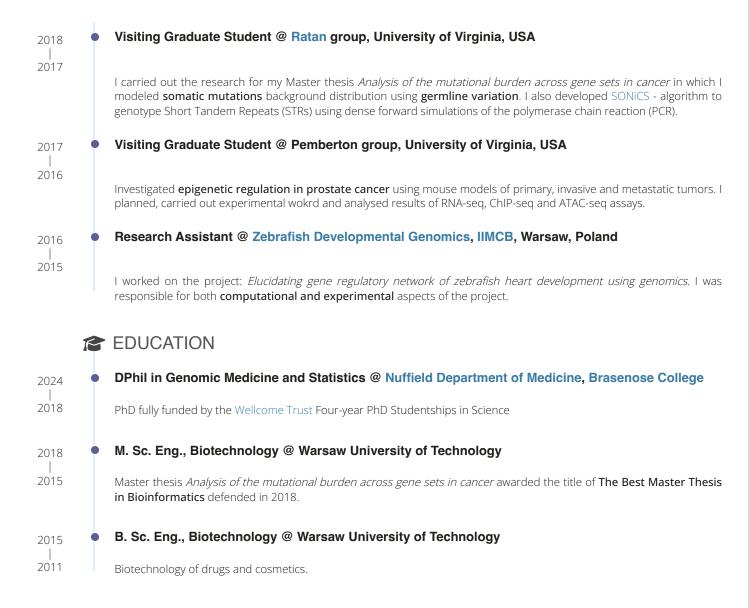tidyverse, data.table, Shiny, plotly, tidymodels, caret and others

High Performance / Cloud Computing:
SLURM, Sun Grid Engine (SGE), Microsoft Azure, Azure Blob Storage, AWS Cloud Storage and more

LANGUAGES

Polish

English

French

2018
|
2017

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).

2017
|
2016

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

Investigated **epigenetic regulation in prostate cancer** using mouse models of primary, invasive and metastatic tumors. I planned, carried out experimental wokrd and analysed results of RNA-seq, ChIP-seq and ATAC-seq assays.

2016
|
2015

Research Assistant @ Zebrafish Developmental Genomics, IIMCB, Warsaw, Poland

I worked on the project: *Elucidating gene regulatory network of zebrafish heart development using genomics.* I was responsible for both **computational and experimental** aspects of the project.

🎓 EDUCATION

2024
|
2018

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

2018
|
2015

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.

2015
|
2011

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

Biotechnology of drugs and cosmetics.

📓 PUBLICATIONS

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

2023

Assessing the limits of zero-shot foundation models in single-cell biology @ bioRxiv

K. Z. Kedzierska, L. Crawford, A. P. Amini, A. X. Lu ✨ This work was featured in The New York Times✨

2023

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, J. Forsyth, **K. Kedzierska**, R. Mestre, Z. Reza, J. Ribeiro, P. Palola, Y. Said

2023 — **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

2022 — **Discordant prognosis of mismatch repair deficiency in colorectal and endometrial cancer reflects variation in antitumour immune response and immune escape** @ The Journal of pathology

M. A. Glaire, N. AJ. Ryan, M. E. Ijsselsteijn, **K. Kedzierska**, S. Obolenski, R. Ali, E. J. Crosbie, T. Bosse, N. F. D. Miranda, D. N. Church

2021 — **Multi-omics analyses of early liver injury reveals cell-type-specific transcriptional and epigenomic shift** @ BMC Genomics

M. Migdal, E. Tralle, K. A. Nahia, L. Bugajski, **K. Z. Kedzierska**, F. Garbicz, K. Piwocka, C. L. Winata, M. Pawlak

2020 — **The *MLH1* polymorphism rs1800734 and risk of endometrial cancer with microsatellite instability** @ Clinical Epigenetics

H. Russell, **K. Kedzierska**, D. D. Buchanan, R. Thomas, E. Tham, M. Mints, A. Keränen, G. G. Giles, M. C. Southey, R. L. Milne, I. Tomlinson, D. Church, A. B. Spurdle, T. A. O'Mara and A. Lewis

2020 — **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

2019 — **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 — **Genomic analysis of DNA repair genes and androgen signaling in prostate cancer** @ BMC Cancer

K. Jividen, **K. Z. Kedzierska**, C.-S. Yang, K. Szlachta, A. Ratan and B. M. Paschal

2018 — **SONiCS: PCR stutter noise correction in genome-scale microsatellites** @ Bioinformatics

K. Z. Kedzierska, L. Gerber, D. Cagnazzi, M. Krützen, A. Ratan, L. Kistler

💬 POSTERS, AND TALKS

2024 — **Invited attendee** @ Chan Zuckerberg Initiative AI in Single-Cell Biology Workshop, Redwood City, CA, USA

2023 ● **Chromatin modfiers in endometrial cancer,** Poster @ Biology of Genomes 2023, **Cold Spring Harbor, NY, USA**

2022 ● **Systematic characterisation of chromatin modifiers in endometrial cancer,** Poster @ European Association for Cancer Research 2022 Congress, **Seville, Spain**

2019 ● **Analysis of the mutational burden across gene sets in cancer,** Invited talk @ Polish Bioinformatics Society Symposium, **Cracow, Poland**

2018 ● **Differential mutation analysis across gene sets in cancers,** Poster @ Biology of Genomes 2018, **Cold Spring Harbor, NY, USA**

2017 ● **Epigenetic regulation of prostate cancer,** Talk @ Visiting Graduate Traineeship Program Grantees Symposium, Charlottesville, VA, USA

🎗 AWARDS

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

Awarded to outstanding graduate students in genomics and data sciences.

2022 ● **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.

2023
|
2021 ● **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.

2019
|
2018 ● **Best Master Thesis in Bioinformatics   @ Polish Bioinformatics Society**

Analysis of the mutational burden across gene sets in cancer - Best Master Thesis defended in Bioinformatics in 2018 in Poland.

2017
|
2016 ● **Visiting Graduate Traineeship Program, now known as BioLAB @ Fulbright Poland**

The Visiting Research Graduate Traineeship Program offered 12-month research traineeships for outstanding, qualified students from the life sciences at selected institutions in the United States.

2015 ● **Grasz o Staz @ PwC Poland**

"Grasz o Staz" competition was a national, prestigious and highly competitive (1:25 success rate) scholarship program in Poland organized by PwC.

TEACHING EXPERIENCE

2022

Unsupervised learning @ NGSchool2022: Machine Learning in Computational Biology, Jablonna, Poland

I co-led, with Kaspar Märtens, lecture and tutorial sessions on unsupervised learning and its use cases in computational biology. All materials are availble at github.com/kzkedzierska/ngs22_unsupervised.

2021

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.

2020

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/.

2019

Unsupervised learning, Introduction to Python @ #NGSchool2019: Machine Learning for Biomedicine, Ostróda, Poland

Tutor for the Introduction to Python (3 h workshop) and for the Unsupervised learning (1,5 h lecture).

2019

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.

2019

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/.

2017

ATAC-seq workshop @ #NGSchool2017: Single-cell Sequencing, Jachranka, Poland

Invited speaker

ATTENDED WORKSHOPS, SUMMER SCHOOLS

2022

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

2019

Machine Learning Summer School @ Imperial College London, University College London, London, UK

€ GRANTS

2022
|
2020

● **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.

2019

● **Visegrad Grant to organise NGSchool2019 @ Visegrad Fund**

23,500 EUR awarded towards organising #NGSchool2019 allowed to keep the cost of attenting the school to the minimum and record the lectures for broader access.

2023
|
2018

● **Scientific grant as part of Wellcome Trust funded DPhil @ Wellcome Trust**

30,000 GBP towards research expenses for the PhD project which allowed me to design and led pilot experiments, managing the grant for successful execution of research objectives.

🤝 NON PROFIT WORK

2022
|
2018

● **President @ NGSchool Society**

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

2016
|
2015

● **President @ Warsaw Society of Biotechnology**

Symbioza's aim is to integrate biotechnology community in Poland starting from the youngest generations, and promotion and popularization of biotechnology.