Charles Harris

+44 7792 541377 | cch57@cam.ac.uk | GitHub | Website | LinkedIn | British Citizen

EDUCATION

University of Cambridge Oct. 2021 – Present
PhD in Computer Science Cambridge, UK

- **Research**: Focus on molecular design, diffusion models and discovering new biology with self-supervised learning
- **Supervisors**: Prof Sir Tom Blundell, Prof Pietro Liò
- **Funding**: Cambridge Centre for AI in Medicine (CCAIM) Studentship, sponsored by AstraZeneca and GSK

Imperial College London Oct. 2020 – Sept. 2021
*MSc in Bioinformatics and Theoretical Systems Biology - **Distinction - 76.5%*** London, UK

- **Supervisors**: Prof Michael Bronstein (Oxford, Twitter), Prof Bruno Correia (EPFL), Prof Michael Sternberg

Imperial College London Oct. 2017 – Sept. 2020
BSc in Biochemistry - 2:1 London, UK

- **Core modules**: Structural Biology, Drug Design, Bioinformatics, Systems Biology

EXPERIENCE

BenevolentAI July. 2022 – Sept. 2022
AI Scientist Intern London, UK

- Worked on developing new AI tools using GNNs and self-supervised learning for the Target Identification team.

SELECTED PUBLICATIONS

Machine Learning-aided Generative Molecular Design. Y. Du, A. R. Jamasb, J. Guo, T. Fu, **C. Harris**, Y. Wang, C. Duan, P. Lio, P. Schwaller, T. Blundell. Nature Machine Intelligence

SynFlowNet: Towards Molecule Design with Guaranteed Synthesis Pathways. M. Cretu, **C. Harris**, E. Bengio, P. Lio - *ICLR 2024 GemBio Workshop*. Link

Evaluating Representation Learning on the Protein Structure Universe A. Jamasb, A. Morehead, Z. Zhang, C. Joshi, K. Didi, S. Mathis, **C. Harris**, J. Tang, J. Cheng, P. Lio, T. Blundell - *ICLR 2024*.

PoseCheck: Generative Models for 3D Structure-based Drug Design Produce Unrealistic Poses. C. Harris, K. Didi, A. Jamasb, C. Joshi, S. Mathis, P. Lio, T. Blundell - *NeurIPS 2023 MLSB Workshop* - [Oral]. Link

DiffHopp: A Graph Diffusion Model for Novel Drug Design via Scaffold Hopping. J. Torge, **C. Harris**, S. Mathis, P. Lio - *ICML WCB 2023* - [Spotlight]. Link

Multi-State RNA Design with Geometric Multi-Graph Neural Networks. C. Joshi, A. Jasamb, R. Vinas **C. Harris**, S. Mathis, P. Lio - *ICML WCB 2023*. Link

Flexible Small-Molecule Design and Optimization with Equivariant Diffusion Models. Charles Harris, K. Didi, A. Schneuing, Y. Du, A. Jamasb, M. Bronstein, B. Correia, P. Lio, T. Blundell - *ICLR MLDD 2023*. Link

Equivariant Diffusion Models for Structure-based Drug Design. A. Schneuing, Y. Du, **C. Harris**, A. R. Jamasb, I. Igashov, W. Du, T. L. Blundell, P. Lió, C. Gomes, M. Welling, M. Bronstein, B. Corriea. - *NeurIPS MLSB Workshop 2022*. Link

Graphein - a Python Library for Geometric Deep Learning and Network Analysis on Protein Structures and Interaction Networks. A. R. Jamasb, R. Viñas Torné, E. J. Ma, **C. Harris**, K. Huang, D. Hall, P. Lió, T. L. Blundell. - *NeurIPS 2022*. Link

COMMUNICATION

Guest Writer - Royal Society of Chemistry CICAG Newsletter Article
Published articles on AI in Drug Discovery and Diffusion Models (viewed over 5,000 times) Feb. 2022

Chair and Founder - 1st Cambridge AI in Drug Discovery Conference Event website
Sold over 3,000 tickets Feb. 2022

Guest - iGEM Synthetic Biology Podcast Link
Discussed AlphaFold2, my research and the impact of computation and AI on biology in general Aug. 2021

Chair and Founder - 1st Imperial AI in Drug Discovery Conference Handbook
Sold over 1,400 tickets Feb. 2021

- Created and chaired Imperial's first **AI in Drug Discovery Conference** with top scientists, business leaders and entrepreneurs, over 1,400 tickets sold
- Moderated two panel discussions (first one with 4 CEOs/Founders of AI in Drug Discovery companies and second with Prof Sir Tom Blundell, Prof Michael Bronstein and Dr Andreas Bender)

INVITED TALKS

Department of Computer Engineering, Sapienza University of Rome - *Rome, Italy* Feb. 2024
VantAI "AI in Drug Discovery Seminar Series" - *NYC, USA* - *Recording* Nov. 2023
AstraZeneca AI Journal Club - *Cambridge, UK* Aug. 2023
Genesis Therapeutics - *San Francisco, USA* Aug. 2023
Merck - *San Francisco, USA* Aug. 2023
Cambridge Chemoinformatics Network meeting - *University of Cambridge* - *Recording* Jun. 2023
Imperial College Computational Biology Conference - *Imperial College London* May 2023
AstraZeneca - *Cambridge, UK* April 2023
IBM Research - *Zurich, Switzerland* April 2023
AI UK Conference (Demonstrator) - *QEII Conference Centre, Westminster* Mar. 2023
Graph Neural Networks and Geometric Deep Learning Course - *University of Cambridge* Feb. 2023

SUPERVISING

Krisztina Sinkovics - MPhil Machine Learning
Theory of conditional flow-matching for small molecule design Jan. 2024 – Present

Jamie Weigold - MPhil Advanced Computer Science Project 87.5% - Top 3 project in cohort
Diffusion models for protein-protein docking Oct. 2022 – Jul. 2023

Keiran Didi - MPhil Computational Biology Project 90% - Top project in cohort
Diffusion models for the protein motif-scaffolding problem Oct. 2022 – Jul. 2023

Jos Torge - Part II Computer Science Tripos Project Spotlight paper at WCB ICML 2024
DiffHopp: A Graph Diffusion Model for Novel Drug Design via Scaffold Hopping Oct. 2022 – Jul. 2023

Mihailo Milosevic - Part II Computer Science Tripos Project
Self-supervised learning of ligand-binding Oct. 2022 – Jul. 2023

Kevalee Shah - Part III Computer Science Tripos Project 81%
Self-supervised learning of ligand-binding Oct. 2021 – Jul. 2022

Zhang Shu - MPhil Advanced Computer Science Project
Equivariant Denoising Diffusion models for Protein Ligand Binding Jan. 2022 – Jul. 2022

Olivier Dietrich - Visiting Masters Project
AI for enzyme kinetics predictions using multi-modal representations Jan. 2022 – Jul. 2022

ACADEMIC SERVICE

Organiser
Machine Learning for Life and Material Sciences Workshop at ICML 2024 Feb. 2024 – July. 2024

- Advised the team on how to leverage different protein structure prediction tools for their project (including AlphaFold2)

Reviewer
Machine Learning for Structural Biology Workshop at NeurIPS 2023 Feb. 2024 – July. 2024

Reviewer
Computational Biology Workshop at ICML 2023 Feb. 2024 – July. 2024

Mentor/Organiser
Catalyse Competition - SynBioUK Dec. 2020 – Present

Advisor Project Website
Imperial College International Directed Evolution Competition (iDEC) Team Jun. 2021 – Oct. 2021

- Advised the team on how to leverage different protein structure prediction tools for their project (including AlphaFold2)

TEACHING

Geometric Deep Learning - *Part III/MPhil CS, University of Cambridge* Jan. 2022 - Present
Bioinformatics - *Part II Computer Science Tripos, University of Cambridge* Nov. 2021 - Dec. 2021

AWARDS - ACADEMIC

Polaris Fellowship - *Entrepreneur First* Dec. 2023
UK-Italy Visiting Researcher Fellowship - *Alan Turing Institute* Oct. 2023
CCAIM PhD Studentship - *University of Cambridge* Oct. 2021
Associateship - *Royal College of Science* Jul. 2020
Gold - *UK Chemistry Olympiad* Jun. 2017
Prefect - *Leighton Park School* June. 2016
David Lean Scholar - *Leighton Park School* Sep. 2015

AWARDS - SPORTING

Cambridge Half Blue (Gliding) - *University of Cambridge* July. 2023
1st Place - *Oxford-Cambridge Gliding Varsity* July. 2023
1st Place, Best Technical Soaring Flight - *National Interuniversity Gliding Competeition* July. 2023

VOLUNTEERING

Ordinary Committee Member
Cambridge University Gliding Club May. 2022 – Jan. 2024

Events and Conference Officer Website
Cambridge University Artificial Intelligence Society Oct. 2021 – Oct. 2022

Chair and Founder Twitter - Instagram
Imperial College Computational Biology Society Oct. 2019 – Aug. 2021

Ambassador
Helen Arkell Dyslexia Charity Oct. 2019 – Present

OTHER SKILLS

Languages: Mandarin (A at GCSE Level)
Interests: Qualified glider pilot (Full Silver Badge, 1st Place - Oxford-Cambridge Varity 2023), Hockey, Running, Drone photography, Science communication