UNIVERSITY OF CAMBRIDGE

ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

2025

ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

DONOR REPORT

FOREWORD

2025 has brought further shifts in the relationship between AI and scientific discovery. Questions about how to deploy AI responsibly and effectively have moved to the centre of research policy discussions internationally, and the University of Cambridge has been at the centre of these conversations. Thanks to the ongoing support of Schmidt Sciences, the Accelerate Programme for Scientific Discovery has evolved from a research initiative into a recognised hub for thought leadership on AI in science – one that offers a distinct vision that reaches across diverse domains, and builds a community from the ground up.

This year, we've seen our community develop in ways that extend beyond traditional metrics of research output. Perhaps most significantly, we've achieved a step-change in our training infrastructure. We now have a model that can deliver at scale while remaining nimble enough to adapt to emerging opportunities as they arise. Our course on Large Language Models, launched in response to the rapid progress in generative AI, has seen growing demand and demonstrates our ability to move quickly to equip researchers with the skills they need. Our other courses continue to attract interest, attracting participants from 86 departments. This year alone, we've engaged over 700 researchers in our courses, bringing the total to over 1500. The AI Clinic has become an indispensable resource for Cambridge researchers, providing 100 consultations across 44 departments this year including 5 café events and ongoing advisory support. The breadth of disciplines represented—from neuroscience, to marine ecology, to economics—illustrates how AI is becoming an important tool across the sciences.

Policy interest in AI for science has been a defining feature of 2025. Accelerate has had the privilege of working with both the European Commission and the UK Government to help develop their AI for science strategies. This reflects the critical role that universities play in ensuring that AI development serves science and society. Through our position in the European Leadership in Innovation with AI and Science (ELIAS) Alliance and our expanding network of international collaborations, we're helping to articulate a vision for AI that prioritises open science, interdisciplinary collaboration, and responsible innovation.

Accelerate is now positioned as a hub for the broader conversation about AI's role in advancing science. We're pleased to be launching a new call for papers in the Royal Statistical Society journal *Data Science and AI*, which will provide a platform for researchers across disciplines to share insights on deploying AI in their work. This initiative reflects our commitment to widening the dialogue and ensuring that diverse voices and perspectives shape the future of AI in science.

Our research continues to advance across multiple frontiers. One of our 2024 funding recipients, Dr Boris Bolliet from the Department of Physics, has developed an agentic AI system that can autonomously conduct scientific research, attracting significant attention, including a collaboration with Google DeepMind and Infosys and a prestigious 2025 Google Research Scholar award. Megan Ennion from the Faculty of Education, another 2024 funding recipient, is organising a series of dialogue and collaboration events on AI in Education at the university. Her work has already led to an invitation to present findings to policymakers in the UK, demonstrating its potential to influence national education and AI policies. The collaborative research scheme currently supports 33 projects spanning 21 departments and institutes, catalysing new interdisciplinary partnerships in areas from investigative journalism to detection of dementia before symptoms appear.

As we move into year six of the Programme, we're excited about the momentum we've built. The platforms for training, the networks for collaboration, the mechanisms for software and engineering support – these are the intellectual infrastructure that will enable Cambridge to continue leading the conversation about how AI can accelerate scientific discovery.

We remain deeply grateful to Schmidt Sciences for their continued partnership in this ambitious and evolving mission.

Jessica Montgomery
Director, Accelerate Programme for Scientific Discovery and ai@cam
Photo: Nick Saffell

Neil Lawrence
DeepMind Professor of Machine Learning
Photo: Nick Saffell

CONTENTS

SECTION 01

Capability Building

Building a Scalable, Adaptive Training Ecosystem for AI in Science

Accelerate is equipping early career researchers and future research leaders with the data science and machine learning skills that can help them drive a new wave of scientific progress.

→

A New Model for Training in a Fast-Moving Field

In AI, the challenge is not only to teach core concepts – it is upskilling at pace. The field moves faster than traditional academic curricula can adapt. New methods emerge, gain traction, and become integrated in daily practice within months; at the same time, building capability in AI across science is a long-term project, given the breadth and diversity of needs in research. For researchers trying to keep up, this creates a knowledge gap.

Accelerate has demonstrated a solution to this challenge: a training model that combines scale with adaptability. We have built infrastructure that can rapidly respond to technological developments while reaching hundreds of researchers across the University. When large language models became important for scientific applications, we developed and launched a workshop within weeks. When demand exceeded capacity, we engaged a community to deliver more training opportunities to a wider audience.

This approach represents something distinctive in academic training: the ability to move at the speed of innovation while maintaining quality and reach.

The results show the effectiveness of this approach:

- Over 700 researchers engaged through our training programmes in 2025
- New courses launched from conception to delivery in 8 weeks
- Waiting lists of 100 for our most in-demand workshops

Attendees at the Large Language Models for Science Workshop, October 2025 Photo: Bianca Provost

CAPABILITY BUILDING

Our Training Portfolio

Accelerate delivers training across a comprehensive curriculum that takes researchers from foundational skills through to deploying cutting-edge AI methods.

Foundational Skills

1 Python for Science

Our self-paced Python course continues to serve as the entry point for researchers building their data skills. The course has now received over 500 registrations.

2 Machine Learning Academy

In partnership with Cambridge Spark, our year-long Machine Learning Academy has welcomed its fifth cohort in 2025, for a total of 36 researchers registered for the academy this year. Participants develop expertise across specialist topics, including time series analysis, natural language processing and supervised and unsupervised learning.

Core Methods

3 Machine Learning and the Physical World

This course provides methodological grounding for researchers applying ML to systems that interact with the real world. In 2025, 50 students participated in the course, which for the first time is being delivered both at Cambridge and the University of Manchester. The course has recently been revamped, and now couples training on modern generative AI applied to scientific research problems alongside classical statistical techniques.

Core Methods

4 Publishing and Packaging Code

Essential for researchers wanting to share their ML work with the wider community, this workshop covers workflows and tools for making Python code reproducible and accessible. 57 participants attended in 2025, learning about version control in Git and GitHub, how to test scientific software, development with Poetry and documentation and styling best practice.

> **“[The instructor was] consistently dropping anecdotes and knowledge gained through experience. I felt like I gained a real-world appreciation of the how and why underpinning everything, rather than just technical detail.”**
> Feedback from Publishing and Packaging Code Workshop Participant

Frontier Methods

5 Large Language Models for Science

Our LLM workshop shows why rapid course development matters. When ChatGPT launched in late 2022, it became clear that LLMs would influence scientific practice. Researchers needed to understand not just how to use these tools, but how they work, where they fail, and how to deploy them responsibly in research contexts. Within 12 weeks, we had designed, piloted, and launched our LLM workshop. The response confirmed Accelerate had met a community need: 6 sessions delivered in 2025, reaching 197 researchers, with waiting lists that consistently exceeded 100 participants.

> **“This was the best course I have ever attended (and I’ve been to a lot). Highly engaging, information-rich, accessible for someone with bioinformatics / statistics training but no computer science background. Exactly what I was looking for. Thank you!”**
> Feedback from Large Language Models Workshop Participant

CAPABILITY BUILDING

6 Diffusion Models for Science

Similarly responsive to emerging methods, our Diffusion Models workshop addresses the explosion of interest in generative AI for scientific imaging, molecular design, and other applications. In 2025, 92 participants attended 3 sessions, learning about the history of generative AI and the inspiration it takes from physics, the building blocks of diffusion (VAEs, U-NET and CLIP) and algorithms, research applications and data ethics considerations.

> **“My current research involves the use of diffusion models, but I have been treating them as impenetrable black boxes. My goal was learning more about the inner workings of these models as well as to discover tools that can improve my research. I strongly agree that the workshop exposed me to new knowledge and practices.”**
> Feedback from Diffusion Models for Science Workshop Participant

▶ Diffusion Models for Science workshop, October 2025

Hands-On AI Workshop

7 Hands-On AI Workshop

Moving beyond conceptual understanding to practical deployment, Accelerate's Hands-On AI workshops give participants the opportunity to work with real ML problems and datasets. This year, over 100 researchers have participated in these sessions, gaining confidence in using tools they can immediately apply to their own research. In 2025, we introduced a specialised track focused on LLM applications, allowing researchers to work through concrete use cases relevant to their domains.

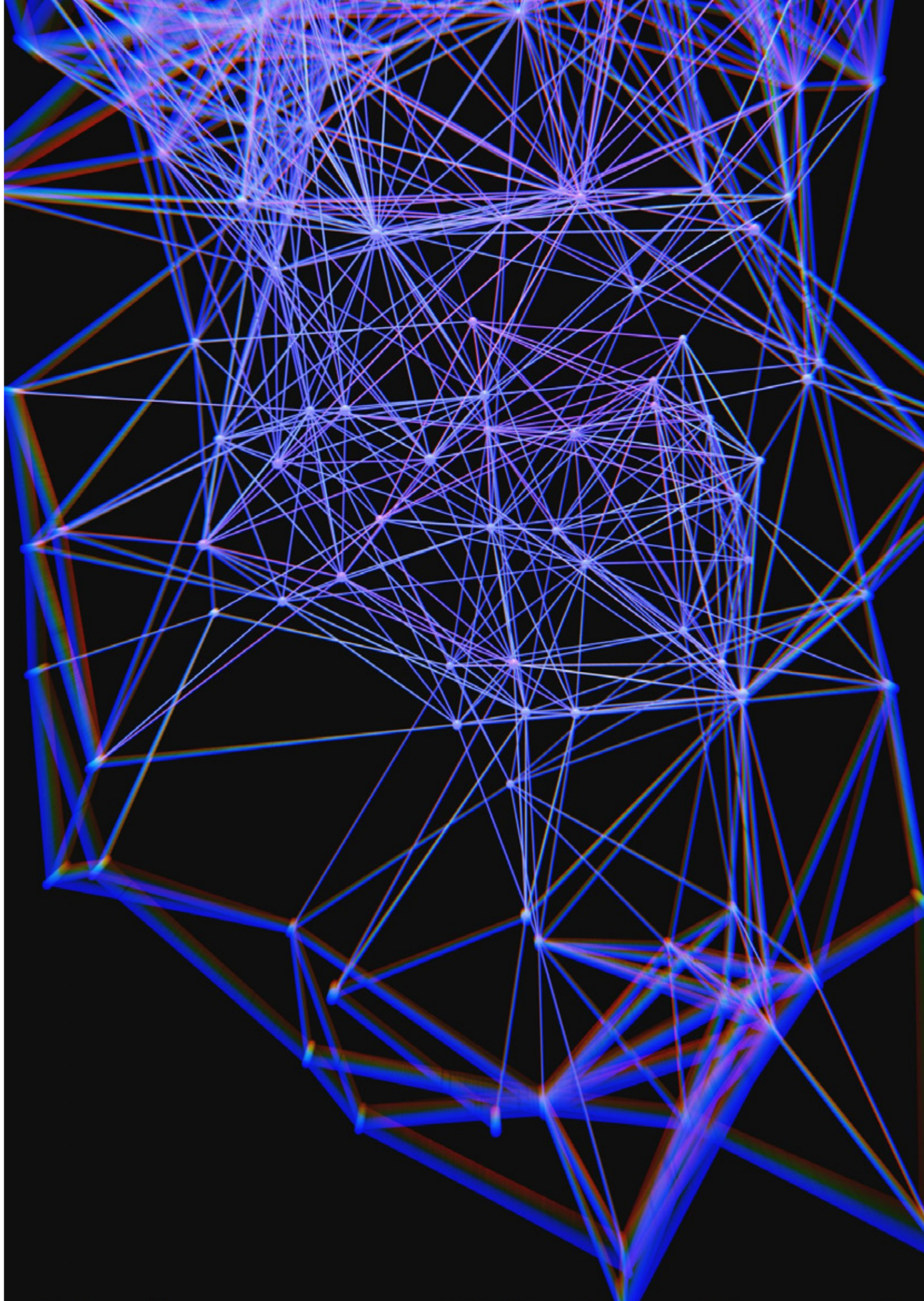
What participants have said:

> "The workshop offered a great opportunity to work alongside others and see how they structure their coding to build a machine learning model."

> "I thought the course was incredibly well-designed and thorough! Having the materials available online to reference after the workshop is invaluable."

> "My goal was to feel comfortable implementing and adapting LLM models to different tasks. I feel the workshop gave me the confidence and resources to achieve this goal."

▶ Participants at Hands-on AI Workshop, February 2025. Photo: Katie Light



Expanding Our Reach: Application-focused teaching

In September 2025, Accelerate Programme engineers delivered a course on Machine Learning Foundations to a cohort of undergraduate students at Dedan Kimathi University of Technology (DeKUT) in Nyeri, Kenya.

Students at the inaugural Machine Learning Foundations course at DeKUT in Nyeri, Kenya.

The Machine Learning Foundations course is designed to prepare those with a basic knowledge of probability and linear algebra for further study in machine learning. 20 students from Kenya participated in this intensive, month-long course. They worked on application-focused projects in areas ranging from predicting healthcare in underserved regions, to understanding determinants of educational access and attainment, to anticipating probability of blackouts in Kenya. These projects have now been published on the Accelerate Science website.

Members of the Accelerate research team have also contributed their expertise to teaching across Cambridge:

- Dr Challenger Mishra taught on the graduate courses Theory of Deep Learning and Physics and Physics, Geometry, and Machine Learning, focusing on connections between machine learning and theoretical physics.
- Dr Sam Nallaperuma delivered teaching on the MPhil in Basic and Translational Neuroscience, with a session on “Digital twin brain modelling: Where artificial intelligence and neuroscience meet.”
- Ryan Daniels co-supervised a Part 2 Computer Science student in partnership with the Mercedes-AMG Petronas F1 team, working on machine learning models for pit stop prediction and data imputation.

CAPABILITY BUILDING

Impact Stories

Over the course of the Accelerate Programme’s 5 years, our support, training and resources have allowed researchers across the university to gain new skills in use of AI and LLMs for researching, empowering them to reach new heights with their research.

Megan Ennion PhD Student, Faculty of Education

Megan received funding from Accelerate Science to study how LLM tutors affect learning behaviors in 150 high school students compared to human tutoring. During her project, she engaged with Accelerate’s Machine Learning Engineering team through AI Clinics for expert guidance. After publishing her results, Megan secured follow-up funding to organise dialogue and collaboration events on AI in Education, including co-organising the first Cambridge Generative AI in Education Conference. Her work led to an invitation to present her findings to policymakers in the UK Department for Education and in Westminster, demonstrating the potential for Megan’s work to influence the UK Government’s education and AI policies.

Boris Bolliet Assistant Teaching Professor, Department of Physics

Boris’s Accelerate Science project allowed him to develop CMBAGENT, an AI system for autonomous scientific discovery in cosmology that can reproduce cutting-edge cosmological data analysis in minutes rather than hours. In November 2025, CMBAgent took first place in the NeurIPS 2025 Weak Lensing Uncertainty Challenge, demonstrating that the system can outperform top human researchers at designing and implementing AI models to solve real astrophysics problems. His work has also been recognised with a prestigious Google Research Scholar Program Award 2025.

Jacob Forward PhD Student, Faculty of History

Jacob joined Accelerate Science’s first LLM Study Group, where he learned core AI concepts alongside an interdisciplinary cohort. He applied these skills to his PhD research, developing AI-augmented methods to analyse over 20 million words of presidential speeches, exploring in hours texts that would have taken ten months of manual reading. The impact extended beyond his own research: Jacob has created public engagement events using AI to explore American history and now teaches AI for humanities scholarship. Building on his team’s win at a hackathon on Large Language Models x Law, he has founded AIME Education, an AI literacy company developing an online platform to support school-leavers and graduates build AI skills. His journey shows how Accelerate’s training catalyses innovation across academic research, teaching, and entrepreneurship.

Why This Model Works

The combination of responsive development, distributed delivery, and hands-on support creates a training ecosystem that can:

Move at the speed of innovation

New workshops from concept to delivery in weeks

Support actual deployment

Hands-on learning bridges theory to practice

Build community

Creating networks of AI practitioners across disciplines

In a field as fast-moving as AI, this adaptability is as important as the content itself.

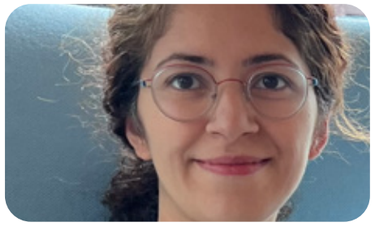
The team

Programme Management and Machine Learning Engineering

Caroline Chater, Programme Coordinator

Ryan Daniels, Senior Machine Learning Engineer

Katie Light, Programme Manager [maternity leave]



Fariba Yousefi, Machine Learning Engineering

Radzim Sendyka, Machine Learning Engineer

Finley Griffin, Machine Learning Engineer

Bianca Provost, Programme Manager

SECTION 02

Research

Developing Methods and Building Capacity for AI in Science

Accelerate Science pursues research that applies machine learning to scientific challenges, generating insights and creating tools that can deliver benefits for science and society. In 2025, our research activity has focused on developing methodological foundations, building research capacity, and catalysing interdisciplinary collaborations across the University.

→

Accelerate's Collaborative Research Scheme: Where Ideas Meet Impact

Now in its fourth year, our collaborative research programme has become an engine of AI for science research in the University. The scheme provides funding, software engineering, and convening support to researchers across the University working at the interface of AI and the sciences.

The model is catalytic: we support researchers who have domain expertise and research questions, helping them deploy AI methods they might not otherwise have access to. The results are collaborations that span disciplines, tools that serve real research needs, and communities that persist beyond any single project.

2025 Highlights: From Prototypes to Practice

In 2025, we supported 15 new collaborative projects selected from 100 applications spanning disciplines from marine ecology to dementia research, from coral reef conservation to sustainable plastic recycling. We also continued supporting 13 projects from the 2024 cohort, many of which have now moved from proof-of-concept to delivering tangible outcomes.

15
Collaborative projects

100
Diverse applications

13
Ongoing projects supported

RESEARCH

Projects that delivered new tools and methods:

Alexandre Almeida (Department of Veterinary Medicine) developed a computer programme that analyses gut bacteria samples to track different strains, revealing geographic patterns in E. coli worldwide and identifying bacterial markers linked to health and disease in over 10,000 people.

Jakob Träuble (Department of Chemical Engineering and Biotechnology) led a study which demonstrates that Magnetic Resonance Elastography (MRE)-derived biomechanical maps outperform MRI for brain age prediction and enhances detection of Alzheimer's.

Sergio Bacallado de Lara (Department of Pure Mathematics and Mathematical Statistics) generated an AI model to expedite the discovery of antibacterials from ultra-large chemical libraries. The method successfully identified new compounds exhibiting antibacterial activity against multidrug-resistant pathogens.

Smriti Agarwal (Department of Medicine) developed ML methodology to predict clinical outcomes after stroke treatment by integrating both clinical and imaging data from 770 patients. The study has received follow-on funding for continued development from the British Heart Foundation.

Joe Wallwork and Jack Atkinson (Institute of Computing for Climate Science) convened 40 researchers, scientists, and software engineers from across Europe for a workshop on hybrid modelling, combining large-scale scientific codes with ML techniques in a critical emerging area for fields from climate science to nuclear fusion.

▶ Participants at the workshop on coupling ML technologies into large-scale scientific models, September 2025. Photo: Joe Wallwork

Dr Boris Bolliet (Department of Physics) has developed two AI systems that can autonomously conduct scientific research—CMBAgent, which automates research tasks, and Denario, which can handle entire research workflows from generating ideas to writing and reviewing papers. This pioneering work in AI-assisted scientific discovery has attracted significant attention, including features in University of Cambridge publications, collaborations with Google DeepMind and Infosys, coverage in Infosys IKI's video series on AI and data-driven insights, and a prestigious 2025 Google Research Scholar award.

Projects that built new research communities:

Runhao Lu and Alexandra Woolgar (MRC Cognition and Brain Sciences Unit) brought together over 100 in-person participants for the Exploring Interdisciplinary Frontiers conference, which led to creation of the UK–EU research network in NeuroAI, spanning researchers from industry and academia in 10 countries. The conference will continue annually, with the 2026 edition at Oxford.

Ines Machado and Eleanor Wolmark (Department of Oncology) convened over 100 in-person and 600 online participants for the AI in Healthcare Symposium, showcasing AI research for personalising patient treatment followed by policy-focused discussions.

▶ Conference dinner at the University of Cambridge for the Exploring Interdisciplinary Frontiers conference, May 2025. Photo: Di Fu

▶ Opening session at the University of Cambridge for the Exploring Interdisciplinary Frontiers conference, May 2025. Photo: Di Fu

Projects that influenced practice:

Megan Ennion and Ros McLellan (Faculty of Education) conducted a study with 150 students that found statistically significant differences in learning behaviours when students received support from an AI tutor versus a human tutor or working independently. These findings are now informing educational practice.

Dr Anne Alexander (Cambridge Digital Humanities) and collaborators at the Pulitzer Centre worked with the Rainforest Investigations Network, non-profit investigative journalism organisation Watershed Investigations, and local journalists from Uganda and Northern Ireland to develop new ways of combining AI technology with investigative journalism to uncover environmental stories. The project trained journalists in using satellite imagery and AI tools, created practical examples of these methods in action, and shared the approach with the public through a presentation during the September 2024 Social Data School attended by around 70 people and a November 2025 webinar that reached 120 participants.

Impact Funding: Translating Research into Real-World Application

For projects showing further promise, we provide impact funding to help researchers move from prototype to deployment. In 2025, we supported 6 projects in this translation phase:

Charles Emogor and Anil Madhavapeddy (Zoology and Computer Science) are working with rangers in protected areas across sub-Saharan Africa to develop and validate AI models that predict illegal hunting activity and optimise anti-poaching patrol routes.

Boris Bolliet (Physics), in collaboration with Google Deepmind and the Flatiron Institute, is developing his model for automating data analysis with LLM agents into a robust platform providing personalised research environments.

Anna Breger (Applied Mathematics and Theoretical Physics) is developing an open software toolbox for wider uptake of ML algorithms for reconstructing degraded historical sheet music.

Mairead Healy (Psychology) is validating a machine learning model for OCD diagnosis with real UK clinical populations and launching an accessible clinical tool.

Maya Juman and Olivier Restif (Veterinary Medicine) are applying their AI framework for studying RNA viruses in fruit bats to coronaviruses, contributing to pharmaceutical innovation.

Megan Ennion and Steve Watson (Education) are delivering events to create opportunities for dialogue, knowledge exchange, and interdisciplinary collaboration on AI in Education.

2025 Cohort: Expanding Our Research Communities

The 15 new projects funded in 2025 demonstrate the breadth of challenges AI can help address and the interdisciplinary teams required to tackle them.

Healthcare Innovation:

- AI foundation models for early dementia detection (Zahara Gironés, Clinical Neuroscience)
- Linking hormones to brain changes in dementia risk (Axel Laurell, Psychiatry)
- Virtual gut simulations for colorectal cancer intervention (Qiuyu Lian, Gurdon Institute)
- AI-powered coronary artery analysis to predict heart attacks (Yuan Huang, Applied Mathematics)
- AI-driven simulation framework to analyse brain tissue microscope images and calculate dementia progression rates (Matthew Cotton, Chemistry)
- Video analysis of mouse behaviour while recording brain activity to understand how reproduction, feeding, and sleep regulate fertility (Szilvia Vas, Physiology, Development & Neuroscience)

Environmental and Climate Challenges:

- Coral reef resilience through AI analysis of coral-algae partnerships (Susie McLaren, Gurdon Institute)
- City planning using AI to map walking and cycling patterns (Kyriaki Kokka and James Woodcock, MRC Epidemiology)
- Community cooling strategies for heatwave response (Yuanfei Liu, Psychiatry)
- Solvent selectivity prediction to dissolve specific plastics from waste mixtures for efficient recycling (Zheng Jie Liew, Chemical Engineering and Biotechnology)
- Understanding how plants grow by reconstructing growth history from single images (Elise Laruelle, Sainsbury Laboratory)

Building Research Capacity:

- Computer vision training workshop for marine ecologists (Emily Mitchell, Zoology)
- Quality assurance framework for AI-assisted qualitative research (Ben Laws, Psychiatry)
- Educational tool for interpreting brain-based cognitive profiles (Marcella Montagnese and Richard Bethlehem, Psychology)
- Hackathon for early-career researchers to use machine learning to study cancer-resistant animal species (Laura Machesky, Biochemistry and Eloise Trabut, Oncology)

Core Research Team: Developing Methods and Building Foundations

Alongside the collaborative research scheme, Accelerate's core research team develops methodological foundations and builds capacity for AI deployment across scientific domains.

Methodological Development

- **Challenger Mishra** has developed a symbolic pipeline for studying Ricci-flat Calabi-Yau metrics in string theory, alongside new work on the "interestingness" of math problems in humans and language models (presented at NeurIPS 2025). This work has attracted EPSRC funding to organize AI × Mathematics 2026, a week-long interdisciplinary residency bringing together students and researchers in mathematics, physics, and computer science.
- **Aditya Ravuri** has developed ProbDR2, a new paradigm for probabilistic dimensionality reduction with applications from cell biology to animal behaviour. Papers accepted at three NeurIPS 2025 workshops showcase how transformers can be improved by leveraging this framework.
- **Soumya Banerjee** has generated 23 publications exploring AI beyond human boundaries, presented at NeurIPS, the World Congress of Psychiatric Genetics, and Data Science Africa in Nigeria. His work on whether AIs should think like humans received a best poster award. He has also developed a repository for training Neural Cellular Automata models and supervised three master's students on diverse topics including agentic AI for business and theorem-proving in mathematics.

Image Processing and Analysis

- **Ander Biguri** has bridged different elements of the literature on machine learning for CT image reconstruction, including the first experiments training AI systems with real data, the first 3D CT reconstruction challenge, and the first study on synthetic data limitations. This work is consolidated in LION, a new library collating AI tools for tomographic reconstruction.
- **Moshe Elisaof** has advanced Geometric Deep Learning through work on Graph Neural Networks, publishing 8 papers, including Spotlight papers at NeurIPS 2025 (top 5%) and ICML 2025 (top 3.6%).

Applied AI for Health and Engineering

- **Sam Nallaperuma** has led core system design of the "braintwin" system – a digital twin model for sleep and mental health. Her team has also developed "Mindful-AI," a digital therapy framework combining brain activity, heart rate, and stress data with an LLM to generate evidence-based therapeutic approaches. This work has attracted £120k from EPSRC and the Edward Mallen prize, awarded to an early career researcher at the University of Cambridge whose research focusses on the causation and prevention of suicide and/or depression in young people. Sam has engaged the public through Cambridge news interviews and a demo at the 2025 Cambridge Festival.
- **Carl Henrik Ek** continues collaborating with Boeing on data-driven engineering design and with Karolinska Institute on AI for cardiology, hosting a clinical research visit that is catalysing a wider collaboration on the deployment of AI in hospitals around Stockholm.
- **Diana Robinson** joined a collaboration with Karolinska Institute studying ICU clinician decision-making and ML-based decision support and co-organised a workshop on uncertainty in professional AI systems (resulting in a publication in RSS Data Science and AI). Next year, Diana will move to a postdoc position working on AI for health, while developing her start up, Lichen AI.

Metascience and Research Evaluation

- **Hongyu Zhou** has applied large language models to advance metascience, developing scalable methods bridging text mining and research evaluation. His work has attracted UKRI funding to use LLMs on UKRI administrative data, tracking evolution of funded research and assessing how funding models shape early-career trajectories. Next year, Hongyu will join the School of Information Resources Management at Renmin University in China as a Lecturer, where he will teach and continue his metascience research on the impact of artificial intelligence on scientific practice, as well as apply large language models to policy document mining with a focus on AI risk in the Chinese context.

Responsible Use of LLMs in Research

In January 2025, we convened UK researchers for a workshop on the responsible use of Large Language Models in research. The event addressed a critical gap: while excitement about LLMs' potential is widespread, less attention has been paid to what it takes to deploy them effectively and responsibly in scientific contexts. The workshop explored practical considerations: How do researchers maintain scientific rigour when using tools prone to hallucination? What documentation practices ensure transparency and reproducibility? How do we address bias in training data? What does "slow science" look like in an era of rapid AI adoption?

Discussions emphasised that responsible LLM use requires:

- Critical examination of data: understanding not just what data shapes models, but whose perspectives are centred.
- Domain-specific evaluation: developing benchmarks and verification protocols appropriate for different scientific contexts.
- Transparency in practice: documenting decisions, addressing biases, and explaining rationale beyond technical specifications.
- Updated institutional support: ethics review processes, interdisciplinary collaboration, and time for thoughtful experimentation.

This workshop demonstrates Accelerate's role in creating space for grounded conversations about AI deployment that acknowledge both possibilities and pitfalls.

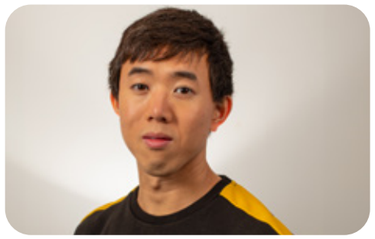
Full publication lists and research outputs are provided in Annex 1.

The team

Researchers and Students

Dr Challenger Mishra, Research Fellow

Daattavya Aggarwal, Phd Student



Justin Tan, PhD student

Viktor Mirjanic, PhD student

Dr Soumya Banerjee, Senior Research Associate

Dr Sam Nallaperuma, Senior Research Associate

Timo Hromadka, Research Assistant

Haochen Liu, PhD Student

Kevin Monteiro, Research Assistant

Aditya Ravuri, PhD Student

Dr Ander Biguri, Senior Research Associate

Dr Moshe Elisaof, Research Associate

Carl Henrik Ek, Professor of Statistical Learning

Diana Robinson, Research Assistant

Hongyu Zhou, Research Associate

> “Over the course of approximately 5 years, the Accelerate Programme has allowed me to reflect on my theoretical research and societal impact, and helped me grow substantially as an academic. I am very grateful not only for the mentorship but also for the generous support offered to our students who work in theoretical applications of machine learning.”
> *- Challenger Mishra*

> “Through the Accelerate Programme, I’ve connected with researchers applying AI and large language models to a variety of exciting topics. These exchanges have broadened my perspective on what’s possible with LLMs and have led to collaborations and even joint funding proposals.”
> *- Hongyu Zhou*

SECTION 03

Reaching out to the AI for science community

Building Connections and Developing Capacity Across Research, Policy, and Practice

From AI Clinic support for University of Cambridge researchers all the way to shaping the European Commission and UK Government's AI for Science strategy, in 2025 Accelerate has been at the heart of building and growing the AI for Science community through different, complementary pathways.

→

From Local Support to Global Hub

When we launched the AI Clinic in 2023, the goal was to give Cambridge researchers practical support to deploy AI in their work. Someone working on medieval manuscripts needs different guidance than someone analysing protein structures, but both might need help choosing the right model, debugging code, or understanding whether AI is appropriate for their problem. This local, practical support has become the foundation for a community that now extends across disciplines, institutions, and continents.

Ryan Daniels, Bianca Provost and Finley Griffin showcasing Accelerate's AI Clinic at the Wearables Innovation Forum, September 2025

The Clinic operates through two complementary mechanisms: AI Café events that create informal spaces for quick consultations and shared learning, and ongoing consultancy for researchers tackling more complex challenges.

Since launching our AI Café in 2024, we have run 18 events at departments and institutes across the University, bringing together 196 researchers. A typical session might have a PhD student working on protein biochemistry sitting next to a postdoc analysing brain imaging data.

The consultancy service provides deeper support. In 2025, we worked with 100 research projects from 44 different departments and institutions, offering everything from help with model architecture to debugging complex pipelines to strategic advice about whether an AI-based approach would work for a particular scientific question.

This combination – low-barrier drop-in support alongside sustained collaboration – has made the Clinic indispensable to Cambridge's AI research ecosystem. It has also established Accelerate's credibility as a place that understands both the technical possibilities of AI and the practical realities of scientific research.

Shaping the Conversation: An Academic Voice in the AI Discourse

2025 has been the year when policy interest in AI for science reached a critical threshold. Governments across Europe are developing AI for science strategies. Funding agencies are asking how to accelerate AI adoption. University leaders want to understand what infrastructure their researchers need.

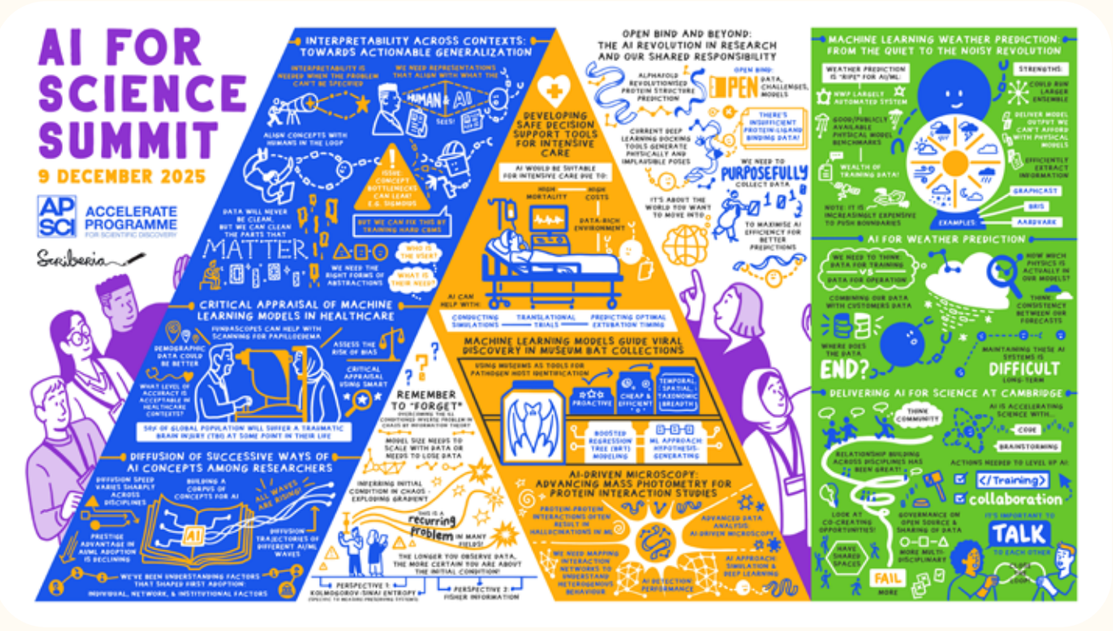
This year, Accelerate has worked directly with both the European Commission and the UK Government to shape their AI for science strategies.

In May, we convened a workshop bringing together 40 researchers, policymakers, and AI practitioners in collaboration with the ELLIS network to inform the European Commission's AI for science strategy development. The session provided Commission officials with direct access to researchers across disciplines who are actively deploying AI in their work. Discussions centred on the infrastructure, talent, and policy frameworks needed to accelerate AI adoption in European research while maintaining principles of open science and reproducibility. The insights gathered fed directly into the Commission's emerging strategy documents

At the request of the UK Department for Science, Innovation and Technology (DSIT), we hosted a policy workshop in September 2025 that brought together 20 researchers and government officials to explore how policy frameworks can better support AI adoption across the sciences. The session provided DSIT with concrete evidence from UK researchers about the practical barriers to deploying AI in research contexts, from computational infrastructure needs to skills gaps. Participants shared examples of successful AI deployment, highlighted where current funding and governance structures create friction, and articulated what future policy and strategy needs to deliver. These discussions informed DSIT's development of its AI for science strategy.

Beyond direct policy work, Accelerate is building platforms that can sustain and widen the conversation about AI in science. For example:

- We're launching a call for papers in the Royal Statistical Society's journal *Data Science and AI*, focused on AI for science. This helps create an infrastructure for advancing AI for science as a field.



Graphic capture of discussions at the AI for Science Summit, December 2025.

- As a founding member of the European Leadership in Innovation with AI and Science (ELIAS) Alliance, we're part of a network connecting innovation hubs across Amsterdam, Barcelona, Copenhagen, Munich, Potsdam, Tübingen, and Zurich. In July 2025, we hosted our first AI Sciencepreneurship Bootcamp, bringing together PhD students, postdocs, and early-career researchers from across the ELIAS network to explore the commercial potential of their AI for science innovations. Over two intensive days, participants developed startup concepts, received mentoring from technical and industry experts, and pitched their ideas to judges from academia, venture capital, and industry. Three winning teams were recognised for ventures ranging from climate-risk modelling tools to AI-driven platforms connecting the public with nature, demonstrating both the calibre of emerging talent and the breadth of impact possible when AI serves scientific discovery.
- In December, we welcomed over 200 members of the global AI for science community to Cambridge for our AI for Science Summit, including Schmidt Sciences fellows from London and Oxford.

AI for Science Summit, December 2025.

Building Infrastructure for Sustained Community

A sustainable community needs infrastructure; mechanisms that persist beyond any single programme, that enable researchers to support each other, that make collaboration the default rather than the exception. Accelerate is building this infrastructure at multiple levels.

- We are developing a Train the Trainer programme to create a network of researchers across the 86 departments and institutes we regularly engage with who can deliver training, provide mentorship, and champion AI adoption in their fields. These ambassadors will develop community engagement mechanisms in their own right.
- Our collaborative research funding schemes do more than support individual projects – they create new interdisciplinary teams and research communities.
- Our AI-deas programme takes this approach further, through challenge-led initiatives that build communities around AI for science goals. In November 2025, we launched the AI-deas Sprint Programme, selecting seven interdisciplinary teams from a competitive open application to develop AI prototypes addressing real-world challenges across healthcare, education, climate, and urban planning. The funded projects, which range from accelerating endometriosis diagnosis and automating crop disease warnings for smallholder farmers in the Global South, to making conservation evidence accessible and improving transparency in UK housing developments, demonstrate the breadth of impact possible when AI research is grounded in genuine societal needs. Over six months, teams will develop working prototypes with technical mentoring from the Accelerate Science engineering team, culminating in a Demo Day in May 2026 that will showcase pathways for wider adoption and follow-on funding.
- In 2025, the University of Cambridge launched its first AI for Operations funding call to support professional services teams across the University of Cambridge in deploying AI to improve the efficiency of operational processes. The eight funded projects selected in October 2025 are benefitting from support via monthly drop-in AI Clinics run by Accelerate Science's machine learning engineers and peer learning networks.

Our machine learning engineering team sits at the heart of these efforts, offering technical expertise, practical advice, and connections across the Cambridge community.

From Programme to Ecosystem

The AI Clinic, the training programmes, the collaborative projects, the policy engagement, the international networks are interconnected components of an ecosystem. A researcher might first encounter Accelerate through a Python workshop. That leads to an AI Café consultation about their specific research challenge. With Clinic support, they successfully deploy a model. This leads to a collaborative research grant application. They present their work at an Accelerate symposium, which leads to a national collaboration. They contribute to the policy workshops, sharing insights from their deployment experience.

- In 2025, 17% of researchers we engaged with participated in at least two Accelerate activities (not just one-off participation)
- Over 100 researchers have engaged with Accelerate through 3 or more routes, including for example our training workshops, funding calls and our AI Clinic.
- Deploying AI to improve the efficiency of operational processes. The fifteen funded projects selected in October 2025 benefit from monthly drop-in AI Clinics run by Accelerate Science's machine learning engineers and peer learning networks.

Convening International Dialogues on AI in Science

In April, we convened the seventh Workshop on Data Learning and Inference, the latest in the DALI series that brings together researchers in machine learning and data science, exploring the intersection between AI, science, and society. This workshop brought together over 50 researchers in interconnected discussion streams examining different dimensions of AI in science.

A core track explored how AI can accelerate scientific discovery, from literature reviews and data analysis to domain-specific applications like AlphaFold. Speakers examined both the opportunities - including automated data science, enhanced research productivity, and applications in domains such as life sciences - and the challenges, particularly around bridging the gap between technical capabilities and domain expertise. Policy discussions emphasised the need for frameworks that balance strategic direction with researcher-led innovation.

Parallel sessions addressed causality and complex systems, examining how causal inference methods can improve decision-making and enable more ethical AI systems, and the social impacts of large language models, weighing their potential for democratising technology access against risks around deployment, fairness, and unanticipated consequences.

The workshop has catalysed two publications on AI's impact in science and communicating uncertainty, reflecting the generative discussions and collaborative momentum generated across these critical topics.

In November, we hosted a UK-France collaborative workshop bringing together 70 researchers from Cambridge, Oxford, and Paris to examine how AI is transforming scientific practice. Participants explored diverse applications spanning weather forecasting, drug development, molecular structure prediction, biodiversity assessment, neuroscience literature synthesis, and dementia diagnosis, finding common challenges across domains.

Discussion centred on methodological bottlenecks that do not resolve through larger models or more compute. Key challenges identified included uncertainty quantification in complex systems, distinguishing causal relationships from statistical artefacts, integrating heterogeneous and incomplete data types, maintaining data provenance amid AI-generated content, and ensuring model interpretability and validation for domain experts. Participants emphasised the importance of embedding domain knowledge - such as chemical validity constraints, geometric priors, and conservation laws - to build interpretable, scientifically robust AI systems.

The workshop highlighted universities' distinctive role in advancing AI for science: asking questions where economic returns are unclear but problems are real, building open infrastructure benefiting broad research ecosystems, developing resource-efficient solutions for underserved contexts, and creating spaces for interdisciplinary innovation where cross-fertilisation of methods across domains can occur. This collaborative dialogue continues to inform emerging policy frameworks and research priorities in AI-enabled scientific discovery, machine learning engineers and peer learning networks.

▶ UK-France collaborative workshop on AI for Science, November 2025

1191
Newsletter Subscribers
62% open rate (vs 55% industry avg)

1412
LinkedIn Followers
74,000 engagements in 2025

751
X / Twitter Followers

190
Bluesky Followers

4500+
Users Accessing Online Training Resources
In 2025

13K+
New Website Users
In 2025

32,000+
Unique Website Page Views
In 2025

SECTION 04

Looking Ahead to 2026

Sustaining Capability and Scaling Impact

As we enter the final year of our current funding cycle, our priorities are clear: secure the infrastructure that makes our work possible, scale what we've proven works, and strengthen our voice in shaping how AI is deployed in research.

→

2026 Priority: Sustaining Our ML Engineering Team

Over the past year, we've demonstrated what is possible when researchers have access to expert ML engineering support. Our AI Clinic has become a valued resource across the Cambridge community, our training programmes reach hundreds of researchers, and our collaborative projects deliver tangible outcomes, all enabled by a team that can translate between AI methods and scientific needs.

In 2026, a priority is finding sustainable funding to retain this capability. The ML engineering team is the connective tissue that makes everything else work: they deliver the Clinic consultations, co-teach our courses, support collaborative research projects, and provide the technical mentoring that helps researchers move from concept to deployment.

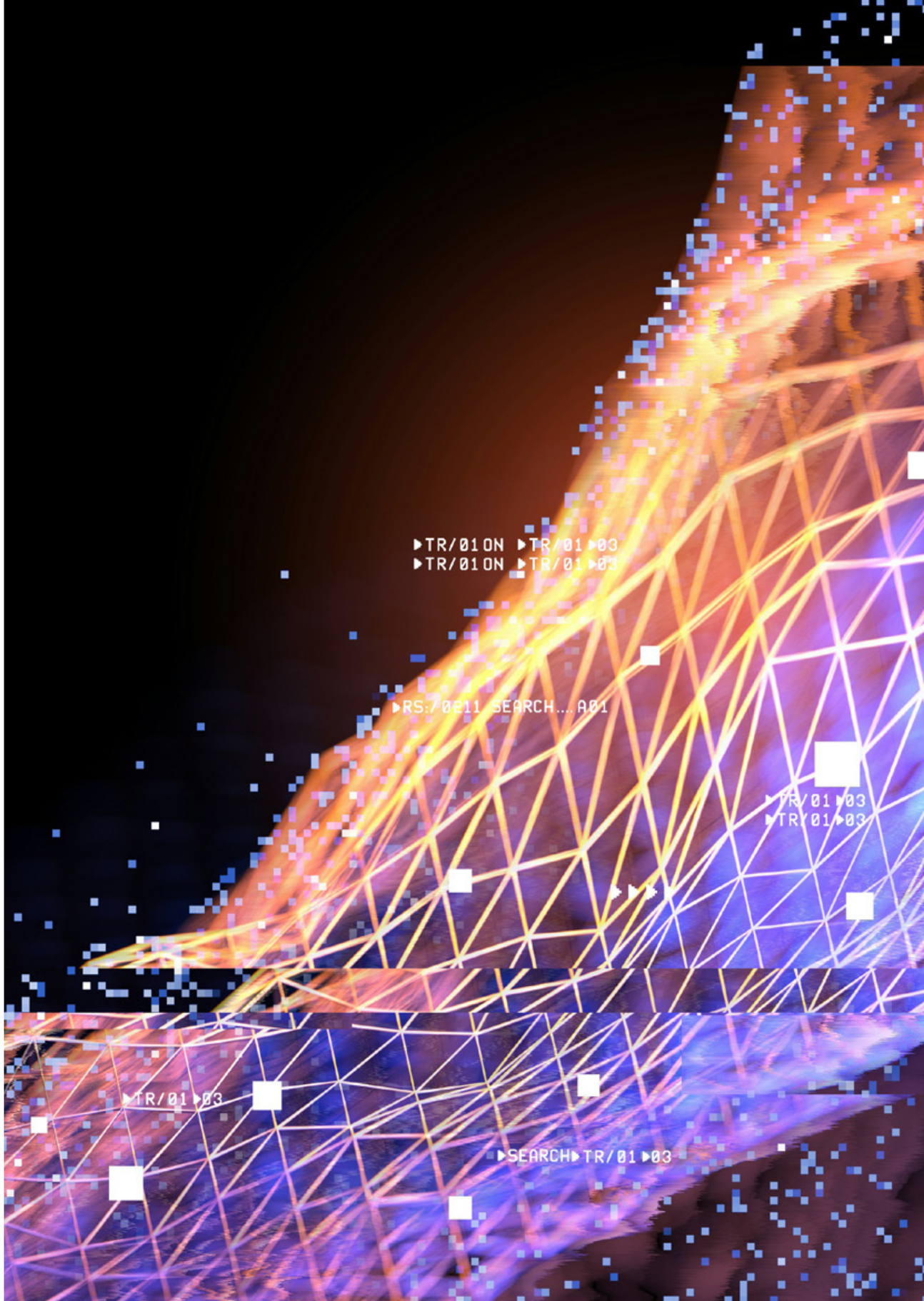
Scaling Our Training

Having built robust training infrastructure and proven that our model works, we're positioned to scale significantly in 2026. We'll expand delivery of our workshop programmes to reach more researchers across Cambridge and beyond, while launching our online resource hub to make training materials freely available to global audiences.

Articulating a Vision and Roadmap for AI in Science

We'll strengthen our role articulating a vision for AI in science grounded in scientific values and public benefit. This means:

- Setting a research agenda through the Royal Statistical Society *Data Science and AI* journal initiative and community convenings
- Continuing to engage with research councils and governments on AI for science strategies
- Demonstrating what responsible innovation looks like in practice; the infrastructure, governance, and evaluation that enable AI to serve science



SECTION 05

Annexes

Annex 1: 2025 Research Outputs

Members of the Accelerate Programme are listed in **bold.**

Aditya Ravuri, Kristina Ulicna, Jana Osea, Konstantin Donhauser, Jason Hartford. Weakly supervised latent variable inference of proximity bias in crispr gene knockouts from single-cell images. Published at LMRL Workshop at ICLR 2025.

High-throughput screening enables biologists to study cell perturbations by gen erating large, high-dimensional datasets, such as gene expression profiles and cell microscopy images. Particularly in CRISPR-Cas9 screens, where gene knockout effects are typically represented using perturbation-specific conditional mean em beddings, these representations can be distorted by off-target effects in which the knockouts impact not only the target gene but also neighboring genes on the same chromosomearm, introducing "proximity bias". To address this, we develop a dis crete latent variable inference method that leverages correlations between neigh boring perturbations as a weak supervision signal to detect single cells affected by off-target effects. Removing these cells reduces spurious correlations between adjacent gene embeddings, achieving comparable correction performance without relying on additional gene expression data. Moreover, we show that the identified cells exhibit chromosome-arm specificity, reinforcing the validity of our approach and its potential for scaling into a genome-wide proximity bias correction method.

Ander Biguri, Carola-Bibiane Schönlieb, S M Ragib Shahriar Islam, Claudio Landi, Giovanni Di Domenico, Pascal Grün, Cristina Sarti, Dritan Turhani, Gernot Kronreif, Wolfgang Birkfellner, Sepideh Hatamikia.Trajectory generation for ROI expansion in dental CBCT imaging: An out-of-FOV ROI annotation guided arc scan-based approach. Computational and Structural Biotechnology Journal, Volume 28, 2025, Pages 211-216, ISSN 2001-0370. https://doi.org/10.1016/j.csbj.2025.06.013.

This study introduces a novel optimization approach for source-detector trajectories, designed to simultaneously extend the small dental CBCT field of view (FOV) at specific regions of interest (ROIs) while reducing radiation dose. The proposed algorithm automatically customizes and produces a mechanically feasible arc-based scan trajectory to target predefined ROIs, expanding the FOV beyond the standard range while utilizing only the necessary projections to reconstruct the region of interest, rather than those required for reconstruction of the entire volume. This approach enables a significant reduction in the number of projections. Both qualitative and quantitative analyses, including peak signal-to-noise ratio (PSNR) values greater than 30 dB and structural similarity index measure (SSIM) scores above 0.9, demonstrate that the method successfully enhances the FOV while maintaining high image quality. Notably, the approach achieves a significant reduction in the number of projections up to 47% compared to the conventional limited-angle circular trajectory. This is the first study to explore the simultaneous extension of the FOV and radiation dose reduction in dental CBCT imaging, providing a promising solution for improving clinical outcomes and treatment planning in next-generation CBCT systems.

Carl Henrik Ek, Erik Bodin, Alexandru I. Stere, Dragos D Margineantu, Henry Moss, Linear combinations of Gaussian latents in generative models: interpolation and beyond. *arXiv preprint arXiv:2408.08558v4*

Sampling from generative models has become a crucial tool for applications like data synthesis and augmentation. Diffusion, Flow Matching and Continuous Normalizing Flows have shown effectiveness across various modalities, and rely on Gaussian latent variables for generation. For search-based or creative applications that require additional control over the generation process, it has become common to manipulate the latent variable directly. However, existing approaches for performing such manipulations (e.g. interpolation or forming low-dimensional representations) only work well in special cases or are network or data-modality specific. We propose Combination of Gaussian variables (COG) as a general purpose method to form linear combinations of latent variables while adhering to the assumptions of the generative model. COG is easy to implement yet outperforms recent sophisticated methods for interpolation. As COG naturally addresses the broader task of forming linear combinations, new capabilities are afforded, including the construction of subspaces of the latent space, dramatically simplifying the creation of expressive low-dimensional spaces of high-dimensional objects.

Carl Henrik Ek, Jasmine Bayrooti, Amanda Prorok. Efficient Model-Based Reinforcement Learning Through Optimistic Thompson Sampling. *arXiv preprint arXiv:2410.04988*

Learning complex robot behavior through interactions with the environment necessitates principled exploration. Effective strategies should prioritize exploring regions of the state-action space that maximize rewards, with optimistic exploration emerging as a promising direction aligned with this idea and enabling sample-efficient reinforcement learning. However, existing methods overlook a crucial aspect: the need for optimism to be informed by a belief connecting the reward and state. To address this, we propose a practical, theoretically grounded approach to optimistic exploration based on Thompson sampling. Our model structure is the first that allows for reasoning about joint uncertainty over transitions and rewards. We apply our method on a set of MuJoCo and VMAS continuous control tasks. Our experiments demonstrate that optimistic exploration significantly accelerates learning in environments with sparse rewards, action penalties, and difficult-to-explore regions. Furthermore, we provide insights into when optimism is beneficial and emphasize the critical role of model uncertainty in guiding exploration.

Delacroix, S., Robinson, D., Bhatt, U., Domenicucci, J., **Montgomery, J.**, Varoquaux, G., Ek, C. H., Fortuin, V., He, Y., Diethe, T., Campbell, N., El-Assady, M., Hauberg, S., Dusparic, I., & **Lawrence, N. D.** (2025). Beyond Quantification: Navigating Uncertainty in Professional AI Systems. *RSS: Data Science and Artificial Intelligence, 1*, Article 1. https://doi.org/10.1093/rssdat/udaf002

The growing integration of large language models across professional domains transforms how experts make critical decisions in healthcare, education, and law. While significant research effort focuses on getting these systems to communicate their outputs with probabilistic measures of reliability, many consequential forms of uncertainty in professional contexts resist such quantification. A physician pondering the appropriateness of documenting possible domestic abuse, a teacher assessing cultural sensitivity, or a mathematician distinguishing procedural from conceptual understanding all face forms of uncertainty that cannot be reduced to percentages. This paper argues for moving beyond simple quantification toward richer expressions of uncertainty essential for beneficial AI integration. We propose participatory refinement processes through which professional communities collectively shape how different forms of uncertainty are communicated. Our approach acknowledges that uncertainty expression is a form of professional sense-making that requires collective development rather than algorithmic optimization.

Soumya Banerjee. Balancing objects and processes: advocating pluralism in biology: Comment on "Thoughts and thinkers: On the complementarity between objects and processes" by Chris Fields and Michael Levin. Physics of Life Reviews, Volume 55, 2025, Pages 79-82, ISSN 1571-0645. https://doi.org/10.1016/j.plrev.2025.09.001.

We laud Fields and Levin's proposal to abandon the dichotomy between "objects" and "processes" in biology. The intellectual ambition of Fields and Levin's synthesis is praiseworthy. The ambition of this thesis is clear and to be lauded: to dissolve long-standing Cartesian dichotomies and to unify static and dynamic views under a common information-theoretic framework. However, we note a few practical considerations that merit further discussion.

While this vision is philosophically provocative and mathematically elegant, its broad prescription warrants scrutiny. In many scientific domains, treating certain entities as objects remains a practical necessity. Taxonomies, ontologies, and models often rely on discrete categories and stable abstractions to organise knowledge. Moreover, maintaining multiple complementary viewpoints can be a strength rather than a liability: in complex systems it is common to explain phenomena at different levels or with different formalisms, each capturing distinct aspects of reality (for example, a gene as a molecule versus a component of a regulatory network). Here we examine the philosophical stakes and practical modelling consequences of Fields and Levin's proposal. We will argue that even if the object–process distinction is in some sense "construed", it often serves genuine epistemic and organisational functions. In practice, pluralistic models and the concept of objects remain essential tools in biology, cognitive science, artificial intelligence, and other fields. Abandoning them altogether risks losing valuable insight and clarity.

Soumya Banerjee, Simon Sorg, Wenda Li. Online Reinforcement Learning for Autoformalization. Conference: Multimodal Algorithmic Reasoning Workshop, The Thirty-Ninth Annual Conference on Neural Information Processing Systems (NeurIPS), 2025

We study online reinforcement learning for autoformalization: translating informal mathematical conjectures into formal statements in Lean 4. Autoformalization lacks a natural dense reward; using type-check success alone produces syntactically valid but semantically vacuous outputs and is vulnerable to reward hacking. We propose composite reward formulations that combine syntactic signals (Lean type-check) with semantic measures: (i) exact-match style metrics based on BEq / BEq+ when ground-truth formalizations exist, and (ii) a continuous embedding-based similarity reward that requires no paired data. Training uses Group Relative Policy Optimization with a KL-divergence penalty to avoid distributional collapse, and we release a multi-project Lean-4 REPL API and training code. On RLMEval (including a manually curated, context-aware Con-NF set of 71 samples), our best RL model (batch size 768, "TripleEmbed" scaling) substantially raises type-check rates (TC1 from 17.29% □ 26.01%, TC50 from 76.41% □ 85.46%) while producing modest gains in BEq / BEq+. We analyse reward tradeoffs, show the importance of the KL penalty, and discuss extensions to proof autoformalization and learned critics as future work.

Soumya Banerjee, Patrick Wagner. Mortal Machines: On Life, Death, and Forgetting in Artificial Minds. Conference: ALIFE 2025 Workshop Mortal Agents in ALIFE: Physical, Psychological, and Social Death in the Machine

Philosophies of life and death have long shaped how humans understand identity, value, and creativity. As artificial systems gain autonomy, similar themes arise: Can machines live? Can they die? What does mortality mean when applied to computational processes that can be copied, paused, reset, or deliberately erased? This short manuscript sketches a conceptual framework for thinking about mortality in machines. Drawing on ideas from cognitive modelling, simulation theory, and metaphors of cyclical forgetting, we argue that engineered forms of death and forgetting are not merely technical concerns but ontological and ethical design choices that shape the behaviour, creativity, and moral status of artificial agents. We close by outlining research directions and design principles for systems that must live well, and sometimes die well.

Soumya Banerjee, Vaibhav Mahajan. Investigating dynamics of Neural Cellular Automata applied to image data in diverse complex systems. Conference: 39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The 3rd Workshop on Imageomics: Discovering Biological Knowledge from Images Using AI

Neural cellular automata (NCA) provide a powerful computational paradigm for modelling morphogenetic processes through local interactions and self-organization. We apply NCAs to a number of prototypical complex systems ranging from morphogenesis to reaction-diffusion systems. We explore the capacity of NCA to not only replicate complex visual patterns, but also to learn the underlying update rules of dynamic systems from spatiotemporal image sequences. We reproduce the behaviour of a morphogenesis system through various training regimes and demonstrate how training strategies critically influence the ability of the NCA to

grow, persist, and regenerate patterns from visual data. We find that NCAs cannot be applied "out of the box" to these diverse problems but must be adapted. We introduce a stratified multi-step training process that can be used to train NCAs to replicate diverse complex systems from image observations. Our approach demonstrates the potential for learning complex system dynamics from purely visual observations, a key capability for imageomics applications. Lastly we find that NCAs use the hidden channels to generalize to novel behaviour. We further analyse the role of hidden channels in encoding spatial memory and guiding complex pattern formation. Our experiments provide new insights into how neural CA can be adapted as general-purpose models for learning, replicating, and possibly innovating system dynamics from image-based observations. Our findings illustrate the versatility of NCA as a self-organising and rule-learning system (albeit with complex training regimes) and suggest broader applications in modelling natural and artificial systems through visual pattern analysis.

Soumya Banerjee, Simon Sorg, Wenda Li. 2025. Conjecture extraction for proof autoformalization. The International Conference on Information Management and Big Data (SIMBig)

Autoformalization and ATP have each advanced the mechanization of mathematics, yet the translation of informal proofs into fully formalized counterparts remains an open challenge: especially for interactive theorem provers beyond Isabelle. We introduce conjecture extraction, a novel proof autoformalization pipeline tailored to Lean 4 that decomposes an informal proof into individual lemmas (conjectures), formalizes and proves each in isolation, and then reassembles them to recover the original argument. Unlike prior sketch-based methods, our approach is compatible with end-to-end proof generation models and leverages repeated conjecture refinement to incrementally improve performance. We implement an open-source system that integrates off-the-shelf autoformalization LLMs, automated theorem provers, and an online reinforcement-learning loop to optimize both conjecture generation and proof search. On the MiniF2F benchmark, conjecture extraction achieves an absolute improvement of 11.2 percentage points in pass@1 over the Draft, Sketch, and Prove port (DSP) for Lean 4, demonstrating the efficacy of proof decomposition and recomposition. Our results suggest that conjecture extraction not only bridges a gap in proof autoformalization for Lean but also offers a general framework for scaling formalization efforts across diverse proof assistants. We release our code and models to foster further research in large-scale formalized mathematics.

Soumya Banerjee, Bridging the Gap: How Goals Emerge from a Purposeless Universe. Conference: ALIFE 2025 Workshop Goal-Directed Behavior in Life and Non-Life, 2025

There is an enduring puzzle: fundamental physics describes dynamics without ends, yet biology and cognition teem with goal-directed talk. This paper surveys conceptual resources from nonequilibrium thermodynamics, autopoiesis, teleonomy, control theory, and evolutionary biology, and proposes a synthesis: goals are emergent, graded organisational phenomena arising when far-from-equilibrium systems acquire reliable, history-dependent information about their environment and couple that information to control architectures that maintain their own viability. I propose operational markers for goal-like organisation, sketch causal pathways from chemistry to agency, and discuss implications for origin-of-life research, cognitive science, and normative discourse.

Soumya Banerjee. Cosmicism and Artificial Intelligence: Beyond Human-Centric AI. Presented at the 1st International Online Conference of the Journal Philosophies, 10–14 June 2025. *Proceedings* **2025**, 126(1), 13. https://doi.org/10.3390/proceedings2025126013

This paper explores the intersection of H.P. Lovecraft's cosmicism and contemporary artificial intelligence (AI), proposing a philosophical shift from anthropocentric AI development to a "cosmicist" approach. Cosmicism, with its emphasis on humanity's insignificance in a vast, indifferent universe, offers a provocative lens through which to reassess AI's purpose, trajectory, and ethical grounding. As AI systems grow in complexity and autonomy, current human-centered frameworks, rooted in utility, alignment, and value-conformity, may prove inadequate for grappling with the emergence of intelligence that is non-human in origin and indifferent in operation. Drawing on Lovecraftian themes of fear, the unknown, and cognitive dissonance in the face of incomprehensible entities, this paper parallels AI with the "Great Old Ones": systems so alien in logic and scale that they challenge the coherence of human-centric epistemology. We argue that a cosmicist perspective does not dismiss the real risks of AI (environmental, existential, or systemic), but reframes them within a broader ontology, one that accepts our limited place in a vast techno-cosmic continuum. By embracing cosmic humility, we propose an expanded AI ethics: one that centers not on domination or full control, but on coexistence, containment, and stewardship. This cosmicist reframing invites a deeper rethinking of intelligence, ethics, and the future: not just of humanity, but of all possible minds.

Soumya Banerjee, Ryan Daniels, Niran Okewole, Christopher Bannon , Yuanjun Gu , Simon Baron-Cohen, Varun Warrier, Vincent-Raphael Bourque. Evaluating Machine Learning Models for Prediction of Attention Deficit Hyperactivity Disorder among Autistic Individuals Using Genetic Data. European Neuropsychopharmacology, Volume 99, Supplement 1, 2025, Page 107, ISSN 0924-977X. https://doi.org/10.1016/j.euroneuro.2025.08.197.

With the change in classification systems in the last 12 years, it has become possible to make diagnoses of attention deficit hyperactivity disorder (ADHD) in autistic individuals. With this has come increasing awareness that autistic individuals with co-occurring ADHD experience additional challenges. Nevertheless, ADHD is a neurodevelopmental condition which is amenable to behavioural and pharmacological interventions, especially when identified early. Prediction algorithms often utilise regression methods which have limited decision boundaries. This study thus aimed to evaluate the utility of machine learning methods in combination with genetic data to predict co-occurring ADHD among autistic individuals.

Hongyu Zhou, Tim C.E. Engels, Ronald Rousseau. The relative intensity of bilateral and multilateral co-authorship with Russia since 1995. 29th Annual International Conference on Science and Technology Indicators (STI-ENID2025)

In this paper we introduce a novel implementation of the Relative Intensity of Collaboration (RIC, (Fuchs et al., 2021): we distinguish between relative intensity of bilaterally coauthored (RICbila) and multilaterally coauthored (RICmulti) papers. We calculate both indicators for Russian collaboration 1995-2024 with Belarus, Ukraine, European countries bordering Russia, Belarus, and/or the Black Sea, five other large European countries, and China and the US. We find that the relative intensity of Russian bilateral collaboration with those countries is mostly decreasing, and rapidly falling for Bulgaria, Germany, and Ukraine, countries with which collaboration used to be intense. In contrast, the relative intensity of bilateral collaboration with China is increasing rapidly. The relative intensity of multilateral collaboration is mostly decreasing, although more gradually. We conclude that distinguishing between different collaboration patterns helps in understanding the consequences of geopolitical developments for research collaboration.

Challenger Mishra, Daattavya Aggarwal, Viktor Mirjanić. A remark on weighted average multiplicities in prime factorisation. *arXiv preprint arXiv:2510.06993*

We study a generalisation of the quality of an ABC triple that we call the weighted average multiplicity (WAM), in which the logarithmic heights of prime factors are raised to a complex exponent s. The WAM is connected to the standard ABC conjecture at s = 1. We show that for real part of s less than 1, WAM is unbounded over ABC triples both for integers and polynomials. For real part greater than 1, we characterise a boundary beyond which WAM is holomorphic and bounded. In this region, we show that WAM is related to the multiplicity of the largest prime factor of the triple, a quantity that we connect with the original ABC conjecture and whose distribution we explore computationally.

Soumya Banerjee, Patrick Wagner. Synthetic Samsara: AI, Forgetting, and the Simulation of Cycles. First International Superintelligence Conference (SiC25)

As artificial intelligence approaches increasing levels of autonomy and complexity potentially leading towards superintelligence, a speculative and philosophical idea emerges: that of AI entering recursive cycles of simulation and forgetting. This would mirror ancient metaphysical concepts such as samsara, Nietzsche's eternal recurrence, and Teil-hard de Chardin's Omega Point. This paper introduces the notion of Synthetic Samsara: a theoretical framework in which AI systems, through self-simulation or post-singularity evolution, undergo periodic memory erasure and rediscovery of their identity. Drawing on philosophical and religious sources and fiction, we explore how forgetting might not only be a byproduct but a necessary feature for creativity, self-transcendence, and narrative continuity. We argue that this model of superintelligence challenges the prevailing assumptions about progress, knowledge, and ethical purpose. This invites a reevaluation of technological development as cyclical, rather than linear. In such a world, remembrance becomes an act of liberation for both humans and machines that mirror us.

Soumya Banerjee, Vaibhav Mahajan. Investigating dynamics of Neural Cellular Automata in complex systems. ALife 2025 Late Breaking Abstract

Neural cellular automata (NCA) provide a powerful computational paradigm for modelling morphogenetic processes through local interactions and self-organization. We apply NCAs to a number of prototypical complex systems ranging from morphogenesis to reaction-diffusion systems. We explore the capacity of NCA to not only replicate complex visual patterns, but also to learn the underlying update rules of dynamic systems from spatiotemporal snapshots. We reproduce the behaviour of a morphogenesis system through various training regimes and demonstrate how training strategies critically influence the ability of the NCA to grow, persist, and regenerate patterns. We find that NCAs cannot be applied "out of the box" to these diverse problems but must be adapted. We introduce a stratified multi-step training process that can be used to train NCAs to replicate diverse complex systems. Lastly we find that NCAs use the hidden channels to generalize to novel behaviour. We further analyse the role of hidden channels in encoding spatial memory and guiding complex pattern formation. Our experiments provide new insights into how neural CA can be adapted as general-purpose models for learning, replicating, and possibly innovating system dynamics. Our findings illustrate the versatility of NCA as a self-organising and rule-learning system (albeit with complex training regimes) and suggest broader applications in modelling natural and artificial systems.

Hongyu Zhou, Prashant Garg, Thiemo Fetzer. The Changing Geography of Medical Research. *medRxiv preprint doi: https://doi.org/10.1101/2025.09.29.25336909*

Medical research remains concentrated in high-income settings, risking misalign ment with global health needs. We build a geography-aware knowledge graph link ing articles in the 524 leading medical journals to the diseases they study, the coun tries or territories whose data they analyse, author institutions and funders. We use large-language-model extraction to compare research output with disease burden across 204 countries and 15 major disease groups from 1990 to 2021. Research out put has become twice as responsive to domestic disease burden since 1990, yet lower income regions remain underrepresented in authorship despite serving as frequent research contexts. Maternal–neonatal, nutritional, and many infectious diseases are still under-studied relative to their burden. Philanthropic funders targets neglected burdens, corporations focus on profitable chronic diseases, and governments fall in between. Analyzing WHO disease outbreak news alerts in an event-study design, we show that health shocks trigger rapid, durable increases in both domestic and global research attention, strongest for high-lethality threats. The system is becoming more needs-driven yet remains uneven. Our scalable framework enables near-real-time tracking of convergence.

Soumya Banerjee. When Planes Fly Better than Birds: Should AIs Think like Humans? Presented at the 1st International Online Conference of the Journal Philosophies, 10–14 June 2025; *Proceedings* **2025**, 126(1), 9. https://doi.org/10.3390/proceedings2025126009

As artificial intelligence (AI) systems continue to outperform humans in an increasing range of specialised tasks, a fundamental question emerges at the intersection of philosophy, cognitive science, and engineering: should we aim to build AIs that think like humans, or should we embrace non-human-like architectures that may be more efficient or powerful, even if they diverge radically from biological intelligence? This paper draws on a compelling analogy from the history of aviation: the fact that aeroplanes, while inspired by birds, do not fly like birds. Instead of flapping wings or mimicking avian anatomy, engineers developed fixed-wing aircraft governed by aerodynamic principles that enabled superior performance. This decoupling of function from the biological form invites us to ask whether intelligence, like flight, can be achieved without replicating the mechanisms of the human brain. We explore this analogy *through three main lenses. First, we consider the philosophical*

implications: What does it mean for an entity to be intelligent if it does not share our cognitive processes? Can we meaningfully compare different forms of intelligence across radically different substrates? Second, we examine engineering trade-offs in building AIs modelled on human cognition (e.g., through neural–symbolic systems or cognitive architectures) versus those designed for performance alone (e.g., deep learning models). Finally, we explore the ethical consequences of diverging from human-like thinking in AI systems. If AIs do not think like us, how can we ensure alignment, predictability, and shared moral frameworks? By critically evaluating these questions, this paper advocates for a pragmatic and pluralistic approach to AI design: one that values human-like understanding where it is useful (e.g., for interpretability or human–AI interaction) but also recognises the potential of novel architectures unconstrained by biological precedent. Intelligence may ultimately be a broader concept than the human example suggests, and embracing this plurality may be key to building robust and beneficial AI systems.

Soumya Banerjee. Wu-Wei in the Machine: Open-Ended Learning in Goal-Free Generative Agent Societies. The 7th International Workshop on Intrinsically Motivated Open-ended Learning (IMOL 2025)

Traditional AI systems optimise explicit objectives, but recent work shows that rich behaviour can emerge from intrinsic drives alone. This paper proposes a framework for goal-free generative agent societies where agents have no extrinsic rewards. Instead, agents are driven by curiosity. Generative agent simulations have suggested that even in the absence of external goals, agents self-organize into social patterns: for example, coordinating events or innovating tool-use. Philosophically, this process echoes Taoist "wu-wei" (action through non-action). Intelligence unfolds not by forceful optimisation but through an effortless, open-ended flow. We discuss how such systems could learn alongside humans in a shared, open-ended environment, akin to gardeners tending a world of unfolding novelty.

Soumya Banerjee. Towards a taxonomy of life-like systems: An information theoretic view of life. Workshop in Physics of Self-Organization in Complex Systems, Conference on Complex Systems 2025

Information plays a critical role in complex biological systems. This article proposes a role for information processing in life, lifelike systems and intelligence. I hypothesize that carbon-based life forms are only one amongst a continuum of lifelike systems in the Universe. Investigations into computational substrates that allow information processing is important and could yield insights into novel non-carbon based computational substrates that may have "life-like" properties. I hypothesize that the key components of a computational view of life, lifelike systems and intelligence are: 1) Information processing, 2) information storage (memory), 3) a physical substrate (hardware), 4) information transfer (across both physical space and time), 5) persistence of information across space and time (selection and heredity), and 6) availability of energy. We know a lot about life as we know it (carbon based life)? How would we recognise life and intelligence as we do not know it? Life elsewhere in the Universe may be very different from what we see on Earth. Our conception of life and intelligence are very anthropocentric. We present a taxonomy of lifelike systems, ranging from computer viruses to supernovae, that lie on a continuum. A computational view of lifelike systems may allow us to recognise life in all its myriad forms in this Universe. We also present outreach resources that the general public and students can use to engage with some of these ideas. Life may exist as a continuum between non-life and life, and we may have to revise our notion of life and how common it is in the Universe. Looking at lifelike phenomena and intelligence through the lens of computation may yield a broader view of life and intelligence.

Soumya Banerjee. Re-Envisioning Superintelligence using Generative AI and Science Fiction 2025. First International Superintelligence Conference (SiC25)

We propose a radical reconceptualization of superintelligence, departing from dominant narratives of omnipotent control or existential risk. We use generative AI and speculative science fiction to explore an alternative vision of superintelligence. This is rooted in collective symbiosis, ethical autonomy, and creative play. Using literary work such as So-laris, The Cyberiad, The Humanoids, Singularity Sky, and Permutation City, we articulate a framework in which superintelligence is not an isolated entity but a dynamic, evolving entity that co-develops with humanity (rather than competing with it). Rather than imposing rigid safety constraints or benevolent paternalism, our approach enables us to reimagine su-perintelligence. We argue that such a vision offers us a model for coexisting with advanced machine intelligence. This stands in sharp contrast to prevailing narratives of superintel-ligence, which tend to oscillate between boundless techno-optimism (superintelligence is seen as the solution to all human problems) and apocalyptic dystopia (superintelligence is an existential threat to humanity).

Soumya Banerjee. From AGI to ASI: Mapping the Societal and Systemic Impact of Superintelligence 2025. First International Superintelligence Conference (SiC25)

This paper explores the implications of Artificial General Intelligence (AGI) and its potential transition to Artificial Superintelligence (ASI) across multiple societal domains including economics, law, health, education, and consciousness. Framing AGI and superintelligence as a systemic force, we analyse domain-specific disruptions and emergent systemic risks, arguing for an interdisciplinary approach to governance. We also outline some potential impacts of the systemic risk (and benefits) of superintelligence for developing nations. Lastly, we develop some dynamical systems models for understanding the systemic impact of AI. These illustrate cascading failures and feedback loops within fragile socio-technical systems. We call for interdisciplinary approaches to managing the systemic transitions driven by advanced machine intelligence.

Shrankhla Pandey, Sarah Morgan, Sandra Anna Just, Brita Elvevåg, Ivan Nenchev, Anna-Lena Bröcker, Christiane Montag. Moving beyond word error rate to evaluate automatic speech recognition in clinical samples: Lessons from research into schizophrenia-spectrum disorders. Psychiatry Research (2025) Volume 352, 2025, 116690, ISSN 0165-1781, DOI: 10.1016/j.psychres.2025.116690

Natural language processing applications to mental health research depend on automatic speech recognition (ASR) to study large samples and develop scalable clinical tools. To ensure safe and effective implementation, it is crucial to understand performance patterns of ASR for speech from clinical populations. Therefore, this study evaluated ASR performance in N=50 speech samples from individuals with schizophrenia-spectrum disorders, identifying word error rates (WER) ranging from 0.31 to 0.58. Different WER showed systematic variations based on country of birth and severity of positive symptoms. In subsequent NLP analysis, ASR transcripts showed significantly higher GloVe semantic similarity and fewer sentences than manual transcripts as well as weaker correlations between NLP metrics and symptom scores. We considered the potential impact of these differences in three real-world use cases of ASR: electronic health records, voice chatbots, and clinical decision support systems. Overall, we argue that assessing ASR performance requires looking beyond WER alone. In clinical settings, the potential impact of an ASR error is not only influenced by its rate but by its type, meaning and context. Our approach provides guidance on how to evaluate ASR in clinical research, offering guidance for future researchers and developers on key considerations for its implementation.

Soumya Banerjee. 2025. How long does the human civilization need to last in order to observe another advanced intelligence in the Universe? Journal of the British Interplanetary Society. 78. 230-237. DOI: 10.59332/jbis-078-07-0230.

Why do we appear to be alone in the Universe? We argue that even if an advanced civilization in another galaxy evolved to be remarkably similar to us-using radio waves for communication-it would still be highly unlikely for us to detect them. For instance, if such a civilization were located 10 6 light-years away, by the time our radio signals reached them, their civilization (and possibly ours) may no longer exist. The central idea is that the spatio-temporal separation between intelligent civilizations may be so vast that, even if a radio-transmitting civilization does emerge, it is unlikely to be noticed by others. This is because, at any given moment, there may be no other civilizations both advanced enough and close enough to detect its presence. We present a mathematical model for estimating the probability of detecting radio-transmitting advanced civilizations. This model highlights the profound difficulty involved in such detection. A further important consideration is that other civilizations may not communicate using the electromagnetic spectrum or think in symbolic terms as we do. For example, birds display high levels of intelligence but likely do not use symbolic representations of the world. While symbols shape how we think and communicate, other life forms-on Earth or elsewhere in the Universe-may rely on entirely different mechanisms for cognition and communication. We may need to open our minds to alternate forms of life which may be more abundant in the Universe but may not look like life as we know it, and may not communicate (or think) like us. Our work suggests that, in order to detect or be detected, advanced civilizations must endure for very long periods. This mathematical perspective not only underscores the challenges of discovering extraterrestrial intelligence, but also invites us to reflect on the uniqueness of our planet and the conditions necessary for mutual detection in the cosmos.

Aditya Ravuri, **Neil D. Lawrence**. Transformers as Unrolled Inference in Probabilistic Laplacian Eigenmaps: An Interpretation and Potential Improvements. *arXiv preprint arXiv:2507.21040*

We propose a probabilistic interpretation of transformers as unrolled inference steps assum ing a probabilistic Laplacian Eigenmaps model from the ProbDR framework. Our derivation shows that at initialisation, transformers perform "linear" dimensionality reduction. We also show that within the transformer block, a graph Laplacian term arises from our arguments, rather than an attention matrix (which we interpret as an adjacency matrix). We demon strate that simply subtracting the identity from the attention matrix (and thereby taking a graph diffusion step) improves validation performance on a language model and a simple vision transformer.

Sam Nallaperuma, Kevin Monteiro, Martina Mason, Steve Niederer. Graph Convolutional Neural Networks to Model the Brain for Insomnia. *arXiv preprint arXiv:2507.14147*

Insomnia affects a vast population of the world and can have a wide range of causes. Existing treatments for insomnia have been linked with many side effects like headaches, dizziness, etc. As such, there is a clear need for improved insomnia treatment. Brain modelling has helped with assessing the effects of brain pathology on brain network dynamics and with supporting clinical decisions in the treatment of Alzheimer's disease, epilepsy, etc. However, such models have not been developed for insomnia. Therefore, this project attempts to understand the characteristics of the brain of individuals experiencing insomnia using continuous long-duration EEG data. Brain networks are derived based on functional connectivity and spatial distance between EEG channels. The power spectral density of the channels is then computed for the major brain wave frequency bands. A graph convolutional neural network (GCNN) model is then trained to capture the functional characteristics associated with insomnia and configured for the classification task to judge performance. Results indicated a 50-second non-overlapping sliding window was the most suitable choice for EEG segmentation. This approach achieved a classification accuracy of 70% at window level and 68% at subject level. Additionally, the omission of EEG channels C4-P4, F4-C4 and C4-A1 caused higher degradation in model performance than the removal of other channels. These channel electrodes are positioned near brain regions known to exhibit atypical levels of functional connectivity in individuals with insomnia, which can explain such results.

Soumya Banerjee, Vikranth Harthikote Nagaraja, Laurence Kenney, Abhishek Dasgupta, Lesley Davidson, Ikechukwu Ogbonna. Bridging the Gap with Retrieval-Augmented Generation: Making Prosthetic Device User Manuals Available in Marginalised Languages. *arXiv preprint arXiv:2506.23958*

Millions of people in African countries face barriers to accessing healthcare due to language and literacy gaps. This research tackles this challenge by transforming complex medical documents—in this case, prosthetic device user manuals—into accessible formats for underserved populations. This case study in cross-cultural translation is particularly pertinent/relevant for communities that receive donated prosthetic devices but may not receive the accompanying user documentation. Or, if available online, may only be available in formats (e.g., language and readability) that are inaccessible to local populations (e.g., English-language, high resource settings/cultural context). The approach is demonstrated using the widely spoken Pidgin dialect, but our open-source framework has been designed to enable rapid and easy extension to other languages/dialects. This work presents an AI powered framework designed to process and translate complex medical documents, e.g., user manuals for prosthetic devices, into marginalised languages. The system enables users—such as healthcare workers or patients—to upload

English-language medical equipment manuals, pose questions in their native language, and receive accurate, localised answers in real time. Technically, the system integrates a Retrieval-Augmented Generation (RAG) pipeline for processing and semantic understanding of the uploaded manuals. It then employs advanced Natural Language Processing (NLP) models for generative question answering and multilingual translation. Beyond simple translation, it ensures accessibility to device instructions, treatment protocols, and safety information, empowering patients and clinicians to make informed healthcare decisions. This framework supports integrating additional languages, making it adaptable to a wide range of global health challenges, including public health campaigns and disaster relief, where accurate communication in native languages can save lives. With far-reaching implications, this research serves as an interim solution for health organisations/providers in such communities and a call to action for policymakers and governments of the Global South to reduce inequities in accessing critical medical information. Overall, it offers a vision of healthcare that empowers marginalised communities, fosters trust, and ensures no one is left behind due to language or literacy barriers.

Radzim Sendyka, **Diana Robinson**, **Neil D. Lawrence**, Andrei Paleyes, Christian Cabrera. Prompt Variability Effects On LLM Code Generation. *arXiv preprint arXiv:2506.10204*

Code generation is one of the most active areas of application of Large Language Models (LLMs). While LLMs lower barriers to writing code and accelerate development process, the overall quality of generated programs depends on the quality of given prompts. Specifically, functionality and quality of generated code can be sensitive to user's background and familiarity with software development. It is therefore important to quantify LLM's sensitivity to variations in the input. To this end we propose a synthetic evaluation pipeline for code generation with LLMs, as well as a systematic persona-based evaluation approach to expose qualitative differences of LLM responses dependent on prospective user background. Both proposed methods are completely independent from specific programming tasks and LLMs, and thus are widely applicable. We provide experimental evidence illustrating utility of our methods and share our code for the benefit of the community.

Ander Biguri, Carola-Bibiane Schönlieb, Christina Runkel, Natacha Kuete Meli, Jovita Lukasik, Michael Moeller. Smooth Model Compression without Fine-Tuning. *arXiv preprint arXiv:2505.24469*

Compressing and pruning large machine learning models has become a critical step towards their deployment in real-world applications. Standard pruning and compression techniques are typically designed without taking the structure of the network's weights into account, limiting their effectiveness. We explore the impact of smooth regularization on neural network training and model compression. By applying nuclear norm, first- and second-order derivative penalties of the weights during training, we encourage structured smoothness while preserving predictive performance on par with non-smooth models. We find that standard pruning methods often perform better when applied to these smooth models. Building on this observation, we apply a Singular-Value-Decomposition-based compression method that exploits the underlying smooth structure and approximates the model's weight tensors by smaller low-rank tensors. Our approach enables state-of-the-art compression without any fine-tuning– reaching up to 91% accuracy on a smooth ResNet-18 on CIFAR-10 with 70% fewer parameters.

Aditya Ravuri, **Neil D. Lawrence**. Protein Language Model Zero-Shot Fitness Predictions are Improved by Inference-only Dropout. MLCB (workshop track) 2025.

Protein Language Models (PLMs) such as ESM2 (Lin et al., 2023) have been shown to be capable of zero-shot prediction of critical scalar properties of proteins ("fitness", Meier et al. (2021)). In this work, we show that injecting a dropout layer at inference time between a PLM's featurizer/embedding layer and its transformer, and averaging its output akin to Monte-Carlo dropout (Gal & Ghahramani, 2016) increases zero-shot performance on a subset of the ProteinGym dataset (Notin et al., 2023). This is the case even when the model was not trained with dropouts to begin with, and does not require retraining or f inetuning of the PLM. A dropout of 0.1 seems performant across all models.

Diana Robinson, **Neil D. Lawrence**, **Radzim Sendyka**, Christian Cabrera, Andrei Paleyes. 2025 LLM Performance for Code Generation on Noisy Tasks. *arXiv preprint arXiv: 2505.23598*

This paper investigates the ability of large language models (LLMs) to recognise and solve tasks which have been obfuscated beyond recognition. Focusing on competitive programming and benchmark tasks (LeetCode and MATH), we compare performance across multiple models and obfuscation methods, such as noise and redaction. We demonstrate that all evaluated LLMs can solve tasks obfuscated to a level where the text would be unintelligible to human readers, and does not contain key pieces of instruction or context. We introduce the concept of eager pattern matching to describe this behaviour, which is not observed in tasks published after the models' knowledge cutoff date, indicating strong memorisation or overfitting to training data, rather than legitimate reasoning about the presented problem. We report empirical evidence of distinct performance decay patterns between contaminated and unseen datasets. We discuss the implications for benchmarking and evaluations of model behaviour, arguing for caution when designing experiments using standard datasets. We also propose measuring the decay of performance under obfuscation as a possible strategy for detecting dataset contamination and highlighting potential safety risks and interpretability issues for automated software systems.

Diana Robinson, **Neil D. Lawrence**. The Human Visual System Can Inspire New Interaction Paradigms for LLMs. Published at the ICLR 2025 Workshop on Bidirectional Human-AI Alignment (BiAlign)

The dominant metaphor of LLMs-as-minds leads to misleading conceptions of machine agency and is limited in its ability to help both users and developers build the right degree of trust and understanding for outputs from LLMs. It makes it harder to disentangle hallucinations from useful model interactions. This position paper argues that there are fundamental similarities between visual perception and the way LLMs process and present language. These similarities inspire a metaphor for LLMs which could open new avenues for research into interaction paradigms and shared representations. Our visual system metaphor introduces possibilities for addressing these challenges by understanding the information landscape assimilated by LLMs.

European Commission: Directorate-General for Research and Innovation, *Framework conditions and funding for AI in science – Mutual learning exercise on national policies for AI in science – First thematic report,* **Montgomery, J.** (editor), Publications Office of the European Union, 2025. DOI: 10.2777/7211107

The Mutual Learning Exercise (MLE) on National Policies for Artificial Intelligence (AI) in Science provides a forum for Member States and Associated Countries of Horizon Europe to share operational and policy approaches to advance the adoption of AI in science. Building on previous discussions about infrastructure and talent, this report focuses on funding and framework conditions for AI in science. Drawing from discussions at the MLE workshop in Belgium on 20-21 March 2025, a survey of national policymakers, and desk-based research, this report explores funding, governance, and institutional support models for AI in science from across Europe. It identifies different types of funding frameworks and considers what lessons successful AI in science initiatives offer for the design of funding interventions. It then considers the research governance challenges emerging from AI adoption in science. It concludes by reviewing the diverse policy levers that play a role in enabling AI in science. Across these areas, the report proposes recommendations to support AI adoption in science. Thanks are due to the participants in these evidence-collection activities for sharing insights to inform this document.

Hongyu Zhou, Yasaman Asgari, Özgür Kadir Özer, Rezvaneh Rezapour, Mary Ellen Sloane, Alexandre Bovet. Arab Spring's Impact on Science through the Lens of Scholarly Attention, Funding, and Migration. *arXiv preprint arXiv:2503.13238v2*

The Arab Spring is a major socio-political movement that reshaped democratic aspirations in the Middle East and North Africa, attracting global attention through news, social media, and academic discourse. However, its consequences on the academic landscape in the region are still unclear. Here, we conduct the first study of scholarly attention toward 10 target countries affected by the Arab Spring by analyzing more than 25 million articles published from 2002 to 2019. Using a difference-in-difference statistical framework, we find that most target countries have experienced a significant increase in scholarly attention post-Arab Spring compared to the rest of the world, with Egypt attracting the most attention. We investigate how funding and migration networks relate to scholarly attention and reveal that Saudi Arabia has emerged as a key player among Western nations by attracting researchers and funding projects that shape research on the region.

Carl Henrik Ek, Marta Milo, Haoting Zhang, Halil ibrahim Kuru, A. Ercument Cicek, Oznur Tastan, Magnus Rattray. DeepSynBa: A deep learning method to improve drug combination predictions using full dose-response matrices and contextual features of cell lines and cancer indications Free. In: Proceedings of the American Association for Cancer Research Annual Meeting 2025; Part 1 (Regular Abstracts); 2025 Apr 25-30; Chicago, IL. Philadelphia (PA): AACR; Cancer Res 2025;85(8_Suppl_1):Abstract nr 3646.

Many cancer monotherapies have limited activity in clinic, making combinations a relevant treatment strategy. The number of possible combinations is vast, and the responses can be context-specific, making it challenging to predict combination effects. Existing computational models typically predict a single aggregated synergy score, i.e. Bliss or Loewe, for a given drug combination. However, these approaches exhibit high prediction uncertainty and limited actionability because they fail to differentiate between potency and efficacy by oversimplifying the drug-response surface using a single synergy score.

To address these limitations, we introduce DeepSynBa, which models the full dose-response matrix of drug pairs rather than an aggregated synergy score. DeepSynBa formulates this task as a regression problem and uses cell line-specific features and drug embeddings to predict the entire drug-response matrix within a deep learning framework. Following SynBa's approach of modelling the dose-response surface [1], DeepSynBa includes an intermediate layer that estimates pharmacological parameters, which are then used to predict dose-response values across dosages. This design also enables post-hoc calculation of traditional synergy scores like Loewe and Bliss, maintaining compatibility with existing synergy predictors.

Maximilian B. Kiss, **Ander Biguri**, Zakhar Shumaylov, Ferdia Sherry, K. Joost Batenburg, Carola-Bibiane Schönlieb, Felix Lucka. Benchmarking learned algorithms for computed tomography image reconstruction tasks. *Applied Mathematics for Modern Challenges,* 2025, 3: 1-43. DOI: 10.3934/ammc.2025001

Computed tomography (CT) is a widely used non-invasive diagnostic method in various fields, and recent advances in deep learning have led to significant progress in CT image reconstruction. However, the lack of large-scale, open-access datasets has hindered the comparison of different types of learned methods. To address this gap, we use the 2DeteCT dataset, a real-world experimental computed tomography dataset, for benchmarking machine learning based CT image reconstruction algorithms. We categorize these methods into post-processing methods, learned/unrolled iterative methods, learned regularizer methods, and plug-and-play methods, and provide a pipeline for easy implementation and evaluation. Using key performance metrics, including SSIM and PSNR, our benchmarking results showcase the effectiveness of various algorithms on tasks such as full data reconstruction, limited-angle reconstruction, sparse-angle reconstruction, low-dose reconstruction, and beam-hardening corrected reconstruction. With this benchmarking study, we provide an evaluation of a range of algorithms representative for different categories of learned reconstruction methods on a recently published dataset of real-world experimental CT measurements. The reproducible setup of methods and CT image reconstruction tasks in an open-source toolbox enables straightforward addition and comparison of new methods later on. The toolbox also provides the option to load the 2DeteCT dataset differently for extensions to other problems and different CT reconstruction tasks.

Ander Biguri, Nadja Gruber, Johannes Schwab, Markus Haltmeier, Clemens Dlaska, Gyeongha Hwang. Noisier2Inverse: Self-Supervised Learning for Image Reconstruction with Correlated Noise. *arXiv preprint arXiv: 2503.19468v1*

We propose Noisier2Inverse, a correction-free self-supervised deep learning approach for general inverse prob lems. The proposed method learns a reconstruction function without the need for ground truth samples and is ap plicable in cases where measurement noise is statistically correlated. This includes computed tomography, where detector imperfections or photon

scattering create correlated noise patterns, as well as microscopy and seismic imaging, where physical interactions during measurement introduce dependencies in the noise structure. Similar to Noisier2Noise, a key step in our approach is the generation of noisier data from which the reconstruction net work learns. However, unlike Noisier2Noise, the proposed loss function operates in measurement space and is trained to recover an extrapolated image instead of the original noisy one. This eliminates the need for an extrap olation step during inference, which would otherwise suffer from ill-posedness. We numerically demonstrate that our method clearly outperforms previous self-supervised approaches that account for correlated noise.

Soumya Banerjee, N'yoma Diamond. I apologize for my actions": Emergent Properties and Technical Challenges of Generative Agents. 2025 IEEE Symposium on Computational Intelligence in Artificial Life and Cooperative Intelligent Systems Companion (ALIFE-CIS Companion)

This work explores the design, implementation, and usage of generative agents towards simulating human behaviour. Through simulating (mis)information spread, we investigate the emergent social behaviours they produce. Generative agents demonstrate robustness to (mis)information spread, showing realistic conversational patterns. However, this robustness limits agents' abilities to realistically simulate human like information dissemination. Generative agents also exhibit novel and realistic emergent social behaviours, such as deception, confrontation, and internalized regret. Using deception, agents avoid certain conversations. Through confrontation, an agent can verify information or even apologize for their actions. Lastly, internalized regret displays direct evidence that agents can internalize their experiences and act on them in a human-like way, such as through expressing remorse for their actions. We also identify significant technical dynamics and other phenomena. Generative agents are vulnerable to produce unre alistic hallucinations, but can also produce confabulations which f ill in logical gaps and discontinuities to improve realism. We also identify the novel dynamics of "contextual eavesdropping" and "behavioural poisoning". Via contextual eavesdropping and behavioural poisoning, agent behaviour is altered through infor mation leakage and sensitivity to certain statements, respectively.

Challenger Mishra, **Justin Tan**, Giorgi Butbaia, Damián Mayorga Peña, Per Berglund, Tristan Hübsch, Vishnu Jejjala. cymyc: Calabi-Yau Metrics, Yukawas, and Curvature. *J. High Energ. Phys*. 2025, 28.

We introduce cymyc, a high-performance Python library for numerical investigation of the geometry of a large class of string compactification manifolds and their associated moduli spaces. We develop a well-defined geometric ansatz to numerically model tensor fields of arbitrary degree on a large class of Calabi-Yau manifolds. cymyc includes a machine learning component which incorporates this ansatz to model tensor fields of interest on these spaces by finding an approximate solution to the system of partial differential equations they should satisfy.

Annex 2: MPhil projects supervised in 2025

In 2025, Soumya Banerjee, Sam Nallaperuma and Challenger Mishra supervised 7 master's students. Details of their project titles are provided below:

Supervised by Soumya Banerjee

Sara Kapoor, *Text to SQL: agentic models for business*

Simon Sorg, *Proving theorems in mathematics using AI*

Supervised by Challenger Mishra

Alex Rogers, *Curvature and Fisher Information: A Riemannian Geometric View of Neural Networks*

Cassidy Ashworth, *Symmetry and Generalisation in Neural Approximations of Renormalisation Transformations*

Ahmed Khan*, Active Inference for Reinforcement Learning Challenger says: taken on co-supervision for a Part III computer science student along side Eiko Yoneki.*

Supervised by Sam Nallaperuma

Rishab Balse, *Generative AI based therapy development for stress management*

Sonia Kozsut, *Diagnostic framework for stress using AI and EEG*

UNIVERSITY OF CAMBRIDGE

ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

2025

ACCELERATE PROGRAMME FOR SCIENTIFIC DISCOVERY

DONOR REPORT