Gavin Leech

g@gleech.org

— Current affiliations —

Head of Research, Dwarkesh Podcast

Founder, Arb Research

Visiting fellow, LCFI, University of Cambridge

Fellow, Foresight Institute

Head of Camp at ESPR

— Some work —

The Scaling Era (2025), Stripe Press.

Ten hard problems in AI (2024), ACM Computing Surveys.

Activation engineering in LLMs (2023), arXiv.

Did masks work? (2022), PNAS.

Collecting all failed replications (2021), Nature blog.

Annual review of breakthroughs in all the sciences

Annual review of AI safety

The ESPR summer camp (2022–).

— Awards —

Fellow, International Strategy Forum, 2024

SIPS commendation for work on the replication crisis, 2023

Open Philanthropy Early-Career Funding Award, 2022

Emergent Ventures grant, 'general career support', 2021

UK Research Innovation Council Studentship: PhD funding, 2019

— Interests —

Getting into things

Forecasting: above superforecaster level

Amateur critic (1100+ book reviews)

Blogger since 2010 (140+ posts)

— Experience —

Head of Research, Dwarkesh Podcast (2025—Now)

Executive producer for the world's only deep podcast.
I lead research for projects and guests, do hiring,
manage a team of 5 editors, and handle tooling (AI integration).

Founder, Arb Research (2021—Now)

Co-leading a research consultancy on AI, forecasting, and other scientific questions for clients including Schmidt Futures, Open Philanthropy, and the Institute for Progress.

Questions answered include: *How does elite education help students?*
Where best to run a vaccine clinical trial?
What ways can AI evals be misleading?
Was Isaac Asimov any good at predicting the future?
When will we get generative AI to work in biology?

Head of Camp, ESPR (2021—Now)

Leading the European Summer Program on Rationality,
an experiential education camp for gifted students.
I do hiring, curriculum, teaching, evals. My classes are on
things like research skills, cosmology, and cultural literacy.
I'm also on the board of FABRIC, the parent organisation.

Volunteer coach, 80,000 Hours (2021—Now)

Advisor for a career advice charity, helping people get into AI safety.

Technical Reviewer, Future of Life Institute (2023—Now)

Reviewing applications for the Vitalik Buterin PhD Fellowship.

Technical Reviewer, PIBBSS (2022—2024)

Reviewing applications for a natural sciences / AI postdoc programme.

Associate, Alignment of Complex Systems (2022—2024)

Comms and concepts at a new research group at Charles University Prague.
Worked on formal theories of neglected problems with AI & other systems.

Strategic Advisor, Mercatus Center (2022—2023)

Led a study of AI economics and immigration; briefed the UK Cabinet Office
with Ronit Kanwar of Schmidt Futures.

Interim Director, Alert (2022—2023)

Setup & executive search, 'Active Longtermist Emergency Response Team'.

Researcher, EpidemicForecasting.org (2020—2022)

Modelled Covid: results in Science, NeurIPS, PNAS. Led a major study from data collection to journal acceptance. `PyMC3, MRP, SEIR, Azure, Seaborn`

Assistant Teacher, University of Bristol (2019—2021)

Solo CS instructor, also trained TAs. Machine learning & algorithms.

Technical Reviewer, AI Safety Camp (2018—2022)

Vetting & interviewing for an early safety research programme.

Data Scientist, AXA Insurance (2016—2019)

Tech lead for a team of 5. Built dozens of ML models (GLMs, GBMs, Transformers) for actuarial pricing, medical risk, & image recognition. Actuarial pricing, image recognition, tooling, etc. Web scraping at scale; migration from SAS to Hadoop, & Hadoop to Azure. `Spark, Tensorflow, bash`

Software developer, Furo (2016)

Developed an Angular SPA for a procurement startup. Sole responsibility for the design & implementation of tendering. TS controllers from C# APIs.
Buzzwords: C# 6, Typescript, Angular, Visual Studio, MS SQL Server

Software developer, Freetobook (2015—2016)

Improved a large legacy accommodation engine: PHP back-end, JS front-end, REST APIs. Solo design of a PCI compliant card storage system. `PHP, LAMP, SQL, Laravel, cURL, Gitlab, New Relic`

Software developer, Stugo (2015)

Jumped into web dev from a cold start. My code was deployed in pharmacies worldwide within 2 weeks. 2nd employee. `C#, WPF, Node.js, MEAN`

Database developer, STAND International (2014–2015)

Built the ops database for an international NGO. Solo greenfield design &

deployment; delivered at zero infrastructure and licence cost.
Buzzwords: SQL, MS Access, Office365, UML, data protection.

Assistant Statistician, The Scottish Government (2014)

Ran official statistics and helped build a data visualisation platform for teachers.
`SAS, SQL, PCA.`

Volunteer, VSO (2012)

International Citizenship Service programme, living and working with locals in Tanzania. Ran workshops on gender and startup grants for vulnerable women.

Intern, Ordinov Inc (2010)

PA at a software company in Beijing. Handled expenses, liaised with foreign clients, looked at inefficiencies in the company.

— Education —

PhD in artificial intelligence, University of Bristol, 2020-2024

BSc in mathematics & statistics, Open University, 2015-2019

MA in econ & philosophy, University of Aberdeen, 2008-2012

Apache Spark Developer, Databricks, 2018

— Publications —

`The Scaling Era` (2025), Stripe Press

`Questionable practices in machine learning` (2024); in review at JMLR.

`Ten Hard Problems in AI` (2024); in review at ACM Computing Surveys.

`Steering Language Models Without Optimization` (2023); arXiv.

`Massively parallel reweighted wake-sleep` (2023); UAI.

`Trees compensate for model misspecification` (2023); arXiv.

`Mass mask wearing reduces Covid transmission` (2022); PNAS.

`Seasonal variation in SARS-CoV-2 transmission in temperate climates` (2022); PLoS Computational Biology.

`Safety Properties of Inductive Logic Programming` (2021); AAAI workshop

`Understanding government interventions against the resurgence of Covid` (2021); Nature Comms.

`Reversals in Psychology` (2021); Nature Blog.

`The effectiveness of government interventions vs COVID-19` (2020); Science.

`Robustness of Nonpharmaceutical Interventions` (2020); spotlight at NeurIPS.

— Software —

TPP: prototype probabilistic programming language using parallel tensors, later completed by Thomas Heap.

Side effects in Gridworlds: Bunch of new Pycolab environments and implementations of DQN and MaxEnt IRL in Tensorflow. Part of a team of 4. Later used in Deepmind papers.

ProlexaPlus: Brought modern language modelling (Flair) into Prolog for some reason.

Masks V Mandates: Policy-grade modern Bayesian workflow in PyMC3. Probabilistic programming for epidemic modelling and effect estimation. End to end with data getters.

Py2HTK: Wee Python wrapper for the Hidden Markov ToolKit.

Project Euler: little computational maths problems done in Haskell. Gimmick is I don't import ∼anything.

Argmin Gravitas: Just a Jekyll site, but a particularly satisfying one.

— Service —

Reviewer. Peer reviewer for PNAS, Nature Scientific Reports, Springer *Machine Learning*, BMJ Global Health, BMC Medicine, AI Safety Camp, PIBBSS, NeurIPS MINT.

Policy advisor. Sep 2021: Briefed the UK Cabinet Office Covid Task Force.

Expert testimony. Briefed the UK Cabinet Office and Department for Trade on AI talent.

Lead organiser. Hosted a small conference, the Bristol Interactive AI Summer School.

— Teaching —

Lead instructor, ESPR.
Course designer and instructor. (Metaphilosophy, metamathematics, meta-science, cultural literacy, speculative cosmology.)

Lead teaching assistant, University of Bristol.
Trained junior TAs for COMS20010: Algorithms 2.

Assistant teacher, University of Bristol.
TA for the fearsome COMS30007: Bayesian Machine Learning.