

RESEARCH PAPERS

Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity,
Becker et al. (2025)

Measuring AI Ability to Complete Long Tasks, *Kwa et al. (2025)*, NeurIPS

HCAST: Human-Calibrated Autonomy Software Tasks, *Rein et al. (2025)*

RE-Bench: Evaluating frontier AI R&D capabilities of language model agents against human experts, *Wijk et al. (2024)*, ICML 2025 (Spotlight)

Resource profile and user guide of the Polygenic Index Repository, *Becker et al. (2021)*, Nature Human Behaviour

PROFESSIONAL EXPERIENCE

METR , <i>Member of Technical Staff</i>	2025 – Present
Qally's , <i>Founder, CEO</i>	2023 – Present
Stealth AI Safety Organization , <i>Project Manager</i>	2023
New York University , <i>Economics Graduate Student</i>	2020 – 2023
FTX Foundation , <i>Program Lead</i>	2022
Global Priorities Institute, University of Oxford , <i>Visiting Researcher</i>	2022
Harvard University, NBER , <i>Predoctoral Fellow</i>	2018 – 2020
University of Bath , <i>Research Assistant</i>	2018
Institute for Fiscal Studies , <i>Communications (and Research) Assistant</i>	2017 – 2018
University of Bristol , <i>Research Assistant</i>	2017 – 2018
Analytics FC , <i>Co-owner, Consultant</i>	2016 – 2018
MessiSeconds, MessiMinutes, The12thMen , <i>Founder, Host, Producer</i>	2011 – 2017

EDUCATION

Harvard University, *Coursework during Predoctoral Fellowship* 2018 – 2020
Details: A (100%, joint top) in Linear Algebra and Real Analysis I (graduate credit).

University of Bristol, *BSc Economics and Econometrics* 2014 – 2018
Details: Top of cohort; first-class honours; thesis prize; highest ever mark in Applied Econometrics.

PERSONAL PROJECTS AND OTHER PROFESSIONAL ACTIVITIES

Presentations: Harvard University, Anthropic, MIT FutureTech, Microsoft Research, OpenAI, Stripe, Constellation, Golden Gate Institute, Foundation for American Innovation, FAR.AI (2025); EU AI Office Workshop on Evaluating General-Purpose AI with Systemic Risk (2024); OptaPro Forum (2017); OptaPro Forum (2016)

Podcasts and radio: Complex Systems, TBPB

Grantmaking: Regrantor, Manifund (2023, 2025)

Refereeing: Journal of the European Economic Association, Journal of Political Economy: Microeconomics

Events: SHELTER Weekend, FTX Fellows Program

Software projects: burden of Long COVID, health data dashboard, financial life web app