Dr Chris Hicks

London, UK | chrishicks@mailbox.org | +44 7938875365
Google scholar | Homepage | Linkedin

Research scientist and technical leader specialising in **machine learning** (ML), **AI security**, and **computer security**. I combine deep expertise in agent-based ML (including generative AI and reinforcement learning) with a strong security background spanning cryptography (esp. key management), identity, systems security, privacy and cyber defence. I have led multi-disciplinary teams, secured significant research funding and **delivered high-impact publications, prototypes and policy-relevant insights** for government, defence and industry partners. I excel at validating new ideas quickly and building teams that deliver high-quality, impactful technical work.

Experience

Principal Research Scientist and Theme Lead

The Alan Turing Institute | October 2022 – Present

Lead a research theme working at the intersection of AI, security and emerging cyber capabilities.

- Secured **£2.8M** funding from a UK defence partner to establish a major research programme on autonomous, AI cyber defence capabilities.
- Built and co-managed a team of 6 full-time researchers, supervised 7 PhD interns and developed a wider community of collaborators, visiting researchers and engineers.
- Delivered 30+ publications with many accepted at top ML and security venues including NeurIPS, TMLR, RLDM, ACM Workshop on AI and Security, AsiaCCS and IEEE Security and Privacy Workshops.
- Identified research opportunities, created roadmaps, validated ideas, prioritised work and managed delivery in close collaboration with government and industry stakeholders.
- Secured an additional **£1.2M+** in funding from UK and US government organisations, including UK's AI Security Institute (AISI) focused on LLM vulnerabilities, reinforcement-learning-enabled exploitation and AI-driven defensive capabilities.
- Co-authored research on LLM poisoning and backdoor vulnerabilities in collaboration with AISI and Anthropic, informing work on frontier-model cyber risks.
- Reviewer for top conferences including ICML, NeurIPS, AAAI, ICLR and AAMAS.

Research Associate: Privacy Enhancing Technologies

The Alan Turing Institute | March 2020 – October 2022

Delivered research on secure system design, privacy and digital identity for the Bill and Melinda Gates Foundation.

- Advised the NHS COVID-19 app data protection impact assessment (DPIA) and featured twice in WIRED on topics including contact tracing and "immunity passports".
- Designed SIMple ID, a privacy-preserving authentication system with hardware-backed security using low-cost feature phones for developing economies. Presented to UIDAI (Government of India).
- Co-founded and organised the Trustworthy Digital Identity Interest Group, **convening 170+ international members** and hosting distinguished speakers.
- Co-awarded funding for projects on:

- Deep RL for systems security (£160k)
- Computer network defence using DRL (£20k)
- Future payment systems (£25k)

Internships, Contracting and Advisory Roles

Various | 2012 – Present
Designed and delivered prototypes including:

- **3D-printer control software** (2013)
- A **mesh-networking** Android application (2014)
- An **embedded authentication system for vehicular networks** (2019)

Provided expert advice for:

- An SME at the intersection of ML, security and human-machine teaming (2023-24).
- BlackHat Europe 2025 as an advisory board member.

Education

Ph.D. Computer Science, Computer Security

University of Birmingham | 2015 – 2020

- Published **3 A-ranked security conference publications** (IEEE Euro S&P, IACR CHES).
- **Reverse engineered** a proprietary cryptographic algorithm from patents and firmware, identifying weaknesses and proposing mitigations.
- Developed **two new cryptographic schemes** providing “unlinkable” authenticated broadcast for secure inter-vehicle communication, using elliptic curve multiplicative properties and anonymous hardware attestation.

M.Eng. Electronic and Software Engineering (First-Class Honours)

University of Birmingham | 2015 – 2020

- **80% overall grade** including an industrial placement year at National Instruments (UK).

Additional Qualifications

Mental Health First Aid (2023) • **Emergency First Aid** at Work (2024) • Level 2 and 3 certificates in **Counselling Skills** (2024-25).

Skills

AI and ML: Machine Learning • Deep Learning • Reinforcement Learning (RL, DRL, MARL) • Generative AI • Agent-based Modelling • ML Security • LLM Security • Model Red-Teaming

Security: Systems Security • Network Security • Cryptography • Provable Security • Secure Systems Design • Identity & Trust • Privacy Enhancing Technologies (PETs) • Threat Modelling

Programming & Tooling: Python (e.g., Numpy, PyTorch, StableBaselines, RLLib, Ray, Wandb, Tensorboard, AutoGen, smolagents, OpenAI API)
Bash • C/C++ • Java • JavaCard • JavaScript • Embedded Systems

Leadership and Delivery: Research Strategy • Public Speaking • Roadmapping • Stakeholder Engagement • Proposal Writing • Project and Team Management