

Matthew J. Cockayne

matthewcockayne2@gmail.com | matt-cockayne.github.io

[LinkedIn: matthew-cockayne-193659199](#) | [ORCID: 0009-0009-8918-3125](#) | [Google Scholar](#)

ABOUT ME

PhD researcher in safe and responsible AI, working on making AI systems interpretable, robust, and equitable. My work includes detecting and mitigating bias, stress-testing models under real-world distribution shift, and building interpretable-by-design architectures that are human-in-the-loop. Applied primarily to skin cancer detection, my methods have delivered state-of-the-art, fairness-improved, and clinically auditable models, published across journals and conferences including an early-accepted (top 9%) MICCAI 2026 paper. I am currently extending this work to cardiac vision-language models and validating clinically. I bring hands-on experience translating research into reproducible, well-documented tools and training materials for interdisciplinary researchers.

SKILLS

Technical Skills

Programming: Python, R, SQL, LaTeX

ML Frameworks: PyTorch, TensorFlow, Keras, scikit-learn, Hugging Face

Frontier Models: Segment Anything Model (SAM/MedSAM), vision-language models (VLMs), zero-shot transfer evaluation, failure mode analysis

Computer Vision: OpenCV, torchvision, TIMM

Data Analysis: Pandas, NumPy, SciPy, statistical modelling

Safe & Responsible AI: Fairlearn, Aequitas, adversarial debiasing, fairness benchmarking, robustness benchmarking, uncertainty quantification

Explainable AI: CAM-methods, SHAP, LIME, LRP, attention visualisation, concept bottleneck models

Development: Git/GitHub, Weights & Biases, TensorBoard

High Performance Computing: HPC cluster management, SLURM job scheduling, batch script optimisation, distributed LLM/VLM training and inference

Certifications: BlueDot Technical AI Safety Course

Research Skills

Systematic literature review and evidence synthesis. Scientific writing, dissemination, and peer review. Technical development and documentation for AI. Bias and fairness auditing. Robustness and uncertainty evaluation. Model interpretability assessment.

EXPERIENCE

Research Assistant - Cardiac Vision-Language Models

Jan 2026-Present

Keele University, School of Computer Science and Mathematics

- Investigating whether cardiac vision-language models learn clinically valid internal representations or instead exploit spurious correlations (scanner artefacts, demographics)
- Validation model representations in collaboration with a practising cardiologist

PhD Researcher - Safe & Responsible AI for Medical Imaging

2023-Present

Keele University, School of Computer Science and Mathematics

- Built a robustness evaluation framework benchmarking transformer architectures under real-world corruptions. Applied to multi-modal skin cancer detection, achieving state-of-the-art performance while quantifying deployment risk via uncertainty estimation.

Code: github.com/xrai keele/DermFormer

- Developed SkinCBM, an interpretable-by-design concept bottleneck model aligned to the 7-point checklist for melanoma, enabling clinicians to inspect and correct intermediate reasoning steps.

Code: github.com/Matt-Cockayne/SkinCBM

- Conducted information analysis of learnt skin lesion concepts, revealing reliance on spurious correlations, and designed sparse polynomial classifiers that preserve full interpretability of decision boundaries (under-review, Image and Vision Computing).

- Built FairCBM, a fairness-aware curriculum learning scheme for concept bottleneck models in dermatology, reducing demographic performance disparities by 44% and improving lowest-group outcomes by 63%. Early accepted (top 9%) at MICCAI 2026.

Code: github.com/Matt-Cockayne/FairCBM

- Evaluated frontier model reliability for skin lesion segmentation under zero-shot transfer, characterising failure modes and developing adaptive selection methods that reduce catastrophic failures.

Code: github.com/xrai keele/Classification-to-Segmentation

- Authored and delivered tutorials on distributed VLM/LLM training and inference on Keele's GreenHPC cluster, supporting large-scale experimentation across research groups.

Code: github.com/Matt-Cockayne/keele-greenhpc-llm

Turing Network Development Research Associate

2021-2022

- Developed educational materials, organised networking events, and set-up hackathons to foster interdisciplinary research across the university as part of a grant from the Alan Turing Institute.

EDUCATION

PhD in Computer Science

2023-2026 (Expected)

Keele University, Keele, UK

Thesis: Responsible AI for Healthcare: Robustness, Interpretability, and Fairness in Skin Lesion Recognition

Viva: Scheduled 3rd July 2026

Funding: Competition-funded Department Studentship

Published 2 first-author papers (Pattern Analysis and Applications journal, AIIH 2025 conference, oral presentation) and 1 first-author paper early accepted (top 9%) at MICCAI 2026. One additional first-author paper under-review at Image and Vision Computing. Co-author on an accepted Discover Computing journal paper, an accepted AIIH 2026 paper and a Machine Vision and Applications paper under review. Developed open-source toolkits MedXAI, SkinCBM, and FairCBM for reproducible research. PGR Representative for 3 consecutive years.

MSc in Artificial Intelligence and Data Science, Distinction

2021-2022

Keele University, Keele, UK

Touchscreen Talent Data Analyst Internship: Developed explainable AI methods for CV-based sentiment analysis in recruitment.

BSc in Physics, First Class Honours

2018-2021

Keele University, Keele, UK

Undergraduate student representative for 3 years.

PUBLICATIONS

Journal Articles

Cockayne, M. J., Ortolani, M., Al-Bander, B. (2025). "DermFormer: Nested Multi-modal Vision Transformers for Robust Skin Cancer Detection." *Pattern Analysis and Applications*, Springer Nature, Volume 28, Article 194.

DOI: [10.1007/s10044-025-01572-0](https://doi.org/10.1007/s10044-025-01572-0)

Cockayne, M. J., Ortolani, M., Al-Bander, B. "Sparse Concept Absence Network for Concept Bottleneck Models." *Image and Vision Computing*, Elsevier. Status: Under Review.

Al-Bander, B., Ortolani, M., Dafaalla, M., **Cockayne, M. J.**, Pergola, G., Mamas, M. A. "Investigating and Mitigating Bias in Cardiovascular Disease Mortality Predictors for Scaling-up AI Fairness." *Discover Computing*, Springer Nature. Status: Accepted.

Guo, S., **Cockayne, M. J.**, Ortolani, M., Al-Hamd, R., Al-Bander, B. "Zero-Shot Crack Segmentation for Structural Health Monitoring Using Class Activation Map Guided Segment Anything Model." *Machine Vision and Applications*. Status: Under Review.

Conference Papers

Cockayne, M. J., Ortolani, M., Al-Bander, B. (2025). "Classification-to-Segmentation: Class Activation Mapping for Zero-Shot Skin Lesion Segmentation." *AIIH 2025*, University of Cambridge. [Oral Presentation] DOI: [10.1007/978-3-032-00656-1_24](https://doi.org/10.1007/978-3-032-00656-1_24)

Cockayne, M. J., Ortolani, M., Al-Bander, B. "Fair Curriculum Learning for Concept Bottleneck Models in Dermatology." *MICCAI 2026*. Status: Early Accepted (Top 9%). Code: github.com/Matt-Cockayne/FairCBM

Raza, A., **Cockayne, M. J.**, Ortolani, M., Hill, J., Al-Bander, B. "ClinAuditAI: Agentic AI for Clinical Guideline Adherence Auditing." *AIIH 2026: Explainability to Accountability*. Status: Accepted.

TEACHING EXPERIENCE

Module Lead - Level 7 Apprenticeship, Image Processing (NHS Partnership)

2026-Present

Keele University, School of Chemical and Physical Sciences

- Leading delivery of the image processing module on Keele's Level 7 technical leadership apprenticeship programme, delivered in partnership with the NHS for practising healthcare and technology professionals

Laboratory Demonstrator

2022-Present

Keele University, School of Computer Science and Mathematics

- Supported undergraduate and master's students across computer science modules for 4 years

- Delivered guest lectures on deep learning, computer vision, and responsible AI

- Mentored students on machine learning and AI ethics research projects

PROFESSIONAL SERVICE

Programme Committee: International Conference on AI in Healthcare (AIIH) 2026

Special Session Organiser: Co-organising a special session on explainability and accountability at AIIH 2026

Peer Review: 1st International Workshop on AI Safety and Security (AI-SS) 2026, Engineering Applications of Artificial Intelligence journal (2024-present)

Leadership: Postgraduate Research Representative, Keele University (2022-2025, 3 years). Student Representative, Physics Programme (2018-2021, 3 years)

Mentorship: Supervised student research on ML and AI ethics, responsible AI and fairness evaluation

Outreach: Data Science Content Creator Internship (2021-2022) - developed AI/data science educational resources.

Dissemination: Departmental research seminars and guest lectures on safe and responsible AI, robustness, fairness-aware learning, and interpretability (2022-present)