MICHAEL J CLARK

\begin{center} Perth, Australia · {NAME}@wassname.org · GitHub · Scholar · ORCID · LinkedIn \end{center}

Summary: ML engineer (8+ yrs) with peer-reviewed publications. Currently principal data scientist at Woodside Energy, Non-Executive Director at Cytophenix. Research interests: representation engineering, model evaluation, steering methods that work when output-level alignment fails.

AI Safety Research

AntiPaSTO: Self-Supervised Steering of Moral Reasoning | arXiv:2601.07473 | GitHub

- Gradient-based representation steering using model's own behavioral consistency as signal
- Works without preference labels; outperforms prompting on out-of-distribution moral dilemmas
- Builds on prior representation alignment work that showed promise but suffered from instability

LLM Ethics Leaderboard | Website

- Evaluation framework for moral reasoning across 12+ frontier models

Eliciting Suppressed Knowledge | GitHub

- Contrastive prompting to surface RLHF-suppressed knowledge (20% improvement on TruthfulQA)

LoRA Lie Detectors | GitHub

- Low-rank adapters for deception detection (87% accuracy on held-out set)

Publications

Ibrahim M, **Clark M**, Castelnau W. "Improving operational efficiency through condition-based monitoring and IoT technologies." *Australian Energy Producers Journal*, 65(2), 2025. DOI: 10.1071/EP24092

Scott NJA, Butler AP, Butler AP, Berg KB, Butler PH, Carr JM, Cook NJ, **Clark MJ**, Anderson NG. "Pilot study to confirm that ovine fat and liver can be distinguished by spectroscopic tissue response on a MARS-CT." *Endocrine Journal*, 57, S421-S422, 2010.

Zeller H, Dufreneix S, **Clark M**, Butler PH, Butler APH, Cook N, Tlustos L. "Charge sharing between pixels in the spectral Medipix2 x-ray detector." *IEEE IVCNZ*, 363-366, 2009.

Berg KB, Carr JM, **Clark MJ**, Cook NJ, Anderson NG, Scott NJ, Butler AP, Butler PH, Butler AP. "Pilot Study to Confirm that Fat and Liver can be Distinguished by Spectroscopic Tissue Response on a MARS-CT." *AIP Conference Proceedings*, 1151(1), 106-109, 2009.

Zainon R, Butler APH, Cook NJ, Butzer JS, Schleich N, De Ruiter N, Tlustos L, **Clark MJ**, Heinz R, Butler PH. "Construction and Operation of the MARS-CT Scanner." University of Canterbury, 2009.

Industry Experience

Woodside Energy — ML SME & Technical Lead | 2023-present

- Lead data science for major Australian energy company, managing team of 4

Cytophenix — Non-Executive Director | 2023-present

- Founding board member of medical AI spinout (Perkins Institute, UWA)
- AI-powered antimicrobial susceptibility testing; TGA/FDA regulatory pathway
- Awarded $1M in CUREator grants (Federal Medical Research Future Fund); raised $1.3M pre-seed (Nov 2025)

Three Springs Technology — Director, Partner | 2019-present

- ML consulting for mining and energy; 15+ projects from research to deployment
- Developed open-source deep learning curriculum for E&P major
- Landgate SPUR grant (2016) for satellite-based water leak detection

ThinkCDS — Technical Director | 2016-2019

- Founded ML consulting firm (merged with Three Springs 2019)
- Point cloud ML, satellite imagery, reinforcement learning for mining

Schlumberger / OMV — Geophysicist | 2011-2017

- Seismic data loading, geophysics workflows; transitioned to ML 2016

Education

MSc Petroleum Geoscience — Victoria University of Wellington | 2013-2014

BSc Physics (1st Class Honours) — University of Canterbury | 2006-2009

Skills

Languages: Python, PyTorch, Transformers, einops, Pandas, NumPy

Infrastructure: Docker, Kubernetes, AWS, Git

Methods: Representation engineering, activation steering, model evaluation, interpretability, experiment design, time-series forecasting

Community

Perth Machine Learning Group (regular speaker). Active contributor on LessWrong and GitHub.