WORK-TRANSFORM-REPEAT
WORKSHOP BERLIN 12 MARCH 2024

Domesticating AI in medical diagnosis

Prof Robin Williams
Director,
Institute for the Study of Science,
Technology and Innovation
The University of Edinburgh

ISSTI
THE INSTITUTE FOR
THE STUDY OF SCIENCE
TECHNOLOGY AND INNOVATION

Pre-script: AI discussions today echo 1980s new technology debate

- Rapid and pervasive adoption of micro-electronics
- Projected technology-driven restructuring of work/economic life
- Predicted outcomes
  - erosion of labour skills and autonomy
  - mass unemployment or 2 hour working week?
- Social scientists called in – *we spent 4 decades investigating the protracted processes through which these technologies were applied, implemented and became embedded in working (and personal) life*
- Outcomes far removed from initial anticipations

BBC Horizon
28 May 1979
The Robots Are Coming!

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Overview

- Present recent findings: Early experiences of developing and implementing AI in medical diagnosis
- Evolutionary understanding based on Sørensen's 1996 Social Learning Perspective
- Implications for how we study emerging technology settings
- Lessons from over three decades of investigating enterprise solutions (e.g. ERP) – the Biography of Artefacts and Practices Perspective

ISSTI
THE INSTITUTE FOR
THE STUDY OF SCIENCE
TECHNOLOGY AND INNOVATION

Early experiences of medical AI
Thanks to Edinburgh research community

- Especially
  Stuart Anderson, Kathrin Cresswell, Andrey Elizondo, Nusa Faric, Stephen Gilbert, Mari Kannelønning, Hajar Mozaffar, David Seibt, Xiao Yang

Nuša Farič, Sue Hinder, Robin Williams, Rishi Ramaesh, Miguel O Bernabeu, Edwin van Beek, Kathrin Cresswell (2023), Early experiences of integrating an artificial intelligence-based diagnostic decision support system into radiology settings: a qualitative study, *Journal of the American Medical Informatics Association,*

Robin Williams, Stuart Anderson, Kathrin Cresswell, Mari Serine Kannelønning, Hajar Mozaffar, Xiao Yang (2024), 'Domesticating AI in medical diagnosis', *Technology in Society,*

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

huge claims around promising emerging technologies? Eg diagnostic AI

- AI achievements => rampant expectations => huge public and private investments – particularly in health (tenfold growth predicted 2021-7); Estimated Global diagnostic AI market USD $1.33 (2024) - 5.6bn (2029)
- UK government policy post-Brexit predicts doubling in size of UK Life Science sector – driven by AI and data
- proliferating partnerships between enthusiastic technologists and medics => start-ups; hungry VC investors create growing list of unicorns (valued over $1bn)
- ***Focus of attention on radiology/diagnostic AI***
  Cf less attention to other potential benefits of Health AI – eg in service optimisation and in targeted population health intervention
- Within this: focus on support for
  eg lung, breast scans; pathology; brain; retina
  –established care pathways/large volume
  –esp. ***large scale screening programmes***
  scope to create viable business model

AI guru Geoff Hinton 2015 "We should stop training radiologists now, it's just completely obvious within five years deep learning is going to do better than radiologists."

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Gulf between promises and practice

Generic overhyped expectations - poorly aligned w emerging prospects

Slow uptake and limited use of AI tools in hospital settings

c.f. frequently claimed laboratory performance of tools exceeding human experts, limited evidence about tool accuracy and productivity in contexts of real-life deployment (needed to fulfil requirements of safety regulators and health service procurement agencies eg UK NICE)

NB current precautionary safety & clinical regulation requires collection of evidence about performance of stabilised artefacts – in contrast to widely circulated visions of real time application of Machine Learning to live clinical data, potentially directly identifying signs not noticed by humans and thereby advancing biomedical research

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

How to navigate huge claims around promising emerging technologies? Eg diagnostic AI

- Risk of uncritically reproducing supplier claims (e.g. Hinton)
- Limited evidence about tool performance (agnotology – funding unbalanced towards tool development c.f. implementation and optimisation eg service enhancement)
- Social scientists are invited to appraise emerging technologies still at embryonic stage? How to do this – cf. await hindsight knowledge and become historians
- Instances of failure and early teething troubles - but avoid snapshot judgements
- Researchers must balance critical NO CAN DO reactions against modernist CAN DO claims

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Social Learning -Domesticating AI

Knut H. Sørensen (1996) `Learning Technology, Constructing Culture. Socio-technical Change as Social Learning' STS Working Paper NTNU

Social learning: encompasses

- collective processes of experimentation, tinkering, sensemaking
- negotiation and conflict as well as knowledge transfer

Sørensen distinguishes different forms:

- ***learning by doing*** - practical activity eg involved in struggle to implement artefact and make it useful in contexts of use
- ***Learning by interacting*** - how local knowledge and experience may be transferred elsewhere;
- ***Learning by regulating*** - how innovation players seek to order other players and interactively establish the "rules of the game"

Domestication - as taming – *bringing in from the wild*:

Our studies highlight dispersed social learning/domestication processes at 2 levels

1) Struggling to get artefacts to be useful in contexts of adoption
2) Developing collective frameworks for appropriate use

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Gathering insights about emerging changes

Bringing together insights from a targeted cluster of independently initiated lines of enquiry – conducted from diverse viewpoints - at different stages and settings of diagnostic AI

- Upstream Biomedical AI research (XY)
- Emerging application of AI in health: electrocardiograms (KC/HM)
- Early roll-out of commercial AI tool for lung MRI scanning (KC/RW)
- Hospital AI procurement strategies: from apps to platform (MK,DS)
- Emerging new models of regulation for AI medical devices (SA/AS/RW))

? starting point for a longer term investigation
A Biography of Medical AI artefacts and practices?

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Early implementation: AI in radiology

- Commercially developed tool being rolled out in UK hospitals;
- scans Veye nodules in chest MRI – can measure size and track changes in size of nodules over time
- Implemented on screen within existing PACS to facilitate usability and avoid breaching information governance
- Limited evidence on performance & cost effectiveness of tool
- NHSX AI Lab funded roll out and evaluation in UK hospitals – to satisfy regulators – NICE – is it safe, effective; cost effective?
- We undertook qualitative evaluation alongside health economics modellers
- Close involvement of elite radiologists from specialised tertiary cancer treatment centre

Issues around how AI tools perform in real-world

- Reliability: tool not good at detecting nodules at edges of scan
- System trained on standard data sets – vendor literature review claims limited genetic variability in signs/outcomes –inner city hospital concerned about validity for Indian and Afro-Caribbean communities. No standards for verifying (e.g. ethnic) fit between training data and target population
- Unstable performance of the tool
  – implementation issues – eg infrastructure and PACS problems;
  - radiologists detect shifting performance when tool upgraded
  NB radiologists exercise forensic scrutiny of tool performance
- Screening strategy: specificity/sensitivity trade-off. Too many false positives could overwhelm radiology service, harming unscreened patients;
- Tool design strategy: focus on one sign of lung cancer c.f. all cancer indications? All lung diseases?

"Getting your algorithm trained on other peoples' data is a huge issue"

An evolving picture

- radiologists learn where they can dependably rely on AI and where tool makes errors – insist they have final responsibility though willing to delegate to machine where it performs well
  *A de facto model of **responsible use of AI** that makes best use of strengths and mitigates weaknesses of human and machine intelligence*
- Second order effects may emerge – reorganisation of radiology service, reconfigurations of expertise (MDTs vs commodification?)?;
  (question – do our findings re elite tertiary centres apply to other parts of health service [district hospital] and countries [USA]?)
- In contrast to widespread portrayals of AI as inscrutable black box that clinicians will defer to/be performed by:
  AI tools are subjected to multiple forms of governance/scrutiny:
  - by health professionals, [radiologists and their societies]
  - by hospital trusts,
  - by health service quality regulators [NICE]
  - by safety [MHRA] regulators
- Continued importance of local clinical validation/data

Emergence of new institutional arrangements for developing, deploying and validating AI

- High costs of data collection, tool development, procurement, accumulating evidence for multiple governance procedures;
  - consideration to post market surveillance infrastructures…
- Cost effectiveness: Software As A Service vendor model is expensive for hospitals : difficult for adopter to create business case – challenge for vendor to cover development and ongoing supply costs
- "Platform model" – platformisation (Seibt et al in prep.)variously conceived: eg intermediaries (Blackford); eg Philips (OEM) to support procurement, implementation/ integration, validation and post-market surveillance of multiple, changing models

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Emergence of new institutional arrangements for developing, deploying and validating AI

- Multiple local solutions from enthusiast/ start-ups.
  Market has pivoted. VC enthusiasm falters. Failures even of unicorns.
- M&A/Exit: US radiology screening company RadNet acquires our developer and prostate & breast cancer and neuro-degeneration
- industrial concentration - one route to multi-tool solutions…– but what shape? some predict handful of companies dominating the field (bigPharma? Health Tech?)
- proprietary models <=> various platform models
  – diverse organisations moving to place themselves at the centre of ecosystem:
  - scanning equipment vendors; eg Cannon
  - PACS suppliers eg Philips
  - service providers; eh RadNet – US radiology service centre
  - platforms/market intermediaries;
  - big IT companies

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Emergence of new institutional arrangements for developing, deploying and validating AI

- C.f. Existing precautionary governance model eg MHRA
  New forms of regulation emerging that address multiple development-operation cycles
  (e.g. FDA "Algorithm Change Protocol.")
  Gilbert S, Anderson S, Daumer M, Li P,Melvin T, Williams R (2023) Learning From Experience and Finding the Right Balance in the Governance of Artificial Intelligence and Digital Health Technologies *J Med Internet Res*
- Model performance varies between target demographic groups and provider sociotechnical configurations and may degrade over time – need for enduring Post-market surveillance
- Various groups developing PMS platforms e.g. American College of Radiologists; NHS; Great Ormond Street

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

An evolving picture of social learning

1) **Coupling AI with clinical settings:** *(learning by doing)*
   e.g. early experimental efforts to create effective models

2) **Embedding, scaling and extending scope** *(learning by doing/)*
   struggle to implement tools in contexts of everyday use
   and demonstrate clinical and economic performance

   Large-scale adoption across multiple settings to collect (learning by interacting)
   evidence needed to satisfy clinical governance and
   procurement (effectiveness and safety) requirements

3) **Developing collective frameworks for** (learning by interacting)
   appropriate implementation and use *(learning by regulating)*

experimentation around institutional arrangements for effectively deploying AI at scope and scale; articulation of new regulatory processes better geared to incremental improvement; shift from upstream precautionary assessment to ongoing post-market surveillance

PLUS Rapidly changing landscape – *actors, relations between them, boundaries change*

Cost etc challenges in sustainably supplying/validating discrete tools;
VC drying-up start-up (even unicorn) failures; exit/M&A - industrial concentration;
platformised models for deployment and validation of multiple tools

ISSTI
THE INSTITUTE FOR
THE STUDY OF SCIENCE
TECHNOLOGY AND INNOVATION

Drawing broader lessons – *potential pitfalls*

Benefits of cross-sector enquiry – but also potential pitfalls.

- Health technology – significant risks as well as benefits
- Strictly regulated setting
- Medical institutions – deep historical roots, not liable to quickly cede authority to claims of technology communities

Strong homologies – but avoid making mechanistic extrapolation to other sectors and settings – however we can make reasoned judgements about relevance of other settings

ISSTI
THE INSTITUTE FOR
THE STUDY OF SCIENCE
TECHNOLOGY AND INNOVATION

Drawing broader lessons – *potential pitfalls*

- Eg BoAP ERP studies - examine how algorithms travel –
  Glaser, V. L., Pollock, N., & D'Adderio, L. (2021). The Biography of an Algorithm: Performing algorithmic technologies in organizations. *Organization Theory*,
- but note huge reconfigurability of ML-based systems – complex assemblages: particular configurations of datasets; standards; physical infrastructure; scanners; metrics
- Need to analyse AI as information infrastructure? *Innovation in Information Infrastructures* international workshop Barcelona, 16th – 18th September 2024
- LLMs – game changer

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

How to position critical social science (Science and Technology Studies) research in this setting?

- How to study settings where compelling visions and predictions informed by powerful interests
- Sensitivity to local contexts (value of local ethnography in unpicking nuance of process/outcomes in particular settings
- Reflecting upon an unfolding array of local developments in a rapidly evolving setting – we are asked to make assessment prior to hindsight
- Attend to longer history and broader context (proposed historical sociology of computer-aided mammography)
- Current studies as part of multi-site, extended enquiry – opening steps in a Biography of Artefacts and Practices investigation (Hyysallo, Pollock and Williams 2019)

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Lessons from extended enquiry into enterprise systems – premature judgement

- 1990 sustained implementation difficulties/failures with packaged enterprise solutions
- 1990 expert consensus: *highly unlikely that generic packaged solutions could meet the needs of diverse organisations and sectors* (Waterlow 1990).
- 1992 a little known Waldorf software house launches client server based version of its enterprise applications suite
- SAP R3 sells to European branch plant cos; extends market from engineering to multiple sectors and becomes the world's largest enterprise resource planning software company - used eg by all the Fortune 1000 companies

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Lessons from extended enquiry into enterprise systems

Insights from extended investigation 1980 - date

- Need to go beyond single workplace ethnography -multi-site ethnographies and longitudinal study
- Encompassing diverse viewpoint; e.g.:
  - Settings of development and use and nexuses linking these
  - Technology support/maintenance as well as supply
  - Comparing early start-ups with established players
  - New actors (eg role of Gartner in enabling operation of market for software)

=> Biography of Artefacts and Practices (BoAP) perspective

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Biography of Artefacts & Practices perspective

Sampsa Hyysalo, Neil Pollock, Robin Williams (2019) Method Matters in the Social Study of Technology: Investigating the Biographies of Artifacts and Practices, *Science and Technology Studies*,

- Performativity of methodology – reflect upon how research design choices affect what may or may not be discovered
- Long-term development and evolution of infrastructures/systems and practices of modern society in contrast to short durations of PhD/funded research projects
- Risks of incorrect extrapolation eg from design to outcomes/ from early to mature products/from early to embedded implementation
- Benefits of extended enquiry – across multiple settings and moments
- Research is a **team task** – multiple discrete studies that contribute to BoAP investigation
- Linking together multiple independently initiated studies
- Various extension Strategies (intensification, extended journey), Strategic opportunism

ISSTI
THE INSTITUTE FOR THE STUDY OF SCIENCE TECHNOLOGY AND INNOVATION

Theoretical-empirical stance

- 1980s Heavy duty global theory (labor process; informatization)
- Balancing bigger visions with detailed empirical investigation
- Distinctive value of work that is
  - Informed by critical theoretical perspectives
  - Fiercely empirical
  - Interdisciplinary, practice oriented
  - Mid-range theory – concepts and methodologies for engaging matters of concern
- Value of mid-range theory. - eg information infrastructures (infrastructural perspectives on AI/digital transformation; AI as a complex assemblage)