Science for Manufacturing Week 6

conversation with…'

Kendra Briken & Prof Dora Scholarios
athclyde Business School
pt of Work, Employment and Organisation

I4 in manufacturing

Can AI ethics be embedded into the innovation lifecycle?

Why we are talking about AI ethics: AI became more potent

Moore's Law was originally derived from an observation by Gordon Moore, the co-founder of Fairchild Semiconductor and later the co-founder and CEO of Intel.

number of components in a dense integrated circuit (i.e., transistors, resistors, diodes, or capacitors) would be doubling occurring every two years – and so would computational capacity

Why we are talking about AI ethics: AI became more autonomous

AI Now Institute (2019)
Discriminating Systems: Gender, Race and Power

Stanford students launch new AI ethics journal

Courtesy of Ruth Starkman

By Adri Kornfein
March 2, 2022, 12:08 a.m.

The New York Times

Is Ethical A.I. Even Possible?

https://www.nytimes.com/2019/03/01/business/ethics-artificial-intelligence.html

Why are we talking about AI ethics?

Equality Act from 2006 defines protected characteristics

- ✓**age**
- ✓**disability**
- ✓**gender reassignment**
- ✓**marriage and civil partnership**
- ✓**pregnancy and maternity**
- ✓**race**
- ✓**religion or belief**
- ✓**sex**
- ✓**sexual orientation**

Regulation as solution?
A route to regulation that reflects the ambition of the UK AI Strategy

‘By virtue of their ability to develop and operate independently of human control, and to make decisions with moral and legal consequences, AI systems present a uniform set of general regulatory and legal challenges concerning agency, causation, accountability and control.

At the same time, the specific regulatory questions posed by AI systems vary considerably across the different domains and industries in which they might be deployed.’

(https://www.adalovelaceinstitute.org/report/regulate-innovate/)

Change to Innovate: Towards AI mindsets – and AI ethics cultures?

Digital Manufacturing/Industry 4.0 is the introduction of technology & management practices **with a shift in mindset**. This shift in mindset can be difficult to achieve and the failure to do it correctly can have a huge cost to the individual business and the national economy.

Positively manufacturing businesses, recognise this with *75%* of Leaders *anticipate a level of staff resistance, or apprehension*, in relation to the changes that I4.0 adoption may bring. (…)

SMAS Industry 4.0 Review findings (https://www.scottish-enterprise.com/support-for-businesses/develop-products-and-services/support-for-manufacturers)

How to regulate ‘second order’ issues, such as decent and fair work?

It is important to understand the potential human and social impacts of AI algorithm-enabled decision making and automation on manufacturing workplaces and the impact on workers’ voice, the decline of human connection to AI-based decision making, and power and surveillance at work (Moore et al 2018, Zuboff 2019). Trade Unions collectively pledge for dignity at work (TUC 2021).

Research focussing on I4 more specifically suggests that AI-driven systems are particularly impactful on autonomy and discretion. Butollo et al.’s (2018) extensive study in German I4 manufacturing concludes that there is little to no integration of machine and data-based analysis with human competence development and work experience in the design of high-performance management systems.

They conclude that although algorithmic control might relieve employees from stressful decision making and improve health and wellbeing, these benefits could also be outweighed by the loss in voice and intrinsic task motivation. Alongside potential benefits, employees lose control over how much of their tacit knowledge, previously embodied in their job-related knowledge and experience and mostly invisible to their employer, necessarily is now recorded and used for decision making. AI-enabled technologies applied to work, thus, extract new (bodily) knowledge, shift the nature of how skills are developed and used, and introduce new forms of management control which extract situated knowledge owned by the worker, which is often crucial for resistance and collective struggle (Briken 2020).

What is Ethics about?

- **Ethics is a 'Should' be approach**: A branch of philosophy concerned with ways of thinking about morality, and moral judgment.
- **Ethics frames the good the bad – and the ugly**: asks basic questions about the good life, about what is better and worse, about whether there is any objective right and wrong, and about how we know it, if there is.
- **Ethics impacts on Morality** – Human conduct and character referring to "those acts which it makes sense to describe as right or wrong, good or bad."

Core areas of concern

- Honesty, openness, responsiveness, accountability, due diligence, and fairness are core ethical principles.
- There is no general or global consensus for defining ethical behaviour for individuals, as it may change from time to time and from one place to another place.
- There is no universal measure or standard as to what constitutes ethical behaviour.

Utilitarism won't save us…

Strands of ethics

- Descriptive ethics: What do people think is right?
- Meta-ethics: What does "right" even mean?
- Normative (prescriptive) ethics: How should people act?

-- > Applied ethics: How do we take moral knowledge and put it into practice?

The idea:

Bridging AI ethics, responsible innovation, and human labour issues

Observation 1: Digital Transformation in Manufacturing led to discussions and frameworks on AI ethics (i.e. Aletheia, Rolls Royce)

Observation 2: Employees seem to resist, or are opposed to digital change – are their concerns related/relatable to ethical issues?

Question 1 : How do different stakeholders in the innovation lifecycle perceive ethical issues? What are the underlying idea that drive engineering and data sciences?

Core Issue: AI ethics as Process

Dimension 1: Perspectives across the Innovation Lifecycle

How do AI ethical frameworks impact on innovation and workplaces in Industry 4.0?

- When are ethical concerns considered? During conception, prototyping, deployment?
- When should AI ethical concerns be considered?

Dimension 2: Perspectives throughout the stack

Industry 4 innovation brings together expertise from multiple partners throughout the stack - users, coders, service providers, libraries, compilers and hardware

- How do ethical considerations vary amongst partners, disciplines and professional background?
- Do AI ethical frameworks impact differently throughout the stack?

Towards Industry Best Practice?

Aletheia Framwork, Rolls Royce

Developed by Rolls Royce, first version published in 2020,

Open Source Tool for AI ethics checks available since December 2021.

Three main areas covered

- Accuracy & Trust
- Governance
- **Social Impact (second order?)**

Social Impact

Benefits	AI and robotics shall be seen as delivering good. Doing good is one of the five key ethical principles of the EU guidelines for ethical AI. Good includes commercial prosperity.	1	Deployment of AI and robotics shall be shown to improve the well-being of employees and/or the general public, such as improved safety, working conditions, job satisfaction.
		2	Additional to 1. (or instead of), deployment of AI and robotics shall be supported by a business case that demonstrates it improves competitiveness and is not just 'AI for the sake of AI'. The business case should include a calculation of energy consumed in the creation and forecasting running of the AI system.
Human impact	AI systems should be used to enhance positive social change and enhance sustainability.	3	For any deployments, it shall be clear where the human boundary/interface/ interaction is with the AI/Analytics/Robotics system; and any negative/positive impact on human factors and/or human behaviours is fully understood and mitigated where necessary.
		4	Early analysis, in conjunction with human resources and employees (or their representatives), shall be undertaken to identify potential job role changes or potential human resource impacts and the opportunities for retraining or redeployment.
		5	Potential for upskilling opportunities or redeployment shall be explored with human resources and employees (or their representatives) when any impact on affected employees is established, to ensure that the organisation has the key capabilities neewded to secure emerging opportunities in AI and robotics.
		6	Analysis shall be undertaken to assess the impact of the deployment on the supply chain – particularly assessing the likelihood for the technology to have a negative impact on the sustainability of any elements of the supply chain. The same assessment should be afforded to customers as appropriate.
		7	Where there is potential for negative impact on the sustainability of the supply chain, this shall be discussed with the supply chain partner as soon as possible to give them maximum opportunity to adapt to remain sustainable. This same opportunity should be afforded to customers as appropriate.
Communication	Knowledge of the human interactions with AI should be provided by key stakeholders.	8	Frequent communication and discussion should be had with all key stakeholders – in particular employees and employee representatives – through a variety of channels.
Loss of skills	AI systems should be used to enhance positive social change and enhance sustainability.	9	Analysis shall be undertaken as to whether any loss/reduction of skills (which result/ cannot be avoided) needs to be sustained, for the good of the business, and how this would be addressed.

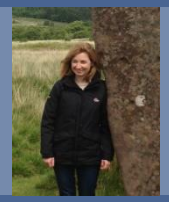
No such thing as perfect solution to AI ethics

‘The Aletheia Framework hasn’t solved all its challenges, it can help reassure organisations, people and communities that the ethical implications of an AI have been fully considered; it is as **fair** as possible; and makes **trustworthy** decisions.’

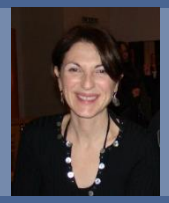
https://www.rolls-royce.com/sustainability/ethics-and-compliance/the-aletheia-framework.aspx

The Research Team -
Bridging disciplines: Engineering, Law, Sociology, Psychology

Dr Kendra Briken
Strathclyde Business School
Dept of Work, Employment & Organisation (Sociology)



Dr Emily Rose
Law School
University of Strathclyde



Prof Dora Scholarios
Strathclyde Business School
Dept of Work, Employment & Organisation (Psychology)

Prof Andrew Sherlock
Director of Data-driven Manufacturing,
National Manufacturing Institute Scotland

Dr Jed Moore
Strathclyde Business School
Dept of Work, Employment & Organisation (Sociology & Engineering)

Industrial Partners

Programme of interviews with relevant partner employees during Q1&2 2022

Community of Interest formed for discussion of AI ethics and its impacts on the workplace

Discussions with partners about a methodology for a future, larger scale, project to gather in-depth data of AI ethics in action

Dissemination of findings via reports and events

University of Strathclyde

NMIS
National Manufacturing Institute Scotland

ANACONDA.

aws

Rolls-Royce

codeplay

dtx Digital Transformation CoLab

Potential Partners

If you are interested in participating contact

kendra.briken@strath.ac.uk