

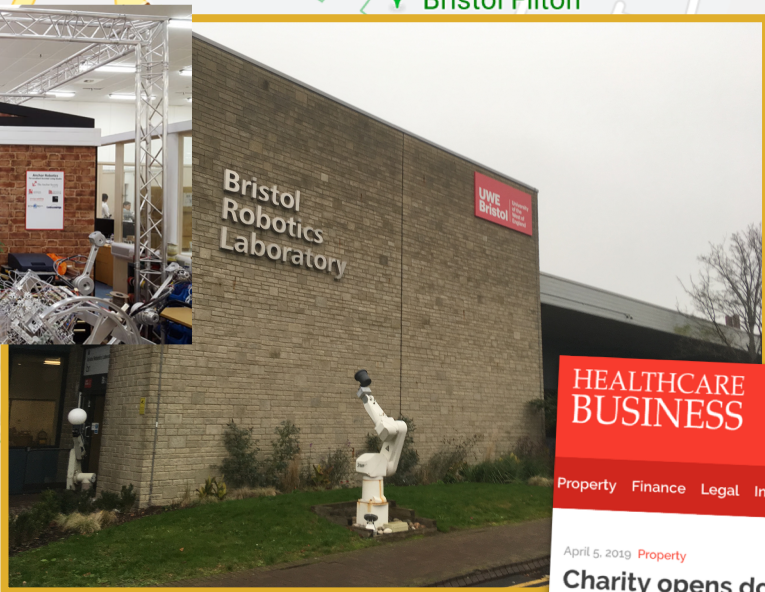
# Ethical standards for robotics: exploring post-automation potentials

Cian O'Donovan Department of Science & Technology Studies, UCL  
c.o'donovan@ucl.ac.uk | @cian | [D/L slides: cianodonovan.com](http://D/L%20slides:cianodonovan.com)

Post-automation? Exploring Democratic Alternatives to Industry 4.0. An International Research Symposium | SPRU – Science Policy Research Unit, University of Sussex  
11<sup>th</sup>-13<sup>th</sup> September, 2019



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 788359



Innovate UK

The ExtraCare Charitable Trust

brl Bristol Robotics Laboratory

HEALTHCARE BUSINESS  
Property Finance Legal Insurance Products Movers & Shakers Mon

### Charity opens doors to Innovation Apartment for Stoke Gifford's older people



Residents from Frenchay, Stoke Gifford and surrounding areas are being invited to visit a new innovation apartment which has opened at Stoke Gifford Retirement Village. The apartment will showcase the use of assistive technologies, including Pepper the robot, and will demonstrate how smart devices can support



Open Bionics  
Bristol Robotics Laboratory  
Students' Union

University of the West of England  
UWE Bristol



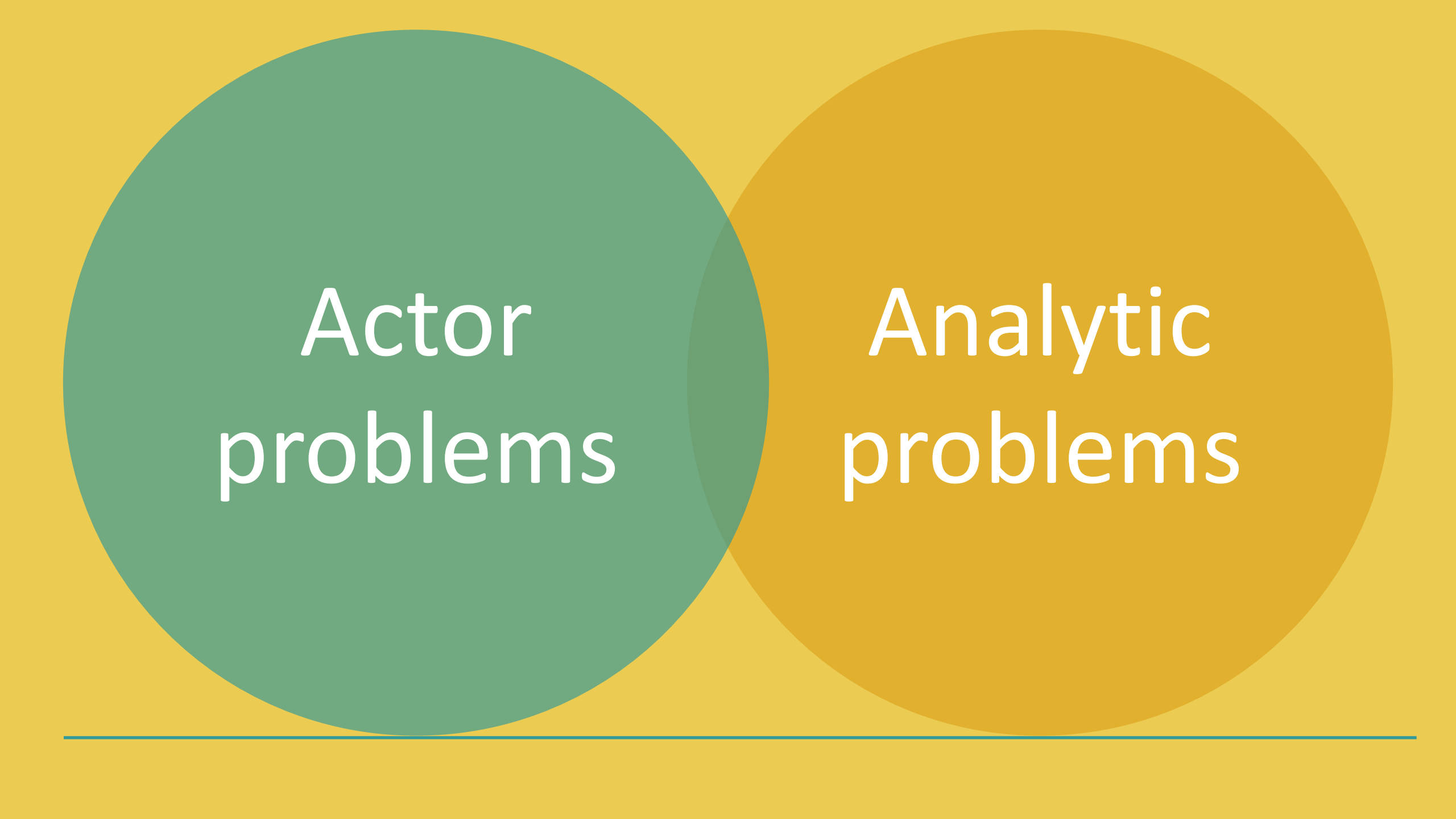
Stoke Gifford Retirement Village



# What's at stake

How are social orders reconfigured by emerging technologies?

What kind of rules are being created through standard making?  
...and what are the motivational and organizational features of this knowledge production?



Actor  
problems

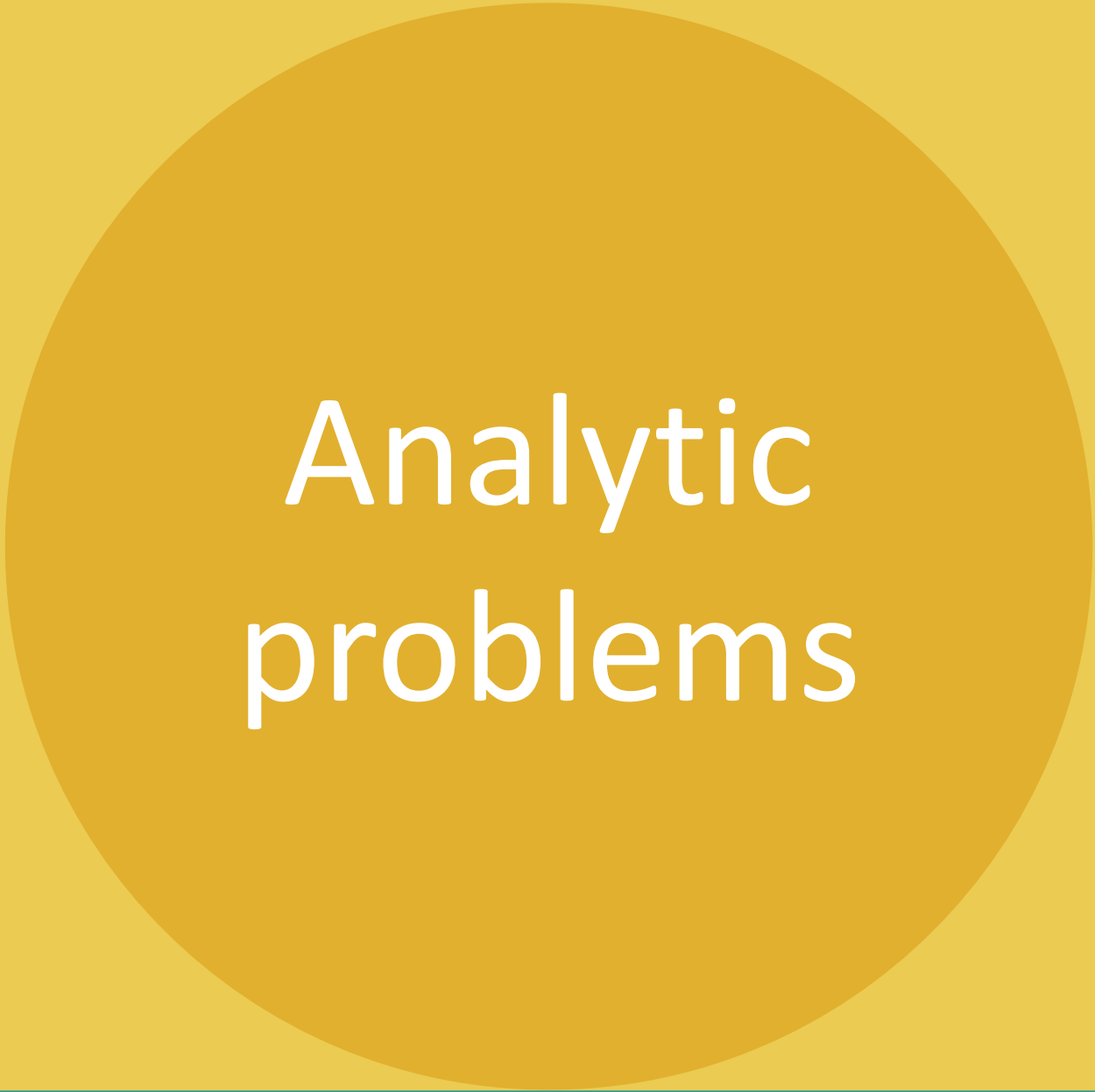
Analytic  
problems

---



## Problems of structure and agency

- Structural problems that local reconfigurations of agency / networks / st-systems can't mitigate.
- Collective action as a way forward?
  - Peer-production of standards
- Alternative way to structuring the world vs. regulation/legislation
  - Often co-terminus and self-referencing processes.



Analytic  
problems

# Actor problems

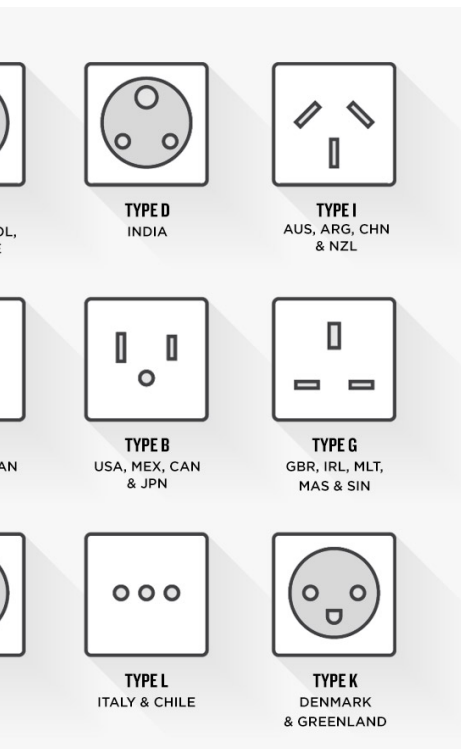
- Set of problems around AI
    - Rise of ethical frameworks
    - Issue: from principles to practice
  - Backgrounded
    - Issue of ethics washing
    - Whose expertise counts
-



# From principles to practice: standards



Standards are the rules, guidelines and procedures by which social and technological worlds are brought together to make things work



**YOUR COMPANY NAME & TEL NUMBER INSERTED HERE**

## SITE SAFETY

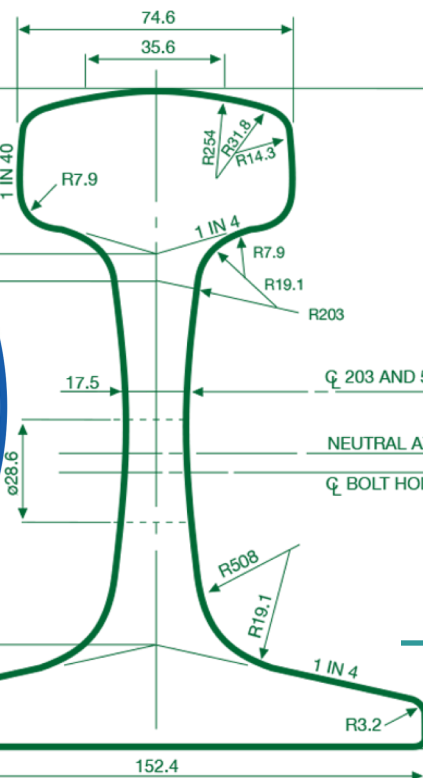
**Unauthorised entry to this site is strictly prohibited**

**Danger**  
Construction work in progress

- 1) All Visitors must report to the site office.
- 2) Permission Must be cleared before entering.
- 3) Safety signs and procedures must be observed.
- 4) Personal protective and safety equipment must be used at all times.
- 5) All persons entering this site must comply with the regulations under the Health and Safety at Work Act 1974.

**Safety helmets must be worn**

**Protective footwear must be worn**



Wi-Fi Standard	Frequency
802.11a (1999)	5 GHz
802.11b (1999)	2.4 GHz
802.11g (2003)	2.4 GHz
802.11n (2009)	2.4/5 GHz
802.11ac (Draft - 2012)	5 GHz

Table 1. IEEE P7000 Series standards in development

Standard	Description
7000	Engineering Methodologies for Ethical Life-Cycle Concerns Working Group
7001	Transparency of Autonomous Systems
7002	Personal Data Privacy Working Group
7003	Algorithmic Bias Working Group
7004	Standard for Child and Student Data Governance
7005	Employer Data Governance Working Group
7006	Personal Data AI Agent Working Group
7007	Ontological Standard for Ethically Driven Robotics and Automation Systems
7008	Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems Working Group
7009	Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems
7010	Well-being Metric for Autonomous and Intelligent Systems (A/IS)
7011	News Site Trustworthiness Working Group
7012	Machine Readable Privacy Terms Working Group
7013	Benchmarking Accuracy, Increasing Transparency, and Governing Use of Automated Facial Analysis Technology Working Group

ETHICAL  
ALIGNED  
First Edition

A Vision for Prioritizing  
with Autonomous ar

IEEE P7005™  
Standard on Employ  
Data Governance

IEEE P7008™  
Standard for Ethically D  
Nudging for Robotic, Inte  
and Autonomous Syste



# P7003 Algorithmic Bias: motivations

- “the IEEE P7000 series empowers innovation across borders and yields societal benefits”  
Havens and Hessami (2019, p. 70)
- "IEEE P7003 is aimed to be used by people/orgs who are developing and/or deploying automated decision (support) systems (which may or may not involve AI/machine learning) that are part of products/services that affect people. Typical examples include anything related to personalization or individual assessment, including systems that perform filtering functions by selecting to prioritize the ease with which people will find some items over others
- Policy drivers e.g. GDPR
- EPSRC Unbias project / ‘ref-able’ impact

## P7003

**Submitter Email:** [johnchavens@email.com](mailto:johnchavens@email.com)  
**Type of Project:** New IEEE Standard  
**PAR Request Date:** 19-Dec-2016  
**PAR Approval Date:** 17-Feb-2017  
**PAR Expiration Date:** 31-Dec-2021  
**Status:** R of a New IEEE Standard

**1.1 Project Number:** P7003  
**1.2 Type of Document:** Standard  
**1.3 Life Cycle:** Full Use

**2.1 Title:** Algorithmic Bias Considerations

**3.1 Working Group:** Algorithmic Bias Working Group  
**Contact Information for Working Group Chair**  
**Name:** Ansgar Koene  
**Email Address:** [ansgar.koene@nottingham.ac.uk](mailto:ansgar.koene@nottingham.ac.uk)  
**Phone:** +447972197518  
**Contact Information for Working Group Vice-Chair**  
None

**3.2 Sponsoring Society and Committee:** IEEE Computer Society/Software & Systems Engineering Standards Committee (C/S2ESC)  
**Contact Information for Sponsor Chair**  
**Name:** Paul Croll  
**Email Address:** [pcroll@computer.org](mailto:pcroll@computer.org)  
**Phone:** 540-644-6224

**Contact Information for Standards Representative**  
None

**4.1 Type of Ballot:** Individual

**4.2 Expected Date of submission of draft to the IEEE-SA for Initial Sponsor Ballot:** 06/2019

**4.3 Projected Completion Date for Submittal to RevCom**

**Note:** Usual minimum time between initial sponsor ballot and submission to Revcom is 6 months.: 10/2019

**5.1 Approximate number of people expected to be actively involved in the development of this project:** 25

**5.2 Scope:** This standard describes specific methodologies to help users **certify how they worked to address and eliminate issues of negative bias in the creation of their algorithms**, where "negative bias" infers the usage of overly subjective or unformed data sets or information known to be inconsistent with legislation concerning certain protected characteristics (such as race, gender, sexuality, etc); or with instances of bias against groups not necessarily protected explicitly by legislation, but otherwise diminishing stakeholder or user well being and for which there are good reasons to be considered inappropriate. Possible elements include (but are not limited to): **benchmarking procedures and criteria for the selection of validation data sets for bias quality control; guidelines on establishing and communicating the application boundaries for which the algorithm has been designed and validated to guard against unintended consequences arising from out-of-bound application of algorithms; suggestions for user expectation management to mitigate bias due to incorrect interpretation of systems outputs by users** (e.g. correlation vs. causation)

**5.3 Is the completion of this standard dependent upon the completion of another standard:** No

**5.4 Purpose:** This standard is designed to provide individuals or organizations creating algorithms, largely in regards to autonomous or intelligent systems, certification oriented methodologies to provide clearly articulated **accountability and clarity** around how algorithms are targeting, assessing and influencing the users and stakeholders of said algorithm. **Certification under this standard will allow algorithm creators to communicate to users, and regulatory authorities, that up-to-date best practices were used in the design, testing and evaluation of the algorithm to avoid unjustified differential impact on users.**

**5.5 Need for the Project:** The rapid growth of algorithm driven services has led to growing concerns among civil society, legislators, industry bodies and academics about potential unintended and undesirable biases within intelligent systems that are largely inscrutable 'black boxes' to the users.

## IEEE P7003™

IEEE Standards Project for [Algorithmic Bias Considerations](#) provides developers of algorithms for autonomous or intelligent systems with protocols to avoid negative bias in their code. Bias could include the use of subjective or incorrect interpretations of data like mistaking correlation with causation. The project offers specific steps to take for eliminating issues of negative bias in the creation of algorithms.



# Post- automation potentials

Robotics innovation pathways  
that are not automatically  
determined  
...arrived at by means of broad  
participation in governance  
decisions and innovation  
processes

## Post-automation potentials

### Reappraising human agency

Standards offer a location at which to re-orientate human agency.

### New appropriation locations

One of the drivers of these standards is the role that A/IS are now playing in locations beyond industrial setting. Actors include children, elderly and ill people.

BUT unclear how these standards might enhance agency for people to deal with A/IS on their own terms.

### New forms of sociotechnical relations

Not clear in what sense groups pursuing creative livelihoods and sustainability with these technologies are impacted.

Possible site of peer-produced governance outside the control of state or firm

### Subverting industrial technology

Explicitly ethical standards offer another site, in addition to strikes and other direct action where tech-workers can reclaim agency in pursuit of steering innovation

### Beyond current social theory

New forms of collective intelligence

### Broadening out and opening up ethical inputs and opportunities for post-automation

Yes, some broadening out of inputs to ethical framework writing.



# Reflections

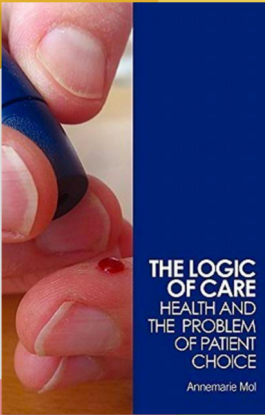
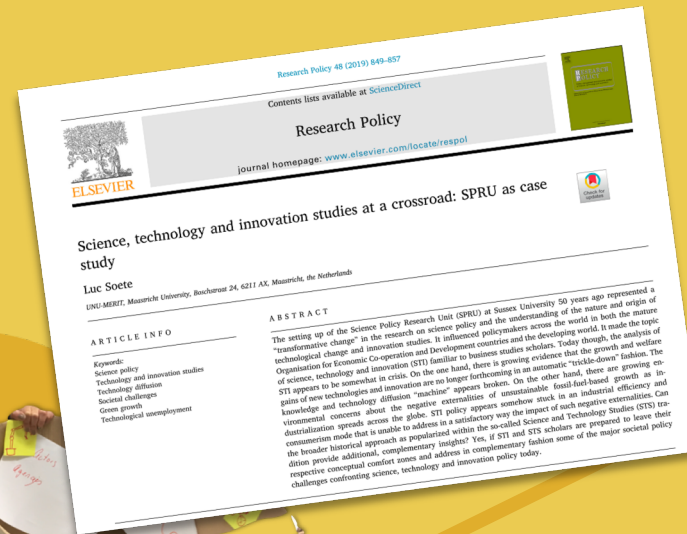
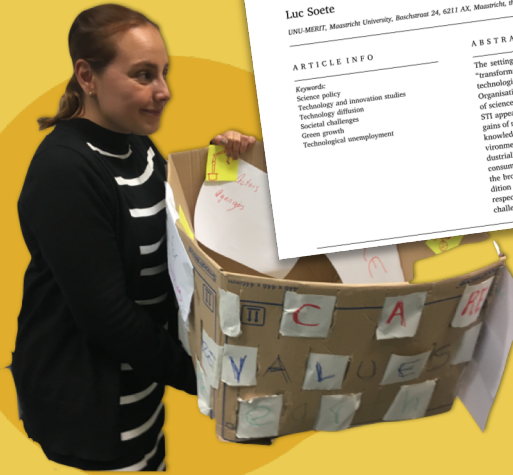
- Does post-automation usefully frame these problems?
  - Is deliberation, diversity and democracy useful here?
  - Does post-automation privilege ...
    - *liberation within* (e.g. global/local platforms)?
    - *liberation from* (alternative social orders)
-



# Speculative reflections

- Augmenting and acknowledging logics of control / logics of care
- Making choice visible

# Care



# Control

DIGITAL INFRASTRUCTURES: ESSAY

---

**Test-Bed Urbanism**

*Orit Halpern, Jesse LeCavalier, Nerea Calvillo, and Wolfgang Pietsch*



**Cian O'Donovan**

c.o'donovan@ucl.ac.uk. | @cian

[www.scalings.eu](http://www.scalings.eu)

[info@scalings.eu](mailto:info@scalings.eu)



**SCALINGS**

SCALING UP  
CO-CREATION



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 788359