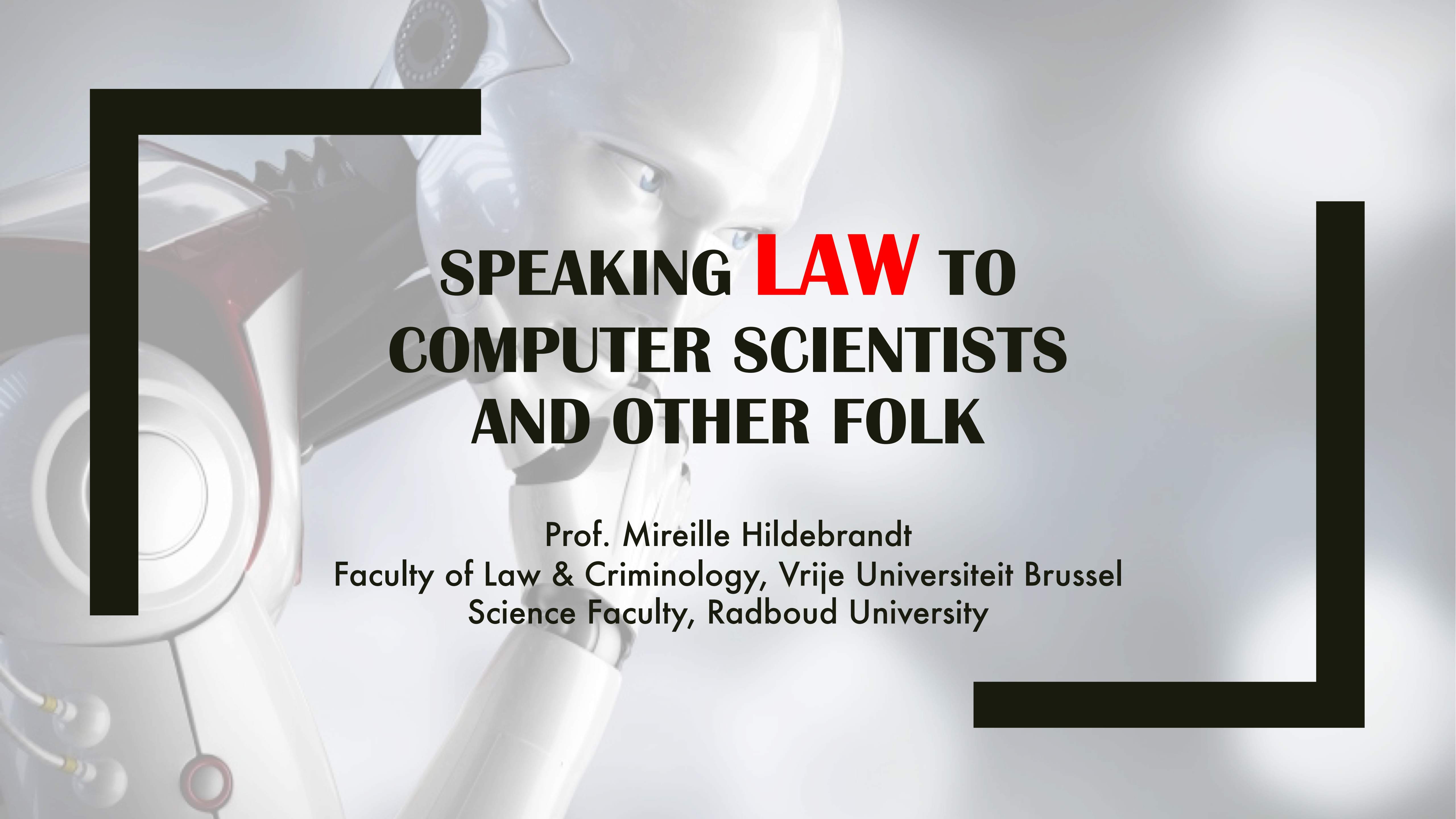
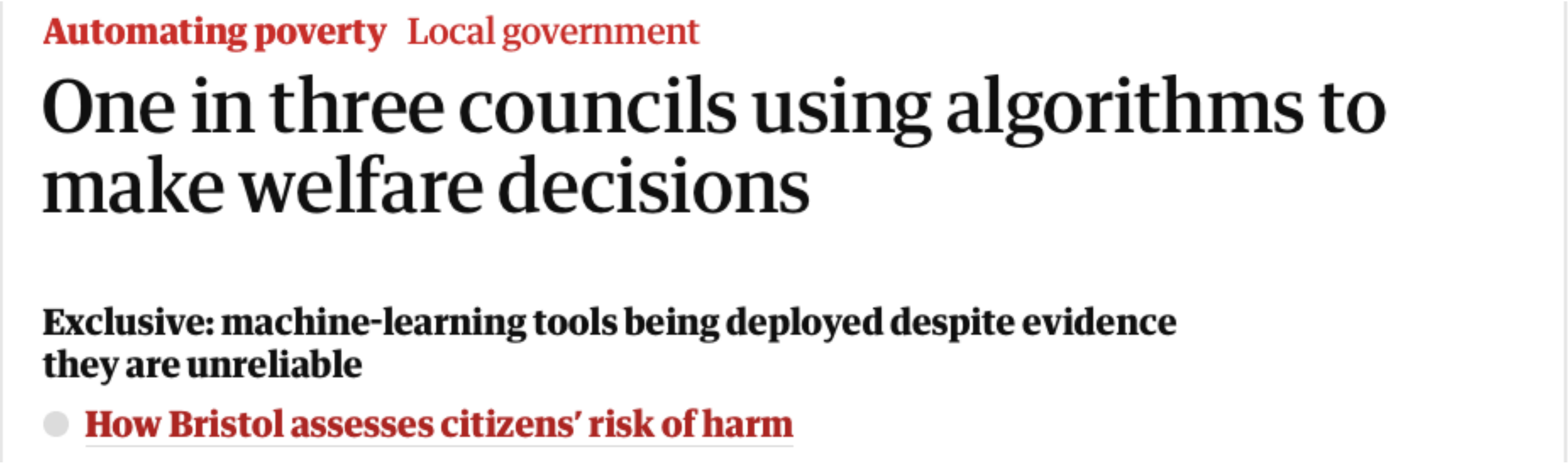


SPEAKING LAW TO COMPUTER SCIENTISTS AND OTHER FOLK

Prof. Mireille Hildebrandt
Faculty of Law & Criminology, Vrije Universiteit Brussel
Science Faculty, Radboud University



Automating poverty Local government

One in three councils using algorithms to make welfare decisions

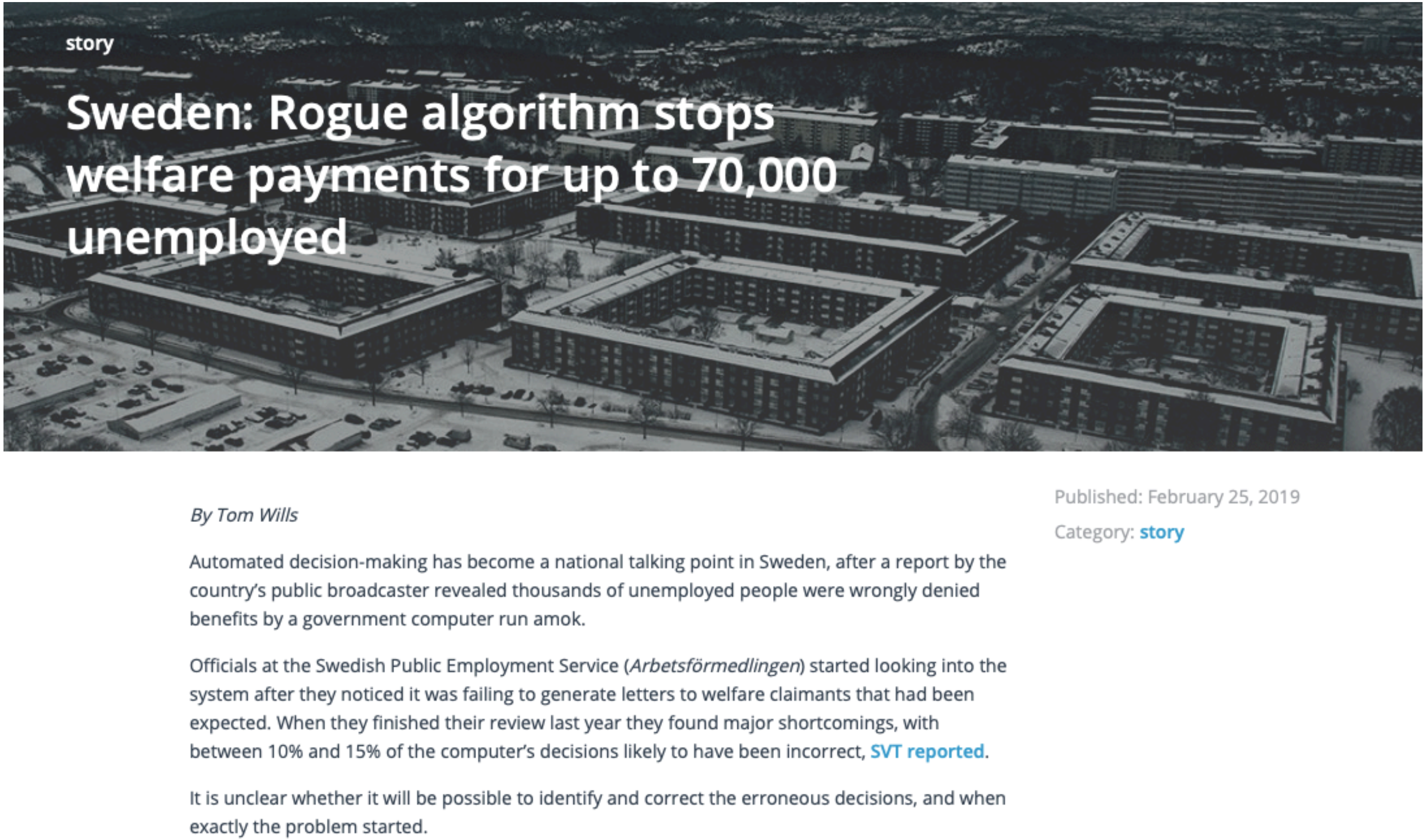
Exclusive: machine-learning tools being deployed despite evidence they are unreliable

- How Bristol assesses citizens' risk of harm

David Spiegelhalter, a former **president of the Royal Statistical Society**, said:

“There is too much hype and mystery surrounding machine learning and algorithms.

- ➢ I feel that councils should demand trustworthy and transparent explanations of
  - **how any system works,**
  - **why it comes to specific conclusions about individuals,**
  - **whether it is fair, and**
  - **whether it will actually help in practice.”**



story

Sweden: Rogue algorithm stops welfare payments for up to 70,000 unemployed

By Tom Wills

Published: February 25, 2019

Category: story

Automated decision-making has become a national talking point in Sweden, after a report by the country's public broadcaster revealed thousands of unemployed people were wrongly denied benefits by a government computer run amok.

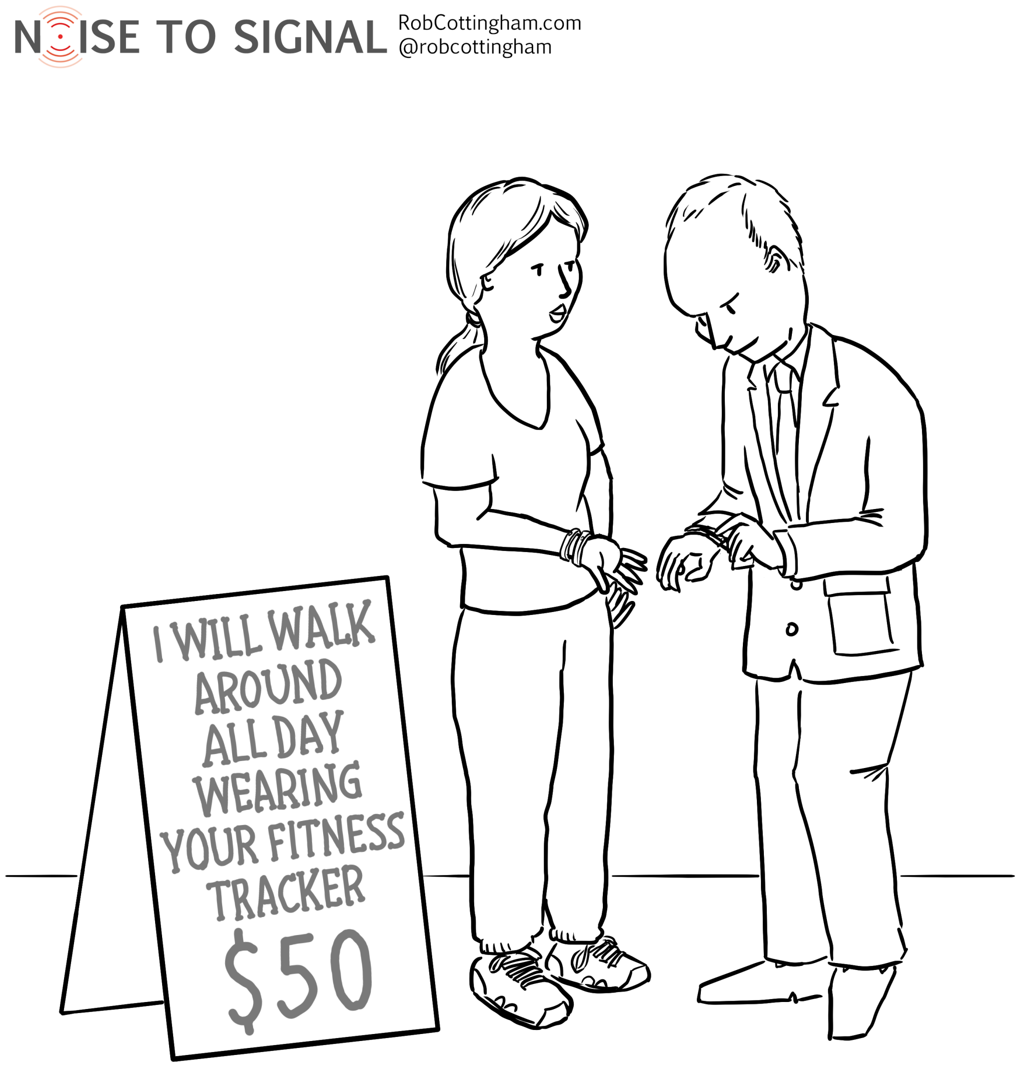
Officials at the Swedish Public Employment Service (*Arbetsförmedlingen*) started looking into the system after they noticed it was failing to generate letters to welfare claimants that had been expected. When they finished their review last year they found major shortcomings, with between 10% and 15% of the computer's decisions likely to have been incorrect, SVT reported.

It is unclear whether it will be possible to identify and correct the erroneous decisions, and when exactly the problem started.

Methodological integrity crises in social and natural sciences

Now emerging in ML research design, because:

- Training data are ‘low hanging fruit’
  - Irrelevant, incomplete, inaccurate
- Test data are gamed
  - When a measure becomes the target, it is no longer a good measure (Goodhart effect)
- P-hacking, data dredging
  - Wrong conclusions drawn with regard to null hypothesis
- Feature space underdeveloped (blind to missing relevant variables)
- Hypothesis space by definition limited (Wolpert NFL)
- Performance metrics chosen that result in high accuracy
- No out of sample testing (only validation on historical data)


NOISE TO SIGNAL
RobCottingham.com
@robcottingham
I WILL WALK AROUND ALL DAY WEARING YOUR FITNESS TRACKER
$50

What’s next?

- Law as architecture
- The choice architecture of the Rule of law
- The GDPR and the Charter of Fundamental Rights
- The methodological integrity of computer science and the GDPR
- Legal protection by design

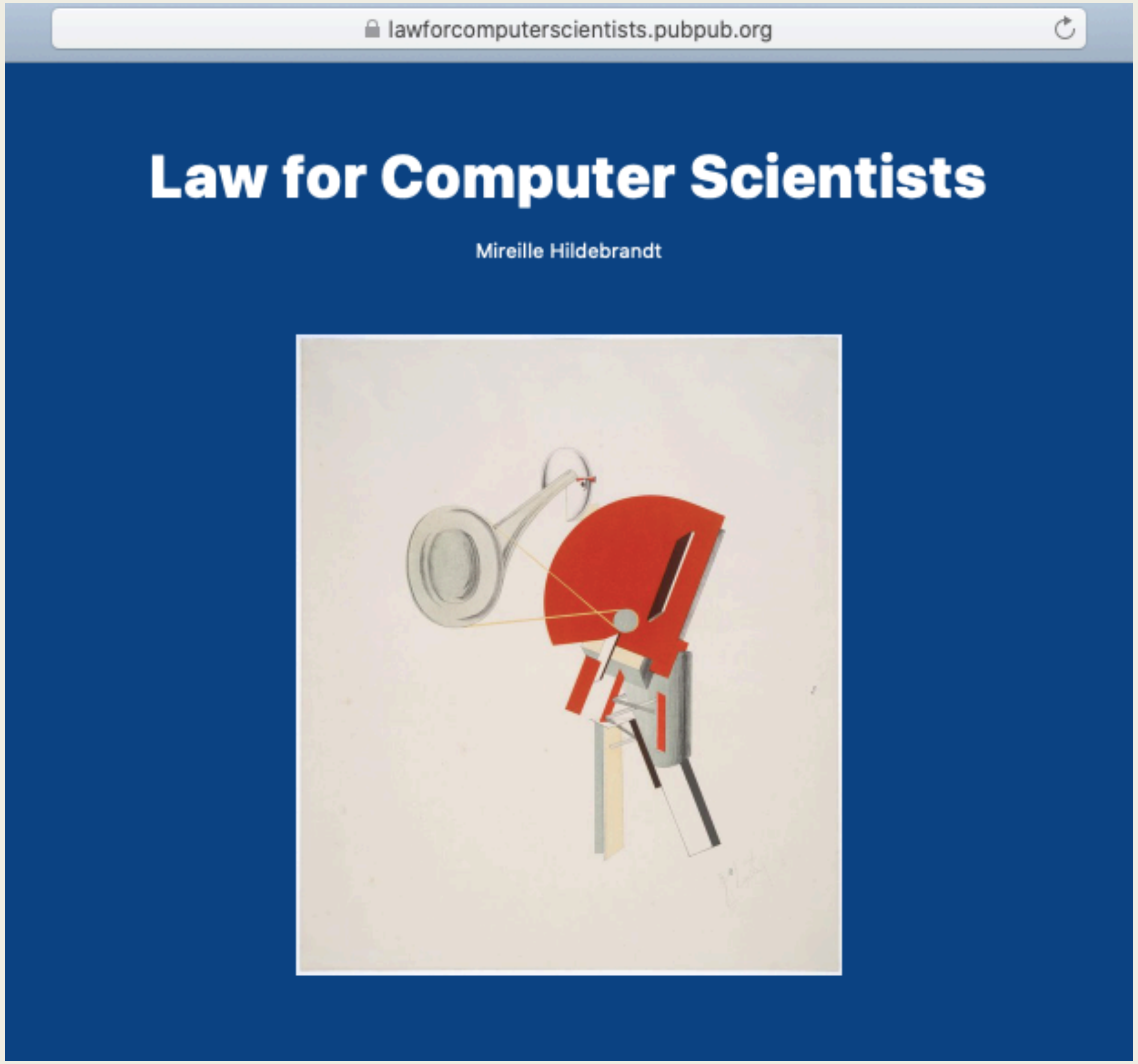
Available at MIT's pubpub:

- https://lawforcomputerscientists.pubpub.org

In print March 2020

Oxford University Press

- Hardcopy
- Ebook in open access


lawforcomputerscientists.pubpub.org

Law for Computer Scientists

Mireille Hildebrandt

Law as architecture

- ‘Positive law’ is a human construction
- Law is multidimensional: legislation, public administration, case law
- **Law is a system defined by and defining human interaction**
  - a system of legal norms (rules, principles)
    - that attribute legal effect
    - that define what counts as a legally relevant action
  - a system of legal relationships (e.g. in contract or property)
    - between legal subjects (natural persons, legal persons)
    - with regard to legal objects (relative and absolute rights)

Law as architecture

- **Legal norms and legal relationships are mutually constitutive:**
  - Law as a system of legal norms (e.g. contracts)
    - that define legal relationships (between the parties of the contract)
  - Law as a system of legal relationships (e.g. the owner of a house and all others)
    - that define legal norms (right to dispose, right to non-interference)

Law as architecture

- Legal norms define:
  - what legal conditions
  - result in what legal effect

- Legal effect is NOT caused but *attributed* by law

Law as architecture

Article 5 GDPR: Principles relating to processing of personal data

1. **Personal data shall be:**

c) adequate, relevant and limited to what is necessary

- in relation to the purposes for which they are processed ('data minimisation');

- If not necessary (in relation to purpose)
- Legal effect is that the processing is unlawful

Law as architecture

Article 5 GDPR: Principles relating to processing of personal data

2. The controller shall be responsible for, and be able to demonstrate compliance with, paragraph 1 ('accountability').

- If not necessary (in relation to purpose)
- Legal effect is that controller responsible

Law as architecture

- **Article 82 GDPR: Right to compensation and liability**

1. Any person who has suffered material or non-material damage
   - as a result of an infringement of this Regulation
   - shall have the right to receive compensation
   - from the controller or processor
   - for the damage suffered.

- If not necessary (in relation to purpose)
- Legal effect: controller is liable

Law as architecture

- So what is ‘necessary’?
  - Whatever is not effective cannot be necessary
  - Necessary in relation to an explicit, legitimate, specified purpose
    - If other means are available the processing is not necessary (subsidiarity)
  - Proportionality test:
    - If processing infringes fundamental rights or freedoms
    - The more serious the infringement, the higher the threshold for ‘necessity’

Law as architecture

- Legal effect is what speech act theory calls a *performative effect*:
  - Not part of propositional or deontological logic
  - Not a matter of causation, but of **meaning**
  - *A speech act* 'does' what it 'says'
    - 'I pronounce you husband and wife' is NOT a description
    - The conditions for transfer of ownership are not a matter of moral choice
    - Liability of the data controller is not caused but *attributed*

Law as architecture

- **Legal norms differ from computer code**
  - Not based on logic (though logic is involved)
  - Not based on causality (though causality is involved)
  - Not based on computation (though complex decision trees may apply)
  - Not based on probability (though no 100% certainty)

Legal certainty: the hallmark of positive law

- Law thrives on a specific type of uncertainty, that is contingent upon:
  - ambiguity of natural language
  - potential enforcement
- Legal certainty depends on:
  - Adaptive nature of norms articulated in human language
  - Potential enforcement depending on the meaning of the norm
- Legal certainty thus:
  - Implies the uncertainty it sustains and resolves: multi-interpretatibility
  - Affords both argumentation and contestation
- This is a feature not a bug, and grounds the Rule of Law

Choice architecture

Law determines the types of choices of those subject to its jurisdiction, e.g.

- Private law ‘makes’ economic markets, e.g.:
  - Freedom to contract & freedom from undue influence
  - Freedom to dispose of one’s property & freedom from interference

  This creates the choice architecture for consumers, businesses etc.

Choice architecture

The Rule of Law determines that those who enact, apply and interpret the law are also subject to the law

- Legality principle: government are not free to act whichever way they want, they must act within their legal competences/legal powers
- This involves a smart system of checks and balances:
  - Those who enact the law are not the same as those who decide on the meaning of the law
  - Cp. sharing a cake: who gets to cut, who gets to choose first

Choice architecture

- Before the Rule of Law was established we had enlightened despots
  - They had the power to decide, without oversight
  - They had good intentions regarding their subjects

- Establishing the Rule of Law meant:
  - Those in charge are subject to the law
  - We do not want to depend on the ethical inclinations of those in charge
  - To reign in the power of our rulers we have countervailing powers

Choice architecture

Computer science applications increasingly determine the types of choices their users have

- Code developers and data-driven platforms behave as enlightened despots
  - They have the power to decide, without oversight
  - They may have good intentions regarding their users

- Establishing the Rule of Law means:
  - Developers and Big Tech under the Rule of Law
  - We (users) do not want to depend on their ethical inclinations
  - We (all of us) need a system of countervailing powers

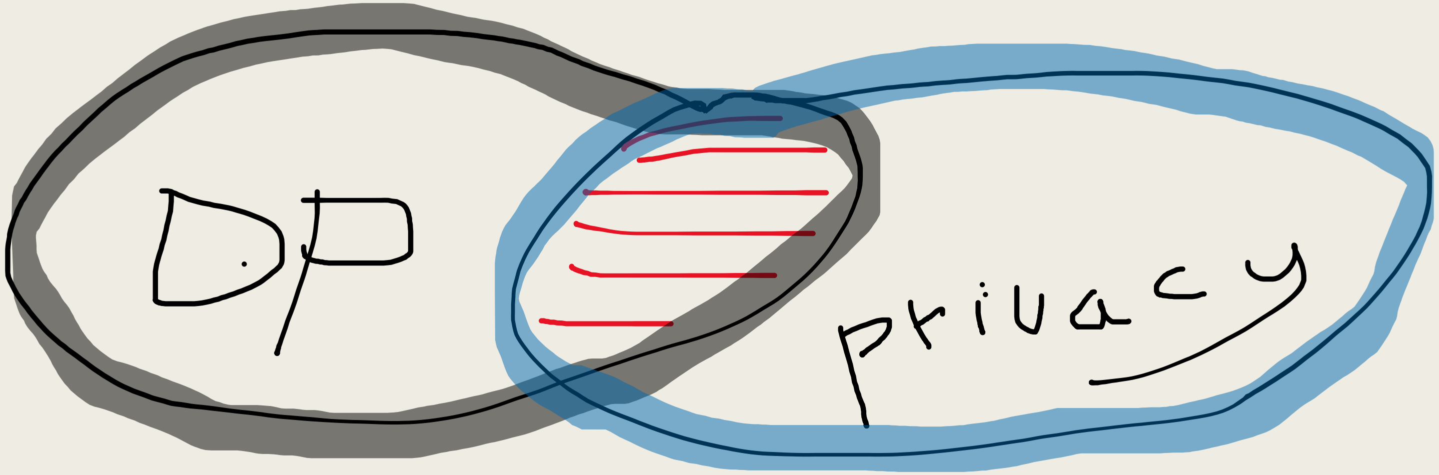
GDPR and the Charter

Data protection law is not equivalent with privacy law

- In Europe (EU) we have two fundamental rights:
  - Art. 7 Charter: right to privacy
  - Art. 8 Charter: right to data protection

GDPR and the Charter

Data protection law is NOT equivalent with privacy law



GDPR and the Charter

Article 7 Respect for private and family life

Everyone has the right to respect for his or her private and family life, home and communications.

GDPR and the Charter

Article 8 Protection of personal data

1. Everyone has the right to the protection of personal data concerning him or her.
2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
3. Compliance with these rules shall be subject to control by an independent authority.

GDPR and the Charter

Article 8 Protection of personal data

2. Such data must be processed
   - fairly
   - for specified purposes and
   - on the basis of the consent of the person concerned
   - or some other legitimate basis laid down by law.

   Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.

GDPR and the Charter

GDPR relevance for CS:

- Personal data are defined as data relating to an identifiable natural person
  - Assume that most sensor, behavioural and textual data are personal data
  - E.g. a dynamic IP address may be personal data

- Controller (liable) is whoever de facto decides the purpose (and the means)
  - Which much be communicated to the data subjects
  - Processor processes on behalf of controller (e.g. cloud provider)

GDPR and the Charter

GDPR relevance for CS:

- To process personal data you always need:
  1. A legal basis (consent, contract, legal obligation, vital interest, public task, legitimate interest of the controller)
  2. Compliance with principles (lawfulness, fairness, transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, accountability)

GDPR and the Charter

GDPR relevance for CS:

- **Content is largely the same as previous data protection directive, but:**
  1. Smart enforcement chapter
  2. New types of legal obligations that ‘speak to’ computational architecture

GDPR and the Charter

GDPR relevance for CS:

- New types of legal obligations that ‘speak to’ computational architecture
  - Data protection by design & default
  - Data protection impact assessment

GDPR and the Charter

GDPR relevance for CS:

- GDPR takes a risk approach:
  1. Controller must assess risk to rights and freedoms natural persons
  2. Such risks must be mitigated:
     - By choosing another way to achieve the purpose
     - By implementing security by design, data protection by default & design

GDPR and the Charter

GDPR relevance for CS:

- GDPR requires proportionality test:
  1. If legitimate interest is used as legal basis
  2. If processing infringes fundamental rights and freedoms

GDPR and the Charter

GDPR relevance for CS:

- GDPR offers a broad exception for scientific research (in the public interest)
  - Secondary purpose is assumed to be compatible
  - But still need a valid legal ground: consent, legitimate interest
  - If sensitive data (e.g. health), pay attention: a more strict regime (though still broad exceptions for e.g. medical research)

GDPR and the Charter

GDPR relevance for CS:

- **Prohibition of automated decisions ex art. 22 GDPR**
  - Both deterministic ('persistent script') and profiling (machine learning)
  - Solely automated (window dressing does not count)
  - Significant effect on data subject

GDPR and the Charter

GDPR relevance for CS:

- **Exceptions prohibition of automated decisions ex art. 22 GDPR**
  - Consent
  - Contract
  - Legal obligation

GDPR and the Charter

GDPR relevance for CS:

- In case of exception:
  - Safeguards, notably right to human intervention
  - Transparency:
    1. Information that decision was automated
    2. Meaningful information on logic of processing
    3. And envisaged consequences
- Goal: to ensure contestability (Rule of Law requirement)

Methodological Integrity ML (applications)

- Both the natural and the social sciences confront a methodological crisis:
  - Reproducibility, replicability
  - E.g. statistical delusions that suggest no further testing is needed
    - If hypothesis A is true then this data is not probable
    - If this data is not probable then hypothesis A is not true

Methodological Integrity ML (applications)

- Both the natural and the social sciences confront a methodological crisis:
  - Reproducibility, replicability
  - E.g. statistical delusions that suggest no further testing is needed
    - If hypothesis A is true then this data is not probable
    - ~~If this data is not probably then hypothesis A is not true~~
  - We actually want to know:
    - Given this data, what is the probability that A is true

Methodological Integrity ML (applications)

Crisis emerging in ML research design, because:

- Training data are ‘low hanging fruit’
  - Irrelevant, incomplete, inaccurate
- Test data are gamed
  - When a measure becomes the target, it is no longer a good measure (Goodhart effect)
- P-hacking, data dredging
  - Wrong conclusions drawn with regard to null hypothesis
- Feature space underdeveloped (blind to missing relevant variables)
- Hypothesis space by definition limited (Wolpert NFL)
- Performance metrics chosen that result in high accuracy
- No out of sample testing (only validation on historical data)

Methodological Integrity ML (applications)

- Maybe this is not a problem for exploratory research?
  i.e. when generating hypotheses?

  [unless this is sold as confirmed claims]

Methodological Integrity ML (applications)

- **Hofman, Sharma, Watts on exploratory and confirmatory research design:**

Exploratory ML researchers are **free** to

- study different tasks,
- fit multiple models,
- try various exclusion rules, and
- test on multiple performance metrics.

When reporting their findings, however, they should:

- transparently declare their full sequence of design choices
  - to avoid creating a false impression of having confirmed a hypothesis rather than simply having generated one,
- report performance in terms of multiple metrics
  - to avoid creating a false appearance of accuracy.

Methodological Integrity ML (applications)

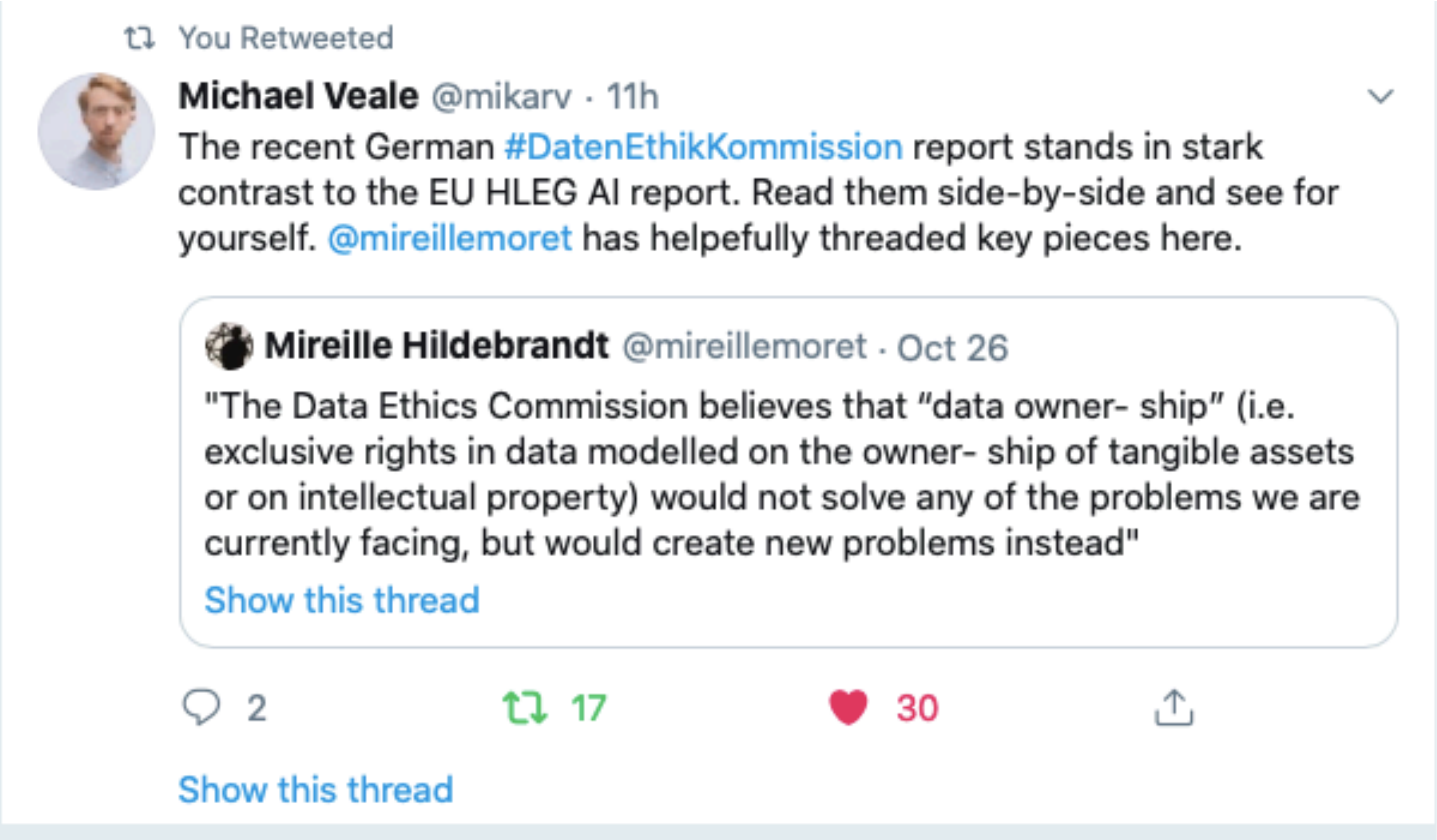
- **Hofman, Sharma, Watts on exploratory and confirmatory research design:**

Confirmatory ML: researchers should be

- **required to preregister their research designs,**
- including data preprocessing choices,
- model specifications,
- evaluation metrics,
- and out-of-sample predictions,
- **in a public forum such as the Open Science Framework (https://osf.io).**

Legal Protection by Design

- This is NOT about ‘values’ by design (depending on unenforceable ethics)
- This is NOT about techno-regulation (brute forcing compliance)



You Retweeted

Michael Veale @mikarv · 11h

The recent German #DatenEthikKommission report stands in stark contrast to the EU HLEG AI report. Read them side-by-side and see for yourself. @mireillemoret has helpefully threaded key pieces here.

Mireille Hildebrandt @mireillemoret · Oct 26

"The Data Ethics Commission believes that "data owner- ship" (i.e. exclusive rights in data modelled on the owner- ship of tangible assets or on intellectual property) would not solve any of the problems we are currently facing, but would create new problems instead"

Show this thread

2 17 30

Show this thread

Legal Protection by Design

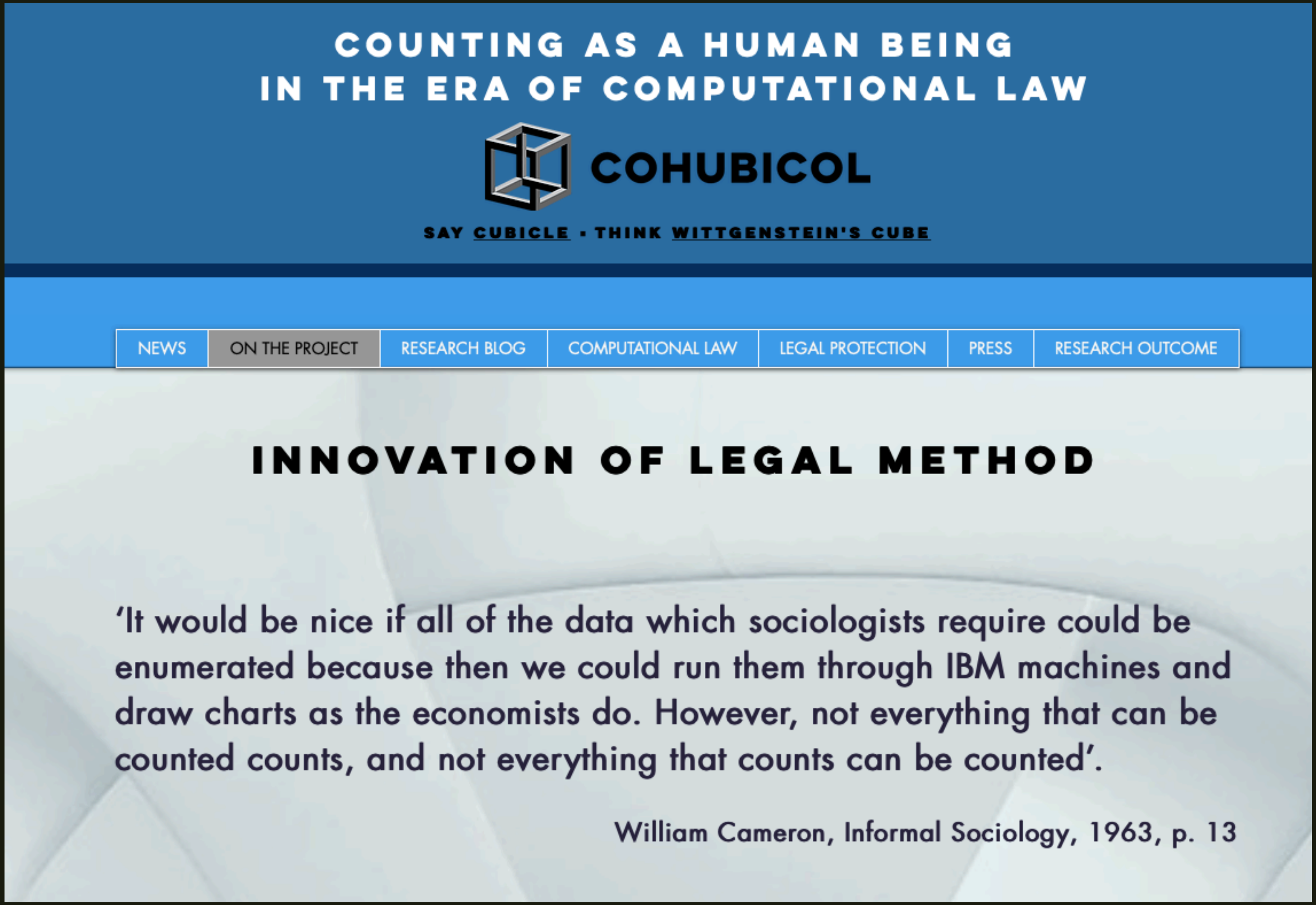
- This is about legal protection by design:
  - Instigated by democratic legislator
  - Resistable in real life and contestable in a court of law

- This is about legal protection by design:
  - Not meant to enable public administration by design
  - Not meant to fuse nudge theory with machine learning to manipulate us

Legal Protection by Design

- Build data minimisation into the computational architecture
- Assess and debate fairness of ML research design, make it contestable
- Develop standards for transparency (preregistration, explanation, contestability)
- Start with determining the purpose, consider its relation to the ML task
- Assess the accuracy of the data, do not confuse it with accuracy of inferences
- Storage limitation will prevent repeating or reinforcing the past
- Integrity and confidentiality will make sure the system respects *who matter*
- Effective accountability (i.e. liability) is crucial for any effective protection

ERC ADVANCED GRANT



COUNTING AS A HUMAN BEING
IN THE ERA OF COMPUTATIONAL LAW

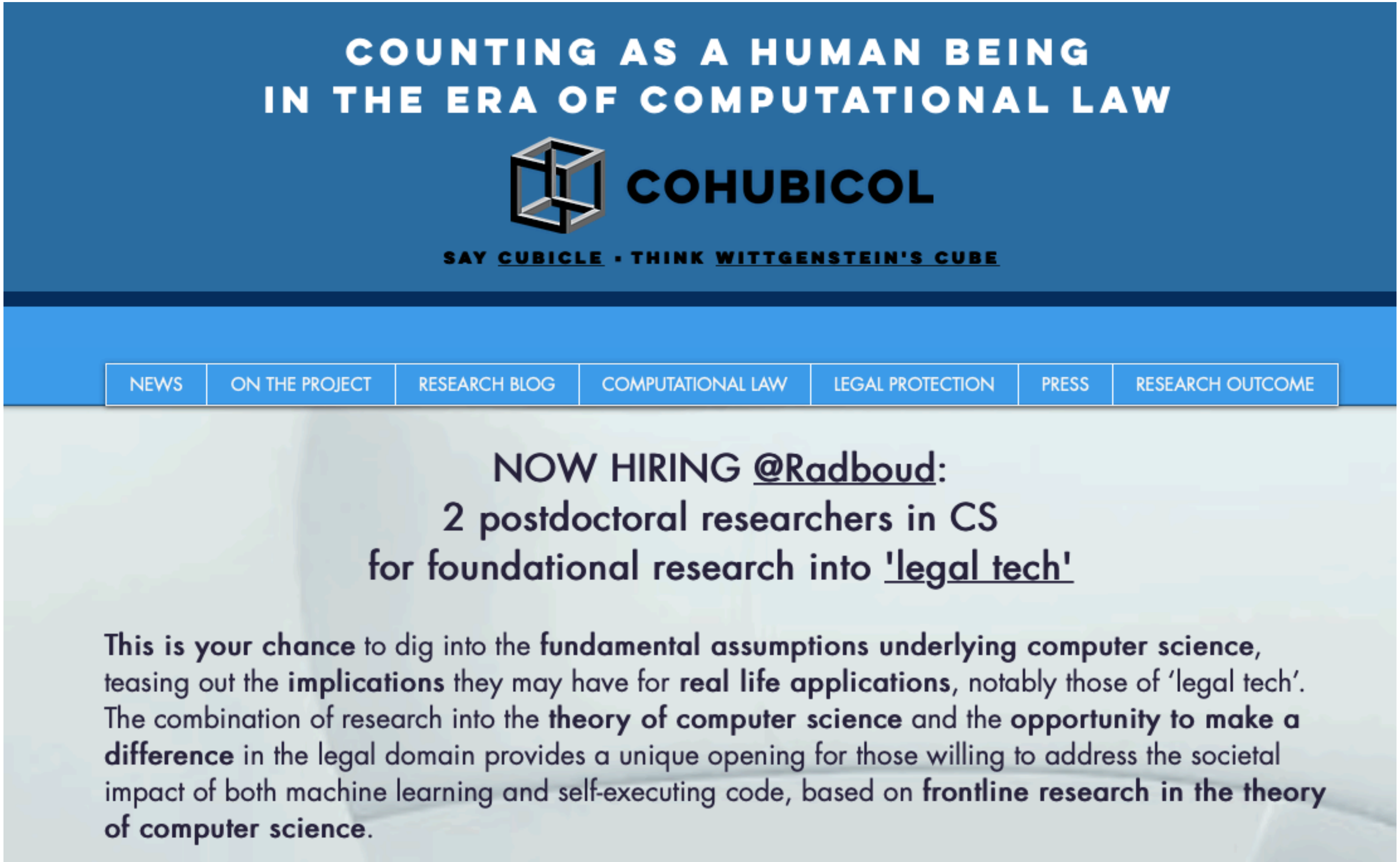
COHUBICOL

SAY CUBICLE • THINK WITTGENSTEIN'S CUBE

INNOVATION OF LEGAL METHOD

‘It would be nice if all of the data which sociologists require could be enumerated because then we could run them through IBM machines and draw charts as the economists do. However, not everything that can be counted counts, and not everything that counts can be counted’.

William Cameron, Informal Sociology, 1963, p. 13



COUNTING AS A HUMAN BEING IN THE ERA OF COMPUTATIONAL LAW

COHUBICOL

SAY CUBICLE • THINK WITTGENSTEIN'S CUBE

NOW HIRING @Radboud:
2 postdoctoral researchers in CS
for foundational research into 'legal tech'

This is your chance to dig into the **fundamental assumptions underlying computer science**, teasing out the **implications** they may have for **real life applications**, notably those of ‘legal tech’. The combination of research into the **theory of computer science** and the **opportunity to make a difference** in the legal domain provides a unique opening for those willing to address the societal impact of both machine learning and self-executing code, based on **frontline research in the theory of computer science**.