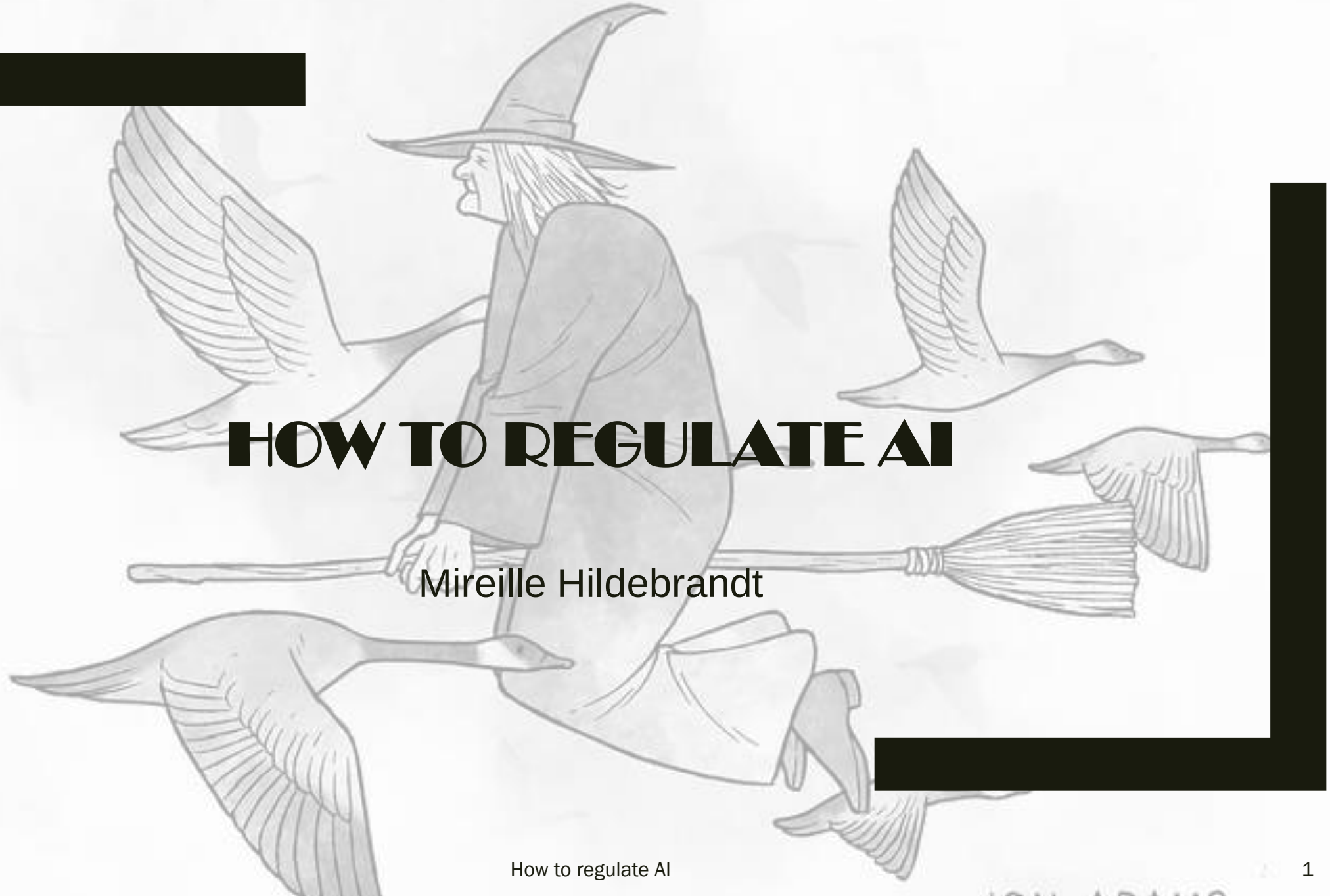


HOW TO REGULATE AI

Mireille Hildebrandt

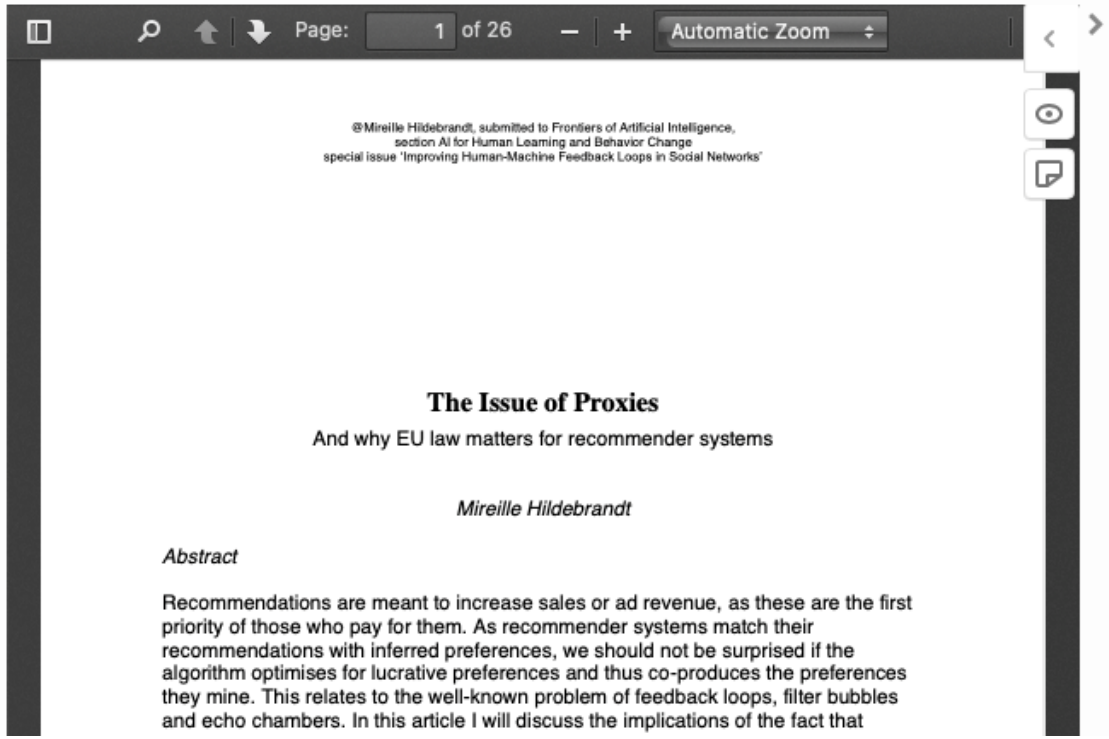
My 3 AI cards on the table

1. Things that matter are not computable
2. They can nevertheless be made computable
3. They can be computed in different ways and the difference matters

SocArXiv Papers | Submit a Paper | Search | Donate | Sign Up | Sign In

The Issue of Proxies, And why EU law matters for recommender systems

AUTHORS
Mireille Hildebrandt



Page: 1 of 26 Automatic Zoom

© Mireille Hildebrandt, submitted to Frontiers of Artificial Intelligence,
section AI for Human Learning and Behavior Change
special issue 'Improving Human-Machine Feedback Loops in Social Networks'

The Issue of Proxies

And why EU law matters for recommender systems

Mireille Hildebrandt

Abstract

Recommendations are meant to increase sales or ad revenue, as these are the first priority of those who pay for them. As recommender systems match their recommendations with inferred preferences, we should not be surprised if the algorithm optimises for lucrative preferences and thus co-produces the preferences they mine. This relates to the well-known problem of feedback loops, filter bubbles and echo chambers. In this article I will discuss the implications of the fact that

Download paper

Downloads: 233

plaudit Be the first to endorse this work

Abstract

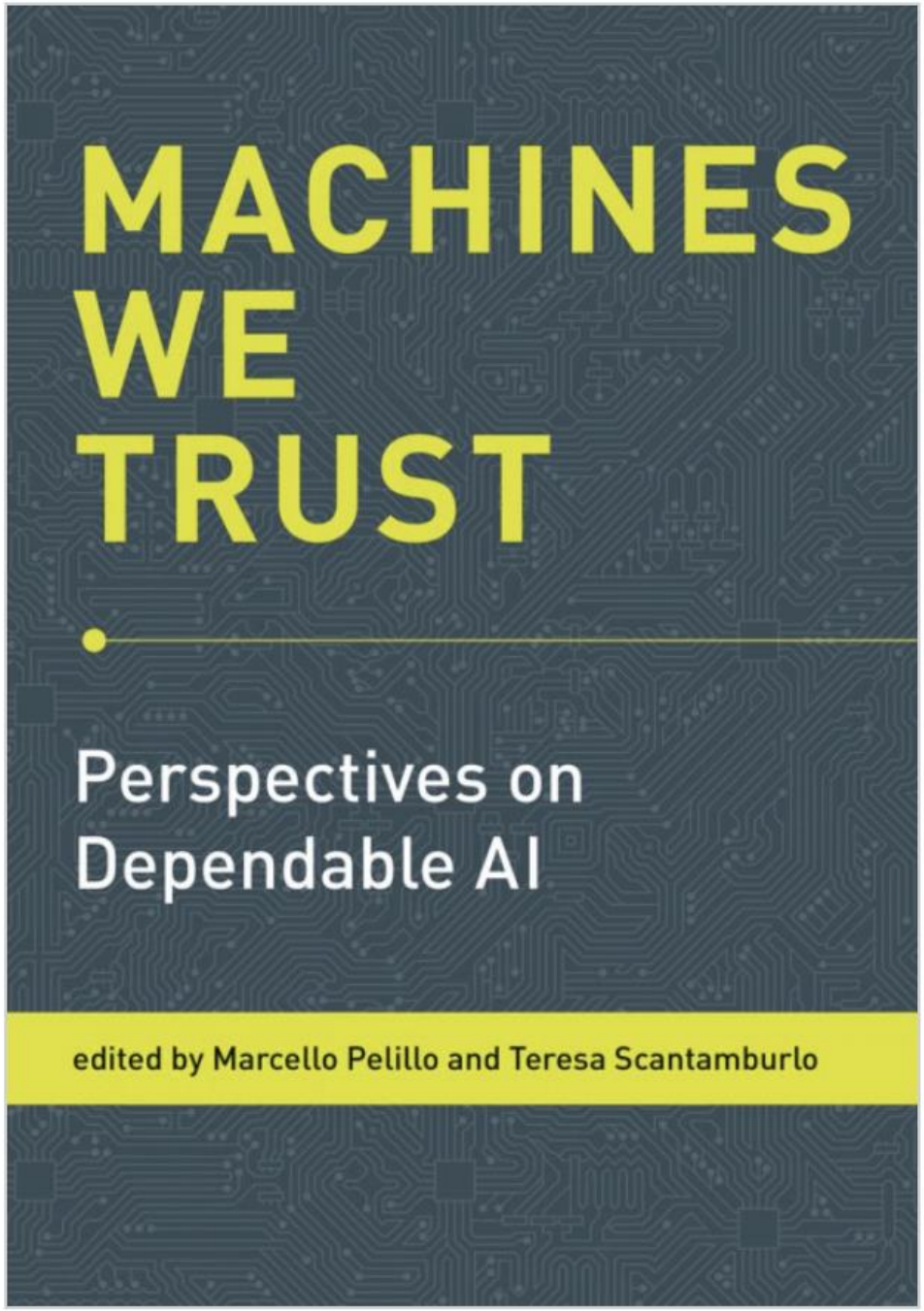
Recommendations are meant to increase sales or ad revenue, as these are the first priority of those who pay for them. As recommender systems match their recommendations with inferred preferences, we should not be surprised if the algorithm optimises for lucrative preferences and thus co-produces the preferences they mine. This relates to the ...

See more

Paper DOI

The issue of proxies

- Both behaviourism and ML depend often depend on:
  - *Inversion of the relationship between a concept/practice/institutional fact*
  - *And the proxy for that concept/practice/institutional fact*



4 The Issue of Bias: The Framing Powers of Machine Learning

Mireille Hildebrandt

4.1 Productive Bias, Wrongful Bias, and Unlawful Bias

In this chapter I will discuss three types of bias and their interrelationship. The first concerns the bias that is inherent in machine learning. This type of inductive bias is inevitable and, though neither good nor bad in itself, is never neutral in real world settings. The second concerns the bias that is problematic from an ethical perspective because it (re)configures the distribution of goods, services, risks, and opportunities or even access to information in ways that are morally problematic. This may regard categorical exclusion of people or the softer tyranny of nudging people into a certain direction based on traits or behaviors. Let's note that these traits or behaviors may be observed (by sensor technologies or online tracking systems) or inferred (by way of machine learning). Bias in observation affects the training data, and bias in inferences affects the throughput of the system; both impact the output. The third type of bias concerns unlawful bias, that is, the targeting of people based on prohibited grounds. This may be a subset of ethical bias, but sometimes bias that is not ethically problematic may nevertheless be unlawful because,[1] for instance, discrimination on the basis of gender may be prohibited categorically, even if some would argue that there is no ethical implication (e.g., charging men a higher car insurance premium because they are found to be more risk prone than women is not necessarily an ethical problem).

The issue of bias

- Wolpert's NFL theorem
- Gadamer's 'prejudice'
- Popper's theory-laden perception
- Hume's scepticism

- Bias is inherent in living agents:
- They need to detect the difference that makes a difference, but …

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

The rules we live by law and the rule of law

1. Modern positive law is text-driven, it in contingent upon a specific ICT infrastructure

- Text (printing press) has certain affordances:
  - *Distantiation over time and space*
    - Allowing to ‘rule’ larger polity beyond face-to-face
    - Allowing the enactment of rules that survive their author
  - *Distantiation between author and text and author and meaning*
    - Kantorowicz: The King’s Two Bodies (Le Roi est Mort, Vive le Roi)
    - Interpretation becomes the hallmark of the law

The rules we live by law and the rule of law

- Modern positive law is text-driven, it in contingent upon a specific ICT infrastructure
- Text (printing press) has certain affordances:
  - *The need for interpretation has two other affordances:*
    - The need for the stabilization of meaning (legal certainty)
    - An inherent instability of meaning that defines law's contestability
  - *Legal norms are externalised and become the focal point of*
    - Coordination, contestation and enforcement

The rules we live by law and the rule of law

- Modern positive law is text-driven, it in contingent upon a specific ICT infrastructure
- Text (printing press) has certain affordances:
  - *The relative autonomy of the law:*
    - The legislator can articulate binding norms
    - But cannot control its interpretation
    - Inviting the system of checks and balances that is key to the rule of law
    - Nonet & Selznick's 'historic bargain' between the political and the legal
    - The lawyers buffer between ruler and ruled

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

The new rules we live by
the marriage of nudge theory and ML

- The internet is driven by code and data, it has other affordances cp to text:
  - *Dis-intermediation of institutions, re-intermediation by code*
  - *ML is data-driven, but obviously also code-driven*
  - *Much ADM is code-driven though based on output of ML*
  - *Both ADM and ML imply formalisation and disambiguation*
  - *Both are based on working with proxies and short cuts*

The new rules we live by
the marriage of nudge theory and ML

- The internet is driven by code and data, it has other affordances cp to text:
  - *The assumptions of code/data-driven differ from those of text-driven systems*
    - discrete, mutually exclusive definitions of variables (code/data)
    - open texture and essentially contestable nature of fundamental concepts
    - explanation (causality) or interpretation (meaning constitution)
    - discrete behaviours or meaningful interaction
    - behaviourism and utilitarianism or hermeneutics and pragmatism

The new rules we live by
the marriage of nudge theory and ML

- This is where nudge theory and ML meet and join forces
  - The inversion of the relationship between proxies and what they stand for
    - *Atomistic primitives of human behaviour as foundational instead of*
    - *The concepts we live by as foundational*

The new rules we live by
the marriage of nudge theory and ML

- Nudge theory and ML **inverse the relationship between** proxy and what it stands for

The new rules we live by
the marriage of nudge theory and ML

- Both treat people as manipulable entities:
  - *Humans may be irrational but at least we can predict them (Ariely)*

- This is both naïve and dangerous:
  - *double contingency, mutually constitutive anticipation*
  - *Goodhart effect, Campbell, Lucas critique*

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

China, US and EU

- China: AI for good (and the state/CP will decide what is good)
- US: market-driven solutions (though this may change, cp Khan at FTC etc.)
- EU: wants to have the best cake and eat it too

China, US and EU

- The race to lead AI at a global scale?
- The race to define AI regulation at the global level?
  - *Geopolitical competition to regulate AI (Smuha)*
  - *EU's Brussels effect (Bradford)*

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

The AI Act

- Defining AI:
  - *Not as a research domain*
  - *Not as an intelligent agent that has something to lose*
  - *Not taking sides in the wars between code- and data-driven*

The AI Act

- Defining AI:
  - *Software (including embedded) systems*
  - *ML, Logic Based, Statistical, Search*
  - *That generate an output that interacts with their environment*
  - ***Focus on impact***

The AI Act

- Focusing on:
  - *decision systems AND robotics*
  - *decisions AND behaviour of AI systems*
  - *health, safety and fundamental rights*
  - *4R: resilient, robust, reliable and responsible AI systems*

The AI Act

- Dedicated chapter on high risk systems, requiring providers to implement:
  - *A quality management system*
    - for when used for the intended purpose
  - *a risk management system*
    - for when used for intended purpose AND other foreseeable use
  - *Data and data governance*
  - *Technical documentation and record keeping*
  - *Transparency to those deploying the system*
  - *Human oversight*
  - *Accuracy robustness and cybersecurity*

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

Data- and code-driven 'law'

- Data-driven:
  - *Advanced legal search (courts, law firms, inhouse)*
  - *Prediction of judgments (case load management)*
  - *Law as text as data (denaturalisation of the law)*
  - *NLLP: proxies and bias*
    - Past decisions as ground truth?

Data- and code-driven 'law'

- Code-driven:
  - *Legal knowledge expert systems*
  - *Smart regulation and smart contracts (blockchain)*
  - *Rules as code, RuleSpeak (RegelSpraak Bd)*
  - *Automated decision making by public administration*

Data- and code-driven 'law'

- Hybrid (e.g. prior knowledge, constraint based ML)

Data- and code-driven 'law'

- Data- and code-driven 'law' have different affordances
- They will transform the 'deep structure' of the law
- Initiating a new mode of existence
  - *Not affording the kind of legal protection that builds on*
    - The ambiguity and multi-interpretability of natural language
  - *Requiring to design for legal protection*

What's new?

- The rules we live by: law and the rule of law
- The new rules we live by: the marriage of nudge theory and machine learning
- China, the US and the EU
- The AI Act
- Data- and code-driven 'law'
- Legal Protection by Design and Question Zero

Question Zero

- Question zero:
  - *What problems does this solve?*
  - *What problems does it NOT solve?*
  - *What problems are created?*

- This is NOT a CBA
  - *which requires proxies, is based on utilitarian calculus*
  - *which requires quantification of what may not be computable*
  - *to quantify you need to qualify*

Question Zero

- If we decide – as a people, as a polity, as a jurisdiction – to
  - *Answer positively Question Zero*
- We will have to build legal protection into
  - *The code- and data-driven architectures we deploy*
- That is why it is crucial that BOTH code- and data-driven ’law’
  - *Are qualified as high risk systems in the AI Act*

Legal Protection by Design

- From legal protection based on **text-driven ICT**
- To legal protection at the level of relevant **data- and code-driven ICT**

This is NOT about ethics

- We do not want to depend on the ethical inclinations of developers
- Or those who fund them (big tech)

Legal Protection by Design

This is NOT about ‘legal by design’ (which is an oxymoron)

Legal protection by design:

- integrating the checks and balances of the rule of law
- based on democratic law-making
- combining **legal certainty** with **contestability**

THE END