# RECSYS AND THE ISSUE OF PROXIES

Mireille Hildebrandt

# THE ISSUE OF PROXIES AND CHOICE ARCHITECTURE

Mireille Hildebrandt

# THE ISSUE OF PROXIES AND CHOICE ARCHITECTURE (THE FRAMING POWERS OF ML)

Mireille Hildebrandt

# Profiling the EU citizen (2008):

https://link.springer.com/book/10.1007/978-1-4020-6914-7

- Putting profiling on the agenda
  - *In terms of data mining and KDD*
  - *Impact on privacy and identity*

- Bringing together CS, engineering, legal, political and social science perspectives

- Doing so by way of cross-disciplinary replies

- Relation with prohibition of ADM in DPD

Mireille Hildebrandt
Serge Gutwirth
*Editors*

## Profiling the European Citizen

*Cross-Disciplinary Perspectives*
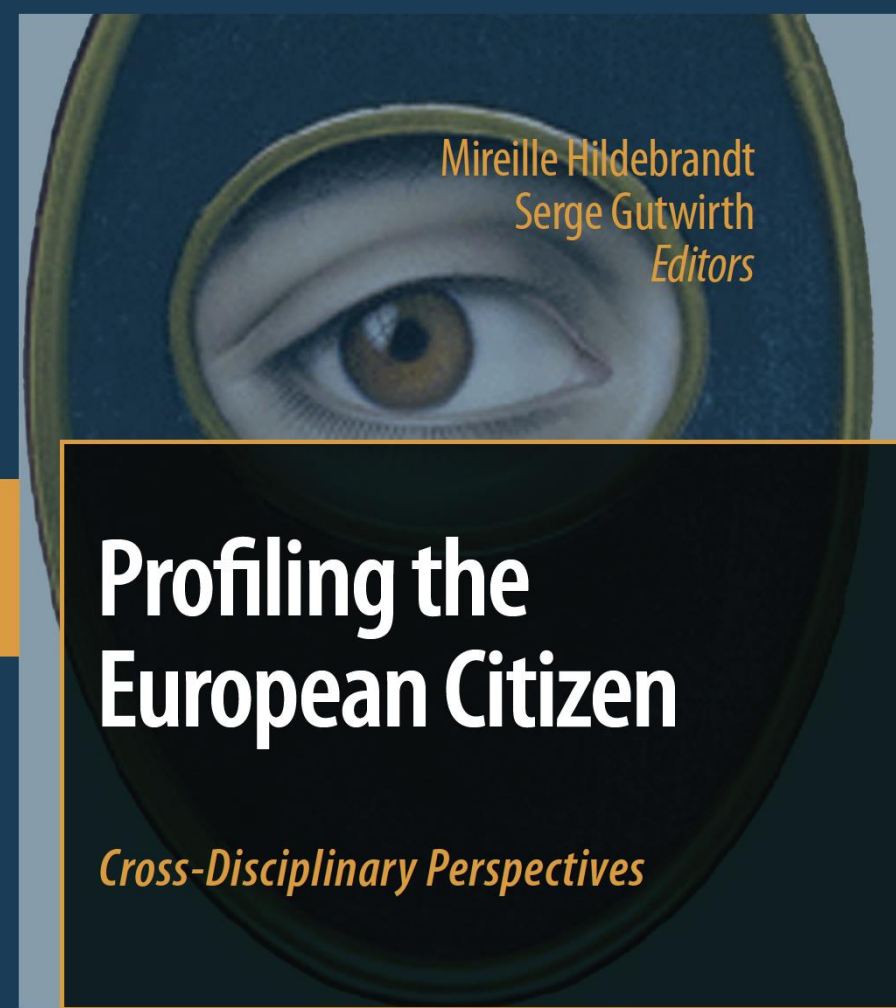
Springer

# The End(s) of Law (2015):

https://www.e-elgar.com/shop/gbp/smart-technologies-and-the-end-s-of-law-9781786430229.html

How do smart technologies

■ undermine, reconfigure and overrule the ends of the law in a constitutional democracy,

■ jeopardizing law as an instrument of justice, legal certainty and the public good (Radbruch).

A Call to

■ not to reject smart technologies,

■ explaining how further engaging them

■ may help reinvent the effective protection of RoL

Smart Technologies and the End(s) of Law

Mireille Hildebrandt

# Why lawyers should get their copy (2019):

https://fdslive.oup.com/www.oup.com/academic/pdf/openaccess/9780198860884.pdf

- Explaining law to non-lawyers offers potentially new self-understanding to lawyers

- Dedicated focus on legal domains that interact with data- and code-driven systems

- Helping lawyers explain the specifics of legal reasoning

- It is open access

# 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

**MACHINES WE TRUST**

Perspectives on Dependable AI

edited by Marcello Pelillo and Teresa Scantamburlo

# 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

## 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

# What's new?

1. Defining RecSyss

2. A political economy of recommender systems (Polanyi or Buchanan)

3. The business and math of persuasion (Goodhart and Ricoeur)

4. Concepts, behavioural atoms and proxies: *the inversification of proxy relations*

5. The choice architecture provided by EU law

# What's new?

1. Defining RecSyss

2. A political economy of recommender systems (Polanyi or Buchanan)

3. The business and math of persuasion (Goodhart and Ricoeur)

4. Concepts, behavioural atoms and proxies: *the inversification of proxy relations*

5. The choice architecture provided by EU law

# Defining RecSyss

■ Solutions to information overload

- *search engines* (the PageRank algorithm)
- *streaming providers* (music, movies, podcasts)
- *social networks* (rankings in timelines, news feeds),
- *webshops* (suggesting similar products or services),
- *platforms mediating the gig economy* (recommending short stays or car rides)
- *cyberphysical infrastructures* (IoT, smart cities, connected cars, smart energy grids)
- *behavioural advertising ecosystem* (that grounds our current infosphere)

# Defining RecSyss

- Solutions to information overload
  - *without RecSyss we could not navigate online and onlife environments*
  - *RecSyss are crucial means to cope with current infosphere*

  - *means co-determine the ends (Dewey), neither good nor bad but never neutral (Kranzberg)*

# Defining RecSyss

■ Solutions to information overload

- *search is a specific type of RecSys*
- *ranking is key to RecSyss*
- *ranking is about relevance*
- *optimisation depends on relevance*


- *Relevance is the vanishing point of information retrieval*

# Defining RecSyss

■ Solutions to information overload

- – *filtering mechanisms tuned to types of users, content or both*
- – *inferencing from behavioural data or based on explicit users preferences*
- – *collaborative filtering (inferences based on aggregate user behaviour),*
- – *content-based (retrieval of relevant documents in the context of a specific domain)*
- – *knowledge-based (often constraint-based, with input provided by domain experts)*
- – *various types of hybrids*

# What's new?

1. Defining RecSyss
2. A political economy of recommender systems (Polanyi or Buchanan)
3. The business and math of persuasion (Goodhart and Ricoeur)
4. Concepts, behavioural atoms and proxies: the inversification of proxy relations
5. The choice architecture provided by EU law

# Political economy of RecSyss

- The 'workflow' or 'pipeline' of RecSyss:

  - *Recommendations are 'made by' RecSyss*
  - *'On behalf of' whoever provides and/or deploys them*
    - Often part of same platform
  - *To deployers or end-users*
    - Recommendations medical treatment: to doctor, consequences affect patient
    - Recommendations credit rating: to insurance company, affecting applicant
    - Recommendations movies: to end-users
    - Recommendations behavioural ads: to advertisers, affecting webvisitors

# Political economy of RecSyss

- 5 types of actors
  - *Developers of RecSyss*
    - Their design decisions determine the upstream choice architecture
  - *Intermediaries that provide RecSyss, often part of big tech platforms*
    - they have access to the behavioural data
  - *Providers of products or services, e.g. advertisers, webshops etc.*
    - In the case of vertical integration they are part of the same platform
  - *Publishers that host recommendations with regard to products or services*
    - They enable large scale personal data collection by allowing tracking
  - *End-users or those targeted*
    - Their choice architecture is determined elsewhere, they often have no idea

# Political economy of RecSyss

■ What incentives 'drive' these actors?
  – *Whose preferences prevail?*
  – *What assumptions are made about human agents?*
  – *How come efficiency trumps relevance, though it depends on relevance?*
  – *Who decides on the proxies for relevance?*

# Political economy of RecSyss

- Economic markets are not 'natural phenomena'
  - *They are created by way of the performative effect of speech acts*
  - *That produce <span style="color:red">institutional facts</span>, such as:*
    - Property rights (erga omnes)
    - Freedom of contract (ad personam)
    - Duties to perform the contract and to compensate damages (breach of contract, tort)
    - Marriage, incorporation of a company (erga omnes)
    - Public law duties based on administrative law (from tax law to environmental law)
    - Human rights (freedom to conduct a business, right to non-discrimination)

# Political economy of RecSyss

- Economic markets are an affordance of the law:
  - *Law is not a brute fact but an institutional fact*
  - *Created by the performative effect of written and spoken legal speech acts*
  - *The relationship between speech act and law =*
    - not causal but constitutive
    - the performative effect is what we call 'the legal effect' (*Rechtsfolge*)

- Legal speech acts enable and constrain markets

# Political economy of RecSyss

- What incentives 'drive' the actors?
    - *Whose preferences prevail?* <span style="color:red">*How do we know?*</span>
    - *What assumptions are made about human agents?* <span style="color:red">*Behaviourist cynicism?*</span>
    - *If efficiency trumps relevance, on which it depends:* <span style="color:red">*relevant for whom*</span>?
    - *Who decides on the proxies for* <span style="color:red">*relevance, why in terms of preferences*</span>?

# What's new?

1. Defining RSs

2. A political economy of recommender systems (Polanyi or Buchanan)

3. The business and math of persuasion (Goodhart and Ricoeur)

4. Concepts, behavioural atoms and proxies: the inversification of proxy relations

5. The choice architecture provided by EU law

# Business of math and persuasion

- From Vance Packert (Hidden Persuaders) to Nir Eyal (Hooked)

- Conflating ML and nudge theory

- The advertisers 'buy' into the narrative, rather than the voters or consumers

- RecSyss are rewarded for influencing preferences while catering to them

# Business of math and persuasion

- Upstream design choices with major downstream implications:
  - *Translating a purpose into a machine consumable task*
  - *This involves formalisation/ground truth/proxy:*
    - input data about what items/services/opinions <span style="color:red">they engaged with previously</span> (proxy for 'engaged' e.g. click or purchasing behaviour) or
    - input data about what items/services/opinions <span style="color:red">similar end-users engaged with</span> (proxy for 'similar end-users' e.g. those using the same platform).

# Business of math and persuasion

■ Upstream design choices with major downstream implications:

  – *Prime example of the Goodhart effect:*

    ■ use of a measure (the distribution of relevant data)

    ■ as a target (steering end-users towards a similar distribution in the future)

[Goodhart/Strathern: when a measure becomes a target it ceases to be a good measure]

# Business of math and persuasion

- Upstream design choices with major downstream implications:
    - *'the distribution of relevant data' can be compressed,*
    - *into different mathematical functions,*
    - *depending on the proxy for relevance (the target variable, the ground truth)*
    - *resulting in different recommendations;*
    - *'the' distribution does not exist, unless in the form of replicating the entire data set*

# Business of math and persuasion

- Human interaction is prone to what Parsons and Luhman called **'double contingency'**

- Their 'double contingency' goes back to **G.H. Mead's 'generalisable other'**

- A more rich and complex concept on **mutual double anticipation (baseball game)**

- Similar to **Ricoeur's 'oneself as another'** (alternating between 'I' and 'me' via the other(s))

- Machine prediction may be used as a proxy for human anticipation

  - But it should never be confused with the latter

# What's new?

1. Defining RSs

2. A political economy of recommender systems (Polanyi or Buchanan)

3. The business and math of persuasion (Goodhart and Ricoeur)

4. Concepts, behavioural atoms and proxies: the inversification of proxy relations

5. The choice architecture provided by EU law

A mathematician, a physicist, and an astronomer were traveling north by train.

They had just crossed the border into Scotland, when the <span style="color:red">astronomer</span> looked out of the window and saw a single black sheep in the middle of a field.

- "Scottish sheep are black," he remarked.

- "No, my friend," replied the <span style="color:red">physicist</span>, "Some Scottish sheep are black."

- At this point the <span style="color:red">mathematician</span> looked up from his paper and said:

    "In Scotland, there exists at least one field, in which there exists at least one sheep, at least one side of which is black."

    – *Tal Yarkoni 2020 The Generalisability Crisis*

- Take a dedicated RecSys that recommends:
    - *Statutory law relevant in a specific case*
    - *Case law relevant in a specific case*
- This is what commercial legal search engines now offer
- They will take legal concepts and find machine consumable proxies
    - *For constellations of facts and for relevant legal issues*
    - *This will be done based on NLP or rather NLLP*
    - *Based on legal text corpora (training data)*
    - *that supposedly provide what ML calls the ground truth*

# Inversification of proxy–principal relations

- Behaviourism (Pavlov, Skinner, Watson) underpinning behavioural economics:
  - *The primitive is an observable behavioural primitive*
  - *The proxy is a natural language concept (vague, imprecise, ambiguous)*
  - *Cognitive bias distracts from the primitives, need to be removed*

- Machine learning
  - *Fairness or justice are impossible concepts: vague, imprecise, ambiguous*
  - *The proxy is a machine readable distribution deemed to be fair or just*
  - *Or fairness/justice are just proxies for a fair distribution?*

# Inversification of proxy–principal relations

- A map is a proxy for a territory, domain or concept

- Developing a map is productive, it helps to navigate

- Mistaking the map (compression) for the territory creates blind spots

- Awareness of the reduction is key to maps being helpful

- Mapping can be done in many ways, it provides framing powers

- Mistaking the proxy for relevance for relevance itself has two implications:
    - *Hiding the framing power (who do the framing, whose options are framed)*
    - *Preempting the discussion on consequences (as these are 'made' inevitable)*

# What's new?

1. Defining RSs

2. A political economy of recommender systems (Polanyi or Buchanan)

3. The business and math of persuasion (Goodhart and Ricoeur)

4. Concepts, behavioural atoms and proxies: the inversification of proxy relations

5. The choice architecture provided by EU law

# Choice architecture of EU law

- Who does the framing? EU legislature

- Whose choices are framed? E.g. Controllers (GDPR), Providers (AI Act)

- To protect whose choices? natural persons in the EU

# Choice architecture of EU law

■ Data Protection

    – *choice architecture of controllers – and processors*

■ Charter of Fundamental Rights (and the European Convention of Human Rights)

    – *choice architecture of natural persons in the EU*

■ AI Act and other parts of the EU Digital Strategy

    – *choice architecture of providers of AI systems, service providers etc.*

# Choice architecture of EU law

Data Protection: GDPR

- Conditions for fair, transparent and lawful processing of personal data

- Principles: purpose limitation, data minimisation, accountability

- Legal basis: 6 ways to 'ground' processing (always based on necessity)

- Transparency requirements

- Risk approach: DPbD and DPIA

- Accountability: fines and private law liability

- Brussels effect: those wishing to compete on the EU market will adapt

# Choice architecture of EU law

Charter of Fundamental Rights of the EU

Art. 7 Privacy

Art. 8 Data Protection

Art. 10 Freedom of thought, conscience and religion

Art. 11 Freedom of expression and information

Art. 16 Freedom to conduct a business

Art. 21 Prohibition of discrimination

Art. 52.3 aligns the scope of rights with that of the ECHR

# Choice architecture of EU law

**Digital Services Act** <span style="color:red">**Recommender system transparency**</span> **art. 24a**

1.  Online platforms shall set out in their terms and conditions and via a designated online resource that can be directly reached and easily found from the online platform's online interface when content is recommended, in a clear, accessible and easily comprehensible manner <span style="color:red">the main parameters used in their recommender systems</span>, as well as any options for the recipient of the service <span style="color:red">to modify or influence those main parameters</span> that they have made available.

2.  The main parameters referred to in paragraph 1 shall include, at a minimum:

    a)  the main criteria used by the relevant system which individually or collectively are most significant in determining recommendations;

    b)  the relative importance of those parameters;

    c)  what objectives the relevant system has been optimised for; and

    d)  if applicable, an explanation of the role that the behaviour of the recipients of the service plays in how the relevant system produces its outputs.

    The requirements set out in paragraph 2 shall be without prejudice to rules on protection of trade secrets and intellectual property rights.

# Choice architecture of EU law

**Digital Services Act <span style="color:red">Recommender system transparency</span> art. 24a**

3. Where several options are available pursuant to paragraph 1, online platforms shall provide a clear and easily accessible function on their online interface allowing the recipient of the service to select and to modify at any time their preferred option for each of the recommender systems that determines the relative order of information presented to them.

# Choice architecture of EU law

Digital Services Act **Recommender system transparency** art. 29

- 1. ***In addition to the requirements set out in Article 24a,*** very large online platforms that use recommender systems shall ***provide*** at least one ***recommender system*** which is not based on profiling, within the meaning of Article 4 (4) of Regulation (EU) 2016/679***, as well as an easily accessible functionality on their online interface allowing the recipient of the service to select and to modify at any time their preferred option for each of the recommender systems that determines the relative order of information presented to them***.

# Choice architecture of EU law

Proposed AI Act

- Mainly targets high risk AI systems:
    - *That have a potentially high impact on physical safety or health (ANNEX II)*
        - Eg medical devices, aircraft, toys etc.
    - *That have a potentially high impact on fundamental rights (ANNEX III)*
        - Eg when intended for deployment in context recruitment, insurance, policing

# Choice architecture of EU law

Proposed AI Act

- Spells out a series of conditions (requirements) that must be met
  - *Before an AI system is placed on the market or put into service in the EU*
  - *Mainly addressing the providers*
  - *Who must <span style="color:red">conduct and document</span> a Conformity Assessment (CE label)*
  - *Violation of the conditions (requirements) can result in*
    - high fines (up to 30 million euro or 6% global turnover)
    - private law liability (still awaiting the update of the Product Liability Directive)

# Choice architecture of EU law

Proposed AI Act

■ Spells out a series of conditions (requirements) that must be met:
- *A dedicated risk management system must be in place*
  - ■ Risk of deployment for intended <span style="color:red">[and other reasonably foreseeable]</span> purposes
- *Data governance*
  - ■ Relevance of training, validation and test data, bias monitoring, GDPR data minimisation
- *Technical documentation and record keeping*
  - ■ Including automated logging
- *Transparency for those who deploy the systems*
- *Human oversight*
  - ■ By design or by way of instruction

Light at the End
of the Tunnel?