

A low-angle, upward-looking photograph of several modern skyscrapers with glass facades, set against a pale, overcast sky. The buildings are arranged in a way that creates a sense of height and scale. The image is framed by a thick black L-shaped border on the left and bottom sides.

THE ISSUE OF PROXIES AND CHOICE ARCHITECTURE

Mireille Hildebrandt



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

Mireille Hildebrandt

COHUBICOL

Counting as a Human Being in the Era of Computational Law

Say cubicle ▪ Think Wittgenstein's cube

[Learn more](#)

**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)


1. What **matters** is incomputable
2. It can nevertheless be made computable
3. In different ways – and that difference **matters**

- To make things computable you need proxies (e.g. justice: fairness)
- You need a **tertium comparationis** (e.g. outcome equality, i.e. a fair share)
- That allows you to compare, rank and calculate (e.g. income = low hanging fruit, Sen)
- This is how economics works (see discussion of BNP, SDG)
- Qualification necessarily precedes quantification (Callon and Law on 'qualculation')

Sabine-Krolak-Schwerdt-Lecture

Luxemburg, 14 June 2022

European Association of Data Scientists (EuADS)

A portrait of Mireille Hildebrandt, a woman with short, wavy grey hair, smiling warmly. She is wearing a red beaded necklace. The background is dark and out of focus.

**The Whisper Challenge - The
Rule of Proxies in Data Science**

Mireille Hildebrandt (Vrije Universiteit Brussel)

My dedicated interest is in legal tech:

- A market with eager investment but few users?
- We @cohubicol have built a **Typology of legal techs** and conclude:
 - *Lots of vague claims with little substantiation?*
 - *A fundamental disconnect between 'reckoning and judgment'?*
 - Discretion and interpretation are key in human judgement
 - Discretion is not about arbitrary decision making, it is the opposite (Dworkin v Hart)
 - Disambiguation and prior interpretation are key to machine judgement
 - Computer code is developed and then it executes [no discretion]
 - Natural language is adaptive due to its 'open texture' - computer code is not

- E.g. Westlaw Edge seems to use (NLP task):
 - **Named entity recognition**
 - *‘Named entity recognition is a subtask of information extraction that seeks to locate and classify named entities mentioned in unstructured text into pre-defined categories such as **person names, organizations, locations, medical codes, time expressions, quantities, monetary values, percentages, etc.***
 - *Named entity recognition aims to find highly variable information (such as names) that tend to occur in similar contexts. For example the word ‘Mr.’ is often followed by a name.’*
 - © cohublicol Typology of Legal Tech

- E.g. Westlaw Edge seems to use (NLP task):
 - A text classification system:
 - ‘an algorithm that adds one or more labels from a predefined set of labels to a piece of text or a document. The most common use cases are 'thematic classification' ("this text talks about this/these topics") but it can also be used for type classification ("this document is a 'court ruling' while another is a 'contract'), or prediction ("based on this description of the trial, the judgement will be X or Y").
 - The underlying model is trained on a manually annotated dataset of examples, i.e. texts with their associated labels. The quality of these annotations has a great impact on the performance of the trained model’.
 - © cohublicol *Typology of Legal Tech*

- E.g. Westlaw Edge seems to use (NLP Task):
 - **A text retrieval system:**
 - ‘a set of algorithms used to help find relevant information in a document or a set of documents. It transforms **the user's input** into a machine-processable representation which is then compared to the precomputed representations of the documents in the collection.
 - This comparison results in a ranked and sorted list of documents that is shown to the user in the interface’
 - © *cohubicol Typology of Legal Tech*

- E.g. Westlaw Edge seems to use various NLP tasks but:
 - 'only a combination of rules, ML approaches and editorial review could ensure high quality output' (referring to litigation analytics)

(Vacek et al., 2019)

- Why pay attention to these **upstream design decisions**?
- They make a difference:
 - *they frame the output*
 - *due to the proxies they select/develop*
 - *resulting in often invisible bias that may be highly problematic*

What's new?

1. The issue of bias (the framing powers of ML)
2. The issue of proxies (the framing powers of ML)
3. The inversification of proxy relations (the pitfalls of behaviourisms)
4. The choice architecture provided by EU law (highjacking the concept)

The issue of bias

- Training *an algorithm on* training data:
 - *Seeking mathematical patterns (unsupervised)*
 - clustering, contingent on the distribution of the data (**training data** = proxy)
 - the choice of the training data and its curation is a key upstream design decision
 - If you train on patient treatment data to improve medical diagnosis:
 - *It may turn out that you got the wrong dataset*
 - *Because doctors use insurance-related codes to document treatment*
 - *And they will often choose a code that results in compensation*

The issue of bias

- Training *an algorithm on* training data:
 - *Seeking relevant mathematical patterns (supervised)*
 - labelling, contingent on labelling instructions and attribution (**label** = proxy)
 - the choice of the label is a key upstream design decision
 - *if you label certain text as 'negative' to measure impact on attitude*
 - *you may get things wrong as human discourse rarely fits the dichotomy*
 - *by using that measure as a target you may nevertheless impact human action*

The issue of bias

- Training *an algorithm on* training data:
 - *Seeking mathematical patterns that result in the right output (reinforcement)*
 - backpropagation, contingent on the goals set (goals = proxy)
 - the choice of the goals is a key upstream design decision
 - *efficient navigation in autonomous driving*
 - *efficient navigation may result in casualties*
 - *mistaking a pedestrian for a plastic bag*

What's new?

1. The issue of bias (the framing powers of ML)
2. The issue of proxies (the framing powers of ML)
3. The inversification of proxy relations
4. The choice architecture provided by EU law

REVIEW article

Front. Artif. Intell., 28 April 2022 | <https://doi.org/10.3389/frai.2022.789076>



The Issue of Proxies and Choice Architectures. Why EU Law Matters for Recommender Systems

 **Mireille Hildebrandt**^{1,2*}

¹Institute of Computing and Information Sciences (iCIS), Science Faculty, Radboud University, Nijmegen, Netherlands

²Research Group Law Science Technology & Society (LSTS), Faculty of Law and Criminology, Vrije Universiteit Brussel, Brussels, Belgium

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 optimizes for lucrative preferences and thus co-produces the preferences they mine. This relates to the well-known problems of feedback loops, filter bubbles, and echo chambers. In this article, I discuss the implications of the fact that computing systems necessarily work with proxies when inferring recommendations and raise a number of

The issue of proxies

- One thing 'standing in' for another:
 - *in mathematics numbers don't necessarily 'stand in for' something else*
 - *E.g. $-6 - 3 = -9$ (what, apples?), or square root of 2*
 - *In statistics and applied math (social science, computer science):*
 - A variable (x, y, z) stands for a feature/category/type with dedicated values:
 - a symbol (usually a letter) standing in for an unknown numerical value in an equation (<https://www.britannica.com/topic/variable-mathematics-and-logic>)
 - algebra (functions, equations)
 - imagine how this enabled abstraction

 **Roger K Moore** @rogerkmoore · 3d

We should never have called it “language modelling” all those years ago; it was (and still is) “word sequence modelling”. Confusion always occurs when you label an algorithm with the name of the problem you’re trying to solve, rather than with what it actually does. @GaryMarcus

What's new?

1. The issue of bias (the framing powers of ML)
2. The issue of proxies (the framing powers of ML)
3. **The inversification of proxy relations**
4. The choice architecture provided by EU law

Inversification of proxy-principal relations

Behaviourism (Pavlov, Skinner, Watson) underpinning **behavioural economics**:

- The primitive (principal) is an observable behaviour
- 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 in the data?**

Inversification of proxy-principal relations

Rational choice theory (Coase, Elstar) underpinning **neoclassical (neoliberal) economics**:

- The primitive (principal) is individual rational choice in the context of game theory
- The proxy is a natural language concept (vague, imprecise, ambiguous)
- Concepts with open texture distract from the primitives, need to disambiguate and discretize

Machine learning

- Fairness or justice are impossible concepts: vague, imprecise, ambiguous
- **The proxy is e.g. a multi agent system based on game theoretical assumptions**
- **Or fairness/justice are just proxies for the outcome of the MAS?**

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. The issue of bias (the framing powers of ML)
2. The issue of proxies (the framing powers of ML)
3. The inversification of proxy relations
4. 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 and institute 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
 - *Risk to fundamental rights and freedoms of natural persons*
- Accountability: fines and private law liability
- Brussels effect: those wishing to compete on the EU market will adapt, because ...

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

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 **conduct and document** 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 **[and other reasonably foreseeable]** 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



End