How 'algoprudence' can contribute to responsible use of ML algorithms

April 16th 2024

Overview

1. Legal framework + AI context

2. Example of 'algoprudence' + legal embedding

3. Q&A

Activities NGO Algorithm Audit

Normative advice commissions
Advising on ethical issues that arise in concrete algorithmic practice through deliberative and diverse normative advice commissions, resulting in algoprudence

Technical tools
Implementing and testing technical tools to detect and mitigate bias, e.g., bias detection tool and synthetic data generation

Knowledge platform
Bringing together experts and knowledge to foster the collective learning process on the responsible use of algorithms, e.g., AI Policy Observatory and white papers

Supported by

SIDNfonds

European **Artificial Intelligence & Society Fund**

Ministerie van Binnenlandse Zaken en Koninkrijksrelaties

Algorithm Audit

Introducing algoprudence

Definition of algoprudence

Specific, case-based and decentralized judgement regarding the responsible use of algorithms

Why algoprudence?
> More democracy, less technocracy in AI
> Deliberative, inclusive and transparent decision-making about normative questions
> Public knowledge about interpretation of open legal norms

algorithms + *jurisprudence*

Algorithm Audit
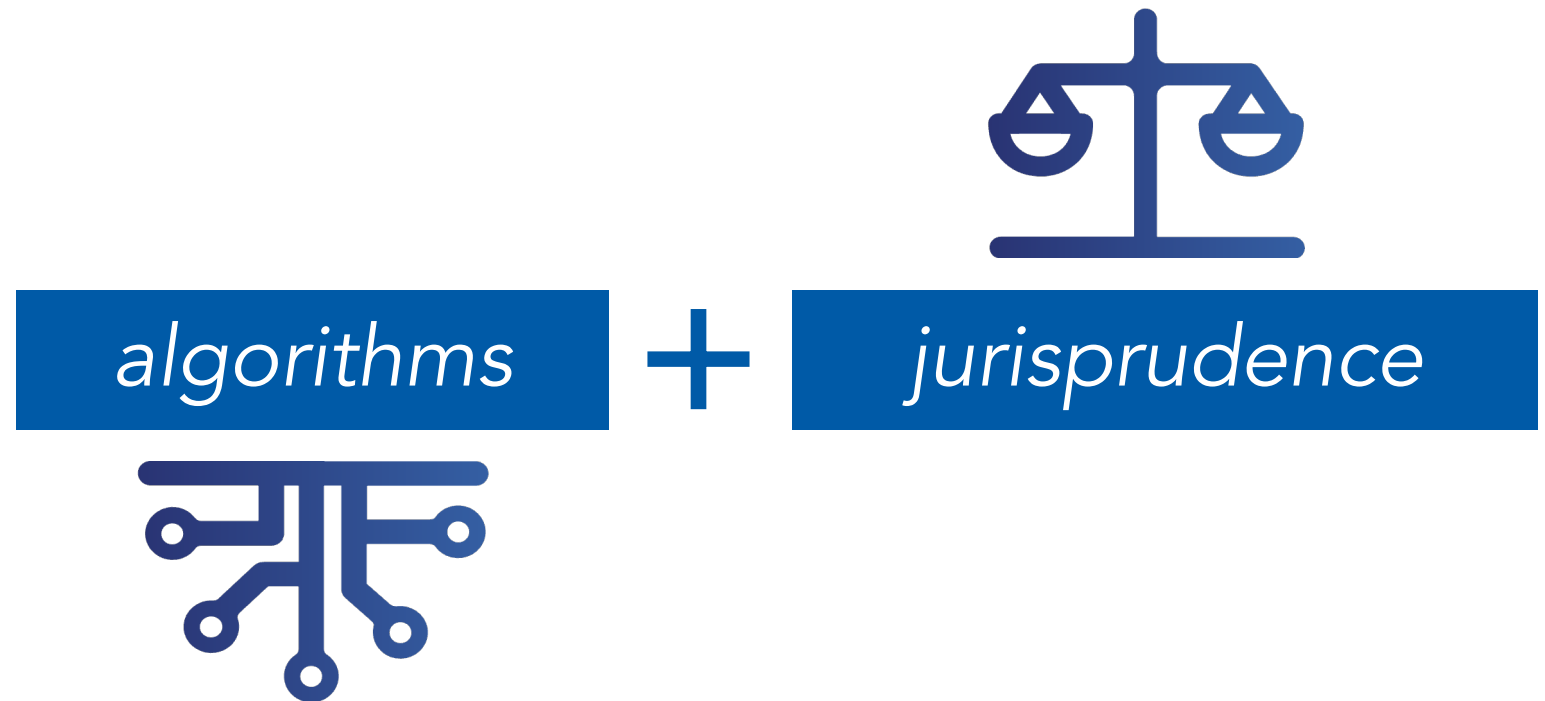
# Dutch General Administrative Law (Awb) codifies some general principles of good administration

General principles of good administration (gpga) (selection)

> Principle of reasoning (*motiveringsbeginsel*)
>> Meta function: whether other principles are satisfied
>> Explain how ML produced an outcome that contributed to decision

> Principle of due diligence (*zorgvuldigheidsbeginsel*)
>> Creating a situation in which all interest can be weighed
>> A suitable method for decision-making should be used

> Principle of fair play (*beginsel van fair play*)
>> Tasks carried out by a PSO without bias
>> Equal treatment of algorithmic-driven decision-support

Use case: ML-driven risk profiling by Dutch public sector organisations (PSOs)

This is just one example of open legal norms that require contextualization. Many other examples exist

Various methods of selecting welfare recipients for re-examination all subjected to gpga

Algorithm Audit
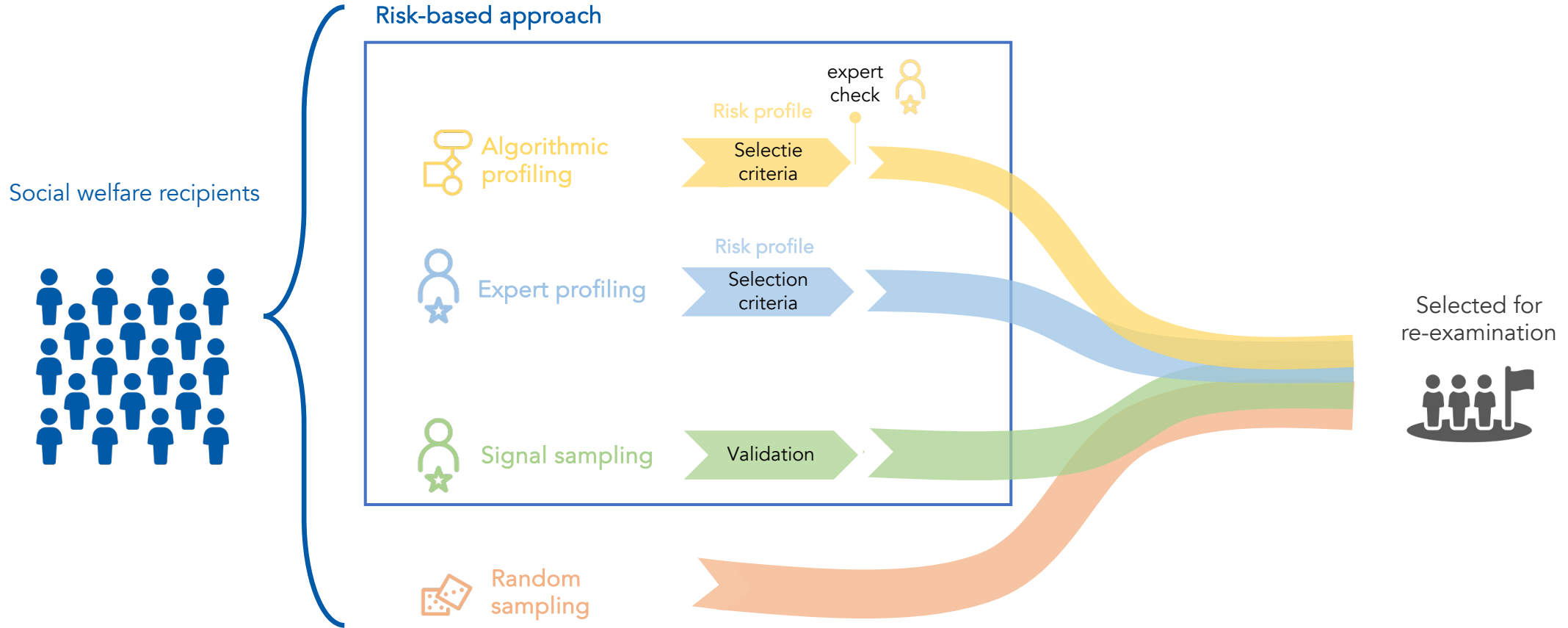
The normative component of expert-driven and algorithmic-driven profiling overlap

Step 1
Available variables in database

Step 2
Variable selection
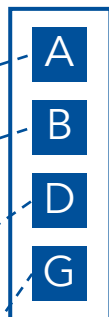
Step 3
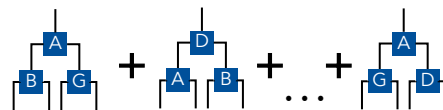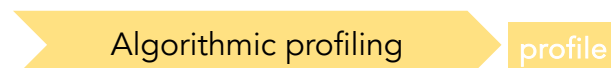Profile composition

Step 4
Expert check on final risk profile

Eligible under certain conditions

Algorithmic profiling → profile

e.g., *explainable boosting* and *extreme gradient boosting*

Expert-driven profiling → profile

Expert check

Ineligible

Algorithm Audit

Three gpga applied on the ML-driven risk profiling use case

> ## Principle of reasoning (*motiveringsbeginsel*)
>> There is no concrete standard for explainability of ML
>> How explainable are gradient boosting and explainable boosting models? ebm is still black box

> ## Principle of due diligence (*zorgvuldigheidsbeginsel*)
>> Is ML a suitable method? If yes, which type of ML and why?
>> Are all relevant data taken into account?
>> What input variables are (in)eligible?
>> Should predictive power of an input variable be taken into account?

> ## Principle of fair play (*beginsel van fair play*)
>> Obligation to detect and mitigate bias in decision-making process
>> Proxy and correlation challenge: higher-dimensional forms of bias resonate with ML logic
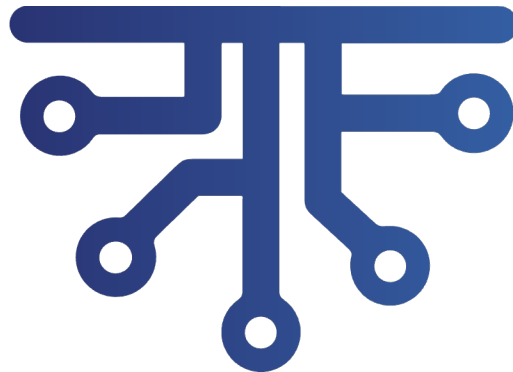>> No silver bullet how to deal with proxies

Institutional impasse

> Courts
>> Limited jurisprudence available
>> Not all cases can be submitted to the courts

> Legislator
>> Too technical to be decided by courts, e.g. which ML method to use
>> Too context-dependent and normative for technical standards

> Supervisory authority
>> Inert
>> Finger pointing

algoprudence

Example of algoprudence: deliberative judgement on Rotterdam case

Citizens subjected to algorithm

Representatives of affected groups

Municipality of Rotterdam

Investigative journalists

Municipal institutions (Rotterdam Court of Auditors and Ombudsperson of Rotterdam)

Legal experts and scientific researchers

1. Initial written reaction

2. Commission gathering
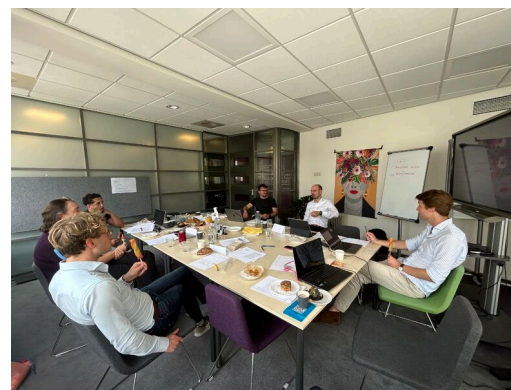
Normative advice commission

Maarten van Asten, Alderman Finance, Digitalisation and Event Municipality of Tilburg

Munish Ramlal, Ombudsperson of Metropole region Amsterdam

Abderrahman Al Aazani, Representative of the Ombusperson of Rotterdam

Francien Dechesne, Associate Professor Law and Digital Technologies, Universiteit Leiden

Oskar Gstrein, Assistant Professor Governance and Innovation, Rijksuniversiteit Groningen

Algoprudence provides flexible and concrete judgements

Step 1

Available variables
in database

Step 2

Variable selection

Eligible under
certain
conditions

Ineligible

Ineligible criteria

ZIP code, city district	⚡
Sex, gender	⊖
Reason for appointment with municipality (annual meeting, intake)	?
Type of contact (mail, phone, text, post)	✂
Literacy rate	⚡
ADHD	⊖
Mental health services	⊖
Number of children	✂
Sectoral (work) experience (hospitality, construction, logistics)	✛
Assertiveness	🔗 ▦
Professional appearance	▦

Eligible criteria

Age	🛞
No show at appointment with municipality	🔗 🛞
Reminders for provinding information	🔗 🛞
Participation in trajectory to work (training, workplace, social duty)	🔗 🛞 ?
Type of living (cohabitation, living together)	🔗 🛞
Cost sharing	🔗 🛞

Legenda

⊖	Legally forbidden	▦	Subjective
🔗	Linkage with aim pursued	✛	Subject to change
✂	No linkage with aim pursued	🛞	Manageable risks
?	Unclear variable	⚡	Proxy discrimination

Rotterdam algoprudence in relation to the general principles of good administration (gpga)

> Principle of reasoning (*motiveringsbeginsel*)
>> Strong explainability requirements, need for discrete categorization

> Principle of due diligence (*zorgvuldigheidsbeginsel*)
>> ML may be used for risk profiling for welfare reexamination under strict conditions

> Principle of fair play (*beginsel van fair play*)
>> List of (in)eligible selection criteria

Algoprudence contextualizes and concretizes open legal norms

Algoprudence as a way of concretizing legal norms

Characteristics of algoprudence

Case-based

Decentralized decision-making

Specific and flexible standards

Considerations are publicly accessible

Non-binding

The effect of algoprudence

Self-regulating	Deviation from the norm

Adds on legislation	Contributes to *state-of-the-art*

Input for positive legal interpretation	> Principle of reasoning (Art. 3:2) > Principle of due diligence (art. 3:46-47) > Principle of fair play (art. 2:4)

Political decision-making	Critical questioning based on an independent standard

Algoprudence: key take-aways

Learning & harmonizing

> Drives collective learning process
> Harmonizes interpretation of national law by local organizations in the context of ethical issues arising when applying algorithms

Questioning & critiquing

> Transparency allows for criticism of normative decision-making in the public use of algorithms
> Opens up space for public debate on normative choices in democratic sight

Inclusion & participation

> Involving various stakeholders in the design of algorithms
> European formula on how AI can be deployed democratically and responsibly by engaging with civil society

Algorithm Audit

What you can do to contribute

1. Comment on algoprudence and share it with those around you

https://algorithmaudit.eu/algoprudence/

2. Submit a case for new algoprudence

https://algorithmaudit.eu/algoprudence/submit-a-case/

3. Participate in our normative advice commissions

info@algorithmaudit.eu

![Algorithm Audit logo]

www.algorithmaudit.eu 🌐

info@algorithmaudit.eu ✉

https://www.linkedin.com/company/algorithm-audit/ 💼

https://github.com/NGO-Algorithm-Audit ⬛