Algorithm Audit

Guidelines for convening normative advice commissions

March 2024

Guidelines for convening normative advice commissions

At Algorithm Audit, we believe decision-making about responsible AI must be transparent and inclusive. This is not only up to technical experts. It's up to all of us.

It is Algorithm Audit's mission to convene diverse and independent normative advice commissions to advise on ethical issues emerging in widely used algorithms. Inspired by established deliberative decision-making processes in other domains, such as ethical review boards emerging in bioethics[1] and academic research[2], and citizens' assemblies covering topics from biometric data usage[3] to climate change[4], our commissions aim to formulate normative advice to navigate complex ethical dilemmas in designing and deploying AI-systems. Since our inception in 2021, we have been experimenting with different formats to gain case-based experience, together with experts and people subjected to AI, in resolving concrete normative dilemmas (Gordian knots). The insights and best practices derived from these experiments are described in this document.

At Algorithm Audit, we uphold an expert-led, inclusive and deliberative approach to formulate case-based normative advice for ethical algorithms. Expertise is crucial in navigating the often complex interactions among legal frameworks, statistical methodologies and ethical considerations inherent in AI-systems. By including diverse stakeholders, such as data subjects and citizen representatives, throughout the entire audit process, we ensure that those most impacted by algorithmic harms have a voice. Although, we realize no silver bullet exists to facilitate inclusive and diverse deliberation. We therefore keep experimenting with new ways of working to provide case-based normative advice.

We hope this document inspires others tasked with normative decisions-making about data modelling. We encourage project teams and ethical review boards to be more transparent about their decisions and the way how these are formed. We welcome feedback on both the guidelines as outlined in this document and secondly on the normative advice we publish on our website, referred to as *algoprudence*.

Together we build public knowledge for ethical algorithms.

Board of Algorithm Audit,

Ariën Voogt, Jurriaan Parie and Samaa Mohammad-Ulenberg

[1] Xafis, V., Schaefer, G. O., Labude, M. K. et al. (2019). 'An ethics framework for big data in health and research'. *Asian Bioethics Review*, 11(3). https://doi.org/10.1007/s41649-019-00099-x

[2] Metcalf, J. and Crawford, K. (2016). 'Where are human subjects in big data research? The emerging ethics divide.' *Big Data & Society*, 3(1). https://doi.org/10.1177/2053951716650211

[3] How to make a Citizens' Biometrics Council (2021) https://www.adalovelaceinstitute.org/feature/how-to-make-citizens-biometrics-council/

[4] For a non-exhaustive overview of various citizens' assemblies see https://www.buergerrat.de/en/citizens-assemblies/citizens-assemblies-worldwide/

Building algoprudence by convening normative advice commissions

Algoprudence refers to the outcome of inclusive and deliberative normative advice commissions. It is a transparent, bottom-up and decentralized way of taking normative decisions how to design and deploy AI-systems[5].

At Algorithm Audit, algoprudence is built in four steps:

1. Case selection – The case committee of Algorithm Audit selects compelling cases for review. A case can be identified by the committee itself or can be submitted by external parties through our website[6]. A compelling case addresses the interaction of a qualitative (legal) and quantitative (statitstical) concept. For instance:
   i. Proxy-discrimination is compelling since the continuous nature of statistical correlation is at odds with the discrete outcome supposed by proportionality assessments as formulated in EU non-discrimination law[7].
   ii. Balancing False Positives (FPs) and False Negatives (FNs) is compelling due to the context-specific nature of interpreting confusion matrix-based evaluation metrics[8]. In a simplified version of reality: in medicine, prioritizing FPs over FNs is common, as it minimizes undiagnosed diseases. Conversely, in the judicial system, priority is given to FNs to minimize wrongful sentencing of innocent people.

   Cases can be the result of AI Fundamental Rights Impact Assessments (FRIAs). Under certain conditions cases submitted by external organisations can be reviewed anonymously. In scouting potential partners for collaboration, Algorithm Audit exercises due diligence in identifying and avoiding potential conflicts of interest with our work.

2. Problem statement –Together with the owner of the algorithm, the case committee collects relevant information about the case under review and proactively reaches out to various stakeholders, among others:
   A. Algorithm developer(s) – Project team of private of public sector organisation designing and deploying the algorithm;
   B. Data subjects – People subjected to the algorithm, for instance inhabitants of Rotterdam that are subjected to risk profiling for social welfare re-examination[9];
   C. Subject matter experts – Experts holding specific domain knowledge about the context in which the algorithm is applied;

[5] Not only the outcome of Algorithm Audit's normative advice commissions is called algoprudence. Other organisations can as well create algoprudence. Other organisations are invited to submit their algoprudence to our case repository. We formally introduce algoprudence in this academic paper.

[6] https://algorithmaudit.eu/algoprudence/submit-a-case/

[7] See section 1.6 Gerards, J., Xenidis, R., Algorithmic discrimination in Europe (2021) https://op.europa.eu/en/publication-detail/-/publication/082f1dbc-821d-11eb-9ac9-01aa75ed71a1.

[8] https://en.wikipedia.org/wiki/Confusion_matrix#Table_of_confusion

[9] AA:2023:02 Risk Profiling for Social Welfare Re-examination https://algorithmaudit.eu/algoprudence/cases/risk-profiling-for-social-welfare-reexamination-aa202302/.

D. Academic experts – Scholars holding specific legal, statistical, ethical or other domain knowledge relevant for the case under review;
E. Civil society organisations – NGOs advocating for specific themes relating to the case under review.

The case committee coordinates the input of the consulted stakeholders. The committee composes an all-encompassing *problem statement* that provides context of the algorithmic application and in which the institutional, technical, legal and ethical dimension are described. A problem statement contains at least the following elements:

i. Introduction – General introduction about the institutional context in which the algorithm is applied and describing the role of the algorithm in this context, e.g., social welfare allowances provided by municipalities, legal obligation for control procedures and how algorithms are used for selecting criteria for risk-profiling;
ii. Focus of case – Scoping the ethical issue occurring when the algorithm is applied, e.g., describing the risk of proxy discrimination given the list of variables that are fed to machine learning algorithm, including explainability properties of the used algorithm;
iii. Legal background – Overview of the applicable legal frameworks to the case, e.g., GDPR, EU non-discrimination law or privacy policy of an organisation. Besides, a summary must be provided of relevant case law. A more rigorous review of jurisprudence can be provided in an appendix.
iv. Description of the ethical issue(s) – Explicitly stating the identified ethical issues as questions.

An example of a problem statement on Risk Profiling for Social Welfare Re-examination can be found in AA:2023:02:P[9]. Problem statements are 6-8 weeks available in Algorithm Audit's case repository for pubic consultation. By submitting a reaction form, everyone can write a respond whether the right questions are addressed. Problem statements require board approval to proceed to the next phase.

3. Convening a normative advice commission – The case committee brings together a normative advice commission that will review the problem statement. As a guideline, the above listed stakeholders A-E under step 2 should be included. The composition of the advice commission requires board approval before the case can proceed to the next step. In scouting potential advice commission members, the case committee of Algorithm Audit exercises due diligence in identifying and avoiding potential conflicts of interest with our work. In our signed agreement with advice commission members, a conflict of interest policy is included that requires them to make known potential conflicts of interest with other roles they may have. The composition of a normative advice commission requires board approval before the commission gathering takes place.

In the composition of its normative advice commissions, special attention is paid to the inclusion of various ethnic and gender backgrounds. This is done to safeguard

Box 1

Experimenting with participation of various stakeholders

There is no universally established method for incorporating people subjected to an algorithm in a normative advice commission. Algorithm Audit therefore experiments with various working formats, among others:

> Include a person subjected to the algorithm as part of the normative advice commission;
> Include people subjected to the algorithm in defining the problem statement, prior to the commission gathering;
> Include people subjected to the algorithm by hosting focus sessions in parallel to the normative advice commission gathering.

The above options are not mutually exclusive. Please reach out to us if you think other options should be taken into account[10].

an inclusive deliberative process and equitable outcome. For evaluating AI systems, the experiences of people subjected to these systems are indispensable. Since those disadvantaged by AI frequently belong to disadvantaged groups that are underrepresented in the class of AI experts, including these disadvantaged groups in the deliberation is imperative. Hence, we experiment with various ways in which their voice is included in the reviews (focus groups, events with local interest groups, inclusion in the advice commission). See also Box 1.

The normative advice commission will execute the following work plan, which is divided in a pre-, in- and post-gathering phase.

> Pre-gathering – Commission members are asked to write an initial reaction on the ethical issues presented in the problem statement. The collection of this primary response is sent to all commission members. Collecting initial reactions is a first, preliminary survey to identify different perspectives on how the ethical issue can be approached. Exchanging viewpoints prior to the gathering streamlines the discussion among attendees during the gathering. No lengthy all-encompassing treatment of the ethical issue is required. Commission members are requested to send an initial reaction of maximum two pages, by an announced date.

> In-gathering: An agenda for the commission gathering will be drafted and shared in advance. The following people are present at the meeting:
  – At least 33% of the board members of Algorithm Audit;
  – At least 66% of the case committee that owns the case review;
  – At least 80% of the normative advice commission members.

  One of us at Algorithm Audit moderates the discussion, with the objective of bringing forth a normative decision about the ethical issue introduced in the problem statement. The meeting is recorded for notetaking and conversation reconstruction purposes during the post-gathering phase.

[10] info@algorithmaudit.eu

> Post-gathering – Based on the notes and recording of the commission gathering, the case committee drafts a preliminary version of the advice document. The document is sent for feedback to all commission members. This process is repeated until the content of the advice document is satisfactory for all. Consensus among commission members regarding the final judgment and advice is not strictly necessary. When no consensus is reached, an elaboration on the existing varying viewpoints will be presented in the advice document, as this provides useful insights regarding the ethical issues discussed. In the final advice document, a disclaimer will be added that the advice document concerns a group effort and that no individual commission member can be held accountable for any statements being made. Normative commission members may not disclose any statements or opinions expressed by individual commission members.

4. Publication of algoprudence – The problem statement and final advice document is published on https://algorithmaudit.eu/algoprudence/.

Oversight of adherence to this policy is assigned to the Supervisory Board of Algorithm Audit.

Timespan for algoprudence creation

Case reviews differ in scale and scope. We apply a multi-tier approach for case reviews:

> Short: timespan 1-2 months, 4 commission members;
> Medium: timespan 2-4 months, 4-6 commission members;
> Long: timespan 4-6 months, 6 commission members.

The number of simultaneous case reviews depends on available resources. The priority of one case over the other can differ depending on urgency of the case under review.

Legal status of algoprudence

Algorithm Audit does not have a mandate to issue legally binding rulings. In our case studies, we provide non-binding ethical advice. Ethical advice often goes beyond advice on what is required for legal compliance. Yet in the absence of regulation or clear standards established by a supervisory body, our independent ethical advice serves as a preliminary signpost for organizations. Our case advice may also help elaborate official standards or support future decisions by legal authorities. In this sense, our ethical advice does have relevance for the legal domain. An in-depth review how Algoprudence is embedded in existing legal frameworks can be found in our scientific legal article: *How 'algoprudence' can contribute to responsible deployment of machine learning algorithms*[11].

[11] [link]

Algorithm Audit

www.algorithmaudit.eu

www.github.com/NGO-Algorithm-Audit

info@algorithmaudit.eu

Stichting Algorithm Audit is registered as a non-profit organisation at the Dutch Chambre of Commerce under license number 83979212