



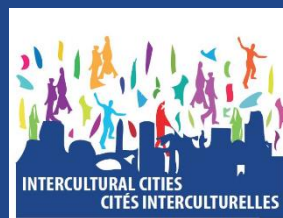
Preventing the potential discriminatory effects of the use of artificial intelligence in local services

Policy Brief

October 2020



ePaństwo
Foundation



COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

*The opinions expressed in this work
are the responsibility of the author
and do not necessarily reflect the official
policy of the Council of Europe.*

Written by Krzysztof Izdebski

Intercultural Cities Unit,
Council of Europe©

Council of Europe, October 2020

Krzysztof Izdebski. Board Member and Policy Director of ePanstwo Foundation (EPF) and Board Member of Consul Democracy Foundation. He is a lawyer specialized in access to public information and re-use of public sector information. He is the author of publications on freedom of information, technology and public administration including "Transparency and Open Data Principles: Why They Are Important and How They Increase Public Participation and Tackle Corruption" and recently published "alGOvrithms. The State of Play. Report on Algorithms Usage in Government-Citizens Relations in Czechia, Georgia, Hungary, Poland, Serbia, and Slovakia.

This policy brief was produced as a background paper based on the webinar organized by the Intercultural Cities Programme of the Council of Europe and Epaństwo Foundation on 21 September 2020.

This is part 3 of the policy brief prepared for the online course Artificial intelligence and anti-discrimination for local authorities by the Intercultural Cities programme. The [full policy brief](#) published online can be found on the Intercultural Cities webpage.

4 How to prevent against discrimination in AI/ADM tools

There are some methods which may help tackle or to minimize the risk of discrimination while using AI/ADM tools.

Examples are **human-centered solutions** embedded in public procurement procedures and **algorithmic impact assessments**.

The World Economic Forum Guidelines for AI procurement put forward the following **10 principles to prevent bias or harm via AI/ADM**.

“Trust-worthy” AI/ADM as defined by the European Commission High-Level Expert Group on AI includes the following principles: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being and accountability. While planning the procurement these principles should also be taken into account.

For example, to secure the transparency of the tool one of the requirements described in the contract notice could include an open-source solution, which means that external experts have the possibility to review software code to reveal potential risks of corruption.

A **practical example** of an introduction to Algorithmic Impact Assessment can be found in the Algorithm Charter For Aotearoa New Zealand risk matrix. The **key elements of a public agency algorithmic impact assessment (AIA)** as described in AI Now Institute Algorithmic Impact Assessments: A Practical Framework For Public Agency Accountability and can be seen below.

This matrix should be used before applying the actual AIA questionnaire which helps to identify risks in more details and is useful to describe concrete discriminatory impact. Participants were presented with an example from Canada. The AIA questionnaire consists of questions such as:

1. Use procurement processes that focus not on prescribing a specific solution but rather on outlining problems and opportunities and allow room for iteration.
2. Define the public benefit of using AI while assessing risks.
3. Align your procurement with relevant existing governmental strategies and contribute to their further improvement.
4. Incorporate potentially relevant legislation and codes of practice in your RFP.
5. Articulate the technical and administrative feasibility of accessing relevant data
6. Highlight the technical and ethical limitations of intended uses of data to avoid issues such as historical data bias.
7. Work with a diverse, multidisciplinary team.
8. Focus throughout the procurement process on mechanisms of algorithmic accountability and of transparency norms.
9. Implement a process for the continued engagement of the AI provider with the acquiring entity for knowledge transfer and long-term risk assessment.
10. Create the conditions for a level and fair playing field among AI solution provider

<ul style="list-style-type: none"> ▪ Does the recommendation or decision made by the system include elements of discretion? 	<ul style="list-style-type: none"> ▪ Does the recommendation or decision made by the system include elements of discretion? 	<ul style="list-style-type: none"> ▪ Will the Automated Decision System use personal information as input data? ▪ What is the highest security classification of the input
--	--	--

<ul style="list-style-type: none"> -Describe what is discretionary about the decision -Is the system used by a different part of the organization than the ones who developed it? -Are the impacts resulting from the decision reversible 	<ul style="list-style-type: none"> Is the system used by a different part of the organization than the ones who developed it? Are the impacts resulting from the decision reversible? How long will impacts from the decision last? 	<p>data used by the system? (Select one)</p> <ul style="list-style-type: none"> Who controls the data? Who collected the data used for training the system? Who collected the input data used by the system?
--	--	---

1. Agencies should conduct a self-assessment of existing and proposed automated decision systems, evaluating potential impacts on fairness, justice, bias, or other concerns across affected communities.
2. Agencies should develop meaningful external researcher review processes to discover, measure, or track impacts over time;
3. Agencies should provide notice to the public disclosing their definition of “automated decision system,” existing and proposed systems, and any related self-assessments and

- researcher review processes before the system has been acquired;
4. Agencies should solicit public comments to clarify concerns and answer outstanding questions; and
 5. Governments should provide enhanced due process mechanisms for affected individuals or communities to challenge inadequate assessments or unfair, biased, or otherwise harmful system uses that agencies have failed to mitigate or correct.

Risk matrix

Likelihood			
<p>Probable Likely to occur often during standard operations</p>			
<p>Occasional Likely to occur some time during standard operations</p>			
<p>Improbable Unlikely but possible to occur during standard operations</p>			
<p>Impact</p>	<p>Low The impact of these decisions is isolated and/or their severity is not serious.</p>	<p>Moderate The impact of these decisions reaches a moderate amount of people and/or their severity is moderate.</p>	<p>High The impact of these decisions is widespread and/or their severity is serious.</p>

Risk rating

<p>Low The Algorithm Charter could be applied.</p>	<p>Moderate The Algorithm Charter should be applied.</p>	<p>High The Algorithm Charter must be applied.</p>
---	---	---

5 Summary

Municipalities which want to prepare for wider implementation of AI/ADM solutions to prevent potential risks of discrimination should:

- Introduce policies on algorithms implementation which described the process and people responsible (ideally multi-disciplinary and diverse team).
- Introduce Algorithmic Impact Assessments.
- Introduce transparency clauses in contracts with companies delivering the software and open access to the source code, if not among the wide public at least among external experts.
- Issue guidelines explaining the operation of algorithms to those who are directly impacted.
- Elaborate on the system of reviewing AI/ADM solutions, again including the multi-disciplinary and diverse team).
- Engage citizens and experts in planning procurement and implementation of AI/ADM which will help to identify potential risks of discrimination.
- Involve knowledge and competencies building schemes for public officials and other municipality employees involved directly or indirectly in using AI/ADM solutions.