47th Closed Session of the Global Privacy Assembly

September 2025

Resolution on meaningful human oversight of decisions involving AI systems

This Resolution is submitted by the Office of the Privacy Commissioner of Canada on behalf of the Working Group on Ethics and Data Protection in Artificial Intelligence (AIWG).

SPONSORS:

- Office of the Privacy Commissioner of Canada

CO-SPONSORS:

- Agencia de Acceso a la Información Pública of Argentina
- Belgian Data Protection Authority
- Office of the Information and Privacy Commissioner of British Columbia
- Bulgarian Commissioner for Personal Data Protection
- European Data Protection Supervisor
- Commission Nationale de l'Informatique et des Libertés, France
- Office of the Data Protection Authority, Bailiwick of Guernsey
- Office of the Privacy Commissioner for Personal Data, Hong Kong, China
- Data Protection Commission of Ireland
- Isle of Man Information Commissioner
- Personal Information Protection Commission, Korea
- Office of the Information and Privacy Commissioner of Ontario, Canada
- Federal Data Protection and Information Commissioner (FDPIC) of Switzerland

The 47th Global Privacy Assembly 2025:

Acknowledging that artificial intelligence (AI) systems are increasingly being integrated into decision-making processes;

Being aware of the potential for a human-centric approach to AI to bring important economic, societal and public interest benefits, including by growing prosperity and addressing pressing global challenges;

Emphasizing the importance of ensuring that the use of AI systems respects human rights, non-discrimination, equity, and justice, and prevents bias;

Recognizing the potential for significant adverse effects that decision-making processes involving AI systems could have on individuals' fundamental rights and freedoms, particularly when the decisions lack meaningful human oversight by individuals able to effectively oversee the relevant AI systems or when there is no effective recourse or opportunity for an individual impacted by the decision ("impacted individual") to effectively challenge such decisions;

Recognizing the importance of ensuring that, where automated decisions may significantly affect individuals' fundamental rights and freedoms, individuals are provided with the possibility to request a timely human review of such decisions;

Acknowledging the benefits of meaningful human oversight of automated decision-making to increase institutional accountability in the development and use of AI systems, improve their reliability, proactively identify and mitigate potential bias in data and algorithms, enhance transparency, contestability, and explainability, and support the improvement and adaptability of AI systems to evolving real-world environments, in order to foster trust amongst data subjects;

Being mindful that without careful consideration in their design and development, automated decisions will reflect patterns found within a system's training data and thus replicate or reinforce past biases or make decisions which assume that the conditions of the past decisions remain true in the present and future;

Recalling that the work plan for the Working Group on Ethics and Data Protection in Artificial Intelligence includes establishing a common understanding of the factors that constitute "meaningful" human oversight of decision-making processes involving AI systems;

Differentiating between the terms "human oversight", which occurs throughout the decision-making process and "human review", which occurs as part of a *post-hoc* review of a decision in which an impacted individual is able to substantiate their point of view, and noting that human review is one activity within the broader human oversight process;

Recognizing that organizations are responsible for establishing their own internal human oversight processes;

Emphasizing that some privacy and data protection laws establish the right for individuals not to be subject to certain fully automated decision-making, and that a decision is considered fully

automated in the absence of sufficient and meaningful human oversight of the decision-making process providing for the data subject's right to obtain human intervention, express their point of view, and contest such decisions;

Emphasizing that intergovernmental standards on AI, such as the OECD Recommendation on Artificial Intelligence (OECD/LEGAL/0449), recognise that organizations that deploy or operate AI systems should implement mechanisms and safeguards, including human oversight, to address risks arising from deviations of the AI system from intended purposes, as well as intentional or unintentional misuse of AI systems;

Recognizing that the type of human oversight required for a decision-making process in which an AI system is used will generally depend on the context for, and potential impacts of, that decision, and that enhanced human oversight and safeguards will be required when AI systems process special categories of personal information or data relating to criminal convictions and offences;

Recalling that the 40th Global Privacy Assembly endorsed the principle that the transparency and intelligibility of AI systems should be improved, including by "providing adequate information on the purpose and effects of artificial intelligence systems in order to verify continuous alignment with expectation of individuals and to enable overall human control on such systems";

Recalling that the 42nd Global Privacy Assembly urged organizations that develop or use AI systems to consider implementing accountability measures including ensuring that a human actor is identified (a) with whom concerns related to automated decisions can be raised and rights can be exercised, and (b) who can trigger evaluation of the decision process and human intervention;

Recalling that, with respect to AI and employment, the 45th Global Privacy Assembly underlined both the importance of enabling a data subject affected by an employer's AI system to obtain recorded, meaningful human review of employment decisions, and the importance of training users of AI tools—including those providing human oversight of decision-making processes involving an AI system;

Taking note that some AI legislation and regulatory guidance identify human oversight as being necessary to effectively mitigate the risks to health, safety, and fundamental rights posed by some AI systems or applications thereof;

Recognizing that the process for human oversight of automated decisions must take into account considerations such as automation bias, in which overseers are overly-trusting of decisions made by AI systems;

Emphasizing that human oversight will not be an adequate remedy for a poorly designed or developed, mis-applied or otherwise fundamentally flawed AI system, and that it remains critical for organizations using AI systems to determine both whether it is appropriate to use an AI system in a given context, and if so, whether the AI system they have chosen will be effective in that context;

Highlighting that human oversight of decision-making processes involving an AI system requires that the overseer has access to all information that is relevant to the decision, presented in a way

that is appropriate to the context of the oversight process, yet limited to what is necessary in the context of the particular oversight taking place (that is, with respect to personal information, the overseer should only have access to that information which is relevant to the decision-making process);

Recognizing that in many instances, the process by which an AI system arrives at a decision will not be readily apparent to humans and that steps must be taken to design the AI system in a way that allows for or increases explainability, with emphasis given to ensuring that both the overseer and affected individuals who may not have technical expertise can understand the decision and ask for its review;

Recognizing also that the ability for individuals to request review of, or exercise other rights or abilities related to, decisions involving AI systems requires that the use of such systems be transparent;

Realizing that in some circumstances meaningful human oversight may not be possible, such as where decision-making processes produce decisions on a scale and timeframe that makes individual monitoring of each impracticable, and that in such situations organizations should consider alternative means of oversight (such as analysis of samples of decisions to ensure the decision-making process is performing as expected);

Recognizing that some or all of the following considerations may impact whether human oversight of a decision involving an AI system is meaningful:

- **Agency**: The organization should design the oversight process in such a way that the overseer has effective control and autonomy that allows them to make decisions and act independently. This includes that the overseer feels comfortable and empowered to exercise their role without fear of repercussion. This may include developing a whistle-blowing procedure;
- **Clarity of role**: The organization should ensure that the overseer is clear about whether their role is to assess a decision made by an AI system, to accept, reject or modify a recommendation made by an AI system, or to consider the output of an AI system as one input among many in their decision-making process, noting that ultimate accountability for the decision remains with the organization;
- **Knowledge and expertise**: The organization should ensure that the overseer has adequate knowledge and expertise to evaluate the AI system's decision, including its appropriateness, accuracy, and the potential impacts the decision may have on the affected individual. In addition, the overseer should also be sufficiently trained to understand the AI system's operations and limitations to be able to identify situations or circumstances in which the AI system's outputs may require additional scrutiny, and to understand factors such as automation bias which may impact their own actions;
- **Resources**: The organization should provide the overseer with the resources necessary to adequately oversee a decision. This should include sufficient time to undertake the oversight and a reasonable workload, information about how the system was trained (including the nature of the training data) and on the logic of the decision-making, relevant data in an interpretable format and appropriate contextual information to support the oversight, and/or access to colleagues, experts, or other resources with whom the overseer can confer. Where appropriate, this may also include allowing the overseer access to the impacted individual to ask clarifying questions;
- **Timing and effectiveness**: The organization should ensure that the oversight occurs at a time and in a manner that permits the overseer to agree with, contest, or mitigate the potential impacts of the

AI system's decision; that is, human oversight of an AI system is unlikely to be meaningful or effective if recourse is only available after the impacts of a decision are experienced by an individual;
- **Evaluation and accountability**: The organization should evaluate overseers based on whether they have diligently performed their prescribed task, not on the outcome of the decision. Accountability for the final decision and its impacts will always remain with the organization.

Recognizing that organizations can take steps to ensure that meaningful human oversight of the decisions involving AI systems is occurring, which may include:

- **Clarifying the intention and value of oversight**: To counter the potential for automation bias and emphasize the value of the oversight role, organizations should make clear to overseers what knowledge or experience they should draw from when reviewing decisions (such as expertise in a field, general life experience, etc.);
- **Training**: In addition to domain expertise and knowledge of the intended uses and limitations of the AI system, organizations should ensure that overseers are trained on concepts such as how to identify and mitigate bias, including both those that impact individuals affected by the AI system and those that impact the overseer themselves (such as automation bias). This training should occur prior to the individual undertaking their oversight role and be regularly revisited;
- **Design of oversight process**: Organizations should design the oversight process to ensure that it is user friendly, including that relevant information is presented in a way that will be useful and understandable in practice. This should be re-examined and, if necessary, improved as overseers gain practical experience, and can be supported by involving overseers in the design process. Organizations should also seek to determine whether overseers are impacted by known issues such as "anchoring", in which an individual's judgement is overly influenced by a particular reference point (such as the first piece of information presented to them);
- **Escalation**: Organizations should put in place measures to appropriately escalate decision-making, including for situations or circumstances identified by the overseer. This may also include designing the automated decision-making process so that above a certain degree of risk, decisions are automatically flagged for mandatory intervention before any action is taken;
- **Documentation**: Organizations should require overseers to document their decisions, particularly in the case in which they reject an AI system's decision. This is done both to allow for the identification of patterns of incorrect decisions from the AI system and in recognition of the fact that bias can be introduced by improperly applied human oversight mechanisms;
- **Assessments**: Organizations should include the nature and extent of human oversight in a data protection impact assessment (DPIA) and any other assessment (such as an Algorithmic Impact Assessment or Fundamental Rights Impact Assessment) of the proposed use of an AI system;
- **Evaluation and testing of oversight process**: Organizations should regularly test the effectiveness of their oversight process. This might include "secret shopping" (in which incorrect decisions are deliberately subjected to oversight to determine whether they are identified);
- **Evaluation of outcomes**: Organizations should regularly evaluate whether there are patterns within AI system decisions that are rejected or overturned by overseers, or incorrect decisions made even after human oversight. This could suggest issues either with the AI system or with the oversight process.

The 47th Global Privacy Assembly therefore resolves to:

1. Promote a common understanding of the notion of meaningful human oversight of decisions that involve AI systems, which includes but is not limited to the considerations set out in this resolution.

2. Urge organizations that use AI systems in decision-making processes to designate overseers who possess the necessary competence, training, resources, and awareness of any contextual information about the decision-making process, as well as understanding of the specific AI system's capabilities, limitations, potential failure modes, and associated risks, including potential biases, and to adopt appropriate human-machine interfaces and interpretability tools to adopt technologies and processes that allow for meaningful human oversight, in particular where these decisions may have significant impacts on individuals' fundamental rights and freedoms.

3. Through the GPA Ethics and Data Protection in Artificial Intelligence Working Group share knowledge and best practices with respect to practical implementation of this notion in the respective applicable legal frameworks and develop resources for DPAs that support their efforts to promote the adoption of the practices described in this resolution by controllers in their respective jurisdictions.

4. Continue to promote the development of technologies or processes that advance explainability for AI systems, recognizing that this will be an important factor in ensuring meaningful human oversight of the decisions of AI systems.

The United States Federal Trade Commission abstains from the adoption of this resolution.