Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems

(Adopted by the Committee of Ministers on 8 April 2020
at the 1373rd meeting of the Ministers' Deputies)

Preamble

The Committee of Ministers, under the terms of Article 15.*b* of the Statute of the Council of Europe,

Considering that member States of the Council of Europe have committed themselves to ensuring the rights and freedoms enshrined in the Convention for the Protection of Human Rights and Fundamental Freedoms (ETS No. 5, "the Convention") to everyone within their jurisdiction and that this commitment stands throughout the continuous processes of technological advancement and digital transformation that European societies are experiencing;

Reaffirming that, as a result, member States must ensure that any design, development and ongoing deployment of algorithmic systems occur in compliance with human rights and fundamental freedoms, which are universal, indivisible, inter-dependent and interrelated, with a view to amplifying positive effects and preventing or minimising possible adverse effects;

Recognising the unprecedented rise in the use of digital applications as essential tools of everyday life, including in communication, education, health, economic activities and transportation, their increasing role in governance structures and the management and distribution of resources and the fact that cross-cutting technologies using algorithmic systems, with the appropriate incentives, have the potential to address important challenges, including climate change and sustainable development;

Conscious therefore of the evolving impact, which may be positive or negative, that the application of algorithmic systems with automated data collection, analytics, decision making, optimisation or machine learning capacities has on the exercise, enjoyment and protection of all human rights and fundamental freedoms, and of the significant challenges, also for democratic societies and the rule of law, attached to the increasing reliance on algorithmic systems in everyday life;

Underlining the need to ensure that racial, gender and other societal and labour force imbalances that have not yet been eliminated from our societies are not deliberately or accidentally perpetuated through algorithmic systems, as well as the desirability of addressing these imbalances through using appropriate technologies;

Bearing in mind that digital technologies hold significant potential for socially beneficial innovation and economic development, and that the achievement of these goals must be rooted in the shared values of democratic societies and subject to full democratic participation and oversight;

Reaffirming therefore that the rule of law standards that govern public and private relations, such as legality, transparency, predictability, accountability and oversight, must also be maintained in the context of algorithmic systems;

Considering that ongoing public and private sector initiatives intended to develop ethical guidelines and standards for the design, development and ongoing deployment of algorithmic systems, while constituting a highly welcome recognition of the risks that these systems pose for normative values, do not relieve Council of Europe member States of their obligations as primary guardians of the Convention;

Recalling the obligation of member States under the Convention to refrain from human rights violations, including through algorithmic systems, whether they are used by themselves or as a result of their actions, and their obligation to establish effective and predictable legislative, regulatory and supervisory frameworks that prevent, detect, prohibit and remedy human rights violations, whether stemming from public or private actors and whether affecting relations between businesses, between businesses and consumers or between businesses and other affected individuals and groups;

Emphasising that member States should ensure compliance with applicable legislative and regulatory frameworks and guarantee procedural, organisational and substantive safeguards and access to effective remedies with regard to all relevant actors, while promoting an environment in which technological innovation respects and enhances human rights and complies with the fundamental obligation that all human rights restrictions be necessary and proportionate in a democratic society and implemented in accordance with the law;

Taking account of, and building on, existing Council of Europe, regional and international norms, standards and recommendations related to the protection of human rights and fundamental freedoms in contemporary societies, as well as the evolving jurisprudence of the European Court of Human Rights;

Reiterating particularly the importance of existing personal data protection standards, notably the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108) as modernised by the Amending Protocol (CETS No. 223), while emphasising that the human rights impacts of algorithmic systems are broader and call for additional protections;

Recalling furthermore that private sector actors, in line with the United Nations Guiding Principles on Business and Human Rights, have the corporate responsibility to respect the human rights of their customers and of all affected parties and that, to this end, flexible governance models should be adopted that guarantee fast and effective reparation and possibilities for redress when incidents occur, ensuring that responsibility and accountability for the protection of human rights are effectively and clearly distributed throughout all stages of the process, from the proposal stage through to task identification, data selection, collection and analysis, system modelling and design, through to ongoing deployment, review and reporting requirements;

Acknowledging the fact that fast-moving socio-technical developments require constant monitoring and the adaptation of applicable governance frameworks to protect human rights effectively in a complex global environment, and recognising the need for regular guidance to be provided to all relevant public and private sector actors,

Recommends that the governments of member States:

1.   review their legislative frameworks and policies as well as their own practices with respect to the procurement, design, development and ongoing deployment of algorithmic systems to ensure that they are in line with the guidelines set out in the appendix to this recommendation; promote their implementation in all relevant areas and evaluate the effectiveness of the measures taken at regular intervals, with the participation of all relevant stakeholders;

2.   ensure that this recommendation, including the guidelines in the appendix, is translated and disseminated as widely as possible and through all accessible means among competent authorities and stakeholders, including parliaments, independent authorities, specialised public agencies, national human rights institutions and equality bodies, civil society organisations and the private sector;

3.   ensure, through appropriate legislative, regulatory and supervisory frameworks related to algorithmic systems, that private sector actors engaged in the design, development and ongoing deployment of such systems comply with the applicable laws and fulfil their responsibilities to respect human rights in line with the UN Guiding Principles on Business and Human Rights and relevant regional and international standards;

4.   endow their relevant national supervisory, oversight, risk assessment and enforcement institutions with the necessary resources and authority to investigate, oversee and co-ordinate compliance with their relevant legislative and regulatory framework, in line with this recommendation;

5.   engage in regular, inclusive and transparent consultation, co-operation and dialogue with all relevant stakeholders (such as civil society, human rights defence organisations, the private sector, the academic and professional community, the media, educational establishments, public libraries, infrastructure providers and basic public services, including social welfare and the police force), paying particular attention to the needs and voices of vulnerable groups, with a view to ensuring that human rights impacts stemming from the design, development and ongoing deployment of algorithmic systems are comprehensively monitored, debated and addressed;

6.   prioritise the building of expertise in public and private institutions involved in integrating algorithmic systems into multiple aspects of societies with a view to effectively protecting human rights;

7.   encourage and promote the implementation of effective and tailored media, information and digital literacy programmes to enable all individuals and groups to

      a.   understand the functions and ramifications of systems employing automated decision making,
      b.   make informed decisions in the use of such systems,
      c.   enjoy the benefits arising from the use of algorithmic systems, and
      d.   minimise the exposure to threats and risks stemming from the use of such systems,

in effective co-operation with all relevant stakeholders, including from the private sector, the media, civil society, educational establishments, and academic and technical institutions;

8.   take account of the environmental impact of the development of large-scale digital services and take the necessary steps to optimise the use and consumption of natural resources and energy;

9.   review regularly, in consultation with all relevant actors, and report domestically and within the Committee of Ministers on the measures taken to implement this recommendation and its guidelines with a view to enhancing their effectiveness and adapting them to evolving challenges.

Appendix to Recommendation CM/Rec(2020)1

Guidelines on addressing the human rights impacts of algorithmic systems

A. Scope and context

1.   These guidelines are designed to advise States, and public and private sector actors, in all their actions regarding the design, development and ongoing deployment of algorithmic systems. To ensure that the human rights and fundamental freedoms of all individuals, as enshrined in the European Convention on Human Rights ("the Convention") and other relevant treaties, are effectively protected throughout the technological evolution, member States of the Council of Europe shall refrain from violating human rights through the use of algorithmic systems, and shall develop legislative and regulatory frameworks that foster an environment where all actors respect and promote human rights and seek to prevent possible infringements. Independently of State obligations and across jurisdictions, public and private sector actors have the responsibility to respect internationally recognised human rights.

2.   For the purposes of this recommendation, "algorithmic systems" are understood as applications that, often using mathematical optimisation techniques, perform one or more tasks such as gathering, combining, cleaning, sorting, classifying and inferring data, as well as selection, prioritisation, the making of recommendations and decision making. Relying on one or more algorithms to fulfill their requirements in the settings in which they are applied, algorithmic systems automate activities in a way that allows the creation of adaptive services at scale and in real time.

3.   Operating typically by detecting patterns in large datasets, algorithmic systems offer the potential to improve the performance of services (particularly through increased precision, targeting and consistency), provide new solutions, and deliver returns in efficiency and effectiveness of task and system performance. They have led to immense improvements in the categorisation and searchability of digital information and have facilitated important advances in fields such as medical diagnostics, transportation and logistics, enabling the broader and faster sharing of information globally and making possible new forms of co-operation and co-ordination. As a result, they permeate many aspects of contemporary human life.

4.   However, there are also significant human rights challenges attached to the increasing reliance on algorithmic systems in everyday life, such as regarding the right to a fair trial; the right to privacy and data protection; the right to freedom of thought, conscience and religion; the right to freedom of expression; the right to freedom of assembly; the right to equal treatment; and economic and social rights. The functionality of algorithmic systems is frequently based on the systematic aggregation and analysis of data collected through the digital tracking at scale of online and offline identity and behaviour of individuals and groups. In addition to the intrusion on individuals' privacy and the increasing potential of highly personalised manipulation, tracking at scale can have a serious adverse effect on the exercise of human rights, which must be considered throughout the entire life cycle of an algorithmic system, from the proposal stage onward.

5.   While it is often argued that the costs are offset by gains in rationalisation and accuracy, it is important to note that most algorithmic systems are based on statistical models in which errors form an inevitable part, sometimes with feedback loops that maintain, replicate and reinforce pre-existing biases, errors and assumptions. Although it may seem as if larger datasets provide better chances of finding recurrent patterns and correlations, accuracy rates do not automatically increase with the size of the dataset. As a result of the large number of people affected by algorithmic systems, the number of errors in the form of false positives and false negatives, and of people who are affected by these errors and inbuilt bias, will also expand, triggering additional interferences with the exercise of human rights in multiple ways.

6.   Algorithmic systems do not process and generate outputs only based on personal data. They can also operate based on non-observational and non-personal data such as simulations, synthetic data, or generalised rules or procedures. However, human rights may still be negatively affected by their use. Individuals and groups whose data is not processed or who have not otherwise been taken into consideration may be directly concerned and significantly impacted, particularly when algorithmic systems are used to inform decision making, adapt recommendations or shape physical environments.

7.   Many algorithmic systems use optimisation techniques where development and implementation stages are closely intertwined. Each use of the algorithmic system can prompt adjustments in its functioning with a view to perfecting results that are based on a narrow range of pre-defined outcomes. Such processes can shape and disrupt environments, particularly when operating at scale. They prioritise certain values over others, for instance general gains over specific losses. This typically happens in ways that are not explicit, transparent, accountable or controllable by the affected individual, and which may generate adverse effects, particularly for minorities and marginalised or disadvantaged groups.

8.   Given the wide range of types and applications of algorithmic systems in everyday life, the level of their impact – positive and negative – on human rights will always depend on the specific purpose for which they are used, their functionality, accuracy, complexity, their effects and the scale on which they are deployed. Their impact will also depend on the broader organisational, thematic, societal and legal context in which they are used, each of which is associated with specific public and ethical values. Applications are very diverse, such as for e-mail spam filters, for health-related data analytics, or for rationalising traffic flows. Algorithmic systems are also employed for predictive purposes in the context of policing and border control, for the purposes of combatting money laundering and fraud, or in labour, employment and educational settings, including as part of public and private recruitment and selection processes.

9.   When assessing a potentially negative human rights impact stemming from the design, development and ongoing deployment of an algorithmic system, it is necessary to evaluate continuously and to document the context, legal ground, purpose, accuracy, side effects and scale of the system's use. The inherent risks of these systems being attacked or confused via adversarial machine-learning or by other means (including cyber-attacks) due to the scale, nature and possible value of the data that are processed should also be considered. The evaluation of the extent of the possible human rights impact of an algorithmic system should take account of the severity, scale and likelihood of giving rise to a human rights violation.

10.   When algorithmic systems have the potential to create an adverse human rights impact for an individual, for a particular group or for the population at large, including effects on democratic processes or the rule of law, these impacts engage State obligations and private sector responsibilities with regard to human rights.

11.   In some cases, the application of an algorithmic system can prompt a particular, higher risk to human rights, for instance because it is used by States for their public service or public policy delivery and the individual does not have a possibility to opt out or suffers negative consequences as a result of the decision to opt out. A similarly heightened risk ensues as a result of use in the context of decision-making processes, by either public authorities or private parties, in situations that carry particular weight or legal consequence. For example, the use of algorithmic systems in the judicial field for the purpose of legal analysis, prediction or individual risk assessment must be introduced with great care and in conformity with the guarantees of a fair trial as enshrined in Article 6 of the Convention. In this recommendation, the term "high risk" is applied when referring to the use of algorithmic systems in processes or decisions that can produce serious consequences for individuals or in situations where the lack of alternatives prompts a particularly high probability of infringement of human rights, including by introducing or amplifying distributive injustice.

12.   Deserving of particular attention in the assessment of potential negative human rights impacts – and resulting questions of responsibility allocation – is the wide range of uses of algorithmic systems that are neither clearly public nor clearly private. This may be the case when parts of a public service are outsourced to private sector providers who may themselves depend on other service providers, when public entities procure algorithmic systems and services from the private sector, or when a company deploys an algorithmic system in order to achieve public policy objectives defined by States.

13.   Cases where functions traditionally performed by public authorities, such as related to transport or telecommunications, become reliant in full or in part on the provision of algorithmic systems by private parties are also complicated. When such systems are then withdrawn for commercial reasons, the result can range from a decrease in quality and/or efficiency to the loss of services that are considered essential by individuals and communities. States should put contingencies in place to ensure that essential services remain available irrespective of their commercial viability, particularly in circumstances where private sector actors dominate the market in ways that place them in positions of influence or even control.

14.   The design, development and ongoing deployment of algorithmic systems involves many actors, including software designers, programmers, data sources, data workers, proprietors, sellers, users or customers, providers of infrastructure, and public and private actors and institutions. In addition, many algorithmic systems, whether learning or non-learning, operate with significant levels of opacity, sometimes even deliberately. Even the designer or operator, who will usually establish the overarching aim and parameters of the system, including the input data, the optimisation target and the model, will often not know what information the system relies upon to make its decision, and is likely to encounter uncertainty about the direct and indirect effects of the system on users and the broader environments in which these systems are intended to operate.

15.   Given this complexity, it is essential that member States be aware of the specific human rights impacts of these processes, and that any investment in such systems contain adequate contingencies for effective monitoring, assessment, review processes and redress for ensuing adverse effects or, where necessary, abandonment of processes that fail to meet human rights standards. Risk-management processes should detect and prevent the detrimental use of algorithmic systems and their negative impacts. Member States should take a precautionary approach and require the refusal of certain systems when their deployment leads to high risks of irreversible damage or when, due to their opacity, human control and oversight become impractical.

B. Obligations of States with respect to the protection and promotion of human rights and fundamental freedoms in the context of algorithmic systems

1. Principles of general application

1.1   **Legislation:** The process of drafting, enacting and evaluating policies and legislation or regulation applicable to the design, development and ongoing deployment of algorithmic systems should be transparent, accountable and inclusive. States should regularly consult with all relevant stakeholders and affected parties, including at sector level where appropriate. States should ensure the enforceability and enforcement of laws, including by demanding that relevant actors produce adequate documentation to verify legal compliance. Where public and private sector actors fail to discharge their legal duties, they should be held responsible.

1.2.   **Ongoing review:** Throughout the entire lifecycle of an algorithmic system, from the proposal stage through to the evaluation of effects, the human rights impacts of individual systems and their interaction with other technologies should be assessed regularly. This is necessary due to the speed and scale at which these systems function and the fast-evolving technological environment in which they operate. This should be done based on broad, effective consultations with those affected or likely to be affected.

1.3   **Democratic participation and awareness:** In order to ensure the full exercise of human rights and democratic freedoms, States should foster general public awareness of the capacity, power and consequential impacts of algorithmic systems, including their potential use to manipulate, exploit, deceive or distribute resources, with a view to enabling all individuals and groups to be aware of their rights and to know how to put them into practice, and how to use digital technologies for their own benefit. In addition, all relevant actors, including those in the public, private and civil society sectors in which algorithmic systems are contemplated or are in use, should promote, encourage and support in a tailored and inclusive manner (taking account of diversity with respect to, for instance, age, gender, race, ethnicity, cultural or socio-economic background) a level of media, digital and information literacy that enables the competent and critical consideration of and use of algorithmic systems.

1.4   **Institutional frameworks:** States should identify and/or develop appropriate institutional and regulatory frameworks and standards that set general or sector-specific benchmarks and safeguards to ensure the compatibility of the design, development and ongoing deployment of algorithmic systems with human rights. Efforts should ensure that direct or indirect risks to human rights, including possible cumulative effects of distinct systems, are promptly identified and that adequate remedial action is initiated. States should invest in relevant expertise to be available in adequately resourced regulatory and supervisory authorities. They should further closely co-operate with independent authorities, equality bodies, national human rights institutions, universities, standard-setting organisations, operators of services, developers of algorithmic systems and relevant non-governmental organisations in various fields, such as, particularly, those engaged in defending human rights.

2. Data management

2.1   **Informational self-determination:** States should ensure that all design, development and ongoing deployment of algorithmic systems provide an avenue for individuals to be informed in advance about the related data processing (including its purposes and possible outcomes) and to control their data, including through interoperability. Deliberate efforts by individuals or groups to make themselves, their physical environment or their activities illegible to automation or other forms of machine reading or manipulation, including through obfuscation, should be recognised as a valid exercise of informational self-determination, subject to possible restrictions necessary in a democratic society and provided for by law.

2.2   **Datasets:** In the design, development, ongoing deployment and procurement of algorithmic systems for or by them, States should carefully assess what human rights and non-discrimination rules may be affected as a result of the quality of data that are being put into and extracted from an algorithmic system, as these often contain bias and may stand in as a proxy for classifiers such as gender, race, religion, political opinion or social origin. The provenance and possible shortcomings of the dataset, the possibility of its inappropriate or decontextualised use, the negative externalities resulting from these shortcomings and inappropriate uses as well as the environments within which the dataset will be or could possibly be used, should also be assessed carefully. Particular attention should be paid to inherent risks, such as the possible identification of individuals using data that were previously processed based on anonymity or pseudonymity, and the generation of new, inferred, potentially sensitive data and forms of categorisation through automated means. Based on these assessments, States should take appropriate action to prevent and effectively minimise adverse effects.

2.3   **Infrastructure:** The increasing centralisation of data and data-processing capacity (including in cloud processing) and the possibility of a lack of choice regarding infrastructure may negatively impact States' ability to discharge their human rights obligations under the Convention. Therefore, States should facilitate the development of alternative, safe and secure infrastructures to ensure that high-quality data-processing and computational capabilities remain available to public and private actors alike.

3. Analysis and modelling

3.1   **Computational experimentation:** States should ensure that computational experimentation likely to trigger significant human rights impacts is conducted only after a human rights impact assessment. The free, specific, informed and unambiguous consent of participating individuals should be sought in advance, with an accessible means of withdrawing consent. Experimentation designed to produce deceptive or exploitative effects should be explicitly prohibited.

3.2.   **Embedding of safeguards:** States should ensure that algorithmic design, development and ongoing deployment processes incorporate safety, privacy, data protection and security safeguards by design, with a view to preventing and mitigating the risk of human rights violations and other adverse effects on individuals and society. Certification schemes based on regional and international standards should be designed and applied to guarantee the provenance and quality of datasets and models. Such safeguards should also form part of procurement processes and should be informed by, and compliant with, regulatory frameworks that ban certain uses of algorithmic systems.

3.3   **Testing:** Regular testing, evaluation, reporting and auditing against state-of-the-art standards related to completeness, relevance, privacy, data protection, other human rights, unjustified discriminatory impacts and security breaches before, during and after production and deployment should form an integral part of testing efforts, particularly where automated systems are tested in live environments and produce real-time effects. State efforts should include public, consultative and independent evaluations of the lawfulness and legitimacy of the goal that the system intends to achieve or optimise, and its possible effects in respect of human rights. Such an evaluation should also form part of procurement processes. Any significant restrictions on human rights that are identified during the testing of such systems should result in immediate rectification and, failing that, in the suspension of the system until such rectifications can take place.

3.4   **Evaluation of datasets and system externalities:** States should ensure that the functioning of the algorithmic systems that they implement is tested and evaluated with due regard to the fact that outputs vary according to the specific context in which they are deployed and the size and nature of the dataset that was used to train the system, including with regard to bias and discriminatory outputs. Depending on the potential impact of the algorithmic system on human rights, testing should, where possible, be performed without using real personal data of individuals, and be guided through a diverse and representative stakeholder process, taking due account of the externalities of the proposed system on populations and their environments, before and after deployment. States should also be aware of the possibility and risks of testing samples or outputs being reused in contexts other than those for which the system was originally developed, including when used for the development of other algorithmic systems. This should not be permitted without new testing and an evaluation of the appropriateness of such uses.

3.5   **Testing on personal data:** States should ensure that the evaluation and testing of algorithmic systems on the personal data of individuals are performed with diverse and sufficiently representative sample populations. Relevant demographic groups should be neither over- nor under-represented. States should also ensure that the staff involved in such activities has sufficiently diverse backgrounds to avoid deliberate or unintentional bias. Furthermore, they should ensure that the development of algorithmic systems is discontinued if testing or deployment involves the externalisation of risks or costs to specific individuals, groups, populations and their environments. Relevant legislative frameworks should disincentivise such externalisation. Special care should be taken in relation to testing in live environments.

3.6   **Alternative and parallel approaches:** As regards the use of algorithmic systems in the delivery of public services and in other high-risk contexts in which States use such technologies, methods such as alternative and parallel modelling should be performed in order to evaluate an algorithmic system and to test its performance and output adequately in comparison to other options.

4. Transparency, accountability and effective remedies

4.1   **Levels of transparency:** States should establish appropriate levels of transparency with regard to the public procurement, use, design and basic processing criteria and methods of algorithmic systems implemented by and for them, or by private sector actors. The legislative frameworks for intellectual property or trade secrets should not preclude such transparency, nor should States or private parties seek to exploit them for this purpose. Transparency levels should be as high as possible and proportionate to the severity of adverse human rights impacts, including ethics labels or seals for algorithmic systems to enable users to navigate between systems. The use of algorithmic systems in decision-making processes that carry high risks to human rights should be subject to particularly high standards as regards the explainability of processes and outputs.

4.2.   **Identifiability of algorithmic decision-making:** States should ensure that all selection processes or decisions taken or aided by algorithmic systems that may significantly impact the exercise of human rights, whether in the public or private sphere, are identifiable and traceable as such at the initial interaction, in a clear and accessible manner.

4.3   **Contestability:** Affected individuals and groups should be afforded effective means to contest relevant determinations and decisions. As a necessary precondition, the existence, process, rationale, reasoning and possible outcome of algorithmic systems at individual and collective levels should be explained and clarified in a timely, impartial, easily-readable and accessible manner to individuals whose rights or legitimate interests may be affected, as well as to relevant public authorities. Contestation should include an opportunity to be heard, a thorough review of the decision and the possibility to obtain a non-automated decision. This right may not be waived, and should be affordable and easily enforceable before, during and after deployment, including through the provision of easily accessible contact points and hotlines.

4.4   **Consultation and adequate oversight:** States should ensure that adequate oversight is maintained by appropriately resourced independent institutions over the number and type of contestations made by affected individuals or groups against certain algorithmic systems that are directly or indirectly implemented by or for them. They should ensure that the results not only lead to remedial action in specific cases but are also fed into the systems themselves to avoid repetition of the offending results, make improvements and possibly discontinue the introduction or ongoing deployment of certain systems due to the likelihood of negative human rights impacts. Information on these contestations and resulting follow-up action should be documented regularly and made publicly available.

4.5   **Effective remedies:** States should ensure equal, accessible, affordable, independent and effective judicial and non-judicial procedures that guarantee an impartial review, in compliance with Articles 6, 13 and 14 of the Convention, of all claims of violations of Convention rights through the use of algorithmic systems, whether stemming from public or private sector actors. Through their legislative frameworks, States should ensure that individuals and groups are provided with access to effective, prompt, transparent and functional and effective remedies with respect to their grievances. Judicial redress should remain available and accessible, when internal and alternative dispute settlement mechanisms prove insufficient or when either of the affected parties opts for judicial review or appeal.

4.6   **Barriers:** States should proactively seek to reduce all legal, practical or other relevant barriers that could lead to directly or indirectly affected individuals and groups being denied an effective remedy for their grievances. This includes the necessity to ensure that adequately trained staff are available to review the case competently and to take appropriate action effectively.

5. Precautionary measures

5.1   **Standards:** States should co-operate with each other and with all relevant stakeholders, including civil society, to develop and implement appropriate guidance (for example, standards, frameworks, indicators, and methods) for state-of-the-art procedures regarding human rights impact assessment. These procedures should be conducted with regard to all algorithmic systems with potentially significant human rights impacts at any stage of the life cycle, with a view to evaluating potential risks and setting out measures, safeguards and mechanisms for preventing or mitigating such risks. Actual harms should be tracked, especially when such systems are applied for non-targeted, explorative purposes. Human rights impact assessments should be made mandatory for all algorithmic systems carrying high risks to these rights.

5.2   **Human rights impact assessments:** States should ensure that they, as well as any private actors engaged to work with them or on their behalf, regularly and consultatively conduct human rights impact assessments prior to public procurement, during development, at regular milestones, and throughout their context-specific deployment in order to identify the risks of rights-adverse outcomes. Confidentiality considerations or trade secrets should not inhibit the implementation of effective human rights impact assessments. Where private sector actors provide services that rely on algorithmic systems and that are considered essential in modern society for the effective enjoyment of human rights, member States should preserve the future viability of alternative solutions and ensure the continued access to such services by affected individuals and groups. For algorithmic systems carrying high risks to human rights, impact assessments should include an evaluation of the possible transformations that these systems may have on existing social, institutional or governance structures, and should contain clear recommendations on how to prevent or mitigate the high risks to human rights.

5.3   **Expertise and oversight:** States should ensure that all human rights impact assessments related to high-risk algorithmic systems are submitted for independent expert review and inspection. Tiered processes should be identified or created where necessary for independent oversight. Human rights impact assessments conducted by or for States should be publicly accessible, have adequate expert input, and be effectively followed up. This may be supported by conducting dynamic testing methods and pre-release trials and by ensuring that potentially affected individuals and groups as well as relevant field experts are consulted and included as actors with real decision-making power, where appropriate, in the design, testing, and review phases.

5.4.   **Follow-up:** In circumstances where the human rights impact assessment identifies significant human rights risks that cannot be mitigated, the algorithmic system should not be implemented or otherwise used by any public authority. If the risk is identified in relation to an algorithmic system that has already been deployed, implementation should be discontinued at least until adequate measures for risk mitigation have been taken. Identified human rights violations should immediately be addressed and remedied, and measures adopted to prevent further violations.

5.5   **Personnel management:** States should ensure that all relevant staff members involved in the procurement, development, implementation, assessment and review of algorithmic systems with significant human rights impacts are adequately trained with respect to applicable human rights and non-discrimination rules and are aware of their duty to ensure not only a thorough technical review but also human rights compliance. Hiring practices should aim for gender parity and diverse workforces to enhance the ability to consider multiple perspectives in the review processes. Such approaches should be documented with a view to promoting them beyond the public sector. States should also work together to share experiences and develop best practices.

5.6   **Interaction of systems:** States should carefully monitor settings where multiple algorithmic systems operate in the same environment in order to identify and prevent negative externalities, particularly where their possible interdependencies and interactions require a precautionary approach. In their public service delivery, States should utilise the mechanism of procurement or engagement of private services with full regard to the need to maintain oversight, know-how, ownership and control over the use of algorithmic systems and their interaction with each other.

5.7   **Public debate:** States should engage in and support ongoing, inclusive, inter-disciplinary, informed and public debates to define what areas of public services affecting the exercise of human rights may not be determined, decided or optimised through algorithmic systems.

6. Research, innovation and public awareness

6.1   **Rights-promoting technology:** States should promote the development of algorithmic systems and technologies that enhance equal access to, and enjoyment of, human rights and fundamental freedoms through the use of tax, procurement, or other incentives. This may include the development of mechanisms to evaluate the impact of algorithmic systems, the development of systems to address the needs of disadvantaged and underrepresented populations, as well as steps to ensure the sustainability of basic services through analogue means, both as a contingency measure and as an effective opportunity for individuals to opt out.

6.2   **Advancement of public benefit:** States should engage in and support independent research aimed at assessing, testing and advancing the potential of algorithmic systems for creating positive human rights effects and advancing public benefit, including by ensuring that the interests of marginalised and vulnerable individuals and groups are adequately taken into account and represented. Where appropriate, this may require the discouragement of influences that may exclusively favour the most commercially viable optimisation processes. States should ensure the adequate protection of whistle-blowing or other actions by employees engaged in the development or ongoing deployment of algorithmic systems who perceive a need to notify regulators and/or the public about present or possible future failures to maintain human rights standards in the systems they have been tasked with building.

6.3   **Human-centric and sustainable innovation:** States should incentivise technological innovation in line with existing human rights, including social rights and internationally recognised labour and employment standards. Efforts to meet the internationally agreed sustainable development goals, notably as regards the extraction and exploitation of natural resources, and to address existing environmental and climate challenges should drive the competitiveness of private sector actors.

6.4   **Independent research:** States should initiate, encourage and publish independent research to monitor the societal and human rights implications of the ongoing deployment of algorithmic systems. In addition, such independent research should study the development of effective accountability mechanisms and solutions to existing responsibility gaps related to the opacity, inexplicability and related incontestability of algorithmic systems. Appropriate mechanisms should be put in place to guarantee the impartiality, global representation and protection of researchers, journalists and academics engaged in such independent research.

C. Responsibilities of private sector actors with respect to human rights and fundamental freedoms in the context of algorithmic systems

1. Principles of general application

1.1   **Responsibility to respect human rights:** Private sector actors engaged in the design, development, sale, deployment, implementation and servicing of algorithmic systems, whether in the public or private sphere, must exercise due diligence in respect of human rights. They have the responsibility to respect the internationally recognised human rights and fundamental freedoms of their customers and of other parties who are affected by their activities. This responsibility exists independently of States' ability or willingness to fulfil their human rights obligations. As part of fulfilling this responsibility, private sector actors should take continuing, proactive and reactive steps to ensure that they do not cause or contribute to human rights abuses and that their actions, including their innovative processes, respect human rights. They should also be mindful of their responsibility towards society and the values of democratic society. Efforts to ensure human rights compliance should be documented.

1.2   **Scale of measures:** The responsibility of private sector actors to respect human rights and to employ adequate measures applies regardless of their size, sector, operational context, ownership structure or nature. The scale and complexity of the means through which they meet their responsibilities may vary, however, according to their resources and the severity of the potential impact that their services and systems have on human rights. Where different sets of private sector actors co-operate and contribute to potential human rights interferences, efforts from all partners are required and should be proportional to their respective impact and abilities.

1.3   **Additional key standards:** Owing to the horizontal effect of human rights and given that design, development and ongoing deployment of algorithmic systems engage private sector actors in very close co-operation with public actors, some of the key provisions that are outlined in Chapter B as obligations of States translate into legal and regulatory requirements at national level and into corporate responsibilities for private sector actors. Irrespective of whether corresponding regulatory action has been taken by States and in addition to the following provisions, private sector actors should uphold the relevant standards contained in paragraphs 1.2, 1.3, 2.1, 3.1, 3.3. and 4.2 of Chapter B above related to ongoing review, democratic participation and awareness, informational self-determination, computational experimentation, testing and identifiability of algorithmic decision making.

1.4   **Discrimination:** Private sector actors that design, develop or implement algorithmic systems should follow a standard framework for human rights due diligence to avoid fostering or entrenching discrimination throughout all life-cycles of their systems. They should seek to ensure that the design, development and ongoing deployment of their algorithmic systems do not have direct or indirect discriminatory effects on individuals or groups that are affected by these systems, including on those who have special needs or disabilities or who may face structural inequalities in their access to human rights.

2. Data management

2.1   **Consent rules:** Private sector actors should ensure that individuals who are affected by their algorithmic systems are informed that they have the choice to give and revoke their consent regarding all uses of their data, including within algorithmic datasets, with both options being equally easily accessible. Users should also be given the possibility to know how their data are being used, what the real and potential impact of the algorithmic system in question is, how to object to the processing of their data, and how to contest and challenge specific outputs. Consent rules for the use of tracking, storage and performance measurement tools of algorithmic systems must be clear, simply phrased and complete, and should not be hidden in the terms of service.

2.2   **Privacy settings:** Private sector actors should facilitate the right of data subjects to protect effectively their privacy while maintaining access to services, in line with all relevant data protection standards. The possibility of choosing from a set of privacy setting options should be presented in an easily visible, neutral and intelligible manner, and facilitate the use of privacy enhancing technologies. Default options should lead only to the collection of data that are necessary for, and proportionate to, the specific legitimate purpose of the data processing, while tracking settings should be set as default in opt-out mode. Any application of mechanisms to block, erase or quarantine user data, for example for security purposes, should be accompanied by due process guarantees and rapid remedies in the event of the erroneous or disproportionate use of data.

3. Analysis and modelling

3.1   **Data and model quality:** Private sector actors should be cognisant of risks relating to the quality, nature and origin of the data they are using for training their algorithmic systems, with a view to ensuring that errors, bias and potential discrimination in datasets and models are adequately responded to within the specific context.

3.2.   **Sample populations:** The evaluation and testing of algorithmic systems on the personal data of individuals should be performed with sufficiently diverse and representative sample populations, and not draw on or discriminate against any particular demographic group. Development of algorithmic systems should be discontinued or adjusted if their development, testing or deployment involves the externalisation of risks or costs to particular individuals, groups, populations and their environments.

3.3.   **Systems and data security:** Private sector actors should configure their algorithmic systems in such a way as to prevent any illegal access, by either their own staff or by third parties, any system interference and any misuse of devices, data and models by third parties, in line with applicable standards.

4. Transparency, accountability and effective remedies

4.1   **Terms of service:** Private sector actors should ensure that the use of algorithmic systems that can trigger significant human rights impacts in the products and services they offer is made known to all affected parties, whether individual or legal entities, as well as to the general public, in clear and simple language and in accessible formats. Adequate information about the nature and functionality of the algorithmic system should be provided to allow for contestation and objection. Terms of service should be reasonably concise, easily understandable and contain clear and succinct language about the possibilities for users to manage settings. They should include information about available options to change the features of the system, applicable complaint mechanisms, the various stages of the procedure, the exact competencies of the contact points, indicative time frames and expected outcomes. All affected parties, new customers or users of products and services whose application rules have been amended should be notified of relevant changes, in a user-friendly format, and requested to consent to the changes where relevant. Failure to consent should not lead to basic services becoming unavailable.

4.2   **Contestability:** In order to facilitate contestability, private sector actors should ensure that human reviewers remain accessible and that direct contact is made possible, including through the provision of easily accessible contact points and hotlines. Individuals and groups should be allowed not only to contest but also to make suggestions for improvements and provide other useful feedback, including with respect to areas where human review is systematically required. All relevant staff involved in the handling of customer complaints should be suitably versed in relevant human rights standards and benefit from regular training opportunities.

4.3.   **Transparency:** Private sector actors should make public information about the number and type of complaints made by affected individuals or groups regarding the products and services they offer, and the outcomes of the complaints, with a view to ensuring that the results not only lead to remedial action in specific cases but are also fed into the systems themselves to draw lessons from complaints and correct errors before harm occurs on a massive scale.

4.4   **Effective remedies:** Private sector actors should ensure that effective remedies and dispute resolution systems, including collective redress mechanisms, are available both online and offline to individuals, groups and legal entities who wish to contest the introduction or ongoing use of a system with potential for human rights violations, or to remedy a violation of rights. The scope of available remedies may not be limited. If prioritisation is necessary and as delays in responding may affect remediability, the most severe human rights impacts should be addressed first. All complaints should allow for an impartial and independent review, should be handled without unwarranted delays and should be conducted in good faith, with respect for due process guarantees. Relevant mechanisms should not negatively impact the opportunities for complainants to seek recourse through independent national, including judicial and regulatory, review mechanisms. No waivers of rights or hindrances to the effective access to remedies should be included in terms of service. Business associations should further invest – in co-operation with trade associations – in the establishment of model complaint mechanisms.

4.5   **Consultation:** Private sector actors should actively engage in participatory processes with consumer associations, human rights advocates and other organisations representing the interests of individuals and affected parties, as well as with data protection and other independent administrative or regulatory authorities, on the design, development, ongoing deployment and evaluation of algorithmic systems, as well as on their complaint mechanisms.

5. Precautionary measures

5.1   **Continuous evaluation:** Private sector actors should develop and document internal processes to ensure that their design, development and ongoing deployment of algorithmic systems are continuously evaluated and tested, not only to detect possible technical errors but also the potential legal, social and ethical impacts that the systems may generate. Where the application of algorithmic systems carries high risks for human rights, including through processes of micro-targeting which they can avoid or mitigate themselves, private sector actors should have the possibility to notify and consult supervisory authorities in all relevant jurisdictions to seek advice and guidance on how to manage these risks, including through the redesign of the services in question. Private sector actors should submit these algorithmic systems for regular independent expert review and oversight.

5.2   **Staff training:** All relevant staff members involved in human rights impact assessments and in the review of algorithmic systems should be adequately trained and made aware of their responsibilities with respect to human rights including, but not limited to, applicable personal data protection and privacy standards.

5.3   **Human rights impact assessments:** Human rights impact assessments should be conducted as openly as possible and with the active engagement of affected individuals and groups. In case of deployment of high-risk algorithmic systems, the results of ongoing human rights impact assessments, identified techniques for risk mitigation, and relevant monitoring and review processes should be made publicly available, without prejudice to secrecy safeguarded by law. When secrecy rules need to be enforced, any confidential information should be provided in a separate appendix to the assessment report. This appendix should be accessible to relevant supervisory authorities.

5.4   **Follow-up:** Private sector actors should ensure appropriate follow-up to their human rights impact assessments by taking adequate action based on the findings recorded throughout the full life cycle of the algorithmic system and monitoring the effectiveness of the responses, with a view to avoiding or mitigating adverse effects on and risks for the exercise of human rights. Identified failures should be resolved as quickly as possible and related activities suspended where appropriate. This requires regular and continued quality assurance checks and real-time auditing throughout design, testing, and deployment stages. It further requires regular consultation with affected individuals to monitor algorithmic systems for human rights impacts in context and *in situ*, and to correct errors and harm appropriately and in a timely manner. This is particularly important given the risk of feedback loops that can exacerbate and entrench adverse human rights impacts.

6. Research, innovation and public awareness

6.1   **Research:** Private sector actors should engage in, fund and publish research, conducted in line with research ethics, aimed at assessing, testing and advancing the potential of algorithmic systems for creating positive human rights impacts and for advancing public benefit. They should also support independent research with this aim and respect the integrity of researchers and research institutions. This may concern the development of mechanisms to evaluate the impact of algorithmic systems, and the development of algorithmic systems to address the needs of disadvantaged and underrepresented populations. Private sector actors should find effective channels of communication with local civil society groups, particularly in geographical areas where human rights concerns are high, in order to identify and respond to possible risks related to the deployment of algorithmic systems.

6.2   **Access to data:** For the purposes of analysing the impacts of algorithmic systems and digitalised services on the exercise of rights, on communication networks, and on democratic systems, private sector actors should extend access to relevant individual data and meta-datasets, including access to data that has been classified for deletion, to appropriate parties, notably independent researchers, the media and civil society organisations. This extension of access should take place with full respect to legally protected interests as well as all applicable privacy and data protection rules.