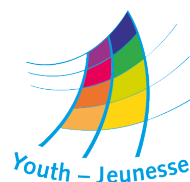




FEMYSO

Debugging Artificial Intelligence –

Policy Recommendations



COUNCIL OF EUROPE



CONSEIL DE L'EUROPE

Introduction.

Young Muslims in Europe are a diverse, hybrid and intersectional group representing a significant portion of the European community. We make up approximately 8.8% and 6.1% of the population in France and Germany respectively, and as a community, we are significantly younger across Europe than all other demographic groups¹, with a median age of 30.4 years. Yet, despite this significant presence, our experiences and concerns are too often underrepresented in policy spaces.

The following policy recommendations are an attempt to begin to include these diverse perspectives and concerns into policy on a vital and emergent issue, that of Artificial Intelligence (AI). As young Europeans, our concerns are shaped by our European contexts, and as young people and Muslims there is also concern for global issues and questions that are seen to impact these global communities. Therefore, some of the concerns implicit in the recommendations relate to issues relevant beyond the boundaries of European countries. These are issues such as climate change, securitisation and criminalisation by AI, migration, and national defence.

The following policy recommendations are the completion of a project that began as the work of many young Muslims from across Europe at the "Debugging Artificial Intelligence : algorithmic justice in minority youth participation" Study Session organised by FEMYSO in collaboration with the Council of Europe, in the European Youth Centre Budapest, Hungary in March 2023. It represents a collective effort by students and early career professionals to engage in the regulation and policy making process of an emergent and evolving technological event, AI. The aim in creating and sharing these recommendations is to begin to diversify voices heard in policy making, with a focus on youth, and racially and religiously minoritised populations across Europe.

Key Themes:

- AI in justice and migration
- AI in the workplace
- AI and the environment
- AI in defence
- AI in culture and media
- AI systems in the public domain

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

¹ *The Global Religious Landscape*, available online at <https://www.pewresearch.org/religion/2012/12/18/global-religious-landscape-muslim/>, 2012

AI and the justice system and migration.

Recognising the historic and ongoing racist and anti-Muslim effects of invasive policing policies and that these communities experience the worst impacts of surveillance technologies and legislation;

Recognising also the current limitations in regulation of data collection, storage, and use in the criminal justice systems and migration systems;

Considering that said limitations in regulation have led to abuses of AI systems in data collection and interpretation for use in the criminal justice system, such as the COMPAS AI programme used in America²;

Further considering that said limitations in regulation have led to abuses of AI systems in data collection and interpretation for use in the migration system, particularly the rights of refugees to seek asylum³, such as the recent controversies involving Frontex⁴;

Stressing the current exemptions from the AIA Article 83 from the scope of the regulation AI systems that are components of large-scale IT systems in the area of migration, security and justice, which is likely to propagate harm against the ethnically and religiously minoritised communities;

Acknowledging the potential of AI systems to bring about public benefit through aiding efficient time management as well as their potential to violate fundamental human rights, particularly the right to privacy and the right to not be discriminated against;

Guided by the European Convention of Human Rights protections on freedom and security, respect for privacy, and a fair trial in civil and criminal matters;

I. Recommends that member states should ensure robustness and clarity in the regulations governing the use of Artificial Intelligence (AI) in the criminal justice system to ensure compliance with and the upholding of the European Convention on Human Rights, and International Human Rights Law. To this effect, member states should:

- A.** Prohibit the default use of AI-based systems in the collection and processing of data used as evidence in criminal justice proceedings;
- B.** Prohibit AI Polygraphs and predictive analytics in criminal justice decision making.
- C.** Allow the use of AI-based systems only in circumstances whereby meaningful human control can be exercised and where accountability structures are well established.

II. Recommends that member states should ensure robustness and clarity in the regulations governing the mass-acquisition, storage, and use of private data in the criminal justice system and migration systems to ensure compliance with the European Convention on Human Rights, and International Human Rights Law. To this effect, member states should:

²The G2 Mesa, N. (2021) *Can the criminal justice system's artificial intelligence ever truly be fair?* Massive Science. Available from: <https://massivesci.com/articles/machine-learning-compass-racism-policing-fairness/>

³Nalbandian, L. (2022) *An eye for an 'I': a critical assessment of artificial intelligence tools in migration and asylum management.* CMS 10(32). <https://doi.org/10.1186/s40878-022-00305-0>

⁴Stavinoha, L., Fotiadis, A., & Zandonimi, G., (2022) *EU'S FRONTEx TRIPPED IN ITS PLAN FOR 'INTRUSIVE' SURVEILLANCE OF MIGRANTS* (online). Available from: <https://balkaninsight.com/2022/07/07/eus-frontex-tripped-in-plan-for-intrusive-surveillance-of-migrants/>

A. Ensure AI for the assessment of evidence is considered 'high-risk'

B. Ensure that AI as part of large-scale EU IT databases for the assessment of evidence is subject to EU data laws and regular, independent oversight.

III. Recommends that member states ban harmful AI practices in the migration context, such as:

A. Prohibit the use of predictive analytics for the purpose of preventing migration, and asylum seeking

B. Prohibit the use of biometric categorisation and identification in publicly accessible spaces.

C. Prohibit the use of AI-driven systems that do not have human oversight and accountability structures accessible to the subjects of the AI system.

IV. Recommends that at an EU-level funding be provided to non-Governmental organisations to monitor the use of AI in 'high-risk' areas including criminal justice and migration systems. Further that the data produced from the monitoring be used to develop mitigation measures and impact assessments of AI on inequalities.

AI in the workplace:

Having examined the current frameworks for ethics approval that many AI providers will currently have to complete or follow before they can be used such as the European High-Level Expert Group on AI's Ethics Guidelines for Trustworthy Artificial Intelligence;

Recognising the potential of AI to widen economic inequality by limiting economic circulation by job displacement and wage suppression as AI reduces the bargaining power of workers, the benefits of AI likely to be captured by tech companies and their shareholders could concentrate wealth and increase inequality, unequal access to AI technology, and regional disparities due to uneven effects of automation are some of the risks already identified by the European Commission⁵;

Recognising that AI-driven technologies in recruitment and professional assessment have the propensity to have led to breaches of privacy, fairness, and anti-discrimination principles, perpetuate existing inequalities and stereotypes in the hiring process, exacerbating anti-Muslim discrimination as well as discrimination based on factors such as gender, race, belief, ability, and age⁶;

Considering the need to prioritise the protection of individual rights, privacy, fairness and anti-discrimination in recruitment and in the workplace whilst engaging with the benefits of AI-driven technologies;

Remembering the European Convention of Human Rights protections on freedom and security, respect for privacy, and anti-discrimination;

This section outlines the following amendments to the European Commission's proposed Artificial Intelligence Act, to ensure fundamental rights protection for people subjected to AI systems in the workplace and in recruitment:

- I. **Recommends** that social and economic inequality be included as factors in Ethical Approval standards for AI providers and AI users which must be passed before the technology can be implemented.
- II. **Recommends** that all relevant stakeholders, including representatives of the companies workers and diverse AI users, be involved in the decision-making for the purchasing and implementation of AI by large companies.
- III. **Recommends** that AI use which replaces human workers should be subject to specific taxation.
- IV. **Recommends** the creation of standardised and mandatory fairness audits for AI-driven systems in the workforce and recruitment at a member state level:

⁵ European Commission (2019) *Liability for Artificial Intelligence and other emerging digital technologies* (online). European Union. <https://doi.org/10.2838/573689>

⁶ European Institute for Gender Equality (2021) *Artificial intelligence, platform work and gender equality* (online). <https://doi.org/10.2839/372863>

- A. An independent body should evaluate the fairness, accuracy, and potential discriminatory impacts of these systems at regular intervals, including their economic and social impacts.
 - B. The standardised audit frameworks should be open and regularly updated and assess the diversity of the data sets used and the models generated by the AI-driven systems in the field of hiring and employee evaluation.
- V. **Recommends** that social and economic inequality be included as factors in Ethical Approval standards for AI providers and AI users which must be passed before the technology can be implemented.
- VI. **Recommends** that all relevant stakeholders, including representatives of the companies workers and diverse AI users, be involved in the decision-making for the purchasing and implementation of AI by large companies.

Environmental concerns of AI.

Recognising the invisible human and environmental impact of technologies through destructive land practices of metal mining and through the electrical demands of the data servers that are required to maintain AI programmes;

Considering by the modern-day slavery that is often used in the mining and assembling processes of the AI hardware⁷;

Guided by the European Convention of Human Rights protections on freedom and security, and prohibition of slavery and forced labour;

This section outlines the following amendments to the European Commission's proposed Artificial Intelligence Act, to ensure fundamental rights protection for people subjected to AI systems and the environmental and global impact of their manufacturing processes:

- I. **Recommends** that all AI providers and AI users must publish and make known their general carbon dioxide emissions to the AI subjects in civilian uses of AI as well as the carbon dioxide emissions caused directly through the running and maintaining of the AI programme.
- II. **Recommends** that the 'lifecycle of AI' include its safe and environmentally conscious construction and decommissioning of the AI hardware, and that this process shall be regulated and monitored by an independent body.

⁷Child Labour Platform (2019) *Child Labour in mining and global supply chains (online)*. International Labour Organization. Available from: https://www.ilo.org/wcmsp5/groups/public/---asia/---ro-bangkok/---ilo-manila/documents/publication/wcms_720743.pdf

National Defence and AI.

Noting with deep concerns the known problems with identifying individuals using AI live facial recognition⁸, especially people of non-white backgrounds, and the dehumanising impact of reducing people to binary data metrics;

Reaffirming the European Convention on Human Rights, in particular the right to life and safety;

This section outlines the following amendments to the European Commission's proposed Artificial Intelligence Act, to ensure fundamental rights protection for people subjected to AI systems in warfare:

- I. **Recommends** the prohibition of AI enabled autonomous target selection and autonomous use of (lethal) kinetic force in all its forms.
- II. **Recommends** that AI-enabled autonomous weapons systems that cannot be used with embodied meaningful human control should be understood to contravene international human rights laws.

⁸ Gentzel, M. *Biased Face Recognition Technology Used by Government: A Problem for Liberal Democracy*. *Philos. Technol.* 34, 1639–1663 (2021). <https://doi.org/10.1007/s13347-021-00478-z>

AI in culture and media:

Having Considered the rapid advancement of AI in social media that has amplified user engagement but also increased the risks of addiction and exposure to inappropriate content, especially for children.

Noting with concern the ability through AI-driven software to target and supply infinite content that is matched to profiles which can expose many to harmful posts and addictive scrolling and that the disabling of these features for minors would reduce the potential for addiction and promote healthier usage habits.

Recognising the European Convention of Human Rights protections on freedom and security, respect for privacy, and freedom of thought and expression;

This section outlines the following amendments to the European Commission's proposed Artificial Intelligence Act, to ensure fundamental rights protection for people subjected to AI systems in media and cultural settings to give AI subjects more control, better protect young people and children, and ensure a safer, more transparent digital future:

- I. **Recommends** social media platforms should offer users the ability to directly modify their interest centres and algorithm preferences. This should include the ability to reset their algorithm and disable content suggestions from unfollowed users.
- II. **Recommends** that social media platforms should provide clear and accessible information on how their algorithms work and users should be able to disable facial recognition and biometric data collection AI services from online platforms.
- III. **Recommends** that features which allow for continuous content flows (such as autoplay, infinite scrolling, content suggestions, reels/shorts, excessive notifications, targeted ads, and integrated games, amongst others) should be possible to disable.
- IV. **Recommends** the creation of special protocols to protect child and adolescent users from harmful content:
 - A. Apply and verify the minimum age required to create an account.
 - B. Automatically configure minors' accounts to private mode.
 - C. Enhance filtering algorithms to avoid inappropriate, violent, or sexual content.
 - D. Introduce a 'child mode' that limits access to age-appropriate features and content.
 - E. Establish a list of addictive practices (those linked to dopamine release) that must be unavailable for accounts held by children. This includes, but is not limited to: autoplay, infinite scrolling, content suggestions, reels/shorts, excessive notifications, reward features such as likes and comments, and integrated games.

AI Systems in the Public Domain.

Recognising the European Convention of Human Rights protections on freedom and security, respect for privacy, freedom of thought and belief and right to not be discriminated against;

Recalling the Dutch childcare benefit scandal as a case study for how lack of transparency and accountability⁹ in decision-making processes involving AI systems raises concerns about discrimination and accountability;

This section outlines the following amendments to the European Commission's proposed Artificial Intelligence Act, to ensure fundamental rights protection for people subjected to AI systems in public spaces and for cases of public benefit to give AI subjects more control, better protections, and ensure a safer, more transparent digital future:

- I. **Recommends** that AI-based systems used in government systems to determine access to state welfare systems including education, healthcare, childcare benefits and unemployment benefits should be open-source and be made publicly accessible.
- II. **Recommends** the creation of an independent regulator responsible for ensuring compliance of AI systems with human rights legislation. In all cases the AI system operator and manufacturer must report to the regulator:
 - A. The intended use of the system
 - B. The intended outcome of the system
 - C. The justification for using an AI system in the specific context
 - D. Details on the system testing and the training models
- III. **Recommends** the prohibition of remote biometric identification systems in publicly accessible spaces.
- IV. **Recommends** that AI-based systems used in the public domain should use open-source code and be made publicly accessible.

⁹ Kuzniacki, B. (2023) 'The Dutch childcare benefit scandal shows that we need explainable AI rules' (online). University of Amsterdam. Available from: <https://www.uva.nl/en/shared-content/faculteiten/en/faculteit-der-rechtsgeleerdheid/news/2023/02/childcare-benefit-scandal-transparency.html>



DEBUGGING
ARTIFICIAL INTELLIGENCE



FEMYSO

Working for a diverse, cohesive,
and vibrant Europe.

Contact us.

@femyso



femyso.org

Rue Archimède 50 1000 Brussels, Belgium