EUROPEAN DIGITAL RIGHTS

Structural Racism, Digital Rights and Technology

European Digital Rights (EDRi) recommendations to inform the European Commission Action Plan on Structural Racism

July 2020

European Digitial Rights (EDRi) is a network of 44 digital rights organisations in Europe working to defend rights and freedoms in the digital age. This briefing outlines core recommendations to advance racial justice and combat racism and related discrimination in the field oftechnology, to inform the upcoming Action Plan on Structural Racism.

I. RACISM, DISCRIMINATION AND TECHNOLOGY: CORE ISSUES

The growing resort to new technologies, including artificial intelligence and other automated decision-making systems in various areas of public life are having a substantial impact on racialised groups in Europe.[1]

AI presents huge potential for exacerbating discrimination in society, at a scale and to a degree of opacity that goes beyond non-automated or 'human' processes. In addition, automated decision making has often been wrongly portrayed as neutral and 'objective', when in fact it embeds and amplifies the underlying structural biases of our societies. This creates a high risk of automation bias and can lead to difficulties for humans to challenge discrimination which is perpetrated by machines or complex systems. In addition to this, however, we see that AI has the potential to pose harms in relation to:

a) discrimination on the basis of grounds not covered in existing discrimination law, such as financial status, such as with examples from targeted advertising and financial credit scoring.
b) collective harms, for example systems which disadvantage certain communities, geographic areas, such as with predictive policing tools.

1 Sarah Chander (2020) "Data Racism: a new frontier" https://www.enar-eu.org/Data-racism-a-new-frontier

c) the deepening existing societal inequalities, such as systems which deploy risk scoring in the criminal justice system, biometric recognition systems deployed disproportionately in lower income or minority areas, or deployments in the field of social welfare.

The below outlines the main ways digital technologies and policy affect racial and ethnic minorities in Europe.

Law Enforcement, Over-Policing, and Surveillance

Increasing evidence demonstrates how new technologies in the field of law enforcement differentiate, target and experiment on communities at the margins.[2] Even where protected identity markers and classes of information are removed from a given dataset, discriminatory outcomes can nonetheless arise.[3] For example, the increased use of both place-based and person-based "predictive policing"[4] technologies to forecast where, and by whom, a narrow type of crimes (petty crimes or crimes derived from situations of ineqaulity or poverty) are likely to be committed repeatedly score racialized communities with a higher likelihood of presumed future criminality. The use of "neutral" factors such as postal code in practice serve as a proxy for race, exacerbating racial biases, affording false objectivity to patterns of racial profiling, and undermining the presumption of innocence in the criminal justice system.[5]

The increased resort to mass surveillance and identification techniques using biometric data (facial recognition, speaker recognition), in combination with the development of highly opaque police databases and matrixes on suspicious individuals also poses severe threats to privacy in particular for racialised groups and communities. The various matrixes dedicated to monitoring and data collection on 'gangs' target young Black, Brown and Roma men and boys, highlighting discriminatory patterns on the base of race and class, with also implications for childrens' rights. Within the EU these technologies remain largely dissimulated and thus uncontrolled, with little cause for redress for the adverse impacts on communities already at risk of surveillance.

Online privacy, content curation and censorship

Marginalised groups face additional risks as a result of the profiling practices of the online advertisement technology industry. Here the business model of tailoring advertising users based on the collection of intimate personal data and sensitive inferences about identities has had serious consequences for the content marginalised groups are likely to see (or not see). These 'filter bubbles' has led to the discriminatory exclusion of women from seeing STEM

2 European Network Against Racism (2019). Data-driven profiling: hardwiring discriminatory policing practices. Available at: https://www.enar-eu.org/IMG/pdf/data-driven-profiling-web-final.pdf

3 Bruno Lepri et al. (2017), "Fair transparent and accountable algorithmic decision-making processes" Philosophy & Technology at page 5.

4 European Network Against Racism (2019). Data-driven profiling: hardwiring discriminatory policing practices. Available at: https://www.enar-eu.org/IMG/pdf/data-driven-profiling-web-final.pdf; In this report, "person-based" predictive policing systems are highlighted as those which purport to forecast , via lists or databases, who pose risks of committing crimes. Examples include the UK Gangs Matrix or the top 400 and 600 (Pro-kid 12) in the Netherlands.Young black and brown men are overrepresented on these matrixes.

5 Aaron Shapiro, "Reform predictive policing," Nature (25 January 2017) . See also David Robinson, and Logan Koepke, "Stuck in a Pattern: Early evidence on "predictive policing" and civil rights" Upturn (August 2016)

jobs online[6], the censoring of Muslim and LGBTQ+ content, to drastically different advertising on recruitment, housing and other results delivered in Google searches by people of colour.[7] Sensitive inferences are generally not protected in data protection law as they are inferred as proxies to protected characteristics, rather than personal data itself.[8]

In addition, greater oversight is needed of the treatment of marginalised groups on social media platforms, who face heightened risks of censorship, content take-downs, and account suspension,[9] and are at the same time more vulnerable to hate speech, online harassment and threats.[10]

Privacy risks and profiling of migrants and racialised groups

Privacy risks are heightened for communities already overpoliced and over-surveilled. For example, there is very little oversight over the access to data that immigration authorities have, leading to potential oversurveillance, and privacy breaches leading to detention and deportation. Oversurveillance also already occurs for racialized groups, undocumented and LGBTQ+ communities, among others.

In Europe, undocumented migrants are generally unable to avail themselves of data protection rights. This vulnerability is heightened due to the development of mass-scale, interoperable repositories of biometric data to facilitate immigration enforcement.[11] In addition, data-sharing agreements between essential government service providers and immigration enforcement hinder the protection and access of undocumented people for fear of deportation. In addition, the proliferation of uses of digital technologies at the border to automate decisions on migration control (such as the Visa Information System, EURODAC, and others) in many cases do not meet requirements of legality, necessity and proportionality, and pose risks of discrimination, inaccurate decision-making, extraction and processing of data without meaningful consent and infringements on the dignity of people on the move.[12]

---

6 Karolina Iwa□ska, "10 Reasons Why Online Advertising is Broken", https://en.panoptykon.org/online-advertising-is-broken

7 Latanya Sweeney (2013) "Discrimination in online ad delivery"

8 Sandra Wachter (2020) "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising" Berkeley Technology Law Journal, Vol. 35, No. 2, 2020, Forthcoming

9 Evan Yoshimoto, 'Supervision or Suppression? How content moderation can uphold racism', Available at: https://www.hertie-school.org/the-governance-post/2020/05/supervision-or-suppression-how-content-moderation-can-uphold-racism/; The Intercept (2020). Invisible censorship. Tiktok told Moderators to Suppress Posts by Ugly People and the Poor to Attract New Users' available: https://theintercept.com/2020/03/16/tiktok-app-moderators-users-discrimination/; The Guardian (2019) ' Instagram murky shadow bans just serve to censor marginalised communities Available at: https://www.theguardian.com/commentisfree/2019/nov/08/instagram-shadow-bans-marginalised-communities-queer-plus-sized-bodies-sexually-suggestive

10 EDRi (2020). 'Platform Regulation Done Right: EDRi position paper on the EU Digital Services Act' https://edri.org/dsa-platform-regulation-done-right/; Amnesty International UK, (2018). 'Troll Patrol' Project. Available at: https://decoders.amnesty.org/projects/troll-patrol ; Amnesty International, (2018), 'Toxic Twitter - The Silencing Effect'. Available at: https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-5/

11 PICUM and Statewatch (2019) "Data Protection, Immigration Enforcement and Fundamental Rights: What the EU's Regulations on Interoperability Mean for People with Irregular Status"

12 StateWatch, Privacy International, Fundaciòn Datos Protegidos and Red en Defensa de los Derechos Digitales (R3D) (2020) Joint Submission to the UN Special Rapporteur on contemporary forms of racism. Available: https://www.statewatch.org/news/2020/june/digital-technologies-and-borders-joint-submission-to-the-un-special-rapporteur-on-contemporary-forms-of-racism/

Uses of AI and other automated systems for migration control disproportionately impact people on the move, including refugees and people living with precarious immigration status. AI is being tested to detect lies for the purposes of immigration applications at European borders, allocate resources at refugee camps through iris scanning, and to (inaccurately) monitor deception in English language tests through voice analysis. Many such experiments violate basic fundamental rights and are based on extraction of data in situatons of significant power imbalances.[13]

Perpetuating inequalities in employment

The use of AI and other algorithmic systems for recruitment poses concerns for historically discriminated groups in employment and is likely to exacerbate existing inequalities experienced by women, racialised groups, those living with disabilities, LGBTQ+ communities, and people with precarious immigration status. Such systems purport to find a good fit for a par-ticular role by screening candidates' applications, based on pre-designed specifications of the ideal candidate. A key concern here is that the "ideal candidate" is often modelled on previous successful employees, likely to reflect and deepen existing privileges, hierarchies and hiring biases.[14] One highly concerning example is the development of technology for hiring which purports to identify whether applicants have a disability, as recently patented by the AI company HireVue.[15]There are also concerns for workers' rights with the growing trend of AI tools for worker surveillance. Such systems have been used in a variety of ways to make automated calculations about worker performance, 'mood assessment', monitoring of task productivity and more.

Profiling in the field of social welfare

AI systems have been deployed in contexts of social welfare resource allocation, eligibility assessment and fraud detection. In a famous case the Dutch government deployed SyRI, a system to detect fraudulent behaviour in benefits creating risk profiles of individuals. In 2019 a Dutch court found that this system violated human rights and privacy law. The court noted that the SyRI program, primarily deployed in poor and migrant neighbourhoods also could lead to discrimination. There are more and more examples of how automated decision-making, profiling and digitalisation more generally are disproportionately affecting poor and working class people. For example, for many years the Polish government has used data-driven systems to profile unemployed people.[16]

13 Petra Molnar, "Technology on the Margins: AI and Migration Management from a Human Rightsw Perspective," Cambridge International Law Journal, 2019, available at https://www.researchgate.net/publication/337780154_Technology_on_the_margins_AI_and_global_migration_management_from_a_human_rights_perspective

14 Institute for the Future of Work 'AI in hiring: Assessing Impacts for Equality' Available at: https://static1.square-space.com/static/5aa269bbd274cb0df1e696c8/t/5ea831fa76be55719d693076/1588081156980/IFOW+-+Assessing+im-pacts+on+equality.pd

15 Loren Larsen, Keith Warnick, Lindsey Zuloaga, and Caleb Rottman, "Detecting Disability and Ensuring Fairness in Automated Scoring of Video Interviews," United States Patent Application Publication, August 20, 2018

16 Panoptykon, "Profiling the Unemployed in Poland: Social and Political Implications of Algorithmic Decision Making", 2015, avilable at https://panoptykon.org/sites/default/files/leadimage-biblioteka/panoptykon_profiling_report_final.pdf

II. RECOMMENDATIONS TO EU INSTITUTIONS

In order to address the increased racial and discriminatory impact in the field of technology, EDRi recommends:

1. For the European Commission to ensure coordination, collaboration and meaningful consultation with racialised communities, anti-racism and digital rights organisations to develop the Action Plan.

2. For the European Commission to implement a review procedure to ensure any new legislation, policy introduced in the field of technology or digital rights does not adversely impact racialised groups.

Following the recommendation of the UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, there must be 'an equality-based approach to human rights governance of emerging digital technologies. This requires moving beyond "colour-blind" or "race neutral" strategies' and instead 'what is required in the context of emerging digital technologies is careful attention to their racialized and ethnic impact.'[17]

In particular, this review procedure (which may form part of ex ante impact assessments) must recognise that discrimination and other harms resulting from the design, development and deployment of technologies are not likely to be addressed with technical adjustments in the design process, but will require holistic, legal, social and policy solutions[18], including the potential for bans for impermissable use (see recommendation 3), but also other policy measures, for example addressing the digital divide for racialised groups.

3. For the European Commission, specifically DG CNECT and JUST, prevent abuses of racialised communities by legallly restricting impermissable uses of artificial intelligence

To address the negative impacts of automated systems at play at the border, mass surveillance technologies and predictive policing systems which increase over-policing of racialised communities, EDRi calls for the European Commission to set clear red-lines for impermissible uses, in particular:

- indiscriminate biometric surveillance and biometric capture and processing in public spaces;
- use of AI to determine access to or delivery of essential public services (such as social security, policing, migration control). Superficial steps to ensure a 'human in the loop' will not suffice to address the harmful consequences of automated decision-making in these fields;

17 UN Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance, 18 June 2020, A/HRC/44/57 'Racial discrimination and emerging digital technologies: a human rights analysis'
18 Ibid.

- uses of AI which purport to identify, analyse and assess emotion, mood, behaviour, and sensitive identity traits (such as race, disability) in the delivery of essential services;
- predictive policing;
- use of AI systems at the border or in testing on marginalised groups, such as undocumented migrants, asylum seekers, refugees, and people on the move;

In addition, EDRi recommends that the upcoming legislative proposal on AI ensures democratic oversight in particular of marginalised groups, recognises collective and community levels of harm posed by AI, and include the strongest possible human rights protection. The full recommendations on artificial intelligence can be found here.

See our position paper:
EDRi (2020) 'Recommendations for a fundamental rights based approach to artificial intelligence regulation' https://edri.org/wp-content/uploads/2020/06/AI_EDRiRecommendations.pdf

4. The European Union and Member States to implement a Ban on biometric mass surveillance in publicly accessible spaces and prevent further proposals that could leas to mass surveillance.

Biometric mass surveillance systems can exacerbate structural inequalities, accelerate unlawful profiling in a context of racialised over-policing, have a chilling effect on people's freedoms of expression and assembly, and put limits on everyone's ability to participate in public and social activities.

In our position paper on this topic, EDRi is therefore calling for an the European Commission to implement, through legislative and non-legislative means and if necessary, infringement proceedings and Court action, an immediate and indefinite ban on biometric processing that leads to mass surveillance in public spaces.

See our position paper here:
EDRi (2020). 'Ban Biometric Mass Surveillance: A set of fundamental rights demands to EU institutions and Member States': https://edri.org/wp-content/uploads/2020/05/Paper-Ban-Biometric-Mass-Surveillance.pdf

5. For the European Commission to ensure adequate legal protection for racialised groups against data-driven profiling

Considering the potential data protection, non-discrimination, broader fundamental rights and collective harms at stake from data-driven profiling, automated-decision making in the field of law enforcement and policing, it is vital that the European Commission ensures adequate legal protection and prevention against harms emenating from data-driven profiling

and policing. In many cases, as stated in recommenation [3] this will require in some cases substantive bans on certain uses of technologies where they undermine fundamental rights. In cases where this threshold does not apply, the European Commission should explore other legal avenues to guarantee remedies in cases of unlawful profiling[19], including exploring the extension of the Racial Equality Directive to policing.

6. Review, evaluate and ensure fundamental rights compliance of EU databases in the fields of police cooperation and migration.

With regards to data extraction and collection, EDRi calls for the evaluation of current EU databases in the fields of police cooperation and migration and their access procedures for law enforcement authorities. The evaluation should assess how antidiscrimination safeguards work in practice and the compliance of existing systems with data protection principles (especially purpose limitation), the necessity and proportionality principles, and the legal requirements of EU judicial cooperation.

7. Review, evaluate and ensure any EU involvement, funding and support for AI and biometric processing in migration control at the border, is consistent with fundamental rights.

Any EU involvement, funding, and support of migration control technologies should be transparent and publicly scrutinized and meaningful mechanisms of oversight and accountability should be implemented. Any biosurveillance technologies introduced as a result of the COVID-19 pandemic should be especially scrutinized for their disproportionate and overbroad reach and differential impact on marginalised groups, limiting freedom of movement among other fundamental rights. Any proposed use of migration related technology, including auto-mated decision-making in immigration and refugee applications should be seen in context of an increasingly xenophobic, racist environment in which policies have exacerbated the exclu-sion, surveillance, and criminalisation of people on the move.

8. Ensure choice, accountability and fundamental rights in the Digital Services Act

EDRi recommends the following measures to improve the functioning of platforms as public space in our democratic societies, to uphold people's rights and freedoms, and to shape the internet as an open, safe and accountable infrastructure for everybody. When drafting and negotiating the Digital Services Act, EU institutions should:

a) Ensure users' choice: mandatory interoperability for gatekeeping platforms with significant network effects would enable the development of a rich and diverse online ecosystem of public spaces where people can freely choose which online community to integrate and to which content moderation policies they want to abide by, in line with their needs and cultural norms;

19 EU Fundamental Rights Agency (2018). 'Preventing unlawful profiling today and in the future' https://fra.europa.eu/en/publication/2018/preventing-unlawful-profiling-today-and-future-guide

b) Provide accountability: when platforms remove content they should follow procedural and transparency rules and users should be able to easily access redress mechanisms regardless of their socio-economic situation or status in the dispute, through the cre-ation of content dispute settlement bodies; illegal racist content inciting violence or discrimination should be referred to competent and properly resourced law enforce-ment authorities for adequate sanctions if they meet the criminal threshold;
c) Address the manipulation business model: European data protection and privacy rules should be enforced (GDPR) and updated (e.g. through the adoption of the ePrivacy Regulation) to regulate the harmful behavioural advertising-based business model that contributes to the prosperity of illegal content online including illegal racist speech. In addition, the new competition tool envisaged by the European Commission should remedy power abuses by digital gatekeepers to put users back in control of their online experiences.

See our position paper here:
EDRi (2020). 'Platform Regulation done Right. EDRi position on the EU Digital Services Act':
https://edri.org/wp-content/uploads/2020/04/DSA_EDRiPositionPaper.pdf

Authors:

European Digital Rights (EDRi) is a network of 44 digital rights organisations from across Europe. We defend rights and freedoms in the digital environment.

Sarah Chander, Senior Policy Adviser, with Petra Molnar, Mozilla Fellow 2019-2020.
For more information, contact sarah.chander@edri.org