UN Special Rapporteur on Freedom of Religion or Belief
Call for Inputs: Report to the UN General Assembly 76th Session on Respecting, Protecting and Fulfilling the Right to Freedom of Thought

30 June 2021

About Access Now

Access Now welcomes this opportunity to provide relevant information to the United Nations (U.N.) Special Rapporteur on Freedom of Religion or Belief (the Special Rapporteur) to inform the Special Rapporteur's report to be presented at the 76th session of the General Assembly on Respecting, Protecting, and Fulfilling the Right to Freedom of Thought (Freedom of Thought Report).[1] Access Now works to defend and extend the digital rights of users at risk around the world through policy, advocacy, technology support, grants, legal interventions and global convenings like RightsCon. As an ECOSOC accredited organisation, Access Now routinely engages with U.N. Special Procedures in support of our mission to extend and defend digital rights of users at risk around the world.[2]

Introduction

Access Now is pleased to provide input on the Special Rapporteur's thematic report by offering a digital rights perspective. In this submission we address (I) the impact of automated decision making on the right to form an opinion, (II) negative impacts of emotion detection technology on the right to freedom of thought (III) encryption is essential to exercise the right to freedom of thought in the digital age, and (IV) recommendations.

I. Impact of automated decision making on the right to form an opinion

The international human rights framework distinguishes between the internal and external dimension of the right to freedom of thought, and the closely related right to freedom of opinion. While the external dimension of these fundamental freedoms can be subject to legitimate restrictions that must be necessary in a democratic society, proportionate, and non-discriminatory, the internal dimension of the freedom of thought and freedom of opinion, so-called *forum internum*, is absolute and non-derogable.[3] Article 19 of the Universal Declaration of Human Rights as well as International Convention on Civil and Political Rights protect these absolute rights from any unjustified restrictions and interferences. The right to form one's opinion is an essential part of freedom of opinion. By the words of the U.N. Special Rapporteur on Freedom of expression and opinion, "any involuntary disclosure of opinions is prohibited and mental autonomy is affirmed."[4]

---

[1] OHCHR, *Call for input: Report on Respecting, Protecting and Fulfilling the RIght to Freedom of Thought,* available online: https://www.ohchr.org/EN/Issues/FreedomReligion/Pages/freedom-of-thought.aspx, 2021.

[2] Access Now, *About Us*, https://www.accessnow.org/, 2021.

[3] Office and the High Commissioner for Human Rights, *CCPR General Comment No. 22: Article 188 (Freedom of Thought, Conscience and Religion)*, available at https://www.refworld.org/docid/453883fb22.html, 1993.

[4] Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Irene Khan, *Disinformation and freedom of opinion and expression*, available at: https://undocs.org/A/HRC/47/25, 2021.

The data-harvesting business models of large online platforms enable the advertising industry to develop data-driven targeting strategies. Through this approach, companies identify and exploit people's behavioural patterns and characteristics. The umbrella term that covers these manipulative techniques is so called "surveillance-based advertising", understood as a blanket term for digital advertising that is targeted to individuals, usually through tracking and profiling based on personal data. The context of where a specific ad is placed can be random, because as it is targeted at individuals, it can follow them around in different contexts.[5] In most cases, surveillance-based advertising is part of a fully-automated process, where each individual ad is chosen and placed in a matter of milliseconds. This means that neither the publisher (e.g. the owner of a website or app) nor the advertiser (e.g. the owner of the brand that is promoted) chooses which ads to show to whom, or where to display them. This is automatically decided by technological systems that are often controlled by third party intermediaries (so-called 'adtech' companies).[6]

Surveillance-based advertisement has significantly contributed to exploitation of people's particular characteristics to increase the persuasiveness of a message and therefore, to unjustifiably interfere with one's absolute freedom to form an opinion. People who are therefore using platforms' services are being manipulated to think or to take decisions they would have otherwise perhaps never made. Surveillance-based advertisement exploits users' vulnerabilities even without directly identifying those vulnerabilities. Through the use of so-called "lookalike audiences," advertisers can duplicate people's groups with certain characteristics in order to reach new individuals that share the same characteristics. Automated tools and the dominance of a few online platforms has enabled greater manipulation as every single individual using their service can be targeted all the time and at any time.

Access Now strongly supports banning practices that adversely impact people's absolute right to freedom of opinion and the freedom of thought, by influencing their thoughts and opinions without their knowledge or consent. This specifically includes targeted behavioural tracking and individual cross-party tracking. Despite online platforms' claims that there is no turning back from surveillance-based advertisement, we need to remember that the internet was not built on a "creepy ad" business model. In fact, quite the opposite. States must not directly or indirectly protect business models of very large online platforms that stand on surveillance based advertisement and that violate international human rights law as described above. Ending abusive models also means opening the door to human rights compliant alternatives, including innovative forms of contextual advertising that relies on minimum personalisation and no individual targeting.[7] This will also enable new players to enter the digital market.

---

[5] Norwegian Consumer Council, *Time to ban surveillance-based advertising: The case against commercial surveillance online*, available at: https://www.forbrukerradet.no/wp-content/uploads/2021/06/20210622-final-report-time-to-ban-surveillance-based-advertising.pdf, 2021.

[6] Norwegian Consumer Council, *Out of control: How consumers are exploited by the online advertising industry*, available at: https://fil.forbrukerradet.no/wp-content/uploads/2020/01/2020-01-14-out-of-control-final-version.pdf, 2020.

[7] Natasha Lomas, *Data from Dutch public broadcaster shows the value of ditching creepy ads*, available at: https://techcrunch.com/2020/07/24/data-from-dutch-public-broadcaster-shows-the-value-of-ditching-creepy-ads/?guccounter=1, 2020.

Finally, surveillance-based advertisement has far reaching impacts on people's personal interactions, choices and participation in democratic debates. Measures intended to increase transparency can help to better understand the scale of the issues, but these are not enough to prevent and mitigate the ongoing human rights abuses. The individual and societal harms created by intrusive targeting and personalisation require a systematic response. From privacy violations to content curation, invasive tracking harms the right to freedom of opinion in tangible ways. It is a positive obligation of States to protect the absolute right to freedom of thought and opinion from activities of large online platforms by creating an adequate regulatory framework establishing and enforcing strong human rights safeguards.

II. Negative impacts of emotion detection technology on the right to freedom of thought

Broadly, the term 'emotion recognition' covers a range of technologies that attempt to infer someone's emotional state from data collected about that person. This can include using Natural Language Processing (NLP) to analyse text to infer sentiment (often called sentiment analysis), but also more obviously invasive techniques such as using images of a person's face, recordings of their voice, or even more fine-grained physiological and behavioral data from wearable devices to make inferences about emotional states.

Many of the 'face-based' emotion recognition applications typically use some form of Paul Eckman's 'basic emotions' theory, which posits a set of 'universal categories' of human emotion and describes how these can be read from facial configurations. However, emotion recognition systems also attempt to infer emotion from other physiological or behavioural data, such as voice or gait, and therefore go beyond detecting something like Eckman's list of basic emotions, to include applications such as "Artificial Intelligence Polygraphs'' that claim to detect deception or systems claiming to detect political or sexual orientation from observable physiological data.[8]

One of the main issues with understanding the impact of emotion recognition systems on the right to freedom of thought is that there are serious doubts about whether current systems, and even future systems, can actually do what they claim. For instance, Lisa Feldman Barret and her co-authors carried out a meta-study to assess the evidence for inferring emotion from facial configurations and concluded that despite"[t]echnology companies [...] investing tremendous resources to figure out how to objectively "read" emotions in people by detecting their presumed facial expressions [...] the science of emotion is ill-equipped to support any of these initiatives."[9] Similar concerns apply to emotion recognition systems that use other physiological or behavioural data, such as voice, although no such comprehensive study has been carried out compared to face-based approaches.

[8] Access Now, *Computers are binary, people are not: how AI systems undermine LGBTQ indentity,* https://www.accessnow.org/how-ai-systems-undermine-lgbtq-identity/ 6 April 2021.

[9] Lisa Feldman Barrett, Ralph Adolphs, Stacy Marsella et al., *Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements*, Journal of Psychological Science in the Public Interest (2019), Vol 20, Issue 1: 1-68, DOI: https://doi.org/10.1177/1529100619832930 at page 48.

A discrepancy therefore exists between the claims coming from those marketing emotion recognition systems, or even researchers developing them, and what the actual systems can do. If emotion recognition systems, at least as they are today, cannot actually detect emotion, or deception and related 'inner states', but simply make flawed or inaccurate inferences about our emotions, the question is whether they can be said to interfere with the right to freedom of thought.

One way to answer this question is to point to the fact that regardless of whether AI systems can actually infer our emotional state, the fact that they claim to do so already has an impact on our freedom of thought. If security forces use emotion recognition to detect potentially aggressive people[10] in crowds or at protests proactively apprehend these people before they have committed any aggressive act, it does not matter whether the inference was flawed or not; the consequences are real.

Moreover, if people believe that such systems are in operation, whether it's in workplaces or public spaces, they will feel pressure to modify their behaviour to be positively evaluated by these systems, especially if rewards or punishments are linked to certain emotional expressions. See, for example, how Canon put cameras in its office that only allow smiling workers to enter certain rooms.[11] If the logic of examples like this are extended, we could have employers demand certain levels of overall "happiness" throughout the workday in order to get bonuses, or workers being punished for being in a "bad mood" as evaluated by these systems.

In such examples, it does not matter whether the system can actually detect one's mood: freedom of thought is infringed regardless. As Susie Alegre points out in her article, *Regulating around freedom in the "forum internum,"* there are three key elements to the rights to freedom of thought and freedom of opinion:

1. the right to keep your thoughts and opinions private;
2. the right not to have your thoughts and opinions manipulated; and
3. the right not to be penalised for your thoughts and opinions.[12]

Emotion recognition systems can undermine our right to keep our thoughts and opinions private by making inferences about our thoughts, feelings and opinions: without our knowledge, without our consent, and in contradiction to our expressed thoughts. The latter occurs in particular with 'deception' detection or AI polygraph systems which, regardless of whether they work, purport to discern our true feelings and intentions, often in a manner that has severe consequences, such as during police interrogation.

---

[10] TNW, *British police to trial facial recognition system that detects your mood*, available online: https://thenextweb.com/news/british-police-to-trial-facial-recognition-system-that-detects-your-mood, 17 August 2020.
[11] The Verge. *Canon put AI cameras in its Chinese offices that only let smiling workers inside*, available online: https://www.theverge.com/2021/6/17/22538160/ai-camera-smile-recognition-office-workers-china-canon, 17 July 2021.
[12] Susie Alegre, Regulating around freedom in the "forum internum," ERA Forum (2021), Vol. 21: 591-604, DOI: https://doi.org/10.1007/s12027-020-00633-7

Emotion recognition systems often have the express, or implicit, intention of manipulating our thoughts by tailoring content to our emotional state.[13] If an emotion recognition system aims to use an individual's emotional state to play music which will calm them down, psych them up, or even make them angry, it explicitly aims to change their thoughts and mood. Such a system could also be used to make political advertising more palatable to them in order to change their feelings about a particular candidate or issues.

Again, whether such systems work or not is not the real issue: they should be regulated — and in this case prohibited — based on what they intend to do, or on the likely consequences of them working as intended. As Susie Alegre points out, many "countries around the world, and the European Union, were quick to ban subliminal advertising on television when it was first raised as a possibility, regardless of whether or not it was effective."[14] The burden of proof of manipulation should not be placed on those affected by a system; if a company sells a product that claims to be able to manipulate our thoughts and opinions, this should be sufficient cause to ban it.

The question of being punished for our thoughts and opinions should be thought anew in the light of emotion recognition. While obvious examples of punishment exist, such as being arrested after being flagged as 'aggressive' by an emotion recognition system at a protest march, there are more subtle forms of consequences that could be conceived as punishment. Take, for instance, targeted advertising that uses emotion recognition. Say an individual is wearing a virtual reality headset, which is making inferences about their emotional state based on data about their eye movement, or even neural signals. In the virtual environment, this data about their emotional state may be used to determine what advertisement, or what content they see. While the stated aim may be to show "more relevant" ads, the determination of what is relevant to them is not an uncontentious decision. If the system infers that they are depressed, or angry, it may not show them ads for jobs (assuming recruiters would not want to target 'negative emotions'), and may even show them content which is likely to worsen their mental state if that means their engagement with the platform will increase. This could certainly be conceived as a form of punishment, in the sense of being denied positive opportunities. On a deeper level, the very act of determining what is "in their interest" or more relevant for them based on inferences about their thoughts and emotions can be conceived as an attack on my autonomy, which robs them of the choice of what content they want to be shown. It may indeed be appropriate to be angry, or sad, in certain situations, for an employer or private company to label such an emotional state as undesirable or negative, and to try to alter that state by technological means, should be seen as an attack on the essence of our autonomy. As Susie Alegre acknowledges, "[p]rotecting the right to freedom of thought is not about directing thoughts in the direction we think

[13] See, for example, this case taken by the Brazilian consumer organisation, IDEC, where such a system was used in a metro line in São Paulo. Access Now intervened, submitting an expert opinion, and the judge ultimately ruled in favour of IDEC: https://www.accessnow.org/sao-paulo-court-bans-facial-recognition-cameras-in-metro/

[14] Susie Alegre, Regulating around freedom in the "forum internum," ERA Forum (2021), Vol. 21: 591-604, DOI: https://doi.org/10.1007/s12027-020-00633-7 at p. 603-604.

is best, it is about recognising and prohibiting the kind of practices and techniques which threaten to undermine the right, no matter who is using them."[15]

Finally, regarding emotion recognition, the idea that AI systems can classify people into discrete emotional categories, which are machine readable, is an attack on inherent human dignity, a cornerstone of human rights. Any such theory of emotions reduces human complexity to machine readable, and predictable, physiological signals, and creates a feedback loop that actually makes human complexity harder to realise in reality. If doors only open when we display cliched outward signs of happiness, and we need to appear jolly in public to see better quality ads, our possibility to think freely, and be fully human, is diminished. Hannah Arendt foresaw this danger in her 1956 work, *The Human Condition*, with the following warning:

> *The trouble with modern theories of behaviorism is not that they are wrong but that they could become true, that they actually are the best possible conceptualization of certain obvious trends in modern society. It is quite conceivable that the modern age—which began with such an unprecedented and promising outburst of human activity—may end in the deadliest, most sterile passivity history has ever known.*[16]

Arendt warns that reductive theories of human behaviour could become self-fulfilling prophecies. This danger reaches a pinnacle with AI emotion recognition, where the most simplistic and reductive conceptualisations of human behaviour are used to build systems of immense scope and consequence.

Access Now therefore joins the growing chorus of civil society organisations, academics, and regulators calling for a complete ban on this dangerous and deeply problematic technology. The intrusion of AI systems into the *forum internum* must be cut off at the root by prohibiting emotion recognition. One such call has come from the European Data Protection Supervisor (EDPS) and European Data Protection Board (EDPB) in a joint opinion from June 2021, which states that "the use of AI to infer emotions of a natural person is highly undesirable and should be prohibited,"[17] although they allow the possibility of exceptions for certain uses in medical contexts. Our partners at Article19 have called for a more thoroughgoing prohibition, demanding a ban on "the conception, design, development, deployment, sale, import and export of emotion recognition technologies, in recognition of their fundamental inconsistency with international human rights standards."[18] We

---

[15] Susie Alegre, Regulating around freedom in the "forum internum," ERA Forum (2021), Vol. 21: 591-604, DOI: https://doi.org/10.1007/s12027-020-00633-7 at p. 594.
[16] Hannah Arendt, *The Human Condition*, (Chicago: The University of Chicago Press) p.322
[17] European Data Protection Board, *EDPRB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)* available online: https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf , 18 June 2021.
[18] Article 19, *Emotion Recognition Technology Report: China's emotion recognition market and its implications for human rights,* available online: https://www.article19.org/emotion-recognition-technology-report/, January 2021.

support this latter, complete prohibition, as the argument for exceptions is undermined by the fact that these technologies cannot be shown to work properly or conform to a high scientific standard. Indeed, the EDPS-EDPB joint opinion confirms this line of reasoning by also calling for a prohibition on "AI systems whose scientific validity is not proven or which are in direct conflict with essential values of the EU." As we have argued here, emotion recognition conforms to both of these criteria, and a complete prohibition is therefore necessary.

III. Encryption is essential to exercise the right to freedom of thought in the digital age

It is widely recognized that people behave differently when there is a possibility that they are being surveilled.[19] Their actions are altered to conform with the observer's viewpoint. Over time, such behavioral changes may well extend to changes in thinking – limiting the freedom to think of ideas and actions that might be perceived unfavorably by those controlling the surveillance lens. In other words, the hesitation to *act* freely could place dangerous limitations on the ability to *think* freely.

Strong encryption is crucial to exercise one's right to freedom of thought in the digital age.[20] In making secure spaces available for online interaction, where access by unauthorized parties is prevented, and individuals remain anonymous, encryption empowers people to think and express themselves without undue impediments. Encryption also serves to protect vulnerable users around the world by promoting confidence that their communications have not been altered or viewed without permission. The former UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, noted in his report to the Human Rights Council's 29th session (2015) that "encryption and anonymity, separately or together, create a zone of privacy to protect opinion and belief. For instance, they enable private communications and can shield an opinion from outside scrutiny, particularly important in hostile political, social, religious and legal environments."[21] Encryption is vital to enabling individuals and communities to exercise their human rights, including their right to privacy, freedom of expression and opinion, which are inextricably linked to the right to freedom of thought. Where privacy and free speech are compromised as a result of encryption being undermined, there is a risk of individuals being deterred from realizing their freedom of thought. This is corrosive to human rights as well as a healthy democracy where dissent and diversity can thrive.

Recommendations

Access Now therefore respectfully urges the Special Rapporteur to consider the following recommendations:

---

[19] Jonathan Shaw, *The Watchers: Assaults on Privacy in America*, Harvard Magazine, available online: https://www.harvardmagazine.com/2017/01/the-watchers, January-February, 2017.

[20] Secure the Internet, *An Open Letter to the Leaders of the World's Governments Signed by Organizations, Companies, and Individuals,* available online: https://securetheinternet.org/

[21] UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Report to the UN Human Rights Council 29th Session: *The Use of Encryption and Anonymity in Digital Communications,* UN Doc. A/HRC/29/32, 22 May 2015, available online: https://www.ohchr.org/en/issues/freedomopinion/pages/callforsubmission.aspx . *See also* submission from Access and PEN America, available online: https://www.ohchr.org/en/issues/freedomopinion/pages/callforsubmission.aspx .

States:

1. Ban human rights abusive methods of surveillance-based advertisement that violate absolute rights to freedom of thought and freedom of opinion. Such a ban should be complemented by stronger enforcement of data protection laws, including the General Data Protection Regulation, competition regulation, and consumer protection legislation. Banning surveillance-based advertising in general will force structural changes to the advertising industry and alleviate a number of significant harms to society at large.
2. Ban the conception, design, development, deployment, sale, import and export of emotion recognition technologies, in recognition of their fundamental non-compliance with international human rights standards and discriminatory impact on historically marginalised and oppressed groups.
3. Encourage the use of strong encryption as a tool to safeguard privacy, anonymity, free expression, and the freedom of thought. States must refrain from adopting any measures, and enacting laws and regulations that, directly or indirectly, break or undermine strong encryption, such as backdoor or exceptional access mechanisms, key escrows and weak encryption standards. Policy proposals that have the effect of placing restrictions or limitations on the use of strong encryption must be subject to public debate and parliamentary scrutiny. Such policies should meet strict standards[22] of legality, necessity and proportionality, and provide for prior judicial authorization and judicial review on a case-by-case basis.

Private Sector:

1. Fully comply with criteria of legally mandated meaningful transparency. In practice and at minimum, private companies should disclose meaningful information about parameters used to determine and target the recipient to whom the advertisement is displayed, including the category and source of personal data uploaded to the online platform, and the legal basis for uploading this personal data pursuant to the legal basis for uploading this personal data as established by data protection regulatory framework. Furthermore, following meaningful information about algorithms to optimise the advertisement, including meaningful explanation of optimization goal, proxy attributes used for its optimization, as well as meaningful explanation of reasons why online platforms optimized and displayed the advertisement to individual users in order to achieve its optimization goal should be made publicly available.
2. Immediately discontinue and refrain from the design, development, and deployment of emotion recognition technologies, as they hold massive potential to negatively affect people's

[22] Privacy and Security Experts, *Necessary & Proportionate: International Principles on the Application of Human Rights to Communications Surveillance,* available online: https://necessaryandproportionate.org/files/en_principles_2014.pdf Necessary and Proportionate, May 2014; *see also* Access, *Universal Implementation Guide For the International Principles on the Application of Human Rights to Communications Surveillance*, available online: https://necessaryandproportionate.org/files/implementation_guide_international_principles_2015.pdf, May 2015.

lives and livelihoods, and are fundamentally and intrinsically incompatible with international human rights standards.

3. Implement effective measures, such as strong encryption, to strengthen secure communication platforms and make them available widely. Companies must uphold human rights responsibilities, prioritize users' rights, and challenge unlawful and overbroad demands by authorities that threaten encryption. They must follow due procedure under laws and international human rights standards in assessing requests from government agencies for access to encrypted data.

International Organizations:

1. Promote the use of strong encryption, particularly in publicly accessible communication technologies which offer online security, privacy and anonymity without discrimination. As secure communication technologies are essential to the promotion and development of human rights, resources must be allocated to ensure that individuals and communities, particularly those at risk, engaged in human rights protection and interacting with international organizations and civil society, have access to encrypted communication channels.

Access Now (https://www.accessnow.org) defends and extends the digital rights of users at risk around the world. By combining direct technical support, comprehensive policy engagement, global advocacy, grassroots grantmaking, legal interventions, and convenings such as RightsCon, we fight for human rights in the digital age.

*For more information, please contact: **un@accessnow.org***