Committee of experts on internet intermediaries (MSI-NET)

COUNCIL OF EUROPE

CONSEIL DE L'EUROPE

6 April 2017 **MSI-NET (2017)04**

MSI-NET 3rd meeting
27-28 March 2017
(Strasbourg, Agora, RoomG05)

Meeting report

1. The meeting was opened by the Chair of the MSI-NET, Prof. Wolfgang Schulz. Jan Kleijssen, Director of Information Society and Action Against Crime then welcomed members and participants, underlining the priority that the Council of Europe attaches to the enabling role of the internet for society, and to the role of internet intermediaries in their various forms. Mr Kleijssen recalled that during the 11th meeting of the CDMSI (Steering Committee on Media and Information Society), which took place on 29 November – 2 December 2016 in Strasbourg, delegates had reviewed the MSI-NET deliverables on the basis of the first drafts submitted by the rapporteurs, had expressed their keen interest in the topic and had given guidance as reflected in the CDMSI meeting report. The need for an appropriate regulatory framework for internet intermediaries that gives due consideration not only to the obligations of states but also to the due diligence duties of intermediaries, including corporate social responsibility standards, had been particularly stressed. In this context Mr Kleijssen noted that the European Court of Human Rights had in its recent Pihl v. Sweden decision (no. 74742/14) appeared to link the limited liability of an online platform for defamatory user-generated content to the small size and non-profit character of the intermediary. Mr Kleijssen further informed the MSI-NET members and participants of the ongoing implementation of the Internet Governance Strategy of the Council of Europe as well as of recent developments in the context of the Council of Europe initiative to create a platform to foster the dialogue between the member states and internet companies to improve the respect for human rights, democracy and the rule of law online. Finally, he underlined the importance for all Council of Europe committees and sub-committees to take the gender dimension into account when formulating policy recommendations, and wished the members and participants a fruitful debate on the highly relevant topics on their agenda.

2. The Chair and Vice-Chair of the MSI-NET, Wolfgang Schulz and Karmen Turk, were unanimously re-elected until 31 December 2017. The agenda Appendix 1) was adopted without any changes. The list of participants appears in Appendix 2. The gender distribution of the 30 participants was 13 women (43%) and 17 men (57%).

Conclusions and decisions

3. With respect to the *draft Committee of Ministers recommendation on internet intermediaries*, the MSI-NET discussed the revised version of the document as presented by the Rapporteur, Matthias Kettemann, in February (Appendix 3). The MSI-NET agreed with the text's new structure that aims to clarify the political and legal context in the Preamble, while consolidating all normative language in the guidelines that differentiate between the negative and positive obligations of member states with respect to the protection and promotion of human rights online, and on the other hand the corporate social responsibility of intermediaries. While the use of a broad and function-based definition of internet intermediaries was welcomed, including the reference to the multifunctionality of many intermediaries, it was also underlined that this recommendation does not address the rights and responsibilities that intermediaries assume when they engage in editorial functions. All members and participants agreed that the limited liability of intermediaries must be reaffirmed through a reassuring regulatory framework, so that fear of liability or sanctions does not lead to preemptive restrictions. While acknowledging the need to act decisively in the face of content that contains hate and incitement to violence, the Committee underlined the need to stress due process safeguards and proportionality considerations in both sections. It was further decided to insert adequate language to remind both states and intermediaries of the importance to support media and information literacy promotion activities. Members also engaged in a discussion regarding the appropriate language to be inserted to promote a gender sensitive approach and to reflect child protection considerations. A number of specific observations, comments and proposals for changes on the draft recommendation where further made and discussed, which will be reflected in the revised draft recommendation.

4. With regard to the *study on the human rights dimensions of algorithms*, the MSI-NET discussed the revised version as presented in February by the rapporteur, Ben Wagner (Appendix 4). The experts supported the revised structure of the study that draws more attention to the concrete human rights that may be affected. They further agreed that the study should also mention possible positive effects for the exercise of human rights, and should include some reflections on thus far not yet fully known human rights impacts of automated data processing techniques. The main characteristics of algorithms that are relevant from a human rights perspective were discussed and agreement found to insert adaptability as one of the notions. It was further agreed to complete the chapter on specific human rights with more concrete examples of problematic practices or side effects. With respect to the desirability of adding recommendations to the study, experts and participants agreed that the aim was not do develop normative provisions but to bring the most important challenges to the attention of the CDMSI and to formulate policy objectives that should be considered in the context of the application of automated data processing techniques and possible regulatory implications. A number of specific observations, comments and proposals for changes on the draft recommendation where further made and discussed, which will be reflected in the revised draft study.

5. The MSI-NET discussed participation in events with a view to ensuring multi-stakeholder input and participation in its work, notably in the context of EuroDIG where the main thrust of the draft recommendation will be presented in a workshop.

Any other business

6. The MSI-NET members agreed to engage, before their next meeting, in consultations with relevant steering and conventional committees as well as other stakeholders on the draft recommendation on internet intermediaries. To this end, the Secretariat was tasked to circulate a revised version of the text, incorporating the discussions and proposals for

change made during their 3rd meeting, during summer 2017, for comments to be obtained prior to the 4th meeting. Agreement on the revised version will be sought beforehand via written procedure. Members of the MSI-NET further agreed to finalise the study on the human rights dimensions of algorithm during their 4th meeting, when in particular the conclusions will be reviewed.

8. The MSI-NET agreed to hold its next meeting in Strasbourg on 18 and 19 September 2017.

9. The Secretariat will prepare a draft meeting report to be sent to the Chair and the Vice-Chair for consideration. Thereafter, the Secretariat will send the draft report to the MSI-NET with a deadline of 5 full working days allowing for comments. In the absence of comments the report will be deemed finalised and will be transmitted to the CDMSI for information. The progress of work of the MSI-NET will be reflected in its draft documents and the reports of its meetings. Therefore, it is considered not necessary to produce abridged reports of meetings.

APPENDIX 1

AGENDA[1]

1. Opening of the meeting

2. Election of Chairperson and Vice-chair person *[Resolution CM/Res(2011)24E]*

3. Adoption of the agenda

4. Information by the Secretariat

5. Discussion on the second draft recommendation by the Committee of Ministers on internet intermediaries ***(doc MSI-NET(2016)05 rev)***

6. Discussion on the revised draft study on human rights dimensions of algorithms ***(doc MSI-NET(2016)06 rev)***

7. Date of next meeting

8. Other business

MSI-NET Terms of Reference

[1] As it appears in document MSI-NET(2017)01

APPENDIX 2

LIST OF PARTICIPANTS

COMMITTEE MEMBERS

Mr Bertrand de la CHAPELLE – Co-founder and Director of the Internet & Jurisdiction, France

Ms Julia HÖRNLE – Professor of Internet Law, Queen Mary University of London

Ms Tanja KERŠEVAN-SMOKVINA – Principal Advisor to Director General, Agency for Communication Networks and Services, Slovenia

Mr Matthias KETTEMANN – Postdoc Fellow, Cluster of Excellence "Normative Orders" University of Frankfurt/Main (Germany) Austria (Rapporteur Recommendation)

Ms Sabine MAASS – Head of Division 'Legal framework for digital services, media industry', Federal Ministry for Economic Affairs and Energy – Germany (apologised)

Mr Arseny NEDYAK – Deputy Director, Department of Media State Policy, Ministry of Telecommunication, Russian Federation

Mr Pēteris PODVINSKIS – Ministry of Foreign Affairs, International Organisations Directorate, Department for Public Policy related to Internet – Latvia

Mr Thomas SCHNEIDER – Deputy Director of International Affairs, International Information Society Coordinator, Federal Department of the Environment, Transport, Energy and Communication DETEC, Federal Office of Communications (OFCOM), Switzerland

Mr Wolfgang SCHULZ – Professor, Faculty of Law, University of Hamburg / Hans-Bredow-Institut (Chair)

Ms Sophie STALLA-BOURDILLON – Associate Professor in Information Technology / Intellectual Property Law, Director of ILAWS, Southampton Law School University of Southampton

Ms Karmen TURK – Trinity Tallinn – Estonia (Vice-Chair)

Mr Dirk VOORHOOF – Lecturer European Media Law, UCPH (Copenhagen University) / Professor at Ghent University / member of the CMPF Scientific Committee Centre for Media Pluralism and Press Freedom

Mr Benjamin WAGNER – Researcher, German Institute for International and Security Affairs (SWP) (Stiftung Wissenschaft und Politik) / Rapporteur Study HR dimensions on Algorithms

COUNCIL OF EUROPE MEMBER STATES

AUSTRIA - Mr Gerhard HOLLEY, Federal Chancellery, constitutional office

AZERBAIJAN - Mr Bakhtiyar MAMMADOV, Chief advisor, Ministry of Communications and High Technologies of the Republic of Azerbaijan *(Apologised)*

GERMANY - Ms Fabienne FUCHSLOCHER, Legal framework for digital services, media industry - Federal Ministry for Economic Affairs and Energy Germany

ITALY - Ms Francesca PELLICANO, Autorità per le Garanzie nelle Comunicazioni, Roma / Napoli

TURKEY - Mr İrfan Dündar ERENTÜRK, Media Specialist, Radio and Television Supreme Council (RTÜK) Ankara

OBSERVERS

EUROPEAN UNION - AGENCY FOR FUNDAMENTAL RIGHTS (FRA) *Apologized*

EUROPEAN COMMISSION - DG CONNECT - Ms Irene ROCHE LAGUNA, Legal officer, DG for Communications Networks, Content & Technology

EUROPEAN AUDIOVISUAL OBSERVATORY - Ms Maja CAPPELLO, Head of Legal Information Department (apologised)

EBU / EUROPEAN BROADCASTING UNION - Mr Giacomo MAZZONE, Head of Institutional Relations, Public Affairs & Communications - Mr Michael WAGNER, Head of Media Law and Communications, Legal Department

OSCE / Office of the Representative on Freedom of the Media: Mr Frane MAROEVIC, Director (apologized)

UNESCO - Ms Xianhong HU, Communication and Information Sector

OBSERVER STATES TO THE COUNCIL OF EUROPE

MEXICO - Ms Lorena ALVARADO QUEZADA, Deputy to the Permanent Observer of Mexico to the Council of Europe (27.03.2017)

REPRESENTATIVES OF CIVIL SOCIETY, ACADEMIC COMMUNITIES AND THE PRIVATE SECTOR

Ms Christina ANGELOPOULOS, Centre for Intellectual Property and Information Law (CIPIL), University of Cambridge (United Kingdom)

Mr Giancarlo FROSIO - Centre for International Intellectual Property Studies (CEIPI) - University of Strasbourg

Ms Catherine KENT - Essex University (apologized)

Ms Aleksandra KUCZERAWY, Legal Researcher, Centre for IT & IP Law – iMinds, Univeristy Leuven, Belgium

Mr Tarlach McGONAGLE - Senior Researcher and Lecturer, Institute for Information Law (IViR) - University of Amsterdam (28.03.2017)

Mr Joe McNAMEE, Executive director, European Digital Rights (EDRi), Brussels, Belgium

NON-MEMBER STATES

MOROCCO

Ms Chanaz El AKRICHI, Head of Cooperation division, Ministry of Communication

Ms Meriem KHATOURI, Director for Media Studies and Development, Ministry of Communication

Mr Jamal Eddine NAJI, Director General, The High Authority for Audio-visual Communication (HACA), RABAT, MAROC

Mr El Mahdi AROUSSI IDRISSI, Director Legal Affairs, The High Authority for Audio-visual Communication (HACA) RABAT, MAROC

SECRETARIAT

Mr Jan KLEIJSSEN, Director, Directorate of Information Society and Action against Crime

Mr Patrick PENNINCKX, Head of Information Society Department

Ms Silvia GRUNDMANN, Head of Media and Internet Division, Information Society Department

Ms Elvana THAÇI, Head of Standard Setting Unit, Media and Internet Division, Information Society Department

Ms Charlotte ALTENHÖNER-DION, Secretary of MSI-NET Committee, Media and Internet Division, Information Society Department

Ms Małgorzata PEĶ, Project Officer, Media and Internet Division, Information Society Department

Ms Elisabeth MAETZ, Assistant, Media and Internet Division, Information Society Department

INTERPRETERS / INTERPRETES

Mr Grégoire DEVICTOR, Mr Luke TILDEN, Mr Nicolas GUITTONNEAU

APPENDIX 3

REVISED VERSION[2] OF THE DRAFT COMMITTEE OF MINISTERS

RECOMMENDATION ON INTERNET INTERMEDIARIES

submitted at the 3rd meeting (27-28 March 2017)

Rapporteur: Matthias C. Kettemann

1. In line with the jurisprudence of the European Court of Human Rights (hereinafter "the Court"), the Council of Europe member States have the obligation to secure to everyone within their jurisdiction the rights and freedoms contained in the Convention for the Protection of Human Rights and Fundamental Freedoms (ETS No. 5, hereinafter "the Convention") both offline and online.

2. Access to the Internet is a precondition for exercising Convention rights online. By enhancing the public's access to information and services and facilitating the dissemination of content, the Internet plays a particularly important role with respect to the freedom of expression, which includes the freedom to receive and impart information and ideas without direct or indirect interference by public authorities and regardless of frontiers.

3. A wide, diverse and rapidly evolving range of actors facilitates interactions between natural and legal persons on the Internet by performing a number of functions. Some connect users to the Internet, enable the processing of information and data, and host and store web-based services. Others aggregate information and enable searches, and give access to, host and index content and services designed and/or operated by third parties. Others facilitate the sale of goods and services and enable other commercial transactions, including payments. Often, they carry on several functions in parallel. The multi-functionality of these actors, commonly referred to as "Internet intermediaries", should be met with a nuanced approach that differentiates between mere hosting or transmitting services and more active, editorial-like functions that may be performed with regard to third-party content.

4. Internet intermediaries fulfill an essential role in the Internet ecosystem as gateways to information and key enablers of the exercise of rights and freedoms online, in particular the right to privacy, including personal data protection, the freedom of assembly and association, the freedom of expression, the prohibition of discrimination, the right to

[2] As contained in document MSI-NET(2016)05rev, dated 20 February 2017.

education, access to knowledge and culture, as well as the participation in public and political debate and in democratic governance.

5. Internet intermediaries may also interfere with the exercise of human rights. Their terms of service and community guidelines often envisage content restrictions based on broad definitions that may lead to unpredictable implementation and contain clauses that facilitate the collection, retention and processing of information from and about users, often without proper notification. Legal remedies may be lacking or provided only through automated processes. Access to justice may further be made difficult through unfavorable jurisdictional clauses. Moreover, intermediaries often moderate and rank third-party content through algorithms, and thereby influence users' access to information online, similar to traditional media.

6. In fulfilling their central role of securing to everyone in their jurisdiction the rights and freedoms protected in the Convention and of guaranteeing public safety and national security, member States should take into account specific features of the Internet, including the end-to-end architecture and global nature of Internet networks and services, the ownership by the private sector, the anonymity of users, the volume of Internet content, and the speed at which it is produced and processed.

7. The regulatory framework and online environment in which Internet intermediaries act is diverse, multi-layered and continuously evolving. As they operate across many countries, they have to comply with conflicting laws of several jurisdictions. In line with Convention rights and the principle of the rule of law, public authorities may request Internet intermediaries to divulge personal data or remove or restrict certain content. The role of the judiciary in relation to such requests ranges in different jurisdictions from prior authorisation to post-implementation review to ensure that the restriction of content or the disclosure of personal data is prescribed by law, proportionate to the legitimate aim pursued, and necessary in a democratic society.

8. The existing legal frameworks that provide for exemptions from liability of intermediaries for third party-content are, however, increasingly being undermined by extra-legal content removal mechanisms and informal co-operation agreements between intermediaries and public authorities. Such agreements may lead to rights violations as they may prompt intermediaries to proactively monitor, identify and remove allegedly illegal content rather than acting upon specific requests from public authority based on the rule of law.

9. Informal agreements or mechanisms may also damage user trust and create legal uncertainty. Intermediaries are increasingly required to assess the validity of requests by State authorities and/or non-state actors to remove content on the basis of vague criteria or their internal content-management policies. Intermediaries are thus tasked with the responsibility of weighing competing fundamental rights and freedoms. User choice is

further limited by the fact that, due to various network effects and mergers, the market is dominated by a small number of highly influential intermediary companies.

10. While the digital era brings about new challenges for the protection of human rights and fundamental freedoms, the fundamental principles of human rights and rule of law apply online as offline. Member States have the primary obligation to protect human rights by refraining from any interference, unless such interference is prescribed by law, necessary in a democratic society, and proportionate to the aim pursued. Any State action that impacts Internet intermediaries must be clearly prescribed by law, predictable, and exercised transparently within the limits conferred by law. Member States further have the positive obligation of promoting the exercise and enjoyment of human rights and freedoms, including by protecting individuals from the actions of private parties. In case of rights violations, procedural guarantees must be in place to provide citizens with easy access to appropriate and effective remedies vis-à-vis States and intermediaries. Internet intermediaries, as all business enterprises, have the corporate responsibility to respect human rights in line with the well-established and internationally accepted UN Guiding Principles on Business and Human Rights.

11. Against this background and in order to provide guidance to all relevant actors, the Committee of Ministers, under the terms of Article 15.b of the Statute of the Council of Europe, recommends that member States:

- implement the Guidelines included in this recommendation in particular when developing and implementing legislative frameworks with regard to Internet intermediaries;
- take all necessary measures to ensure that Internet intermediaries fulfill their role and responsibilities to respect human rights in line with the UN Guiding Principles on Business and Human Rights and the Recommendation of the Committee of Ministers to member States on Human Rights and Business;
- engage in a regular dialogue with stakeholders from the private sector, civil society, academia and the technical community, with a view to sharing information and discussing emerging technological developments related to Internet intermediaries that impact the exercise and enjoyment of human rights and related legal and policy issues;
- promote these Guidelines in international and regional forums that deal with the roles and responsibilities of Internet intermediaries.

Guidelines on the promotion and protection of human rights and fundamental freedoms with regard to Internet intermediaries

I – Duties and responsibilities of States

1.1 Legality

1.1.1. Any request, demand or other action by public authorities addressed to Internet intermediaries that interferes with human rights and fundamental freedoms must be based on law. The law must be easily accessible, non-arbitrary and otherwise in accordance with international law.

1.1.2. Laws, regulations and policies applicable to Internet intermediaries, regardless of their objective or scope of application, including commercial and non-commercial activities, shall guarantee effective protection of individuals' human rights and fundamental freedoms vis-à-vis potential infringements by Internet intermediaries, as well as sufficient guarantees against arbitrary application in practice.

1.1.3. States shall not exercise pressure on Internet intermediaries through extra-legal means, if such action is likely to lead to interferences that violate human rights or fundamental freedoms.

1.1.4. States cannot absolve themselves from their obligation to secure human rights and fundamental freedoms online by delegating it or parts of it to Internet intermediaries. States shall refrain from delegating through legislation or other means such authority or tasks to Internet intermediaries that oblige them to introduce procedures for balancing fundamental rights and freedoms.

1.1.5. The process of enacting legislation or other regulations applicable to Internet intermediaries should be transparent, accountable and inclusive, and should respect the multi-stakeholder nature of Internet governance and the various interests involved. To that end, States should regularly consult with all affected parties. Before passing legislation, and in regular intervals thereafter, States should conduct impact assessments with regard to potential negative impacts on human rights.

1.1.6. Taking into account the substantial differences in size and organizational structure of intermediaries, States should ensure that legislation, regulation, and policies related to Internet intermediaries are interpreted, applied and enforced without discrimination on any grounds, including residence, nationality, or gender as well as multiple or intersecting forms of discrimination.

1.1.7. States should ensure that legislation, regulation and policies relating to Internet intermediaries are effectively implementable, do not lead to extraterritorial effects in violation of international law and do not challenge the operation of Internet-based trans-border communication.

1.2. Legal certainty, proportionality, necessity, and transparency

1.2.1. Any legislation applicable to Internet intermediaries and to their relations with States and individual users should be accessible and predictable. All laws should be clear and sufficiently precise to enable intermediaries and individuals to regulate their conduct.

1.2.2. Any legislation should include clear restrictions to discretionary powers granted to public authorities in relation to Internet intermediaries, in particular when exercised by the executive branch and law enforcement. The law must indicate the scope of such discretion to protect against arbitrary application. Abuse of discretionary power should be controlled by judicial or other independent and transparent review.

1.2.3. States should make available in a timely manner comprehensive information on the number, nature and legal basis of requests submitted by State authorities to Internet intermediaries that have implications for the exercise of rights and freedoms. These include content removal requests and requests for disclosure of personally identifiable information. States should not prohibit intermediaries from disclosing anonymised or aggregated information about interferences with the exercise of rights and freedoms online, whether based on court or administrative orders, private complainants' requests, or enforcement of their own content restriction policies.

1.2.4. States should as a general rule exercise their jurisdiction only with respect to Internet intermediaries established within their jurisdiction for the services provided to users in that jurisdiction. States should assert jurisdiction over Internet intermediaries not established within their jurisdiction or content made available by individuals located outside their territory only in limited circumstances, such as when such content is clearly unlawful under international law, in cases of universal jurisdiction, or when there is substantial connection between the content or the content-producer to that State. With a view to avoiding legal uncertainty and

conflicts of laws, States shall commit to cooperating amongst themselves and with all relevant stakeholders in order to develop common jurisdictional principles and cross-border procedures, including through appropriate non-state forums.

1.3. Safeguards for freedom of expression

1.3.1. All laws that may lead to interferences with the freedom of expression, including when applied by intermediaries, must respect the established jurisprudence of the Court with regard to freedom of expression, specifically on the Internet. In particular must the legal framework be precise and provide specific rules for the scope of and procedures for monitoring, removing and restricting content as well as for effective judicial review of all such actions.

1.3.2. Any request by State authorities addressed to Internet intermediaries to restrict access to or remove content must be based on law and pursue one of the legitimate aims foreseen in Article 10.2 of the Convention. Any such restriction must be necessary in a democratic society for the pursuit of a legitimate public good and proportionate to the aim pursued. Any legal terms used to designate content to be restricted must be clearly described by law. State authorities must carefully evaluate any restrictions before applying them and seek to apply the least restrictive measure. In doing so, States should recognise that in a democratic society not only information and ideas that are favorably received or regarded as inoffensive are protected, but also those that offend, shock or disturb, including political dissent and protest.

1.3.4. State authorities should not, through legal or extra-legal means, compel or incentivise Internet intermediaries to determine the lawfulness of third-party content or to censor lawful communication, including content that offends, shocks, or disturbs. State authorities shall seek to obtain an order by a court or an independent authority to establish the unlawfulness of content before demanding intermediaries to restrict access.

1.3.4. States should ensure in law and practice that intermediaries are not held liable for the content on their platforms. In cases where the functions of Intermediaries consist in storing content from third parties, they may be held liable only if they do not act expeditiously in reaction to standardised notification procedures, and remove illegal content or disable access thereto as soon as they are made aware of its illegality. Takedown procedures should not be designed in a manner that creates incentives to remove or block lawful content, for instance by providing very short timeframes.

1.3.5. The removal of content or restriction of access to content can only be justified by law if there is a pressing social need for the removal of the content or the restriction of access. All content restrictions should allow notice of such restriction to both the content producer/issuer and users seeking access to the content, including information on how to proceed in order to challenge the removal/restriction order.

1.3.6. In cases where intermediaries perform different functions, State authorities should apply an approach that is differentiated and graduated in line with Recommendation CM/Rec(2011)7 of the Committee of Ministers to member States on a new notion of media. They should acknowledge that rights and duties of intermediaries, in particular liability for third-party content, depends on the role and position an intermediary takes both de jure and de facto.

1.3.7. While notice-and-takedown is a well-established approach to limiting liability of intermediaries, States may apply a more graduated approach in relation to specific content. Notice-and-(counter) notice procedures may be more sensible for copyright issues, notice-wait-and-takedown approaches for defamation, notice-and-takedown and notice-and-suspension for serious cases of hate speech. Notice-and-judicial-take-down should only serve as complementary solutions. Automatic takedown should only be applied to content prohibited by international law.

1.3.8. State authorities should not directly or indirectly impose an obligation on Intermediaries to systematically monitor the activities of their users in order to prevent unlawful activities or unlawful third-party content, be it by automated means or not. Before addressing any request to Internet intermediaries or promoting, alone or with other States or international organisations, co-regulatory approaches by Internet intermediaries, State authorities shall consider their duty to minimise such monitoring, as well as the limits of automated means of content monitoring that are unable to assess context.

1.4. Safeguards for privacy and data protection

1.4.1. Any demand or request by State authorities addressed to Internet intermediaries to access personal information or other data of their users, or any other measure which interferes with the right to privacy, must be based on law and pursue one of the legitimate aims foreseen in Article 8.2 of the Convention and must be necessary and proportionate to the aim pursued. The protection of the right to privacy and data protection extends to devices used to access the Internet or store data.

1.4.2. State authorities should ensure that Intermediaries' policies and practices uphold the principles of data processing (lawfulness, fairness and transparency, purpose limitation, data minimisation, accuracy, storage time limitations, integrity and confidentiality) and guarantee the rights of the data subject in full compliance with the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No.108).

1.4.3. Surveillance measures undertaken by States, in cooperation with Internet intermediaries or not, must be targeted and comply with Article 8.2 of the Convention. In particular they must be mandated by law and must include sufficient procedural and oversight safeguards. All surveillance must be authorised by a judge or other independent body. States authorities should ensure that intermediaries appropriately confine, in compliance with the principles and purposes of the Convention, data linkage practices across services.

1.5. Access to an effective remedy

1.5.1. States should proactively seek to reduce all legal, practical or other relevant barriers that could lead to a denial of access to an effective remedy for grievances of individual users

1.5.2. States should guarantee easily accessible and effective mechanisms for all individuals to challenge all legal or extra-legal actions that interfere with the right to freedom of expression or the right to privacy, or other Convention rights, in compliance with Articles 6 and 13 of the Convention.

1.5.3. States should guarantee an effective remedy for all violations of human rights and fundamental freedoms by Internet intermediaries in compliance with Article 6 and 13 of the Convention. This includes ensuring that intermediaries ensure prompt and effective review of user grievances and alleged terms of service violations, and provide for effective remedies, including judicial review, when internal and alternative dispute settlement mechanisms prove insufficient or where the individual(s) concerned opt for judicial redress as their preferred option.

II - Responsibilities of Internet intermediaries with regard to human rights and fundamental freedoms

2.1. Respect for human rights and fundamental freedoms

2.1.1. Internet intermediaries shall in all their actions respect the internationally recognised human rights and fundamental freedoms of their users and of third parties who are affected by their activities. The responsibility to respect human

rights exists independently of the States' duty ability or willingness to fulfill their own human rights obligations.

2.1.2. The responsibility of intermediaries to respect human rights applies regardless of their size, sector, operational context, ownership or structure, impact and nature of the intermediary service. Nevertheless, the scale and complexity of the means through which intermediaries meet their responsibility may vary according to these factors and the human rights impact of an intermediary's business model and practices.

2.1.3. Internet intermediaries should engage in regular due diligence assessments regarding human rights and gender equality. These should include an assessment of actual and potential direct and indirect human rights impacts of their actions, both on users and third parties, and an appropriate follow-up to these assessments by acting upon the findings and monitoring and evaluating the effectiveness of identified responses. Intermediaries should conduct these assessments as open as possible and encourage user engagement.

2.1.4. Intermediaries should ensure that their terms of service and any contractual relations with other parties respect their human rights obligations. They shall further ensure that their terms of service agreements and internal policies are applied and enforced consistently and in compliance with applicable due process safeguards, including notification and access to effective remedies, and that their actions do not have discriminatory effects on users or third parties, including actual or potential users who may have special needs. The prohibition of discrimination may require under certain circumstances that intermediaries make special provisions for users or groups of users that face factual inequality in their access to rights in order to correct this inequality and prevent discriminatory effects.

2.2. Accountability and transparency

2.2.1. Internet intermediaries should apply due diligence in all their actions. All interference by intermediaries with free and open data traffic and communications should be based on clear policy and transparent criteria with sufficient procedural guarantees and must be limited to specific legitimate purposes, such as to preserve the integrity and security of the network, in line with the human rights and fundamental freedom guaranteed in the Convention.

2.2.2. Internet intermediaries should ensure that all terms of service agreements and especially policies specifying the rights of users and the content moderation tools, standards and practices for content moderation and disclosure of user data are publicly available in clear, plain language and accessible formats. They should notify

users of all changes in relevant policies as applicable and without delay (and, if possible, well in advance), and in formats that enable individuals to process and understand the changes without unreasonable effort. Continued use of a service should not be made contingent on accepting terms of service that are more restrictive of privacy, data protection or freedom of expression rights.

2.2.3. The process of developing and applying private law terms of service agreements and content restriction policies should be transparent, accountable and inclusive. Intermediaries should seek to engage in negotiations with consumer associations and other organisations representing the interests of users before adopting policies and undertake human rights impact assessments for all of them, and regularly after adoption. Any such assessments should be made public. Internet intermediaries should seek to empower their users to engage in processes of monitoring, evaluating, reviewing and revising, where appropriate, intermediaries' policies and practices to better reflect a commitment to human rights and fundamental freedoms.

2.2.4. Internet intermediaries should clearly and transparently inform their users about the operation of automated data processing techniques in the performance of their functions, including through algorithms that facilitate searches based on user profiles and predicted preferences, or the distribution of algorithmically selected and curated news. They must also inform users clearly about the monetisation of their data and communications, including identification of the parties involved so as to enable individuals to adapt their conduct. Processing of user data should be limited to the purpose consented to and services existing at the time of agreement by users.

2.2.5. Intermediaries should regularly publish transparency reports that provide specific anonymised information about all interference with free and open data traffic and communications and about all requests received for such interference. Such reports should cover requests for disclosure of user data and content removal, whether based on court orders, private complainants' requests, or enforcement of their own content restriction policies.

2.3. Safeguards for freedom of expression

2.3.1. Internet intermediaries shall respect the rights of users to receive and impart information and ideas. Due consideration must be given to the size of the intermediary and the substitutability of the service and forum it provides. They should not on a general basis conduct ex ante monitoring or filtering to detect unlawful content, except regarding content prohibited by international law. All measures taken to restrict access to, remove, or block content on behalf of a State

must be based on an order by a court or an independent authority, and must be effectuated through the least restrictive technical means. All restriction of content should be limited in scope to the precise remit of the order, whose validity must be reviewed periodically. Procedural safeguards must further be in place to inform the user whose content is challenged, including information with respect to access to effective remedies.

2.3.2. Intermediaries should seek to protect the rights to freedom of expression of their users when confronted with government requests for content restriction that are inconsistent with internationally recognised laws and standards. If the content in question is in compliance with the content restriction policies of intermediaries, these should challenge the order in view of its legality, necessity and proportionality in a democratic society.

2.3.3. When restricting access to certain content in line with their content restriction policies, intermediaries should do so in a transparent and non-discriminatory manner, and by the least restrictive technical means. They should further ensure that users are fully aware of the nature of the content restriction, including with regard to the use of automated flagging techniques, are notified and have a possibility to challenge the restriction. If an internal redress process does not lead to a satisfactory conclusion, they should cooperate in any subsequent judicial proceedings. Content should be reinstated without delay if the appeal against the restriction of content is successful or if there is no longer a pressing social need to restrict the access to the content at issue.

2.3.4. Recognising that automated means of content restrictions may be necessary to prevent similar content from reappearing, intermediaries should carefully assess the human rights impacts of automated content management, for example through predictive profiling, and the importance of considering an expression's context.

2.3.5. Where access to content is restricted or denied, or content removed, the intermediary should display a notice that is visible when attempts to access the content are made, that clearly explains what content has been restricted on what legal basis.

2.4. Safeguards for privacy and data protection

2.4.1. Internet intermediaries should limit the collection of personal data from individual users to what is directly necessary in the context of a clearly defined and explicitly communicated purpose. The collection, retention, aggregation or sharing of personal data must be based on a legitimate interest and in almost all cases the informed and unambiguous consent of the individual user with respect to the

specific purpose in line with Convention 108. Convention 108. The aggregation of data through multiple services or devices must be specifically permitted by users who have to be informed about the nature and purpose of any aggregation in order to properly give consent. Users maintain the right to review, modify, and delete personal data and may withdraw their consent at any time, which shall prevent any further processing of that data.

2.4.2. Intermediaries shall respect the rights to privacy of their users when confronted with government demands that compromise these rights in a manner inconsistent with internationally recognised laws and standards.

2.4.3. Intermediaries should not disclose personally identifiable information about a user unless requested to do so by a court or other competent national authority that has determined with sufficient evidence that the disclosure is necessary in a democratic society and proportionate to the legitimate aim pursued.

2.5. Access to an effective remedy

2.5.1. Internet intermediaries shall make available effective complaint mechanisms and dispute resolution systems that provide prompt and direct redress in cases of user grievances and alleged violations of terms of service. While the complaint mechanisms and their procedural implementation may vary with the size, impact and role of the Internet intermediary, they shall be easily accessible, transparent and meet the principles enshrined in Article 13 of the Convention. Intermediary-based complaint mechanisms shall not supplant state-based judicial and non-judicial review mechanisms.

2.5.2. All complaint mechanisms shall comply with due process safeguards and must include the right to be heard in an independent and impartial process that leads to a reasoned decision which is open to appeal.

2.5.3. Intermediaries should ensure that all users and third parties affected by their actions have full and easy access to information about applicable complaints mechanisms, the various stages of the procedure, indicative time frames, and expected outcomes.

2..5.4. Intermediaries should not include in their terms of service waivers of rights or hindrances to the effective access to remedies, such as mandatory jurisdiction outside of a user's country of residence or non-derogable arbitration clauses.

2.5.5. Intermediaries should seek to provide access to alternative review mechanisms that can facilitate the resolution of disputes that may arise between individual users. Intermediaries should not, however, make alternative dispute mechanism obligatory as the only means of dispute resolution.

2.5.6. Intermediaries should regularly analyse the frequency, patterns and causes of complaints received in order to learn lessons for improving their policies, procedures and practices and for preventing future grievances.

2.5.7. Intermediaries should engage in dialogue with consumer associations and other organisations representing the interests of users in order to ensure that their complaint mechanisms are designed, implemented, and evaluated through a participatory process.

APPENDIX 4

REVISED VERSION[3] OF THE

STUDY ON THE HUMAN RIGHTS DIMENSIONS OF ALGORITHMS

submitted at the 3rd meeting (27-28 March 2017)

Rapporteur : Ben Wagner

1. INTRODUCTION

What information can you see on your Facebook feed? Who is a criminal or a terrorist? Will you get health insurance? Are we going to give you a job? Algorithms are increasingly answering questions that human beings used to answer, typically through automated decision-making processes. These algorithms may not take decisions themselves but may only prepare and present decisions to human decision-makers. The way in which this takes place, however, often leads to quasi-automated decision making, blurring the boundary between human and automated decision-making. These algorithms raise considerable challenges not only in each policy area where they are used, but also for society as a whole on how to safeguard fundamental rights and human dignity in the face of rapidly changing technology. The right to free elections, workers' rights, the right to life, freedom of expression, privacy and even the rule of law itself are all impacted. Responding to challenges associated with 'algorithms' used by the public and private sector, in particular by internet intermediaries is currently one of the most hotly debated questions for human rights.

There is an increasing perception that "software is eating the world" (Andreessen 2011), as human beings feel that they have no control over and do not understand the technical systems that surround them. While disconcerting, it is not always negative. It is a by-product of this phase of modern life in which globalised economic and technological developments produce large numbers of software-driven technical artefacts and "coded objects" (Kitchin and Dodge 2011) embed key human rights relevant decision-making capacities. Which split-second choices should a software-driven vehicle make if it knows it is going to crash? Do the algorithms of quasi-monopolistic internet companies have the power to tip elections? What rights do workers have whose entire relationship with their employer is automated? Who will receive health insurance and what information is provided in Facebook newsfeeds? Is racial, ethnic or gender bias more likely in an automated system and how much bias should be considered acceptable?

Historically, private companies in line with the economic, legal and ethical frameworks they deemed appropriate decided on how to develop software. There is no normative framework

[3] As contained in document MSI-NET(2016)06rev, dated 20 February 2017

for the development of systems and processes that lead to algorithmic decision-making or for the implementation thereof. In fact, it is unclear whether a normative framework regarding the use of algorithms or an effective regulation of automated data processing techniques is feasible as many technologies based on algorithms are still in their infancy. Issues arising from the use of algorithms as part of the decision-making process are manifold and complex and include concerns about data quality, privacy and unfair discrimination. At the same time, the debate about algorithms and their possible consequences for individuals, groups and societies is at an early stage. This should not, however, prevent efforts towards understanding what algorithms actually do, which consequences for society flow from them and how possible human rights concerns could be addressed.

This report identifies some human rights concerns raised through the increasing dominance of algorithms. Depending on the types of functions performed by algorithms, their impact on the exercise of human rights will vary. When algorithms infringe human rights, who is responsible? The person who programmed the algorithm, the operator of the algorithm, or the human being who implemented an algorithmically-prepared decision? Is there a difference between such a decision and a human-made decision? What effects does it have on the way in which human rights are accessed, enjoyed and guaranteed in accordance with well-established human rights standards, including rule of law principles and judiciary processes?

Challenges related to the human rights impact of algorithms and automated data processing techniques are bound to grow as related systems are increasing in complexity and interact with each other's outputs in ways that become progressively impenetrable to the human mind. This report does not intend to comprehensively address the topic but rather seeks to map out some of the main current concerns from the Council of Europe's human rights perspective, and to consider possible regulatory options that member states may have to minimise adverse effects. A number of related themes will require more detailed research to more systematically assess their challenges and potential from a human rights point of view, including questions related to big data processing, machine learning, artificial intelligence or the internet of things.

2. THE SCOPE OF THE REPORT

When looking at algorithms and the automated data processing techniques they engage in, it is important to be clear what types of algorithms are being discussed here. This study will build on existing well-established definitions, in particular the work of Tarleton Gillespie (2014), Nicholas Diakopoulos (2015) and Frank Pasquale (2015). It is further important to keep in mind that the term 'algorithm' is applied widely and has a varied set of meanings, depending on whether it is used in the computer science community, among mathematicians and information technologists, or in public, including political, discourse. Mapping out the human rights dimensions of algorithms must also consider the divergence between formal definitions of algorithms and the popular usage of the term. In fact, many of the debates about algorithms focus less on algorithms themselves and more broadly on the role of technology in society (Bucher 2016).

The definition used here starts from Tarleton Gillespie's assumption that "algorithms need not be software: in the broadest sense, they are encoded procedures for transforming input data into a desired output, based on specified calculations. The procedures name both a problem and the steps by which it should be solved." (Gillespie 2014:167) Thus it can be suggested that algorithms are "a series of steps undertaken in order to solve a particular problem or accomplish a defined outcome." (Diakopoulos 2015:400)

This report will not discuss algorithms that automate manufacturing processes or perform other such routine tasks. Rather, it seems reasonable to limit the discussion to algorithms that are digital and are of "public relevance". This report will focus on algorithmic decision-making with implications for human rights. Without being exhaustive or aiming to predict all possible potential iterations of algorithms and their decision-making in the future, the following characteristics of algorithms that engage in automated data processing and (semi-)automated decision making are considered key from a human rights perspective for this report: automation, data analysis, and adaptability.

A. AUTOMATION

Automation is one of the core challenges associated with algorithmic decision-making. The ability of automated computing systems to replace human beings in a growing number of situations is a key characteristic of the practical implementation of algorithms. Automated decision-making algorithms are used across a variety of domains, from simplistic models that help online service providers to carry out operations on behalf of their users (Kim et al., 2014) to more complex profiling algorithms (Hildebrandt, 2008) that filter systems for personalised content. Automated algorithmic decision-making is usually difficult to predict for a human being and its logic will be difficult to explain after the fact.

B. DATA ANALYSIS

Data analysis algorithms are applied to large amounts of data to find patterns of correlation within the dataset without making a statement on causation (Grindrod, 2014). Their use of data mining and pattern recognition without "understanding" causal relationships may lead to errors and raise concerns about data quality. These algorithms replicate the functions previously performed by human beings but involve a quantitatively different decision-making logic to much larger amounts of data input.

C. SOCIAL CONSTRUCTS AROUND ALGORITHMS

While algorithmic decision-making is increasingly adept at mimicking human decision making, important elements (such as discretion) of decision-making processes cannot be automated and often become lost when human decision-making processes are automated (Spiekermann 2015). Without judging their respective "quality", decision-making processes by humans and by algorithms are fundamentally and categorically different, have different consequences and make different mistakes. While society and governments have considerable experience understanding human decision-making and its failures, they are only beginning to understand the flaws of algorithmic decision-making. One key challenge is the frequent perception that algorithms are able to create neutral and independent

predictions about future events.[4] This challenge, however, relates less to algorithms and more to the human perception and interpretation of their implementation and results.

Traditionally, developers have programmed algorithms by hand "to process and transform input data into a desired output, based on specified calculations." (Gillespie, 2014). With technological evolution, however, the output of algorithms is becoming increasingly opaque, in particular when relying on learning capacities that obscure to human beings not only the pattern of decision-making but also the rationale behind it. Even when a human being formally takes a decision, for instance the decision to remove certain content from a social media platform (see below C.), the human being will often 'rubber stamp' an algorithmically prepared decision, having neither time, context or skills to make an adequate decision in the individual case. Thus, while it may seem logical to draw a distinction between fully automated decision-making and semi-automated decision-making, in practice the boundaries between the two are blurred. In neither case will a human being be able to provide a reasoned argument why a certain decision needed to be taken in the specific case. This has repercussions for the right of the concerned individual to seek an effective remedy against a human rights violation (see below E.)

It should be noted that algorithms as discussed here do not exist meaningfully without interaction with human beings. Mathematic or computational constructs may not have adverse human rights impacts but their implementation and application to human interaction may have. It is nonetheless misleading to claim that computing systems are or can be neutral. Technologies – in their application to human interaction - are deeply social constructs (Winner 1980, 1986) with considerable political implications (Denardis 2012). While a decision-making software may be "biased but ambivalent" (McCarthy 2011:90), it has no meaning without a social system around it which provides meaning. It is thus too simple to blame the algorithm or to suggest to no longer resort to computers or computing. Rather, it is the social construct and the specific norms and values embedded in algorithms that need to be questioned, criticised and challenged. Indeed, it is not the algorithms themselves but the decision-making processes around algorithms that must be scrutinised in terms of how they affect human rights.

The question whether the quality of decisions with respect to human rights differs between those taken by human and those taken by or based on algorithmic calculation can only be answered if we know how human decision-making functions. There is evidence that it is special as regards the use of tacit knowledge and tacit norms (Schulz and Dankert 2016). This, to take an example, enables humans to notice exceptional cases where the application of a rule is not appropriate even though the case falls within its scope. The increasing importance of algorithms in decision making calls for a better understanding of the design and characteristics of decision making procedures.

[4] The excitement surrounding Google Flu trends in 2011 which later turned out to be unjustified as their prediction ability was far lower than had been claimed is one example of the ongoing struggle with assertions regarding the accuracy of predictive algorithms (Lazer et al. 2014; Lazer and Kennedy 2015).

3. IMPACTS OF ALGORITHMS ON HUMAN RIGHTS

The principle reservations towards algorithms and automated data processing techniques usually point to their opacity and unpredictability.[5] Beyond these general concerns, specific human rights are particularly affected. These are referenced below with some case studies as to how and why the use of algorithms may lead to rights violations.

A. FAIR TRIAL – ARTICLE 6 ECHR

The trend towards using automated decision-making embedded in algorithms in national security and crime prevention is growing. Following a string of violent attacks in the US and Europe, politicians have called for online social media platforms to use their algorithms to identify terrorists (Rifkind 2014; Toor 2016). Some such platforms are seemingly already using algorithms to identify accounts that generate extremist content, and governments are asking for the results. Apart from the significant impact such application of algorithms has for the freedom of expression (see below C.), it also raises concerns for fair trial standards contained in Article 6 of the ECHR, notably the presumption of innocence, the right to be informed promptly of the cause and nature of an accusation, and the right to defend oneself in person.

In the field of crime prevention, the main policy debates regarding use of algorithms relate to predictive policing. This approach goes beyond the ability of human beings to draw conclusions from past offences to predict possible future patterns of crime. It includes developed automated systems that predict which individuals are likely to become involved in a crime (Perry 2013), or are likely to become repeat offenders and therefore require more severe sentencing.[6]

In addition, considerable concerns exist that the operation of such assessments in the context of crime prevention is likely to create echo chambers within which pre-existing prejudice may be further cemented. Bias or prejudice, related, for example, to racial or ethnic background, may not be recognised as such by the police, when integrated into an automated computer program that is deemed independent and neutral (see also F.). As a result, bias may become standardised and may then less be likely to be questioned as racially motivated than if based on a human decision. While it is unclear how prevalent such decisions created by algorithms are in the criminal justice system generally, the mere potential of their use raises serious concerns with regard to Article 6 of the ECHR and the principle of equality of arms as established by the European Court of Human Rights.[7]

---

[5]See *The great question of the 21st century: Whose black box do you trust?* at https://www.linkedin.com/pulse/great-question-21st-century-whose-black-box-do-you-trust-tim-o-reilly?trk=eml-b2_content_ecosystem_digest-hero-22-null&midToken=AQGexvwxq0Q3iQ&fromEmail=fromEmail&ut=2SrYDZ8lkCS7o1.

[6] See also Article 19, *Algorithms and Automated Decision-Making in the Content of Crime Prevention: A Briefing paper,* 2016.

[7] See, for instance, in *Jespers v. Belgium* (application 8404/78) of 15 October 1980.

B. THE RIGHT TO PRIVACY – ARTICLE 8 ECHR

The longest and most sustained human rights debate on algorithms and automated data processing relates to the right to privacy.[8] Algorithms facilitate the collection, processing and repurposing of vast amounts of data and images, which may have serious repercussions on the enjoyment of the right to private and family life as guaranteed in Article 8 of the ECHR as well as European personal data protection standards.

Algorithms play a role in online tracking and profiling of individuals whose browsing patterns are recorded by "cookies"[9] and similar technologies such as digital fingerprinting, and aggregated with search queries (search engines) and other data (e.g. social media tracking and data collection through apps on mobile devices) (Tene and Polonetsky 2012). One of the main applications of online tracking and profiling is targeted advertising based on the profile of a person's presumed interests.

Efforts are ongoing to modernise the 1981 Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data in line with the technological evolution, and to further define the rights of the data subject with respect to the implications for privacy of contemporary tools for data collection, processing, repurposing and profiling. Article 8 of the draft modernised Convention for the Protection of Individuals with Regard to the Processing of Personal Data establishes the explicit right of every individual not to be subjected to a decision significantly affecting him or her based solely on an automated processing of data without having his or her views taken into consideration; the right to obtain knowledge of the reasoning underlying data processing where the results of such processing are applied to him or her; and to object at any time, on grounds relating to his or her situation, to the processing of personal data concerning him or her unless the controller demonstrates legitimate grounds for the processing which override his or her interests or rights and fundamental freedoms.[10]

Data protection regulatory frameworks at EU level, such as the General Data Protection Regulation of April 2016, which is to enter into force in May 2018, also establish standards for the use of algorithms in data collection, including possibly a "right to explanation" (Goodman and Flaxman 2016) and the right to access to "knowledge of the logic involved in any automatic processing of data concerning him" (EU Directive 95/46/EC).[11]

---

[8] See Sills 1970.

[9] A cookie is a small amount of data generated by a website and saved by the web browser with the purpose to remember information about the user, similar to a preference file created by a software application. While cookies may serve many functions, their most common purpose is to store login information for a specific site. Cookies are also used to store user preferences for a specific site. For example, a search engine may store search settings in a cookie.

[10] See https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806a616c

[11] See https://secure.edps.europa.eu/EDPSWEB/edps/EDPS/Ethics for further details.

Particular concerns arise from the use of data brokers who aggregate the information contained in personal profiles. This information may then be mined through the use of algorithms, which creates a risk of large-scale surveillance ("dataveillance") by private entities and governments alike (Rubinstein, Lee, and Schwartz 2008). The main concern of using data from profiles for different purposes through algorithms is that the data loses its original context. Repurposing of data is likely to affect a person's informational self-determination. Search engines may have a similar effect on the right to privacy and data protection as they also facilitate the aggregation of data about a specific individual and make it easier to find information by reducing the practical obscurity of anonymous data.

Another key aspect related to the usage of algorithms for automated data processing focusses on 'cloud' data storage. This refers to solutions whereby files and other data are no longer stored on local storage but are stored remotely on servers accessible via the Internet. However, by virtue of engaging in non-local storage practices, the data of users may be processed by algorithms while stored remotely in intrusive ways that would not usually be practiced. Such automated data processing can take place in two places: (1) in transit to the remote network storage location and (2) on the remote servers where the data is stored. It may be increasingly difficult for users to ascertain whether they are using local or remote services, as modern operating systems are gradually becoming more deeply enmeshed with 'cloud' remote services. With regard to data in transit, it may therefore be difficult to determine whether it is sufficiently protected through technologies such as strong end-to-end encryption, and whether it is not manipulated in some form.[12]

C. FREEDOM OF EXPRESSION – ARTICLE 10 ECHR

The operation of algorithms also affects the right to freedom of expression. While the positive impact of search algorithms and search engines for the fundamental right to freedom of expression has been repeatedly discussed,[13] their potential for harming the freedom of information and freedom of expression of individuals, groups and whole segments of societies is increasingly being underlined.[14]

Content which is not indexed or ranked highly by an internet search engine is less likely to reach a large audience. A search algorithm might also be biased towards certain types of

[12] For example, Microsoft's cloud service 'SkyDrive' operates an automated process designed to remove certain content (such as nudity). See Clay 2012.

[13] See, for instance, Council of Europe, *Recommendation of the Committee of Ministers to member States on the protection of human rights with regard to search engines*, CM/Rec(2012)3, Adopted by the Committee of Ministers on 4 April 2012 at the 1139th meeting of the Ministers' Deputies, paragraph 1, available at https://wcd.coe.int/ViewDoc.jsp?id=1929429, observing that s*earch engines "enable a worldwide public to seek, receive and impart information and ideas and other content in particular to acquire knowledge, engage in debate and participate in democratic processes."*

[14] See, for instance, the 2016 Report of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, David Kaye, to the 32nd session of the Human Rights Council (A/HRC/32/ 38), pointing out that "search engine algorithms dictate what users see and in what priority, and they may be manipulated to restrict or prioritise content".

content or content providers, thereby risking to affect related values such as media pluralism and diversity.[15] Here the question is how the results that search engines provide should respond to the wishes of its users and to what extend such responses should promote media pluralism and promote diversity.

Social media platforms also deploy algorithmic predictions of user preferences and consequently guide the advertisements individuals might see, how their social media feeds are arranged and the order in which search results appear, thereby substantially compromising the freedom of expression and right to information of users. Given the size of platforms such as Google or Facebook, their centrality for many experience of the internet as a quasi-public sphere (York 2010) and their ability to massively amplify certain voices (Bucher 2012), this is by no means a trivial matter.

According to Article 10 of the ECHR, any measure that blocks access to content through filtering or removal of content must be prescribed by law, pursue one of the legitimate aims foreseen in Article 10.2, and must be necessary in a democratic society. In line with the jurisprudence of the European Court of Human Rights, any restriction of the freedom of expression must correspond to a "pressing social need" and be proportionate to the legitimate aim(s) pursued.

Algorithms are widely used for content filtering and content removal processes (Urban, Karaganis, and Schofield 2016), including on social media platforms, directly impacting on the freedom of expression and raising rule of law concerns (questions of legality, legitimacy and proportionality). Content removal on social media platforms often takes place through semi-automated or automated processes. While large social media platforms like Google or Facebook frequently claim that human beings remove all content (Buni and Chemaly 2016), large parts of the process are automated (Wagner 2016) and based on semi-automated processes. According to a report from the British Intelligence and Security Committee of Parliament, [16] various automated techniques exist for identifying content believed to break the terms of service of the respective provider, be it because of extremist content, child exploitation or illegal acts such as the incitement to violence. These techniques may also be used to disable or automatically suspend user accounts (Rifkind 2014).

In the US, the Obama administration has advocated for the use of automated detection and removal of extremist videos and images.[17] Additionally, there have been proposals to modify search algorithms in order to "hide" websites that would incite and support extremism. The automated filtering mechanism for extremist videos has been adopted by

[15] Submission from Aleksandra Kuczerawy, Brendan van Alsenoy and Jef Ausloos.

[16] See http://isc.independent.gov.uk/committee-reports/special-reports.

[17]See https://www.article19.org/resources.php/resource/38579/en/algorithms-and-automated-decision-making-in-the-context-of-crime-prevention

Facebook and YouTube for videos. However, no information has been released about the process or about the criteria adopted to establish which videos are 'extremist.'[18]

Similar initiatives have been developed in Europe. The Europol Internet Referral Unit had, one year after its launch in July 2015, assessed and processed 11.000 messages containing violent extremist content materials across 31 online platforms in eight languages, leading to the removal of 91.4% of the total content from the platforms.[19] The system has reportedly been automated with the introduction of the Joint Referral Platform announced in April 2016.[20]

Such practices raise considerable human rights concerns related to foreseeability and legality of interferences with the freedom of expression. Notably the data on extremist online content that Europol is processing refers not just to content that is illegal in Council of Europe member states, but also to material that violates the terms of service of an internet intermediary. Moreover, in many situations extremist content or material inciting terrorism is difficult to identify because of the complexity of disentangling factors such as cultural context and humor. According to the European Court of Human Rights, Article 10 also protects shocking, offensive or disturbing content. Algorithmic blocking, filtering or removal of content is likely to have a significant adverse impact on legitimate content. The already highly prevalent dilemma of large amounts of legal content being removed because of the terms of service of internet intermediaries is further exacerbated by the pressure placed on intermediaries to actively filter according to vague notions such as "extremist".

Public concern in Europe and the U.S. has grown following the U.S. elections in 2016 with respect to the creation and dissemination of fake news, including through automated techniques and on social media platforms, thereby possibly having significant influence over democratic decision-making processes (see also below H.). As a result, there have been renewed calls for traditional media responsibility standards to be applied to social media platforms. Some scholars have likened Facebook to be acting as a "news editor [that] has editorial responsibility for its trending topics" (Helberger and Trilling 2016). The question follows, whether social media platforms, through their algorithms that rank and curate

[18] See http://www.reuters.com/article/us-internet-extremism-video-exclusive-idUSKCN0ZB00M

[19] See https://www.europol.europa.eu/newsroom/news/europol-internet-referral-unit-one-year.

[20] See EC communication from the Commission to the European Parliament, the European Council and the Council delivering on the European Agenda on Security to fight against terrorism and pave the way towards an effective and genuine Security Union https://ec.europa.eu/home-affairs/sites/homeaffairs/files/what-we-do/policies/european-agenda-security/legislative-documents/docs/20160420/communication_eas_progress_since_april_2015_en.pdf. See also Article 19, *Algorithms and Automated Decision-Making in the Context of Crime Prevention: A briefing paper,* 2016.

third-party submissions, exert a form of editorial control traditionally performed by media professionals and therefore create specific media responsibilities.[21]

D. FREEDOM OF ASSEMBLY AND ASSOCIATION – ARTICLE 11 ECHR

The internet and in particular social networking services are a vital tool for the exercise and enjoyment of the right to freedom of assembly and association, offering great possibilities for enhancing the potential for participation of individuals in political, social and cultural life.[22] The freedom of individuals to use internet platforms, such as social media, to associate with each other and to establish associations, and to organise themselves for purposes of peaceful assembly, including protest, in line with Article 11 of the ECHR has equally been emphasised.[23]

In line with Article 11, any restriction to the right to freedom of peaceful assembly and to freedom of association must be prescribed by law, pursue a legitimate aim and be necessary in a democratic society. The operation of algorithms on social media platforms that can lead to automatic sorting out of certain individuals or groups from calls for assemblies, for instance, may have a significant impact on the freedom of assembly, as users who rely on social media platforms for their contacts, may without knowledge not be receiving certain communications. The operation of algorithmic filters by public authorities may also prevent peaceful protests from gathering.

E. RIGHT TO AN EFFECTIVE REMEDY – ARTICLE 13 ECHR

Article 13 of the ECHR states that everyone, whose rights are violated shall have an effective remedy before a national authority. States must therefore ensure that individuals have access to judicial or administrative procedures that can impartially decide on their claims concerning violations of human rights online, including effective non-judicial mechanisms, administrative or other means for seeking remedy such as through national human rights institutions. As primary responsible entity for all rights contained in the ECHR, states must take appropriate steps to protect against human rights violations, including by private-sector actors, and must ensure that those affected have access to an effective remedy. They should therefore encourage all private-sector actors to respect human rights throughout their operations, in particular by establishing effective complaint mechanisms to address early and remedy directly grievances of individuals.

An increasing number of companies, especially larger ones, use algorithms and automated data processing techniques for running their complaints procedures. This can have a

[21]See also http://reutersinstitute.politics.ox.ac.uk/news/editors-vs-algorithms-who-do-you-want-choosing-your-news

[22] See Recommendation CM/Rec(2012)4 of the Committee of Ministers to member States on the protection of human rights with regard to social networking services.

[23] See Recommendation CM/Rec(2016)5 of the Committee of Ministers to member States on Internet freedom and Recommendation CM/Rec(2014)6 of the Committee of Ministers to member States on a Guide to human rights for internet users.

significant effect on the amount of time lapsed until an individual receives a satisfactory response. In the context of automated content removal processes on social media platforms (see above C.), the use of algorithms is particularly evident in the response times that different types of content receive and how content is prioritised, a process that is evidently automated. The same goes for the threshold of user complaints that are required before a piece of content is reviewed. There are strong suggestions that the complete responses of internet intermediaries such as Facebook, Google or Microsoft to user queries are automated for many types of inquiries and complaints (Wagner 2016; Zhang, Stalla-Bourdillon, and Gilbert 2016). Often, many users will need to complain about a specific type of content before an automated algorithm identifies it as relevant to be referred to a human operator for content review. These operators are reported to be working often under time pressure and with minimal instructions as to what specifically to remove in line with internal "deletion rules".[24] The right to an effective remedy explicitly implies the right to a reasoned and individual decision. Thus far, all such decisions have been taken by human beings who, in the exercise of their functions and based on comprehensive training, have been granted a considerable margin of discretion. In principle, it is a judge or administrative official's discretion to decide how the balancing of individual rights, such as the freedom of expression and the protection from violence, shall be put into practice based on a careful case-by-case analysis of the individual context, condition and character of the situation at hand. As a result of the increased use of algorithmic data processing techniques in complaints procedures, however, algorithms are gradually replacing humans.

In addition, serious concerns exist as to whether automatic response processes to complaints constitute an effective remedy. While the famous removal of a YouTube video on a European Parliament debate related to torture was reinstated after only few hours, following an MEP complaint, who even received a public apology from Google, there are considerable doubts as to whether all complaints are treated with such attentiveness.[25] Rather, algorithms often obscure access to a reasoned explanation as to why certain steps were taken in a particular case.

Orders made by public authorities to restrict access to a specific website or content are often based on vague terms such as "hate speech" or "extremist" that appear often not to have been assessed in terms of their human rights compliance (Husovec 2014). In doing so, the public authority may pass the choice of tools and measures onto a private party, which can only then implement solutions (such as content removal or restriction) that the public authorities themselves could not legally prescribe. Public-private partnerships may thus allow public actors "to impose regulations on expression that could fail to pass constitutional muster" (Mueller 2010:213) in contravention of rule of law standards. Moreover, these kinds of demands by public institutions of private actors lead to overbroad and automated filtering of content, as these are the most cost-effective responses to a public request to "remove all hate speech".

---

[24] See http://international.sueddeutsche.de/post/154513473995/inside-facebook.

[25] See https://www.marietjeschaake.eu/en/when-youtube-took-down-my-video.

With respect to the right to privacy, automated techniques and algorithms facilitate forms of secret surveillance and "dataveillance" that are impossible for the affected individual to know about. The European Court of Human Rights has underlined that the absence of notification at any point undermines the effectiveness of remedies against such measures.[26]

F. PROHIBITION OF DISCRIMINATION – ARTICLE 14 ECHR

Another key fundamental freedom that is frequently cited in relation to the operation of algorithms is the right to protection against discrimination.

Search algorithms and search engines by definition do not treat all information equally. While processes used to select and index information may be applied consistently, the search results will typically be ranked according to perceived relevance. Accordingly, different items of information will receive different degrees of visibility, depending on which factors the ranking algorithm takes into account.[27] As a result of data aggregation and profiling, it is possible that search algorithms and search engines rank the advertisement of smaller companies that are registered in less affluent neighbourhoods lower than those of large entities which may put them at a disadvantage. Search engines and search algorithms may also not treat all users equally. Different users may be presented with different results, on the basis of behavioural or other profiles, including personal risk profiles that may be developed for the purpose of insurance or credit scoring or more generally for differential pricing, i.e., offering different prices for the same goods or services to different consumers based on their profile.[28]

A biased algorithm within a large quasi-monopolistic search engine that systematically discriminates one group in society, for example based on their age, sexuality, race, gender or socio-economic standing, may raise considerable concerns not just in terms of the access to rights of the individual end-users or customers affected by these decisions, but also for society as a whole.[29] It can be argued as a result that individuals should have the right to view an 'unbiased' and not personally targeted version of their search results. This can be seen as a way for an individual to exit their own 'filter bubble' and see an untargeted version of the search content, social media timeline or other internet-based service or product that they are using. As a matter of fact, algorithms may be useful tools to reduce bias in places where it is common, such as in hiring processes.

---

[26] See *Roman Zakharov v. Russia* (application 47143/06) of 4 December 2015.

[27] The algorithm may also – deliberately or not – be impacted by a variety of external factors, which may relate to business models, legal constraints (e.g. copyright) or other contextual factors.

[28] Submission from Aleksandra Kuczerawy, Brendan van Alsenoy and Jef Ausloos.

[29] Submission from Sophie Stalla-Bourdillon, Steffen Staab and Laura Carmichael.

G. SOCIAL RIGHTS AND ACCESS TO PUBLIC SERVICES

The workplace is another key area where automated decision-making has become increasingly common in recent years. Algorithms may be involved in decisions on both hiring and firing staff, staff organisation and management, as well as the individual evaluations of employees. Automated feedback loops, sometimes linked to customer input, may decide over the performance evaluation of staff (Kocher and Hensel 2016). These decision-making processes are by no means perfect when humans conduct them. Bias related to race (Bertrand and Mullainathan 2004) class and gender (Altonji and Blank 1999; Goldin and Rouse 1997) has been demonstrated repeatedly in human resources management practices and processes. With more and more companies moving towards algorithmic recruitment methods (Rosenblat, Kneese, and others 2014), however, new concerns related to the lack of transparency in the decisions they make, both in the hiring process and beyond, have been raised. Moreover many of these automated decision-making processes are based on data received via internet intermediaries. Allowing the 'wisdom of the crowd' to make decisions about individuals' employment is not only highly questionable from an ethical point of view, it also limits the ability of workers to contest such decisions as they seem to be an 'objective' measures of their performance (Tufekci et al. 2015). This may raise concerns with respect to the rights contained in the Revised European Social Charter.

As individual employment platforms are "transforming people into Human Computation," (Irani 2015:227) questions arise about workers' rights, employee self-determination and how societies as a whole believe that human beings should be treated at the workplace. Notably the increased automation in the workplace also raises considerable challenges in relation to privacy rights (Hendrickx and van Bever 2013) of employees and how they can be safeguarded in the workplace. As more and more systems are automated and more and more data is collected at the workplace, employees' rights under Article 8 are evidently in danger even if they are not directly targeted by such data collection measures (see above B.) Finally, there are additional challenges related to the usage of algorithms by both public and private sector organisations to monitor staff communications. Such practices are typically employed to ensure that staff represent well either a company or a bureaucracy and have evident implications for the freedom of expression of the employees (Voorhoof and Humblet 2013) and their human rights under Article 10 of the Convention (see above C.).

Government agencies and services are increasingly automating their decision-making with the use of algorithms (van Haastert 2016). While it is heavily debated whether such systems can increase efficiency or not, what is evident is that the operation of such systems poses considerable questions for transparency and accountability of public decision-making, which must be held to a higher standard in their decision-making than the private or non-profit sector. At present the public sector in Europe is employing automated decision-making in areas as diverse as social security, taxation, health care and the justice system (van Haastert 2016; Tufekci et al. 2015). There is considerable danger of social sorting in medical data as algorithms can sort out specific citizen groups or human profiles, thereby possibly preventing their access to social services. Another example relates to the practice of *Profiling the Unemployed in Poland,* which was analysed by researchers in an effort to

assess the social and political implications of algorithmic decision-making associated with social benefits (Jędrzej Niklas, Karolina Sztandar-Sztanderska, and Katarzyna Szymielewicz 2015). This analysis identified several challenges which are relevant also for the use of algorithms in other areas of the public sector service delivery, such as non-transparent and algorithmic rules being applied in the distribution of public services and computational shortcomings triggering arbitrary decisions, for instance, with respect to receipt of social benefits.

H. THE RIGHT TO FREE ELECTIONS

The operation of algorithms and automated recommender systems, that may create 'filter bubbles' - fully-automated echo chambers in which individuals only see pieces of information that confirm their own opinions or match their profile (Bozdag 2013; Pariser 2011; Zuckerman 2013) - can have momentous effects for democratic processes in society. Fully-automated echo chambers pose the danger of creating "ideological bubbles" (O'Callaghan et al. 2015), that may be relatively easy to enter but hard to exit (Salamatian 2014), and they may have crucial effects in particular in the context of elections.

While it has been argued since the advent of the internet that online campaigning and social media networks were likely to change the way in which politics and elections were run, it is only more recently that academic research has revealed the extent to which the curation and manipulation of online content on social media platforms may 'tip' elections. During U.S. elections, researchers reportedly manipulated the Facebook platform to influence users voting behaviour by telling them how their friends had said they had voted, without users' knowledge, and were able to convince a statistically significant segment of the population to vote in the congressional mid-term elections on 2 November 2010 (Bond et al. 2012). [30] There are strong indications that since then Facebook has been selling related political advertising services to political parties around the world, with similar behaviour observed during the UK local elections in 2016 (Griffin 2016). Whether Facebook and similar quasi-monopolistic online platforms are using their power to influence human voting benevolently or not is less the point than the fact that they – in principle – have the ability to massively influence elections.

The right to free elections, as established by Article 3 of Protocol 1 has been acknowledged by the European Court of Human Rights as fundamental principle in a truly democratic political regime. Importantly as noted in the draft feasibility study on the use of internet in elections by the Committee of Experts on Media Pluralism and Transparency of Media Ownership (MSI-MED) at the Council of Europe, regulatory challenges related to elections are not due to the rise of intermediaries but rather a lack of adequate regulation. As the

[30] In an experiment, Facebook researchers showed a graphic to some users in their news feed, indicating how many of their friends had voted that day and providing a button to click that they had voted as well. Users who were prompted with news of their friends' voting turned out to be 0.39% more likely to vote than the others, and their decision had a further effect on the voting behavior of their friends. The researchers concluded that their single message on Facebook, strategically delivered, increased turnout directly by 60,000 voters, and thanks to the ripple effect, ultimately caused an additional 340,000 votes to be cast (amongst an overall 82 million) that day. See Jonathan Zittrain, *Engineering an election,* Harvard Law Review Forum Vol. 127, 335 – 339 (2014).

study notes the "most fundamental, pernicious, and simultaneously difficult to detect implication of the shift to social media is not the rising power of intermediaries but the inability of regulation to level the playing field for political contest and limit the role of money in elections."(Cross-reference to CoE MSI-MED Report by Damian Tambini). Use and effects of bots, fake news, effects on pluralism and social cohesion

4. MECHANISMS OF GOVERNANCE, ACCOUNTABILITY, TRANSPARENCY AND ETHICS

A. INTRODUCTION

Historically, challenges related to automated data processing have been addressed through data protection legislation. Today, relevant and innovative approaches such as the introduction of the "right to explanation" (Goodman and Flaxman 2016) are also the product of data protection legislation. However, there is a significant difference between the right to privacy and data protection regulation, which is in the end still a governance mechanism to safeguard privacy and other rights. While it is clear that the challenges around discrimination of content or the manipulation of elections go beyond privacy and data protection and raise questions of competition law or may be relevant to electoral commissions and parliaments, the expertise of the data protection community may well be drawn from when attempting to identify suitable regulatory responses to algorithmic governance.

Finally, very fundamental legal and ethical questions surround the legal personhood of automated systems such as algorithms that cannot easily be resolved in this report. While not wishing to exculpate those involved in development, programming and implementation of autonomous systems, it must be acknowledged that automation, vast data analysis and adaptability and self-learning create considerable challenges for accountability of algorithmic decisions. This has led some authors to suggest that many forms of algorithmic transparency, accountability and regulation are impossible because the programmers themselves are unable to predict or fully understand how the algorithm makes the decisions that it makes (Kroll 2016).

B. IS IT REASONABLE AND POSSIBLE TO REGULATE ALGORITHMS?

There is growing concern in Europe at the political and public level that the use of algorithms as such poses considerable challenges to human rights and should therefore be regulated.[31] While there is no consensus in terms of what mechanisms would be appropriate for regulating the use of algorithms, there are already numerous cases in which governments and independent auditors regulate algorithms before implementing them.

[31] See, for instance, the vote on 26 January 2016 in the French National Assembly for a new Bill on digital rights. The Bill includes provisions relating to algorithmic transparency and the duty of 'loyalty', or fairness, of online platforms and algorithmic decision-making" (Rosnay 2016).

The software and algorithms used in 'slot machines' in Australia and New Zealand must, by government regulation, be "fair, secure and auditable" (Woolley et al. 2013). As part of this process, the developers of such machines are required to submit their algorithms to regulators before they can be presented to consumers. The Australian/New Zealand Gaming Machine National Standard in its most recent revision 10.3 defines in extraordinary technical detail how such machines should operate. For example the "Nominal Standard Deviation (NSD) of a game must be no greater than 15" and "the hashing algorithm for the verification of gaming equipment software, firmware and PSDs is the HMAC-SHA1 algorithm".[32] Gambling equipment in the United Kingdom is also controlled by a specific licensing regime. There is further an ongoing debate in the financial sector about the regulation of high-speed trading algorithms as these are seen to have a strong potential destabilising effect on the overall financial system. One leading politician suggested in 2012 that financial trading "algorithms will have to undergo a stress test to ascertain their stability" (Steinbrück 2012). Similar regulation has been portended in the area of online content regulation and internet hotlines. The British Police Child Exploitation and Online Protection Centre demanded that their 'Facebook button' be provided by default to all internet users (Wagner 2016). While this attempt to pressure Facebook into changing its default code on the British Facebook website was unsuccessful, it suggests what kind of regulatory responses could be expected if states begin to define the content of algorithms on large online platforms.

C. TRANSPARENCY

To many consumers and regulators, algorithms seem like black boxes to both consumers and regulators (Pasquale 2015). As Tufekci et al note: "a common ethical concern about algorithmic decision-making is the opaque nature of many algorithms. When algorithms are employed to make straightforward decisions, such as in the case of medical diagnostics or aviation, a lack of transparency raises important questions of accountability" (Tufekci et al. 2015:11). Thus there is a frequent and growing debate about algorithmic transparency, including government requests to companies which algorithms should be reviewed by independent auditors, regulators or the general public (Diakopoulos 2015; Rosnay 2016) before their introduction.

Provision of entire algorithms to the public is unlikely, as private companies regard their algorithm as their key trade secret.[33] However, there is also a debate around the possibility of providing key subsets of information about the algorithms to the public, for example which variables are in use, the average values and standard deviations of the results produced or the amount and type of data being processed by the algorithm.

---

[32] The Australian/New Zealand Gaming Machine National Standard which is available here: https://publications.qld.gov.au/dataset/a-nz-gaming-machine-national-standards

[33] In a decision of 28 January 2014, the German Federal Supreme Court (Bundesgerichtshof) rejected a claim for information concerning a credit agency's algorithm as it was a protected business secret. It, however, allowed a claim for information concerning the data used to calculate creditworthiness through the means of the algorithm. (SOURCE?)

All of these measures aim to increase transparency of automated systems, complicated by the frequent changes in the algorithms used. Google, for example, changes its algorithm hundreds of times per year (Tufekci et al. 2015). There is also the frequent danger of manipulation and 'gaming' of algorithms if they are made public. At the same time, machine learning techniques complicate transparency to a point where provision of all of the source codes of an algorithm may not even be sufficient, and instead there is a need for an actual explanation of how the results of an algorithm were produced. Initial steps towards a right to *effective* transparency can be drawn from the European General Data Protection Regulation (GDPR), including a possible right to explanation (Goodman and Flaxman 2016).

As the use of algorithms in decision-making potentially prejudices the rights of individuals, an oversight mechanism may ensure that the algorithm operates in a fair and sustainable manner. For example, section 28 b of the German Federal Law on Data Protection provides that there has to be a scientifically proven mathematical-statistical process for the calculation of the probability of a specific behaviour of an individual before such an algorithm can be used for making a decision about a contract.

D. ACCOUNTABILITY

What accountability do individuals or companies have for the algorithms they implement? This depends very much on the nature of the algorithms and their outputs. In many cases, if the outputs are defamatory, infringe copyright or raise other legal concerns, existing governance mechanisms ensure that these kinds of outputs are limited (Staab, Stalla-Bourdillon, and Carmichael 2016). The case of Max Mosley taking action against Google is just one of many examples (Stanley 2011). However, such mechanisms typically only affect second order rules, i.e. changes to the outputs of algorithms. By contrast, there is a general lack of regulatory frameworks to influence first order rules and ensure that algorithms in the first place are producing results that uphold and protect fundamental values or basic ethical and societal principles.

However, it has been suggested that "[t]echnologists think about trust and assurance for computer systems a bit differently from policymakers, seeking strong formal guarantees or trustworthy digital evidence that a system works as it is intended to or complies with a rule or policy objective rather than simple assurances that a piece of software acts in a certain way." (Kroll et al. 2016)

This in turn feeds into the wider debate on auditing of algorithms by which 'zero knowledge proofs' could conceivably be generated by algorithms to demonstrate that they conform to certain properties without the individual engaging in the proof being able to see the actual algorithm (Kroll 2016).

E. ETHICAL FRAMEWORKS AND IMPROVED RISK ASSESSMENT

Aside from direct regulatory mechanisms to influence the code of algorithms, indirect mechanisms to influence algorithm codes could also be considered. These address the production process or the producers of algorithms and attempt to ensure that they are aware of the legal challenges, ethical dilemmas and human rights concerns raised by automated decision-making. An instrument to achieve such goal could consist of standardised professional ethics or forms of licensing system for data engineers and algorithm designers similar to those that exist for professions like doctors, lawyers or architects.[34] Another suggestion frequently made is that existing mechanisms for the management and development processes of software could be improved (Spiekermann 2015). This may particularly concern agile software development techniques where modularity, temporality and capture pose considerable challenges for privacy (Gürses and Hoboken 2017) as well as other human rights (Mannaro 2008).

Importantly these challenges exist not just for professionals who develop algorithms, but also for 'data scientists' who use them. It has been frequently argued that much of the usage of algorithms in machine learning takes places without "understanding" causal relationships (correlation instead of causation), which may lead to bias and errors and raise concerns about data quality (O'Neil 2016). The challenge, however, relates less to the algorithms themselves and more to the way human beings perceive and interpret their results. The belief that computer algorithms produce neutral unbiased results (Chun 2006) without any form of politics (Denardis 2008) is at the heart of this problem. Accordingly, it would be more helpful to ensure more critical engagement in public debates about algorithms than to attempt to change them.

The direct regulation of algorithms or software codes should be approached with extreme care. It is the regulatory approach that provides the most pitfalls and is most likely to exacerbate problems. Notably, direct regulation raises considerable concerns about freedom of opinion and expression and the right to privacy. Moreover, given the fact that regulators typically do not have comprehensive knowledge about algorithms, greater steps towards transparency and accountability of algorithms would seem far more appropriate.

5. CONCLUSIONS

Understanding how automated decision-making systems operate is fraught with great difficulty and raises numerous human rights challenges. Many of these challenges are so difficult to assess because the field is comparatively new and finding effective solutions remains difficult. As a first step, policy-makers should seek to learn more about the implementation of automated decision-making systems in their respective countries. As a second step, they should try to ensure that existing laws and legal frameworks are effectively implemented in response to the challenges posed by automated decision-making in the various spheres of their application. Here the findings of this study are similar to the MSI-MED draft feasibility study on the use of internet in elections, which suggests that the

[34] Submission from Markus Oermann, University of Hamburg.

key challenges faced are not related to the rising importance in the role of intermediaries but rather due to regulatory failures of governance.

The findings in this report should not be understood as calls for regulating the development of algorithms or other software codes. Interference with the right of individuals to research, develop and test could constitute in itself a violation of their freedom of opinion, expression, thought and research. Aside from the significant human rights impacts of regulating research and development of algorithms, it would prevent a deeper understanding of how algorithms operate and what effects they have.

Nonetheless policy discussions related to algorithms and automated data-processing techniques should be guided by legal, social and ethical considerations that are interrelated and interdependent and, broadly, related to the issues of effective transparency, accountability and the need for continued research and development.

SUMMARY AND MAIN CONCLUSIONS

Summary will be inserted here once a final version of the study has been completed

BIBLIOGRAPHY

Altonji, JG and RM Blank. 1999. '*Race and Gender in the Labor Market*'. Pp. 3143–3259 in Handbook of labor economics. Elsevier B.V. Retrieved (http://www.sciencedirect.com/science/article/pii/S1573446399300390).

Andreessen, Marc. 2011. '*Why Software Is Eating The World'*. Wall Street Journal, August 20. Retrieved 1 September 2016 (http://www.wsj.com/articles/SB10001424053111903480904576512250915629460).

Bertrand, Marianne and Sendhil Mullainathan. 2004. '*Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*'. The American Economic Review 94(4):991–1013.

Bond, Robert M. et al. 2012. '*A 61-Million-Person Experiment in Social Influence and Political Mobilization'.* Nature 489(7415):295–298.

Bozdag, Engin. 2013. '*Bias in Algorithmic Filtering and Personalization*'. Ethics and Information Technology 15(3):209–227.

Bucher, Taina. 2012. '*Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook*'. New Media & Society 1461444812440159.

Bucher, Taina. 2016. '*The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms*'. Information, Communication & Society 1–15.

Buni, Catherine and Soraya Chemaly. 2016. '*The Secret Rules of the Internet*'. The Verge. Retrieved 9 September 2016 (http://www.theverge.com/2016/4/13/11387934/internet-moderator-history-youtube-facebook-reddit-censorship-free-speech).

Chun, Wendy Hui Kyong. 2006. C*ontrol and Freedom : Power and Paranoia in the Age of Fiber Optic*s. Cambridge Mass.: MIT Press.

Denardis, Laura. 2008. '*Architecting Civil Liberties*'. in Global Internet Governance Academic Network Annual Meeting. Hyderabad (Andra Pradesh), India: GIGANET. Retrieved (http://worldcat.org/oclc/619234880/viewonline).

Denardis, Laura. 2012. '*Hidden Levers of Internet Control*'. Information, Communication & Society (September):37–41.

Diakopoulos, Nicholas. 2015. '*Algorithmic Accountability*'. Digital Journalism 3(3):398–415.

Gillespie, Tarleton. 2014. '*The Relevance of Algorithms*'. Pp. 167–94 in Media technologies: Essays on communication, materiality, and society, edited by T. Gillespie, P. J. Boczkowski, and K. A. Foot. Cambridge Mass.: MIT Press.

Goldin, Claudia and Cecilia Rouse. 1997. *Orchestrating Impartiality: The Impact Of' blind' auditions on Female Musicians.* National bureau of economic research. Retrieved 9 September 2016 (http://www.nber.org/papers/w5903).

Goodman, Bryce and Seth Flaxman. 2016. '*European Union Regulations on Algorithmic Decision-Making and a Right to Explanation*'. in 2016 ICML Workshop on Human Interpretability in Machine Learning. New York, NY: ArXiv e-prints.

Griffin, Andrew. 2016. '*How Facebook Is Manipulating You to Vote'*. The Independent. Retrieved 31 August 2016 (http://www.independent.co.uk/life-style/gadgets-and-tech/news/uk-elections-2016-how-facebook-is-manipulating-you-to-vote-a7015196.html).

Gürses, Seda and Joris Hoboken. 2017. '*Privacy After the Agile Turn*'. in The Cambridge Handbook of Consumer Privacy, edited by Selinger. Retrieved (https://osf.io/ufdvb/).

van Haastert, Hugo. 2016. '*Government as a Platform: Public Values in the Age of Big Data*'. Oxford Internet Institute.

Helberger, Natali and Damian Trilling. 2016. '*Facebook Is a News Editor: The Real Issues to Be Concerned about*'. Media Policy Project. Retrieved 9 September 2016 (http://blogs.lse.ac.uk/mediapolicyproject/2016/05/26/facebook-is-a-news-editor-the-real-issues-to-be-concerned-about/).

Hendrickx, Frank and Aline van Bever. 2013. '*Article 8 ECHR: Judicial Patterns of Employment Privacy Protection*'. Pp. 183–208 in The European Convention on Human Rights and the Employment Relation, edited by F. Dorssemont, K. Lörcher, and I. Schömann. Oxford: Hart Publishing.

Husovec, Martin. 2014. '*CJEU Allowed Website-Blocking Injunctions with Some Reservations*'. Journal of Intellectual Property Law & Practice jpu101.

Irani, L. 2015. '*Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk*'. South Atlantic Quarterly 114(1):225–234.

Jędrzej Niklas, Karolina Sztandar-Sztanderska, and Katarzyna Szymielewicz. 2015.Warsaw, Poland: Panoptykon Foundation. Retrieved (https://en.panoptykon.org/articles/profiling-unemployed-poland-%E2%80%93-report).

Kitchin, R. and M. Dodge. 2011. *Code/Space Software and Everyday Life*.

Kocher, Eva and Isabell Hensel. 2016. '*Herausforderungen Des Arbeitsrechts Durch Digitale Plattformen – Ein Neuer Koordinationsmodus von Erwerbsarbeit*'. Neue Zeitschrift Für Arbeitsrecht (16/2016):984–89.

Kroll, Joshua A. et al. 2016. '*Accountable Algorithms*'. Retrieved 1 September 2016 (http://balkin.blogspot.co.at/2016/03/accountable-algorithms.html).

Kroll, Joshua A. 2016. '*Accountable Algorithms (A Provocation)'*. Media Policy Project. Retrieved 9 September 2016 (http://blogs.lse.ac.uk/mediapolicyproject/2016/02/10/accountable-algorithms-a-provocation/).

Lazer, David and Ryan Kennedy. 2015. *What We Can Learn from the Epic Failure of Google Flu Trends.*

Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. '*The Parable of Google Flu: Traps in Big Data Analysis*'. Science 343(6176):1203–5.

Mannaro, Katiuscia. 2008. '*Adopting Agile Methodologies in Distributed Software Development*'. Universita` degli Studi di Cagliari, Cagliari. Italy. Retrieved (http://le.uwpress.org/content/87/2/284.short).

McCarthy, Daniel R. 2011. '*Open Networks and the Open Door: American Foreign Policy and the Narration of the Internet*'. Foreign Policy Analysis 7(1):89–111.

Mueller, Milton. 2010. *Networks and States: The Global Politics of Internet Governance*. MIT Press.

O'Callaghan, D., D. Greene, M. Conway, J. Carthy, and P. Cunningham. 2015. '*Down the (White) Rabbit Hole: The Extreme Right and Online Recommender Systems'*. Social Science Computer Review Social Science Computer Review 33(4):459–78.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Perry, Walt L. 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Rand Corporation. Retrieved 9 September 2016 (https://books.google.com/books?hl=en&lr=&id=ZdstAQAAQBAJ&oi=fnd&pg=PP1&dq=Perry,+Walter,+and+Brian+McInnis.+2013.+Predictive+Policing:+The+Role+of+Crime+Forecasting+in+Law+Enforcement+Operations.+Santa+Monica,+CA:+RAND.&ots=924yNa6Vct&sig=N3HnEi1FBr9YyMXV77GsgPbovYc).

Rifkind, Malcolm. 2014. *Report on the Intelligence Relating to the Murder of Fusilier* Lee Rigby.

Rosenblat, Alex, Tamara Kneese, and others. 2014. '*Networked Employment Discrimination'. Open Society Foundations*' Future of Work Commissioned Research Papers. Retrieved 9 September 2016 (http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2543507).

Rosnay, Mélanie Dulong de. 2016. '*Algorithmic Transparency and Platform Loyalty or Fairness in the French Digital Republic Bill*'. Media Policy Project. Retrieved 1 September 2016 (http://blogs.lse.ac.uk/mediapolicyproject/2016/04/22/algorithmic-transparency-and-platform-loyalty-or-fairness-in-the-french-digital-republic-bill/).

Rubinstein, Ira, Ronald D. Lee, and Paul M. Schwartz. 2008. *Data Mining and Internet Profiling: Emerging Regulatory and Technological Approaches*. Rochester, NY: Social Science Research Network. Retrieved 9 September 2016 (http://papers.ssrn.com/abstract=1116728).

Salamatian, Kavé. 2014. '*From Big Data to Banality of Evil'*. Retrieved 9 September 2016 (https://www.oximity.com/article/Vortrag-Big-Data-und-Ethik-1).

Schulz, Wolfgang and Kevin Dankert. 2016. ''*Governance by Things' as a Challenge to Regulation by Law*'. Internet Policy Review 5(2).

Sills, Arthur J. 1970. '*Automated Data Processing and the Issue of Privacy*'. Seton Hall Law Review 1.

Spiekermann, Sarah. 2015. *Ethical IT Innovation: A Value-Based System Design Approach*. CRC Press.

Staab, Steffen, Sophie Stalla-Bourdillon, and Laura Carmichael. 2016. '*Observing and Recommending from a Social Web with Biases*'. arXiv Preprint arXiv:1604.07180. Retrieved 9 September 2016 (http://arxiv.org/abs/1604.07180).

Stanley, JE. 2011. '*Max Mosley and the English Right to Privacy*'. Wash. U. Global Stud. L. Rev. 10(3). Retrieved (http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/wasglo10&section=25).

Steinbrück, Peer. 2012. *Vertrauen Zurückgewinnen: Ein Neuer Anlauf Zur Bändigung Der Finanzmärkt*e. Berlin, Germany: Deutscher Bundestag - German Federal Parliament.

Tene, Omer and Jules Polonetsky. 2012. '*To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioral Advertising*'. Retrieved 9 September 2016 (http://conservancy.umn.edu/handle/11299/155947).

Toor, Amar. 2016. '*Automated Systems Fight ISIS Propaganda, but at What Cost?*' The Verge. Retrieved 9 September 2016 (http://www.theverge.com/2016/9/6/12811680/isis-propaganda-algorithm-facebook-twitter-google).

Tufekci, Zeynep, Jillian C. York, Ben Wagner, and Frederike Kaltheuner. 2015. *The Ethics of Algorithms: From Radical Content to Self-Driving Cars.* Berlin, Germany: European University Viadrina. Retrieved (https://cihr.eu/publication-the-ethics-of-algorithms/).

Voorhoof, Dirk and P. Humblet, eds. 2013. '*The Right to Freedom of Expression in the Workplace under Article 10 ECHR*'. Pp. 183–208 in The European Convention on Human Rights and the Employment Relation. Oxford: Hart Publishing.

Wagner, Ben. 2016. *Global Free Expression: Governing the Boundaries of Internet Content.* Cham, Switzerland: Springer International Publishing.

Williamson, Ben. 2016. '*Computing Brains: Learning Algorithms and Neurocomputation in the Smart City*'. Information, Communication & Society 0(0):1–19.

Winner, L. 1980. '*Do Artifacts Have Politics?*' Daedalus.

Winner, L. 1986. '*The Whale and the Reactor: A Search for Limits in an Age of High Technology*'.

Woolley, Richard, Charles Livingstone, Kevin Harrigan, and Angela Rintoul. 2013. '*House Edge: Hold Percentage and the Cost of EGM Gambling'*. International Gambling Studies 13(3):388–402.

York, Jillian C. 2010. '*Policing Content in the Quasi-Public Sphere*'. Boston, MA: Open Net Initiative Bulletin. Berkman Center. Harvard University.

Zhang, Pei, Sophie Stalla-Bourdillon, and Lester Gilbert. 2016. '*A Content-Linking-Context Model For "notice-and-Take-Down" procedures'.* Pp. 161–65 in. ACM Press. Retrieved 9 September 2016 (http://dl.acm.org/citation.cfm?doid=2908131.2908171).

Zuckerman, Ethan. 2013. *Digital Cosmopolitans: Why We Think the Internet Connects Us, Why It Doesn't, and How to Rewire It.* W. W. Norton & Company.