20 February 2017

MSI-NET(2016)06 rev

STUDY ON THE HUMAN RIGHTS DIMENSIONS OF ALGORITHMS

SECOND DRAFT (20 FEBRUARY 2017)

SUBMITTED FOR THE 3ᴿᴰ MEETING (27-28 MARCH 2017)

Rapporteur: Ben Wagner

TABLE OF CONTENTS

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 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

4

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.[1] 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

---

[1] 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).

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.

3. IMPACTS OF ALGORITHMS ON HUMAN RIGHTS

The principle reservations towards algorithms and automated data processing techniques usually point to their opacity and unpredictability.[2] 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.[3]

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.).

---

[2]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.

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

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.[4]

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.[5] 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"[6] 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

---

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

[5] See Sills 1970.

[6] 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.

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.[7]

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).[8]

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

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

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

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.[9]

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,[10] their potential for harming the freedom of information and freedom of expression of individuals, groups and whole segments of societies is increasingly being underlined.[11]

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 content or content providers, thereby risking to affect related values such as media pluralism and diversity.[12] 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.

---

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

[10] 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."*

[11] 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".

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

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, [13] 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.[14] 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 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.'[15]

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.[16] The

---

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

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

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

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

system has reportedly been automated with the introduction of the Joint Referral Platform announced in April 2016.[17]

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 third-party submissions, exert a form of editorial control traditionally performed by media professionals and therefore create specific media responsibilities.[18]

D. FREEDOM OF ASSEMBLY AND ASSOCIATION – ARTICLE 11 ECHR

---

[17] 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.

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

13

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.[19] 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.[20]

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.

---

[19] 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.

[20] 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.

An increasing number of companies, especially larger ones, use algorithms and automated data processing techniques for running their complaints procedures. This can have a 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".[21] 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.[22] Rather, algorithms often obscure access to a reasoned explanation as to why certain steps were taken in a particular case.

---

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

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

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".

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.[23]

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.[24] 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

---

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

[24] 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.

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.[25]

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.[26] 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.

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.

---

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

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

17

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). [27] 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 study notes the "most fundamental, pernicious, and simultaneously

---

[27] 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).

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.[28] 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.

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".[29] 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

---

[28] 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).

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

21

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.[30] 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.

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

---

[30] 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?)

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).

23

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.[31] 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

---

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

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 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

25

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.

26

Study on the human rights dimensions of algorithms - 2nd draft (20/02/2017)

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. 'T*he 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. 'T*o 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.

28