LSE Media and Communications

Media@LSE Working Paper Series

Editors: Bart Cammaerts, Nick Anstead and Richard Stupart

‘Algorithmic Bias’ through the Media Lens

A Content Analysis of the Framing of Discourse

Rocío Izar Oyarzun Peralta

Published by Media@LSE, London School of Economics and Political Science ("LSE") Houghton Street, London WC2A 2AE. The LSE is a School of the University of London. It is a Charity and is incorporated in England as a company limited by guarantee under the Companies Act (Reg number 70527).

ABSTRACT

This research project investigated the 'algorithmic bias' phenomenon from a social sciences perspective, through the analysis of media framing. Concepts belonging to the critical literature on the study of algorithms were considered alongside academic work on media framing and theories on the reproduction of dominant discourses in order to design a theoretical framework for the evaluation of how the media, in this case the mainstream media in the United Kingdom and technology publications, cover and frame the topic of 'algorithmic bias'.

The study understands framing as a process that tends to reproduce dominant perspectives, which in this case is considered to be the corporate perspective. For that reason, this work attempts to detect whether the media reproduces corporate discourse in its coverage or contests it with other possible discourses. The study looked for contrasts in coverage patterns, aiming to detect the presence of potential different framings of the topic, in a two-way comparison between the two kinds of media and two specific time periods: before and after the Cambridge Analytica scandal. The method used was quantitative content analysis: the project analysed a sample of 187 news articles to reveal the extent to which the framing of discourses varied and analysed these findings in light of the theoretical framework. The main finding of this work was that the framing did not vary abruptly after the occurrence of the scandal.

1. INTRODUCTION

Algorithms are understood as ‘socio-technical assemblages’ and the ultimate ‘power brokers’ of society, embodying a legitimate way to produce the information upon which society relies nowadays (Ananny, 2015; 93; Diakopoulos, 2013: 2). In this context of growing relevance, some claim algorithms are producing pernicious effects over society, one of them being the ‘algorithmic bias phenomenon’. ‘Algorithmic bias’ is defined by this work as ‘*the output of an algorithm that systematically presents results that discriminate against certain groups of individuals or that systematically generates information that is flawed or tends to be skewed towards particular topics or areas of interest*’.

The objective of this work is to analyse the content of the coverage of ‘algorithmic bias’ in the UK mainstream media and in technology publications so as to contribute to the understanding of the way media frames and reports this topic. The main question that concerns this research is: ‘*How has the UK mainstream media and the technology publications framed the discussion on algorithmic bias in two time periods selected?’* The two sub-questions are: *‘How prominent is corporate discourse in the coverage on algorithmic bias as compared to other discourses?’* and *‘What are the most salient issues in the discussion on algorithmic bias?’*

This work is divided into four sections. The theoretical framework section incorporates specific concepts from the literature presented in order to define the particular approach this social sciences study will take. Furthermore, the section will also present an understanding of the dominant discourse regarding algorithms, the corporate discourse, which is based on the critical literature on algorithms. Framing is understood by this work as a media process that allows for the highlighting of particular aspects of reality and that tends to reproduce the dominant perspective, which this work defines as the corporate perspective. However, the way in which media frames topics is also considered to be evolutionary and changing. Theories of the media framing of technology will help illustrate how the media tends to frame these topics. Even though framing is contended to reproduce mainly dominant perspectives, this work argues that media can be a ‘two-fold’ actor, which usually reproduces the dominant perspectives but might also contest such discourse with others. An additional factor is introduced at the end of the theory section: the occurrence of the Cambridge Analytica scandal, which is seen as a potential turning point for the way media frames algorithmic bias topics.

The way in which this is work attempts to respond to the research questions is explained in the ‘Methodology’ section, where the design of the quantitative content analysis approach is explained. The study presents a two-way comparison: between mainstream UK media and technology publications and before and after the Cambridge Analytica scandal. While doing this comparison, it attempts to detect the presence of the corporate dominant discourse while also considering that there might be other discourses present in the coverage. The codebook, the main methodological tool this work employs, opens up these possibilities and allows us to see which of these topics are more present in the coverage and with which kind of discourse these topics are associated. The quantitative results from the content analysis analysis are presented in the ‘Results’ section, followed by the ‘Discussion’ section, the final part, which assesses the empirical results in relation to the theoretical framework and discusses the main findings.

2. THEORY

2.1 Characterising Algorithms

Definitions of algorithms range from computer to social sciences approaches. For computer science, algorithms are rules designed for computers to ‘do things’: ‘command structures’ (Goffey, 2008: 17) sequences of computational steps that transform inputs into outputs (Cormen, et al. 2009: 5, 13). They are conceived as purely rational tools bringing together mathematics with technology to produce certain and objective results (Seaver, 2013: 2). They are considered equal to any other computer software: they are pieces of technology. Above all, for computer science, algorithmic essence has to do with the automatic character that allows algorithms to act without continuous human monitoring (Winner, 1977).

Computer science characterisation of algorithms is limited to the features above. For social sciences, the algorithm as defined in the paragraph above is not an object of interest in itself: that characterisation falls short for explaining its sociological effects. Algorithms are better understood as ‘algorithmic systems’, which intertwine the complex dynamics of people and code (Seaver, 2013: 9). Further scholars have understood algorithms as wider ‘assemblages’ (Kitchin, 2017: 14; Ananny, 2015: 93) marked by their social construction (Gillespie, 2014: 192) which involves different sorts of individuals, code, data, hardware, regulations, laws, among many other factors (Kitchin, 2017: 20). In that line, algorithms are claimed to be married to the conditions in which they are developed (Geiger, 2014).

A considerable part of the literature focuses on how ‘profoundly performative’ algorithms are (Mackenzie & Vurdubakis, 2011). For some, the main reason why algorithms are a concerning object of study is because they are showing pervasive effects within the social realm (Kitchin, 2014: 26). In recent years, the critical algorithmic literature has been focused on their uses and misuses: the way in which they work is claimed to be harmful, discriminatory and inscrutable (Rubel, Castro, and Pham, 2018: 9). Algorithms have been defined as ‘actants’ (Tufekci, 2015: 207): agents that are not alive but that have ‘computational agency’ over the world. Due to their increasing use, scholars even claim they are ‘the new power brokers in society’ (Diakopoulos, 2013:2). Algorithms are being constituted as entities of public knowledge that have increasing power over the construction of discourse (Gillespie, 2014: 169; Kitchin, 2014: 18). They are claimed to be the fundamental logic behind the information upon which society now depends (Langlois, 2013) fixating structures of power and knowledge (Kushner, 2013) while remodelling the ways which social and economic structures function (Kitchin, 2016: 16). For this stream of literature, the study of algorithms becomes key to understanding available ecosystems of information (Anderson 2011). Algorithms are now seen as the most relevant medium for communications (Gillespie, 2014: 3) and are claimed to be the most ‘protocological’ and computationally organised medium ever (Galloway, 2004). To the social sciences, there is central challenge to the study of algorithms: it is claimed there is not enough visibility of the ways in which they exercise all this agency and power (Diakopoulos, 2013: 2)

2.2 Algorithmic Bias

The social sciences literature on algorithms varies greatly, focusing on different topics: economic consequences (Reich, 2016) labour (Bauer, 2016) commodification (Fuchs, 2015) surveillance (Mann, 2003; Van Dijck, 2014; Mansell, 2015, 2016) user generated content (van Dijck, 2009) and user perspective (Bucher, 2015) just to mention a few. Nevertheless, this work will only focus on the topic of ‘algorithmic bias’: due to their increasing relevance in ecosystems of information, there is a shared concern in the critical literature over the effects algorithms have on the dynamics of participation and integration of individuals in public life because algorithms are argued to produce skewed results – or information – that affect these dynamics.

Algorithmic biases are contented to have different roots. Early literature on computer systems indicates there are three main sources: ‘pre-existing’, ‘technical’ and ‘emergent’ (Friedman & Nissenbaum, 1996). Even though this refers to computer systems in general and might seem outdated, this work finds great value in these categories as they can be used to describe some of the current issues with bias. The ‘pre-existing’ category is linked to how biases already present in social institutions and practices are embedded in the algorithmic design (p.333): biases are introduced, either consciously or unconsciously, into the system. This might occur when data is formalised, so that algorithms can act on it automatically (Gillespie, 2014:170). On the other hand, the work acknowledges ‘technical biases’ (Friedman & Nissenbaum, 1996: 335) constraints of the technology, including hardware, software and peripherals (Diakopoulos, 2013; Drucker, 2013; Kitchin & Dodge, 2011; Neyland, 2015). For example, if a search result is shown alphabetically, one might tend to prioritise the results that are first available. Lastly ‘emergent bias’ (p. 336) results from changes in the context of use of the original algorithms, for instance, when the actual users of the system vary from the users for which the technology was originally designed. As for the definition of ‘algorithmic bias’ this early work also defines biased computer systems as technologies that ‘systematically and unfairly discriminate against certain individuals or groups of individuals in favour of others’ (p.333). This turns algorithms into ‘instruments of injustice’ (p.345). In that sense, the concept of bias is attached to a ‘pejorative’ rather than to a ‘descriptive’ conception, which some argue can lead to confusion about what the actual issue is and how to respond to it: some other academics prefer to understand the algorithmic bias and its roots as simple deviations from standards (Danks & London, 2017: 4691, 92).

In line with the idea of algorithms as ‘instruments of injustice’, highly critical publications have risen in the discussion around algorithmic bias controversial levels (see Noble, 2018; Eubanks, 2018; Wachter-Boettcher, 2017; O’Neil, 2016). Algorithms are argued to be becoming ‘weapons of math destruction’ (O’Neil, 2016:31): it is claimed that the combination of big data with biased algorithms is increasing inequality and threatening democracy. A great part of the literature focuses on how algorithms may generate discrimination in many arenas: loan allocation, emergency responses, medical diagnostics, labour selection, judicial decisions, law enforcement, and education performance indicators, just to mention a few (Rubel, Castro, and Pham, 2018; Noble, 2018; Eubanks, 2018; Wachter-Boettcher, 2017; O’Neil, 2016 amongst others). This line of literature believes that algorithmic developers prioritise the choosing of ‘some aspects’ of over others, simplifying reality in the models (Kitchin, 2017: 16) and ‘camouflaging it with math’. The concern is that these choices are contended to be diverting away from serving ‘real people’ and have moral repercussions for society (O’Neil, 2016: 48). While this literature emphasises the ‘social inequality’ effects of algorithms, other scholars put the attention on how platforms that rely on the use of biased algorithms may become a concern for citizenship as well (Greiger, 2009; van Dijck, 2009; Morozov, 2011; Bozdag, 2013; Mansell, 2015; Tufekci, 2015). Algorithms are here presented as ‘gatekeeping’ tools, that perform an obscure filtering, blocking and processing of the digital information, which is consumed by the public (such as news, social media, blogs, etc). Furthermore, this filtering is claimed to be functional to the revenue purposes of the corporations that own such algorithms (Mansell, 2015: 120) generating concern regarding how profit-driven decision-making affects the diversity and availability of information needed for the full exercise of citizenship (Tufekci, 2015: 208; Mansell, 2015:11).

These cases are related to the claim that algorithms can embody values (Nissenbaum, 2001 :120). Some even claim that these values installed into the systems might be driven by ‘propriety, commercial, institutional, or political gains’ (Gillespie, 2014:191). Furthermore, the critical literature argues that the ‘ill-conceived’ results produced by algorithms usually go unquestioned (O’Neil. 2016: 7). As seen, biased algorithms are linked to unfairness (Friedman & Nissenbaum, 1996: 345; O’Neil, 2016). Some claim that fairness is not likely to be calculated into algorithms (Pitt et al., 2013) as there is a tendency to prioritise the ‘datification of everything’ (Millington & Millington, 2015) and the ‘optimisation’ of millions of people (O’Neil, 2016: 12, 95) and processes. It is helpful to remember that biased algorithms are usually produced by corporations that operate in the marketplace, and this marketplace is characterised by the existence of power asymmetries, in which corporations are claimed to carry a great weight (Turow, 2011). It is argued that algorithms, just like any knowledge system, are shaped to reflect the objectives and principles of those who aim to profit from them (Hesmondhalgh 2006; Mager, 2012; van Couvering, 2010; Gillespie, 2014: 187). Critical approaches contend that fairness is discarded for the sake of profit: corporations are expanding the production and selling of algorithms and as long as the product is cost-effective, little importance is given by their providers and users (governments, corporations, institutions) to the collateral effects on the less-empowered sectors of society (O’Neil, 2016:12).

Even though it will not be developed further, it is important to acknowledge that bias is also discussed, on a lower level, by non-social scientists: pernicious effects are recognised, yet still, there is a tendency to focus on the need to develop knowledge on code and algorithms as a way to improve the status of technology (Steiner, 2012: 218). This contrasts with the critical social science approach: the focus is on the need for more efficiency and optimisation rather than on the causes, consequences and mitigation of algorithms and the power asymmetries that might be embedded in them. Moreover, part of the critical literature relativises the idea of bias. While not denying there is something about algorithms that produces skewed results, there is also a concern about the ways in which bias is defined: how can someone know whether something is biased or not? It is even possible to produce non-biased knowledge through algorithms? (Draude, Klumbyte, Treusch, 2018: 15). According to some, this puts algorithmic scrutiny in a complicated situation (Gillespie, 2014: 75): accusing algorithmic technologies of being biased necessarily implies there is some sort of ‘unbiased’ measure that can be applied to algorithms to make them neutral. Since this is not available, it is contended that all accusations about biased algorithms lack ‘solid ground’ to sustain themselves. Also, even though this work will not be engaging in a user perspective, it is worth noticing the claims that algorithms do not have a ‘one-way’ influence only: the ways in which they prioritise some results generates a ‘loop’ effect between the algorithmic calculations and the users’ ‘calculations’ (Gillespie, 2014: 83) meaning that, somehow, users are also biased in their use of algorithms.

2.3 Corporations and the Dominant Discourse

Even though the critical literature on the topic is growing, algorithms have plenty of ‘evangelists’ (O’Neil, 2016: 13) most of them being corporations, the main providers of these technologies. Based on what the critical literature argues, there is a case for arguing that the dominant discourse about algorithms is the corporate discourse. Even though this work will not engage in a critical discourse analysis methodology of corporate discourse, this section will provide a summary of how the literature conceives the discourse of corporations about the issue of bias.

The critical literature contends that there is an overall denial amongst corporations regarding the existence of bias in algorithms. It is argued that corporations claim that algorithms are nothing but empowering and beneficial to their users and for society (Smart and Shadbolt, 2014; Mansell, 2015: 121); they are portrayed as ‘purely formal beings of reason’ (Goffey, 2008, p. 16; Morozov, 2011). Their main advertised goal is often cost-efficiency: they can generate millions of operations in seconds. Furthermore, their rational character and automation helps them to be characterised as bias-reducing tools, replacing and displacing the interference that might come from potential self-serving intermediaries (Pasquale, 2014:5; Kitchin & Dodge, 2011: 15; Arthur, 2009).

Because of the growing relevance of algorithms in every aspect of life (Gillespie, 2014) it is claimed there is a need for companies to ‘defend’ them publicly, using the promise of ‘algorithmic objectivity’ (Gillespie, 2014: 169): the technical and automational aspects of algorithms is sold as guarantee of neutrality and impartiality. ‘Algorithmic objectivity’ is argued to be a key discursive tool for algorithm providers which allows them to defend algorithmic legitimacy given the current controversy around them (p.180). This discourse is a functional strategy to argue against reputational threats (Carroll and McCombs, 2003) of bias accusations or mishandling (Gillespie, 2014:183): corporations need to seem ‘hands-off’ from any error that might occur from the use of algorithms. Moreover, it is argued that corporations often resort to ‘performed backstage’ narratives (Hilgartner 2000) ‘public-friendly’ versions of how algorithms work, in order to maintain the promise of objectivity. In sum, it is the discursive crafting of algorithms as impartial tools that legitimises them as trustworthy and helps maintain the constructed neutrality of the provider in the process (Gillespie, 2014: 179).

2.4 The Media and Framing

As seen, there is a case to argue that the dominant discourse on algorithms is the corporate discourse. Dominant discourse has been defined as the sort of language that is most relevant amongst societies and tends to mirror the ways of communication of those among whom the most power is concentrated (Foucault, 1969; Hall, 1997) which in this case, is argued to be corporations (O’Neil, 2016:12, Turow, 2011). As mentioned, for some, the issue of bias has to do with the inequalities that derive from the use of algorithms. It has been claimed that one of the best approaches to dealing with inequality issues is to study the role of reproduction and contestation of dominant discourses (Van Dijk, 1993: 249).

As seen, algorithms are discursively constructed as legitimate socio-technical tools. Algorithms have become a medium in itself (Gillespie, 2014:3, Galloway, 2004) but they also rely on ‘the media’ for legitimation in the face of society (Gillespie, 2014: 182). The literature has suggested that effects of the media coverage are linked to slight differences in how articles are presented: framing (see Entman, 1993; Iyengar, 1996; Scheufele, 1999, 2000; Callaghan, 2005; Chong, 2007b). Framing is the selection and highlighting of particular aspects of a ‘perceived reality’ to make them more prominent in a communicating text (Entman, 1993: 52; Yioutas and Segvic, 2003) and is understood to be an evolving process that fluctuates over time (Chong, 2007: 108). Framing also shapes attributions of liability and establishes solutions to issues (Cooper, 2002). Moreover, framing is claimed to be related to the reproduction of dominant perspectives (Entman, 2004: 164): if power is the ability to make others do what one wants them to do (Nagel, 1975) having the ability to ‘tell people what to think about’ is one way of exerting power and influence (Cohen, 1963: 13; Entman, 2004:165). Analysing the reproduction of dominant discourses can thus shed light upon which discourses are mirrored or contested and about whether socio-political inequalities and asymmetries are being resisted by the media (Fairclough, 1995; 2003).

In that sense, the media is claimed to be a central tool to ‘strengthen or undermine’ the efforts behind corporate discourse on algorithmic bias (Gillespie, 2014: 182): the ways in which society perceives algorithms do not depend that much upon what technical experts or the academia may say, but rather on the ‘mundane realities’ of the news cycles promoted by providers and critics of algorithms (Gillespie, 2014: 182). In general, literature claims the media can act more as a ‘lapdog’ than as a ‘watchdog’ (Bednar, 2012) for corporations, not acting as ‘major opponents’(van Dijk, 1995: 28) portraying only the goodness of algorithms to the public eye (Mansell, 2015: 121; Tufekci, 2015b: 9). However, some claim recent technology cases in the news, highlighting social concerns about the use of algorithms, might have fuelled a discussion in the media on their negative effects, ‘dispelling’ the naive-optimistic view that venerates the mathematical character of algorithms as a guarantee for objective and unbiased outcomes (Koene, 2018: 1) potentially changing the framing of technology topics. Putting the media focus on algorithmic wrongdoings, for some, implies that the media could exercise a ‘countervailing’ power against corporate discourse (O’Neil, 2016: 130).

Additionally, this research did not find a broad framing literature that applies to the case of technologies such as algorithms, nevertheless, studies on other areas of science and technology were found (see Rössler, 2001; Nisbet et al. 2003; Nisbet & Lewenstein, 2002; Gaskell et al., 2004; Scheufele & Lewenstein, 2005; Fisk, 2016). One of these studies argues that while technical media tends to highlight technological benefits and present positive framings (Scheufele & Lewenstein, 2005: 665) the case is likely to be different in mainstream outlets, as technology topics tend to be picked up later, framing controversial aspects of technologies rather than technical discourses (Nisbet et al., 2003). The result of this contrast has been characterised as a ‘war of words’, a competition of frames (Scheufele & Lewenstein, 2005: 665).

2.5 Timeline of the Study

The ‘Facebook–Cambridge Analytica’ scandal will be central to the study. This involved Facebook and Cambridge Analytica, a British political consulting company. A breech in Facebook users’ data privacy occurred when Cambridge Analytica recollected data ‘inappropriately’ through the platform and used it with commercial objectives, to influence public opinion in favour of political clients. On 17/03/2017, *The New York Times* and *The Guardian* published investigations on the episode (by Graham-Harrison & Cadwalladr and Rosenberg, Confessore & Cadwalladr). The controversy reached the level of parliamentary inquiries into the corporations involved and had worldwide repercussions in the press, which might have fuelled the discussion (Koene, 2018: 1) about the ethical and legal standards for technology companies. Since the issue of algorithmic bias involves technology companies, the occurrence of the scandal and the press attention it had might be helpful to the purpose of this work: evaluating how the media frames the corporate discourse about ‘algorithmic bias’.

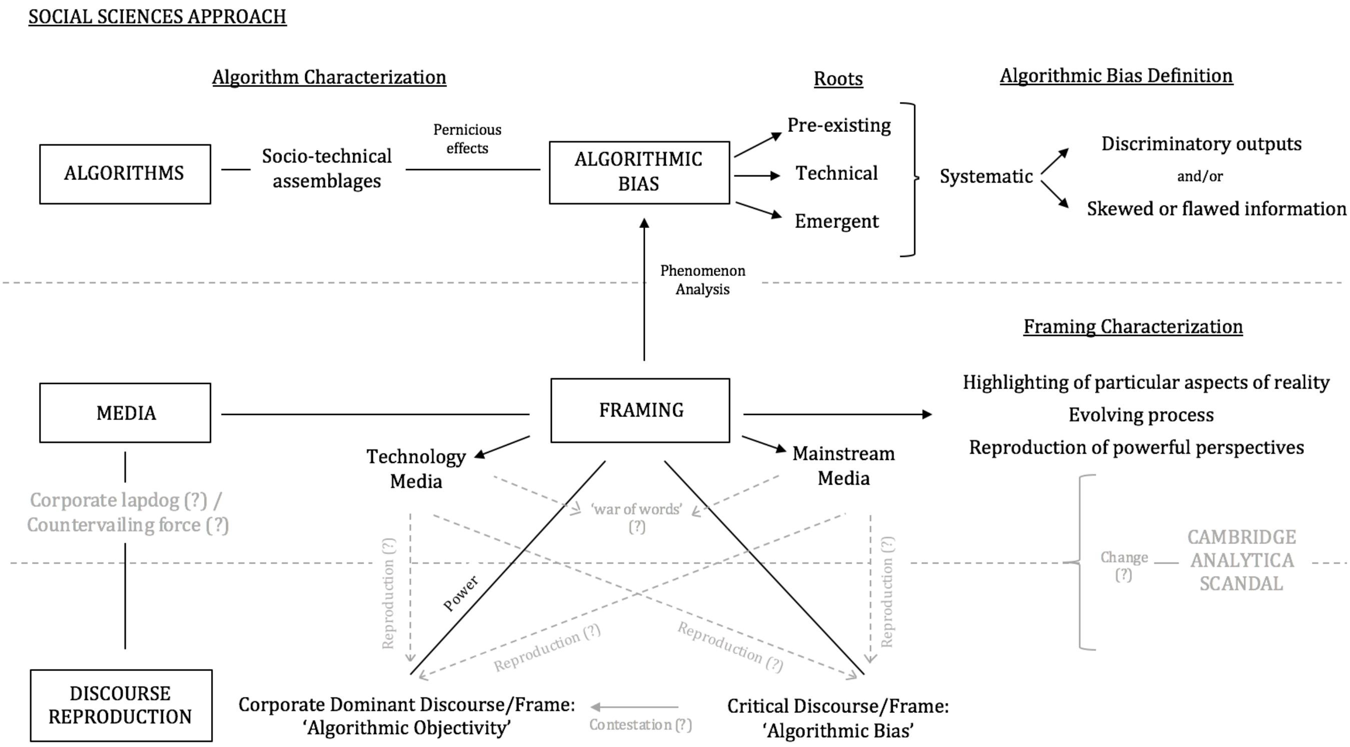
2.6 CONCEPTUAL FRAMEWORK

This work will take a social science approach as it focuses on algorithmic bias and media framing. Algorithms will be understood as ‘socio-technical assemblages’ (Kitchin, 2017: 14; Gillespie, 2014: 192; Ananny, 2015: 93): the result of the intertwining of individuals, data, hardware, regulations, laws, among many other factors (Kitchin, 2017: 20). As for the definition of ‘algorithmic bias’, neither of the options presented above seem fully suitable for encapsulating the phenomenon this work is focused on. While acknowledging there might be various divides in the algorithmic literature, such as those who characterise bias ‘pejoratively’ and ‘objectively’ and between those who focus on the need for efficiency and optimisation rather than on the analysis of power asymmetries as a way to mitigate algorithmic bias (amongst others) this project highlights one significant division between those who focus on discrimination and the 'social inequality' effects of algorithmic bias (Rubel, Castro, and Pham, 2018:9, O’Neil, 2016; etc) and others that focus on issues of 'gatekeeping' related to the information that is made available to the public through algorithms (Tufekci, 2015 ; Mansell, 2015; Morozov, 2011; van Dijck, 2009; Helberger, et al.,2015). As this study will be looking empirically at both aspects, based on the Friedman and Nissenbaum (1996) characterisation, it proposes a redefinition of ‘algorithmic bias’ as ‘*the output of an algorithm that systematically presents results that discriminate against certain groups of individuals or that systematically generates information that is flawed or tends to be skewed towards particular topics or areas of interest’.*

Also, this work will consider three main forms of bias: pre-existing, technical, and emergent (Friedman & Nissenbaum, 1996). The bias phenomenon will be analysed through media framing, which is understood as the highlighting of particular aspects of a ‘perceived reality’ (Entman, 1993: 52) a process that tends to reproduce dominant perspectives (Entman, 2007: 164) in an evolving way (Chong, 2007:108). The reproduction of dominant discourses is seen as the mirroring of the ways of communication of those with the most power in societies (Foucault, 1969; Hall, 1997). The dominant discourse for this work is the corporate discourse: the ‘promise of algorithmic objectivity’, the idea that the technical character of algorithms guarantees neutrality (Gillespie, 2014: 169). This research redefines the dominant discourse as the ‘dominant discourse of algorithmic objectivity’. Furthermore, this project contends that the study of the way media frames the algorithmic bias topic is relevant as media tends ‘not act as a major opponent’ of corporations (van Dijk, 1995: 28). Nevertheless, it is also considered that critical literature contends that the press has the ability to undermine the corporate efforts behind the construction of the dominant discourse of algorithmic objectivity (Gillespie, 2014: 182). This work defines two main framings: a ‘critical’ framing, coverage on the bias topic strongly focused on corporations mishandling of algorithms, contesting the dominant discourse, and a ‘dominant’ framing, coverage that does not critically approach the bias issue and rather mirrors and crystallises the dominant discourse of algorithmic objectivity’.

Additionally, it will consider the concept of the ‘war of words’, in which mainstream and technology media seem to associate with critical and dominant framings, respectively (Scheufele & Lewenstein, 2005). Finally, it is contended in this project that the Facebook–Cambridge Analytica scandal might have triggered in the media an overall call for more attention on the wrongdoings of corporations (Koene, 2018:1) regarding topics such as bias, potentially affecting how media frames the algorithmic topics (see Figure 1 below):

Figure 1 - Conceptual Framework



2.7 RESEARCH OBJECTIVES

The purpose of this research is to content analyse the coverage of the ‘algorithmic bias’ in the UK mainstream media and in technology publications so as to contribute to the understanding of the way media frames and reports the problem of algorithmic bias. The main research question of this work is: *‘how has the UK mainstream media and the technology publications framed the discussion on algorithmic bias in two time periods selected?’*. The objective is to see whether the media framing reproduces or critically contests the dominant discourse on the topic and, furthermore, will try to detect whether the algorithmic bias framing has varied before and after the Cambridge Analytica scandal and, in case that occurred, how and to what extent the framings have changed.

Furthermore, this research project also intends to respond to two sub-questions: *a) ‘how prominent is the corporate discourse in the coverage on algorithmic bias as compared to other discourses?’* and *b) ‘what are the most salient issues in the discussion on algorithmic bias?’*. The first question will contrast the dominant discourse described in theory against other possible discourses presented by the media (critical, governmental, technical discourses). The second question is based partially in the divide found in the literature regarding the social inequality effects of algorithms and the ‘gatekeeping’ of information. Since the work attempts to consider both empirically, it is of interest to see whether the most salient topics presented in the coverage are associated with either of these approaches, to the dominant discourse or to other discourses.

3. METHODOLOGY

The project is designed to present a comparison of the coverage of the UK mainstream media with technology publications to find a nuanced view in the latest publications regarding the topic of algorithmic bias.

3.1 Methodological Procedure

Quantitative content analysis was selected as the most suitable method for this study. This has been defined as '*a research technique for making replicable and valid inferences from texts to the contexts of their use*' (Krippendorf, 2004: 18). This project looks for the prominence or contestation of the 'dominant discourse of algorithmic objectivity' in the media. Content analysis is suitable for the pursuit of this objective, as this method's strength is to detect and quantify the presence of certain characteristics and prominent features of texts (Deacon, 1999: 116). Other methods, such as interviews, have been discarded as a way to detect the reproduction of corporate discourse as their content is unable to be analysed on a large scale, which is what this research attempts to do (Krippendorf, 2004; Hansen, 1998). Quantitative content analysis has been used for the analysis of media framing in different areas (see Appendix 5) nevertheless, considering time and space limitations, this research has not found a wide corpus of literature that applies this method to the study of framing of technologies in the media.

3.2 Types of Media Selected

This research will compare coverage on two types of media: the UK 'mainstream media' and technology publications. Mainstream media is claimed to influence large audiences, while reflecting and shaping the prevailing 'currents of thought' (Chomsky, 1997) which is aligned with the work's objective – to detect the reproduction of dominant discourses. The UK press was selected over others as, even though the Cambridge Analytica controversy had international repercussions, the company was based in the UK and the country is often considered one of the biggest technology hubs in the world. Therefore, it is expected for this specific topic to be of interest for the mainstream media of the country.

As for the technology publications, their origin is not taken as a relevant factor as the focus on technology is a given. Technology publications are niche, which implies that mainly audiences with specific interests consume them: since the academic literature often refers to an interest in algorithms in the corporate sector, corporations are seen as suitable audience for technology publications. Therefore, it is believed there is a case for arguing that it is highly likely that technology magazines will cover an emerging issue like algorithmic bias and while doing so, represent the corporate discourse in a greater extent than the mainstream media does.

3.3 Outlets Selection

Since the objective is to analyse two kinds of publications, a variety of outlets were selected to gather a significant and equative sample: 13 publications (7 mainstream, 6 technology). To facilitate the sampling process and accessibility, only online publications were considered. The specific outlets selected are as follows: For mainstream media, the criteria was based on the top seven most used online news brands in the country by the percentage of weekly usage: *BBC News, The Guardian, The Daily Mail, The Huffington Post, Sky News, MSN News,* and *The Daily Telegraph* (Reuters, 2017). As for technology, the criteria used was based on a 2017 selection of the top ten most respectable online technology publications made by a technology expert, according to one of the selected mainstream media (Solomon, September 2017). Six publications were selected randomly from the list, in no particular order of relevance: *MIT Technology Review, Wired, Fast Company, Gizmodo, TechCrunch and Recode* (see Appendix 5).

3.4 Time Period

The period analysed is before and after the Facebook–Cambridge Analytica scandal, which broke on March 17th, 2018. The study considers two and a half months of coverage before and after the scandal: January 1st–June 1st, 2018. The length period has to do with the recent character of the events in relation to the development of this research and with the desire to study periods of equal length. The reason why the periods of time before and after this story broke are relevant to this project is that, as mentioned, framing is considered to be an evolving process (Chong, 2007: 108) and it is contended that the scandal might have affected the ways (Koene, 2018: 1) in which different sorts of media frame the issues around the algorithmic bias topic.

3.5 Selection of Articles

This sample is composed of 187 articles (n=187). The decision to analyse full articles, titles included, is linked to claims in the theory that argue that framing is concerned with the general character of the texts (Entman, 1993) rather than with individual words or paragraphs. Other framing resources, such as the use of photography, illustration or positioning are not considered.

It is important to note that the total sample number is the sum of the total results obtained from each site search: the sample includes all data pulled from the selected media in the time frame indicated showed on the day the search was performed (7/6/2018). Since the pool of data was quite limited, there was no need for systematic sampling techniques (Krippendorff 2004; Neuendorf 2002).

The study relied on Google Search as it provided a bigger data universe than other sources, such as LexisNexis. The researcher searched the terms ‘bias + algorithm’, ‘bias + artificial intelligence’ and ‘bias + machine learning’ (all usually used as equivalent to algorithms by the media). Further details on the selection of articles can be found in Appendix 5.

3.6 Coding

The codebook is the way in which this work has chosen to operationalise the research question and sub questions, in order to answer them and, furthermore, to detect further possible topics of interest. While designing the code, there was an intention to balance the presence of items related to both the critical literature and dominant discourse. The code book is composed of a series of codes (Hansen, 1998: 116) designed in a way that allows one to content analyse the articles and establish relationships between the codes (or variables).

The cluster ‘Explanatory Variables’, classifies the articles according to categories that are not to be affected by any other variables. The cluster ‘Discourses’ was designed in order to answer sub-question A, it contains one of the key codes, #10 (‘posture’) which will help categorise articles according to the frames and contains other variables that operationalise and contrast different discourses. The cluster ‘Main Topics’ helps to answer sub-question B, presenting different topics associated with critical and dominant frames, as well as the Cambridge Analytica scandal. The cluster ‘Regulation’ intends to detect whether regulation topics are relevant in the coverage, which will help answer some hypotheses that will be proposed later on in the work. Please see Appendix 3 for an extended coding book offering the full description and the rationale behind each variable and Appendix 5 for piloting procedures.

3.7 Inter-coder Reliability

Inter-coder reliability test reports reliability by variable (Riffe and Freitag, 1997) and ensures the repeatability and quality of the code book (Deacon, 1999: 128). The test implies the replication of the reading and coding by a second coder not involved in the original design of the study (Neuendorf, 2002:9). The tests were conducted with the online tool ReCal2 for percent agreement as an indicator for reliability (Macnamara, 2005: 11).

The ICR test was conducted for each of the codes and for each of the categories that were coded as dummies (excluding explanatory variables) presenting an overall percent agreement of 94%, while showing wide range of percent agreements ranging from 68% to 100%. After conducting the full test on the 19 articles, there was an overall disagreement on two codes (‘posture’ presenting a 74% and ‘polemical’, 68%)’ and two dummy categories inside one code (‘d_MultiReg’ and ‘e_Mixed’, corresponding to code 32 ‘regulation’, presenting both 74% as well). After discussion, the code ‘polemical’ was discarded from the codebook as it was argued that some other variables could respond to the same objective of answering the hypothesis regarding the ‘level’ of coverage. As for the codes ‘posture’ and ‘regulation’, the disagreements found were considered not to be of crucial relevance, as it was detected that they had to do with blurriness in similar categories like ‘positive’ and ‘very positive’ and ‘mixed’ and ‘multi stakeholder’ regulation, categories that could easily have been considered part of the same broader categories. Please see Appendix 5 for details on the development of the final version of the code and ICR tests and pilots.

3.8 Ethics and Reflexivity

Ethical critiques are not often found for quantitative methods, given that the information used is publicly available (Neuendorf, 2016:130). Nevertheless, this research does acknowledge a few limitations that might affect this study. The first is the requirement for objectivity (Berger & Luckman, 1966; Hansen et al.,1998 :94). This challenge has been quite present in this research: not only the reading of the articles is open to varied interpretations; regardless of ICR levels, the overall design of the codebook might be marked by the inherent ‘pre-existing’ biases of the research designer, as the coding of frames relies heavily how the researcher perceives the topics (Matthes & Kohring, 2008: 262). Even though the code was thoroughly revised to be open to balanced results, it is possible that the code itself might be biased or that the coder is biased in the interpretation of results. In order to mitigate the latter, all results were analysed by a second coder independently to check for coherence; nevertheless, defining which results are ‘significant’ and which are not in light of the literature is also quite subjective.

Moreover, it has been claimed that because the method is ‘intensive and time-consuming’ for coders and co-coders, many studies often involve relatively ‘small samples’, such as the one that this study presents (n=187) which have been criticised by researchers as ‘unscientific and unreliable’ (Macnamara, 2005: 5). Nevertheless, it must be pointed out that the small sample number of this project has to do with the lack of available information on the topic rather than with an intention to diminish the workload. The selection of keywords used intends to mitigate this as much as possible, while at the same time trying not to overly ‘stretch’ the search and lose focus of the main topic.

Finally, it could be claimed there is a paradox: while this study is concerned about the presence of bias in algorithms, it has used a technology accused of bias to obtain the sample, namely, Google Search. Personalisation algorithms such as the one Google presents affect one’s experience based on one’s previous interactions with the platform. Therefore, it is argued that it is very difficult, if not impossible, for a lone researcher to abandon the ‘user’ position and get a neutral perspective on the topics she/he intends to analyse (Seaver, 2013: 5). If this research was to be developed further, the addition of print media and other sources for data might be useful for the mitigation of this possible issue.

4. RESULTS

This section presents the content analysis results for the 187 articles. The database was analysed in Excel and SPSS for frequencies, percentages, correlations and regressions. The raw coding data can be found in Appendix 6 and 7, the summary analysis of results in Appendix 3 and details on SPSS procedures in Appendix 4. Because of extension limitations, only the most significant results will be exposed. Here, the sub-questions will be quantitatively assessed to later respond to the general research question. In the process of responding to these questions, other hypotheses will follow, resulting from the combination of intuition and deduction based on literature arguments or because of interesting data results.

Sub-question a) ‘how prominent is corporate discourse in the coverage on algorithmic bias as compared to other discourses?’

The code ‘Posture’ helped categorise the articles according to the two main framings, from ‘Very Negative’ (resonating critical frames) to ‘Very Positive’ (resonating corporate discourse). Analysing the results from the frequency table below, we see that critical frames were slightly predominant in both media. Corporate frames were not. In both periods, the results show that most of the sample was framed as ‘Negative’. 53% of the articles are overall negative (99/187) (very negative + negative) while only 31% are positive (58/187) and the rest neutral. Moreover, technology publications present an overall higher presence of dominant frames: 37% (41/111) vs. mainstream 22% (17/76).

Table 1 - Posture - Frequency by type of media and before and after

	Mainstream		Technology		Grand total (both media)	Grand total % (both media)
	Before	After	Before	After		
Very Negative	9	16	22	17	64	34%
Negative	6	9	10	10	35	19%
Neutral	2	17	8	3	30	16%
Positive	4	4	7	10	25	13%
Very Positive	5	4	14	10	33	18%
Grand total of articles	26	50	61	50	187	100%
	76		111			

Nevertheless, in mainstream media, the percentage of negative frames in relation to the total articles published in each period decreased – from 58% (15/26) to 50% (25/50) in favour of more neutral postures. For technology, negative framings increased slightly: 52% (32/61) to 54% (27/50). After the scandal, the case is different for positive frames: for mainstream, the overall positive framings decreased from 35% (9/26) to 16% (8/50) and for technology they increased from 34% (21/6 to 40% (20/50).

To conduct regression analyses to check on the significance of the change in framings, two additional variables were created: ‘CriticalFrames’ and ‘DominantFrames’, conglomerating negative (negative + very negative) and positive (positive + very positive) frames. The code ‘After_CA’ was the independent variable. Further hypotheses follow from the assessment of the results above, originating from a reading of the literature that suggests that the polemical character of the scandal might have affected the way media frames the topic:

a) Critical frames increased after the occurrence of the scandal

b) Dominant frames decreased after the occurrence of the scandal

For Hypothesis a) changes in both types of media turned out to be non-significant (p-values>0.10): the analyses did not find evidence to suggest that after the scandal there was a significantly higher presence of critical framings in the sample for either type of media. As for Hypothesis b) the changes are significant (p-value<0.10) only for mainstream media: evidence was found to suggest that articles published after the scandal are associated with a reduction of 19 percentage points of dominant framings in the sample. This is not the case for technology outlets (p-value>0.10) as seen, for tech dominant framings actually increased after, but not significantly.

Additionally, the codes ‘Corpo_Quote’, ‘GovPar_quote’ and ‘Sources’ show the percentage of repetitions of each sort of discourse: corporate, governmental and technical. The comparison of percentages presented in the crosstab below indicate that corporate discourse is considerably more repeated than others. These tendencies are quite similar for both media. For mainstream, after the scandal, we see a considerable increase of governmental discourse against a decrease in technical discourse. After conducting a regression analysis, this result proved to be significant: there is evidence to suggest that the amount of mainstream articles that quote government discourse increased after the scandal in 30 percentage points.

<u>Table 2 - Comparison of discourses - Before and after by media</u>

	Corpo_Quote		GovPar_Quote		Sources (tech)	
	Before	After	Before	After	Before	After
Mainstream	41%	56%	11%	38%	56%	26%
Technology	52%	54%	15%	22%	N/A	N/A

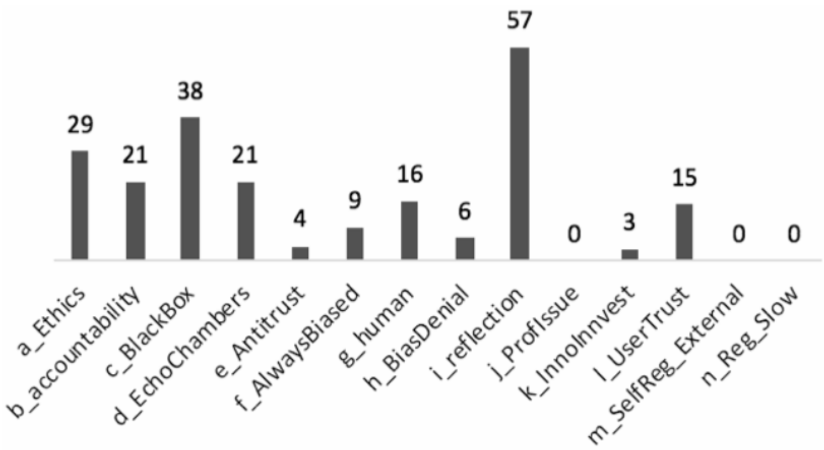
<u>*Sub-question b) ‘what are the most salient issues in the discussion on algorithmic bias*</u>*?’*

In the majority of the sample, the relevance given to the algorithmic bias topic was marginal: less than half of the articles, 41.7% (78/187) presented the issue as main topic (Code 28: ‘Main_Topic’). This remained fairly constant between periods, before: 50% (39/78) vs. after: 50% (39/78) and across media, mainstream: 39% (30/77) vs tech: 43.6% (48/110). Additionally, it is worth mentioning that, in the second period, the percentages of articles that made references to the topic of Cambridge Analytica scandal indicate fair to little relevance given to the topic in both media – mainstream: 30% (15/50) vs. 20% technology (10/50) – (Code 27 ‘CA_mention’).

Furthermore, overall the most salient issues (present in more than 10% of the sample) are associated with critical discourse and were highlighted more in technology publications (Code 23: ‘Issues’) and overall did not vary significantly before and after the Cambridge Analytica scandal (with the exception of the effects on political life, which increased significantly (from 23 to 32 mentions after the scandal, see Figure 18 Appendix 3).

In relation to Code 23: ‘Issues’ the main topics were ‘the reflection pre-existing biases’, the issue of ‘black boxes’, the lack of corporate ethics and accountability and algorithms as ‘echo-chambers’ (see Figure 1).

Figure 1 - Issues - Overall count of mentions



As for Code 25: ‘Suggestions’, the sample also emphasised the need for technology corporations to have more transparency and diversity (see Figure 21, Appendix 3) and Code 24: ‘Effects’ indicated interest in the following social consequences: ethnic discrimination, negative effects on political life, gender discrimination, the dissemination of ‘fake news’, discrimination in personnel selection and discriminatory police practices.

Also, according to Code 31: ‘Profit’, it might be worth noticing that no significant importance was given to the topic of biased algorithms being functional to profit for corporations: only 8.6% of the overall sample presented this argument and it was more emphasised in technology publications, 10.9% (12/110) vs. mainstream: 5.2% (4/77).

Therefore, based on the critical literature arguing for external regulation to mitigate the effects of algorithms (such as the ones listed above) an additional hypothesis arises:

c) Critical framings of algorithmic biases are associated with external regulation narratives

From all the articles that acknowledged the issue of algorithmic bias (excluding ‘Very Positive’ frames, which amounted to 33 articles) the majority of the sample, 58% (89/154) did not make references towards any kind of external regulation (‘f_NoMention’). 38% (59/154) made reference to external regulation (Codes: ‘a_ProReg’, ‘d_MultiReg’, ‘e_Mixed’) and 4% (6/154) self-regulation (‘c_SelfReg’). Regarding critical frames only, 45% (45/99) mentioned external regulation (mainstream 43%, (17/40) and tech 47%, (28/59). Only two articles showed explicit anti-regulation mentions (‘b_AntiReg’) but they were very positive. Nevertheless, there is one key finding about regulation: from frequency tables we see that the percentage of articles that do mention external regulation increased considerably after the scandal: from 31% (21/68) to 44% (38/86) (excluding Very Positive frames).

<u>Table 3 - Regulation - Overall percentages of mentions – Before and After</u>

	Articles Published	Mentions of External Regulation	% of mentions in relation to articles published
Before	87	21	24%
After	100	41	41%

Given the later finding, it might be worth testing the following hypothesis for each type of media:

d) The amount of articles mentioning external regulation increased after the scandal

For easier analysis on SPSS, a new variable was created: ‘RegulationTalk’, grouping the 38% of the sample that discusses external regulation (‘a_ProReg’, ‘d_MultiReg’ and ‘e_Mixed’). Regarding Hypothesis d) the regression between ‘RegulationTalk’ and ‘After_CA’ showed significant results for technology media (p-value<0.10). There is a positive association between articles published after the scandal and articles that discuss external regulation: data suggests that in technology publications, the number of articles discussing regulation increased after the scandal by 19 percentage points. This is not the case for mainstream media (p-value>0.10). As for Hypothesis c) the correlation between variables for both media is significant (p-value<0.1) and it is positive: in the sample there is evidence to suggest that in both media critical framings of algorithmic biases are associated with external regulation narratives. Both correlations are significant at a 90% confidence interval, but for mainstream media is significant at 95% and for technology at 99%.

At this point, a final hypothesis follows based on the literature that argues that mainstream coverage tends to be sensationalistic:

e) Different kinds of mainstream media present different ‘levels’ of coverage of the algorithmic bias topic

Table 4 - Count of Posture by mainstream media outlet

Outlet	Very Negative	Negative	Neutral	Positive	Very Positive	Grand Total
BBC News	2	1	6	1	0	10
Daily Mail	6	5	8	1	3	23
Huffington Post	2	0	0	1	1	4
MSN News	4	1	0	3	1	9
Sky News	0	1	1	0	0	2
The Guardian	11	6	4	0	2	23
The Telegraph	0	1	0	2	3	6
Grand Total	25	15	19	8	10	77

Codes 30, 31 and 32 help to answer this question. Code 30, ‘sources’, is considered as an indicator of the ‘level’ (quality) of coverage. From table below, we can see that the ‘Daily Mail’ presented the highest percentage of reliance on computer/technical experts to address the issue of bias, 57%. Additionally, it presented the highest frequency of technical explanations for the issue of bias (‘technical’) the greatest number of articles included in the ‘Technology’ section (‘TechSection’) and neutral to positive postures. This makes *The Daily Mail* the mainstream media with the highest level of coverage in the technical dimension and the highest neutral to positive posture.

Table 5 - Sources, technical and TechSection - Frequencies and percentages by media

	Articles count	Sources count	% sources	technical count	% technical	TechSection count	% TechSection
BBC News	10	3	30%	2	20%	5	50%
Daily Mail	23	13	57%	12	52%	17	74%
Huffington Post	4	0	0%	0	0%	1	25%
MSN News	9	3	33%	2	22%	0	0%
Sky News	2	0	0%	0	0%	2	100%
The Guardian	23	7	30%	4	17%	3	13%
The Telegraph	5	2	40%	0	0%	1	20%
Grand Total	76	28	37%	20	26%	29	38%

Also, the discussion of external regulation is understood as an indicator of a deep-level conversation on the topic. The media showing the highest levels of ‘RegulationTalk’ (considering the total number of articles) is *The Guardian*, which presents also the highest frequency of negative frames.

Table 6 - RegulationTalk - Frequencies and Percentages by Media

	Mentions of External Regulation (RegulationTalk)	Count of Articles	% Mentions of Regulation
BBC News	4	10	40%
Daily Mail	4	23	17%
Huffington Post	2	4	50%
MSN News	3	9	33%
Sky News	1	2	50%
The Guardian	8	23	35%
The Telegraph	1	6	17%
Grand Total	23	77	30%

Based on this information, different kinds of mainstream media do present different ‘levels’ of coverage of the topic. According to the two different dimensions (technical and regulation) *The Daily Mail* and *The Guardian* presented the highest levels.

The information presented above together with additional results below will help answer the RQ: *‘how has the UK mainstream media and the technology publications framed the discussion on algorithmic bias in two time periods selected?’*

We can see that, overall, after the scandal, there was an increase in coverage on the algorithmic bias topic in mainstream media (due to a considerable number of articles being published after the scandal by the mainstream media, see Table 7).

Table 7 - Coverage of the issue - Before and After and by type of media

	Before	After	Total count by media type	Total % by media type
Mainstream	26	50	76	41%
Technology	61	50	111	59%
Total count by before and after	87	100	Total sample: 187	
Total % before and after	47%	53%		

Furthermore, a correlation test proved that there is a strong positive correlation (p-value<0.1) between the articles that were published after the event and references to Cambridge Analytica for both media. Also, even though technology publications tended to cover the bias issue more than the mainstream media, for tech, the number of articles covering the topic actually decreased after the scandal, from 55% (61/111) to 45% (50/111). As seen, the mentions of external regulation increased considerably afterward for technology media and, furthermore, a significant positive correlation (p-value<0.1) was found between critical framings and the mention of Cambridge Analytica: critical tech articles were more likely to mention the scandal.

5. DISCUSSION

This section will address the main findings from the results section in relation to the overall research question and literature. There will be an interpretation of data with the aim of understanding what it they imply in terms of the reproduction of the dominant discourse on the topic and will try to detect whether the algorithmic bias framing has varied before and after the scandal. The following key findings will be interpreted:

In relation so sub-question a) critical frames were slightly predominant in both media. Corporate/Dominant frames were not. This does not support the theory that media is a ‘lapdog’ of corporations (Bednar, 2012) and their discourse. Nevertheless, the predominance of the critical coverage was not overwhelming: only 53% of the total sample. Furthermore, the results are in line with the literature on framings of technology (Scheufele & Lewenstein, 2005): technology publications present an overall higher presence of corporate/dominant frames. Moreover, after the scandal, there was an increase in coverage on the algorithmic bias topic, coming from mainstream media, equating the volume of articles produced by technology publications in the second period (50 articles in each). This evidence supports the theory that technology topics tend to be picked up later by mainstream media (Nisbet et al., 2003) and that the occurrence of scandals such as Cambridge Analytica might have fuelled the discussion on the bias topic for the mainstream case (Koene, 2018: 1).

Nevertheless, overall, there was minimal change in framings between the two time periods: no evidence was found to suggest that after the scandal there was a significantly higher presence of critical framings in the sample for either type of media. This does not go in line with the supposition that the Cambridge Analytica scandal might have had an impact on the framing of algorithmic bias. Still, there were some small changes: only for mainstream, articles published after the scandal were associated with a reduction of corporate/dominant framings and for technology, corporate/dominant framings actually increased after, but not significantly. These slightly contradictory trends in the data could be understood under the light of the ‘war of words’ theory (Scheufele & Lewenstein, 2005: 665): given the considerable increase of critical coverage on the topic after the scandal in mainstream media, technology publications ramped up: to a small extent, the production of news articles portrayed the benefits of algorithmic technologies. This is confirmation is reinforced by the finding that, after the event, only critical tech articles were more likely to mention the Cambridge Analytica scandal. Still, the topic of Cambridge Analytica was given fair to little relevance in both media: 30% of the articles mentioned the issue in mainstream media and 20% in the technology media. This finding might not be striking for tech, but for mainstream media it might imply that technological coverage is not mainly controversial, as the literature suggests (Scheufele & Lewenstein, 2005).

In relation to Sub-question b) the most salient issues are associated with critical discourse, were highlighted more in technology publications and did not vary significantly after the Cambridge Analytica scandal. Mentions of issues present in the dominant discourse (such as the need for innovation, the importance of user trust and self-regulation and the threat of external regulation) presented low to non-existing mentions in the sample. The fact that these topics were emphasised more in technology publications is in line with the theory that claims that mainstream media tends to cover topics more 'controversially'. In this context, the redefinition of algorithmic bias that this work proposes proves to be useful, as it encapsulates both kinds of phenomena. It might be worth noticing that these topics below will not be developed to their full extent in this section, rather, they will be introduced in order to illustrate the themes that characterise the media coverage on the algorithmic bias topic.

The reflection of pre-existing biases was by far the issue most mentioned in relation to algorithmic bias. This empirical evidence sustains the claim that the early work published on biases in computer systems is still in force (Friedman & Nissenbaum, 1996). Depending on how the data fed into the system is 'curated', this might lead to a lock-in of existing asymmetries and discriminatory outputs. Moreover, preconceptions might also affect the design of the algorithms themselves, leading to a reproduction of skewed information, which is related to the theories of 'gatekeeping'.

The issue of 'black-boxes' is related the literature about the lack visibility over the working of algorithms (Diakopoulos, 2013: 2) and their complexity (Tufekci, 2005: 206). The term refers to a 'system whose workings are mysterious; we can observe its inputs and outputs, but we cannot tell how one becomes the other' (Pasquale, 2015: 3). It could be said there are two dimensions to the 'black-box issue': one is technical and refers to how neural networks generate outputs in ways not even developers fully comprehend (Shwartz-Ziv & Tishby, 2017). The other is social: algorithms are claimed to obscurely embed power relationships (Mansell, 2012: 186, 2016: 4) which is linked to theories that raise concerns regarding the sort of values algorithms prioritise (Nissenbaum, 2001; Gillespie, 2014; Napoli, 2014, O’Neil, 2016). A fragment from tech media illustrates the issue:

> *‘Flaws in most AI systems aren’t easy to fix, in large part because they are black boxes: the data goes in and the answer comes out without any explanation for the decision. Compounding the issue is that the most advanced systems are jealously guarded by the firms that create them’ (Hill, 2018)*

An additional highly mentioned topic was the issue of ‘echo-chambers’. A wide range of literature focuses on how algorithms, especially in blogs, social media and search platforms, generate echo-chambers and filter bubbles and the effects that this might have on the public (Pariser, 2011): based on user consumption, algorithms can amplify these preferences, delivering content that is always ‘more of the same’ (Rohgalf, 2017: 90). Critical literature argues that this is concerning, as it might create clusters of ‘like-minded’ users, which ultimately might increase the risks of polarisation and radicalisation amongst users, affecting also the space for public deliberation (Chadwick, 2013).

The dissemination of ‘fake news’ and negative effects on political life are in some ways related to the issue above. As seen, the filtering of information can affect the space for public deliberation, and therefore, political life. Moreover, in recent cases like the alleged international meddling in the 2016 elections in the US and UK, the dissemination of fake news (Alcott & Gentzkow, 2017; Guess et al., 2018) and fake political advertising (Wood, et al., 2018) were claimed to have been amplified by the echo-chamber effects social media algorithms produce. Again, the early literature on computer systems proves to be useful: these biases could be characterised as ‘emergent’ (Friedman & Nissenbaum, 1996: 336) resulting from negative changes in the original purpose of use of social platforms. The issue was well-emphasised by the media in texts such as this:

> *‘Filter bubbles (...) have strained the social fabric and polarized the political realm. The fallout has been so intense that Mark Zuckerberg recently went against the wishes of Facebook’s investors, changing their algorithms to facilitate “deeper, more meaningful exchange”. He even*

> *apologized "for the ways my work was used to divide people rather than bring us together"'.*
> *(Byrne, 2018)*

As for discriminatory consequences, the algorithmic bias coverage highlighted a series of arenas of concern which are also present in the critical literature. First, algorithms have been denounced for reinforcing ethnic prejudices in decision-making (Nakamura, 2009: 158) specifically, the use of ethnically biased algorithms in law enforcement (Introna & Wood, 2004; Joh, 2017) judicial decisions (Angwin et al., 2016; Flores et al.,2016; Butler, 2018) and search platforms (van Couvering, 2010; Noble, 2018) and face recognition software (Howard & Borenstein, 2017). Headlines below help illustrate the coverage on the topic:

> *'Photo algorithms ID white men fine – black women, not so much' (Simonite, 2018);*

> *'Google Censors Gorillas Rather Than Risk Them Being Mislabelled as Black People – But Who Does That Help?' (Fussel, 2018);*

> *'Concerns in the courtroom as algorithms are used to help judges rule on jail time which can flag black people as twice as likely to re-offend compared to white defendants' (Associated Press, 2018)*

Gender discrimination is also a prominent topic (Caliskan et al., 2017) it affects search engines (Otterbaher et al., 2017, 2018) face and voice recognition (Yapo & Weiss, 2018) and virtual assistance (Payne et al, 2013) amongst other areas. Below some headlines from the media:

> *'Women must act now, or male-designed robots will take over our lives' (Graham, 2018);*

> *'Are Alexa and Siri sexist? Expert says AI assistants struggle to understand female speech because it is quieter and more "breathy"' (Pinkstone, 2018);*

> *'Bank of America's bot is "Erica" because apparently all digital assistants are women' (Guthrie-Weissman, 2018).*

The use of biased algorithms by in hiring processes and the workplace (Kim, 2017; O'Neil, 2016; Marshall, 2018) has also been made prominent by the media:

> *'How to persuade a robot that you should get the job' (Buranyi, 2018);*

> *'Are computers turning into bigots? Machines that learn to think for themselves are revolutionising the workplace. But there's growing evidence they are acquiring racial and class prejudices too' (Naish, 2018);*

> *'Companies using controversial AI software to help hire or fire employees - and gauge whether they like their boss' (Palmer, 2018).*

Finally, the use of biased algorithms in police forces (Joh, 2017, 2018) was also a quite dominant topic:

> *'UK accused of flouting human rights in "racialised" war on gangs' (Dodd, 2018);*

> *'Facial recognition AI built into police body cameras will suffer from racial bias and could lead to false arrests, civil liberties experts warn' (Pinkstone, 2018);*

> *'Police gang database breaches human rights and discriminates against young black men who are fans of grime music' (Bartlett, 2018).*

The biases presented above are seen products of algorithmic wrongdoings, the literature leads us to argue that the origins of these biases are often associated with broader issues that affect the development of algorithms. The data supports this theory: some of the most mentioned topics are the need for algorithmic ethics (Ananny, 2016) which encapsulate the highly mentioned issues of transparency and diversity, and algorithmic accountability. Ethics has been defined as 'the study of what we ought to do' (Merrill, 2011:3). In relation to transparency, the media tends to claim there is ought to be more visibility, which is in line with this quote from the sample:

> *'We (...) need to start querying the outcomes of the decisions made by algorithms and demand transparency in the process that leads to them (...) pressure must be applied to ensure AI algorithms are not just powerful and scalable but also transparent to inspection' (Bartoletti, 2018)*

Nevertheless, key literature on the topic argues that ethics should go beyond the mere examination of codes and search for a holistic assessment of algorithms as socio-technical assemblages (Ananny, 2016: 102,109) there is an argument for the ‘disentanglement’ of such conglomerates (Diakopoulos & Koliska, 2016:882). Indeed, it is claimed that the solution to ethics issues cannot be resolved ‘technically’, as it implies the questioning of values (Danks & London, 2017: 4692). This can be related to the academic claims that the issue of ‘black-boxes’ is not just technical, as it involves the ways in which preferences and power asymmetries might be hidden in the assemblage. Furthermore, diversity issues were also mentioned by the coverage in the following ways:

> *‘Many AI problems, like bias, stem from the fact that such a narrow group of people are building the technology’ (Snow, 2018)*

> *‘Having a diverse tech workforce is crucial for the fair development of technologies like AI – to help prevent bias and ensure it’s solving the problems of the wider society, not just the issues of a single (white, male) group’ (Winick, 2018)*

Even though some academic articles argue for the importance of diversity for the mitigation of bias (Kotsiantis, 2007) literature research suggests the academia does not seem to focus on this topic as much as it does on the need for transparency. The mention of these topics, nevertheless, is linked to another concern, algorithmic accountability: who can or should be held responsible for changes (or effects) that these technologies imply? (Mansell, 2017:46). Significant mentions of this topic in the coverage seem to mirror the interest present in the literature on the topic (Diakopoulos, 2013, 2015, 2016). As one of the articles puts it:

> *‘Taking responsibility for handling AI can’t, and won’t, happen automatically’ (Terdiman, 2018)*

Given all the biases and defects that were presented in relation to algorithmic use, the critical literature often and strongly argues in favour of measurements to tackle such issues. The sample and the literature suggest there is a need to make the outputs of algorithms more fair and reliable. Critical strands of social sciences tend to suggest the way to do so is through *regulation*. Nonetheless, the tech industry and regulation are often portrayed as enemies: regulation is seen as an inhibitor of progress (Wiener, 2004: 483). Regulation, though, is claimed to have many forms: in one end of the regulation spectrum, we can find ‘self-regulation’, which could be defined as corporate attempts to rule and constrain behaviour to avoid direct state intervention or intervention by any other actors (Cohen & Sundararajan, 2015: 869) which this study associates with the dominant corporate discourse. On the other end, we find ‘external regulation’, which might encompass state or other sorts of external regulation by third parties, including governments. The critical literature argues that if technologies are presenting pervasive effects, there is a case for arguing in favour of external forms of regulation (Mansell, 2014; Goodman & Flaxman, 2016: 1). The evidence in the data sustains that: in both media critical framings of algorithmic biases are associated with external regulation narratives, that in technology publications, the number of articles discussing regulation increased after the scandal and that the amount of mainstream articles that quote government discourse increased significantly after. However, a central finding leads us to a relativisation of the ‘criticalness’ of the negative frames in the coverage: the majority of the sample, in both media, did not make references to any kind of external regulation. As seen, the literature establishes that the dominant corporate discourse rejects external regulation. Given the evidence that the issue of bias is acknowledged by the coverage but external regulation is not mentioned, it could be claimed that the level of criticism present is quite tepid, not firmly defying the dominant corporate discourse. An additional finding sustains the argument above: no significant importance was given to how biased algorithms are claimed to be functional to corporate profit.

Finally, the sample suggests different kinds of mainstream media do present different ‘levels’ of coverage of the topic: *The Daily Mail* and *The Guardian* presented the highest levels (technical and regulation, respectively). As seen, academic works on the framing of technologies in the media argue that, in mass media, controversial aspects of technologies tend to be more stressed than technical discourse (Nisbet et al., 2003). Of all articles sampled, *The Daily Mail* was an outlier, presenting significantly high levels of coverage in the technical dimension accompanied by high neutral to positive postures. This is in accordance with information in Appendix 5: this newspaper presents the higher conservative bias in the sample, which could be claimed to be associated with a non-contestation to the dominant discourse. Additionally, *The Guardian*, which is characterised as liberal and left-biased, presented the highest levels of mention for corporate regulation. The rest of the evidence supports the theory that ‘low level’ technical coverage tends to be prioritised in mass media. This is supported by the finding that overall, none of the media in the sample presented significantly high coverage discussing regulation, which was considered as an indicator for coverage level. Furthermore, the fact that the amount of mainstream articles published on the topic increased considerably after the event (from 26 to 50) while technical outlets’ coverage remained quite constant, is in accordance with the claim that mass media does tend to pick up on technology topics with some delay (Scheufele & Lewenstein, 2005).

The discussion above leads us to the answering of the research question:

How has the UK mainstream media and the technology publications framed the discussion on algorithmic bias in the two time periods selected?

Overall, while both media presented more critical than dominant frames, evidence suggests that mainstream media contests dominant discourses to a greater extent than technology publications. This is in line with the technology framing literature. Additionally, according to the conceptual framework that informed this empirical study, a fundamental indicator for contestation was not present in either case: a solid call for external regulation, which leads us to qualify the ‘strength’ of the critique as ‘mild’. The work attempted to detect, furthermore, whether the algorithmic bias framing varied before and after the Cambridge Analytica scandal. The study found that even though there was an increase in coverage on the topic after the scandal, generated by mainstream media, the framing did not fluctuate abruptly in favour of more critical frames (or vice versa). This evidence contrasts with theories that argue that framing evolves over time (Chong, 2007: 108) but is, nevertheless, in line with the arguments of the framing of technology literature: even though the scandal did seem to catch the late attention of mainstream media, it did not increase deep-level critical discussions on the issue of bias. Moreover, given the results, the arguments of the critical algorithmic bias literature about the possibility of media to act as a ‘countervailing’ force or as a ‘watchdog’ for corporations are difficult to affirm solidly: even though most of the coverage was indeed critical, it did not prove to be critical ‘enough’. This, however, is in line with the literature that suggest that media, in fact, do not tend to act as major opponents of corporations, ‘major’ being the key word.

6. FURTHER RESEARCH

Since the topic of framing is related to the formation of public opinion (see Cobb, 2005; Lecheler & de Vreese, 2012; Druckman, 2001) further research on the impact of the framing of technologies such as algorithms might be useful to the assessment of the debate over the regulation of technology corporations. Interviewing methodologies with a focus on audiences might be functional to the evaluation of the effects of framing over the formation of public opinion. Additionally, it would be interesting to see how this public opinion on the topic might be intertwined with decision-making processes at a corporate and governmental level.

Furthermore, it might be interesting to study exactly how these discourses get translated into frames in the media agenda and the role stakeholders might have in influencing these processes. In this case, the interviewing of journalists and stakeholders from technology and academia might help illustrate the case, together with a content analysis on the coverage, focusing on the production of the news, analysing both patterns in sources and journalists. Furthermore, it would be interesting to contrast experts’ perspectives with the media coverage.

Finally, as the production and use of algorithms expand, it might be worth putting away the focus from the corporate production of algorithms to examine how governments might be getting involved in the development of such technologies and to see how the discourses that are presented for the case examined in this research might be adapted to fit state narratives, and how the principles that concern the external regulation of technology corporations (ethics, accountability) apply for the case of the state production of algorithms.

7. CONCLUSION

Overall, the focus on the framing of the discourse about algorithmic bias allowed for the detection of what could be claimed to be three main interesting findings. The first finding is that the critical discourse and issues associated with the critical narratives dominated the coverage in the sample, rather than corporate or other discourses. The second finding is that the robustness of this critical discourse was quite low if considering other factors like references to external regulation and the fact that most of the mainstream coverage on the topic happened after the Cambridge Analytica scandal. And, third, as seen, after the scandal, critical framings did not increase for either media and dominant frames presented minimal variations (decreasing slightly in mainstream and increasing non-significantly in tech) which leads us to the one fundamental finding: that the framing of the articles did not present abrupt changes after the occurrence of the scandal.

Given these results, it could be claimed that the literature and method used to develop this empirical research proved to be useful: quantitative content analysis provided rich results and all the empirical evidence was able to be meaningfully assessed under the light of the theoretical framework. Regarding the role of media in the reproduction of the dominant discourse, it could be said that a combination of theories helped to illustrate the case this research presented: as the coverage tended to present more critical frames, it could be claimed that the arguments depicting the media as fully functional to corporate interests cannot be sustained fully, but since the strength of the critique is mild, the case is the same for arguments from the critical literature that argue that the media can exert a strong countervailing force against corporations. In that sense, the most suitable theory to explain framing of the algorithmic bias topic are those who contend the media tends not to act as a ‘major’ opponent of corporations. In that line, theories of media framings of technologies do seem to apply fully for this particular research. Nevertheless, since there was no considerable difference in framing between the two periods, it could be claimed the idea of framing as an evolving process cannot be applied to the evidence presented here. Still, it is worth wondering whether this has to do with the limitations of this project in terms of the time period selected: perhaps the analysis was too constrained to allow such fluctuations to be seen.

Finally, even though it is not the purpose of this work to engage in a deep discussion about the mitigation of algorithmic bias, having in mind the theoretical framework that informed this work, it might be worth mentioning that it is contended here that arguing that algorithms can eventually be dissociated from the issue of bias is a naive interpretation of the working of these technologies. This work argues that the roots of such issues are situated in the human realm rather than in the technical. For instance, just as one should not expect editing in journalism to be free of value, it should not be expected that algorithmic gatekeeping will disappear. The same applies to discrimination: as long as prejudices and unfairness are present in society, they might be present in algorithms as well. To expect algorithms to be free of bias would imply the existence of a 'universal' standard for neutrality. This by no means should be interpreted as a view that the future of algorithms is doomed: while not disregarding the benefits of their use and while not expecting the issue of bias to disappear entirely, this work does acknowledge their possible pernicious consequences and contends that regulation tools such as state legislation should be used to guarantee that the power and effects algorithms have are in line, to the fullest extent possible, with the best interest of the public. This will inevitably imply the adjustment and managing of many factors involved in the 'algorithmic assemblages', but it is contended this should be done for and in front of society.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Robin, for her great ideas and generous amount of time dedicated to the design of this project. Also, I would like to thank Natalia, Lucía and Morgane. Without their time and patience this dissertation would have never been possible.

REFERENCES

Agresti, A. & Finlay, B. (2009) *Statistical methods for the social sciences*, Upper Saddle River, N.J: Pearson Prentice Hall.

Allcott, H. & Gentzkow, M. (2017) *Social Media and Fake News in the 2016 Election*, Journal of Economic Perspectives 31 (2): 211–236.

Amoore, L. (2006) *Biometric borders: Governing mobilities in the war on terror*, Political Geography, 25: 336–351.

Amoore, L. (2009) *Algorithmic war: Everyday geographies of the war on terror*, Antipode, 41: 49–69.

Ananny, M. (2016) *Toward an Ethics of Algorithms Convening, Observation, Probability, and Timeliness*, *Science, Technology, & Human Values*, 41 (1): 93-117.

Anderson, C.W. (2011) *Deliberative, agonistic, and algorithmic audiences: Journalism's vision of its public in an age of audience*, Journal of Communication, (5): 529–547.

Anderson, C.W. (2012) *Towards a sociology of computational and algorithmic journalism*, New Media & Society, 15 (7): 1005-102.

Angwin, J.; Larson, J.; Mattu, S. and Kirchner, L. (2016) *Machine bias: There's software used across the country to predict future criminals. and it's biased against blacks*, ProPublica, May 23.

Arthur, W.B. (2009) *The nature of technology: What it is and how it evolves*, New York: Allen Lane.

Bauer, J.M. (2016) *The Internet and income inequality: Socio-economic challenges in a hyperconnected society*,Telecommunications Policy, 1-11.

Bednar, M.K. (2012) *Watchdog or Lapdog? A Behavioral View of the Media as a Corporate Governance Mechanism*, Academy of Management Journal, 55 (1).

Bengsston, S. (2018) *Ethics and morality beyond the Actor-Network: Doing the right thing in an algorithmic culture*, Presented at the 68th Annual Conference of the International Communication Association, Prague, Czech Republic, May 24-28, 2018.

Berriman, L. (2017) *Book review. Framing internet safety: the governance of youth online*, Information, Communication & Society, 20 (12): 1829–1830.

Bijker, W.E. (1993) *Do Not Despair: There Is Life after Constructivism*, Science, Technology, & Human, Values, 18: 113-138.

Bijker, W.E. (1995) *Of bicycles, bakelites, and bulbs: toward a theory of sociotechnical change*, Cambridge, Mass.: MIT Press.

Bijker, W.E. & Law, J. (1992) *Shaping technology/building society: studies in sociotechnical change*, Cambridge, Mass.: MIT Press.

Bowker, G. & Star, S.L. (2000) *Sorting Things Out: Classification and Its Consequences*, Cambridge, Mass.: MIT Press.

Bozdag, E. (2013) *Bias in algorithmic filtering and personalization*, Ethics and Information Technology, 15 (3): 209-227.

Bucher, T. (2012) *‘Want to be on the top?’ Algorithmic power and the threat of invisibility on Facebook,* New Media and Society, 14(7): 1164–1180.

Bucher, T. (2015) *The algorithmic imaginary: exploring the ordinary effects of Facebook algorithms*, Information, Communication & Society, 20 (1): The Social Power of Algorithms.

Budak, C.; Goel, S. and Rao, J.M. (2014) *Fair and Balanced? Quantifying Media Bias through Crowdsourced Content Analysis* (November 17, 2014) Available at SSRN: https://ssrn.com/abstract=2526461 or http://dx.doi.org/10.2139/ssrn.2526461

Butler, P. (2018) *Technocratic Automation and Contemplative Overlays in Artificially Intelligent Criminal Sentencing,* In: Giorgino V., Walsh Z. (eds) Co-Designing Economies in Transition. Palgrave Macmillan, Cham.

Caliskan, A.; Bryson, J.J.; Narayanan, A. (2017) *Semantics derived automatically from language corpora contain human-like biases*, Science, 356 (6334): 183-186.

Carroll, C.E. & McCombs, M. (2003) *Agenda-setting effects of business news on the public's images and opinions about major corporations*, Corporate Reputation Review, (16): 36–46.

Chadwick, A. (2013) *The Hybrid Media System,* Politics and Power, Oxford: Paperback.

Chomsky, N. (October, 1997) *What makes mainstream media mainstream*, Z Magazine.

Chong, D. & Druckman, J. (2007b) *Framing Theory*, The Annual Review of Political Science, 103-126.

Chong, D. & Druckman, J.N. (2007a) *A theory of framing and opinion formation in competitive elite environments,* Journal of Communication, 57 (1): 99-118.

Cohen, M. and Sundararajan, A. (2015) *Self-Regulation and Innovation in the Peer-to-Peer Sharing Economy,* University of Chicago Law Review Online, 82(1): Article 8.

Cooper, A.H. (2002) *Media framing and social movement mobilization: German peace protest against INF missels, the Gulf War, and NATO peace enforcement in Bosnia*, European Journal of Political Research, 41: 37–80.

Cormen, T.H.; Leiserson, C.E.; Rivest, R.L. and Stein, C. (2009) *Introduction to Algorithms,* (3rd ed.) Mass.: MIT Press and McGraw-Hill.

Cox, G. (2013) *Speaking code: Coding as aesthetic and political expression*, Cambridge, Mass.: MIT Press.

Crawford, K. (2015) *Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics,* Science, Technology, & Human Values, 41 (1): 77–92.

Danks, D. & London, A.J. (2017) 'Algorithmic Bias in Autonomous Systems', Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI-17).

de Vreese, C.H.; Peter, J. & Semetko, H.A. (2001) *Framing politics at the launch of the Euro: A cross-national comparative study of frames in the news,* Political Communication, 18: 107–122.

Deacon, D., Pickering, M.; Golding, P. and Murdock, G. (1999) *Researching Communications: A Practical Guide to Methods in Media and Cultural Analysis,* London: Arnold.

Diakopoulos, N. & Koliska, M. (2016) *Algorithmic transparency in the news media,* Digital Journalism, 5 (7): 809-828.

Diakopoulos, N. (2013) *Algorithmic accountability reporting: On the investigation of black boxes*, Tow/Knight Brief, Tow Center for Digital Journalism, Columbia Journalism School.

Diakopoulos, N. (2015) *Algorithmic accountability: Journalistic investigation of computational power structures*, Digital Journalism, 3 (3): Journalism in an Era of Big Data: Cases, Concepts, and Critiques, 398-415.

Dijck, J. van (2009) *Users Like You? Theorizing Agency in User-generated Content*, Media Culture & Society, 31 (1): 41–58.

Dijck, J. van (2014) *Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology,* Surveillance & Society, 12 (2): 197–208.

Dimitrova, D.V.; Kaid, L.L.; Williams, A. & Trammell, K.D. (2005) *War on the web: The immediate news framing of Gulf War II,* The Harvard International Journal of Press/Politics, 10: 22–44.

Drucker, J. (2013) *Performative materiality and theoretical approaches to interface,* Digital Humanities Quarterly, 7(1) Retrieved July 24, 2018, from http://www.digitalhumanities.org/dhq/vol/7/1/000143/000143.html

Druckman, J.N. (2001) *The implications of framing effects for citizen competence*, Political Behavior, 23 (3): 225-256.

Entman, R.M. (2007) *Framing Bias: Media in the Distribution of Power*, Journal of Communication, 57 (1): 163–173.

Eubanks, V. (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, London: St. Martin's Press

Fairclough, N. (1995) *Media Discourse*, London: Edward Arnold.

Fairclough, N. (2003) *Analysing Discourse: Textual Analysis for Social Research*, London: Routledge.

Feinstein, A. & Cicchetti, D. (1990) *High agreement but low kappa: I. The problems of two paradoxes,* J Clin Epidemiol, 43 (6): 543–9.

Fisk, N. (2016) *Framing internet safety: the governance of youth online*, Cambridge, Mass.: MIT Press.

Flores, A.W.; Bechtel, K.; and Lowenkamp, C.T. (2016) *False positives, false negatives, and false analyses: A rejoinder to machine bias: There’s software used across the country to predict future criminals. and it’s biased against blacks,* Federal Probation, 80 (38).

Foucault, M. (1969) *The archaeology of knowledge*, Social Science Information, 9 (1): 175 - 185. London: Sage.

Friedman, B.; Nissenbaum, H. (1996) *Bias in Computer Systems*, ACM Transactions on Information Systems, 14 (3): 330–347.

Fuchs, C. (2015) *Dallas Smythe today - the audience commodity, the digital labour debate, Marxist political economy and critical theory. Prolegomena to a digital labour theory of value,* TripleC - Cognition, Communication and Co-operation, 10 (2): 692-740.

Galloway, A. (2004) P*rotocol: How Control Exists after Decentralization*, Cambridge, Mass.: MIT Pres

Gardner, M.J. & Altman, D.G. (1986) *Confidence intervals rather than P values: estimation rather than hypothesis testing,* British Medical Journal, 292.

Gaskell G.; Ten Eyck, T.; Jackson J. & Veltri, G. (2004) *Public attitudes to nanotechnology in Europe and the united states,* Nature Materials, 3 (8): 496.

Geiger, R.S. (2009) *Does Habermas Understand the Internet? The Algorithmic Construction of the Blogo/Public Sphere*, Gnovis, A Journal of Communication, Culture, and Technology, 10 (1): 1-29.

Geiger, S.R. (2014) B*ots, bespoke, code and the materiality of software platforms*, Information, Communication & Society, 17 (3): 342–356.

Gillespie, T. (2014b, June 25) *Algorithm [draft] [#digitalkeyword]*, Culture Digitally. Retrieved July 26, 2018, from http://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/

Gillespie, T. (2014) T*he relevance of algorithms*, In Gillespie,T; Boczkowski, PT & Foot, KA (Eds.) *Media technologies: Essays on communication, materiality, and society* (pp. 167–193) Cambridge: MIT Press.

Gillespie, T. & Postigo, H. (May 4, 2012) *Five more points*, Culture Digitally, Available at http://culturedigitally.org/2012/05/five-more-points/.

Gitelman, L. (2006) *Always Already New: Media, History, and the Data of Culture*, Cambridge, Mass.: MIT Press.

Goffey, A. (2008) Algorithm, in *Software studies: A Lexicon*, ed. Matthew Fuller, 15-20. Cambridge, Mass.: MIT Press.

Goodman, B. & Flaxman, S. (2016) *EU regulations on algorithmic decision-making and a 'right to explanation,* Ai Magazine, 38 (3), Presented at 2016 ICML Workshop on Human Interpretability in Machine Learning (WHI 2016) New York, NY.

Guess, A.; Nyhan, B. and Reifler, J. (2018) *Selective Exposure to Disinformation: Evidence from the Consumption of Fake News During the 2016 US Presidential Campaign*, European Research Council.

Hall, S. (1997) The work of representation, in Hall, S (Ed.) *Representation: Cultural Representations and Signifying Practices*, 133,74. London: Sage.

Hansen, A. (1998) *Mass Communication Research Methods*, New York: New York University Press.

Helberger, N.; Kleinen-von K.; van der Noll, R.K. (2015) *Regulating the new information intermediaries as gatekeepers of information diversity,* info, 17 (6): 50-71.

Hesmondhalgh, D. (2006) *Bourdieu, the media and cultural production,* Media, Culture and Society, 28, (2): 211-231.

Hilgartner, S. (2000) *Science on Stage: Expert Advice as Public Drama*, Stanford, CA: Stanford University Press.

Howard, A. & Borenstein, J. (2017) *The Ugly Truth About Ourselves and Our Robot Creations: The Problem of Bias and Social Inequity*, Science and Engineering Ethics, September 2017, 1-16.

Introna, L.D. (2011) *The Enframing of Code: Agency, Originality and the Plagiarist'*, Theory, Culture & Society, 28 (6): 113–141.

Introna, L.D. & Wood, D. (2004) *Picturing algorithmic surveillance: the politics of facial recognition systems,* Surveillance & Society, 2 (2/3): 177-198.

Iyengar S., (1996) *Framing responsibility for political Issues*, Ann. Amer. Acad. Polit. SS 546, 59–70.

Joh, E.E. (2017) *A Certain Dangerous Engine': Private Security Robots, Artificial Intelligence, and Deadly Force,* UC Davis L. Rev., 2017.

Joh, E.E. (2017) *Feeding the Machine: Policing, Crime Data, & Algorithms*, William & Mary Bill of Rights Journal, 26 (2): Article, 287-302.

Joh, E.E. (2018) *Artificial Intelligence and Policing: First Questions,* Seattle Univ. L. Rev., 2018. Available at SSRN: https://ssrn.com/abstract=3168779

Kim, P.T. (2017) *Auditing Algorithms for Discrimination,* 166 University of Pennsylvania Law Review, Online 189.

Kitchin, R. (2014) *The data revolution: Big data, open data, data infrastructures and their consequences,* London: Sage.

Kitchin, R. (2016) *Thinking critically about and researching algorithms,* Information, Communication & Society, 20 (1): 14–29.

Kitchin, R. & Dodge, M. (2011) *Code/space: Software and everyday life*, Cambridge: MIT Press.

Kotsiantis, A.B. (2007) Supervised machine learning: A review of classification techniques' Maglogiannis, I et al., in *Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems,* St. Raphael, CA: Morgan and Claypool.

Krippendorff, K. (2004) *Content Analysis: An Introduction to Its Methodology,* London: Sage.

Kushner, S. (2013) *The freelance translation machine: Algorithmic culture and the invisible industry,* New Media & Society, 15(8): 1241–1258.

Langlois, G.M. (2012) *Participatory culture and the new governance of communication: The paradox of participatory media,* Television and New Media, 14 (2): 91-105.

Latzer, M.; Just, N.; Saurwein, F. and Slominski, P. (2003) *Regulation remixed: institutional change through self and co-regulation in the mediamatics sector,* Communications & Strategies, 50 (2): 127-15.

Lecheler, S.; de Vreese, C.H. (2012) *News Framing and Public Opinion. A Mediation Analysis of Framing Effects on Political Attitudes*, Journalism & Mass Communication Quarterly, 89 (2): 185-204.

Lipartito, K. (2011) *The Narrative and the Algorithm: Genres of Credit Reporting from the Nineteenth Century to Today*, Miami: Florida International University.

Livingstone, R. (March 13, 2017) *The future of online advertising is big data and algorithms,* The Conversation. Retrieved 24 July 2018. From: https://theconversation.com/the-future-of-online-advertising-is-big-data-and-algorithms-69297

Lombard, M., Snyder-Duch, J., & Bracken, C.C. (2004) *Practical resources for assessing and reporting intercoder reliability in content analysis research projects,* Retrieved April 28, 2004, from www.temple.edu/mmc/reliability

MacCormick, J. (2013) *Nine algorithms that changed the future: The ingenious ideas that drive today’s computers,* Princeton, NJ: Princeton University Press.

Mackenzie, A. (2007) *Protocols and the irreducible traces of embodiment: The Viterbi algorithm and the mosaic of machine time,* In R. Hassan & R. E. Purser (Eds.) 24/7: Time and temporality in the network society (pp. 89–106) Stanford, CA: Stanford University Press.

Mackenzie, A. & Vurdubakis, T. (2011) *Code and codings in Crisis: Signification, performativity and excess,* Theory, Culture & Society, 28 (6): 3–23.

Macnamara, J. (2005) *Media content analysis: Its uses, benefits and Best Practice Methodology,* Asia Pacific Public Relations Journal, 6 (1): 1–34.

Mager, A. (2012) *Algorithmic ideology: How capitalist society shapes search engines,* Information, Communication, and Society, 15 (5): 769-787.

Mann, B.W. (2018) *Autism Narratives in Media Coverage of the MMR Vaccine-Autism Controversy under a Crip Futurism Framework,* Health Communication 0:0, 1-7.

Mann, S.; Nolan, J. and Wellman, B. (2003) *Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments,* Surveillance & Society, 1 (3): 331-355.

Mansell, R. (2015) *The public's interest in intermediaries,* Info: The Journal of Policy, Regulation and Strategy for Telecommunication, Information and Media, 17 (6): 8-18.

Mansell, R. (2016) *Unpacking Black Boxes: Understanding Digital Platform Innovation*, Submitted to Information, Communication and Society journal, 25 November 2016.

Mansell, R. (2017) The rise of computational power: accountability, imagination, and choice, in George, Cherian, (ed.) *Communicating with Power*. ICA International Communication Association. Annual Conference Theme Book Series. Peter Lang, New York, USA, pp. 69-83.

Mansell, R. (November 27, 2014) *Governing the gatekeepers: is formal regulation needed,* LSE Media Policy Blog, Blog Entry, URL: http://blogs.lse.ac.uk/mediapolicyproject/2014/11/27/governing-the-gatekeepers-is-formal-regulation-needed/

Marshall, P. (2018) *Algorithms Can Mask Biases in Hiring,* Sage: Business Researcher.

Matthes, J. & Kohring, M. (2008) *The Content Analysis of Media Frames: Toward Improving Reliability and Validity', Journal of Communication,* 58 (2): 258-279.

Merrill, J.C. (2011) Theoretical Foundations for Media Ethics, in *Controversies in Media Ethics*, 3rd ed., edited by A. D. Gordon, J. M. Kittross, J. C. C. Merrill, W. Babcock, and M. Dorsher, 3-32. New York: Routledge.

Millington, B. & Millington, R. (2015) *The datafication of everything: Toward a sociology of sport and big data,* Sociology of Sport Journal, 32 (2): 140-160.

Miyazaki, S. (2012) *Algorhythmics: Understanding micro-temporality in computational cultures*, Computational Culture, Issue 2. Retrieved July 24, 2018, from http://computationalculture.net/article/algorhythmics-understanding-micro-temporality-in-computational-cultures

Montfort, N.; Baudoin, P.; Bell, J.; Bogost, I.; Douglass, J.; Marino, M.C., Vawter, N. (2012) *10 PRINT CHR$ (205.5 + RND (1)) : GOTO 10*'. Cambridge: MIT Press.

Morozov, E. (2011) *Don't be evil,* The New Republic, July 13. Available at http://www.tnr.com/article/books/magazine/91916/google-schmidt-obama-gates-technocrats.

Morozov, E. (2012) *The Net Delusion: The Dark Side of Internet Freedom,* London: Paperback.

Muhamad, J. & Yang, F. (2017) *Framing Autism: A Content Analysis of Five Major News Frames in U.S.-Based Newspapers,* Journal of Health Communication, 22 (3): 190-197.

Nagel, J. (1975) *The descriptive analysis of power*, New Haven, CT: Yale University Press.

Nakamura, L. (2009) *The new media of surveillance,* London: Routledge. pp. 149–162.

Napoli, P.M. (2013) *The algorithm as institution: Toward a theoretical framework for automated media production and consumption,* Presented at the Media in Transition Conference, Massachusetts Institute of Technology, Cambridge, MA.

Neuendorf, K. (2002) *The Content Analysis Guidebook*, Thousand Oaks, CA: Sage Publications.

Neyland, D. (2015) *On organizing algorithms,* Theory, Culture & Society, 32 (1): 119–132.

Nisbet, M.C. & Lewenstein, B.V. (2002) *Biotechnology and the American media – the policy process and the elite press, 1970 to 1999,* Sci. Commun., 23 (4): 359–391.

Nissenbaum, H. (1994) *Computing and accountability*, Commun ACM, 37 (1): 72–80.

Nissenbaum, H. (2001) *How computer systems embody values*, Computer, 34 (3): 120–119.

O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Cambridge, Mass.: Harvard University Press.

Otterbaher, J.; Bates, J.; Clough, P. (2017) *Competent Men and Warm Women: Gender Stereotypes and Backlash in Image Search Results,* CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 6620-6631.

Otterbaher, J.; Checco, A.; Demartini, G.; Clough, P. (2018) *Investigating User Perception of Gender Bias in Image Search: The Role of Sexism,* SIGIR '18 The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, 933-936.

Pariser, E. (2011) *The Filter Bubble. What the Internet is Hiding from Yo,.* London: Paperback.

Pasquale, F. (2014) *The emperor’s new codes: Reputation and search algorithms in the finance sector,* Draft for discussion at the NYU ‘Governing Algorithms’ conference. Retrieved July 24, 2018, from http://governingalgorithms.org/wp-content/uploads/2013/05/2-paper-pasquale.pdf

Pasquale, F. (November 18, 2009) *Assessing algorithmic authority*, Madisonian.net, Available at http://madisonian.net/2009/11/18/assessing-algorithmic-authority/.

Payne, J.; Szymkowiak, A.; Robertson, P.; Johnson, G. (2013) Gendering the Machine: Preferred Virtual Assistant Gender and Realism in Self-Service, in Aylett R., Krenn B., Pelachaud C., Shimodaira H. (eds) *Intelligent Virtual Agents*. IVA 2013. Lecture Notes in Computer Science, 8108, Springer, Berlin, Heidelberg.

Pinch, T.J., & Bijker, W.E. (1984) *The social construction of facts and artifacts: Or how the sociology of science and technology might benefit each other*, Social Studies of Science, 14: 399-442.

Pitt, J.; Busquets, D. and Riveret, R. (2013) *The pursuit of computational justice in open systems,* AI & Society, Dec.: 1-19.

Porter, T.M. (1995) *Trust in numbers: The pursuit of objectivity in science and public life,* Princeton, NJ: Princeton University Press.

Prell, C. (2009) *Rethinking the Social Construction of Technology through 'Following the Actors': A Reappraisal of Technological Frames,* University of Sheffield, Sociological Research Online, 14 (2).

Reich, R. (2016) When Robots Take Over, ch. 22 in *Saving Capitalism: For the Many, Not the Few* (pp. 203-210) New York: Vintage.

Reuters (2017) *Reuters Institute Digital News Report 2017,* Reuters Institute for the Study of Journalism. URL:

https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf

Riffe, D. & Freitag, A. (1997) *A content analysis of content analyses: twenty-five years of Journalism Quarterly*, Journalism and Mass communication Quarterly, 74: 873–882.

Rössler, P. (2001) *Between Online Heaven and Cyberhell: The Framing of `The Internet' by Traditional Media Coverage in Germany,* New Media and Society, 3 (1): 49-66

Sandvig, C.; Hamilton, K.; Karahalios, K.; Langbort, C. (2014) Gangadharan, SP; Eubanks, V; Barocas, S, eds. *An Algorithm Audit* (PDF) Data and Discrimination: Collected Essays. Open Technology Institute.

Saurwein, F.; Just, N. & Latzer, M. (2015) *Governance of algorithms: options and limitations,* info, 17 (6): 35-49.

Scheufele, D.A. (1999) *Framing as a theory of media effects*, J. Commun., 49 (1): 103–122.

Scheufele, D.A. (2000) *Agenda-setting, priming, and framing revisited: Another look at cognitive effects of political communication*, Mass Communication & Society (2 & 3): 297–316.

Scheufele, D.A. & Lewenstein B.V. (2005) *The public and nanotechnology: How citizens make sense of emerging technologies,* Journal of Nanoparticle Research, 7: 659–667.

Seaver, N. (2013) *Knowing Algorithms,* Media in Transition, Volume 8, Cambridge, Mass.

Seymour, W. (2018) Detecting Bias: Does an Algorithm Have to Be Transparent in Order to Be Fair?, p. 2-8 in *BIAS, Proceedings of the Workshop on Bias in Information, Algorithms, and Systems,* Sheffield, United Kingdom, March 25, 2018.

Shneiderman, B. (November 29, 2016) *Opinion: The dangers of faulty, biased, or malicious algorithms requires independent oversight,* PNAS. 113 (48) 13538-13540; published ahead of print November 23, 2016. https://doi.org/10.1073/pnas.1618211113

Smart, P.R. (2013) The web-extended mind, in Halpin, H. and Monnin, A. (Eds) *Philosophical Engineer: Toward a Philosophy of the Web,* Wiley Blackwell, Chichester, 116-133.

Smart, P.R. & Shadbolt, N.R. (2014) Social machines, in Khosrow-Pour, M. (Ed.) *Encyclopedia of Information Science and Technology*, Third Edition. IGI Global, Hershey, PA, pp. 6855-6862.

Steiner, C. (2012) *Automate this: How algorithms took over our markets, our jobs, and the world,* New York, NY: Portfolio.

Striphas, T. (2010) *How to have culture in an algorithmic age,* Differences and Repetitions, July 13. Available at http://www.diffandrep.org/2010/07/13/how-to-have-culture-in-an-algorithmic-age/.

Striphas, T. (2015) *Algorithmic culture,* European Journal of Cultural Studies, 18 (4/5): 395-412.

Takhteyev, Y. (2012) *Coding places: Software practice in a South American City,* Cambridge, Mass.: MIT Press.

Thompson, J. (2005) *The New Visibility,*Theory, Culture and Society, 22 (6): 31-51.

Trottier, D. (2012) *Social Media as Surveillance*, Ashgate: Farnham.

Tufekci, Z. (2015) *Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency*, Colorado Technology Law Journal, 13: 203–218.

Tufekci, Z. (2015b) *Facebook Said Its Algorithms Do Help Form Echo Chambers, and the Tech Press Missed It,* New Perspectives Quarterly, 32 (3): 9-12.

Tushnet, R. (2008) *Power without responsibility: Intermediaries and the First Amendment*, The George Washington Law Review, 76 (4): 986-1016.

Van Couvering, E. (2007) *Is relevance relevant? Market, science, and war: Discourses of search engine quality,* Journal of Computer-Mediated Communication, 12 (3).

Van Couvering, E. (2010) *Search engine bias: the structuration of traffic on the World-Wide Web,* PhD thesis, The London School of Economics and Political Science (LSE).

Wachter-Boettcher, S. (2017) *Technically Wrong: Sexist Apps, Biased Algorithms, and Other Threats of Toxic Tech,* London: W. W. Norton.

Weizenbaum, J. (1976) *Computer power and human reason: from judgment to calculation,* San Francisco: W.H. Freeman.

Wiener, J.B. (2004) *The regulation oftechnology, and the technology of regulation,* Technology in Society, 26: 483–500.

Winner, L. (1978) *Autonomous Technology: Technics-out-of-Control as a Theme in Political Thought,* Cambridge, Mass.: MIT Press.

Wood, A.K.; Ravel, A.M.; Dykhne, I. (2018) *Fool Me Once: Regulating 'Fake News' and Other Online Advertising,* Southern California Law Review, 91(6) 2018, Forthcoming.

Yapo, A; Weiss, J (2018) *Ethical Implications of Bias in Machine Learning*, Proceedings of the 51st Hawaii International Conference on System Sciences, 5365-5371.

Yioutas, J. & Segvic, I. (2003) *Revisiting the Clinton/Lewinsky scandal: The convergence of agenda setting and framing,* Journal and Mass Communication Quarterly, 80 (3): 567–582.

NON-ACADEMIC REFERENCES

Angwin, J. & Mattu, S. (20 September 2016) Amazon Says It Puts Customers First. But Its Pricing Algorithm Doesn't. *ProPublica*, ProPublica [Last accessed: July 27, 2018]

Associated Press, (January 31, 2018) Concerns in the courtroom as algorithms are used to help judges rule on jail time which can flag black people as twice as likely to re-offend compared to white defendants. *Daily Mail,* URL: http://www.dailymail.co.uk/news/article-5333757/AI-court-When-algorithms-rule-jail-time.html [Last accessed: July 27, 2018]

Bartlett, N. (May 8, 2018) Police gang database 'breaches human rights and discriminates against young black men who are fans of grime music. *MSN News*, URL: https://www.msn.com/en-gb/news/uknews/police-gang-database-breaches-human-rights-and-discriminates-against-young-black-men-who-are-fans-of-grime-music/ar-AAwYrKL [Last accessed: July 27, 2018]

Bartoletti, I. (April 4, 2018) Women must act now, or male-designed robots will take over our lives. *The Guardian*, URL: https://www.theguardian.com/commentisfree/2018/mar/13/women-robots-ai-male-artificial-intelligence-automation [Last accessed: July 27, 2018]

Buranyi, S. (March 4, 2018) How to persuade a robot that you should get the job. *The Guardian,* URL: https://www.theguardian.com/technology/2018/mar/04/robots-screen-candidates-for-jobs-artificial-intelligence [Last accessed: July 27, 2018]

Byrne, W. (February 28, 2018) Now Is The Time To Act To End Bias In AI. *FastCompany,* URL: https://www.fastcompany.com/40536485/now-is-the-time-to-act-to-stop-bias-in-aI [Last accessed: July 27, 2018]

Dodd, V. (May 13, 2018) UK accused of flouting human rights in 'racialised' war on gangs. *The Guardian,* URL: https://www.theguardian.com/uk-news/2018/may/09/uk-accused-flouting-human-rights-racialised-war-gangs [Last accessed: July 27, 2018]

Fussel, S. (Januray 11, 2018) Google Censors Gorillas Rather Than Risk Them Being Mislabeled As Black People—But Who Does That Help?. *Gizmodo,* URL: https://gizmodo.com/google-censors-gorillas-rather-then-risk-them-being-mis-1821994460 [Last accessed: July 27, 2018]

Graham-Harrison, E. & Cadwalladr, C. (March 17, 2018) Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian,* Archived from the original on March 18, 2018. URL: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

Guthrie Weissman, C. (March 28, 2018) Bank of America’s bot is “Erica” because apparently all digital assistants are women. *FastCompany,* URL: https://www.fastcompany.com/40551011/bank-of-americas-bot-is-erica-because-apparently-all-digital-assistants-are-women [Last accessed: July 27, 2018]

Hills, T.T. (March 26, 2018) Algorithms Can’t Tell When They’re Broken–And Neither Can We. *FastCompany*, URL: https://www.fastcompany.com/40549744/algorithms-cant-tell-when-theyre-broken-and-neither-can-we [Last accessed: July 27, 2018]

Media Bias/Fact Check (2016) BBC, URL: https://mediabiasfactcheck.com/bbc/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2016) Gizmodo, URL:https://mediabiasfactcheck.com/gizmodo/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2016) MIT Technology Review, URL:https://mediabiasfactcheck.com/mit-technology-review/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2016) Wired, URL:https://mediabiasfactcheck.com/wired-magazine/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) Daily Mail, URL: https://mediabiasfactcheck.com/daily-mail/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) Daily Telegraph, URL: https://mediabiasfactcheck.com/daily-telegraph/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) Fast Company, URL:https://mediabiasfactcheck.com/fast-company-magazine/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) MSN News, URL: https://mediabiasfactcheck.com/msn-com/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) ReCode, URL: https://mediabiasfactcheck.com/recode/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) Sky News, URL: https://mediabiasfactcheck.com/?s=sky+news [Last accessed: July 27, 2018]

Media Bias/Fact Check (2017) TechCrunch, URL:https://mediabiasfactcheck.com/techcrunch/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2018) The Guardian, URL: https://mediabiasfactcheck.com/theguardian/ [Last accessed: July 27, 2018]

Media Bias/Fact Check (2018) Huffington Post, URL: https://mediabiasfactcheck.com/huffington-post/ [Last accessed: July 27, 2018]

Naish, J. (Januray 25, 2018) Are computers turning into bigots? Machines that learn to think for themselves are revolutionising the workplace. But there's growing evidence they are acquiring racial and class prejudices too. *Daily Mail*, URL: http://www.dailymail.co.uk/sciencetech/article-5310031/Evidence-robots-acquiring-racial-class-prejudices.html [Last accessed: July 27, 2018]

Palmer, A. (March 29, 2018) Companies using controversial AI software to help hire or fire employees - and gauge whether they like their boss. *Daily Mail*, URL: http://www.dailymail.co.uk/sciencetech/article-5556221/New-AI-software-help-companies-hire-fire-employees-gauge-like-job.html [Last accessed: July 27, 2018]

Pinkstone, J. (April 27, 2018) Facial recognition AI built into police body cameras will suffer from racial bias and could lead to false arrests, civil liberties experts warn. *Daily Mail*, URL: http://www.dailymail.co.uk/sciencetech/article-5663953/Facial-recognition-AI-built-police-body-cameras-lead-FALSE-ARRESTS-experts-warn.html [Last accessed: July 27, 2018]

Pinkstone, J. (March 14, 2018) Are Alexa and Siri SEXIST? Expert says AI assistants struggle to understand female speech because it is quieter and more 'breathy'. *Daily Mail*, URL: http://www.dailymail.co.uk/sciencetech/article-5499339/AI-assistants-sexist-understand-men-better.html [Last accessed: July 27, 2018]

Rosenberg, M.; Confessore, N.; Cadwalladr, C. (March 17, 2018) How Trump Consultants Exploited the Facebook Data of Millions. *The New York Times,* Archived from the original on

March 17, 2018. From: https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html [Last accessed: July 27, 2018]

Simonite, T. (February 6, 2018) Photo algorithms ID white men fine - black women, not so much. *Wired*, URL: https://www.wired.com/story/photo-algorithms-id-white-men-fineblack-women-not-so-much/ [Last accessed: July 27, 2018]

Snow, J. (January 26, 2018) Here are the secrets to useful AI, according to designers. *MIT Technology Review*, URL: https://www.technologyreview.com/the-download/610077/here-are-the-secrets-to-useful-ai-according-to-designers/ [Last accessed: July 27, 2018]

Solomon, M. (September, 2017) Top 10 Tech Blogs & Publications (For Your Inbox). *The Huffington Post,* available at: https://www.huffingtonpost.com/michael-s-solomon/top-10-tech-blogs-publica_b_14062006.html [Last accessed: July 27, 2018]

Winick, E (March 8, 2018) Tech talent actually shows promise for a more female future. *MIT Technology Review*, URL: https://www.technologyreview.com/the-download/610450/tech-talent-actually-shows-promise-for-a-more-female-future/ [Last accessed: July 27, 2018]

APPENDIX 1 – FINAL CODEBOOK

EXPLANATORY VARIABLES

1. ID number (ID) [start with 1 and number onwards]

2. Media in which article is published (Media)

    1. BBC News
    2. The Guardian
    3. Daily Mail
    4. Huffington Post
    5. Sky News
    6. MSN News
    7. The Telegraph
    8. MIT Tech Review
    9. Wired
    10. Fast Company
    11. Gizmodo
    12. TechCrunch
    13. Recode

3. Search Criteria (Search)

    1. Bias+ algorithm
    2. Bias + algorithms
    3. Bias + artificial intelligence
    4. Bias + AI
    5. Bias + Machine learning

4. URL [Open]

5. Headline[Open]

6. Date of publication (Date) [day/month/year]

7. Who is the author of the article? (Author) [Open]

8. Was the article published after the Cambridge Analytica scandal (March 17)? (after_CA) [No=0, Yes=1]

9. What kind of publication is it? (kind)
   1. Mainstream Media - UK consumption [0]
   2. Technology publications [1]

<u>DISCOURSES</u>

10. What is the overall posture of the article towards al? (from 1 to 5)(Posture)
   1. Very Negative
   2. Negative
   3. Neutral
   4. Positive
   5. Very Positive

11. Is either of these general framings more evident in the article? (Algo_ReduceReinforce)
   1. Algorithms can reduce biases [0]
   2. Algorithms reinforce biases [1]
   3. N/A

12. Are governments/parliaments mentioned in any part of the article? (GovPar_Mention) [No=0, Yes=1]

13. Are representatives of governments/parliaments directly quoted? (GovPar_quote) [No=0, Yes=1]

14. Does the article paraphrase something government/parliaments stated? (no direct from specific person quote but referring to something governments/parliaments stated) (GovPar_ref) [No=0, Yes=1]

15. Are any of these ideas presented in relation to algorithms in general?
    1. Algorithms are empowering and beneficial to users (a_Empower) [No=0, Yes=1]
    2. Algorithms are cost/effective for companies (b_CostEffective) [No=0, Yes=1]

16. Are corporations mentioned in any part of the article? (Corpo_Mention) [No=0, Yes=1]

17. Are any corporate representatives directly quoted (Corpo_Quote) [No=0, Yes=1]

18. Does the article paraphrase something corporations satated? (no direct from specific person quote but references to general statements made by corporations) (Corpo_Ref) [No=0, Yes=1]

19. Does the article mention corporations’ efforts to solve algorithmic biases? (CorpoEffort) [No=0, Yes=1]

20. Did the journalist depend on his/her source on computer/technical experts? (sources) [No=0, Yes=1]

21. Does the article give a technical explanation for the bias? (technical) [No=0, Yes=1]

22. For Mainstream Media, is the article in the technology/science section? (TechSection) [No=0, Yes=1]

MAIN TOPICS

23. Are any of the following framings towards 'algorithmic bias' present in the article? (issues) (select more than one option)

1. The issue of algorithmic bias shows the lack of corporate ethics (a_Ethics) [No=0, Yes=1]
2. The issue of algorithmic bias shows the lack of corporate accountability (b_accountability) [No=0, Yes=1]
3. Algorithms are "black boxes" that should be more transparent for the understanding of the publics. (c_BlackBox) [No=0, Yes=1]
4. Algorithms reproduce and enhance content biases for users (d_EchoChambers) [No=0, Yes=1]
5. The algorithmic bias issue is related to the lack of antitrust legislation (e_Antitrust) [No=0, Yes=1]
6. There is no way to mitigate algorithmic biases (algorithms will always be biased) (f_AlwaysBiased) [No=0, Yes=1]
7. Algorithms always need to be controlled by humans (g_human) [No=0, Yes=1]
8. Algorithms are not biased (h_BiasDenial) [No=0, Yes=1]
9. Algorithms reflect pre-existing biases in societies (i_reflection) [No=0, Yes=1]
10. Government regulation reduces profitability (more regulation leading to less interest in innovation investment, therefore less possibilities to innovate remove the bias from product, no chances to improve product imply a decline in profitability for corporations) (j_ProfIssue) [No=0, Yes=1]
11. Innovation investment is key to get rid of algorithmic biases (k_InnoInnvest) [No=0, Yes=1]
12. Corporate self-regulation is important to preserve user trust (l_UserTrust) [No=0, Yes=1]

13. Corporate self-regulation is important to prevent government regulation (m_SelfReg_External) [No=0, Yes=1]
14. Government regulation is too slow to catch up with innovation (n_Reg_Slow) [No=0, Yes=1]

24. Are any of the following consequences of algorithmic bias discussed? (effects) (more than one option possible)

1. The dissemination of ‘fake news’ (a_Fakenews) [No=0, Yes=1]
2. Political interference (algorithms have influenced electoral processes) (b_PolInter) [No=0, Yes=1]
3. Negative effects on personnel selection processes (c_personnel) [No=0, Yes=1]
4. Negative effects on judicial decisions (d_Jud) [No=0, Yes=1]
5. Negative effects on police decisions (e_pol) [No=0, Yes=1]
6. Negative effects on welfare allocations (f_welfare) [No=0, Yes=1]
7. Negative effects on credit scoring (g_credit) [No=0, Yes=1]
8. Negative effects on education systems (h_educ) [No=0, Yes=1]
9. Gender discrimination (i_gender) [No=0, Yes=1]
10. Ethnic discrimination (j_ethnic) [No=0, Yes=1]
11. Negative effects on privacy (k_Privacy) [No=0, Yes=1]
12. Other (l_other) [open]

25. Does the article make any of the following suggestions to deal with the issue of ‘Algorithmic Biases’? (suggestions) (more than one option possible)

1. Corporations need higher ethical standards to prevent algorithmic bias (a_MoreEthics) [No=0, Yes=1]
2. Corporations should be more diverse to mitigate algorithmic bias (b_diversity) [No=0, Yes=1]

3. Corporations should be more transparent in order to detect algorithmic bias (c_MoreTrans) [No=0, Yes=1]
4. Corporations should change their business models because of algorithmic bias (d_ChangeModels) [No=0, Yes=1]
5. Corporations should promote user understanding on how their algorithms work. (e_ConsumerKnow) [No=0, Yes=1]

26. Is Cambridge Analytica the main topic of the article (CA_MainTopic) [No=0, Yes=1]

27. Is Cambridge Analytica referenced in the article? (CA_Mention) [No=0, Yes=1]

28. Is the algorithmic bias issue the main topic of the article? (Main_topic) [No=0, Yes=1]

29. Does the article mention any of these companies: Google, Facebook (or CA) Twitter, Apple or Amazon? (MainCompanies) [No=0, Yes=1]

30. Are any of the below mentioned as users of biased algorithms? (users) (do they use an algorithm that is biased?) (more than one possible)

1. Medical (a_medical) [No=0, Yes=1]
2. Social Media (b_SoMe) [No=0, Yes=1]
3. Finance (c_Fin) [No=0, Yes=1]
4. Governments (d_Gov) [No=0, Yes=1]
5. Legal (e_legal) [No=0, Yes=1]
6. Police System (f_police) [No=0, Yes=1]
7. Education (g_EducSys) [No=0, Yes=1]
8. Search Platform Companies (h_SearchPlat) [No=0, Yes=1]
9. Other (i_other) [open]

31. Is there a direct reference in the article to algorithms benefiting advertising revenue for corporations? (AdRevenue) [No=0, Yes=1]

<u>REGULATION</u>

32. Does the article make any reference that could be related to any of these postures towards regulation? (regulation) (search for keywords regulation, audit, control or monitoring in the text)

1. Pro-Government regulation (External regulation) (a_ProReg) [No=0, Yes=1]
2. Anti-Government regulation (b_AntiReg) [No=0, Yes=1]
3. Corporate self-regulation (companies must regulate themselves) (c_SelfReg) [No=0, Yes=1]
4. Multi-Stakeholder Regulation (stakeholders participate in the dialogue together) (d_MultiReg) [No=0, Yes=1]
5. Mixed (mentions both self-regulation and government regulation) (e_Mixed) [No=0, Yes=1]
6. Does not mention regulation (f_NoMention) [No=0, Yes=1]

APPENDIX 2 - INICIAL CODEBOOKS

- VERSION 1

Identification and Explanatory variables

1. ID number [start with 1 and number onwards]

2. Media in which article is published

   1

   2

   3

   4

   5

   6

3. Date of publication [day/month/year]

4. Was the article published before or after the Cambridge Analytica scandal? [No=0, Yes=1]

   a. Before

   b. After

5. Headline/title of text [Open]

6. How long is the article [word count]

7. Who is the author of the article? [Open]

8. What kind of publication is it?

1. Mainstream - tech section
2. Tech Niche

<u>Narrative classification</u>

9. Overall tone of the article

a. Positive
b. Negative
c. Neutral

10. How could the framing towards the ‘algorithmic bias’ be characterised? [No=0, Yes=1]

a. Positive
b. Neutral
c. Negative

11. How could the overall framing be characterised? [No=0, Yes=1]

a. Pro-regulation
a. Neutral
b. Anti-regulation

<u>Corporate Narrative</u>

12. Are corporate representatives directly quoted or referenced in the article? [No=0, Yes=1]

13. Are the following concepts mentioned as a consequence of the event? [No=0, Yes=1]
a. Corporations acquiring an overall milder attitude
b. Acknowledgement of need for further corporate ethics

c. Change of business models
d. Need for consumers to improve their knowledge on data use.

14. If present, are the items above mentioned by the article as a critique or as a reproduction of corporate discourse? [No=0, Yes=1]
a. Critique
b. Reproduction

15. Could the piece be classified as a rant of as a deep level discussion on the topic? [No=0, Yes=1]
a. Controversial
b. Deep level discussion

16. Did the journalist depend on his source for expert analysis and interpretation of the events? [No=0, Yes=1]

- <u>VERSION 2</u>

<u>Explanatory Variables</u>

1. ID number (ID) [start with 1 and number onwards]

2. Media in which article is published (Media)
   1. BBC News
   2. The Guardian
   3. Daily Mail
   4. Huffington Post
   5. Sky News

6. MSN News
7. The Telegraph
8. MIT Tech Review
9. Wired
10. Fast Company
11. Gizmodo
12. TechCrunch
13. Recode

3. Search Criteria (Search)
1. Bias+ algorithm
1. Bias + algorithms
1. Bias + artificial intelligence
2. Bias + AI
3. Bias + Machine learning

4. URL [Open]

3. Headline[Open]

4. Date of publication (Date) [day/month/year]

5. Who is the author of the article? (author) [Open]

6. Was the article published after the Cambridge Analytica scandal (March 17)? (after_CA) [No=0, Yes=1]

7. What kind of publication is it? (kind)

1. Mainstream Media - UK consumption
2. Technology publications

<u>Algorithmic Bias (AB) Narratives</u>

8. What is the overall posture of the article towards algorithmic bias?
   1. Very Negative (a_VeryNegative)
   2. Negative (b_Negative)
   3. Neutral (c_neutral)
   4. Positive (d_positive)
   5. Very Positive (e_VeryPositive)

9. Is the algorithmic bias issue the main topic of the article? (Main_topic) [No=0, Yes=1]

10. Is either of these general framings more evident in the article? (Algo_ReduceReinforce)
    1. Algorithms can reduce biases
    2. Algorithms reinforce biases
    3. N/A

11. Are any of the following CRITICAL framings towards ‘algorithmic bias’ present in the article? [No=0, Yes=1] (more than one option possible)

    1. The issue of AB evidences the lack of corporate ethics and accountability (a_Ethics)
    2. Algorithms are “black boxes” that should be more transparent for the understanding of the publics. (b_BlackBox)
    3. Algorithms are ‘Echo-chambers’ that enhance existing biases. (c_EchoChambers)
    4. The AB issue is related to the lack of antitrust legislation. (d_Antitrust)
    5. There is no way to mitigate ABs (algorithms will always be biased). (e_AlwaysBiased)

6. Because AB exists, algorithms always need to be controlled by humans. (f_human)
7. Other (g_Other)

12. Are any of the following social consequences of AB highlighted? [No=0, Yes=1] (more than one option possible)

1. The dissemination of ‘fake news’ (a_Fakenews)
2. Political interference (algorithms have influenced electoral processes) (b_PolInter)
3. Negative effects on personnel selection processes (c_personnel)
4. Negative effects on judicial/police decisions (d_JudPol)
5. Negative effects on welfare allocations (e_welfare)
6. Negative effects on credit scoring (f_credit)
7. Negative effects on education systems (g_educ)
8. Gender discrimination (h_gender)
9. Ethnic discrimination. (i_ethnic)
10. Negative effects on privacy (j_Privacy)
11. Other [mention] (k_Other)

13. Are any of the following NON-CRITICAL framings related ‘algorithmic bias’ present in the article? [No=0, Yes=1] (more than one option possible)

1. Algorithms are not biased (a_BiasDenial)
2. Algorithms are not inherently biased, they only reflect pre-existing biases in societies. (b_reflection)
3. External regulation implies profitability issues (more regulation leading to less interest in innovation investment, therefore less possibilities to innovate remove the bias from product, no chances to improve product imply a decline in profitability for corporations) (c_ProfIssue)
4. Innovation investment is key to get rid of algorithmic bias(d_InnoInnvest)

5. Self-regulation is important to improve algorithmic bias and therefore preserve user trust. (e_UserTrust)
6. Self-regulation is important to prevent external regulation. (f_SelfReg_External)
7. External regulation is too slow to catch up with innovation (g_Reg_Slow)
8. Other (h_Other)

14. Does the article make any of the following suggestions to deal with the issue of ‘Algorithmic Biases’? (more than one option possible) [No=0, Yes=1]

1. Corporations need higher ethical standards to prevent algorithmic bias. (a_MoreEthics)
2. Corporations should be more diverse to mitigate algorithmic bias (b_diversity)
3. Corporations should be more transparent in order to detect algorithmic bias (c_MoreTrans)
4. Corporations should change their business models because of algorithmic bias. (d_ChangeModels)
5. Corporations should promote user understanding on how their algorithms work. (e_ConsumerKnow)
6. Other (f_other)

<u>Sector Narratives</u>

15. Is Cambridge Analytica the main topic of the article (CA_MainTopic)

16. Is Cambridge Analytica mentioned in the article? (CA_Mention) [No=0, Yes=1]

17. Does the article mention any of these companies: Google, Facebook (or CA) Twitter, Apple or Amazon? (MainCompanies) [No=0, Yes=1]

18. Are any of these industries/organisations mentioned as users of biased algorithms? (do they use an algorithm that is biased?) (more than one possible) [No=0, Yes=1]

1. Medical (a_medical)
2. Social Media (b_SoMe)
3. Finance (c_Fin)
4. Governments (d_Gov)
5. Legal/Police System (e_legal)
6. Education (f_EducSys)
7. Search Platform Companies (g_SearchPlat)
8. Other [name industry] (h_other)

19. Is there a direct reference in the article to algorithms benefiting advertising revenue for corporations? (AdRevenue) [No=0, Yes=1]

<u>Regulation Narratives</u>

20. Does the article make any reference that could be related to any of these postures towards regulation? (search for keywords regulation, audit, control or monitoring in the text) [No=0, Yes=1]

1. Pro- External regulation (a_ProReg)
2. Anti-regulation (b_AntiReg)
3. Self-regulation (companies must regulate themselves) (c_SelfReg)
4. Multi-Stakeholder Regulation (stakeholders participate in the dialogue together) (d_MultiReg)
5. Mixed (mentions both self-regulation and external regulation) (e_Mixed)
6. Does not mention regulation (f_NoMention)

21. Are governments/parliaments mentioned in any part of the article? (GovPar_Mention)

22. Are representatives of governments/parliaments directly quoted? (GovPar_quote) [No=0, Yes=1]

23. Does the article paraphrase something government/parliaments stated? (no direct from specific person quote but referring to something governments/parliaments stated) (GovPar_ref) [No=0, Yes=1]

<u>Corporate Narratives</u>

24. Are any of these ideas presented in relation to algorithms in general? [No=0, Yes=1]:
    1. Algorithms are empowering and beneficial to users (a_Empower)
    2. Algorithms are cost/effective for companies (b_CostEffective)

25. Are corporations mentioned in any part of the article? (Corpo_Mention) [No=0, Yes=1]

26. Are any corporate representatives directly quoted (Corpo_Quote) [No=0, Yes=1]

27. Does the article paraphrase something corporations satated? (no direct from specific person quote but references to general statements made by corporations) (Corpo_Ref) [No=0, Yes=1]

28. Does the article mention corporations engaging in any of the following attitudes towards the issue? (more than one option possible) [No=0, Yes=1]
    1. Corporations are trying to improve the algorithmic bias (a_CorpImprove)
    2. Corporations are recognising they made mistakes with algoirithmic bias (b_CorpoEthics)
    3. Corporations are changing their business models because of algorithmic bias. (c_CorpoModels)

4. Corporations are promoting user understanding on how their algorithms work. (d_ConsumerKnowledge)
5. Corporations directly opposing external regulation to algorithmic bias (e_Corpo_NoReg).
6. Corporations are showing poor responses towards the algorithmic bias issue (f_PoorResponses)
7. Other (f_other)

<u>Technical Narratives</u>

29. Is the term ‘algorithmic bias’ explicitly mentioned? (AlgoBias_mention)

30. Could the piece be classified as a polemical discussion on the topic (versus technical or ongoing debate)? (polemical) [No=0, Yes=1]

31. Did the journalist depend on his/her source on computer/technical experts? (sources) [No=0, Yes=1]

32. Does the article give a technical explanation for the bias? [No=0, Yes=1] (technical)

33. For Mainstream Media, is the article in the technology section? [No=0, Yes=1] (TechSection)

APPENDIX 3 – RATIONALE FOR CODE DESIGN (IN BOXES) AND SUMMARY ANALYSIS

Explanatory Variables

1. ID number (ID) [start with 1 and number onwards]

Short ID for identification of articles

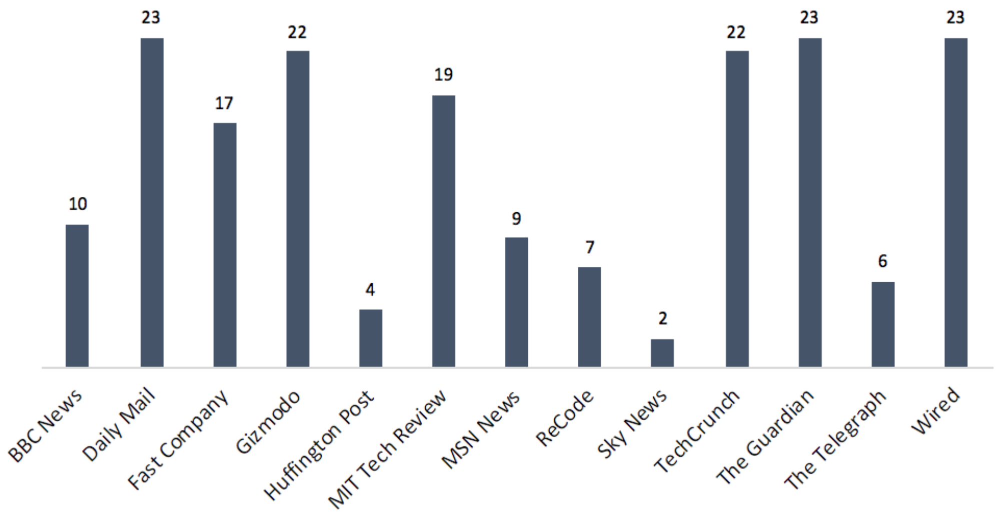
2. Media in which article is published (Media)

Short ID for name of media

Table 1 – Media - Frequency by media

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	BBC News	10	5.3	5.3	5.3
	Daily Mail	23	12.3	12.3	17.6
	Fast Company	17	9.1	9.1	26.7
	Gizmodo	22	11.8	11.8	38.5
	Huffington Post	4	2.1	2.1	40.6
	MIT Tech Review	19	10.2	10.2	50.8
	MSN News	9	4.8	4.8	55.6
	ReCode	7	3.7	3.7	59.4
	Sky News	2	1.1	1.1	60.4
	TechCrunch	22	11.8	11.8	72.2
	The Guardian	23	12.3	12.3	84.5
	The Telegraph	6	3.2	3.2	87.7
	Wired	23	12.3	12.3	100.0
	Total	187	100.0	100.0	

Figure 1 – Media - Frequency of articles by media



3. Search Criteria (Search)

Short ID for easy identification of search criteria

1. Bias+ algorithm
2. Bias + algorithms
3. Bias + artificial intelligence
4. Bias + AI
5. Bias + Machine learning

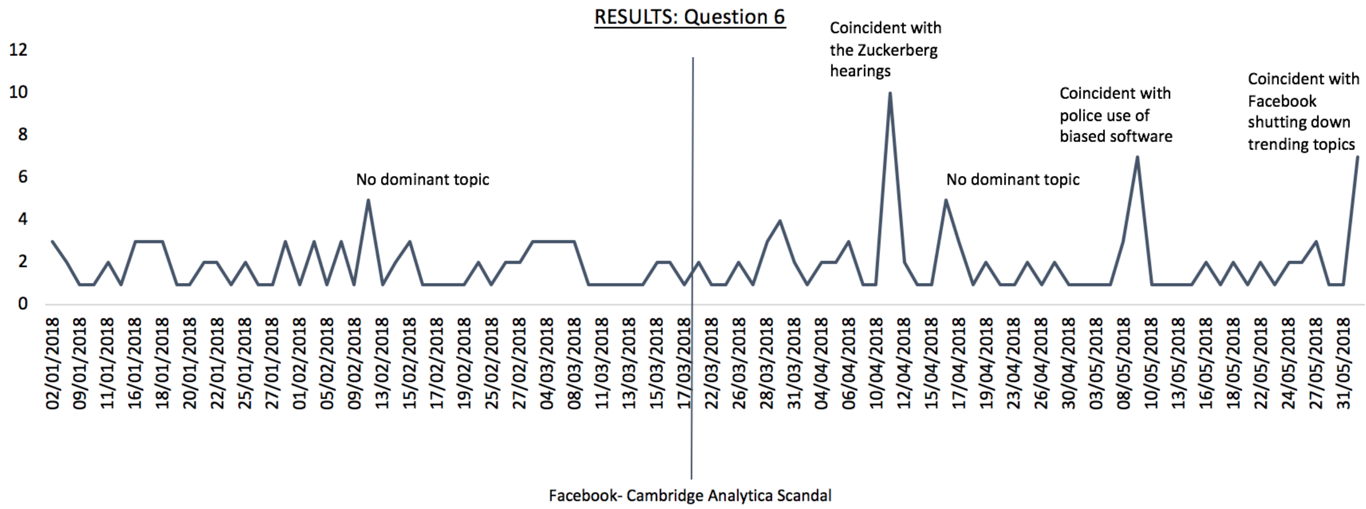
4. URL [Open]

5. Headline [Open]

6. Date of publication (Date) [day/month/year]

Intended to detect any possible peaks in article publications

Figure 2 – Date - Patterns of publication by date



7. Who is the author of the article? (Author) [Open]

Intends to detect possible bias due to non-diversity of authors.

RESULTS: non-significant

8. Was the article published after the Cambridge Analytica scandal (March 17)? (after_CA)

[No=0, Yes=1]

The relevance of this question is directly related to the RQ, this code is fundamental to answer if or how each type of publication has covered the algorithmic bias topic before and after the scandal. It will be used as main Independent Variable for statistical analyses.

Table 2 – After_CA - Coverage by kind of media - Before and After

		after_CA 0	after_CA 1	Total
kind	0	26	50	76
	1	61	50	111
Total		87	100	187

Figure 3 – After_CA - Coverage - Before and After

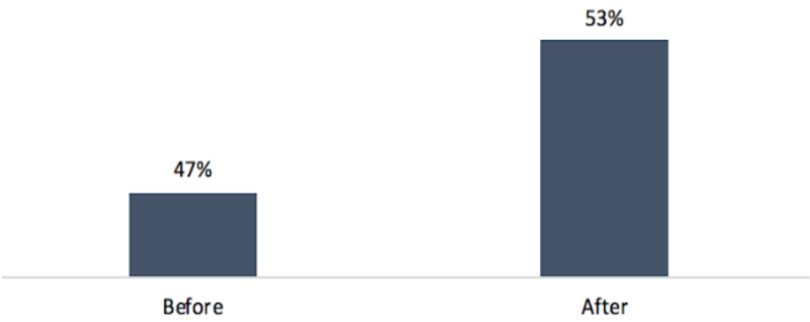
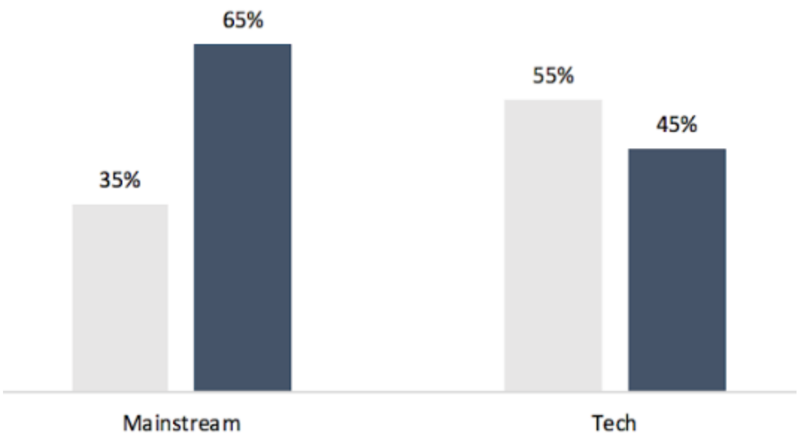


Figure 4 – After_CA - Comparative coverage by types of media



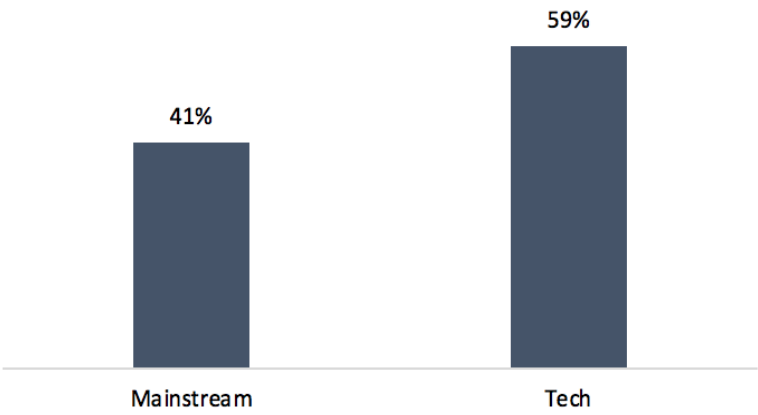
9. What kind of publication is it? (kind)

1. Mainstream Media - UK consumption [0]
2. Technology publications [1]

Table 4 – Kind - Frequencies and Percentages by type of media

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	77	41.2	41.2	41.2
	1	110	58.8	58.8	100.0
	Total	187	100.0	100.0	

Figure 5 – Kind - Percentage of coverage type of media



DISCOURSES

10. What is the overall posture of the article towards algorithmic bias? (From 1-5) (Posture)

This code is central, it will give the most important indicator towards the overall framing of each of the articles regarding the topic: from a highly critical perspective to ones aligned with the dominant discourse.

1. **Very Negative**

An article that could be considered as representative of a highly critical and contestatory position towards the issue of bias, aligned with the terms and lexicon present in the critical literature.

2. **Negative**

An article that frames the issue of bias according to the concepts related to the critical literature but that does not make heavy use of critically loaded lexicon towards the issue of bias.

3. **Neutral**

An article that acknowledges the issue of bias, but it does not present a clear posture towards it.

4. **Positive**

The issue is acknowledged by the article but presents quite optimistic prospects for solving the situation in the future.

5. **Very Positive**

The algorithmic bias issue is mentioned by the article but not highlighted as a negative feature of algorithms. Overall the article could be aligned with the ‘the dominant discourse of ‘algorithmic objectivity’, presenting highly optimistic prospects for the use of algorithms.

Table 5 – Posture - Overall Frequencies and Percentages

N	Valid	187
	Missing	0
Mean		2.61
Median		2.00
Mode		1
Range		4
Percentiles	25	1.00
	50	2.00
	75	4.00

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	64	34.2	34.2	34.2
	2	35	18.7	18.7	52.9
	3	30	16.0	16.0	69.0
	4	25	13.4	13.4	82.4
	5	33	17.6	17.6	100.0
	Total	187	100.0	100.0	

Table 6 - Posture - Overall percentages – before and after

Posture	Before	After
1	36%	33%
2	18%	19%
3	11%	20%
4	13%	14%
5	22%	14%
Grand Total	100%	100%

Figure 6 - Posture - Comparative percentages by type of media

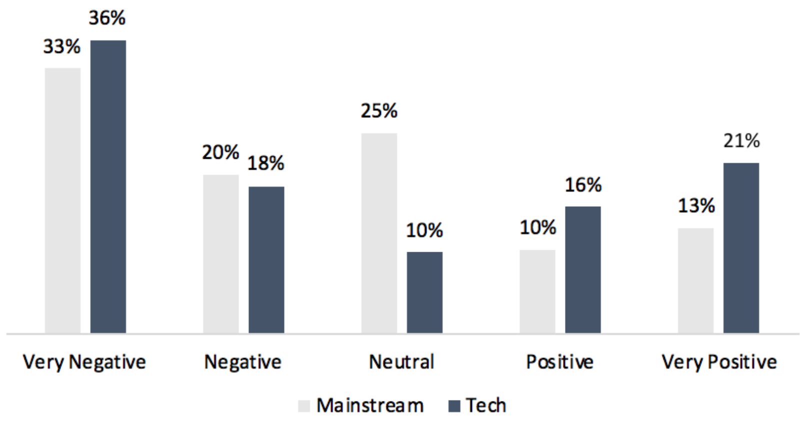


Table 7 - Posture – Mainstream Publications – Frequencies and Percentages

N	Valid	77
	Missing	0
Mean		2.52
Median		2.00
Mode		1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	25	.0	32.5	32.5
	2	15	.0	19.5	51.9
	3	19	.0	24.7	76.6
	4	8	.0	10.4	87.0
	5	10	.0	13.0	100.0
	Total	77	.0	100.0	

Table 8 - Posture – Tech Publications – Frequencies and Percentages

N	Valid	110
	Missing	0
Mean		2.68
Median		2.00
Mode		1

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	39	35.5	35.5	35.5
	2	20	18.2	18.2	53.6
	3	11	10.0	10.0	63.6
	4	17	15.5	15.5	79.1
	5	23	20.9	20.9	100.0
	Total	110	100.0	100.0	

11. Is either of these general framings more evident in the article?

(Algo_ReduceReinforce) [No=0, Yes=1]

1. Algorithms can reduce biases

The category is related to the dominant discourse that venerates algorithmic objectivity and the benefits of the technology.

2. Algorithms reinforce biases [1]

The category is related to the critical social sciences literature that focuses on the social consequences of algorithms and the way in which they might carry biases.

3. N/A

Does not present clear posture.

Table 9 – Algo_ReduceReinforce - Overall Frequencies and Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	28	15.0	15.0	15.0
	1	147	78.6	78.6	93.6
	N/A	12	6.4	6.4	100.0
	Total	187	100.0	100.0	

Table 10 - Algo_ReduceReinforce - Overall percentages – Before and After

	Before	After
Algorithms Reinforce Bias	75%	82%

Table 11 – Algo_ReduceReinforce - Mainstream

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	10	13.0	13.0	13.0
	1	60	77.9	77.9	90.9
	N/A	7	9.1	9.1	100.0
	Total	77	100.0	100.0	

Table 12 – Algo_ReduceReinforce - Technology

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	18	16.4	17.1	17.1
	1	87	79.1	82.9	100.0
	Total	105	95.5	100.0	
Missing	System	5	4.5		
Total		110	100.0		

12. Are governments/parliaments mentioned in any part of the article?

(GovPar_Mention) [No=0, Yes=1]

13. Are representatives of governments/parliaments directly quoted? (GovPar_quote) [No=0, Yes=1]

14. Does the article paraphrase something government/parliaments stated? (no direct from specific person quote but referring to something governments/parliaments stated) (GovPar_ref) [No=0, Yes=1]

Codes 12, 13 and 14 relate to possible references to governmental discourse in the discussion around algorithmic bias. Given the reading of the critical literature, they relevant as the mention of governments and regulation in the discussion is considered as an indicator that the dominant corporate discourse is being contested.

Table 12 - GovPar_Mention – GovPar_Quote – GovPar_Ref - Overall percentages

GovPar_Mention	GovPar_quote	GovPar_ref
35%	22%	24%

Table 13 - GovPar_quote – Before and After by media type

	GovPar_Quote	
	Before	After
Mainstream	11%	38%
Technology	15%	22%

Figure 7 -GovPar_Mention – GovPar_Quote – GovPar_Ref - Overall percentages

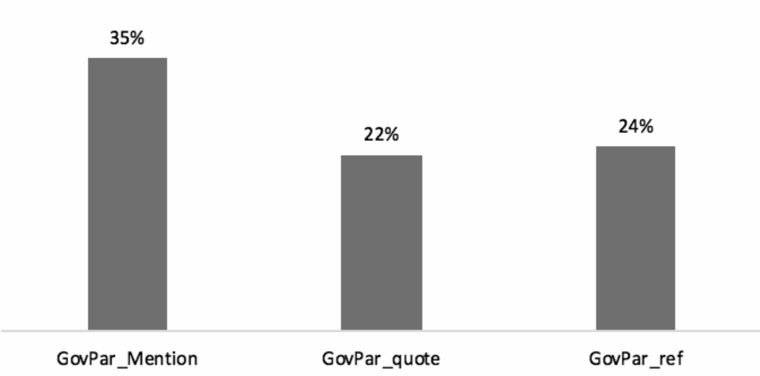
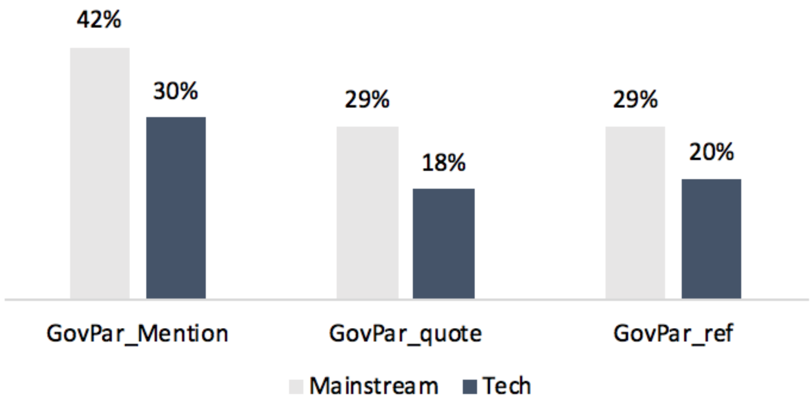


Figure 8 - GovPar_Mention – GovPar_Quote – GovPar_Ref – Percentages by media type



15. Are any of these ideas presented in relation to algorithms in general?

1. Algorithms are empowering and beneficial to users (a_Empower)

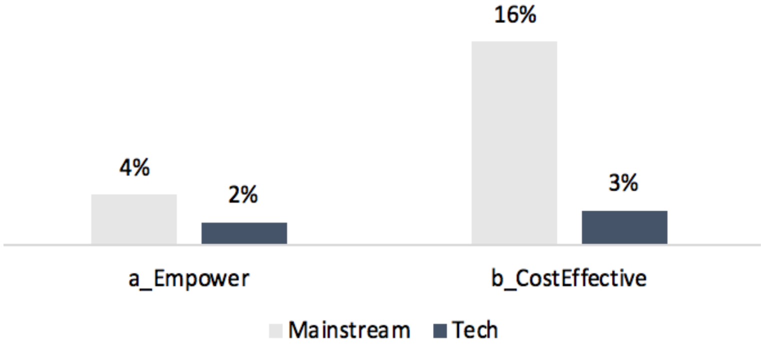
2. Algorithms are cost/effective for companies (b_CostEffective)

These two arguments presented in the code correspond to the dominant discourse of corporations and they are included with the aim to detect the reproduction of such discourse. They are categorised as dummies for more flexibility and relatability of codes.

Table 14 - a_Empower – b_CostEffective - Overall percentages

a_Empower	b_CostEffective
3%	8%

Figure 9 - a_Empower – b_CostEffective - Percentages by media type



16. Are corporations mentioned in any part of the article? (Corpo_Mention) [No=0, Yes=1]

17. Are any corporate representatives directly quoted (Corpo_Quote) [No=0, Yes=1]

18. Does the article paraphrase something corporations stated? (no direct from specific person quote but references to general statements made by corporations) (Corpo_Ref) [No=0, Yes=1]

Codes 16, 17 and 18 are designed with the intention to determine the prominence and reproduction of corporate discourse, in relation to others, helping answer SQ-a.

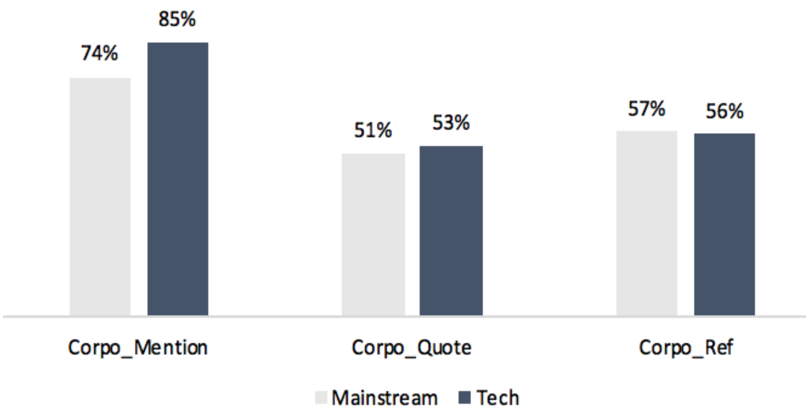
Table 15 - Corpo_Mention – Corpo_Quote – Corpo_Ref - Overall percentages

Corpo_Mention	Corpo_Quote	Corpo_Ref
80%	52%	57%

Table 16 - Corpo_Quote - Before and After

	Corpo_Quote	
	Before	After
Mainstream	41%	56%
Technology	52%	54%

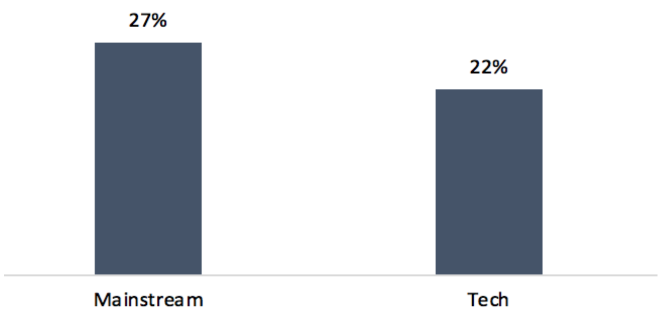
Figure 10 - Corpo_Mention – Corpo_Quote – Corpo_Ref - Percentages by media type



19. Does the article mention corporations’ efforts to solve algorithmic biases? (CorpoEffort) [No=0, Yes=1]

The presence of an acknowledgement towards the corporate efforts to improve the algorithmic bias issue is seen as a form of reproduction of the dominant corporate discourse.

Figure 11 - CorpoEffort - Percentages by media type



20. Did the journalist depend on his/her source on computer/technical experts? (sources) [No=0, Yes=1]

21. Does the article give a technical explanation for the bias? (technical) [No=0, Yes=1]

22. Is the article in the technology/science section? (TechSection) [No=0, Yes=1]

Codes 20, 21 and 22 are designed with the intention to determine the prominence and reproduction of technical discourse, in relation to others, helping answer SQ-a and with to help answer the hypothesis related to the technology framing literature that mainstream media presents low levels of coverage on tech topics.

Figure 12 - Sources – Technical – TechSection – Overall percentages

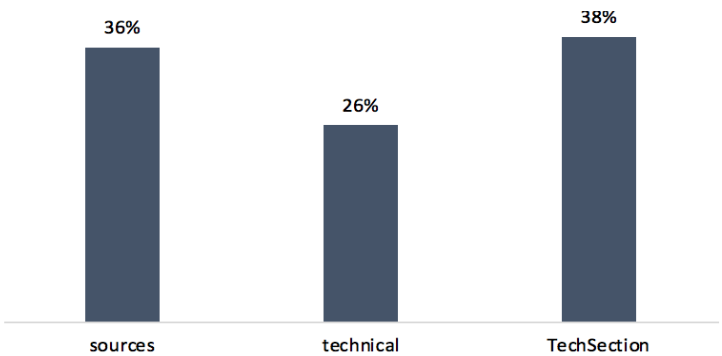
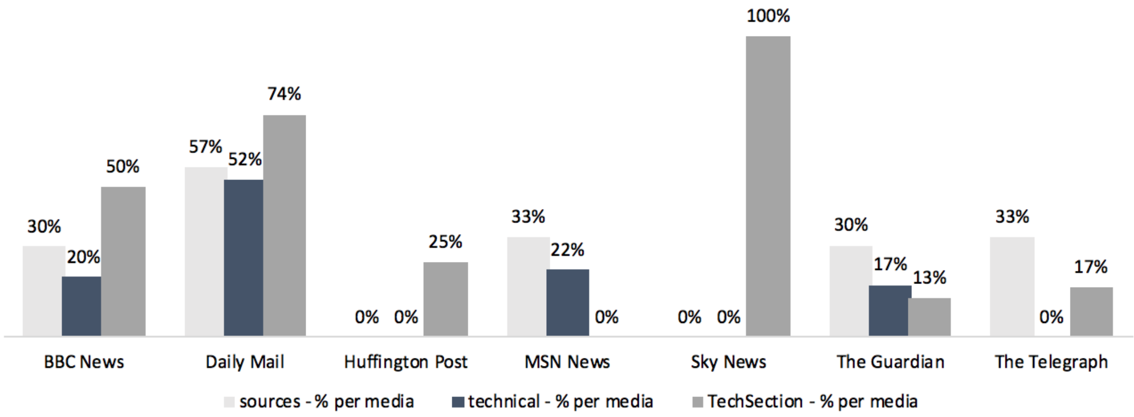


Table 17 - Sources – Technical – TechSection – Percentages by media

	count of articles	sources count	sources %	technical count	technical %	TechSection count	TechSection %
BBC News	10	3	30%	2	20%	5	50%
Daily Mail	23	13	57%	12	52%	17	74%
Huffington Post	4	0	0%	0	0%	1	25%
MSN News	9	3	33%	2	22%	0	0%
Sky News	2	0	0%	0	0%	2	100%
The Guardian	23	7	30%	4	17%	3	13%
The Telegraph	6	2	33%	0	0%	1	17%
Grand Total	77	28	36%	20	26%	29	38%

Figure 13 - Sources – Technical – TechSection – Percentages by media



MAIN TOPICS

23. Are any of the following framings towards ‘algorithmic bias’ present in the article? (issues) (select more than one option)

This code proposes different framings related to arguments that belong to the critical literature and dominant corporate discourse. It is meant to identify which are the arguments are most present in the media framings. These results will help answer the sub-question referring to the topics most present on the discussion on algorithmic bias. The different categories are coded as dummies as more than one option are possible to appear in each article.

1. **The issue of algorithmic bias shows the lack of corporate ethics** (a_Ethics) [No=0, Yes=1]

Related to critical literature that refers to how the industry should act in order to avoid social consequences of technologies such as algorithmic bias.

2. **The issue of algorithmic bias shows the lack of corporate accountability** (b_accountability) [No=0, Yes=1]

Related to critical literature that refers to how the industry take responsibility for pervasive effects of technologies such as algorithmic bias.

3. **Algorithms are “black boxes” that should be more transparent for the understanding of the publics**. (c_BlackBox) [No=0, Yes=1]

Related to critical literature that refers to how algorithms by design are inscrutable tools not easy to interpret for analysis.

4. **Algorithms reproduce and enhance content biases for user** (d_EchoChambers) [No=0, Yes=1]

Related to critical literature that refers to how algorithms can reproduce existing views amongst clusters of users and might generate cleavages and socio-political division as consequence.

5. **The algorithmic bias issue is related to the lack of antitrust legislation** (e_Antitrust) [No=0, Yes=1]

Related to critical literature that argues that because there is no regulation for fair competition in the industry, tech corporations do not show a concern to improve social consequences of their technologies.

6. **There is no way to mitigate algorithmic biases** (algorithms will always be biased) (f_AlwaysBiased) [No=0, Yes=1]

Related to critical literature that argues algorithms are a pernicious technology that, because of the way they are conceived, they inherently present biases that are not to be solved as they are not technological issue but rather social.

7. **Algorithms reflect pre-existing biases in societies** (i_reflection) [No=0, Yes=1]

Related to critical literature that contends that biases are introduced to algorithms as they are designed, because they reflect preconceptions already present in the broader society and institutions that produce them. It reflects an acknowledgment of the issue but it is less deterministic than the item above.

8. **Algorithms always need to be controlled by humans** (g_human) [No=0, Yes=1]

Related to critical literature that contends that because algorithms tend to reproduce and enhance bias, they cannot be ‘left alone’ to work in a fully automated way without human monitoring.

9. **Algorithms are not biased** (h_BiasDenial) [No=0, Yes=1]

Related to the dominant discourse that contends that because algorithms are purely rational and objective tools, they cannot be biased.

10. **Government regulation reduces profitability** (j_ProfIssue) [No=0, Yes=1]

Related to the dominant discourse that argues that external regulation leads to less interest in innovation investment, therefore less possibilities to innovate remove the bias from product, no chances to improve product imply a decline in profitability for corporations.

11. **Innovation investment is key to get rid of algorithmic biases** (k_InnoInnvest) [No=0, Yes=1]

Related to the dominant discourse that argues that without more innovation in algorithmic technology the mitigation of bias issues will not be possible.

12. **Corporate self-regulation is important to preserve user trust** (l_UserTrust) [No=0, Yes=1]

Related to the dominant discourse that contends that tech companies should find ways to regulate themselves in order to prevent issues such as bias that might deter users from using their products.

13. **Corporate self-regulation is important to prevent government regulation** (m_SelfReg_External) [No=0, Yes=1]

Related to the dominant discourse that contends that external governmental regulation is pernicious as it might affect innovation, therefore, in order to show commitment and avoid external parties regulating tech business, corporations need to establish regulations on themselves.

14. **Government regulation is too slow to catch up with innovation** (n_Reg_Slow) [No=0, Yes=1]

Related to the dominant discourse that argues that because governmental regulation moves at an extremely slow burocratic pace, it will never be able to catch up with the rhythm of innovation in the tech industry. Therefore, imposing government regulation implies imposing regulation that is inherently behind new developments.

Figure 14 - Issues - Overall count of mentions

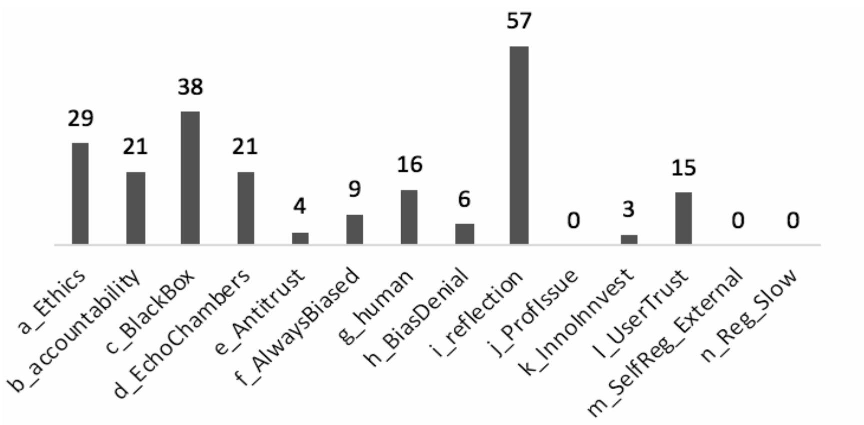


Figure 15 - Issues - overall count of mentions – Before and After

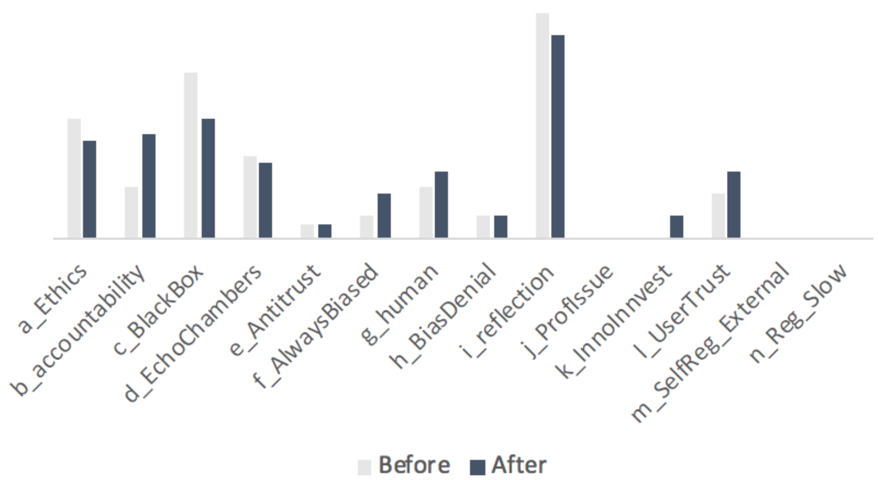
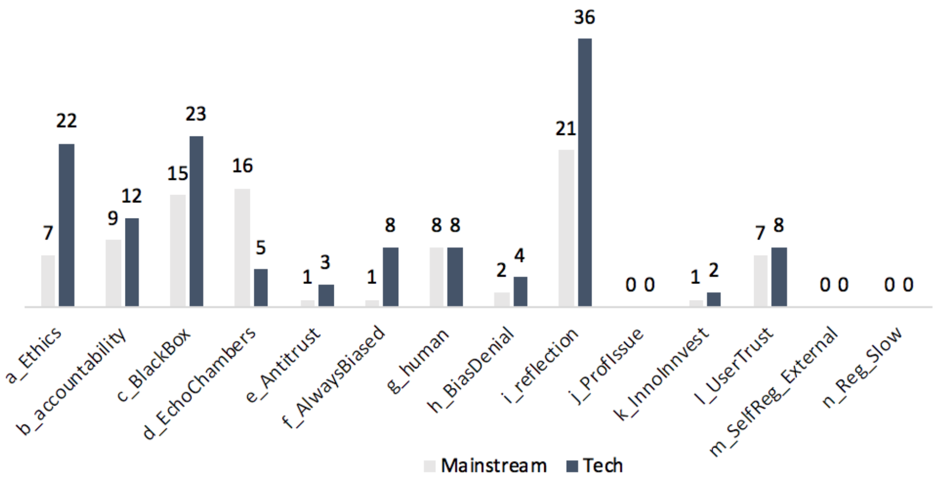


Figure 16 - Issues - overall count of mentions by type of media



24. Are any of the following consequences of algorithmic bias discussed? (effects) (more than one option possible)

All the possible options proposed by this code are taken from the critical literature, that focus both on the ‘gatekeeping’ and ‘discriminatory’ effects of algorithms. This will help answer SQ-2: what are the most salient issues in the discussion on algorithmic bias? The categories are coded as dummies for more flexible interpretation of results.

1. The dissemination of 'fake news' (a_Fakenews) [No=0, Yes=1]
2. Negative effects on political life (b_PolInter) [No=0, Yes=1]
3. Negative effects on personnel selection processes (c_personnel) [No=0, Yes=1]
4. Negative effects on judicial decisions (d_Jud) [No=0, Yes=1]
5. Negative effects on police decisions (e_pol) [No=0, Yes=1]
6. Negative effects on welfare allocations (f_welfare) [No=0, Yes=1]
7. Negative effects on credit scoring (g_credit) [No=0, Yes=1]
8. Negative effects on education systems (h_educ) [No=0, Yes=1]
9. Gender discrimination (i_gender) [No=0, Yes=1]
10. Ethnic discrimination (j_ethnic) [No=0, Yes=1]
11. Negative effects on privacy (k_Privacy) [No=0, Yes=1]
12. Other (l_other) [open]

Figure 17 - Effects - Overall count of mentions

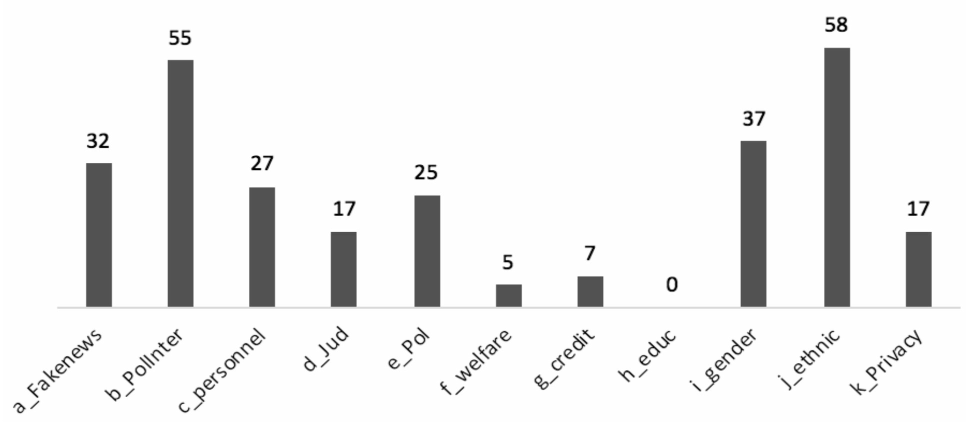


Figure 18 - Effects - Overall count of mentions – Before and After

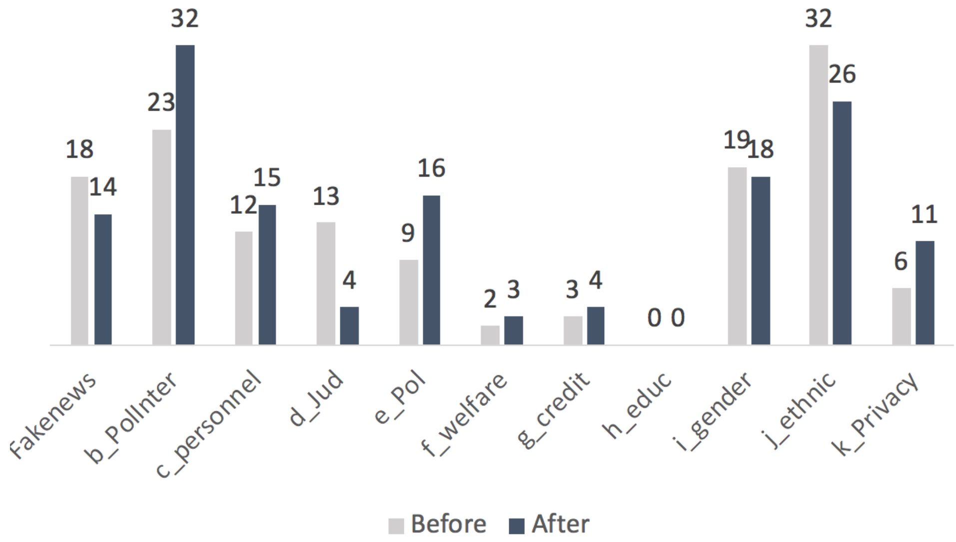
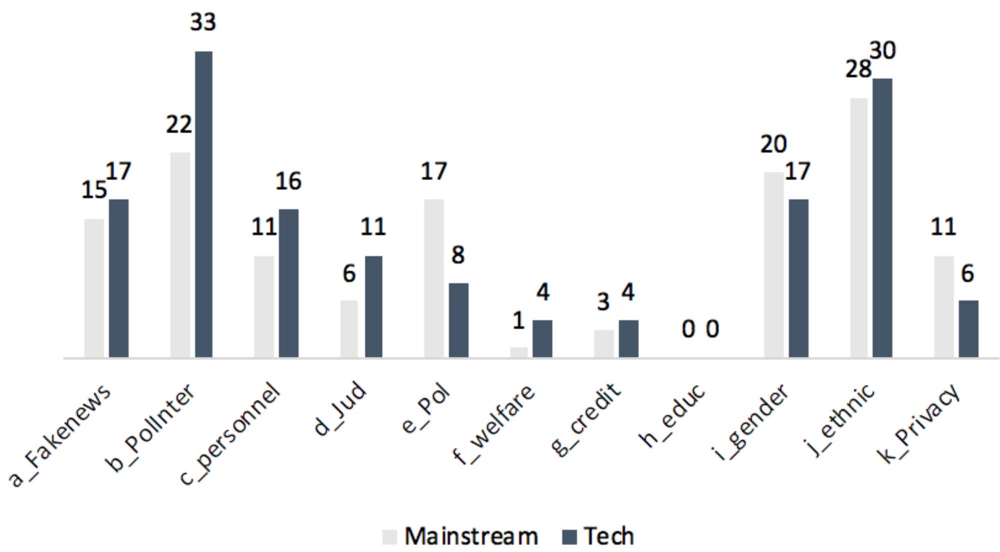


Figure 19 - Effects - Overall count of mentions by type of media



25. Does the article make any of the following suggestions to deal with the issue of ‘Algorithmic Biases’? (suggestions) (more than one option possible)

All the possible options proposed by this code are taken from suggestions present in the critical literature regarding how to deal with the issue of bias. The code is intended to detect the prominence of reproduction of critical arguments. The categories are coded as dummies for more flexibility in interpretation as more than one option might be present in the articles.

1. Corporations need higher ethical standards to prevent algorithmic bias (a_MoreEthics) [No=0, Yes=1]
2. Corporations should be more diverse to mitigate algorithmic bias (b_diversity) [No=0, Yes=1]
3. Corporations should be more transparent in order to detect algorithmic bias (c_MoreTrans) [No=0, Yes=1]
4. Corporations should change their business models because of algorithmic bias (d_ChangeModels) [No=0, Yes=1]
5. Corporations should promote user understanding on how their algorithms work. (e_ConsumerKnow) [No=0, Yes=1]

Figure 20 - Suggestions - Overall count of mentions

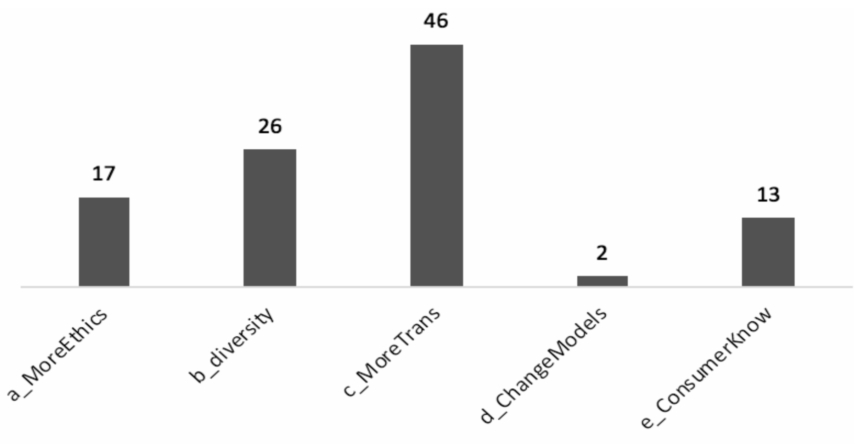
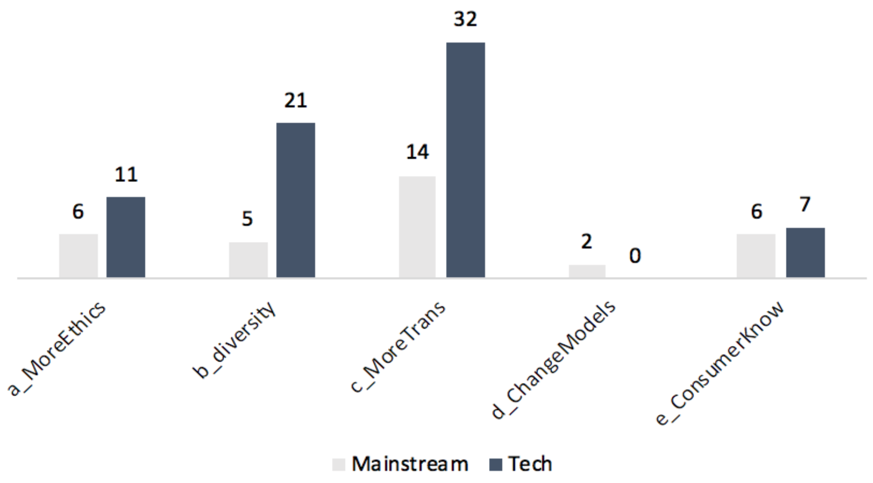


Figure 21 - Suggestions - Overall count of mentions by type of media



26. Is Cambridge Analytica the main topic of the article (CA_MainTopic) [No=0, Yes=1] *()*

This code is intended to help answer the RQ, evaluating whether the Cambridge Analytica scandal might have acted as a catalyst of publications on the algorithmic bias topic.

Table 18 - CA_MainTopic - Overall frequencies and Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	182	97.3	97.3	97.3
	1	5	2.7	2.7	100.0
	Total	187	100.0	100.0	

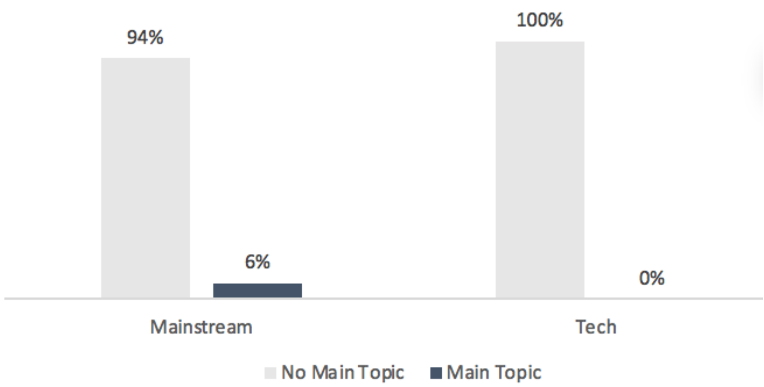
Table 19 - CA_MainTopic - Mainstream - frequencies and Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	72	93.5	93.5	93.5
	1	5	6.5	6.5	100.0
	Total	77	100.0	100.0	

Table 20 - CA_MainTopic - Technology - frequencies and Percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	110	100.0	100.0	100.0

Figure 22 - CA_MainTopic - comparative by type of media



27. Is the Cambridge Analytica scandal referenced in the article? (CA_Mention) [No=0, Yes=1]

> This code is intended to help answer the RQ, evaluating whether the Cambridge Analytica scandal might have acted as a catalyst of publications on the algorithmic bias topic.

Table 21 - CA_Mention – Overall frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	162	86.6	86.6	86.6
	1	25	13.4	13.4	100.0
	Total	187	100.0	100.0	

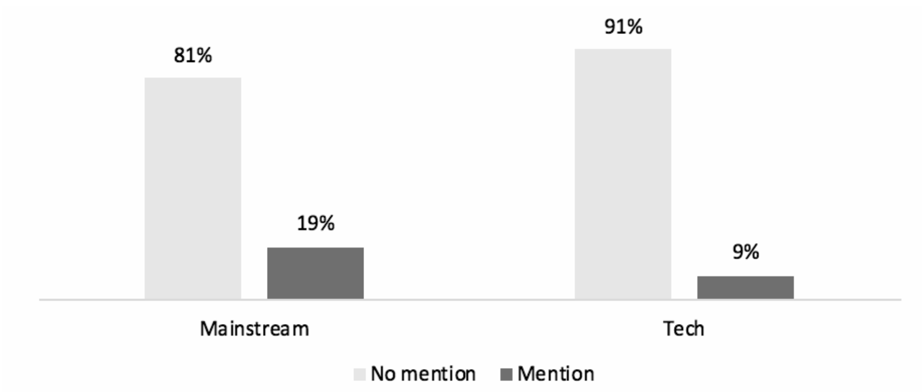
Table 22 - CA Mention – Mainstream - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	62	80.5	80.5	80.5
	1	15	19.5	19.5	100.0
	Total	77	100.0	100.0	

Table 23 - CA Mention – Technology - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	100	90.9	90.9	90.9
	1	10	9.1	9.1	100.0
	Total	110	100.0	100.0	

Figure 23 - CA Mention – comparative percentages by type of media



28. Is the algorithmic bias issue the main topic of the article? (Main_topic) [No=0, Yes=1]

This code is intended to help answer the RQ, to indicate whether the algorithmic bias issue is discussed in depth or not, relating to the literature that claims mainstream media tends to discuss tech topics controversially.

Table 24 - MainTopic – Overall frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	109	58.3	58.3	58.3
	1	78	41.7	41.7	100.0
	Total	187	100.0	100.0	

Table 25 - CA Mention – Overall – Frequencies Before and After

	Before	After	Grand Total
Algorithmic Bias Main Topic	39	39	78

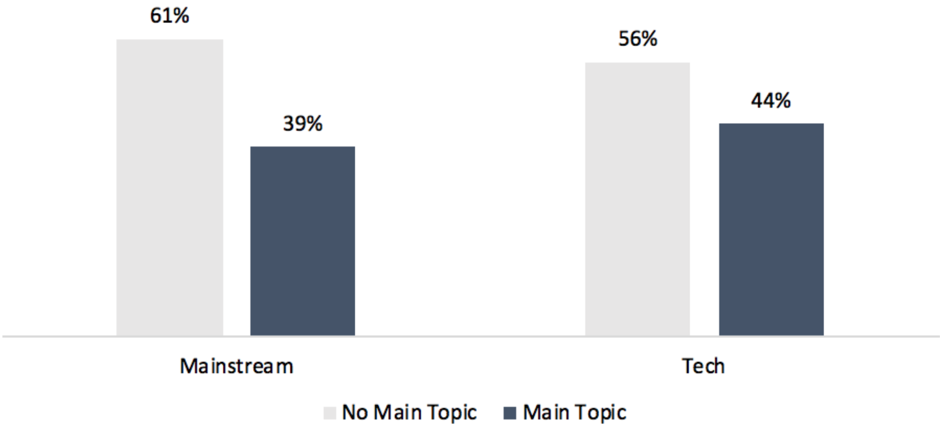
Table 26 - MainTopic – Mainstream - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	47	.0	61.0	61.0
	1	30	.0	39.0	100.0
	Total	77	.0	100.0	

Table 27 - MainTopic – Technology- frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	62	56.4	56.4	56.4
	1	48	43.6	43.6	100.0
	Total	110	100.0	100.0	

Figure 24 - CA Mention – comparative percentages by type of media



29. Does the article mention any of these companies: Google, Facebook (or CA) Twitter, Apple or Amazon? (MainCompanies) [No=0, Yes=1]

This code is aimed at identifying whether the media framing associates the issue of algorithmic bias to Big Tech - this code might come useful if regulation themes are present.

Table 28 - MainCompanies – Overall frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	81	43.3	43.3	43.3
	1	106	56.7	56.7	100.0
	Total	187	100.0	100.0	

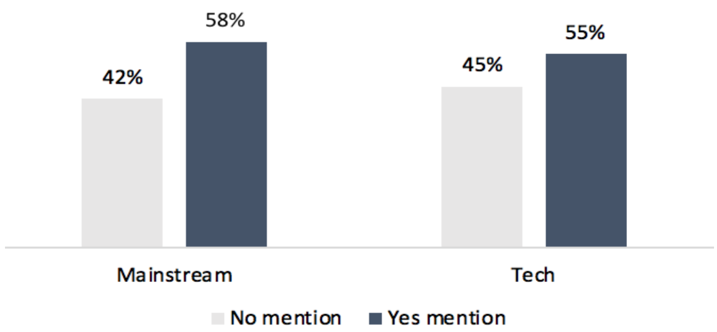
Table 29 - MainCompanies – Mainstream - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	32	41.6	41.6	41.6
	1	45	58.4	58.4	100.0
	Total	77	100.0	100.0	

Table 30 - MainCompanies – Technology - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	49	44.5	44.5	44.5
	1	61	55.5	55.5	100.0
	Total	110	100.0	100.0	

Figure 25 - MainCompanies – comparative percentages by type of media



30. Are any of the below mentioned as users of biased algorithms? (do they use an algorithm that is biased?) (more than one possible) [No=0, Yes=1]

The code is designed to identify which industries are more related to the use of biased algorithms in case regulation themes are present.

1. Medical (a_medical) [No=0, Yes=1]
2. Social Media (b_SoMe) [No=0, Yes=1]
3. Finance (c_Fin) [No=0, Yes=1]
4. Governments (d_Gov) [No=0, Yes=1]
5. Legal (e_legal) [No=0, Yes=1]
6. Police System (f_police) [No=0, Yes=1]
7. Education (g_EducSys) [No=0, Yes=1]
8. Search Platform Companies (h_SearchPlat) [No=0, Yes=1]
9. Other (other_i)[open]

Figure 26 - Users - Overall count of mentions

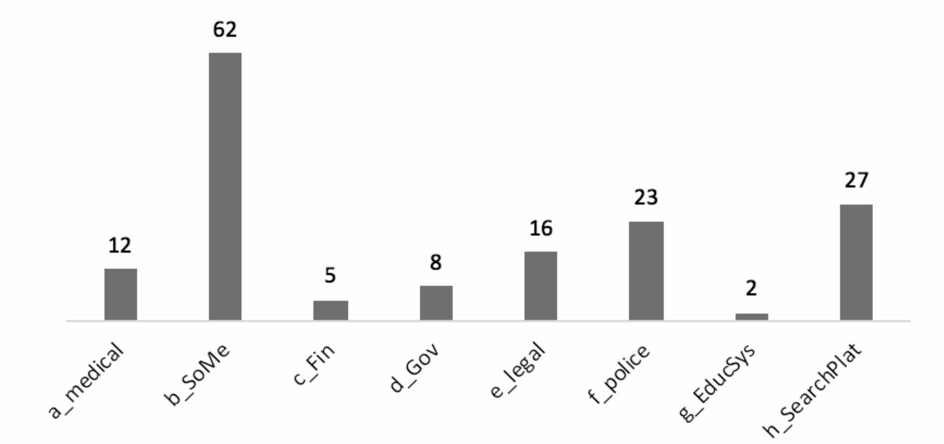


Figure 28 - Users - Overall count of mentions – Before and After

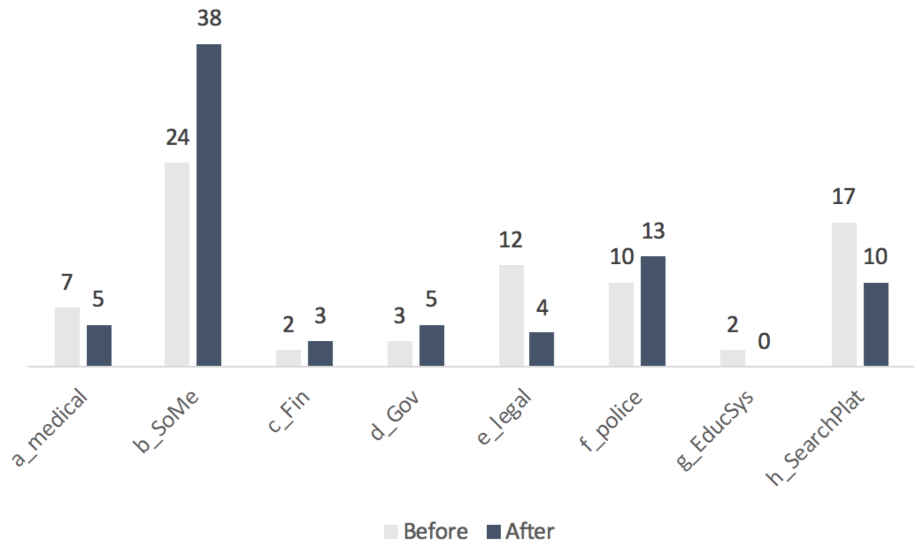
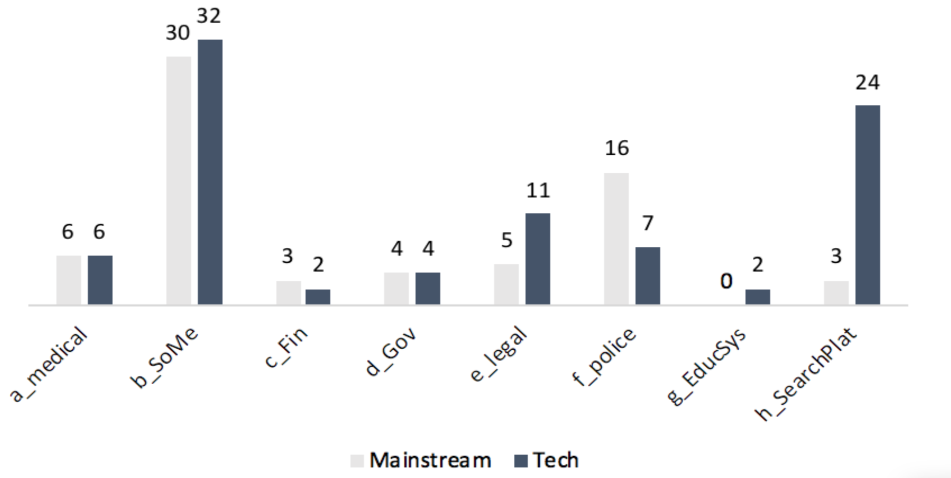


Figure 29 - Users - Overall count of mentions by type of media



31. Is there a direct reference in the article to algorithms benefiting revenue for corporations? (Profit) [No=0, Yes=1]

The code is in line with the critical literature that discusses how platforms manipulate algorithms, user preferences and the information they consume in order to benefit corporate revenue.

Table 31 - Profit - Overall frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	171	91.4	91.4	91.4
	1	16	8.6	8.6	100.0
	Total	187	100.0	100.0	

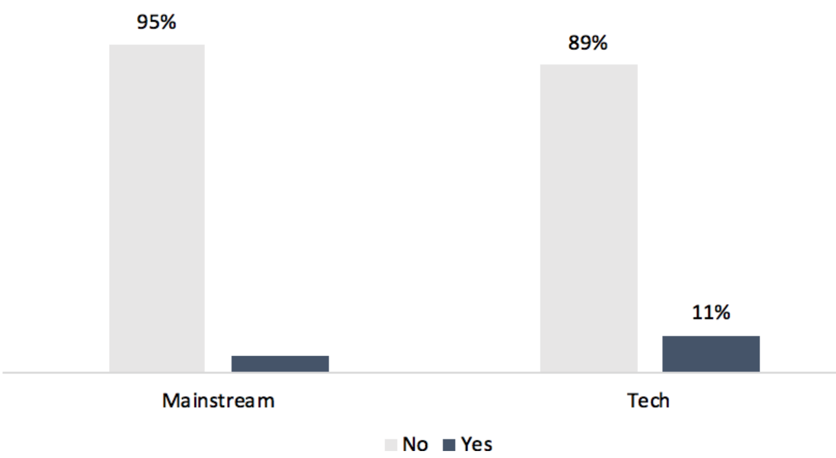
Table 32 - Profit - Mainstream - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	73	94.8	94.8	94.8
	1	4	5.2	5.2	100.0
	Total	77	100.0	100.0	

Table 33 - Profit - Technology - frequencies and percentages

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	98	89.1	89.1	89.1
	1	12	10.9	10.9	100.0
	Total	110	100.0	100.0	

Figure 30 - Profit - Comparative percentages by type of media



REGULATION

32. Does the article make any reference that could be related to any of these postures towards regulation? (search for keywords regulation, audit, control or monitoring in the text)

This code is related to the technology regulation literature. Since algorithms are claimed to have negative social consequences by social sciences critical approaches, an important section of the literature argues that in order to develop sustainable technological developments that are in line with the development and progress of society, regulation is needed. This will help answer hypothesis that link the critical literature with external regulation narratives. The categories are coded as dummies for easier manipulation of results.

1. Pro-Government regulation (a_ProReg) (external, government regulation) [No=0, Yes=1]
2. Anti-Government regulation (b_AntiReg) (no one should regulate the companies) [No=0, Yes=1]
3. Corporate self-regulation (c_SelfReg) (companies must regulate themselves) [No=0, Yes=1]
4. Multi-Stakeholder Regulation (d_MultiReg) (stakeholders participate in the dialogue together) [No=0, Yes=1]
5. Mixed (e_Mixed) (mentions regulation but does not specify the kind) [No=0, Yes=1]
6. Does not mention regulation (f_NoMention) [No=0, Yes=1]
7. N/A (as the article does not highlight an issue around the algorithmic bias, the topic for regulation does not apply) (g_NA_reg) [No=0, Yes=1]

Table 34 - Regulation - Overall count of mentions

a_ProReg	b_AntiReg	c_SelfReg	d_MultiReg	e_Mixed	f_NoMention	N/A_Reg
36	2	8	4	22	85	30

Figure 31 - Regulation - Overall percentages of mentions (excluding N/A, Very Positive Frames)

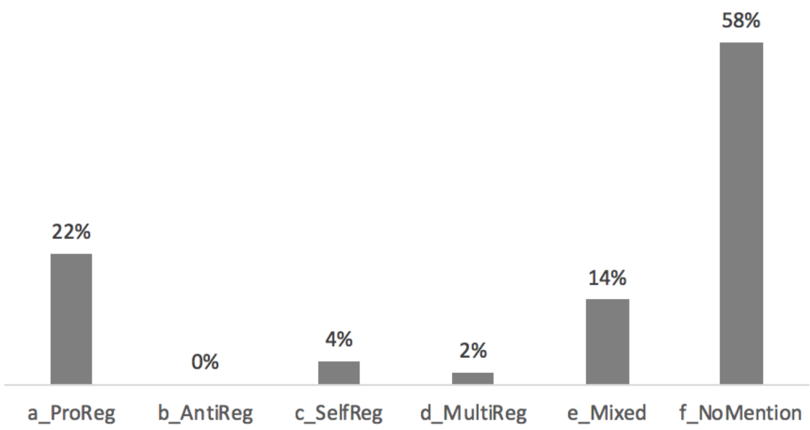


Table 35 - Regulation - Overall percentages of mentions – Before and After

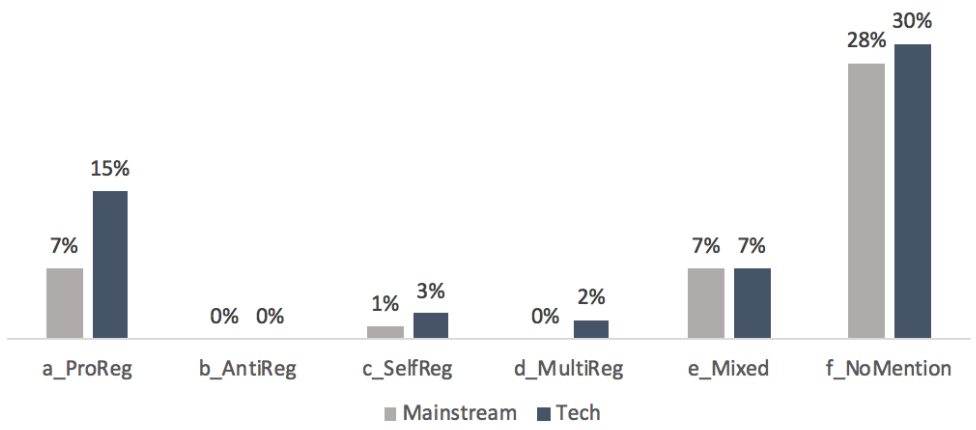
	Articles Published	Mentions of External Regulation	% of mentions in relation to articles published
Before	87	21	24%
After	100	41	41%

Table 36 - Regulation - Overall count of mentions by media type

	a_ProReg	b_AntiReg	c_SelfReg	d_MultiReg	e_Mixed	f_NoMention	N/A_Reg
Mainstream	11	0	3	1	11	47	4
Tech	25	2	5	3	11	38	26

Figure 32 - Regulation - Overall percentages of mentions by media type

(excluding N/A, Very Positives)



ADDITIONAL VARIABLES (created after results, conglomerating categories)

Dominant and Critical Frames (DominantFrames; CriticalFrames)

Table 37 – Mainstream - Postures- Overall count of postures

(excluding neutral, 19 counts)

	CriticalFrames	DominantFrames	Total
Before	15	9	24
After	25	9	34
Grand Total	40	18	58

Table 38 – Technology - Postures- Overall count of postures

(excluding neutral, 11 counts)

	CriticalFrames	DominantFrames	Total
Before	32	21	53
After	27	19	46
Grand Total	59	40	99

RegulationTalk

Table 39 – Mainstream - RegulationTalk- Overall count of regulation talk

Row Labels	RegulationTalk	Articles published
Before	5	26
After	18	51
Grand Total	23	77

Table 40 – Mainstream - RegulationTalk- Overall count of regulation talk

Row Labels	RegulationTalk	Articles published
Before	16	61
After	23	49
Grand Total	39	110

APPENDIX 4 – SPSS – HYPOTHESIS TESTING - LINEAR REGRESSIONS AND CORRELATION RESULTS

Regression analysis were used to test whether after the Cambridge Analytica scandal (independent variable) there were significant changes in the coverage (Agresti & Findlay: 2009: 265). Correlation tests were conducted to establish whether there might be possible relationships between variables (p.269). A 90% confidence interval was applied as the sample number is quite limited (n= 187, 0.10 threshold) (Gardner & Altman, 1986).

- Hypothesis A: (H1) The presence of critical frames increased after the occurrence of the scandal (H0: The presence of critical frames did not increase after the occurrence of the scandal)

REGRESSION 1: Mainstream - Variables regressed: Critical Frames - After CA

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	90.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	.577	.099		5.829	.000	.412	.742
	after_CA	-.077	.122	-.073	-.630	.530	-.280	.126

a. Dependent Variable: CriticalFrames

H0 is not rejected as there is insufficient evidence to support H1: the results are not significant. We did not find evidence to suggest that after CA we have more critical framings in Mainstream in the sample.

REGRESSION 2: Technology - Variables Regressed: Critical Frames - After CA

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	90.0% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	.525	.064		8.137	.000	.418	.632
	after_CA	.015	.096	.015	.160	.873	-.144	.175

a. Dependent Variable: CriticalFrames

H0 is not rejected as there is insufficient evidence to support H1: the results are not significant. We did not find evidence that suggest that after CA we have more critical framings in Tech in the sample.

- Hypothesis B: (H1) the presence of dominant frames decrease after the occurrence of the scandal (H0: the presence of dominant frames did not decrease after the occurrence of the scandal)

REGRESSION 3: Mainstream - Variables Regressed: DominantFrames - After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.346	.081		4.277	.000	.211	.481
	after_CA	-.186	.100	-.212	-1.865	.066	-.352	-.020

a. Dependent Variable: DominantFrames

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the results are significant ($p < 0.10$). We did find evidence to suggest that articles published after the scandal are associated with a reduction of 19 percentage points of dominant framings in Mainstream in the sample.

REGRESSION 4: Tech - Variables Regressed: DominantFrames - After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.344	.062		5.530	.000	.241	.448
	after_CA	.056	.093	.057	.601	.549	-.098	.210

a. Dependent Variable: DominantFrames

H0 is not rejected as there is insufficient evidence to support H1: the results are not significant ($p > 0.10$). We did not find evidence to suggest that after CA we have more dominant frames in Tech in the sample.

- Hypothesis D: (H1) the amount of articles mentioning external regulation increased after the scandal (H0: the amount of articles mentioning external regulation did not increase after the scandal)

REGRESSION 6: Mainstream - Variables regressed: RegulationTalk - After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.185	.088		2.105	.039	.039	.332
	after_CA	.182	.110	.190	1.663	.101	.000	.365

a. Dependent Variable: RegulationTalk

H0 is not rejected as there is insufficient evidence to support H1: the results are not significant in mainstream media ($p > 0.10$). We did not find evidence that suggest that after CA there are more articles that discuss external regulation in Mainstream media in the sample.

REGRESSION 7: Tech - Variables regresses: RegulationTalk - After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.267	.061		4.368	.000	.165	.368
	after_CA	.193	.091	.201	2.135	.035	.043	.344

a. Dependent Variable: RegulationTalk

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the regression showed significant results in tech media (p-value<0.10). There a positive association between articles published after the event and articles that discuss external regulation. This suggests that the percentage of articles that discuss regulation increased after the scandal in 19 percentage points.

- Hypothesis C: (H1) critical framings of the issue are associated with discussions that mention external regulation (H0: critical framings of the issue are not associated with discussions that mention external regulation)

CORRELATION 1: Mainstream - Variables correlated: RegulationTalk - CriticalFrames

Correlations

		RegulationTalk	CriticalFrames
RegulationTalk	Pearson Correlation	1	.270*
	Sig. (2-tailed)		.017
	N	77	77
CriticalFrames	Pearson Correlation	.270*	1
	Sig. (2-tailed)	.017	
	N	77	77

*. Correlation is significant at the 0.05 level (2-tailed).

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the correlation between variables for mainstream media is significant (p-value<0.1 threshold) and it is positive: for the higher the presence of regulation talk, the higher the presence of critical postures (or vice versa) in the sample. There is a correlation between critical framings and external regulation in Mainstream.

CORRELATION 2: Technology- Variables correlated: RegulationTalk - CriticalFrames

Correlations

		RegulationTalk	CriticalFrames
RegulationTalk	Pearson Correlation	1	.283**
	Sig. (2-tailed)		.003
	N	110	110
CriticalFrames	Pearson Correlation	.283**	1
	Sig. (2-tailed)	.003	
	N	110	110

**. Correlation is significant at the 0.01 level (2-tailed).

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the correlation between variables for mainstream media is significant (p-value<0.1 threshold) and it is positive: for the higher the presence of regulation talk, the higher the presence of critical postures (or vice versa) in the sample. There is a correlation between critical framings and external regulation in Tech.

- Hypothesis E: (H1) articles published after are associated with mentions of Cambridge Analytica (H0: articles published after are not associated with mentions of Cambridge Analytica)

CORRELATION 3: Mainstream - Variables correlated: after_CA and CA_mention

Correlations

		after_CA	CA_Mention
after_CA	Pearson Correlation	1	.351**
	Sig. (2-tailed)		.002
	N	77	77
CA_Mention	Pearson Correlation	.351**	1
	Sig. (2-tailed)	.002	
	N	77	77

**. Correlation is significant at the 0.01 level (2-tailed).

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the correlation between variables for mainstream media is significant (p-value<0.1 threshold) and it is positive. There is a strong correlation between articles published after the scandal and mentions of Cambridge Analytica.

CORRELATION 4: Tech - Variables correlated: after_CA and CA_mention

Correlations

		CA_Mention	after_CA
CA_Mention	Pearson Correlation	1	.348**
	Sig. (2-tailed)		.000
	N	111	111
after_CA	Pearson Correlation	.348**	1
	Sig. (2-tailed)	.000	
	N	111	111

**. Correlation is significant at the 0.01 level (2-tailed).

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the correlation between variables for tech media is significant (p-value<0.1 threshold) and it is positive. There is a strong correlation between articles published after the scandal and mentions of Cambridge Analytica.

- Hypothesis F: (H1) Negative frames are associated with mentions of Cambridge Analytica (H0 Negative frames are not associated with mentions of Cambridge Analytica)

CORRELATION 5: Mainstream - Variables correlated: CriticalFrames and CA_mention

Correlations

		CriticalFrames	CA_Mention
CriticalFrames	Pearson Correlation	1	-.059
	Sig. (2-tailed)		.611
	N	76	76
CA_Mention	Pearson Correlation	-.059	1
	Sig. (2-tailed)	.611	
	N	76	76

H0 is not rejected as there is insufficient evidence to support H1: the correlation between variables for mainstream media is non- significant (p-value<0.1 threshold) and it is positive. There is no evidence to sustain there is an association between critical frames and mentions of Cambridge Analytica.

CORRELATION 6: Tech - Variables correlated: CriticalFrames and CA_mention

Correlations

		CriticalFrames	CA_Mention
CriticalFrames	Pearson Correlation	1	.169
	Sig. (2-tailed)		.076
	N	111	111
CA_Mention	Pearson Correlation	.169	1
	Sig. (2-tailed)	.076	
	N	111	111

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the correlation between variables for tech media is significant (p-value<0.1 threshold) and it is positive. There is a correlation between critical frames and mentions of Cambridge Analytica for tech in the sample.

- In relation to Sub-question A) ‘how prominent is corporate discourse in the coverage on algorithmic bias is as compared to other discourses?

Hypothesis: (H1) government discourse increased after the Cambridge Analytica Scandal (H0: government discourse did not increase after the Cambridge Analytica Scandal)

REGRESSION 8: Mainstream - Variables regresses: GovPar_Quote- After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.077	.084		.914	.364	-.063	.217
	after_CA	.303	.104	.322	2.921	.005	.130	.476

a. Dependent Variable: GovPar_quote

H0 can be rejected and H1 accepted because of sufficient evidence in the sample in favour or H1: the regression showed significant results (p-value<0.10). There a positive association between mainstream articles published after the event and articles that quoting governments or parliaments. This suggests that the percentage of articles that quote government discourse increased after the scandal in 30 percentage points in the sample.

REGRESSION 9: Tech - Variables regresses: GovPar_Quote- After_CA

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	90.0% Confidence Interval for B Lower Bound	90.0% Confidence Interval for B Upper Bound
1	(Constant)	.164	.050		3.248	.002	.080	.248
	after_CA	.056	.075	.071	.745	.458	-.069	.181

a. Dependent Variable: GovPar_quote

H0 is not rejected as there is insufficient evidence to support H1: the results are not significant in tech media (p > 0.10). We did not find evidence that suggest that after CA there are more articles that quoting governments or parliaments in the sample.

APPENDIX 5 – METHODOLOGY

Examples of Content Analysis studies

Politics (Akhavan-Majid & Ramaprasad, 1998; de Vreese, et al., 2001) gender (Boni, 2002; Lind & Salo, 2002) international conflicts (Entman, 1991; Dimitrova, et al., 2005) and health (Muhamad & Yang, 2017; Mann, 2018) just to mention a few.

Description of media

BBC News: it is the British public service broadcaster, the most distinguished British media conglomerate worldwide. Independent third-party “Media Bias Fact Check”, states that BBC News has a slight left-centre bias, showing Very High levels of factual reportings (Media Bias Fact Check, BBC, 2018). With 47% of weekly use, it is the most consumed online brand in the UK (Reuters, 2017)

The Guardian: one of the most relevant daily British newspapers. Its story selection it has been characterised as liberal, leaning towards the left, but is generally highly factual (Media Bias Fact Check, The Guardian, 2018). With 14% of weekly use, it is amongst most consumed online news brand in the UK (Reuters, 2017)

Daily Mail: British daily newspaper. This media source has been characterised as s ‘moderately to strongly biased toward conservative causes’, presenting a mixed level of factual reporting

(Media Bias Fact Check, Daily Mail, 2017). With 14% of weekly use, it is amongst most consumed online news brand in the UK (Reuters, 2017).

Huffington Post: US originary, it's an online news and opinion site and blog that has local and international editions (Budak, et al., 2014). It presents a moderate to strong bias towards liberalism and a high factual reporting (Media Bias Fact Check, 2018). Even though it is not originally from the UK it represents 14% of weekly use in the country, therefore being amongst most consumed online news brand in the UK (Reuters, 2017).

Sky News: initially a 24/7 British television news channel, it now provides news online as well. It presents a slight to moderate liberal bias and high levels of factual reporting (Media Bias Fact Check, 2017). With 10% of weekly use, it occupies the third place amongst the most consumed online news brand in the UK (Reuters, 2017).

MSN News: characterised as a web portal that recollects news from an ample variety of mainstream media, mostly coming from left/center or low biased sources. High factual reporting (Media Bias Fact Check, 2017). With 7% of weekly use, it occupies the fourth place amongst the most consumed online news brand in the UK (Reuters, 2017).

Daily Telegraph: is a broadsheet newspaper published in London distributed in the United Kingdom and also internationally. Presents a slight to moderate conservative in bias with high factual reporting (Media Bias Fact Check, 2016). With 6% of weekly use, it occupies the fifth place amongst the most consumed online news brand in the UK (Reuters, 2017).

MIT Technology Review: it is published by the Massachusetts Institute of Technology. It tends to highly scientific and evidence, based in the use of credible scientific sources. Very high factual reporting (Media Bias Fact Check, 2016).

Wired: the magazine covers technology industry topics like the internet and digital culture, science and security. It is claimed to rely on credible sources coming from newswires but also technological and scientific sources. It is argued to have a slight to moderate liberal bias and high factual reporting (Media Bias Fact Check, 2018)

Fast Company: the focus of this magazine is on technology, business, and design. It is argued to be bias slightly towards liberalism and the center-left, it presents high levels of factual reporting (Media Bias Fact Check, 2017).

Gizmodo: it is characterised as a design, technology and science fiction digital magazine, sometimes reporting on politics concerns technology. It presents a left bias and high factual reporting (Media Bias Fact Check, 2016).

TechCrunch: tech industry news, covering business, breaking news, product opinions and emerging trends in tech, it does not tend to cover a great deal of politics, but it has presented eft-center biases when doing so. Factual reporting characterised as high (Media Bias Fact Check, 2017).

Recode: tech news site focusing on business in Silicon Valley. It does not tend to cover politics topics, presenting very little bias. This source has been claimed to be one of the least rate this least biased and more factual (Media Bias Fact Check, 2017).

It might be worth clarifying three things about the selection criteria for mainstream media from the Reuters 2017 report on Digital News: first, that the category ‘website of local paper’ was discarded as it is not very specific. Second, that the Huffington Post, while not being originary from the UK, was considered as it has a significantly high digital consumption in the country and, third, that Buzzfeed News was discarded from the selection as for the researcher did not meet the ‘mainstream media’ criteria, it is rather understood as a niche publication.

Keywords and Article Selection

A ‘census’ sort of sampling was exercised due to the small amount of data found: all units in the sampling frame that met the criterion were taken (Macnamara, 2005: 13). Originally, there was a limit of sampling of 30 articles per outlet. Nevertheless, given the fact that the sampling universe was quite small, there was no need to recur to this limit. Instead, all the articles containing the relevant terms were selected for the sample. The researcher searched the terms ‘bias+algorithm’ first and select all the available results for the sample. In case of not obtaining sufficient information, it was required to search the combination of ‘bias+machine learning’ and then ‘bias+artificial’ and ‘bias+AI’ (as all these technologies are driven by algorithms and often used as equivalent). All articles selected must contain the word ‘bias’ and any of the variations above. The need to perform different searches in order to find relevant articles is related to the overall quite limited presence of information about the topic in the media.

Furthermore, there was a deliberate decision not to look for the term “algorithmic bias” as it might bias the sample towards more critical approaches. By searching by bias + algorithms (and the variations) the search allowed for a more varied results and the presence of corporate discourse, which can all be examined under the light of the ‘algorithmic bias’ literature.

As for the search procedure, the results were sorted by relevance on Google Search (acknowledging that this might be biasing the sample itself). The keywords (‘bias+algorithm’; ‘bias+algorithms’; ‘bias+machine learning’; ‘bias+artificial’ and ‘bias+AI’) were searched in the whole text of the page. These keywords were selected as they were considered more neutral than searching for terms like ‘algorithmic bias’ -which is claimed to have a more critical connotation towards the issue-.

<u>Search Procedure</u>

Route to search the articles: Google Search → Tools: date 1/1/18-1/6/18 → Settings → Advanced Search → Find pages with.. → “all these words” → ‘algorithm’ + ‘bias’ → “site or domain” → site URL.

All articles must contain the word bias. There were cases in which results would come up from the search but were not relevant for the study. As a first step, all articles were manually checked for the word ‘bias’ and approved in case they were proven to be worthy of consideration. Given the case an article presented the word bias, but not in significant relation to either of the other keywords, they will were discarded. Example of discarded articles:

https://goo.gl/nqfNt8; https://goo.gl/d3b6fs; https://goo.gl/3n1y5B. Repeated articles were also excluded from the search results. Articles that contained all the keywords but are not relevant to the search will be excluded as well. Examples: https://goo.gl/MFTVmm; https://goo.gl/DroCBF. If the articles contained al keywords but referred to a body of content that is not available in text (Ex.: videos) they were be discarded too. Example: https://goo.gl/4RUFZJ.

Selection of co-coder

One of the challenges the researcher faced during the development of this project was how to assess its reliability and ensure rigurosity, beyond the selection of the specific ICR test -per cent agreement, basic assessment; Scott’s pi; Cohen’s kappa; Spearman’s rho; Pearson’s correlation coefficient (r); Krippendorf’s apha; and Lin’s concordance correlation coefficient, etc (Lombard et al., 2004)-. What was most challenging was the actual finding and selection of a suitable co-coder. For the present study, being a solo researcher at a master’s level implied a significant challenge at the time to find someone willing to dedicate time and attention to the project. The researcher was able to find two willing and suitable subjects eager to dedicate their time to this project, however, at this point, the researcher faced an additional decision: choosing between a ‘neutral’ native English-speaking co-coder, not involved in the academia, which would ensure the coding is replicable at the ‘simplest’ of levels and a doctoral level co-coder, non-English native, which might risk lowering the replicability given her familiarity with the lexicon and methodology. Quantitative content analysis is claimed to be quite a time-consuming methodology (Newbold et al., 2002: 249; Macnamara, 2005: 5) furthermore, it can be quite an arid task for someone not involved in the actual research. Due to time constrains of the project, it was decided that there would be only one co-coder. To ensure maximum focus, dedication and grasping of fundamental concepts, the researcher opted to select the doctoral-level co-coder to carry on the with task. Nevertheless, the ‘neutral’ potential co-coder proved to be instrumental as well for the proof-reading of the codebook, making useful suggestions on the rephrasing of code questions for easy understanding.

Pilots and Intercoder Reliability test

Two reading pilots were conducted on the codebook before testing it for reliability. After the pilots, alterations were made in order to make the codebook stronger (Deacon, 1999) and clear for the understanding of others.

The first reading pilot involved the reading of the original codebook by an assistant professor, who suggested more granularity on the codes to allow more flexibility at the time of the analysis of results (find first codebook version attached to Appendix section). Some of the codes, like the main code 'Posture' were added more categories in order to make them more comprehensive and inclusive (Krippendorff, 2004: 132) and additional codes were included in the code book. At the end, the codebook ended up containing 32 codes (in contrast with the original 16) (find second version of the codebook attached to Appendix section).

Strategies to maximize agreement were performed (Macnamara, 2005: 12). A second reading pilot of the expanded version was done by two assistant researchers (one neutral and one academic). At this instance, suggestions were made on the rephrasing of the questions (simplification of language) and about the clusters in which the codes were organised, with the purpose of facilitating the coding. Furthermore, suggestions were made about the nature of some of the codes in order to facilitate the analysis on SPSS: for instance, code 'posture' was changed from dummy to ordinal in order to perform the regressions needed for the testing of the hypotheses. All these changes resulted in the creation of the final codebook (found in Appendix section) which was used for the coding pilots.

A first coding pilot was conducted on two articles (one considered to be critical and one dominant) to test the understanding of the codes. While proving that it was overall functional for coding, the test evidenced that some codes were not being understood equally by the coder and co-coder, like ‘Posture’, ‘Algo_ReduceReinforce’, ‘a_Ethics’, ‘b_PolInter’, ‘c_Personnel’, ‘k_Privacy’, ‘c_MoreTrans’, ‘b_AntiReg’, ‘d_MultiReg’, ‘e_Mixed’, ‘a_Empower’, ‘b_CostEffective’, ‘CorpoEffort’ and ‘Polemical’. These results were discussed with the co-coder in a ‘re-briefing” (p.12) to clarify descriptions and instructions in order to proceed with the full test.

For the ICR test, 10% of the sample was tested (19 articles) which were randomly selected (selecting one every 10 articles). Percent agreement was taken as indicator for reliability. Despite overall percentage agreements proved to be acceptable, Scott's Pi, Cohen's Kappa and Krippendorff's Alpha show low levels of reliability (Feinstein & Cicchetti, 1990): when one code is predominantly assigned by both coders (variability lack) these coefficients can show low results that can be misleading, that is why this research project used percent agreement as it proves to be the most straightforward indicator.

After conducting the full test on the 19 articles, there was an overall disagreement on two codes (‘posture’ presenting a 74% and ‘polemical’, 68%)’ and 2 categories inside one code (‘d_MultiReg’ and ‘e_Mixed’, both categories coded as dummies, corresponding to code 32 ‘regulation’, presenting both 74% as well). After discussion, code ‘polemical’ was discarded from the codebook as it was argued some other variables could respond to the same objective of answering hypothesis regarding ‘level’ of coverage. As for codes ‘posture’ and ‘regulation’, the disagreements found were considered not to be of crucial relevance, as it was detected that had to do with blurriness in similar categories like ‘positive’ and ‘very positive’ and ‘mixed’ and ‘multi stakeholder’ regulation. In that sense, these classifications ‘fell’ into the same broader categories that were afterwards grouped into variables ‘CriticalFrames’/’DominantFrames’ and ‘RegulationTalk’. The ICR test was conducted for each of the variables and for each of the categories that were coded as dummies, presenting an overall percent agreement of 94%, while showing wide range of percent agreements ranging from 68% to 100%. Coding details for both coders are available in the Appendix section, together with the ReCal2 results for each code and category.

APPENDIX 6 - CODING MAINSTREAM ARTICLES

_ID	Media	Media Name	Seach	URL	Headline	Date	Author	after_CA	kind	Posture	CriticalFrames	DominantFrames	Main_topic
1	1	BBC News	1	o.uk/news/b	I didn't ever	06/02/2018	Sarah Finley	1	0	3	0	0	0
2	1	BBC News	1	uk/news/te	YouTube to	17/01/2018	Leo Kelion	0	0	3	0	0	0
3	1	BBC News	1	uk/news/te	Are you sca	01/06/2018	ane Wakefie	1	0	1	1	0	1
4	1	BBC News	1	uk/news/te	Facebook t	19/01/2018	Dave Lee	0	0	3	0	0	0
5	1	BBC News	1	0180316-w	Four things	19/03/2018	Zaria Gorve	1	0	4	0	1	0
6	1	BBC News	2	ticle/8bbe07	Are sex rob	06/04/2018	Ciaran Varle	1	0	3	0	0	0
7	1	BBC News	2	ews/live/wo	As it happe	10/04/2018	and Taylor	1	0	3	0	0	0
8	1	BBC News	1	uk/news/te	Facebook's	11/04/2018	N/A	1	0	3	0	0	0
9	1	BBC News	1	uk/news/te	Mark Zuck	12/04/2018	Dave Lee	1	0	2	1	0	0
10	1	BBC News	1	o.uk/news/b	Is this the y	01/05/2018	Matthew Wa	1	0	1	1	0	0
11	2	The Guardia	2	018/jan/21/t	As technolo	21/01/2018	aul Chadwic	0	0	2	1	0	0
12	2	The Guardia	3	02/youtube-a	How an ex-	02/02/2018	s & Erin Mc	0	0	1	1	0	1
13	2	The Guardia	2	8/feb/17/lan	Enemies, tr	17/02/2018	ımes Graha	0	0	2	1	0	0
14	2	The Guardia	2	ing-impene	Dehumanis	04/03/2018	ephen Burar	0	0	1	1	0	1
15	2	The Guardia	2	mar/04/robo	How to per	04/03/2018	ephen Burar	0	0	1	1	0	1
16	2	The Guardia	2	/mar/13/wo	Women mu	12/03/2018	ana Bartole	0	0	1	1	0	1
17	2	The Guardia	2	18/apr/04/a	Algorithms	04/04/2018	rietje Schaa	1	0	1	1	0	1
18	2	The Guardia	1	-live-update	Mark Zuck	11/04/2018	a Carrie W	1	0	3	0	0	0
19	2	The Guardia	2	-zuckerberg	The Facebo	12/04/2018	sta and Jona	1	0	2	1	0	0
20	2	The Guardia	2	16/the-peop	The People	16/04/2018	Emily Bell	1	0	2	1	0	0
21	2	The Guardia	2	te-guardian-	The Guardi	16/04/2018	Editorial	1	0	1	1	0	1
22	2	The Guardia	2	8/apr/19/art	Artificial int	19/04/2018	, Matt Mey	1	0	5	0	1	0
23	2	The Guardia	1	ay/09/uk-ac	UK accuse	09/05/2018	Vikram Dod	1	0	1	1	0	1
24	2	The Guardia	2	ay/13/we-c	We created	13/05/2018	ginia Eubar	1	0	1	1	0	1
25	2	The Guardia	1	14/is-your-h	Employers	14/05/2018	Emine Sane	1	0	3	0	0	0
26	2	The Guardia	2	tificial-intell	Artificial int	16/05/2018	ane Howare	1	0	1	1	0	0
27	2	The Guardia	2	oscientist-h	Neuroscien	18/05/2018	Emine Sane	1	0	3	0	0	0
28	2	The Guardia	1	gs-we-learn	Five things	22/05/2018	im Waterso	1	0	2	1	0	0
29	2	The Guardia	2	-appears-b	Mark Zuck	22/05/2018	Alex Hern	1	0	2	1	0	0
30	2	The Guardia	1	/27/jaron-la	Six reasons	27/05/2018	Jaron Lanie	1	0	1	1	0	0
31	2	The Guardia	1	police-trial-	Police trial	27/05/2018	cott and Ha	1	0	1	1	0	1
32	2	The Guardia	1	new-publis	New publis	01/06/2018	Alison Flood	1	0	5	0	1	0
33	2	The Guardia	1	8/jun/01/face	Facebook s	01/06/2018	ss Associat	1	0	3	0	0	1
34	3	Daily Mail	1	3847/The-r	The REAL	04/01/2018	a Konstanti	0	0	5	0	1	0
35	3	Daily Mail	1	3337/Comp	Minority Re	18/01/2018	oe Pinkston	0	0	2	1	0	1
36	3	Daily Mail	1	310031/Evi	Are compu	25/01/2018	John Naish	0	0	1	1	0	1
37	3	Daily Mail	1	5333757/AI	Concerns i	31/01/2018	sociated Pre	0	0	1	1	0	1
38	3	Daily Mail	2	3/Critics-Fo	Will Ford's	06/02/2018	and Cather	0	0	5	0	1	0
39	3	Daily Mail	1	82979/Stud	Is facial rec	12/02/2018	nnie Palme	0	0	2	1	0	1
40	3	Daily Mail	1	3/By-2050-h	Humanity w	16/02/2018	Zoe Nauma	0	0	5	0	1	0
41	3	Daily Mail	1	37/Secretive	Secretive b	28/02/2018	nnie Palme	0	0	1	1	0	1
42	3	Daily Mail	2	469/Google	Google is le	06/03/2018	nnie Palme	0	0	4	0	1	0
43	3	Daily Mail	1	e-5499339/	Are Alexa a	14/03/2018	oe Pinkston	0	0	2	1	0	1
44	3	Daily Mail	1	ew-AI-softw	Companies	29/03/2018	nnie Palme	1	0	3	0	0	0
45	3	Daily Mail	1	gle-employ	Google sho	05/04/2018	nnie Palme	1	0	3	0	0	0
46	3	Daily Mail	1	367/AI-stud	The real MI	11/04/2018	hoebe West	1	0	3	0	0	0
47	3	Daily Mail	1	ho-Diamon	Who are Di	11/04/2018	Caitlyn Hitt	1	0	3	0	0	1
48	3	Daily Mail	2	-Elon-Musk	Elon Musk	11/04/2018	nnie Palme	1	0	1	1	0	0
49	3	Daily Mail	2	99/The-Late	The Latest:	11/04/2018	sociated Pre	1	0	3	0	0	0
50	3	Daily Mail	1	urveillance-	Secret surv	17/04/2018	Tim Collins	1	0	3	0	0	0
51	3	Daily Mail	1	ows-algorit	Are social r	23/04/2018	nnie Palme	1	0	1	1	0	1
52	3	Daily Mail	1	cognition-AI	Facial reco	27/04/2018	oe Pinkston	1	0	2	1	0	1
53	3	Daily Mail	1	oversial-Ga	Controvers	09/05/2018	ss Associat	1	0	1	1	0	1
54	3	Daily Mail	1	zons-facial-	Is Amazon'	24/05/2018	Tim Collins	1	0	2	1	0	1
55	3	Daily Mail	1	5796239/Fa	Facebook s	01/06/2018	ss Associat	1	0	3	0	0	1
56	3	Daily Mail	1	3/Facebook	Facebook k	01/06/2018	sociated Pre	1	0	3	0	0	1
57	4	uffington Po	1	or-2018-fro	My Five Pr	02/01/2018	David Sable	0	0	1	1	0	0
58	4	uffington Po	2	-tech-trends	Top 5 Tech	04/01/2018	Alice Bonasi	0	0	5	0	1	0
59	4	uffington Po	1	lways-untru	Fake news	09/01/2018	ns Erik Gou	0	0	1	1	0	0
60	4	uffington Po	1	intelligence	Artificial Int	04/05/2018	James Ball	1	0	4	0	1	0
61	5	Sky News	1	erberg-face	Zuckerberg	11/04/2018	N/A	1	0	3	0	0	0
62	5	Sky News	1	at-centre-o	Ethics mus	16/04/2018	ick Styliano	1	0	2	1	0	1
63	6	MSN News	2	powered-bo	Can an AI-p	04/03/2018	nny Anders	0	0	4	0	1	0
64	6	MSN News	1	es-travel-fa	False news	08/03/2018	rian Resnic	0	0	2	1	0	1
65	6	MSN News	1	80%99s-cri	Facebook's	31/03/2018	chael J. Co	1	0	4	0	1	0
66	6	MSN News	1	ssons-from	9 essential	14/04/2018	rian Resnic	1	0	1	1	0	0
67	6	MSN News	1	bnb%E2%8	How AI is p	17/04/2018	ela Heathm	1	0	5	0	1	0
68	6	MSN News	1	l-of-flouting	UK accuse	08/05/2018	Vikram Dod	1	0	1	1	0	1
69	6	MSN News	1	n-rights-and	Police gang	08/05/2018	Nicola Bartle	1	0	1	1	0	1
70	6	MSN News	1	hatrix-unde	Scotland Ya	08/05/2018	ean Morriso	1	0	1	1	0	1
71	6	MSN News	1	down-troubl	Facebook s	01/06/2018	Martin Coulte	1	0	4	0	1	1
72	7	The Telegrap	2	25/big-bank	Big banks s	25/02/2018	Lucy Burton	0	0	5	0	1	0
73	7	The Telegrap	1	17/firms-cal	The robot w	17/03/2018	aroline Bullo	0	0	4	0	1	0
74	7	The Telegrap	2	s/essential-	Want better	23/03/2018	mma Kendre	1	0	2	1	0	0
76	7	The Telegrap	1	-interviewi	Robots inte	21/04/2018	amilla Turn	1	0	5	0	1	0
77	7	The Telegrap	2	obots-will-n	Robots will	27/05/2018	mes Titcom	0	0	4	0	1	0

APPENDIX 7 - CODING TECHNOLOGY ARTICLES

APPENDIX 8 - TECHNOLOGY ARTICLES AND IDs

_ID	Media	Media Name	Headline	Date	URL
78	8	MIT Tech Review	The US military wants AI to dream up weird new helicopters	19/03/2018	https://www.technologyreview.com/s/610454/the-us-military-wants-ai-to-dream-up-weird-new-helicopters/
79	8	MIT Tech Review	Google Photos Still Has a Problem with Gorillas	11/01/2018	https://www.technologyreview.com/the-download/609959/google-photos-still-has-a-problem-with-gorillas/
80	8	MIT Tech Review	Here are the secrets to useful AI, according to designers	26/01/2018	https://www.technologyreview.com/the-download/610077/here-are-the-secrets-to-useful-ai-according-to-designers/
81	8	MIT Tech Review	MIT Technology Review Announces 2018 EmTech Digital Conference March 26-27	31/01/2018	https://www.technologyreview.com/press-room/press-release/20180131-mit-technology-review-announces-2018-emtech-digital-conference-march/
82	8	MIT Tech Review	“We’re in a diversity crisis”: cofounder of Black in AI on what’s poisoning algorithms in our lives	14/02/2018	https://www.technologyreview.com/s/610192/were-in-a-diversity-crisis-black-in-ais-founder-on-whats-poisoning-the-algorithms-in-our/
83	8	MIT Tech Review	Terrible people have learned to exploit the internet. Yasmin Green is fighting back.	19/02/2018	https://www.technologyreview.com/s/610176/terrible-people-have-learned-to-exploit-the-internet-yasmin-green-is-fighting-back/
84	8	MIT Tech Review	Bias already exists in search engine results, and it’s only going to get worse	26/02/2018	https://www.technologyreview.com/s/610275/meet-the-woman-who-searches-out-search-engines-bias-against-women-and-minorities/
85	8	MIT Tech Review	The high-tech medicine of the future may be biased in favor of well-off white men	26/02/2018	https://www.technologyreview.com/s/610367/the-high-tech-medicine-of-the-future-may-be-biased-in-favor-of-well-off-white-men/
86	8	MIT Tech Review	Tech talent actually shows promise for a more female future	08/03/2018	https://www.technologyreview.com/the-download/610450/tech-talent-actually-shows-promise-for-a-more-female-future/
87	8	MIT Tech Review	China wants to shape the global future of artificial intelligence	16/03/2018	https://www.technologyreview.com/s/610546/china-wants-to-shape-the-global-future-of-artificial-intelligence/
88	8	MIT Tech Review	AI savants, recognizing bias, and building machines that think like people	26/03/2018	https://www.technologyreview.com/s/610621/emtech-digital-oren-etzioni-brenden-lake-timnit-gebru/
89	8	MIT Tech Review	The startup diversifying the AI workforce beyond just “techies”	27/03/2018	https://www.technologyreview.com/s/610633/the-startup-diversifying-the-ai-workforce-beyond-just-techies/
90	8	MIT Tech Review	For better AI, diversify the people building it	28/03/2018	https://www.technologyreview.com/s/610637/for-better-ai-diversify-the-people-building-it/
91	8	MIT Tech Review	Here’s how the US needs to prepare for the age of artificial intelligence	06/04/2018	https://www.technologyreview.com/s/610379/heres-how-the-us-needs-to-prepare-for-the-age-of-artificial-intelligence/
92	8	MIT Tech Review	What Mark Zuckerberg’s testimony told us about the past, present, and future of Facebook and its data	11/04/2018	https://www.technologyreview.com/the-download/610854/what-mark-zuckerbergs-testimony-told-us-about-the-past-present-and-future-of/
93	8	MIT Tech Review	UK lawmakers want to bring good old British decorum to the AI industry	16/04/2018	https://www.technologyreview.com/the-download/610892/uk-lawmakers-want-to-bring-good-old-british-decorum-to-the-ai-industry/
94	8	MIT Tech Review	The Growing Impact of AI on Business	30/04/2018	https://www.technologyreview.com/s/611013/the-growing-impact-of-ai-on-business/
95	8	MIT Tech Review	This company audits algorithms to see how biased they are	09/05/2018	https://www.technologyreview.com/the-download/611113/a-new-company-audits-algorithms-to-see-how-biased-they-are/
96	8	MIT Tech Review	Microsoft is creating an oracle for catching biased AI algorithms	25/05/2018	https://www.technologyreview.com/s/611138/microsoft-is-creating-an-oracle-for-catching-biased-ai-algorithms/
97	9	Wired	A CHILD ABUSE PREDICTION MODEL FAILS POOR FAMILIES	15/01/2018	https://www.wired.com/story/excerpt-from-automating-inequality/
98	9	Wired	TWITTER TRIED TO CURB ABUSE. NOW IT HAS TO HANDLE THE BACKLASH	16/01/2018	https://www.wired.com/story/twitter-project-veritas-videos-backlash/
99	9	Wired	CRIME-PREDICTING ALGORITHMS MAY NOT FARE MUCH BETTER THAN UNTRAINED HUMANS	17/01/2018	https://www.wired.com/story/crime-predicting-algorithms-may-not-outperform-untrained-humans/
100	9	Wired	SHOULD DATA SCIENTISTS ADHERE TO A HIPPOCRATIC OATH?	02/02/2018	https://www.wired.com/story/should-data-scientists-adhere-to-a-hippocratic-oath/
101	9	Wired	THE TECHIES RUNNING FOR CONGRESS WALK A FINE LINE	02/02/2018	https://www.wired.com/story/techies-running-for-congress-walk-fine-line/
102	9	Wired	PHOTO ALGORITHMS ID WHITE MEN FINE—BLACK WOMEN, NOT SO MUCH	06/02/2018	https://www.wired.com/story/photo-algorithms-id-white-men-fineblack-women-not-so-much/
103	9	Wired	INSIDE THE TWO YEARS THAT SHOOK FACEBOOK—AND THE WORLD	18/02/2018	https://www.wired.com/story/inside-facebook-mark-zuckerberg-2-years-of-hell/
104	9	Wired	BAD ACTORS ARE USING SOCIAL MEDIA EXACTLY AS DESIGNED	11/03/2018	https://www.wired.com/story/bad-actors-are-using-social-media-exactly-as-designed/
105	9	Wired	SUSAN WOJCICKI ON YOUTUBE'S FIGHT AGAINST MISINFORMATION	15/03/2018	https://www.wired.com/story/susan-wojcicki-on-youtubes-fight-against-misinformation/
106	9	Wired	AI and the Future of Work	15/03/2018	https://www.wired.com/wiredinsider/2018/04/ai-future-work/
107	9	Wired	DON'T ASK WIKIPEDIA TO CURE THE INTERNET	16/03/2018	https://www.wired.com/story/youtube-wikipedia-content-moderation-internet/
108	9	Wired	HOW CODERS ARE FIGHTING BIAS IN FACIAL RECOGNITION SOFTWARE	29/03/2018	https://www.wired.com/story/how-coders-are-fighting-bias-in-facial-recognition-software/
109	9	Wired	EMMANUEL MACRON TALKS TO WIRED ABOUT FRANCE'S AI STRATEGY	31/03/2018	https://www.wired.com/story/emmanuel-macron-talks-to-wired-about-frances-ai-strategy/
110	9	Wired	MARK ZUCKERBERG AND THE TALE OF TWO HEARINGS	11/04/2018	https://www.wired.com/story/mark-zuckerberg-congress-day-two/
111	9	Wired	THE STARTUP THAT WILL VET YOU FOR YOUR NEXT JOB	27/04/2018	https://www.wired.com/story/the-startup-that-will-vet-you-for-your-next-job/
112	9	Wired	WANT TO PROVE YOUR BUSINESS IS FAIR? AUDIT YOUR ALGORITHM	09/05/2018	https://www.wired.com/story/want-to-prove-your-business-is-fair-audit-your-algorithm/
113	9	Wired	AI ISN’T A CRYSTAL BALL, BUT IT MIGHT BE A MIRROR	09/05/2018	https://www.wired.com/story/ideas-ai-as-mirror-not-crystal-ball/
114	9	Wired	FACIAL RECOGNITION TECH IS CREEPY WHEN IT WORKS—AND CREEPIER WHEN IT DOESN’T	09/05/2018	https://www.wired.com/story/facial-recognition-tech-creepy-works-or-not/
115	9	Wired	GOOGLE AND THE RISE OF 'DIGITAL WELL-BEING'	09/05/2018	https://www.wired.com/story/google-and-the-rise-of-digital-wellbeing/
116	9	Wired	SAM HARRIS AND THE MYTH OF PERFECTLY RATIONAL THOUGHT	17/05/2018	https://www.wired.com/story/sam-harris-and-the-myth-of-perfectly-rational-thought/
117	9	Wired	FRANCE, CHINA, AND THE EU ALL HAVE AN AI STRATEGY. SHOULDN’T THE US?	20/05/2018	https://www.wired.com/story/the-us-needs-an-ai-strategy/
118	9	Wired	THE REAL REASON GOOGLE SEARCH LABELED THE CALIFORNIA GOP AS NAZIS	31/05/2018	https://www.wired.com/story/google-search-california-gop-nazism/
119	9	Wired	FACEBOOK IS KILLING TRENDING TOPICS	01/06/2018	https://www.wired.com/story/facebook-killed-trending-topics/
120	10	Fast Company	This year your first job interview might be with a robot	10/01/2018	https://www.fastcompany.com/40515583/this-year-your-first-interview-might-be-with-a-robot
121	10	Fast Company	Warner’s warning: YouTube’s recommendation algorithm is susceptible to “bad actors”	05/02/2018	https://www.fastcompany.com/40526693/warners-warning-youtubes-recommendation-algorithm-is-susceptible-to-bad-actors
122	10	Fast Company	8 Ways To Make Real Progress On Tech’s Diversity Problem	12/02/2018	https://www.fastcompany.com/40528096/8-ways-to-make-real-progress-on-techs-diversity-problem
123	10	Fast Company	This Investment Platform Funds More Diverse Companies By Focusing On Data, Not Founders	14/02/2018	https://www.fastcompany.com/40527386/this-investment-platform-funds-more-diverse-companies-by-focusing-on-data-not-founders
124	10	Fast Company	This AI Job Site Promises to Find Diverse, Qualified Candidates	15/02/2018	https://www.fastcompany.com/40531805/this-ai-job-site-promises-to-find-diverse-qualified-candidates-for-5
125	10	Fast Company	This Chrome extension matches LinkedIn profiles to open tech jobs	21/02/2018	https://www.fastcompany.com/40533516/this-chrome-extension-matches-linkedin-profiles-to-open-tech-jobs
126	10	Fast Company	Now Is The Time To Act To End Bias In AI	28/02/2018	https://www.fastcompany.com/40536485/now-is-the-time-to-act-to-stop-bias-in-ai
127	10	Fast Company	A New Border Security App Uses AI To Flag Suspicious People In Seconds	06/03/2018	https://www.fastcompany.com/40539812/new-unisys-tool-brings-machine-learning-to-border-security
128	10	Fast Company	Algorithms Can’t Tell When They’re Broken–And Neither Can We	26/03/2018	https://www.fastcompany.com/40549744/algorithms-cant-tell-when-theyre-broken-and-neither-can-we
129	10	Fast Company	Bank of America’s bot is “Erica” because apparently all digital assistants are women	28/03/2018	https://www.fastcompany.com/40551011/bank-of-americas-bot-is-erica-because-apparently-all-digital-assistants-are-women
130	10	Fast Company	“Did We Create This Monster?” How Twitter Turned Toxic	04/04/2018	https://www.fastcompany.com/40547818/did-we-create-this-monster-how-twitter-turned-toxic
131	10	Fast Company	It’s Time To Regulate Big Tech, Says VC Who Helped Create It	07/04/2018	https://www.fastcompany.com/40554409/its-time-to-regulate-big-tech-says-vc-who-helped-create-it
132	10	Fast Company	What HR Is Doing To Make Sure There Aren’t More #MeToo Moments	18/04/2018	https://www.fastcompany.com/40558858/what-hr-is-doing-to-make-sure-there-arent-more-metoo-moments
133	10	Fast Company	Are Your Slack Chats Accidentally Destroying Your Work Culture?	25/04/2018	https://www.fastcompany.com/40560336/are-your-slack-chats-accidentally-destroying-your-work-culture
134	10	Fast Company	When It Comes To AI, Facebook Says It Wants Systems That Don’t Reflect Our Biases	03/05/2018	https://www.fastcompany.com/40567330/when-it-comes-to-ai-facebook-says-all-that-power-comes-with-great-responsibility
135	10	Fast Company	5 ways to ensure AI unlocks its full potential to serve humanity	10/05/2018	https://www.fastcompany.com/40570350/5-ways-to-ensure-ai-unlocks-its-full-potential-to-serve-humanity
136	10	Fast Company	Why the gender wage gap is persistent for freelancers, too	24/05/2018	https://www.fastcompany.com/40573469/why-the-gender-wage-gap-is-persistent-for-freelancers-too
137	11	Gizmodo	Canada Is Using AI to Study ‘Suicide-Related Behavior’ on Social Media	02/01/2018	https://gizmodo.com/canada-is-using-ai-to-study-suicide-related-behavior-1821704710
138	11	Gizmodo	Google Censors Gorillas Rather Than Risk Them Being Mislabeled As Black People—But Who Does That Help?	11/01/2018	https://gizmodo.com/google-censors-gorillas-rather-then-risk-them-being-mis-1821994460
139	11	Gizmodo	Why Google's Half-Assed Fact-Checking Widget Was Destined to Piss Off Conservatives	16/01/2018	https://gizmodo.com/why-googles-half-assed-fact-checking-widget-was-destine-1822005133
140	11	Gizmodo	Study Finds Crime-Predicting Algorithm Is No Smarter Than Online Poll Takers	18/01/2018	https://gizmodo.com/study-finds-crime-predicting-algorithm-is-no-smarter-th-1822173965
141	11	Gizmodo	New AI System Predicts How Long Patients Will Live With Startling Accuracy	18/01/2018	https://gizmodo.com/new-ai-system-predicts-how-long-patients-will-live-with-1822157278
142	11	Gizmodo	How Algorithmic Experiments Harm People Living in Poverty	23/01/2018	https://gizmodo.com/how-algorithmic-experiments-harm-people-living-in-pover-1822311248
143	11	Gizmodo	Ford Files Patent for Autonomous Robocop Car That Learns How to Hide From Drivers	25/01/2018	https://gizmodo.com/ford-files-patent-for-autonomous-robocop-car-that-learn-1822423481
144	11	Gizmodo	Rupert Murdoch's MySpace Apparently Still Haunts Mark Zuckerberg	12/02/2018	https://gizmodo.com/rupert-murdochs-myspace-apparently-still-haunts-mark-zu-1822924885
145	11	Gizmodo	Even When Spotting Gender, Current Face Recognition Tech Works Better for White Dudes	12/02/2018	https://gizmodo.com/even-when-spotting-gender-current-face-recognition-tec-1822929750
146	11	Gizmodo	Tech History Group Dedicated to Preserving Information Busted Deleting Apology Tweets [Updated]	12/02/2018	https://gizmodo.com/tech-history-group-dedicated-to-preserving-information-1822931470
147	11	Gizmodo	Why This AI Chatbot Is a Risky Way to Document Workplace Harassment	15/02/2018	https://gizmodo.com/why-this-ai-chatbot-is-a-risky-way-to-document-workplac-1823033499
148	11	Gizmodo	Conservative Twitter Users Lose Thousands of Followers, Mass Purge of Bots Suspected [Updated]	21/02/2018	https://gizmodo.com/conservative-twitter-users-lose-thousands-of-followers-1823185428
149	11	Gizmodo	Justice Department Drops $2 Million to Research Crime-Fighting AI	27/02/2018	https://gizmodo.com/justice-department-drops-2-million-to-research-crime-f-1823367404
150	11	Gizmodo	Google Is Helping the Pentagon Build AI for Drones	06/03/2018	https://gizmodo.com/google-is-helping-the-pentagon-build-ai-for-drones-1823464533
151	11	Gizmodo	The University of Arizona Tracked Students’ ID Card Swipes to Predict Who Would Drop Out	09/03/2018	https://gizmodo.com/the-university-of-arizona-tracked-students-id-card-swi-1823654183
152	11	Gizmodo	Study Finds Predictive Policing No More Racist Than Regular Policing	13/03/2018	https://gizmodo.com/study-finds-predictive-policing-no-more-racist-than-reg-1823733844
153	11	Gizmodo	Apple Recruits Google's Search and AI Chief John Giannandrea to Help Improve Siri	03/04/2018	https://gizmodo.com/apple-lures-googles-search-and-ai-chief-john-giannandre-1824304602
154	11	Gizmodo	Mark Zuckerberg Is Deluded If He Thinks AI Can Solve Facebook’s Hate Speech Problem	11/04/2018	https://gizmodo.com/mark-zuckerberg-is-deluded-if-he-thinks-ai-can-solve-fa-1825177589
155	11	Gizmodo	Searching for ‘Cat Meat’ on Yelp Brought Up Chinese Restaurants Until Today	25/04/2018	https://gizmodo.com/searching-for-cat-meat-on-yelp-brought-up-chinese-res-1825546069
156	11	Gizmodo	Blame Facebook for That Moronic Diamond and Silk Hearing	26/04/2018	https://gizmodo.com/blame-facebook-for-that-moronic-diamond-and-silk-hearin-1825570032
157	11	Gizmodo	NYC Launches Task Force to Study How Government Algorithms Impact Your Life	16/05/2018	https://gizmodo.com/nyc-launches-task-force-to-study-how-government-algorit-1826087643
158	11	Gizmodo	Stop This BLATANT CENSORSHIP': The Poor, Confused Souls Sending Their YouTube Complaints to the FCC	23/05/2018	https://gizmodo.com/stop-this-blatant-censorship-the-poor-confused-souls-1826197652
159	12	TechCrunch	Twitter hits back again at claims that its employees monitor direct messages	16/01/2018	https://techcrunch.com/2018/01/15/twitter-hits-back-again-at-claims-that-its-employees-monitor-direct-messages/
160	12	TechCrunch	Study shows software used to predict repeat offenders is no better at it than untrained humans	17/01/2018	https://techcrunch.com/2018/01/17/study-shows-software-used-to-predict-repeat-offenders-is-no-better-at-it-than-untrained-humans/
161	12	TechCrunch	WTF is GDPR?	20/01/2018	https://techcrunch.com/2018/01/20/wtf-is-gdpr/
162	12	TechCrunch	Why inclusion in the Google Arts & Culture selfie feature matters	21/01/2018	https://techcrunch.com/2018/01/21/why-inclusion-in-the-google-arts-culture-selfie-feature-matters/
163	12	TechCrunch	Rupert Murdoch wants Facebook to pay for the news	22/01/2018	https://techcrunch.com/2018/01/22/rupert-murdoch-wants-facebook-to-pay-for-the-news/
164	12	TechCrunch	Google tweaks search snippets to try to stop serving wrong, stupid and biased answers	31/01/2018	https://techcrunch.com/2018/01/31/google-tweaks-search-snippets-to-try-to-stop-serving-wrong-stupid-and-biased-answers/
165	12	TechCrunch	Factmata closes $1M seed round as it seeks to build an ‘anti fake news’ media platform	01/02/2018	https://techcrunch.com/2018/02/01/factmata-closes-1m-seed-round-as-it-seeks-to-build-an-anti-fake-news-media-platform/
166	12	TechCrunch	Google fined $21.1M for search bias in India	09/02/2018	https://techcrunch.com/2018/02/09/google-fined-21-1m-for-search-bias-in-india/
167	12	TechCrunch	Dbrain pitches a new token paying users crypto to train artificial intelligence	13/02/2018	https://techcrunch.com/2018/02/13/dbrain-pitches-a-new-token-paying-users-crypto-to-train-artificial-intelligence/
168	12	TechCrunch	Uncommon.co launches and raises $18m to bring objectivity and efficiency to hiring	15/02/2018	https://techcrunch.com/2018/02/15/uncommon-iq-launches-and-raises-18m-to-bring-objectivity-and-efficiency-to-hiring/
169	12	TechCrunch	IBM Watson CTO Rob High on bias and other challenges in machine learning	27/02/2018	https://techcrunch.com/2018/02/27/ibm-watson-cto-rob-high-on-bias-and-other-challenges-in-machine-learning/
170	12	TechCrunch	AI will create new jobs but skills must shift, say tech giants	28/02/2018	https://techcrunch.com/2018/02/28/ai-will-create-new-jobs-but-skills-must-shift-say-tech-giants/
171	12	TechCrunch	False news spreads faster than truth online thanks to human nature	08/03/2018	https://techcrunch.com/2018/03/08/false-news-spreads-faster-than-truth-online-thanks-to-human-nature/
172	12	TechCrunch	Why achieving diversity in tech requires firing the industry’s ‘cultural fit’ mentality	22/03/2018	https://techcrunch.com/2018/03/22/why-achieving-diversity-in-tech-requires-firing-the-industrys-cultural-fit-mentality/
173	12	TechCrunch	France wants to become an artificial intelligence hub	29/03/2018	https://techcrunch.com/2018/03/29/france-wants-to-become-an-artificial-intelligence-hub/
174	12	TechCrunch	Terry Myerson, EVP of Windows and Devices, is leaving Microsoft, prompting a big AI, Azure and devices reorganization	29/03/2018	https://techcrunch.com/2018/03/29/terry-myerson-evp-of-windows-and-devices-is-leaving-microsoft-prompting-a-big-ai-azure-and-devices-reorganization/
175	12	TechCrunch	News startup Knowhere aims to break through partisan echo chambers	05/04/2018	https://techcrunch.com/2018/04/04/news-startup-knowhere-aims-to-break-through-partisan-echo-chambers/
176	12	TechCrunch	UK report urges action to combat AI bias	16/04/2018	https://techcrunch.com/2018/04/16/uk-report-urges-action-to-combat-ai-bias/
177	12	TechCrunch	Two Facebook and Google geniuses are combining search and AI to transform HR	17/04/2018	https://techcrunch.com/2018/04/17/two-facebook-and-google-geniuses-are-combining-search-and-ai-to-transform-hr/
178	12	TechCrunch	What do AI and blockchain mean for the rule of law?	12/05/2018	https://techcrunch.com/2018/05/12/what-do-ai-and-blockchain-mean-for-the-rule-of-law/
179	12	TechCrunch	Riminder raises $2.3 million for its AI recruitment service	25/05/2018	https://techcrunch.com/2018/05/25/riminder-raises-2-3-million-for-its-ai-recruitment-service/
180	13	ReCode	Full transcript: Recode Executive Editor Kara Swisher on Recode Media	02/01/2018	https://www.recode.net/2018/1/2/16842052/transcript-kara-swisher-13-media-peter-kafka-dan-frommer-uber-susan-fowler-trump-2017-journalism
181	13	ReCode	Rupert Murdoch says Facebook needs to pay publishers the way cable companies do	22/01/2018	https://www.recode.net/2018/1/22/16920264/rupert-murdoch-facebook-google-pay-publishers-cable-carriage-fees
182	13	ReCode	Full transcript: Journalist and author Joanne Lipman on Recode Decode	28/03/2018	https://www.recode.net/2018/3/28/17171222/transcript-journalist-author-joanne-lipman-what-she-said-book-13-decode
183	13	ReCode	Full transcript: Yelp CEO Jeremy Stoppelman on Recode Decode	15/04/2018	https://www.recode.net/2018/4/15/17240588/yelp-ceo-jeremy-stoppelman-13-decode-transcript
184	13	ReCode	McKinsey’s latest AI research predicts it could create trillions worth of value ... Someday	19/04/2018	https://www.recode.net/2018/4/19/17251788/mckinsey-artificial-intelligence-research-report-michael-chui
185	13	ReCode	Full transcript: ‘Microtrends’ author and political strategist Mark Penn on Recode Decode	18/05/2018	https://www.recode.net/2018/5/18/17366344/transcript-microtrends-author-politics-mark-penn-13-decode
186	13	ReCode	Full video and transcript: U.S. Senator Mark Warner at Code 2018	30/05/2018	https://www.recode.net/2018/5/30/17397134/senator-mark-warner-virginia-transcript-code-2018
187	12	TechCrunch	Social media is giving us trypophobia	27/01/2018	https://techcrunch.com/2018/01/27/social-media-is-giving-us-trypophobia/

APPENDIX 9 - MAINSTREAM ARTICLES AND IDs

_ID	Media	Media Name	Headline	Date	URL
1	1	BBC News	I didn't even meet my potential employers'	06/02/2018	https://www.bbc.co.uk/news/business-42905515
2	1	BBC News	YouTube toughens advert payment rules	17/01/2018	https://www.bbc.co.uk/news/technology-42716393
3	1	BBC News	Are you scared yet? Meet Norman, the psychopathic AI	01/06/2018	https://www.bbc.co.uk/news/technology-44040008
4	1	BBC News	Facebook to use surveys to boost 'trustworthy' news	19/01/2018	https://www.bbc.co.uk/news/technology-42755832
5	1	BBC News	Four things to make us understand our AI colleagues	19/03/2018	http://www.bbc.com/capital/story/20180316-why-a-robot-wont-steal-your-job-yet
6	1	BBC News	Are sex robots just turning women into literal objects?	06/04/2018	https://www.bbc.co.uk/bbcthree/article/8bbe0749-62ee-40f9-a8ac-a2d751c474f6
7	1	BBC News	As it happens: Zuckerberg takes the blame	10/04/2018	https://www.bbc.co.uk/news/live/world-us-canada-43706880
8	1	BBC News	Facebook's Zuckerberg says his data was harvested	11/04/2018	https://www.bbc.co.uk/news/technology-43725643
9	1	BBC News	Mark Zuckerberg's dreaded homework assignments	12/04/2018	https://www.bbc.co.uk/news/technology-43735385
10	1	BBC News	Is this the year 'weaponised' AI bots do battle?	01/05/2018	https://www.bbc.co.uk/news/business-42559967
11	2	The Guardian	As technology develops, so must journalists' codes of ethics	21/01/2018	https://www.theguardian.com/commentisfree/2018/jan/21/technology-codes-ethics-ai-artificial-intelligence
12	2	The Guardian	How an ex-YouTube insider investigated its secret algorithm	02/02/2018	https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot
13	2	The Guardian	Enemies, traitors, saboteurs: how can we face the future with this anger in our politics?	17/02/2018	https://www.theguardian.com/commentisfree/2018/feb/17/language-public-social-media-politics-repercussions
14	2	The Guardian	Dehumanising, impenetrable, frustrating': the grim reality of job hunting in the age of AI	04/03/2018	https://www.theguardian.com/inequality/2018/mar/04/dehumanising-impenetrable-frustrating-the-grim-reality-of-job-hunting-in-the-age-of-ai
15	2	The Guardian	How to persuade a robot that you should get the job	04/03/2018	https://www.theguardian.com/technology/2018/mar/04/robots-screen-candidates-for-jobs-artificial-intelligence
16	2	The Guardian	Women must act now, or male-designed robots will take over our lives	12/03/2018	https://www.theguardian.com/commentisfree/2018/mar/13/women-robots-ai-male-artificial-intelligence-automation
17	2	The Guardian	Algorithms have become so powerful we need a robust, Europe-wide response	04/04/2018	https://www.theguardian.com/commentisfree/2018/apr/04/algorithms-powerful-europe-response-social-media
18	2	The Guardian	Mark Zuckerberg faces tough questions in two-day congressional testimony – as it happened	11/04/2018	https://www.theguardian.com/technology/live/2018/apr/11/mark-zuckerberg-testimony-live-updates-house-congress-cambridge-analytica?page=with:block-5ac
19	2	The Guardian	The Facebook hearings remind us: information warfare is here to stay	12/04/2018	https://www.theguardian.com/commentisfree/2018/apr/12/mark-zuckerberg-facebook-congressional-hearing-information-warfare-normal
20	2	The Guardian	The People vs Tech by Jamie Bartlett review – once more into the digital apocalypse	16/04/2018	https://www.theguardian.com/books/2018/apr/16/the-people-vs-tech-review-jamie-bartlett-silicon-valley
21	2	The Guardian	The Guardian view on artificial intelligence: not a technological problem	16/04/2018	https://www.theguardian.com/commentisfree/2018/apr/16/the-guardian-view-on-artificial-intelligence-not-a-technological-problem
22	2	The Guardian	Artificial intelligence, robots and a human touch	19/04/2018	https://www.theguardian.com/technology/2018/apr/19/artificial-intelligence-robots-and-a-human-touch
23	2	The Guardian	UK accused of flouting human rights in 'racialised' war on gangs	09/05/2018	https://www.theguardian.com/uk-news/2018/may/09/uk-accused-flouting-human-rights-racialised-war-gangs
24	2	The Guardian	We created poverty. Algorithms won't make that go away	13/05/2018	https://www.theguardian.com/commentisfree/2018/may/13/we-created-poverty-algorithms-wont-make-that-go-away
25	2	The Guardian	Employers are monitoring computers, toilet breaks – even emotions. Is your boss watching you?	14/05/2018	https://www.theguardian.com/world/2018/may/14/is-your-boss-secretly-or-not-so-secretly-watching-you
26	2	The Guardian	Artificial intelligence: it may be our future but can it write a play?	16/05/2018	https://www.theguardian.com/stage/2018/may/16/artificial-intelligence-it-may-be-our-future-but-can-it-write-a-play
27	2	The Guardian	Neuroscientist Hannah Critchlow: 'Consciousness is a really funny word'	18/05/2018	https://www.theguardian.com/science/2018/may/18/neuroscientist-hannah-critchlow-cambridge-consciousness-funny-word
28	2	The Guardian	Five things we learned from Mark Zuckerberg's European parliament appearance	22/05/2018	https://www.theguardian.com/technology/2018/may/22/five-things-we-learned-from-mark-zuckerbergs-european-parliament-appearance
29	2	The Guardian	Mark Zuckerberg appears before European parliament – as it happened	22/05/2018	https://www.theguardian.com/technology/live/2018/may/22/mark-zuckerberg-facebook-appears-before-european-parliament-live-updates?page=with:block-5b(
30	2	The Guardian	Six reasons why social media is a Bummer	27/05/2018	https://www.theguardian.com/technology/2018/may/27/jaron-lanier-six-reasons-why-social-media-is-a-bummer
31	2	The Guardian	Police trial AI software to help process mobile phone evidence	27/05/2018	https://www.theguardian.com/uk-news/2018/may/27/police-trial-ai-software-to-help-process-mobile-phone-evidence
32	2	The Guardian	New publisher plans to offer budding authors £24,000 salary	01/06/2018	https://www.theguardian.com/books/2018/jun/01/new-publisher-plans-to-offer-budding-authors-24000-salary
33	2	The Guardian	Facebook scraps 'outdated' trending news section	01/06/2018	https://www.theguardian.com/technology/2018/jun/01/facebook-scraps-outdated-trending-news-section
34	3	Daily Mail	The REAL reason women get passed over for promotions and the best jobs – and what to do about it	04/01/2018	http://www.dailymail.co.uk/femail/article-5233847/The-reason-women-passed-promotions-jobs.html
35	3	Daily Mail	Minority Report-style AI used by courts to predict whether criminals will re-offend is 'no more accurate than untrained humans'	18/01/2018	http://www.dailymail.co.uk/sciencetech/article-5283337/Computer-software-accurate-UNTRAINED-people.html
36	3	Daily Mail	Are computers turning into bigots? Machines that learn to think for themselves are revolutionising the workplace. But there's growing evidence they are acquiring racial and class prejudices too	25/01/2018	http://www.dailymail.co.uk/sciencetech/article-5310031/Evidence-robots-acquiring-racial-class-prejudices.html
37	3	Daily Mail	Concerns in the courtroom as algorithms are used to help judges rule on jail time which can flag black people as twice as likely to re-offend compared to white defendants	31/01/2018	http://www.dailymail.co.uk/news/article-5333757/AI-court-When-algorithms-rule-jail-time.html
38	3	Daily Mail	Will Ford's driverless 'Robocops' lead to bias-free policing? Experts say the system could improve accuracy in law enforcement - but others warn it could lead to dangerous abuses of power	06/02/2018	http://www.dailymail.co.uk/sciencetech/article-5359463/Critics-Fords-driverless-cop-car-raises-privacy-concerns.html
39	3	Daily Mail	Is facial recognition technology RACIST? Study finds popular face ID systems are more likely to work for white men	12/02/2018	http://www.dailymail.co.uk/sciencetech/article-5382979/Study-finds-popular-face-ID-systems-racial-bias.html
40	3	Daily Mail	Humanity will abandon speech and communicate through a 'collective AI consciousness' using nothing but THOUGHTS by 2050	16/02/2018	http://www.dailymail.co.uk/sciencetech/article-5380573/By-2050-humans-communicate-completely-without-words.html
41	3	Daily Mail	Secretive big-data startup Palantir is testing a controversial system in New Orleans that can predict the likelihood of someone committing a crime, report claims	28/02/2018	http://www.dailymail.co.uk/sciencetech/article-5447037/Secretive-startup-uses-predictive-policing-New-Orleans.html
42	3	Daily Mail	Google is letting the Pentagon use its AI to analyze drone footage in secretive 'Project Maven' deal, report claims	06/03/2018	http://www.dailymail.co.uk/sciencetech/article-5469469/Google-working-Pentagon-secretive-AI-drone-project.html
43	3	Daily Mail	Are Alexa and Siri SEXIST? Expert says AI assistants struggle to understand female speech because it is quieter and more 'breathy'	14/03/2018	http://www.dailymail.co.uk/sciencetech/article-5499339/AI-assistants-sexist-understand-men-better.html
44	3	Daily Mail	Companies using controversial AI software to help hire or fire employees - and gauge whether they like their boss	29/03/2018	http://www.dailymail.co.uk/sciencetech/article-5556221/New-AI-software-help-companies-hire-fire-employees-gauge-like-job.html
45	3	Daily Mail	Google should not be in the business of war': Over 3,000 employees pen letter urging CEO to pull out of the Pentagon's controversial AI drone research, citing firm's 'Don't Be Evil' motto	05/04/2018	http://www.dailymail.co.uk/sciencetech/article-5583707/Thousands-Google-employees-pen-letter-urging-CEO-pull-controversial-Pentagon-AI-project.html
46	3	Daily Mail	The real Minority Report: AI that studies CCTV to predict crime BEFORE it happens will be rolled out in India	11/04/2018	http://www.dailymail.co.uk/sciencetech/article-5603367/AI-studies-CCTV-predict-crime-happens-rolled-India.html
47	3	Daily Mail	Who are Diamond and Silk and why is Mark Zuckerberg being questioned about them in Congress?	11/04/2018	http://www.dailymail.co.uk/news/article-5605041/Who-Diamond-Silk-Mark-Zuckerberg-questioned-Congress.html
48	3	Daily Mail	Elon Musk slams Silicon Valley's lack of regulation on social media, citing 'willy nilly proliferation of fake news' among the ways it 'negatively affects the public good'	11/04/2018	http://www.dailymail.co.uk/sciencetech/article-5604291/Teslas-Elon-Musk-says-AI-social-media-regulated-stop-spread-fake-news.html
49	3	Daily Mail	The Latest: Reddit says it banned 944 suspicious accounts	11/04/2018	http://www.dailymail.co.uk/wires/ap/article-5597699/The-Latest-Is-Facebook-really-changing-just-stalling.html
50	3	Daily Mail	Secret surveillance software created by EX-SPIES is harvesting Facebook photos to create a huge facial recognition database that could be used to monitor people worldwide	17/04/2018	http://www.dailymail.co.uk/sciencetech/article-5624383/Surveillance-company-run-ex-spies-harvesting-Facebook-photos.html
51	3	Daily Mail	Are social networks SEXIST? New research shows how algorithms on popular platforms like Instagram may discriminate against women	23/04/2018	http://www.dailymail.co.uk/sciencetech/article-5649291/New-research-shows-algorithms-popular-platforms-like-Instagram-discriminate-against-women.html
52	3	Daily Mail	Facial recognition AI built into police body cameras will suffer from racial bias and could lead to false arrests, civil liberties experts warn	27/04/2018	http://www.dailymail.co.uk/sciencetech/article-5663953/Facial-recognition-AI-built-police-body-cameras-lead-FALSE-ARRESTS-experts-warn.html
53	3	Daily Mail	Controversial Gangs Matrix database `not the answer´ to violent crime wave	09/05/2018	http://www.dailymail.co.uk/wires/pa/article-5706905/Controversial-Gangs-Matrix-database-not-answer-violent-crime-wave.html
54	3	Daily Mail	Is Amazon's facial recognition system RACIST? Expert claims AI 'discriminates against black faces'	24/05/2018	http://www.dailymail.co.uk/sciencetech/article-5768539/Is-Amazons-facial-recognition-RACIST-Expert-says-distinguish-black-faces.html
55	3	Daily Mail	Facebook shuts down troubled `trending´ section	01/06/2018	http://www.dailymail.co.uk/wires/pa/article-5796239/Facebook-shuts-troubled-trending-section.html
56	3	Daily Mail	Facebook kills off trending topics in battle to stop 'fake news'	01/06/2018	http://www.dailymail.co.uk/sciencetech/article-5795693/Facebook-kills-trending-topics-tests-breaking-news-label.html
57	4	Huffington Post	My Five Predictions for 2018 From an Ever-Cloudy Crystal Ball	02/01/2018	https://www.huffingtonpost.com/entry/my-five-predictions-for-2018-from-an-ever-cloudy-crystal_us_5a4babbbe4b06cd2bd03e28a
58	4	Huffington Post	Top 5 Tech Trends for 2018	04/01/2018	https://www.huffingtonpost.com/entry/top-5-tech-trends-for-2018_us_5a4e6be1e4b0ee59d41c090e
59	4	Huffington Post	Fake news isn't always untrue: The coverage of 'Fire and Fury' and Trump's mental fitness | The Knife Media	09/01/2018	https://www.huffingtonpost.com/entry/fake-news-isnt-always-untrue-the-coverage-of-fire_us_5a552b3fe4b0f9b24bf31b61
60	4	Huffington Post	Artificial Intelligence Will Affect The News We Consume. Whether That's A Good Thing Is Up To Humans.	04/05/2018	https://www.huffingtonpost.com/entry/artificial-intelligence-news-ethics_us_5ac49e86e4b0ac473edb9084
61	5	Sky News	Zuckerberg's own data 'sold to malicious third parties'	11/04/2018	https://news.sky.com/story/live-zuckerberg-faces-second-day-of-questions-11326338
62	5	Sky News	Ethics must be at centre of AI technology, says Lords report	16/04/2018	https://news.sky.com/story/ethics-must-be-at-centre-of-ai-technology-says-lords-report-11333333
63	6	MSN News	Can an AI-powered bot help parents raise better humans?	04/03/2018	https://www.msn.com/en-gb/news/techandscience/can-an-ai-powered-bot-help-parents-raise-better-humans/ar-BBKdgeL?li=AAnZ9Ug
64	6	MSN News	False news stories travel faster and farther on Twitter than the truth	08/03/2018	https://www.msn.com/en-gb/news/techandscience/false-news-stories-travel-faster-and-farther-on-twitter-than-the-truth/ar-BBK1JGW?li=AAnZ9Ug
65	6	MSN News	Facebook's crisis demands a reevaluation of computer science itself	31/03/2018	https://www.msn.com/en-gb/news/techandscience/facebook%E2%80%99s-crisis-demands-a-reevaluation-of-computer-science-itself/ar-AAvj5qf
66	6	MSN News	9 essential lessons from psychology to understand the Trump era	14/04/2018	https://www.msn.com/en-gb/news/world/9-essential-lessons-from-psychology-to-understand-the-trump-era/ar-AAvO4xX
67	6	MSN News	How AI is powering Airbnb's mission to change how we travel forever	17/04/2018	https://www.msn.com/en-gb/news/other/how-ai-is-powering-airbnb%E2%80%99s-mission-to-change-how-we-travel-forever/ar-AAvZ5wV
68	6	MSN News	UK accused of flouting human rights in 'racialised' war on gangs	08/05/2018	https://www.msn.com/en-gb/news/uknews/uk-accused-of-flouting-human-rights-in-racialised-war-on-gangs/ar-AAwYtlb
69	6	MSN News	Police gang database 'breaches human rights and discriminates against young black men who are fans of grime music'	08/05/2018	https://www.msn.com/en-gb/news/uknews/police-gang-database-breaches-human-rights-and-discriminates-against-young-black-men-who-are-fans-of-grime-music/ar-AAwYrKL
70	6	MSN News	Scotland Yard's gang 'matrix' under fire for targeting people 'who pose no risk of violence'	08/05/2018	https://www.msn.com/en-gb/news/newslondon/scotland-yards-gang-matrix-under-fire-for-targeting-people-who-pose-no-risk-of-violence/ar-AAwYDhk
71	6	MSN News	Facebook shuts down troubled 'trending' section amid fake news controversies	01/06/2018	https://www.msn.com/en-gb/news/newslondon/facebook-shuts-down-troubled-trending-section-amid-fake-news-controversies/ar-AAy7grY
72	7	The Telegraph	Big banks struggling to cater for clued up Generation Z's needs	25/02/2018	https://www.telegraph.co.uk/business/2018/02/25/big-banks-struggling-cater-clued-generation-zs-needs/
73	7	The Telegraph	The robot will interview you now: AI could revolutionise recruitment by weeding out bias	17/03/2018	https://www.telegraph.co.uk/business/2018/03/17/firms-calling-robots-navigate-recruitment-metoo-world/
74	7	The Telegraph	Want better AI? Start educating it like a child	23/03/2018	https://www.telegraph.co.uk/business/essential-insights/artificial-intelligence-future/
75	7	The Telegraph	Eight tools to help increase business productivity	06/04/2018	https://www.telegraph.co.uk/business/future-technologies/tools-to-increase-productivity/
76	7	The Telegraph	Robots interviewing graduates for jobs at top city firms as students practice how to impress AI	21/04/2018	https://www.telegraph.co.uk/news/2018/04/21/robots-interviewing-graduates-jobs-top-city-firms-students-practice/
77	7	The Telegraph	Robots will never take jobs because work gives people 'dignity', Microsoft boss says	27/05/2018	https://www.telegraph.co.uk/technology/2018/05/27/robots-will-never-take-jobs-work-gives-people-dignity-microsoft/

APPENDIX 10 - ICR TEST RESULTS - 2 ARTICLES PILOT

ReCal2 - Intercoder Reliability Test - Results for Initial pilot with 2 articles

n columns	132
n variables	66
n coders per var	2

Variable	Column	Percent Agreement	Scott's Pi	Cohen's Kappa	Krippendorff's Alpha	N Agreements	N Disagreements	N Cases	N Decisions
Posture	Variable 1 (cols 1 & 2)	0	-0.333333333	0	0	0	2	2	4
Main_topic	Variable 2 (cols 3 & 4)	100	undefined*	undefined*	undefined*	2	0	2	4
Algo_ReduceReinforce	Variable 3 (cols 5 & 6)	50	-0.333333333	0	0	1	1	2	4
a_Ethics	Variable 4 (cols 7 & 8)	50	-0.333333333	0	0	1	1	2	4
b_accountability	Variable 5 (cols 9 & 10)	100	undefined*	undefined*	undefined*	2	0	2	4
c_BlackBox	Variable 6 (cols 11 & 12)	100	undefined*	undefined*	undefined*	2	0	2	4
d_EchoChambers	Variable 7 (cols 13 & 14)	100	1	1	1	2	0	2	4
e_Antitrust	Variable 8 (cols 15 & 16)	100	undefined*	undefined*	undefined*	2	0	2	4
f_AlwaysBiased	Variable 9 (cols 17 & 18)	100	undefined*	undefined*	undefined*	2	0	2	4
g_human	Variable 10 (cols 19 & 20)	100	undefined*	undefined*	undefined*	2	0	2	4
h_BiasDenial	Variable 11 (cols 21 & 22)	100	undefined*	undefined*	undefined*	2	0	2	4
i_reflection	Variable 12 (cols 23 & 24)	100	1	1	1	2	0	2	4
j_ProfIssue	Variable 13 (cols 25 & 26)	100	undefined*	undefined*	undefined*	2	0	2	4
k_InnoInnvest	Variable 14 (cols 27 & 28)	100	undefined*	undefined*	undefined*	2	0	2	4
l_UserTrust	Variable 15 (cols 29 & 30)	100	undefined*	undefined*	undefined*	2	0	2	4
m_SelfReg_External	Variable 16 (cols 31 & 32)	100	undefined*	undefined*	undefined*	2	0	2	4
n_Reg_Slow	Variable 17 (cols 33 & 34)	100	undefined*	undefined*	undefined*	2	0	2	4
a_Fakenews	Variable 18 (cols 35 & 36)	100	undefined*	undefined*	undefined*	2	0	2	4
b_PolInter	Variable 19 (cols 37 & 38)	50	-0.333333333	0	0	1	1	2	4
c_personnel	Variable 20 (cols 39 & 40)	50	-0.333333333	0	0	1	1	2	4
d_Jud	Variable 21 (cols 41 & 42)	100	undefined*	undefined*	undefined*	2	0	2	4
e_Pol	Variable 22 (cols 43 & 44)	100	undefined*	undefined*	undefined*	2	0	2	4
f_welfare	Variable 23 (cols 45 & 46)	100	undefined*	undefined*	undefined*	2	0	2	4
g_credit	Variable 24 (cols 47 & 48)	100	undefined*	undefined*	undefined*	2	0	2	4
h_educ	Variable 25 (cols 49 & 50)	100	undefined*	undefined*	undefined*	2	0	2	4
i_gender	Variable 26 (cols 51 & 52)	100	undefined*	undefined*	undefined*	2	0	2	4
j_ethnic	Variable 27 (cols 53 & 54)	100	undefined*	undefined*	undefined*	2	0	2	4
k_Privacy	Variable 28 (cols 55 & 56)	50	-0.333333333	0	0	1	1	2	4
a_MoreEthics	Variable 29 (cols 57 & 58)	100	undefined*	undefined*	undefined*	2	0	2	4
b_diversity	Variable 30 (cols 59 & 60)	100	undefined*	undefined*	undefined*	2	0	2	4
c_MoreTrans	Variable 31 (cols 61 & 62)	50	-0.333333333	0	0	1	1	2	4
d_ChangeModels	Variable 32 (cols 63 & 64)	100	undefined*	undefined*	undefined*	2	0	2	4
e_ConsumerKnow	Variable 33 (cols 65 & 66)	100	1	1	1	2	0	2	4
CA_MainTopic	Variable 34 (cols 67 & 68)	100	undefined*	undefined*	undefined*	2	0	2	4
CA_Mention	Variable 35 (cols 69 & 70)	100	1	1	1	2	0	2	4
MainCompanies	Variable 36 (cols 71 & 72)	100	undefined*	undefined*	undefined*	2	0	2	4
a_medical	Variable 37 (cols 73 & 74)	100	undefined*	undefined*	undefined*	2	0	2	4
b_SoMe	Variable 38 (cols 75 & 76)	100	1	1	1	2	0	2	4
c_Fin	Variable 39 (cols 77 & 78)	100	undefined*	undefined*	undefined*	2	0	2	4
d_Gov	Variable 40 (cols 79 & 80)	100	undefined*	undefined*	undefined*	2	0	2	4
e_legal	Variable 41 (cols 81 & 82)	100	undefined*	undefined*	undefined*	2	0	2	4
f_police	Variable 42 (cols 83 & 84)	100	undefined*	undefined*	undefined*	2	0	2	4
g_EducSys	Variable 43 (cols 85 & 86)	100	undefined*	undefined*	undefined*	2	0	2	4
h_SearchPlat	Variable 44 (cols 87 & 88)	100	undefined*	undefined*	undefined*	2	0	2	4
Profit	Variable 45 (cols 89 & 90)	100	undefined*	undefined*	undefined*	2	0	2	4
a_ProReg	Variable 46 (cols 91 & 92)	100	undefined*	undefined*	undefined*	2	0	2	4
b_AntiReg	Variable 47 (cols 93 & 94)	50	-0.333333333	0	0	1	1	2	4
c_SelfReg	Variable 48 (cols 95 & 96)	100	undefined*	undefined*	undefined*	2	0	2	4
d_MultiReg	Variable 49 (cols 97 & 98)	50	-0.333333333	0	0	1	1	2	4
e_Mixed	Variable 50 (cols 99 & 100)	50	-0.333333333	0	0	1	1	2	4
f_NoMention	Variable 51 (cols 101 & 102)	100	1	1	1	2	0	2	4
99_Reg	Variable 52 (cols 103 & 104)	100	undefined*	undefined*	undefined*	2	0	2	4
GovPar_Mention	Variable 53 (cols 105 & 106)	100	1	1	1	2	0	2	4
GovPar_quote	Variable 54 (cols 107 & 108)	100	undefined*	undefined*	undefined*	2	0	2	4
GovPar_ref	Variable 55 (cols 109 & 110)	100	undefined*	undefined*	undefined*	2	0	2	4
a_Empower	Variable 56 (cols 111 & 112)	50	-0.333333333	0	0	1	1	2	4
b_CostEffective	Variable 57 (cols 113 & 114)	50	-0.333333333	0	0	1	1	2	4
Corpo_Mention	Variable 58 (cols 115 & 116)	100	undefined*	undefined*	undefined*	2	0	2	4
Corpo_Quote	Variable 59 (cols 117 & 118)	100	1	1	1	2	0	2	4
Corpo_Ref	Variable 60 (cols 119 & 120)	100	1	1	1	2	0	2	4
CorpoEffort	Variable 61 (cols 121 & 122)	50	-0.333333333	0	0	1	1	2	4
AlgoBias_mention	Variable 62 (cols 123 & 124)	100	undefined*	undefined*	undefined*	2	0	2	4
TechSection	Variable 63 (cols 125 & 126)	100	undefined*	undefined*	undefined*	2	0	2	4
sources	Variable 64 (cols 127 & 128)	100	undefined*	undefined*	undefined*	2	0	2	4
technical	Variable 65 (cols 129 & 130)	100	undefined*	undefined*	undefined*	2	0	2	4
Polemical	Variable 66 (cols 131 & 132)	50	-0.333333333	0	0	1	1	2	4
Total									

APPENDIX 11 - CODING DETAILS BY CODER - 2 ARTICLES PILOT

Coder	Variable	Article ID	
		20	170
1	Posture	2	4
2	Posture	1	5
1	Main_topic	0	0
2	Main_topic	0	0
1	Algo_ReduceReinforce	1	1
2	Algo_ReduceReinforce	1	0
1	a_Ethics	0	1
2	a_Ethics	0	0
1	b_accountability	0	0
2	b_accountability	0	0
1	c_BlackBox	0	0
2	c_BlackBox	0	0
1	d_EchoChambers	1	0
2	d_EchoChambers	1	0
1	e_Antitrust	0	0
2	e_Antitrust	0	0
1	f_AlwaysBiased	0	0
2	f_AlwaysBiased	0	0
1	g_human	0	0
2	g_human	0	0
1	h_BiasDenial	0	0
2	h_BiasDenial	0	0
1	i_reflection	0	1
2	i_reflection	0	1
1	j_ProfIssue	0	0
2	j_ProfIssue	0	0
1	k_InnoInnvest	0	0
2	k_InnoInnvest	0	0
1	l_UserTrust	0	0
2	l_UserTrust	0	0
1	m_SelfReg_External	0	0
2	m_SelfReg_External	0	0
1	n_Reg_Slow	0	0
2	n_Reg_Slow	0	0
1	a_Fakenews	0	0
2	a_Fakenews	0	0
1	b_PolInter	1	0
2	b_PolInter	0	0
1	c_personnel	1	0
2	c_personnel	0	0
1	d_Jud	0	0
2	d_Jud	0	0
1	e_Pol	0	0
2	e_Pol	0	0
1	f_welfare	0	0
2	f_welfare	0	0
1	g_credit	0	0
2	g_credit	0	0
1	h_educ	0	0
2	h_educ	0	0
1	i_gender	0	0
2	i_gender	0	0
1	j_ethnic	0	0
2	j_ethnic	0	0
1	k_Privacy	0	1
2	k_Privacy	0	0
1	a_MoreEthics	0	0
2	a_MoreEthics	0	0
1	b_diversity	0	0
2	b_diversity	0	0
1	c_MoreTrans	0	1
2	c_MoreTrans	0	0
1	d_ChangeModels	0	0
2	d_ChangeModels	0	0
1	e_ConsumerKnow	0	1
2	e_ConsumerKnow	0	1
1	CA_MainTopic	0	0
2	CA_MainTopic	0	0
1	CA_Mention	1	0
2	CA_Mention	1	0
1	MainCompanies	1	1
2	MainCompanies	1	1
1	a_medical	0	0
2	a_medical	0	0
1	b_SoMe	1	0
2	b_SoMe	1	0
1	c_Fin	0	0
2	c_Fin	0	0
1	d_Gov	0	0
2	d_Gov	0	0
1	e_legal	0	0
2	e_legal	0	0
1	f_police	0	0
2	f_police	0	0
1	g_EducSys	0	0
2	g_EducSys	0	0
1	h_SearchPlat	0	0
2	h_SearchPlat	0	0
1	Profit	0	0
2	Profit	0	0
1	a_ProReg	0	0
2	a_ProReg	0	0
1	b_AntiReg	0	0
2	b_AntiReg	0	1
1	c_SelfReg	0	0
2	c_SelfReg	0	0
1	d_MultiReg	0	0
2	d_MultiReg	0	1
1	e_Mixed	0	1
2	e_Mixed	0	0
1	f_NoMention	1	0
2	f_NoMention	1	0
1	N/A_Reg	0	0
2	N/A_Reg	0	0
1	GovPar_Mention	1	0
2	GovPar_Mention	1	0
1	GovPar_quote	0	0
2	GovPar_quote	0	0
1	GovPar_ref	0	0
2	GovPar_ref	0	0
1	a_Empower	0	0
2	a_Empower	0	1
1	b_CostEffective	0	0
2	b_CostEffective	0	1
1	Corpo_Mention	1	1
2	Corpo_Mention	1	1
1	Corpo_Quote	0	1
2	Corpo_Quote	0	1
1	Corpo_Ref	0	1
2	Corpo_Ref	0	1
1	CorpoEffort	0	0
2	CorpoEffort	0	1
1	AlgoBias_mention	0	0
2	AlgoBias_mention	0	0
1	sources	0	0
2	sources	0	0
1	technical	0	0
2	technical	0	0
1	TechSection	0	0
2	TechSection	0	0
1	polemical	1	0
2	polemical	1	1

APPENDIX 12 - ICR TEST RESULTS - 10% SAMPLE

ReCal2 - Intercoder Reliability Test - Results

n columns	132
n variables	66
n coders per var	2

Variable	Column	Percent Agreement	Scott's Pi	Cohen's Kappa	Krippendorff's Alpha	N Agreements	N Disagreements	N Cases	N Decisions
Posture	Variable 1 (cols 1 & 2)	73.68	0.63	0.63	0.64	14	5	19	38
Main_topic	Variable 2 (cols 3 & 4)	89.47	0.68	0.69	0.69	17	2	19	38
Algo_ReduceReinforce	Variable 3 (cols 5 & 6)	100	1	1	1	19	0	19	38
a_Ethics	Variable 4 (cols 7 & 8)	84.21	0.62	0.63	0.63	16	3	19	38
b_accountability	Variable 5 (cols 9 & 10)	89.47	0.6	0.61	0.61	17	2	19	38
c_BlackBox	Variable 6 (cols 11 & 12)	100	1	1	1	19	0	19	38
d_EchoChambers	Variable 7 (cols 13 & 14)	100	1	1	1	19	0	19	38
e_Antitrust	Variable 8 (cols 15 & 16)	100	undefined*	undefined*	undefined*	19	0	19	38
f_AlwaysBiased	Variable 9 (cols 17 & 18)	89.47	-0.06	-0.06	-0.03	17	2	19	38
g_human	Variable 10 (cols 19 & 20)	100	1	1	1	19	0	19	38
h_BiasDenial	Variable 11 (cols 21 & 22)	100	undefined*	undefined*	undefined*	19	0	19	38
i_reflection	Variable 12 (cols 23 & 24)	100	1	1	1	19	0	19	38
j_ProfIssue	Variable 13 (cols 25 & 26)	100	undefined*	undefined*	undefined*	19	0	19	38
k_InnoInnvest	Variable 14 (cols 27 & 28)	100	undefined*	undefined*	undefined*	19	0	19	38
l_UserTrust	Variable 15 (cols 29 & 30)	100	undefined*	undefined*	undefined*	19	0	19	38
m_SelfReg_External	Variable 16 (cols 31 & 32)	100	undefined*	undefined*	undefined*	19	0	19	38
n_Reg_Slow	Variable 17 (cols 33 & 34)	100	undefined*	undefined*	undefined*	19	0	19	38
a_Fakenews	Variable 18 (cols 35 & 36)	94.74	0.64	0.64	0.65	18	1	19	38
b_PolInter	Variable 19 (cols 37 & 38)	89.47	0.76	0.76	0.76	17	2	19	38
c_personnel	Variable 20 (cols 39 & 40)	89.47	0.6	0.61	0.61	17	2	19	38
d_Jud	Variable 21 (cols 41 & 42)	100	1	1	1	19	0	19	38
e_Pol	Variable 22 (cols 43 & 44)	94.74	0.77	0.77	0.78	18	1	19	38
f_welfare	Variable 23 (cols 45 & 46)	100	undefined*	undefined*	undefined*	19	0	19	38
g_credit	Variable 24 (cols 47 & 48)	100	1	1	1	19	0	19	38
h_educ	Variable 25 (cols 49 & 50)	100	undefined*	undefined*	undefined*	19	0	19	38
i_gender	Variable 26 (cols 51 & 52)	100	1	1	1	19	0	19	38
j_ethnic	Variable 27 (cols 53 & 54)	89.47	0.73	0.73	0.74	17	2	19	38
k_Privacy	Variable 28 (cols 55 & 56)	89.47	0.6	0.61	0.61	17	2	19	38
a_MoreEthics	Variable 29 (cols 57 & 58)	84.21	0.31	0.31	0.33	16	3	19	38
b_diversity	Variable 30 (cols 59 & 60)	89.47	0.44	0.46	0.46	17	2	19	38
c_MoreTrans	Variable 31 (cols 61 & 62)	89.47	0.73	0.73	0.74	17	2	19	38
d_ChangeModels	Variable 32 (cols 63 & 64)	100	undefined*	undefined*	undefined*	19	0	19	38
e_ConsumerKnow	Variable 33 (cols 65 & 66)	100	1	1	1	19	0	19	38
CA_MainTopic	Variable 34 (cols 67 & 68)	100	undefined*	undefined*	undefined*	19	0	19	38
CA_Mention	Variable 35 (cols 69 & 70)	100	1	1	1	19	0	19	38
MainCompanies	Variable 36 (cols 71 & 72)	94.74	0.89	0.89	0.89	18	1	19	38
a_medical	Variable 37 (cols 73 & 74)	100	undefined*	undefined*	undefined*	19	0	19	38
b_SoMe	Variable 38 (cols 75 & 76)	94.74	0.85	0.85	0.86	18	1	19	38
c_Fin	Variable 39 (cols 77 & 78)	100	undefined*	undefined*	undefined*	19	0	19	38
d_Gov	Variable 40 (cols 79 & 80)	100	undefined*	undefined*	undefined*	19	0	19	38
e_legal	Variable 41 (cols 81 & 82)	89.47	0.68	0.69	0.69	17	2	19	38
f_police	Variable 42 (cols 83 & 84)	100	1	1	1	19	0	19	38
g_EducSys	Variable 43 (cols 85 & 86)	100	undefined*	undefined*	undefined*	19	0	19	38
h_SearchPlat	Variable 44 (cols 87 & 88)	94.74	0.77	0.77	0.78	18	1	19	38
Profit	Variable 45 (cols 89 & 90)	89.47	-0.06	0	-0.03	17	2	19	38
a_ProReg	Variable 46 (cols 91 & 92)	89.47	0.44	0.46	0.46	17	2	19	38
b_AntiReg	Variable 47 (cols 93 & 94)	94.74	-0.03	0	0	18	1	19	38
c_SelfReg	Variable 48 (cols 95 & 96)	100	undefined*	undefined*	undefined*	19	0	19	38
d_MultiReg	Variable 49 (cols 97 & 98)	73.68	-0.15	0	-0.12	14	5	19	38
e_Mixed	Variable 50 (cols 99 & 100)	73.68	0.12	0.16	0.15	14	5	19	38
f_NoMention	Variable 51 (cols 101 & 102)	89.47	0.79	0.79	0.79	17	2	19	38
99_Reg	Variable 52 (cols 103 & 104)	100	1	1	1	19	0	19	38
GovPar_Mention	Variable 53 (cols 105 & 106)	100	1	1	1	19	0	19	38
GovPar_quote	Variable 54 (cols 107 & 108)	100	1	1	1	19	0	19	38
GovPar_ref	Variable 55 (cols 109 & 110)	89.47	0.6	0.61	0.61	17	2	19	38
a_Empower	Variable 56 (cols 111 & 112)	89.47	0.6	0.6	0.61	17	2	19	38
b_CostEffective	Variable 57 (cols 113 & 114)	94.74	0.77	0.77	0.78	18	1	19	38
Corpo_Mention	Variable 58 (cols 115 & 116)	100	1	1	1	19	0	19	38
Corpo_Quote	Variable 59 (cols 117 & 118)	89.47	0.77	0.78	0.78	17	2	19	38
Corpo_Ref	Variable 60 (cols 119 & 120)	89.47	0.77	0.78	0.78	17	2	19	38
CorpoEffort	Variable 61 (cols 121 & 122)	89.47	0.68	0.68	0.69	17	2	19	38
AlgoBias_mention	Variable 62 (cols 123 & 124)	100	undefined*	undefined*	undefined*	19	0	19	38
TechSection	Variable 63 (cols 125 & 126)	94.74	0.77	0.77	0.78	18	1	19	38
sources	Variable 64 (cols 127 & 128)	94.74	-0.03	0	0	18	1	19	38
technical	Variable 65 (cols 129 & 130)	100	1	1	1	19	0	19	38
Polemical	Variable 66 (cols 131 & 132)	68.42	0.27	0.34	0.29	13	6	19	38
Total		94.26							

APPENDIX 13 - ICR CODING DETAILS BY CODER - 10% SAMPLE

Coder	Variable	Article ID																		
		1	10	20	30	40	50	60	70	80	90	100	110	120	130	140	150	160	170	180
1	Posture	3	1	2	1	5	3	4	1	3	4	2	1	5	1	1	2	1	4	1
2	Posture	2	1	2	1	5	2	4	1	3	4	1	1	5	1	1	1	1	5	1
1	Main_topic	0	0	0	0	0	0	0	1	0	1	1	0	0	1	1	0	0	0	0
2	Main_topic	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0
1	Algo_ReduceReinforce	0	1	1	1	99	99	1	1	1	1	1	1	0	1	1	1	0	1	1
2	Algo_ReduceReinforce	0	1	1	1	99	99	1	1	1	1	1	1	0	1	1	1	0	1	1
1	a_Ethics	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	1	1	1	1
2	a_Ethics	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1
1	b_accountability	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0
2	b_accountability	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0
1	c_BlackBox	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0
2	c_BlackBox	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0
1	d_EchoChambers	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	d_EchoChambers	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	e_Antitrust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	e_Antitrust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	f_AlwaysBiased	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	f_AlwaysBiased	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
1	g_human	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	g_human	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	h_BiasDenial	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	h_BiasDenial	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	i_reflection	1	1	0	0	1	0	1	0	0	0	0	0	0	1	1	0	0	1	0
2	i_reflection	1	1	0	0	1	0	1	0	0	0	0	0	0	1	1	0	0	1	0
1	j_ProfIssue	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	j_ProfIssue	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	k_InnoInnvest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	k_InnoInnvest	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	l_UserTrust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	l_UserTrust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	m_SelfReg_External	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	m_SelfReg_External	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	n_Reg_Slow	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	n_Reg_Slow	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	a_Fakenews	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
2	a_Fakenews	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
1	b_PolInter	0	0	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	0	1
2	b_PolInter	0	0	1	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	1
1	c_personnel	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
2	c_personnel	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	d_Jud	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0
2	d_Jud	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	1	0	0
1	e_Pol	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0
2	e_Pol	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
1	f_welfare	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	f_welfare	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	g_credit	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
2	g_credit	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
1	h_educ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	h_educ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	i_gender	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1
2	i_gender	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1
1	j_ethnic	0	0	0	0	0	0	0	1	0	0	0	1	0	1	1	1	1	0	0
2	j_ethnic	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	0	0
1	k_Privacy	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0
2	k_Privacy	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
1	a_MoreEthics	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	1
2	a_MoreEthics	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
1	b_diversity	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	1
2	b_diversity	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
1	c_MoreTrans	0	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	1	1	0
2	c_MoreTrans	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	1	1	0
1	d_ChangeModels	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	d_ChangeModels	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	e_ConsumerKnow	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
2	e_ConsumerKnow	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	CA_MainTopic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	CA_MainTopic	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	CA_Mention	0	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0
2	CA_Mention	0	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0
1	MainCompanies	0	1	1	1	1	1	0	0	1	1	1	1	0	1	0	1	0	1	0
2	MainCompanies	0	1	1	1	1	1	0	0	1	1	1	1	0	0	0	1	0	1	0
1	a_medical	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	a_medical	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	b_SoMe	0	0	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0
2	b_SoMe	0	0	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
1	c_Fin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	c_Fin	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	d_Gov	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	d_Gov	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	e_legal	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	0
2	e_legal	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0
1	f_police	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
2	f_police	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
1	g_EducSys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	g_EducSys	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	h_SearchPlat	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
2	h_SearchPlat	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
1	Profit	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	Profit	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	a_ProReg	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0	0	0
2	a_ProReg	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
1	b_AntiReg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	b_AntiReg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	c_SelfReg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	c_SelfReg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	d_MultiReg	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	d_MultiReg	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	1	0
1	e_Mixed	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	1
2	e_Mixed	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
1	f_NoMention	1	1	1	1	1	0	0	0	1	1	0	1	1	0	1	1	0	0	0
2	f_NoMention	1	1	1	1	1	0	0	0	1	0	0	1	0	0	1	1	0	0	0
1	N/A_Reg	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
2	N/A_Reg	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	GovPar_Mention	0	0	1	0	0	1	0	1	0	0	0	1	0	1	0	1	0	0	1
2	GovPar_Mention	0	0	1	0	0	1	0	1	0	0	0	1	0	1	0	1	0	0	1
1	GovPar_quote	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	0
2	GovPar_quote	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	0
1	GovPar_ref	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	0	0
2	GovPar_ref	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
1	a_Empower	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	a_Empower	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	b_CostEffective	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	b_CostEffective	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
1	Corpo_Mention	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1
2	Corpo_Mention	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1
1	Corpo_Quote	1	1	0	0	1	1	0	0	1	1	1	1	0	1	0	1	1	1	1
2	Corpo_Quote	1	1	0	0	0	1	0	0	1	1	1	0	0	1	0	1	1	1	1
1	Corpo_Ref	1	1	0	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	0
2	Corpo_Ref	1	1	0	0	0	1	0	0	1	0	1	1	1	1	0	1	1	1	0
1	CorpoEffort	1	0	0	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0
2	CorpoEffort	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0
1	AlgoBias_mention	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	AlgoBias_mention	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	sources	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	sources	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	technical	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	technical	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	TechSection	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	TechSection	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
1	polemical	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	polemical	1	0	1	1	0	1	0	0	1	0	1	0	0	1	0	1	0	1	0