Artificial intelligence in the audiovisual sector

IRIS *Special*

A publication
of the European Audiovisual Observatory

European Audiovisual Observatory

COUNCIL OF EUROPE
CONSEIL DE L'EUROPE

2. The stuff AI dreams are made of – big data

Andrea Pin, Associate Professor of Comparative Public Law, University of Padua

2.1. Introduction

It is commonly said that big data is the oil of the AI revolution.[92] Since data science and technological engineering joined forces, a massive flow of information has flooded the globe, affecting how we live and understand politics, the economy and culture. Thanks to AI's capabilities, the phenomenon of big data has had an enormous, and probably enduring, impact on how individuals and groups make plans, obtain information about themselves and the world, entertain themselves, and socialise.

Nowadays' computers are technologically capacious. Their algorithms are extremely sophisticated. Their neural networks replicate the intellectual processing of human beings and enable them to make complex analyses. By processing big data, firms can anticipate customers' choices and preferences at such an early stage that they can predict what customers want even before *they* do. Thanks to big data, business processes are moving from a "reactive" to a "proactive" approach.[93]

The Internet is playing a fundamental role within this scenario. As individuals use the Internet to share information, even about themselves and their lives, practically without interruption, the web gathers the raw materials from which AI will draw inferences, make guesses, and find out responses to queries. Oxford philosopher Luciano Floridi coined the concept of "onlife" to describe how frequently and unconsciously human beings transition between the real world and the online world.[94]

This phenomenon is escalating. In 2023 it is estimated there will be more than five billion Internet users and 3,6 devices per capita, and 70% of world population will

[92] Pan S. B., "Get to know me: Protecting privacy and autonomy under big data's penetrating gaze", *Harvard Journal of Law and Technology* 30, 2016, p. 239, https://jolt.law.harvard.edu/assets/articlePDFs/v30/30HarvJLTech239.pdf; Surden H., "Artificial intelligence and law: An overview", *Georgia State University Law Review* 35, 2019, pp. 1311 and 1315.
[93] Microsoft Dynamics 365, *Delivering personalized experiences in times of change*, 2007, p. 3, https://www.hso.com/wp-content/uploads/2020/03/Digitally-transforming-customer-experiences-ebook.pdf.
[94] Floridi L., "Soft ethics and the governance of the digital", *Philosophy & Technology* 31, 1, 2018, p. 1.

have mobile connectivity.[95] The more the world is connected, the more big data will be produced. It is not by chance that one of the most hotly currently debated issues is the introduction of 5G networks, since they can provide considerable informational advantage to their owners.

The media field and industry are big players in this scenario. Their job has always consisted in collecting, processing, and disseminating information. Thanks to big data, now they can profile their audience and learn what it expects, how to couch news or to tell a story, or what would be a good finale for a certain movie. Big data allows customisation of the offering through identification of potential news-readers, or movie-goers, as "computers are more accurate than humans at predicting from 'digital footprints' personality traits [or] political attitudes".[96]

The novelty brought about by big data is also changing the media landscape. "… [D]igital TV/movies/music and a myriad of online distribution models have been challenging incumbent distributors (CDs, cable) for years … Online publishers are mining consumer signals from what they read, where they are, the social signals they send –for example what articles they share, what topics are trending on Facebook and Twitter – to serve up personalised, relevant content while not being too repetitive and predictable, thus automating and surpassing what human editors can do".[97] Traditional media now compete in generating news with non-professional information providers that sift through the web searching for news or bloggers that share their views on social media platforms within which distribution and consumption of content are virtually indistinguishable.[98]

This chapter addresses the most relevant legal ramifications of such a global shift in the media world. It touches upon the crucial issue of privacy protection. It then deals with the potential discriminations and bias that a big data-driven strategy can run into and considers the risks of misinformation, polarisation of politics, and the media field becoming a mass surveillance system. Later on, the chapter casts a bird's eye view at how media markets and strategies are changing in light of big data dynamics. Finally, it briefly addresses the debates on the correct regulatory approach to big data.

Overall, the need to regulate AI has gained much traction throughout the years. Although technologies are global and know no border, the regulatory purpose, approach,

---

[95] Cisco, *Cisco Annual International Report (2018-2023) White Paper*, 9 March 2020, https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html?fbclid=IwAR31-e732ws1p1cIW5PYHQjVOJkPSzV0dGt3sq_qkX_P8wb9Q4Yn0Ez0a0Y.

[96] European Data Protection Supervisor, Opinion 7/2015 Meeting the challenges of big data, 19 November 2015, p. 16, https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf.

[97] Byers A., "Big data, big economic impact", 10, 2015, https://kb.osu.edu/bitstream/handle/1811/75420/ISJLP_V10N3_757.pdf?sequence=1&isAllowed=y.
See also Bruckner M. A., "The promise and perils of algorithmic lenders' use of big Data", *Chicago-Kent Law Review* 93, 2018, p. 8, https://scholarship.kentlaw.iit.edu/cklawreview/vol93/iss1/1/ or Ambrose M. L., "Lessons from the Avalanche of Numbers: Big Data in Historical Perspective", *ISJLP*, 11, 2015, p. 213, ("Netflix predicts our movies").

[98] Perritt H. H. Jr., "Technologies of storytelling: New models for movies", *Virginia Sports & Entertainment Law Journal*, 10, 2010, p. 153, http://blogs.kentlaw.iit.edu/perrittseminar/files/2016/07/perritt-technologies-of-storytelling-Westlaw_Document_05_56_44.pdf.

and scheme of the big legal players within this scenario – the United States, the European Union and China – diverge deeply. The US approach is committed to ensuring that markets within which AI is massively deployed remain open and efficient; the EU's paramount concern seems to consist in ensuring that the dignity of the individual is respected; China is mostly preoccupied with social peace, stability, and the ordered development of its economy. Each of these approaches accords big data a specific legal treatment.

2.2. Privacy as the big data gatekeeper

Concerns proliferate that big data-driven tools may integrate in a pervasive system of mass surveillance and manipulation. One of the main safeguards against this threat is privacy. Many countries and supranational legal systems have put in place regulations that limit and monitor what and how information is collected and processed, also with the purpose of constraining big data analytics and preventing social disruption. In this respect, privacy laws serve as a shield against big data's overreach.

2.2.1. The United States of America

The Western world is split in its understanding and protection of privacy. The approaches of the United States and the European Union are far from aligned. Despite its historical sensitiveness to privacy, the United States lacks comprehensive regulation of the collection and gathering of information on the web. Several legal regimes coexist, each regulating a specific sector, without any comprehensive nationwide regulation.[99] The US approach, however, usually sees information as a new, huge market, with positive ramifications for the national economy. While certain states have started implementing pieces of legislation that protect and regulate privacy, with California in a leading position, the collection and gathering of personal data is largely allowed and even promoted. A quite general legal baseline is that the subjects who confer their data should be merely *aware* that their information will be processed in various ways, including for profiling and the trading of their preferences. Since most of the protagonists of the AI-based global industry are based in the US, such a favourable regulatory scheme allows them to fully exploit the advantages of the new oil of data.

[99] Houser K. A. & Voss W. G., "The end of Google and Facebook or a new paradigm in data privacy", *Richmond Journal of Law and Technology*, 25, 2018, p. 18, https://jolt.richmond.edu/files/2018/11/Houser_Voss-FE.pdf.

2.2.2. The European Union

Privacy protection within the European Union is based on the General Data Protection Regulation (GDPR),[100] which was adopted on 27 April 2016 and became applicable as of 25 May 2018. The GDPR itself is the peak of a longer process that has enhanced the protection of personal data over the decades, and represents a very different journey from that of the United States. Although the European Union is committed to making it "easier for business and public authorities to access high quality data to boost growth and create value",[101] the European Union's overall attitude rests on a rejection of the commodification of personal data.[102] The GDPR's legal baseline is that a subject must give his/her *consent* to data processing.[103] Consent itself must be unambiguous, freely given, and well informed:[104] the subject must be given the details about the scope and the purpose of the processing.[105] The GDPR's protection covers EU citizens as well as any other natural persons' data, as long as the processing takes place within the EU. In other words, it protects anyone within its territories.[106]

The gap between the US and the European approaches has created a rift in the exchange of data across the Atlantic. The GDPR is very conservative as to the sharing of information gathered within the European Union, and requires that any data transfer outside EU borders comply with EU standards.[107] The EU regulatory philosophy has been perceived to be so protective of privacy that many non-EU citizens tend to prefer EU-based companies over entities not subject to the jurisdiction of the European Union. Conformance with the GDPR has therefore become a reputation asset for companies working in the field of AI even outside the European Union, pushing them to implement privacy protection rules spontaneously.[108]

---

[100] Consolidated text: Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02016R0679-20160504.

[101] European Data Protection Supervisor, Opinion 3/2020 on the European strategy for data, 16 June 2020, p. 4, https://edps.europa.eu/sites/edp/files/publication/20-06-16_opinion_data_strategy_en.pdf. See also Council of the European Union, Shaping Europe's Digital Future – Council Conclusions, 9 June 2020, https://data.consilium.europa.eu/doc/document/ST-8711-2020-INIT/en/pdf.

[102] European Data Protection Board, Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects, Version 2.0, 8 October 2019, No. 54, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf

[103] Art. 6 GDPR.

[104] Manheim K. & Kaplan L., "Artificial intelligence: Risks to privacy and democracy", *Yale Journal of Law & Technology*, 106, 2019, p. 1069, https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf.

[105] Art. 6, par. 4, and 7, GDPR.

[106] European Data Protection Supervisor, Opinion 3/2018 EDPS Opinion on online manipulation and personal data, 19 March 2018, p. 14, https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf.

[107] Art. 45 GDPR.

[108] Moerel L. & Lyon C., "Commoditization of data is the problem, not the solution – Why placing a price tag on personal information may harm rather than protect consumer privacy, *Future of Privacy Forum*, 24 June

Such a high level of privacy protection from the GDPR comes, however, at a cost. The companies' need to obtain consent from the Internet users who visit their websites translates into a plethora of repetitious, and sometimes obscure, requests for consent that traditionally pop up as soon as a webpage is displayed.[109] This phenomenon has flooded the Internet to the extent that most users simply click "yes" and keep navigating the website without paying attention to how their information is collected, processed and disseminated.[110] This course of action is certainly risky but understandable. Some have made the estimation that a normal person – not a skilled lawyer or a maniacally meticulous Internet user – would waste 76 working days per year reading all the privacy warnings that pop up while he/she is online.[111] Too much privacy protection can be counter-productive: individuals may give away all the protection by consenting in too superficial a manner, thereby allowing massive harvesting of their information.

Moreover, the potentials of big data analysis can weaken the privacy protection accorded by the GDPR on many fronts. First, the GDPR imposes fewer restrictions on anonymised data, as anonymisation is supposed to protect privacy. Thanks to increasing AI capabilities, however, "it is becoming ever easier to infer a person's identity by combining allegedly 'anonymous' data with other datasets including publicly available information for example on social media"[112] … "The bigger and the more comprehensive" a data collection, the more likely it is that an individual whose data has been anonymised will be re-identified.[113]

On top of this, EU privacy rules require that individuals be given detailed information regarding the purpose and scope of the processing of the data they confer. Through neural networks and deep learning, AI-based systems draw inferences that even software developers cannot fully anticipate. This very capacity of big data jeopardises how EU privacy regulation is construed. As big data processing returns results that cannot be fully foreseen, it is extremely difficult to provide individuals with a detailed picture of what their information will be used for.[114]

2020, https://fpf.org/2020/06/24/commoditization-of-data-is-the-problem-not-the-solution-why-placing-a-price-tag-on-personal-information-may-harm-rather-than-protect-consumer-privacy.

[109] European Data Protection Supervisor, Opinion 7/2015 Meeting the challenges of big data, *op. cit.*, p. 11.

[110] Tsesis A., "Marketplace of ideas, privacy, and the digital audience", *Notre Dame Law Review*, 94, 2019, p. 1590, https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=4845&context=ndlr.

[111] Hartzog W., *Privacy's blueprint,* Harvard University Press, 2018.

[112] European Data Protection Supervisor, Opinion 4/2015. Towards a new digital ethics, September 11, 2015, p. 6, https://edps.europa.eu/sites/edp/files/publication/15-09-11_data_ethics_en.pdf.

[113] European Data Protection Supervisor, Opinion 7/2015, "Meeting the challenges of big data", *op. cit.*, p. 15.

[114] AGCM, AGCOM, and Garante per la protezione dei dati personali, Indagine conoscitiva sui *Big Data*, p. 25-26, https://www.agcom.it/documents/10179/17633816/Documento+generico+10-02-2020+1581346981452/39c08bbe-1c02-43dc-bb8e-6d1cc9ec0fcf?version=1.0. The document explains how "dynamic consent" is taking off as a viable option within the EU privacy regulatory scheme. This concept understands consent as a gradual process, during which the subject can be contacted more than once to ask whether he or she consents to a certain usage of his or her information.

2.2.3. China

Chinese public and private institutions draw massive amounts of data from a wealth of sources to profile individuals with the highest degree of accuracy. Collecting and processing personal data about the Chinese population is instrumental to China's grand civic plan, which foresees the implementation of a wide-ranging surveillance and monitoring scheme that exploits AI to profile and predict individuals' and groups' behaviours.[115] The overall goal of this plan consists in the construction of a pervasive social credit system – an AI-based mechanism that gathers information from personal records, smartphones, and mass-surveillance systems, and then ranks individuals and accords them privileges and rights based on their previous conduct.[116]

In China, public institutions are trying to make everyone's life transparent, and not private. To this end , they partner with Chinese private firms. A handful of big tech companies such as WeChat and Alibaba thus operate as digital hubs for the lives of Chinese citizens.[117] The Chinese are encouraged to use the same mobile app for a wide array of activities – from reserving a taxi to paying for a restaurant, socialising or interacting with a public administration. A huge amount of information about anyone is thus gathered and passed over to public institutions for profiling.[118]

2.2.4. Three different approaches?

Odd as it may seem, some have speculated that a similar social credit system is already in place also in the private sector of the United States.[119] Private companies don't merely profile their clients to make them loyal. They also sell the information about them to other companies. Personal preferences and purchase habits are thus matched to better profile users, anticipate their decisions, and nudge them.[120] A bank or an insurance

---

[115] State Council, Notice of the State Council Issuing the New Generation of Artificial Intelligence Development Plan, No. 358 July 2017, pp. 2-5, and 18-21, https://flia.org/notice-state-council-issuing-new-generation-artificial-intelligence-development-plan.

[116] State Council, Notice concerning Issuance of the Planning Outline for the Construction of a Social Credit System (2014-2020), No. 21, 14 June 2014, https://chinacopyrightandmedia.wordpress.com/2014/06/14/planning-outline-for-the-construction-of-a-social-credit-system-2014-2020.

[117] Pieranni S., *Red Mirror*, Laterza, 2020, pp. 22-23.

[118] Ibid, pp. 40 and 115.

[119] Baker L. C., "Next generation law: Data-driven governance and accountability-based regulatory systems in the West, and social credit regimes in China", *Southern California Interdisciplinary Law Journal*, 28, 2018, pp. 170-171, https://lcbackerblog.blogspot.com/2019/05/just-published-next-generation-law-data.html.

[120] The European Parliament has recently called on the European Commission to "ban platforms from displaying micro-targeted advertisements": European Parliament, Resolution of 18 June 2020 on competition policy – annual report 2019, https://www.europarl.europa.eu/doceo/document/TA-9-2020-0158_EN.html. According to Morozov E., "Digital socialism?", *New Left Review*, 116/117, March-June 2019, p. 62, https://newleftreview.org/issues/II116/articles/evgeny-morozov-digital-socialism, "Amazon got a patent on 'anticipatory shipping' – allowing it to ship products to us before we even know we want them".

company can accurately assess an individual's financial risk based on a variety of information, ranging from his/her education, his/her lifestyle, or the places and people he/she visits. A political party can assess the political inclination of an individual based on the movies he/she watches, the media channels he/she prefers, or his/her family records.

It should be of little or no surprise that the overall US approach to data protection overlooks the negative potential of such a private accumulation of personal data. The US culture of rights has traditionally focused on keeping public powers under check. This approach is still lively, and keeps the US attention focused on the threats of public powers, whereas Europe has always been more attentive to private companies' capacity to violate fundamental rights.[121] The paradoxical result is that the US is the global hub for big data innovation, but does not see the big data threat to fundamental rights the way Europe appears to do.

Such different approaches to privacy have powerful consequences for the ordinary lives of citizens and media companies alike. As will become apparent below, the exploitation of AI-based technologies transforms media corporations into more than information givers. They can become information gatherers and participate in profiling individuals.

2.3. Big data bias and discrimination

Although one would not expect software to be biased, one of the biggest challenges for data-driven technologies is their discriminatory potential. The gathering, processing, and dissemination of information can incorporate, embed and amplify prejudices. The most famous example probably is the Microsoft chatbot Tay. In 2016, Microsoft created a Facebook profile for innovative software capable of interacting on the media platform with other Facebook users by gathering information from the web, identifying trends, and exchanging opinions accordingly.

In the span of 16 hours, the Facebook account was opened and then shut down, after its creators realised it was engaging in sexist and racist posts.[122] The software developers certainly did not provide their bot with the set of prejudices it later displayed on the web. Its makers simply used the web itself to teach the bot, which evidently found racism and sexism to be widespread and attention-drawing. Tay shaped its language and

[121] As to the European attentiveness to private companies' harmful potential, see European Data Protection Officer, Opinion 8/2016 EDPS Opinion on coherent enforcement of fundamental rights in the age of big data, 23 September 2016, p. 5, https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf. *See* also Pollicino O., "L'"autunno caldo' della Corte di giustizia in tema di tutela dei diritti fondamentali in rete e le sfide del costituzionalismo alle prese con i nuovi poteri privati in ambito digitale", *Federalismi*, 15 October 2019, https://www.federalismi.it/nv14/editoriale.cfm?eid=533.

[122] "Microsoft 'deeply sorry' for racist and sexist tweets by AI chatbot", *The Guardian*, 26 March 2016, https://www.theguardian.com/technology/2016/mar/26/microsoft-deeply-sorry-for-offensive-tweets-by-ai-chatbot.

themes based on the training it was subject to. It learned and adopted prejudices on its own.

Tay's ephemeral life explains the importance of training for AI. AI-based systems require a lot of data in order to learn. The more information they gather, the more capable they become of making inferences and choices. Unfortunately, big datasets to train algorithms are often unavailable, so software programmers often exploit what is already available on the web. This choice is extremely problematic, because human beings cannot fully supervise the learning process, and AI can take unforeseen or even unwelcome directions. It can draw and incorporate biases from society, boosting them with its activity.[123]

Unbalanced datasets can unintentionally create biases, as the case of facial recognition exemplifies. Western AI systems of face recognition often fail to correctly identify non-Caucasian individuals because other ethnic groups appear on the web less often than Caucasians, while AI software developed in China suffers from the reverse problem.[124] As a result, there is a higher probability that, say, in Western countries an African individual is mistaken for someone else than a Caucasian is. Media systems that incorporate big data-based processes therefore face a formidable challenge, as by exploiting AI they may incorporate prejudices and social imbalances.

Fighting discrimination is very difficult in the field of big data and neural networks because of the dangers of "proxy discrimination".[125] Proxy discrimination is a private or public policy that includes a requisite or factor that is facially neutral but actually embeds a discriminatory tradition, practice, or belief. For example, in socially or territorially divided societies, the zip code or the housing price can serve as a proxy discrimination for insurance policies or zoning, as it may deprioritise some ethnicities while preferring others. Even if software developers expressly prohibit AI from considering ethnicity while making inferences, other factors can serve as proxies for discrimination.[126] Within a given society, big data-driven market strategies, political campaigns, or welfare providers can – even involuntarily – isolate and systematically discriminate worse-off groups by proxy.

---

[123] Stevenson M. T. & Doleac J. L., *Algorithmic Risk Assessment in the Hands of Humans*, Institute of Labor Economics, 1 December 2019, p. 1, http://ftp.iza.org/dp12853.pdf; Bruckner M. A., *op. cit*, p. 25.

[124] Grother P., Ngan M., Hanaoka K., "Face recognition vendor test (FRVT) Part III. Demographic effects", National Institute of Standards and Technology Interagency 8280, December 2019, https://doi.org/10.6028/NIST.IR.8280.

[125] Prince A. E. R. & Schwarcz D., "Proxy discrimination in the age of artificial intelligence and big data" *Iowa Law Review* 105, 2020, p. 1260, https://ilr.law.uiowa.edu/print/volume-105-issue-3/proxy-discrimination-in-the-age-of-artificial-intelligence-and-big-data.

[126] Idem.

2.4. Informing the people: Media, misinformation, and illegal content

AI is a powerful media tool. It can discover facts, detect preferences, profile users and anticipate social trends. In a few words, it can provide people with more of what they want to receive. Customising media offerings through big data has a price, though.

AI is a very good tool for the pre-selection of content that media users may find of interest. Given the overflow of information, AI's capacity to profile a user can predict his/her interests in a piece of information, making the media's work more effective and the user's experience more enjoyable. However, AI exploitation may make media users unaware of the fact that their horizons are narrowing – that the type of information they receive may not portray reality accurately, but only the "reality" of what AI understands their interests to be.

Feeding users with more of what they already prefer, know, or are interested in, tends to create social bubbles. Big data technologies can filter information depending on what a media user supposedly likes or believes. Instead of widening the horizon of users, AI is thus able to boost individuals' intellectual selectiveness. A user-friendly news industry may lose sight of its purpose of providing society with broad perspectives, fully informed news and challenging viewpoints.

Big data-driven media strategies can thus unwillingly trigger the creation of informational bubbles. There is the additional risk, however, that a bubble is generated intentionally. Big tech companies can profile users and information to boost or hinder the spread of certain information depending on their market strategies or agendas.[127]

Big data also pits traditional media against social media. Social media exploit the strong protection normally accorded to freedom of speech, and live off their continuous presence on the web and their capacity to feed the audience with more news.[128] They therefore offer a cheap and easily accessible alternative to professional media operators and outlets. Such asymmetric competition has triggered a dangerous "race to the bottom" in the field of news providers.[129] In order to avoid losing the audience, traditional media try to keep up with the speed of non-professional services such as blogs, often at the expense of accuracy.[130]

AI-based media platforms' bubbles often participate in spreading "fake news". A plague in today's news industry, according to some statistics "fake news" is capable of

---

[127] Singer H., "How Washington should regulate Facebook", *Forbes*, 18 October 2017, https://www.forbes.com/sites/washingtonbytes/2017/10/18/what-to-do-about-facebook.

[128] Shefa M. C., "First Amendment 2.0: Revisiting Marsh and the quasi-public forum in the age of social media", *University of Hawaii Law Review*, 41, 2018, p. 160.

[129] AGCM, AGCOM, and Garante per la protezione dei dati personali, Indagine conoscitiva sui Big Data, *op. cit.*, p. 30.

[130] European Data Protection Supervisor, Opinion 3/2018 EDPS Opinion on online manipulation and personal data, *op. cit.*, p. 13 ("There is evidence that … concentration and elimination of local journalism facilitates the spread of disinformation").

reaching more people and more quickly than curated, fact-checked information,[131] giving life to what Cass Sunstein has called "cybercascades".[132] The bubble system aggravates the process, as it filters out facts and different viewpoints, thereby reinforcing deeply held viewpoints and even prejudices.

Big data-driven strategies are calling into question the historical role that the media system and freedom of speech have played in democratic regimes. Instead of broadening horizons, challenging viewpoints, exposing biases and making society progress, contemporary media platforms run the risk of mutually insulating social groups and reinforcing deeply held opinions. Traditionally, liberal constitutionalism values and protects freedom of speech greatly because different viewpoints make societies progress through the free exchange of opinions. Contrarily, big data technologies are capable of creating "echo chambers",[133] which expel dissent and gravitate around unchallenged beliefs. Opinions that challenge deeply seated worldviews are ejected from a bubble and will probably find their place within another bubble, which offers virtually no exchange outside itself. [134] Big data can thus narrow perspectives and immunise prejudices from the benefits of freedom of speech.

Private and public institutions have grown aware of the distortions that big data can cause to media and broader society. For example, Twitter recently created a contentious fact-checker tool with the purpose of detecting "fake news" or tweets that harm identifiable groups.[135] The EU's *Code of Practice on Disinformation*[136] has urged a comprehensive consideration of the phenomenon, emphasising that "all stakeholders have roles to play in countering the spread of disinformation". A list of signatories to the code that includes Facebook, Google, Mozilla, TikTok and Twitter has thus promised to "[d]ilute the visibility of disinformation by improving the findability of trustworthy content", and to "facilitate content discovery and access to different news sources representing alternative viewpoints". Overall, many are calling for regulation of the deployment of AI in a way that would bring Internet service providers closer to the "traditional media responsibility standards".[137]

EU policies especially target terrorist content, child sexual abuse material, racism, and xenophobic and hate speech,[138] which are usually topics of great concern for today's

---

[131] Idem.

[132] Sunstein C. R., "*#republic: Divided democracy in the age of Social Media*", Princeton University Press, 2017, p. 57.

[133] Sasahara K. et al., "*On the inevitability of online echo chambers*", https://arxiv.org/abs/1905.03919.

[134] Jones R. L., "Can you have too much of a good thing: The modern marketplace of ideas", *Missouri Law Review*, 83, 2018, p. 987, https://scholarship.law.missouri.edu/mlr/vol83/iss4/8/.

[135] Pham S., "Twitter says it labels tweets to provide 'context, not fact-checking'", *CNN Business*, https://edition.cnn.com/2020/06/03/tech/twitter-enforcement-policy/index.html.

[136] EU Code of Practice on Disinformation, https://ec.europa.eu/digital-single-market/en/news/code-practice-disinformation.

[137] European Data Protection Supervisor, Opinion 3/2018 EDPS Opinion on online manipulation and personal data, *op. cit.*, p. 16.

[138] Policy Department for Economic, Scientific and Quality of Life Policies, "Online platforms' moderation of illegal content online", June 2020, p. 9, https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652718/IPOL_STU(2020)652718_EN.pdf.

social media. In fact, given the massive inflow of data, filtering information before deciding whether to host it is technically unrealistic. Online platforms thus normally blend two different schemes: on the one hand, they adopt a "notice-and-takedown" system - anyone can complain that a specific display of content is in breach of the law and have the medial platform make an assessment; on the other hand, most platforms adopt big data-based filtering systems that sift through the materials automatically and pervasively, making decisions on what should be concealed from the public.[139] Most platforms have an additional safeguard against such automated decisions, allowing individuals to challenge a software decision to remove some material.[140]

Within the US and the EU, which has "one of the most comprehensive regulatory frameworks for tracking illegal content online",[141] service providers enjoy broad liability exemptions. Such exemptions aim to preserve their positive role in connecting people and disseminating information.[142] EU law has reinforced this rule by prohibiting its member states from imposing general obligations on hosting platforms to monitor the material they host.[143] The scenario is in flux, however.[144] In interpreting the Directive on electronic commerce, the Court of Justice of the European Union has stated that service providers that do not simply passively display materials are expected to do more than simply review and remove materials when necessary once they are requested to do so.[145] In fact, the court stated, a judicial order of removal extends "to information, the content of which, whilst essentially conveying the same message [to which the judicial order refers], is worded slightly differently, because of the words used or their combination,

[139] Ibid, p. 45 .
[140] Ibid, p. 10.
[141] Ibid, p. 66.
[142] For the United States, see Title 47, Section 230 of the Communication Decency Act, https://www.fcc.gov/general/telecommunications-act-1996; For the EU, see Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce'), https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0031&from=EN, Art. 14: "1. Where an information society service is provided that consists of the storage of information provided by a recipient of the service, Member States shall ensure that the service provider is not liable for the information stored at the request of a recipient of the service, on condition that: (a) the provider does not have actual knowledge of illegal activity or information and, as regards claims for damages, is not aware of facts or circumstances from which the illegal activity or information is apparent; or (b) the provider, upon obtaining such knowledge or awareness, acts expeditiously to remove or to disable access to the information." As for the protection of minors, see Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive; codified version; text with EEA relevance). A consolidated version including the amendments introduced in 2018 is available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02010L0013-20181218.
[143] Policy Department for Economic, Scientific and Quality of Life Policies, *op. cit.*, p. 21.
[144] Nunziato D. C., "The marketplace of ideas online", *Notre Dame Law Review*, 94, 2019, p. 1521, https://scholarship.law.nd.edu/cgi/viewcontent.cgi?article=4844&context=ndlr.
[145] C-324/09, *L'Oréal et al. v. eBay International AG*, paras. 113-115, http://curia.europa.eu/juris/document/document.jsf?text=&docid=107261&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=12642628.

compared with the information whose content was declared to be illegal".[146] Some have criticised this sensible principle because it would result in the "Good Samaritan paradox": the more a platform is committed to patrolling the information it publishes, the more it becomes liable. There are concerns that such a judicial approach would encourage providers to remain passive and limit their monitoring activity in order to avoid liability risks.[147] It is now a matter of debate whether the EU should revise its policy and imitate the US approach, which has preserved the liability exemption for platforms, as this would encourage them to become more proactive, or whether this would jeopardise the protection of individuals and groups.[148]

In the context of illegal materials posted on online platforms, AI can certainly play an important role. Given the huge amount of data exchanged and the tendency to create bubbles within which media users hardly find information they do not like or viewpoints they disagree with, illegal materials may not be detected by human beings for a long time. Developing AI-based systems that filter content may therefore become advisable or even necessary. AI and big data are not just part of the problem – they can be part of the solution. Obviously, AI-based monitoring should not become a form of automated censorship. Providers may exploit AI systems to filter out materials that are simply controversial, thereby insulating the public sphere from minoritarian opinions or information that many would find hard to engage with . This risk should be kept in check.

2.5. Big data politics and the political bubble[149]

Democracies need a sound public sphere to survive and flourish.[150] The existence and exchange of alternative worldviews and political opinions is crucial for their survival. More generally, within democracies "people should be exposed to materials that they would not have chosen in advance",[151] as one of the benefits historically associated with democracies is that "biases are filtered out in the large republic".[152]

Social media have flooded contemporary politics. Legal academia and courts have responded by slowly but steadily developing the classical idea of public forums to

[146] C-18/18, *Eva Glawischnig-Piesczek v. Facebook Ireland Ltd*., par. 41, http://curia.europa.eu/juris/document/document.jsf?text=&docid=218621&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=12642666.
[147] Policy Department for Economic, Scientific and Quality of Life Policies, *op. cit*., p. 20; Policy Department Economic and Scientific Policy, "Liability of Online Service Providers for Copyrighted Content – Regulatory Action Needed?", January 2018, p. 10, https://www.europarl.europa.eu/RegData/etudes/IDAN/2017/614207/IPOL_IDA(2017)614207_EN.pdf.
[148] Ibid, p. 67.
[149] For a different viewpoint on the filter-bubble/echo chamber issue see chapter 5 of this publication.
[150] Wischmeyer T., "Making social media an instrument of democracy", *European Law Journal*, 25, 2019, p. 172, https://onlinelibrary.wiley.com/doi/abs/10.1111/eulj.12312.
[151] Sunstein C. R., *op. cit.* p. 6.
[152] McGinnis J. O., *Accelerating Democracy*, Princeton University Press, 2013, p. 127.

incorporate also social media sites that are privately owned.[153] Because of their pervasive social role and their pivotal importance in providing the public with news feeds and political opinions, the US Supreme Court has dubbed social media sites as "the modern public square".[154] They are so essential to social and political life – the court has argued – that they must be accessible to the general public.[155] Since 2001, US courts have also "treated computers and Internet access as 'virtually indispensable in the modern world of communications and information gathering'."[156]

Social media are not universally accessible places within which everybody is welcomed and able to make an argument, however. Big data analysis allows social media to segment the public sphere in self-referential bubbles.[157] Even the media platforms that do not intentionally filter information, still tailor their news feeds to their users' needs and choices, therefore creating informational bubbles. Such bubbles are capable of dividing public opinion into impenetrable, homogenous spheres of influence.[158]

The creation of homogenous, partisan, non-conversational echo chambers is no substitute for democratic pluralism[159] and can even threaten it.[160] The scandal of Cambridge Analytica, which allegedly harvested data of Facebook users without their consent to develop "psychographic profiles" and then target selected individuals to nudge their voting behaviours,[161] is just one example of how big data can affect politics.[162] And there is wider evidence of the deployment of big data-fed bots to influence political agendas.[163]

Harvard Law Professor Cass Sunstein has explored the impact of AI-based social media platforms in the political sphere in his acclaimed volume *#Republic*.[164] Sunstein has persuasively shown AI's capacity to generate informational clusters and polarise politics. Political campaigns can target well-profiled users, exposing them to certain opinions or facts while silencing or downplaying the statements of political opponents or facts that

---

[153] Nunziato D. C., *op. cit.*, p. 3.
[154] *Packingam v. North Carolina* 582 U.S. ___ (2017), https://www.supremecourt.gov/opinions/16pdf/15-1194_08l1.pdf.
[155] Ibid.
[156] Shefa M. C., *op. cit.*, p. 164.
[157] Sunstein C. R., *op. cit.*
[158] Sasahara K. et al., *op. cit.*
[159] Wischmeyer T., *op. cit.*, p. 173-174.
[160] Manheim K. & Kaplan L., *op. cit.* , p. 109.
[161] Ibid, p. 139.
[162] For more examples drawn from various countries, see Gurumurthy A. and Bharthur D., "Democracy and the algorithmic turn", *Sur International Journal of Human Rights*, 27, 2018, pp. 43-44, https://sur.conectas.org/en/democracy-and-the-algorithmic-turn, and Tenove C., Buffie J., McKay S. and Moscrop D., *Digital threats to democratic elections: how foreign actors use digital techniques to undermine democracy*, January 2018, passim, https://democracy2017.sites.olt.ubc.ca/files/2018/01/DigitalThreats_Report-FINAL.pdf.
[163] When the Federal Communication Commission considered repealing some rules regulating the Internet in 2017, 21 out of 22 million commentsthe Commission received on its website were fake news (Manheim K. & Kaplan L., *op. cit.*, p. 145.)
[164] Sunstein C. R., op. cit.

would call into question their own platform and agenda.[165] AI thus splinters the public sphere into homogenous environments which hardly interact together. Successful politicians often go to extremes to galvanise their supporters and reinforce the bubble system.

Big data politics often blurs the line between personal and institutional capacity. Many political figures prefer using their personal social media profiles rather than institutional profiles also to communicate with the general public on institutional matters. By using their personal profiles, they force the public – which would normally follow institutional media pages and profiles - into their sphere of supporters.

Some legal systems have deployed countermeasures to fight this privatisation of the public sphere into separate media echo chambers. The US experience provides the most telling example of this development. Many public figures – including President Donald Trump – who have used personal websites for institutional purposes have blocked individuals making critical comments about their posts, therefore walling them out from their briefing activity to citizens.[166] Some citizens thus ejected from the audience sued the politicians - and won in court. Judges considered the structure of media platforms and how politicians were using them, and concluded that such platforms had to be considered public places that should remain open to everyone. Politicians could still "mute" their followers, thereby preventing them from engaging in a conversation within their own profile, but not "block" them, as this would have prevented some citizens from being informed on matters of public interest.[167]

2.6. Media as surveillance watchdogs?

Big data analysis has been instrumental to the development of artificial face recognition techniques. Thanks to AI capabilities, software can peruse and compare an enormous amount of images, to find matches. Differently from old-fashioned close-circuit cameras, which human agents scrutinise looking for matches, today's computer vision has the capacity to process images almost instantly. In a 2019 decision, a Welsh court dealt with artificial face recognition.[168] The software that the Welsh police had deployed at several public events was able to process up to 40 faces per second. The total figure is impressive: in roughly 50 deployments, the software processed roughly 500 000 individuals – one out of six of the total population of Wales. AI can become a powerful tool of mass surveillance, as has already happened in countries such as China, where a

[165] Mor N., "No Longer Private: On Human Rights and the Public Facet of Social Network Sites", *Hofstra Law Review* 47 (2018), p. 669, https://www.hofstralawreview.org/wp-content/uploads/2019/04/bb.7.mor_.pdf (6 August 2020).
[166] *Ibidem*, p. 42 ff.
[167] *Knight First Amendment Inst. at Columbia Univ. v. Trump* 302 F. Supp. 3d 541 (SDNY 2018), https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=2780&context=historical (6 August 2020).
[168] *(Bridges) v. The Chief Constable of South Wales Police et al.*, [2019] EWHC 2341, https://www.judiciary.uk/wp-content/uploads/2019/09/bridges-swp-judgment-Final03-09-19-1.pdf.

project of a systematic AI-based surveillance system, with more than half a billion of cameras deployed, is ongoing.[169]

Face recognition cuts across a variety of issues seen above. First, face recognition techniques are a matter of privacy. They process human faces – not just of those in a database, but of everyone. In fact, in order to exclude someone from the group of persons of interest, a software must process their face first. According to the European legal culture, such a massive privacy intrusion must be properly justified. As the European Court of Human Rights has repeatedly insisted, public interests do not override privacy concerns – on the contrary, they require a preliminary assessment of the expected benefits and costs to ensure that any deployment is proportionate to the task.[170]

Second, face recognition techniques runs the risk of being biased. As noted above, "false positives" – wrong matches – are more frequent in ethnic groups that are underrepresented in the training materials.[171] False positives often have practical consequences, as they may reinforce racial prejudices and nudge public institutions, such as police patrols, to target ethnic minorities for which software returns more false positives.[172]

Third, face recognition can be misleading on a variety of grounds. Some software programmes are able to exploit the immense AI capabilities by using live and recorded images coming from any Internet source.[173] Such technology can exploit the media industry to gather more materials and increase its database. A debate is ongoing on the pros and cons of developing or adopting software that sifts through the web to find matches of people, as has happened in many local police agencies of the U.S. to track down suspects. Such a huge dataset draws on a variety of materials that can be spurious, incorporate bias,[174] and transform any single bit of social life or media broadcast into a record.

---

[169] Carter W. M., "Big Brother facial recognition needs ethical regulations", *The Conversation*, 22 July 2018, https://theconversation.com/big-brother-facial-recognition-needs-ethical-regulations-99983.

[170] *Lopez Ribalda and others v. Spain* (apps. No. 1874/13 and 8567/13: http://hudoc.echr.coe.int/fre?i=001-197098); *Gorlov and others v. Russia* (app. no. 27057/06; 56443/09; 25147/14: http://hudoc.echr.coe.int/spa?i=001-194247); *Antovic and Mirkovic v. Montenegro* (app. no. 70838/13: http://hudoc.echr.coe.int/fre?i=001-178904); *Bărbulescu v Romania* (app. no. 61496/08: http://hudoc.echr.coe.int/spa?i=001-177082).

[171] Buolamwini J. & Gebru T., "Gender shades: Intersectional accuracy disparities in commercial gender classification" *Proceedings of Machine Learning Research* 81, 2018, pp. 1 and 15, http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

[172] Fung B. and Metz R., "This may be America's first known wrongful arrest involving facial recognition", 24 June 2020, *CNN Business*, https://edition.cnn.com/2020/06/24/tech/aclu-mistaken-facial-recognition/index.html.

[173] Hill K., "The secretive company that might end privacy as we know it", *New York Times*, 18 January 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html; Ducklin P., "Clearview AI facial recognition sued again – this time by ACLU", *Naked Security*, 29 May 2020, https://nakedsecurity.sophos.com/2020/05/29/clearview-ai-facial-recogition-sued-again-this-time-by-aclu.

[174] Geiger R. S. et al., "Garbage in, garbage out? Do machine learning application papers in social computing report where human-labeled training data comes from?", https://arxiv.org/abs/1912.08320.

It is no surprise that IBM,[175] Microsoft[176] and Amazon[177] have recently issued statements that they will not offer their face recognition technologies to the police anymore. Many US states are considering banning artificial face recognition or have already implemented legislation that limits or prohibits it.[178] There is therefore a growing consensus in Western countries that even public interests cannot justify pervasive mass surveillance systems that exploit the web.

2.7. The media market: Big data-driven market strategies

Big data has revolutionised the universe of media. Many players in the media industry now depend on big tech companies to better connect with their audiences.[179] In fact, gathering and processing huge amounts of data in a fruitful way requires capabilities that few own. The pool of companies that can harvest big data is very limited, and the majority of market players rely on this pool to better understand who their clients are, what type of market strategy they should implement or how to gain more visibility. Some big tech companies in the field, such as Amazon, even produce media content themselves. Thanks to their technological capabilities, big tech companies thus now operate either (or both) as media makers and as mediators between the media industry and its consumers.

The Court of Justice of the European Union's landmark Google Spain case[180] encapsulates the paramount role that big tech companies now play in the news field and their resistance to the laws governing it. When an individual complained that a Google search of his name returned a list of results at the top of which was a very old newspaper item about him that could still ruin his reputation, Google's first line of defence was that it did not handle personal data; it only connected searches with results.[181] In other words, Google made the argument that it was not responsible for what it made available through Google search. The court responded with a historical judgement, showing its awareness of the unique role of Google in Internet searches. It found that Google was responsible for how it ranked its answers to a query, as it could resurrect long forgotten pieces of information that would not have been accessible to the general public otherwise.

---

[175] Krishna A., "IBM CEO's Letter to Congress on Racial Justice Reform", 8 June 2020, https://www.ibm.com/blogs/policy/facial-recognition-susset-racial-justice-reforms/.

[176] Greene J. Microsoft won't sell police its technology, following similar moves by Amazon and IBM", *The Washington Post*, 11 June 2020, https://www.washingtonpost.com/technology/2020/06/11/microsoft-facial-recognition/.

[177] Hao K., "The two-year fight to stop Amazon from selling face recognition to the police", *MIT Technology Review*, 12 June 2020, https://www.technologyreview.com/2020/06/12/1003482/amazon-stopped-selling-police-face-recognition-fight. See also Hartzog W., *op. cit.*, p. 76-77.

[178] See the Illinois Biometric Information Privacy Act, https://www.termsfeed.com/blog/bipa/.

[179] Tsesis T., *op. cit.*, p. 1589.

[180] *Google Spain SL et al. v. Agencia Española de Protección de Datos*, C-131/12, https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62012CJ0131&from=EN.

[181] Ibid., para. 22.

Big tech companies do not simply populate the media market. They deeply affect its dynamics, too. Their unique ability to profile the market entraps their users in a "lock in" phenomenon and generates a quasi-market monopoly.[182] They are so pervasive and indispensable that those who do not want to use them often have to leave the market altogether. Many Internet users know that "visiting a single website results typically in the disclosure of browsing behaviour to over 100 third parties who seek to limit their own legal liability by means of dense 'privacy policies' which can run to hundreds of pages", but they cannot avoid visiting the same websites time and again.[183] The few companies that exploit the potentials of big data may patrol their territories even further by engaging in "killer acquisitions", through which they purchase innovative start-ups to either mine the data they have collected[184] or protect their dominant position.[185] In Frank Pasquale's words, like "Pharaoh trying to kill off the baby Moses", big tech companies can deny their rivals "the chance to scale".[186]

The simultaneous presence of more than one company that uses big data does not ensure that a market is competitive.[187] Big data can help the development of market strategies, including pricing, that benefit the competitors, not the customers. There is evidence that algorithms of different companies can maximise pricing through an implicit collusive strategy, simply by processing information about the market itself.[188] An algorithm can suggest a company raise prices because it predicts that its competitors will decide to do the same. Thanks to user profiling and clustering, they can also "segment … the market" and charge each user according to their willingness to pay. These practices create the "maximum revenue [for firms] but no consumer welfare".[189] Such a data-driven market strategy is usually not punishable, as there is no collusion, but has the benefits that normally attach to collusive behaviours.[190]

---

[182] AGCM, AGCOM, and Garante per la protezione dei dati personali, Indagine conoscitiva sui Big Data, *op. cit*., p. 26 and 78.

[183] European Data Protection Supervisor, Opinion 3/2018 EDPS Opinion on online manipulation and personal data, *op. cit*., p. 7.

[184] Zuboff S., *The Age of Surveillance Capitalism*, Profile Books, 2019, pp. 102-103.

[185] AGCM, AGCOM, and Garante per la protezione dei dati personali, "Indagine conoscitiva sui Big Data", *op. cit*., p. 81, See also Hughes C., *op. cit*.

[186] Pasquale F., *The Black Box Society*, Harvard University Press, 2015, p. 67.

[187] European Data Protection Supervisor, Opinion 3/2020 on the European strategy for data, op. cit., p. 8 (where it is warned against the creation or reinforcement of "situations of data oligopoly").

[188] Den Boer A. V., "Dynamic pricing and learning: Historical origins, current research, and new directions", *Surveys in operations research and management science*, 20, 2015, p. 1, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2334429); AGCM, AGCOM, and Garante per la protezione dei dati personali, "Indagine conoscitiva sui Big Data", op. cit.

[189] European Data Protection Officer, Opinion 8/2016 EDPS Opinion on coherent enforcement of fundamental rights in the age of big data, op. cit., p. 6.

[190] Harrington, J. E. Jr., "Developing competition law for collusion by autonomous artificial agents", *Journal of Competition Law & Economics*, 14, 2019, pp. 349-351, https://academic.oup.com/jcle/article-abstract/14/3/331/5292366?redirectedFrom=fulltext.

2.8. Regulatory approaches to AI-based systems

Many have voiced the need for new regulatory schemes in order to ensure that AI is utilised in a way that respects the rule of law, fundamental rights and ethical values. Big tech companies have long resisted public efforts to regulate the field,[191] but now appear to have come to terms with the necessity of constraining AI, although they push for company self-regulation rather than state rules.

Most constraints, however, do not aim to depress the utilisation of AI; in fact, they are expected to boost its role by making it more trustworthy and reliable.[192] There is wide consensus, in fact, that AI needs to be "lawful" (law-compliant), "ethical" (committed to respecting ethical principles and values) and "robust" (technologically and sociologically safe), in order to successfully integrate with human societies.[193]

Debates often emphasise that big data analyses need a new approach to legal regulation. Traditional tools may not be sufficient to ensure that the world of big data respects basic human values. Because of AI's black box structure and large-scale effects, legal sanctions are hardly capable of constraining big data-based technologies and strategies. Lawsuits may arrive late, when one's reputation or a company is in ruins, and liabilities may be hard to locate. AI needs to incorporate legal values within its data processing, in order to make sure that it protects them while it is operating.

Because of the wealth of information it gathers, its pervasive deployment and its capacity to replace human operators with robots, AI also poses ethical questions. *Digital ethics* is a new frontier for AI regulation and has drawn considerable attention especially in the US, in Canada and in Europe, where ethical codes have mushroomed.[194] As a field, digital ethics covers a wealth of topics, including "moral problems relating to *data* and *information* ..., *algorithms* ... and corresponding *practices* and *infrastructures*",[195] in a way that cuts across different disciplines and perspectives. Albeit extremely lively, the situation is magmatic at the moment, also because of the difficulties in drawing lines between the legal and the ethical components of AI regulation.[196]

---

[191] Zuboff S., op. cit., p. 105.

[192] Van Dijk N. & Casiraghi S., "The ethification of privacy and data protection law in the European Union: The case of artificial intelligence", *Brussels Privacy Hub*, 6, 22, May 2020, p. 5, https://brusselsprivacyhub.eu/publications/BPH-Working-Paper-VOL6-N22.pdf.

[193] High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, p. 2, https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai. See N. van Dijk & S. Casiraghi, *op. cit.*, p. 14.

[194] Jobin A., Ienca M. and Vayena E., "The global landscape of AI ethics guidelines", *Nature Machine Intelligence*, 1, 2019, pp. 393-395, https://www.nature.com/articles/s42256-019-0088-2.

[195] Floridi L., *op. cit.*, p. 3.

[196] For example, see the Council of Europe's *Recommendation CM/Rec(2020)1 of the Committee of Ministers to member States on the human rights impacts of algorithmic systems*, 8 April 2020, https://search.coe.int/cm/pages/result_details.aspx?objectid=09000016809e1154, which showcases the variety of regulatory layers necessary for the development of sound AI-based systems.

2.9. Conclusion

Big data is a big reason for the societal, economic, and political success of AI. Processing vast amounts of data is crucial for big tech companies. It has not been just a blessing, however, and it requires people working in the field to take action to ensure that AI is beneficial to human beings.[197] Chris Hughes, co-founder of Facebook, has warned that the digitalisation of the economy may contribute to what he perceives to be "a decline in entrepreneurship, stalled productivity growth, and higher prices and fewer choices for consumers".[198] The stakes are so high that a member of the National Assembly, the lower house of the French Parliament, has even submitted a proposal to entrench a *Charter of artificial intelligence and of algorithms* within the preamble of the French constitution, to better protect human rights.[199]

AI maximises people engagement. Eliciting "as much response as possible from as many people as possible" is a key factor of success, as it provides feedback and allows companies to adjust their business plans and models to their customers in real time.[200] Political players and social influencers exploit this phenomenon by triggering emotional responses from their potential audience. Big data politics and economy place media at the centre stage, as they spread news, gather information, process emotions, and connect social spheres.

Big data aggrandises the role of the media for contemporary societies. Companies, politicians, influencers and other political figures exploit big data to market their ideas, agendas and opinions, as well as to shape their audiences.[201] Internet platforms allow legacy media to spread their content and generate new competition between traditional and new outlets.

Media players can also play a negative role. Through profiling the "thinking patterns and psychological makeup," they can deliberately misinform and mislead an audience.[202] Moreover, in countries where few media players operate, or where there are only or almost exclusively state-run social media,[203] a political regime can effectively control the news and also how people react to it, by disseminating fabricated favourable feedback and insulating unfavourable comments.[204] Within the scenario generated by big

---

[197] See the Asilomar Principles, developed in conjunction with the 2017 Asilomar conference. *Future of Life Institute*, https://futureoflife.org/ai-principles.

[198] See also Hughes C., *op. cit.*

[199] http://www.assemblee-nationale.fr/dyn/15/textes/l15b2585_proposition-loi.

[200] Akin Unver H., "Artificial intelligence, authoritarianism and the future of political systems", Centre for Economics and Foreign Policy Studies, July 2019, p. 3, https://edam.org.tr/wp-content/uploads/2018/07/AKIN-Artificial-Intelligence_Bosch-3.pdf.

[201] Idem.

[202] European Data Protection Supervisor, Opinion 4/2015. Towards a new digital ethics, *op. cit.*, p. 7.

[203] Pasquale F., *op. cit.*, p. 10, notes that "the distinction between state and market is fading" because of massive AI deployment in strategic sectors of public and private interest.

[204] Akin Unver H., *op. cit.*, p. 8. See also Meaker M., "How governments use the Internet to crush online dissent", *The Correspondent*, 27 November 2019, https://thecorrespondent.com/142/how-governments-use-the-internet-to-crush-online-dissent/18607103196-db0c0dab.

data, media can discharge a critical role in protecting democracy, equality, minority groups and open societies - or in undermining them.[205]

Finally, mass surveillance can have a chilling effect on creativity and innovation. Despite earlier expectations that AI would simply boost inventiveness,[206] some have detected "a tendency to discourage or penalise spontaneity, experimentation or deviation from the statistical 'norm', and to reward conformist behaviour".[207]

The vast deployment of AI nowadays requires that the media sphere become aware of its unique role. The media sector should strive to use AI in a lawful, ethical, and robust way. Thanks to their connecting role, the media could encourage the wider world of AI-based businesses to embrace the same values and become lawful, ethical, and robust. In particular, an ethical commitment may encourage media platforms to go beyond a merely passive role. While many regulations limit providers' legal liability for the content they host,[208] and more burdens imposed on media have not succeeded in encouraging more policing, it can still be a worthwhile ethical goal for media platforms to patrol their content.[209]

[205] High-Level Expert Group on Artificial Intelligence, "Ethics guidelines for trustworthy AI, *op. cit.*, p. 11.
[206] Perritt, H. H., Jr., *op. cit.*, p. 107.
[207] European Data Protection Supervisor, Opinion 4/2015. Towards a new digital ethics, *op. cit.*, p. 9. See also Pan S. B., *op. cit.*, p. 257 ("The goal of big data is to generalize") and Pasquale F., *op. cit.*, p. 188.
[208] Perritt H. H., Jr., *op. cit.*, p. 149.
[209] ERGA2020 Subgroup 1 – Enforcement, ERGA Position Paper on the Digital Services Act, p. 6, https://nellyo.files.wordpress.com/2020/06/erga_sg1_dsa_position-paper_adopted-1.pdf.