RESEARCH ARTICLE

OPEN ACCESS

PEER REVIEWED

Navigating the AI frontier: European parliamentary insights on bias and regulation, preceding the AI Act

Allessia Chiappetta *York University*

Abstract: Understanding Members of the European Parliament's (MEPs) attitudes and perceptions towards AI is crucial for aligning technological development with European values. This research paper focuses on the attitudes and perspectives of MEPs within the Special Committee on Artificial Intelligence in a Digital Age (AIDA) towards bias and discrimination in AI, as well as their views on regulatory measures. By conducting a critical discourse analysis of AIDA hearing transcripts, this study uncovers how MEPs perceive and comprehend bias and discrimination in AI and their stance on regulatory measures. The research argues that MEPs need to expand their understanding of AI due to their current limited comprehension. Findings reveal that MEPs view AI as both a risk and a source of innovation, with a prevailing sense of distrust. Some MEPs consider AI sentient and self-regulating, yet they all acknowledge the consequences of AI, including inherent biases and discriminatory practices – leading them to advocate for regulatory intervention. The insights gained from this study contribute to a deeper comprehension of the relationship between policymakers and emerging technologies, paving the way for informed decision-making and policy development within the European Union and beyond.

Introduction

The integration of artificial intelligence (AI) in European society has sparked deliberation in the EU Parliament. AI's rapid advancement holds transformative potential across various domains, including healthcare, transportation, education and law. Policymakers face intricate ethical, social and legal challenges to ensure fair, transparent and beneficial AI usage. While AI offers improved efficiency, accuracy and decision-making, it also harbours inherent biases originating from design, operation and training data (Lum & Isaac, 2016, p.19), with potential for self-generated biases (O'Neil, 2016). Consequently, it is vital for Members of the European Parliament (MEPs) – indeed policymakers from all countries – to possess a comprehensive understanding of AI's profound societal and economic impacts. This understanding will empower MEPs to make informed decisions on policies, regulations and funding, shaping the development and deployment of AI. By establishing a regulatory framework that fosters innovation, safeguards individual rights and maximises AI's benefits for European citizens and businesses, MEPs can navigate the path ahead.

The EU is grappling with a broad spectrum of legal concerns related to AI, spanning data privacy protection, potential biases and discrimination in AI outcomes, liability and accountability challenges for AI errors, transparency and explainability in AI decision-making processes, consumer protection issues tied to AI-powered products and services and the lack of clear, standardised AI-specific regulations. The EU has made efforts to address these challenges through existing legal frameworks such as the General Data Protection Regulation (GDPR), the Product Liability Directive and the Charter of Fundamental Rights. However, these frameworks have limitations, primarily due to their lack of AI-specific regulations, complexity, enforcement challenges and inadequate terminology. This recognition of the limitations of existing EU frameworks has led to the creation of the *AI Act*.

The EU *AI Act*, first proposed in April 2021 and subsequently updated in June 2023, represents a significant legislative effort to regulate AI within the EU. The core objective of this proposed legislation is to strengthen Europe's position as a global leader in AI innovation while ensuring that AI technologies adhere to European values and regulations (European Commission, 2021). Addressing legal and social concerns is pivotal to ensuring AI technologies align with legal standards, fostering innovation and safeguarding individuals and society from potential harm. The introduction of the *Act* seeks to provide a dedicated legal framework to navigate these intricate issues, guiding the ethical and responsible development and integration of AI technologies, thereby bridging existing legal gaps while addressing

emerging challenges (European Parliament, 2023).

Understanding MEPs' perspectives and attitudes toward AI is essential to grasp the depth and significance of the legislative framework behind the *AI Act*. Throughout the legislative process, MEPs played a pivotal role in shaping the final draft of the *Act*, representing a diverse range of interests and concerns from their constituents and stakeholders. These discussions span a spectrum of viewpoints, from those advocating robust AI regulations to protect fundamental rights and safety, to those emphasising the importance of balanced rules that promote innovation and competitiveness. In the context of this research, the examination of specific AIDA hearings, namely "AI and Bias" and "AI and the Data Strategy," meticulously chosen for this study, provides a nuanced glimpse into the diverse perspectives held by MEPs regarding the intricacies and potential biases inherent in AI technologies. These hearings serve as pivotal focal points for unravelling the evolving understanding of MEPs concerning AI's multifaceted challenges, underscoring the imperative for comprehensive regulations in the AI domain. The *AI Act* ultimately represents a compromise forged through these deliberations and negotiations, aiming to strike a balance between fostering AI innovation and ensuring alignment with European values and regulations. As such, understanding MEPs' deliberations and viewpoints is integral to comprehending the legislation's underlying principles and goals. This research provides valuable insights into how MEPs engaged in debates surrounding AI, enriching the EU's AI development and utilisation by addressing ethical, social and legal challenges and aligning AI deployment with the values and goals of European citizens.

Throughout this article, I argue for the need for MEPs to enhance their understanding of AI and its associated risks, biases and negative impacts through further research and call for the development of a stringent regulatory framework that promotes ethical and responsible AI development. Insights from MEPs' conceptualisations and understandings of AI, revealed during the hearings, highlight their recognition of potential biases and discrimination in its creation and operation. To regulate AI effectively, I support an approach that considers the experiences of marginalised communities and actively tackles various forms of discrimination in AI system design and deployment. Policymakers must strike a balance between leveraging private sector innovation and expertise, while ensuring public service accountability and prioritising the public interest[1]. Integrating private sector tech-

---

1. The term "European public interest" lacks a universally accepted definition (Hossfeld, Muller-Lagarde, Zevounou, 2020, p. 7). Its interpretation can vary and has evolved over time, encompassing economic, political and, more recently, sustainable development aspects (Hossfeld, Muller-Lagarde, Zevounou, 2020, p. 4). The flexibility in its interpretation reflects the dynamic nature of the concept

nologies into the public sector requires careful navigation to maximise benefits while maintaining transparency, fairness and outcomes aligned with the common good.

This research adopts a Critical Race Theory (CRT) and Social Construction of Technology (SCOT) meta-methodology to examine the implications of bias and the context of AI technologies used in the public sector. With AI's increasing presence in law enforcement, public services and the judiciary, it is vital to prevent the perpetuation of inequalities and injustices within marginalised communities. Racial bias is a significant focus, as AI often reinforces discrimination against racialised individuals. Employing the SCOT principles, this research highlights the sociotechnical aspects of AI development and the influence of policymakers. Interpretations of technological artefacts, including AI, are culturally constructed and shaped by persuasive discourse propagated by the technology industry and public/private actors (Yousefikhah, 2017, p. 36). Policymakers' understandings and perceptions of authority, legitimacy and accuracy regarding AI are crucial, as they influence the political context and discourse surrounding the technology (Brey, 2005, p. 68). This understanding is essential to critically assess the deployment and societal impact of AI, while advocating for equitable and accountable technological practices.

The article is organised as follows: Firstly, a comprehensive overview of the Special Committee on Artificial Intelligence in a Digital Age (AIDA) is outlined, including their mandate and role in the broader landscape of policymaking concerning AI. Next, a review of existing AI literature is presented, highlighting the challenges identified by scholars. Subsequently, the methodology employed to investigate AIDA hearings is described, followed by a discussion of my research findings. Through this study, I argue that MEPs must deepen their understanding of AI and its associated pitfalls, as their current comprehension is limited. The analysis reveals that MEPs perceive AI both as a potential risk and a source of innovation. Furthermore, the findings demonstrate MEPs' scepticism towards AI, however some perceive it as possessing sentient and self-regulating capabilities. MEPs also exhibit an awareness of AI's adverse effects, including inherent biases and discriminatory practices, which underscores the need for regulatory intervention. The article concludes by examining the implications of these findings for the effective regulation of AI and offering policy recommendations.

and its adaptability to changing political objectives and priorities within the EU. In this paper, the term "public interest" will be employed with the intent of emphasising the common good.

Section 1: AIDA background

The European Parliament established the AIDA on June 18th, 2020, with the intent to develop a long-term EU road-map on AI – analysing the impact and challenges of AI deployment, identifying common EU-wide objectives and proposing recommendations. Their task included producing a draft report in November 2021 and a final report in March 2022 on AI in the EU (European Parliament, n.d.). The AIDA was made up of 34 full members (see Figure 1 for full list of MEPs) and was chaired by Dragoş Tudorache (Renew Europe, Romania). With a mandate extended from 12 to 18 months, the committee conducted 30 hearings, eight coordinators' meetings and ten workshops from September 2020 to March 2022. These sessions covered a wide range of topics related to AI, including skills, employment, education, health, transport, environment, industry, e-government and third-country approaches. The hearings served as the primary source for gathering oral evidence from experts, policymakers and the business community. Sixteen expert witnesses were carefully selected based on their qualifications, expertise and diverse backgrounds, representing academia, industry, non-governmental organisations and research institutions. Their expertise enabled in-depth analysis of complex issues – such as data privacy and security, the ethical implications of AI algorithms and the regulation of AI in the public sector, where intricate challenges, including balancing innovation with citizen well-being, required in-depth analysis and discussion – ultimately leading to evidence-based recommendations and solutions. The presence of these witnesses was vital in shaping the debates, offering MEPs a broader perspective and deeper understanding of the topics at hand. MEPs engaged in interactive discussions with witnesses, fostering comprehensive debates. MEPs showed interest in harnessing AI's transformative potential while prioritising risk mitigation. The sessions aimed to strike a balance, exploring ways to achieve responsible AI that maximises its benefits while minimising harm, addressing concerns such as bias, discrimination and ethical implications (European Parliament, 2020). The objective was to achieve an equitable and EU-wide standard for AI regulation while finding the right balance between ethics, innovation and safeguarding rights (European Parliament, 2020).

Section 2: Literature review

This literature review delves into scholars' viewpoints on AI, highlighting the prevailing consensus on the existence of bias, lack of transparent decision-making and the need for regulation, while also recognising contrasting perspectives that emphasise the positive impact and transformative economic potential of AI in soci-

ety. The term "artificial intelligence" was coined by Professor John McCarthy in 1955, defining it as "the science and engineering of making intelligent machines" (Manning, 2020, p. 1). AI systems consist of algorithms (Babuta, Oswald & Rinik, 2018, p. 2) and serve as sets of instructions used in various industries and by the government, enabling specific outcomes such as depositing funds or granting access through swipe cards.

The role of AI systems

AI systems are pivotal in monitoring and predicting human behaviour through automated decision-making, using vast amounts of data to discern patterns and predict individual preferences, habits and future actions (Zuiderveen Borgesius, 2018, p. 13). Data drives algorithms and AI (Haggart and Tusikov, 2023, p. 6), facilitating their development, training and application in machine learning models. Industry representatives often portray algorithms as authoritative and influential, endorsing the concept of algorithmic neutrality (Haggart and Tusikov, 2023, p. 143). This fosters the belief that data, regarded as objective and unbiased, shapes human behaviour, leading to the perception that technology's utilisation of data is neutral (Beer, 2016, p. 7).

Challenges to algorithmic neutrality

The perception of algorithms as inherently neutral – thus endowed with considerable legitimacy and influence – is somewhat overstated, given that they are human-created rules and are susceptible to human biases (Haggart and Tusikov, 2023, p. 111). Scholars challenge the prevailing assumption of objectivity and neutrality in AI data, emphasising the growing recognition that complete objectivity in data-driven AI is unattainable (Leavy et al., 2020, p. 1). AI developers' limited understanding of AI systems' internal mechanisms complicates issues of accountability and comprehension (Siapka, 2018). Machine learning algorithms analyse vast datasets, detecting patterns and extrapolating beyond training set examples (Carney, 2020, p. 5). Advanced automated decision-making algorithms, including those for government service eligibility and recidivism risk evaluations, may oversimplify contextual complexities (Leavy et al., 2020, p. 3), with profound implications such as access limitations to essential services and unfavourable outcomes like benefit denials, increased scrutiny or incarceration (Haggart & Tusikov, 2023, p. 124).

AI and bias

AI systems, despite promises of higher intelligence and knowledge (Hwang, 2020), often exhibit inaccuracies and biases. AI models can perpetuate and amplify hu-

man prejudice and bias, leading to discriminatory outcomes (Leavy et al., 2020; O'Neil, 2016; Siapka, 2018). Bias in AI technology remains hidden but impactful, as software designers unintentionally embed bias into the design and operation of the systems (Haggart and Tusikov, 2023; O'Neil, 2016). Moreover, biased training data, such as racially biased police datasets, can result in the reproduction and amplification of those biases in predictive AI models (Lum & Isaac, 2016; Satzewich & Shaffir, 2009), perpetuating a cycle of discrimination that disproportionately affects marginalised groups (Leavy et al., 2020; Siapka, 2018; Zuiderveen Borgesius, 2018). Governments and companies employ 'neutral' datasets and algorithms that ultimately discriminate against specific groups, particularly women and racialised individuals (Eubanks, 2018; Zuiderveen Borgesius, 2018). The negative impact of AI systems extends beyond marginalised groups and affects society as a whole. For instance, technologies like automated resume readers and hiring surveys create barriers for applicants from racialised backgrounds or with mental health challenges, making it harder for them to secure job interviews (O'Neil, 2016, p. 101).

AI in decision-making

Pedro Domingos (2015) notes the increasing role of computers in decision-making processes across various domains, including credit, hiring, insurance rates, policing and arrests (p. 274). O'Neil (2016) highlights that while high-end decision-making involves human input, the majority of decisions in the public sector and lower sections of the economy are automated (p. 126). The crucial distinction lies in the fact that human decision-making can evolve with society, whereas AI systems remain stagnant, perpetuating past biases unless deliberately modified by engineers (O'Neil, 2016, p. 168). Furthermore, AI systems project historical patterns into the future, leading to harmful outcomes such as perpetuating poverty, increased incarceration rates, discriminatory practices in recidivism sentencing and predatory loan algorithms (Leavy et al., 2020; O'Neil, 2016; Siapka, 2018). These disparities created by public sector AI systems infringe upon individuals' human rights and affect the daily lives of the majority of the population, emphasising the need for concern and scrutiny regarding eligibility determination for government services. AI systems lack transparency, where opaque and invisible models are the norm, visible only to their developers (O'Neil, 2016; Rossi, 2018).

Positive impact of AI

While literature often highlights the potential harms of AI systems, some scholars acknowledge the multifaceted uses and benefits. For instance, AI-powered tools in

healthcare have shown promise in detecting eye diseases, identifying health risks and improving cancer screening (Haggart and Tusikov, 2023, p. 163). AI's positive impact is also evident in finance, retail, transportation, manufacturing and agriculture (O'Neil, 2016). The EU is hopeful that AI's positive impacts on various fronts will include stimulating innovation by driving advancements in technology and helping to boost economic growth and competitiveness through increased productivity and efficiency (Roberts et al., 2021, p. 6). However, Domingos (2015) takes a more optimistic stance, envisioning a "Master Algorithm" that can derive all knowledge from data (p. 40). His optimism contrasts with other scholars' concerns about the dangers associated with all-knowing technology.

AI regulation and the European Union

Scholars argue for AI regulation to counter dataism, challenging the belief in objective data-driven regulation (Haggart & Tusikov, 2023; Siapka, 2018). They recognise AI's limitations in ensuring fairness and propose measures such as algorithmic audits, impact assessments, legislative enforcement and avenues for challenging biased AI outputs (Balayn & Gürses, 2021; Bridges, 2001; O'Neil, 2016). However, the focus on de-biasing techniques is criticised for neglecting broader systemic issues and socio-technical contexts (Balayn & Gürses, 2021).

An examination of GDPR and biased AI reveals tensions between data protection and AI's reliance on diverse datasets (Siapka, 2018). Data curators play a pivotal role, underscoring the urgent necessity of democratising data through a wide array of perspectives and emphasising the intricate interplay of ethical and pragmatic factors in shaping the development of AI regulations and practices (Leavy et al., 2020, p. 2). Proposed legal solutions like value-sensitive design and regulatory sandboxes aim to navigate these challenges (Siapka, 2018). Trustworthy AI, rooted in fundamental rights and ethics, seeks to build public trust and ensure compliance (Siapka, 2018; Zuiderveen Borgesius, 2018). In the EU, discussions about the *AI Act* centre on balancing innovation and individual rights, with regulatory sandboxes gaining prominence (Ponce Del Castillo, 2021; Roberts et al., 2023). Scholars underscore the urgency of implementing measures to prevent discrimination in AI applications, aligning with the discussions within the EU about the *AI Act* (Leavy et al., 2020; Rossi, 2018). The EU's unique approach emphasises ethical boundaries, prohibiting high-risk AI systems and emphasising equality and redress (Csernatoni, 2019; High-Level Expert Group on AI, 2020; MacCarthy & Propp, 2021). Further revisions are suggested to address systemic risks such as collaboration among policymakers, regulators, AI makers and users, with recommendations including assessing technology capabilities, ensuring explainability, formulating redress

processes and supporting education curricula – ultimately emphasising the need for a comprehensive approach involving all stakeholders to ensure AI's positive societal impact (Roberts et al., 2023; Rossi, 2018).

To gain a comprehensive understanding of the potential consequences of AI technology and effectively mitigate its negative impacts on European society as a whole, it is imperative to conduct a comprehensive evaluation of policymakers' understanding of AI systems, bias, transparency and regulation.

Section 3: Methodological framework

I employed critical discourse analysis (CDA) to examine MEP speeches in my research. CDA enabled the identification and exploration of power dynamics and underlying meanings embedded in MEP language use. By analysing how MEPs frame arguments and employ language to shape discourse, CDA provided insights into the political and social contexts influencing decision-making processes (Bloor & Bloor, 2007/2013, p. 11). This method was appropriate for examining how the AIDA perceives and defines AI through their language use, highlighting opposing views and their impact on legislation and regulation in EU Parliament. The research focused on MEPs, their narrative, the intended audience (e.g. public/private sector, society at large) and the presentation of their views.

Critical Race Theory (CRT) was employed as a meta-methodology to complement CDA in examining MEP speeches. CRT provides a framework for understanding power dynamics and systemic inequalities within language and discourse, which is particularly pertinent when exploring issues of bias and discrimination in AI. Additionally, I incorporated the principles of the Social Construction of Technology (SCOT) – which emphasises the role of social factors in shaping technological development – as an additional meta-methodology to analyse how MEPs' language and attitudes contribute to the construction of AI within the AIDA committee and policymaking, shedding light on the social, political and economic factors shaping technology perception and regulation. These meta-methodologies, along with CDA, enriched the research's multifaceted approach to understanding MEPs' perspectives on AI.

The scope of this research paper is to uncover the answer to the following question: How do MEPs within the AIDA perceive and comprehend bias and discrimination in AI and what are their perspectives on the regulatory measures concerning AI?

Data collection for this study involved accessing European Parliamentary websites, which provided documents from the AIDA, including verbatim committee hearing transcripts, draft and final reports on AI and party statements. The primary data sources were speeches from the EU Parliament found in the AIDA committee hearing transcripts. Two key hearings of the AIDA were analysed, titled: *AI and Bias* (30 November 2021) and *AI and the Data Strategy* (30 September 2021). These debates involved 24 MEPs engaging in a 40-minute Q&A session with experts, totalling 3.3 hours. In the *AI and Bias* hearing, 18 MEP statements were analysed and in the *AI and the Data Strategy* hearing, 24 MEP statements were analysed. These debates were selected due to their relevance to the discussion on bias and discrimination in AI technologies. Each debate featured two panels comprising AI scholars, NGO directors and industry representatives as witnesses who presented their perspectives, concerns and suggestions for the EU's final AI report. MEPs then posed questions to the experts, resulting in dynamic debates that informed the drafting and amendment of the final AI report.

This study aimed to extract MEPs' views on AI and its implementation by analysing their language and attitudes expressed in questions and statements during the debates. The language used and the content of the statements offer valuable insights into MEPs' positions and perspectives on AI. By analysing the main themes and ideas emerging from witness testimonies and MEPs' questions, a deeper understanding of how the AIDA committee and European policymakers perceive AI was gained. In my discourse analysis, I applied the Procedural Approach by Strauss and Corbin (1990; 2015) to codify and thematically analyse statements in the transcripts. I employed recurring themes and established connections among them, continuously refining my codes. The analysis revealed three main themes: risk, bias and regulation. Eleven codes were utilised, including: discrimination, profit prioritisation, ethics and fairness, system regulation, government intervention, types of bias, accountability, self-regulation, industry regulation, state legislation and citizen input. These themes and codes were instrumental in understanding how AI was conceptualised in the AIDA debates.

Section 4: Discussion

In these two hearings, the issues of bias and discrimination, as well as regulation, took centre stage in the debate. Although no further issues arose from these specific hearings, it is worth noting that other hearings addressed various AI-related topics. For the purpose of this study, our focus remains on these two pivotal debates. MEPs face a paradoxical situation, recognising AI as a powerful societal tool

while acknowledging undemocratic consequences and inequitable outcomes within a capitalist system that prioritises private data ownership. MEPs aim to maintain Europe's global leadership in regulation and innovation, despite the accompanying challenges. Achieving an equilibrium between combating discrimination and fostering innovation in AI, especially when utilised by private entities for public interests, poses complex challenges in navigating conflicts between public usage and private control.

Bias and discrimination

Understanding the bias issue and its complexities

MEPs initially did not prioritise bias as a central topic in AI discussions, but in response to expert input and the need for further discussion on the matter, they dedicated a separate hearing to AI and bias. During discussions on addressing bias in AI, MEPs outlined various types of bias and their origins. Maria-Manuel Leitão-Marques (S&D) highlighted biases arising from non-representative training data, while noting that societal biases such as racism and sexism can still manifest even with representative data.

Challenges in detecting bias: The struggle with AI discrimination

Amongst the MEPs, there was a clear consensus that AI has the ability to discriminate, whether it be a result of biased data or in its biased application. Pilar del Castillo Vera (PPE) acknowledged the existence of bias and added that even a rigorously tested AI system may produce biased outcomes in real-world deployment. Furthermore, the experts emphasised that AI systems are trained on historical data, which fails to capture societal progress in combating discrimination. Bias in AI systems can stem from unintentional factors, negligence, or intentional actions and these variations necessitate tailored approaches to address them effectively. Elena Kountoura (The Left) discusses algorithmic biases, stating:

> we are aware that a number of systems incorporate algorithmic bias that generally targets the most vulnerable populations, thereby exacerbating inequalities and discrimination... The main problem, however, is the difficulty of detecting such bias, given the inbuilt opacity of certain AI-based systems.(*AIDA Public Hearing on AI and Bias*, p. 25)

The reference to "opacity" in the quote pertains to the system's lack of transparency and interpretability in its decision-making processes. Alessandra Basso (ID) was one of the MEPs who asked experts what can be done by legislators, "to protect

the most vulnerable, such as persons with disabilities, from possible prejudice arising from the misinterpretation or misuse of available data?" (*AIDA Public Hearing on AI and Bias*, p. 13).

While most MEPs acknowledged the presence of bias in AI systems and its role in discrimination, Kosma Złotowski (ECR) questioned whether a decision made by an AI system using objective and well-structured data could still be considered discriminatory. Złotowski suggested that disregarding AI decisions based on their favouritism or discrimination towards certain groups could be seen as "manipulating the technology" to conform to specific social or political views (*AIDA Public Hearing on AI and Bias*, p. 14). However, expert statements revealed that both the datasets used to train the technology and the technology itself exhibit bias. This underscores the inherent opacity of AI systems and the challenges encountered when attempting to discern bias within these systems.

Public sector concerns: AI's impact on structural inequalities

AI's impact on existing structural inequalities was another point of interest highlighted by Pernando Barrena Arza (The Left). Arza provided examples of Austria's use of AI-powered algorithms to offer social services, including scoring people based on their employment prospects and prioritising services based on that ranking and the Netherlands' use of algorithms to penalise people, predominantly in lower-income neighbourhoods, based on whether they were likely to have committed benefit fraud. His questions demonstrated concern for AI use in the public sector:

> How do we ensure that the algorithms used by the public services themselves are not biased? What sort of intervention is required in the public sector to ensure that artificial intelligence systems that are used are ethical, unbiased and do not penalise the most vulnerable in society? (*AIDA Public Hearing on AI and Bias*, p. 15)

This question demonstrates that MEPs are concerned about the potential ethical and social implications of AI used in the public sector. Although MEPs are considering possible interventions that could be taken to ensure that ethical standards are met, they are still unsure of the best route to regulating it.

Recognising AI risks: MEPs' understanding and concerns

MEPs understand that AI is risky, particularly in relation to the potential for the technology to perpetuate discrimination and inequality. MEPs unanimously recog-

nised that certain populations face heightened vulnerability in AI decision-making, specifically low-income individuals and racialised communities. MEPs demonstrated their understanding of AI being a risk to marginalised populations by providing examples of AI-powered algorithms discriminating against people when used in social services and examples that AI systems have delivered inaccurate, biased results because of biased datasets (Arza, *AIDA Public Hearing on AI and Bias*, p. 15). They also demonstrated their understanding of where this risk originates from, i.e. datasets and how the algorithms are trained (Leitão-Marques, S&D), existing structures of inequality (Kountoura, The Left) or the technology itself (Arza, The Left).

Addressing bias: Proposed solutions and regulatory frameworks

AI's widespread use and numerous applications make it crucial to address biases that can harm and violate the human rights of large populations within a short period. MEPs recognise the need to distinguish between technological biases and different forms of bias that require distinct remedies. Proposed solutions by MEPs include enhancing transparency and explainability of AI systems, fostering diversity and inclusion in AI development teams and strengthening regulatory frameworks for ethical and responsible AI development and use. However, these proposed solutions, while important, only scratch the surface and fail to fully address systemic issues, perpetuation of socio-economic inequalities, marginalisation of disadvantaged communities and reinforcement of discriminatory practices that disproportionately affect marginalised groups.

B. Regulation

The role of the EU and the need for accountability

MEPs emphasised the EU's responsibility to establish balanced government regulation for AI. MEPs generally support the use of AI in public sectors for its potential to enhance efficiency, effectiveness and accuracy, leading to improved services. However, they are not in favour of completely replacing human decision-making with AI. Kountoura (The Left) emphasised the need for algorithmic accountability, remarking that:

> Algorithmic accountability should include the obligation to report, explain or justify algorithmic decision-making and mitigate any adverse social impacts or potential damage. (*AIDA Public Hearing on AI and Bias*, p.25)

Kountoura's emphasis on algorithmic accountability highlights a conflict within

this stance. While she calls for reporting and explaining AI decisions to mitigate social impacts, this approach does not seek permission but rather aims to provide transparency. This implies that while AI's growing presence is accepted, MEPs view human involvement as a safeguard rather than a total replacement for human decision-making, underscoring a nuanced position.

Human involvement in decision-making: Balancing AI and human oversight

Kim Van Sparrentak (Verts/ALE) questioned whether algorithms should be completely excluded from certain decision-making scenarios due to the lack of accountability compared to humans. She asked:

> Should we perhaps not rely on algorithms in certain situations at all but require a human to be able to explain exactly why certain decisions are made and why they are justified, rather than an algorithm with a human somewhere in the loop or even human oversight? (*AIDA Public Hearing on AI and Bias*, p. 24)

Van Sparrentak suggested relying on humans to provide explicit justifications for decisions, acknowledging that humans are also prone to bias. This proposal highlights a broader distrust among MEPs regarding the impartiality of AI decision-making and emphasises the desire for greater accountability and transparency through human involvement.

Despite MEPs recognising the need for diverse datasets and diverse oversight, this solution may perpetuate unbalanced power structures. Assuming that "diverse" individuals can fully understand the discrimination faced by all groups is flawed, as people with intersecting identities experience different forms of discrimination (Crenshaw, 1991). Relying on diverse human oversight in AI can be seen as a form of identity politics, reducing individuals to their social identities rather than acknowledging their individual experiences and perspectives (Crenshaw, 1991; Crenshaw, 1989).

Sergey Lagodinsky (Verts/ALE) expressed concerns about the privatisation of public functions and the possibility of human actors being replaced by privately-controlled AI systems. This emphasises the need for legislation to achieve a balanced approach, preserving human decision-making and preventing undue influence from private entities, thereby protecting democratic values and public interest (*AIDA Public Hearing on AI and the Data Strategy*, p. 35). Basso (ID) used the example of facial recognition – a high-risk AI system that, "if left in the hands of a system of self-checks, places great power in the hands of those providing the service it-

self" (*AIDA Public Hearing on AI and the Data Strategy*, p. 37) – to propose a more democratic form of oversight that involves public participation in decision-making processes. These inquiries demonstrate MEPs' awareness of the importance of safeguarding human decision-making, mitigating private influence and ensuring democratic values and public interests are upheld. The call for democratic oversight aims to address concerns about potential misuse or abuse of AI, engaging various stakeholders in well-informed debates and responsible deployment. This highlights MEPs' unease regarding the privatisation of the AI ecosystem, rather than the AI itself, which makes them cautious about AI's use and therefore stresses the importance of human involvement in the decision-making processes.

AI as the engine, data as the fuel: The role of data in AI

Miapetra Kumpula-Natri stresses the importance of regulated AI systems and data, stating:

> One way to think of the relation between the data and AI is that if AI algorithms are the engine, the data is the fuel. If we have the finest engine, it's useless if we do not have the necessary fuel. (*AIDA Public Hearing on AI and the Data Strategy*, p. 4)

This perspective, shared by other MEPs, advocates for a trustworthy and transparent data economy where the EU regulates data usage, access and personal control (*AIDA Public Hearing on AI and the Data Strategy*, p. 4). However, this understanding overlooks critical aspects. By treating data as fuel, MEPs risk reducing individuals to mere inputs, neglecting the complex social, economic and political contexts of data generation and its potential consequences. Furthermore, the fuel analogy fails to acknowledge the potential harms of unregulated data collection and use. Examples like biased facial recognition databases, resulting from the disproportionate policing of racialized communities, highlight the risk of discriminatory practices and criminalisation of racialised individuals (Lum & Isaac, 2016, p. 19). This reflects the prioritisation of technological development over marginalised communities' interests.

Importance of citizen involvement: Engaging citizens in AI regulation

MEPs emphasised the importance of citizen involvement in AI regulation, recognising that citizens are the most affected by these technologies. Alexandra Geese (Verts/ALE) advocated for involving affected groups in AI development to address the lack of diversity in the tech industry, drawing inspiration from Germany's coali-

tion agreement (*AIDA Public Hearing on AI and Bias*, p. 12). Kountoura (The Left) emphasised the need for transparency and understanding, stating, "details should be provided regarding the workings of mass data analysis, thereby helping individuals to understand and keep track of the decisions affecting them" (*AIDA Public Hearing on AI and Bias*, p. 25).

Maria da Graça Carvalho (PPE) emphasised the importance of data literacy, stating, "data literacy will be critical to guarantee that citizens embrace the opportunities of data... and understand the environment and its risks." (*AIDA Public Hearing on AI and the Data Strategy*, p. 33). However, I argue that merely promoting diversity and individual responsibility is insufficient in addressing systemic inequalities and power imbalances underlying technological innovation. Individual responsibilisation highlights a common trend in policymaking where the burden for addressing issues is placed on the individual rather than on the institutions and systems that create and perpetuate these issues (Haggart and Tusikov, 2023, p. 112). This approach neglects the systemic issues surrounding data collection, usage and power dynamics. It also fails to acknowledge the limited resources and knowledge individuals may possess. Individual responsibilisation allows policymakers to evade accountability for the harm caused and places the burden on those being harmed.

Challenges in comprehending AI: Understanding AI's decision-making

MEPs face challenges in comprehending the decision-making processes of AI systems due to their complexity. We can see that some MEPs understand AI to be a sentient technology. Gilles Lebreton (ID) raised a question:

> But don't you think that artificial intelligence has now evolved to such an extent that it is capable of creating its own selection criteria, and therefore that bias of artificial origin is also a possibility? (*AIDA Public Hearing on AI and Bias*, p. 24)

Attributing agency and autonomy to AI is a significant phenomenon that has gained prevalence. Some MEPs perceive AI's ability to make decisions and act as sentient, ascribing agency to it. Adriana Maldonado López (S&D) spoke to her confidence in AI systems being able to self-regulate and self-correct biases inherent to the technology.

Algorithms, despite being created by humans, have been endowed with autonomy, minimising human agency in their creation and use (Ziewitz, 2016, p. 5). This shift in responsibility has implications, absolving creators and organisations from accountability for automated regulation's impacts (Haggart and Tusikov, 2023, p.

124). While MEPs acknowledge AI's decision-making abilities, the technology itself has not reached a point where it can be referred to as sentient (Husain, 2017, p. 27). MEPs, including Basso (ID) and Kountoura (The Left), raised doubts about whether AI can be considered a "self-sentient machine". Their questions reveal a lack of clarity regarding AI's sentience, agency and decision-making abilities, indicating that MEPs have not yet reached a consensus or a unified understanding on these aspects of AI. While humans attribute autonomy and power to AI, it is not truly autonomous as it lacks self-awareness, consciousness and independent decision-making. However, software programs can exert social power by shaping and directing human behaviour, exhibiting a form of intelligence and adaptability. The belief in AI's self-correcting abilities may downplay its potential threats, but the actual capabilities of AI in this regard remain uncertain. Creating self-aware or sentient machines is still a speculative and distant prospect. Furthermore, ethical and social concerns arise even if such machines were possible, including the need to ensure their alignment with public interests. MEPs' emphasis on human-centric and policymaker-led regulation implies scepticism towards AI's current ability to self-regulate effectively without causing harm.

Accountability and definitions: Grappling with AI regulation

MEPs acknowledge the inability of algorithms and AI systems to be held accountable, especially when their decisions directly impact individuals' livelihoods, reflecting a deep distrust in the fairness and impartiality of AI. Concerns regarding bias and discrimination in AI decision-making have fueled this distrust. MEPs are therefore advocating for human oversight of "high-risk" AI systems that directly affect humans, including impact assessments throughout the development process (see Bridges, 2001). Discussions differentiating between high-risk and low-risk applications reveal ongoing deliberations on the appropriateness of AI in specific contexts, reflecting evolving debates on AI regulation and governance. However, the feasibility of implementing a rights-based approach and ensuring accountability in AI development raises important questions. MEPs have yet to determine the party responsible for bearing the burden of accountability – whether it is the system designer, operator, or regulators.

A rights-based approach necessitates clear definitions of human rights violations and "fair outcomes" in AI decision-making, along with monitoring and enforcement mechanisms. It requires substantial resources, expertise, robust accountability and auditing processes to ensure AI systems align with human rights principles. MEPs need to provide clear definitions for key terms such as "fairness", "transparency" and "accountability" in their efforts to regulate AI. Castillo Vera (PPE) underscored

the need for MEPs to define what a 'fair' outcome is and that there must be an evaluation of the nature of the AI systems to determine which is the best metric for mitigating potential risks. The absence of precise definitions raises concerns about their strategies for addressing bias and implementing standards. This lack of clarity may stem from conflicting views or a lack of awareness. It hampers the development of effective regulations, resulting in legal ambiguity and challenges in enforcement. Superficial understanding of these terms impedes the establishment of standardised guidelines for AI systems, hindering efforts to align regulation with EU values. Stakeholders must establish shared understandings and definitions to develop a robust regulatory framework. Failure to define these terms can have detrimental consequences, as corporate interests may be prioritised over citizen protection.

Limitations

Party politics introduce complex limitations to the study, influencing MEPs' questioning strategies. Party affiliations shape the topics, tone and framing of questions, potentially compromising the reflection of individual attitudes and knowledge. This can prioritise party objectives over transparency. Recognising party politics as a significant limitation when inferring MEPs' true attitudes from their hearing questions is crucial.

The statements released by each of the main political parties delineate distinct party-driven perspectives on AI development, bias mitigation and the management of the data economy. Notably, the EPP Group underscores the importance of AI bias detection and mitigation, data quality and international collaboration, in line with their party's proclivity for fostering ethical and human-centric AI solutions in a global market context. In contrast, the S&D Group prioritises strict adherence to EU legislation, human rights and non-discrimination in AI, consistent with their commitment to ensuring equitable treatment and engendering trust among citizens. Renew Europe's strong emphasis on fundamental rights and data fairness mirrors their party's overarching values. The Greens/EFA, meanwhile, place a strong focus on averting AI bias and granting citizens greater empowerment, a reflection of their commitment to a human-centric approach to AI.

These politically-driven priorities exert a substantial influence over the nature of MEP questions during AI-related hearings and the overall trajectory of these deliberations. This necessitates a comprehensive understanding when interpreting MEPs' true attitudes based solely on their hearing inquiries. Additionally, the analysis suggests a notable shift towards greater public involvement and human-

centric AI policy, reflecting parties' responsiveness to debate concerns. Recognising the nuances introduced by party politics is crucial when discussing the limitations of this study, aiming to discern MEPs' genuine sentiments from their hearing engagement

Additionally, in this study, it is important to note that I exclusively examine MEPs' views and comprehension of AI as expressed during the AIDA hearings leading up to the enactment of the *AI Act*. While these hearings provide a valuable snapshot of their perspectives within that specific context, it is vital to acknowledge that the study's scope is confined to this particular time frame and does not encompass MEPs' views beyond these hearings or in response to subsequent developments in the AI field. The rapidly evolving nature of AI necessitates recognising this temporal limitation when interpreting MEPs' views and understanding of AI in the broader context.

Conclusion

The EU seeks a delicate balance between AI's economic benefits and social implications, as observed in the AIDA hearings. While AI offers economic advancement and improved social outcomes, addressing bias and discrimination is crucial to prevent exacerbating societal inequalities. To safeguard vulnerable individuals and minimise the influence of technology companies, it is essential to comprehend the significance of prioritising the social costs and value of AI over corporate interests.

The European Parliament's negotiating position on the *AI Act*, as reflected in the final draft of the *AI Act,* bears a clear imprint of the discussions and recommendations made during the AIDA hearings, encompassing several key modifications to the proposed legislation. First, they have aligned the definition of AI systems with the Organisation for Economic Co-operation and Development (OECD) standards, aiming for global consistency (European Parliament, 2023). Additionally, they advocate for the prohibition of AI systems encompassing both real-time and ex-post use as well as biometric categorisation systems using sensitive characteristics, predictive policing systems, emotion recognition systems and indiscriminate scraping of biometric data (European Parliament, 2023). High-risk AI systems, according to the Parliament's position, must not only fall within certain areas or use cases, but also pose a "significant risk" to health, safety, fundamental rights, or the environment. They emphasise a layered approach to regulate general-purpose AI systems, ensuring robust protection of fundamental rights and imposing transparency obligations on generative foundation AI models (European Parliament, 2023). These key positions directly stem from nuanced discussions during the AIDA hear-

ings. The strengthening of governance, the establishment of an AI Office and the empowerment of national authorities also align with the need for robust enforcement mechanisms highlighted in these discussions (European Parliament, 2023). Moreover, the Parliament's dedication to fostering research and innovation, evident in the exemption of research activities and open-source AI components from certain regulations, underscores the careful balance between regulation and advancement discussed during the AIDA hearings (European Parliament, 2023). Overall, the *AI Act*'s evolution reflects a dynamic and adaptive response to the multifaceted issues explored in the AIDA hearings.

MEPs acknowledge AI's pivotal role in Europe's economy, appreciating its potential to enhance efficiency and competitiveness. Nonetheless, they acknowledge the associated risks, as highlighted by the comprehensive 65-page report on AI and the implementation of the *AI Act*. This report underscores the urgent need for regulation, transparency and accountability. Given the data-driven nature of the economy, the state assumes a central role in resource allocation and regulatory measures. The economic impetus of EU MEPs to foster innovation and maintain a competitive edge in the technology sector contrasts with their social objective of ensuring fairness, transparency and accountability in AI. These dual motivations are not compatible, representing a complex challenge.

MEPs display scepticism towards AI's abilities, perceiving it as potentially sentient and recognising the need for intervention to address biases and regulate its use. The EU's establishment of AIDA and the *AI Act* reflects its intention to regulate AI through government intervention. However, there are differing opinions among MEPs regarding the regulatory entities for different types of AI. Trust in AI use in the public sector is low, potentially leading to strict regulations for "high-risk" AI in this domain. Conversely, MEPs may adopt an industry-centric approach for "low-risk" AI, setting standards for private AI development to foster innovation and competitiveness. This cooperative framework aims to address concerns and promote responsible AI practices. Nonetheless, achieving a balanced approach may prove challenging, as profit-seeking often conflicts with societal well-being objectives.

CRT highlights power imbalances (Delgado & Stefancic, 2001) and how evolving technology perpetuates racialised power structures (O'Neil, 2016). Marginalised groups remain impacted by technology controlled by more powerful social groups, reinforcing hierarchies. This reveals power imbalances through identity politics (Crenshaw, 1991), with CRT showing how AI systems sustain systemic discrimination (Crenshaw, 1989). The hearings missed the depth of these issues, emphasising the need for MEPs to deepen their understanding of AI's societal implications and

address AI's complex ramifications.

MEPs recognise AI's potential as an ally in combating inherent biases by developing new technologies. CRT and SCOT acknowledge that bias in AI is not solely the result of individual actions, but a systemic phenomenon embedded in legal, social and economic structures. Eliminating biases in AI thus requires addressing the underlying social and economic structures shaping its development. This entails reevaluating system design, tackling systemic issues such as unequal access to education and resources and upholding fairness and social justice principles. Centring marginalised perspectives in AI development, along with ongoing monitoring and regulation, is crucial for achieving equitable outcomes and mitigating bias. CRT should guide AI system development, considering the dynamics of power, privilege, discrimination and exclusion in society.

MEPs highlight the role of citizen participation in mitigating AI risks, emphasising individual and social responsibility. They stress the need for a trustworthy and transparent data economy where individuals have control over their data and influence over its usage. However, this places the onus on individuals to educate themselves and demand accountability. The question arises: is this enough? While data literacy is crucial, European regulations are lagging, potentially compromising citizen protection. Without effective regulations, relying solely on individual responsibility will not prevent data misuse or ensure citizen safety.

The EU *AI Act* serves as a pioneering model for other countries. Insights from this research can guide future policy development within the EU and globally. As democratic nations navigate AI's challenges and opportunities, further research will examine policymaking approaches in different contexts, informing effective and equitable regulations. This study, exploring MEPs' understanding and perception of AI, establishes a valuable foundation for future research, illuminating policymakers' attitudes towards emerging technologies.

References

AIDA public hearing on AI and bias. (2021). [Verbatim report]. European Parliament. https://www.europarl.europa.eu/committees/en/aida/home/publications?tabCode=vebatim-reports

AIDA public hearing on AI and the data strategy. (2021). [Verbatim report]. European Parliament. https://www.europarl.europa.eu/committees/en/aida/home/publications?tabCode=vebatim-reports

Balayn, A., & Gürses, S. (2021). *Beyond debiasing: Regulating AI and its inequalities* [Report]. European Digital Rights- EDRi. https://edri.org/our-work/if-ai-is-the-problem-is-debiasing-the-solut

ion/

Beer, D. (2016). How should we do the history of big data? *Big Data & Society*, *3*(1), 1–10. https://do i.org/10.1177/2053951716646135

Beer, D. (2017). The social power of algorithms. *Information, Communication & Society*, *20*(1), 1–13. h ttps://doi.org/10.1080/1369118X.2016.1216147

Bloor, M., & Bloor, T. (2013). *The practice of critical discourse analysis: An introduction* (Reprint). Routledge. https://doi.org/10.4324/9780203775660

Brey, P. (2005). Artifacts as social agents. In H. Harbers (Ed.), *Inside the politics of technology: Agency and normativity in the co-production of technology and society* (pp. 61–84). Amsterdam University Press. https://www.jstor.org/stable/j.ctt45kcv7.6

Bridges, L. (2001). Race, law and the state. *Race & Class*, *43*(2), 61–76. https://doi.org/10.1177/0306 396801432005

Carney, T. (2020). Artificial intelligence in welfare: Striking the vulnerability balance. *Monash University Law Review*, *46*(2), 1–30. https://doi.org/10.26180/13370369.v1

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, *1*, Article 8.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against Women of Color. *Stanford Law Review*, *43*(6), 1241–1299. https://doi.org/10.2307/1229039

Csernatoni, R. (2019). *An ambitious agenda or big words? Developing a European approach to AI* (Policy Brief 117). Egmont Institute. https://www.jstor.org/stable/resrep21397

de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & Society*, *35*(4), 917–926. https://doi.org/10.1007/s00146-020-00960-w

Delgado, R., & Stefancic, J. (2001). *Critical race theory: An introduction* (1st ed.). New York University Press. http://www.jstor.org/stable/j.ctt9qg26k.

Domingos, P. (2015). *The master algorithm: How the quest for the ultimate learning machine will remake our world*. Basic Books.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor* (First Edition). St. Martin's Press.

European Commission. (2021). *Europe fit for the digital age: Commission proposes new rules and actions for excellence and trust in artificial intelligence* [Press release]. https://ec.europa.eu/commissio n/presscorner/detail/en/ip_21_1682)

European Parliament. (2020). *Setting up a special committee on artificial intelligence in a digital age, and defining its responsibilities, numerical strength and term of office* [Decision]. https://www.europar l.europa.eu/doceo/document/TA-9-2020-0162_EN.html

European Parliament. (2023). *Parliament's negotiating position on the artificial intelligence act* [Plenary report]. EPRS | European Parliamentary Research Service. https://www.europarl.europa.eu/ RegData/etudes/ATAG/2023/747926/EPRS_ATA(2023)747926_EN.pdf

European Parliament. (n.d.). *About: Welcome words*. AIDA | Committees | European Parliament. http

s://www.europarl.europa.eu/committees/en/aida/about

Haggart, B., & Tusikov, N. (2023). *The new knowledge: Information, data and the remaking of global power*. Rowman & Littlefield.

High-Level Expert Group on AI (AI HLEG). (2020). *Assessment list for trustworthy artificial intelligence (ALTAI) for self-assessment* [Independent report]. European Commission. https://digital-strategy.ec.eu ropa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment

Hossfeld, C., Muller-Lagarde, Y., & Zevounou, L. (2020). The evolution of the European public good assessment in the EU endorsement process of IFRS. *Accounting in Europe*, *17*(3), 314–333. https://do i.org/10.1080/17449480.2020.1818799

Husain, A. (2017). *The sentient machine: The coming age of artificial intelligence*. Scribner.

Hwang, T. (2020). *Subprime attention crisis: Advertising and the time bomb at the heart of the internet*. FSG Originals.

Klein, H. K., & Kleinman, D. L. (2002). The social construction of technology: Structural considerations. *Science, Technology, and Human Values*, *27*(1), 28–52. https://doi.org/10.1177/01622 4390202700102

Leavy, S., O'Sullivan, B., & Siapera, E. (2020). *Data, power and bias in artificial intelligence* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2008.07341

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19. https://doi.org/10.111 1/j.1740-9713.2016.00960.x

MacCarthy, M., & Propp, K. (2021, April 28). Machines learn that Brussels writes the rules: The EU's new AI regulation. *Lawfare*. https://www.lawfareblog.com/machines-learn-brussels-writes-rules-eu s-new-ai-regulation.

Manning, C. (2020). *Artificial intelligence definitions* [Glossary]. Human Centered Artificial Intelligence, Stanford University. https://hai.stanford.edu/sites/default/files/2020-09/AI-Definitions-HAI.pdf

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Ponce Del Castillo, A. (2021). *The AI regulation: Entering an AI regulatory winter? Why an ad hoc directive on AI in employment is required* (European Economic, Employment and Social Policy, pp. 1–9) [Policy brief]. European Trade Union Institute (ETUI). https://doi.org/10.2139/ssrn.3873786

Roberts, H., Cowls, J., Hine, E., Mazzi, F., Tsamados, A., Taddeo, M., & Floridi, L. (2021). Achieving a 'good AI society': Comparing the aims and progress of the EU and the US. *Science and Engineering Ethics*, *27*(6), Article 68. https://doi.org/10.1007/s11948-021-00340-7

Roberts, H., Cowls, J., Hine, E., Morley, J., Wang, V., Taddeo, M., & Floridi, L. (2023). Governing artificial intelligence in China and the European Union: Comparing aims and promoting ethical outcomes. *The Information Society*, *39*(2), 79–97. https://doi.org/10.1080/01972243.2022.2124565

Rossi, F. (2018). Building trust in artificial intelligence. *Journal of International Affairs*, *72*(1), 127–134.

Satzewich, V., & Shaffir, W. (2009). Racism versus professionalism: Claims and counter-claims about racial profiling. *Canadian Journal of Criminology and Criminal Justice*, *51*(2), 199–226. https://doi.org/

10.3138/cjccj.51.2.199

Siapka, A. (2018). *The ethical and legal challenges of artificial intelligence: The EU response to biased and discriminatory AI* [Thesis, Panteion University of Athens]. doi.org/10.2139/ssrn.3408773

Yousefikhah, S. (2017). Sociology of innovation: Social construction of technology perspective. *AdMinister*, *30*, 31–43. https://doi.org/10.17230/ad-minister.30.2

Ziewitz, M. (2016). Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*, *41*(1), 3–16. https://doi.org/10.1177/0162243915608948

Zuiderveen Borgesius, F. (2018). *Discrimination, artificial intelligence, and algorithmic decision-making* (pp. 1–91) [Study]. Directorate General of Democracy, Council of Europe. https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decisionmaking/1680925d73

Appendix

GROUP	TOTAL	MEMBERS	SUBSTITUTES
epp	18	ASIMAKOPOULOU Anna-Michelle CARVALHO Maria da Graça DEL CASTILLO VERA Pilar DIDIER Geoffroy MAYDELL Eva RESSLER Karlo SIKORSKI Radosław TERRAS Riho VOSS Axel	ARIAS ECHEVERRÍA Pablo METSOLA Roberta NIEBLER Angelika SALINI Massimiliano ŠTEFANEC Ivan VIRKKUNEN Henna WALSMANN Marion WARBORN Jörgen WISELER-LIMA Isabel
S&D	14	BENIFEI Brando GARCÍA DEL BLANCO Ibán KUMPULA-NATRI Miapetra MIKSER Sven NICA Dan SCHALDEMOSE Christel SIPPEL Birgit	AGIUS SALIBA Alex GEBHARDT Evelyne HRISTOV Ivo KAILI Eva LEITÃO-MARQUES Maria-Manuel MALDONADO LÓPEZ Adriana ROBERTI Franco
renew europe.	10	ANSIP Andrus HAHN Svenja SÉJOURNÉ Stéphane SOLÍS PÉREZ Susana TUDORACHE Dragoș	CHARANZOVÁ Dita DANTI Nicola IN'T VELD Sophia MELCHIOR Karen SCHREINEMACHER Liesje
THE GREENS/EFA	6	CORMAND David GEESE Alexandra VAN SPARRENTAK Kim	BOESELAGER Damian KOLAJA Marcel LAGODINSKY Sergey
ID	6	ANDERSON Christine BARDELLA Jordan BASSO Alessandra	KRAH Maximilian LEBRETON Gilles PANZA Alessandro
ECR	6	BIELAN Adam BOURGEOIS Geert ZŁOTOWSKI Kosma	KARSKI Karol STANCANELLI Raffaele TOŠENOVSKÝ Evžen
	4	KOUNTOURA Elena MAUREL Emmanuel	BARRENA ARZA Pernando PEREIRA Sandra
NI	3	RADAČOVSKÝ Miroslav SCHALLER-BAROSS Ernő	PIGNEDOLI Sabrina

FIGURE 1: Chart listing MEPs and their group membership.