IJRAR.ORG E-ISSN: 2348-1269, P-ISSN: 2349-5138

INTERNATIONAL JOURNAL OF RESEARCH AND ANALYTICAL REVIEWS (IJRAR) | IJRAR.ORG

An International Open Access, Peer-reviewed, Refereed Journal

Evaluating the Impact of Data Bias in AI on Everyday Decision-Making

Ashwin Purohit[1] Asst. Prof. Mr. Sumitkumar Tripathi[2]

[1]Department of Information Technology, [2]Department of Information Technology,
SHETH L.U.J. & SIR M.V. COLLEGE OF ARTS, SCIENCE & COMMERCE,
University of Mumbai, Mumbai, India.

ABSTRACT:

This research embarks on a critical exploration of the often-overlooked aspect of Artificial Intelligence – the inherent biases in data that shape AI decisions. This study is rooted in a comprehensive document analysis, aiming to unravel the extent to which data bias infiltrates various sectors, including employment, healthcare, and finance, and how it influences the perceptions and decisions of individuals within these realms.

The core objective of this research is to discern whether the influence of data bias in AI systems leads to predominantly negative or positive outcomes in everyday decision-making. It delves into the public's awareness and understanding of data bias in AI, gathering insights through surveys and interviews to gauge whether individuals are apprehensive or accepting of AI's role in their professional and personal environments.

A significant focus of this study is to articulate the reasons behind any scepticism or acceptance observed among the participants. It seeks to identify the characteristics and strategies that individuals can adopt to adapt to and thrive in an AI-driven world, where biassed data could potentially skew AI decisions.

By the conclusion of this research, we aim to offer a nuanced perspective on the impact of data bias in AI, enriching the discourse on AI's role in society. The findings are intended to enlighten a diverse audience, from AI developers to the general public, about the intricacies of AI and its implications, fostering a more informed and balanced view of AI's integration into daily life.

Keyword: Data bias in AI, Artificial intelligence

INTRODUCTION:

In the contemporary landscape of technological advancement, Artificial Intelligence (AI) stands at the forefront, heralding a new era of innovation and efficiency. AI's integration into various sectors – from healthcare and finance to education and employment – has been transformative, reshaping how decisions are made and actions are executed. However, with great power comes great responsibility, and the burgeoning influence of AI brings to light critical concerns, particularly regarding the biases inherent in the data that fuel these intelligent systems. The research aims to dissect and understand the multifaceted implications of data bias in AI and its ripple effects on daily life.

The genesis of this study lies in the growing recognition that AI, despite its potential for objectivity and precision, is not immune to the prejudices and partialities that plague human decision-making. Data bias in AI emerges from various sources – historical data sets reflecting past inequalities, skewed user-generated content, or unrepresentative sampling – and these biases can inadvertently become embedded in AI algorithms. The consequences are far-reaching, influencing decisions in areas as critical as job recruitment, loan approvals, medical diagnoses, and legal judgments.

This research is predicated on the hypothesis that data bias in AI does not merely exist in a vacuum but actively impacts the decisions and perceptions of individuals in their everyday lives. Through a methodical document analysis complemented by surveys and interviews, this study seeks to capture the pulse of public sentiment regarding AI's role in their professional and personal spheres. Are people cognizant of the biases present in AI? Do they view AI's decision-making capabilities with scepticism or acceptance? And crucially, what are the underlying reasons for their attitudes?

Moreover, this study endeavours to go beyond mere observation and analysis. It aims to chart a course for how individuals can adapt to an AI-driven world, identifying skills and strategies to mitigate the impact of data bias and secure their place in an increasingly automated future.

In essence, this research is not just an academic exercise; it is a journey to unravel the complex interplay between AI and human life, seeking to provide clarity and guidance in an age where AI's influence is both undeniable and enigmatic.

This research aims to dissect and understand the multifaceted nature of data bias in AI systems and its consequential effects on individuals and society. The following are key components of this research:

1. **Understanding the Origins of Data Bias:** At the heart of this study is an exploration of where and how data biases originate. This includes biases in historical datasets, user-generated content, and algorithmic processing. Recognizing the sources of bias is crucial for understanding its impact on AI decisions.

2. **Analysing the Impact on Decision-Making:** A significant focus of this research is to examine how data bias in AI affects decision-making in critical sectors like employment, healthcare, and finance. This involves investigating the extent to which biassed AI decisions influence individual choices and societal outcomes.

3. **Public Perception and Awareness:** This study seeks to gauge public awareness and attitudes towards data bias in AI. Are individuals cognizant of the biases present in AI systems? How do these perceptions shape their trust and reliance on AI in their daily lives?

4. **Mitigation and Adaptation Strategies:** A pivotal aspect of this research is identifying and evaluating strategies to mitigate the impact of data bias in AI. This includes exploring diverse data collection practices, algorithmic fairness techniques, and policy recommendations to ensure more equitable AI systems.

5. **Ethical and Social Implications:** The ethical dimensions of data bias in AI are a critical component of this study. It aims to address the broader ethical questions and social implications of biassed AI, including issues of fairness, privacy, and societal impact.

6. **Informing Policy and Practice:** By understanding the impact of data bias in AI, this research aims to inform policymakers, AI developers, and stakeholders. The goal is to contribute to the development of policies and practices that address data bias, promoting more responsible and inclusive AI.

7. **Long-Term Societal Impact:** Beyond immediate effects, this research considers the long-term societal impact of data bias in AI. Understanding these implications is vital for preparing society for an AI-driven future, ensuring that AI technologies are developed and used in ways that benefit all members of society.

OBJECTIVE OF RESEARCH:

While AI systems are designed to enhance decision-making processes, the presence of data bias can significantly alter their effectiveness and fairness. Data biases in AI, stemming from skewed datasets or prejudiced algorithmic design, can lead to decisions that inadvertently perpetuate societal inequalities. These biases can be particularly impactful in sectors where AI plays a critical role, such as healthcare, law enforcement, and financial services. The effectiveness of AI in these areas is often contingent on the neutrality and representativeness of the data it processes.

This study will investigate the extent to which individuals are aware of and concerned about data bias in AI and its impact on their daily lives. It will explore the specific aspects of decision-making that people believe are most susceptible to data bias and the potential consequences of these biases.

Additionally, the study will examine what measures individuals think can be implemented to prevent or mitigate the effects of data bias in AI. This includes exploring public opinion on strategies such as diversifying data sources, implementing fairness algorithms, and establishing regulatory frameworks. The research aims to provide a comprehensive understanding of public perception regarding data bias in AI and to propose actionable solutions for creating more equitable and unbiased AI systems.

LITERATURE:

Overview of Data Bias in AI

Causal effect of racial bias in data and machine learning algorithms on user persuasiveness & discriminatory decision making: An Empirical Study (Publication Date: January 22, 2022)
This paper delves into the specific impacts of racial bias present in data sets and machine learning algorithms. It examines how these biases can undermine user trust and lead to decisions that are discriminatory. The study is significant in bridging the gap between theoretical understanding of bias and its practical, harmful outcomes, particularly in customer-related decisions. It underscores the importance of creating AI systems that are not only technically proficient but also ethically responsible and equitable.

AI Bias and Decision-Making Models

AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry by Belenguer, E. (Publication Date: February 10, 2022)
Belenguer's research explores the complex landscape of AI bias, particularly focusing on algorithmic decision-making models that exhibit discriminatory tendencies. The paper argues for the establishment of a transnational, independent body empowered to enforce solutions for AI bias. It draws an interesting parallel with the pharmaceutical industry, suggesting that machine-centric solutions used there could be adapted to

address biases in AI. This approach underscores the need for robust, globally recognized standards and practices to combat AI bias.

AI in Organisational Decision-Making

Impact of Artificial Intelligence on Decision-making in Organisations (Publication Date: August 9, 2023)
This study provides an insightful analysis of how AI impacts decision-making within organisations. It emphasises the dual role of AI in enhancing decision-making capabilities while also presenting ethical challenges. The paper discusses the potential of AI to transform organisational processes but also cautions about the need for a deep understanding of AI functions and ethical implications. It suggests that the future of organisational decision-making will likely be a blend of AI and human intelligence, necessitating a balanced approach to AI integration.

Accountability in AI-Based Decision Systems

System Cards for AI-Based Decision-Making for Public Policy (Publication Date: March 1, 2022)
The paper introduces an innovative concept of 'system cards' as a tool for auditing AI-based decision-making systems, especially in the context of public policy. These system cards are designed to act as comprehensive scorecards that present the outcomes of formal audits, providing a clear and structured overview of AI system performance and accountability. This framework is proposed as a means to ensure transparency and responsibility in AI decision-making, highlighting the importance of regular and formal evaluations of AI systems used in public policy.

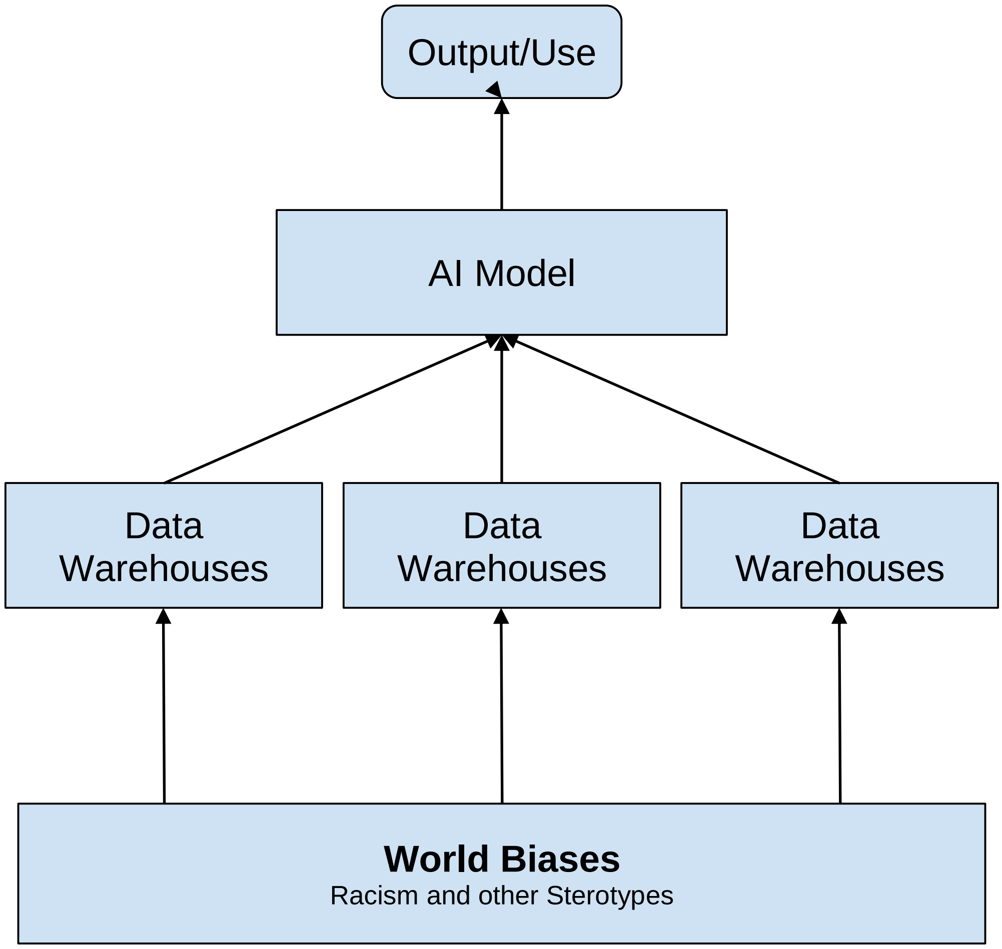
Addressing and Mitigating Bias in AI

Bias in AI-systems: A multi-step approach by Eirini Ntoutsi (Publication Date: 2020)
Ntoutsi's paper takes a detailed look at the issue of bias in AI systems, emphasising that bias can be introduced and amplified at various stages of the AI learning process. The study explores the different sources of bias and proposes a multi-step approach to identify, address, and mitigate these biases. It emphasises the need for continuous vigilance and intervention at each stage of AI development to prevent the propagation of bias. The paper is particularly valuable for its comprehensive approach to understanding and tackling bias in AI, offering practical strategies for developers and practitioners.

Bias in data-driven artificial intelligence systems—An introductory survey
Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, Ioannis Kompatsiaris, Katharina Kinder-Kurlanda, Claudia Wagner, Fariba Karimi, Miriam Fernandez, Harith Alani, Bettina Berendt, Tina Kruegel, Christian Heinze, Klaus Broelemann, Gjergji Kasneci, Thanassis Tiropanis, Steffen Staab

MATERIALS & METHODS:

For this research and understand the origin of Bias we are going to use a methodology, that is Categorising the Biases based on the source they origin from.



These biases, categorised as World, Data, Use, and Design biases, manifest at different stages of AI development and deployment, painting a comprehensive picture of the origins and impacts of bias in AI systems.

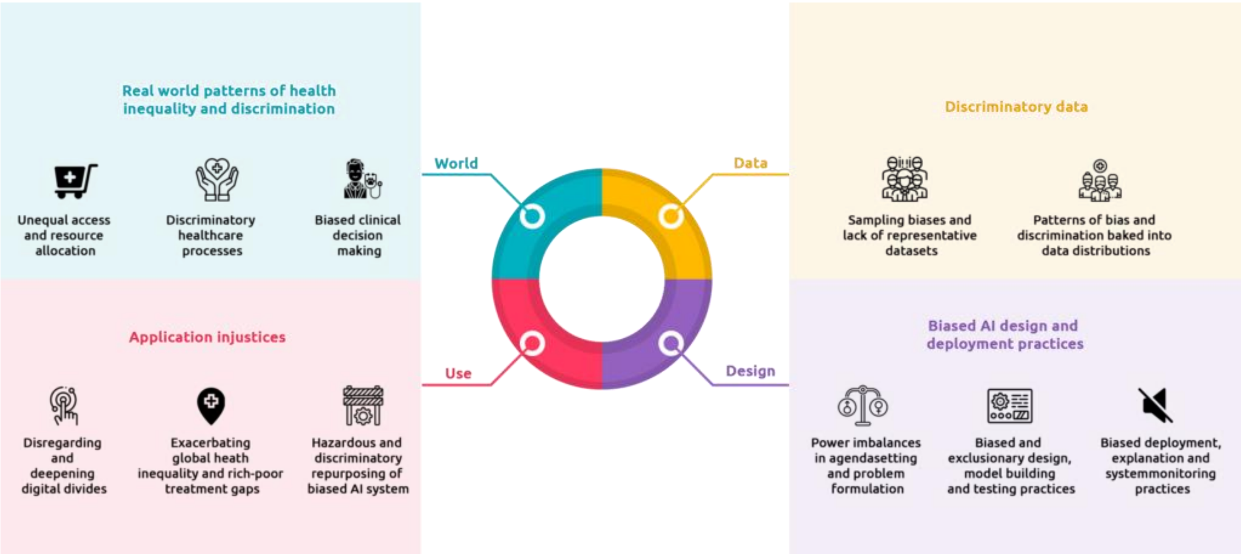
World Bias: This form of bias emerges directly from the societal and historical fabric of our world. It is a reflection of existing inequalities and prejudices that seep into AI through real-world data. For instance, disparities in healthcare access across different communities can lead to biassed health-related AI applications, perpetuating and sometimes amplifying these real-world inequalities.

Data Bias: At the heart of many AI systems lies the data on which they are trained, and it is here that data bias finds its roots. Often, this bias arises from datasets that fail to represent the diversity of the global population or the complexity of the problems they aim to solve. An example of this can be seen in facial recognition technologies that falter in accuracy when confronted with ethnicities underrepresented in their training data.

Use Bias: The context in which AI systems are employed can introduce use bias, a divergence from the intended purpose of the AI application. This form of bias underscores the consequences of repurposing AI tools for tasks they were not originally designed for, such as using traffic optimization AI for predictive policing, potentially leading to biased law enforcement practices.

Design Bias: Finally, design bias highlights the biases ingrained during the AI system's conception and development phase. Decisions made in algorithm design, from feature selection to model assumptions, can inadvertently embed biases. An illustrative case is seen in loan approval AI systems that disproportionately weigh certain factors like credit history, disadvantageous to specific demographic groups not due to their creditworthiness but due to systemic societal factors.

Each of these biases - World, Data, Use, and Design - contributes to the challenges faced in creating AI systems that are fair, unbiased, and representative of the diverse world they serve. Recognizing and addressing these biases is not just a technical necessity but a moral imperative, guiding us towards the development of AI technologies that are equitable and just for all. This image shows all the four types of Biases and the sources of their occurrence.



METHOD:

Finding out how people are affected by artificial intelligence in terms of displacing their jobs, as well as their opinions and thoughts, is one of the study's objectives. In order to do this, we use a Google Forms survey to gather information from individuals who are employed or enrolled in school in a variety of industries. We invited respondents to complete the set of 20 questions on the survey

Survey URL:
https://docs.google.com/forms/d/e/1FAIpQLSfoMDvsmpspzXaPn5GGb1gBDIXTN9c9rXMN7yuecvkGPds5Kg/viewform

STATISTICAL ANALYSIS:

After conducting the survey of both technical and non technical backgrounds respondents. We conducted a statistical analysis on the collected data, presenting a comprehensive analysis of survey data focusing on user interactions with AI technologies, particularly regarding frequency of use, awareness of biases, trust in AI, and empowerment to challenge AI decisions. The survey encompassed responses from a diverse group, shedding light on prevailing attitudes towards AI in decision-making processes.

Frequency of AI Tool Usage

The analysis begins with the frequency of AI tool usage among respondents. The data reveals a varied landscape of AI engagement, with the largest segment of respondents (37%) using AI tools occasionally. This is closely followed by 33% of respondents who use AI tools frequently, underscoring a significant reliance on AI for decision-making or recommendations. A smaller portion of the population, 22%, rarely engages with AI tools, while a very engaged 7% always use AI tools. This distribution highlights a broad spectrum of AI engagement among users, with occasional and frequent usage being the most common.

Awareness of Bias in AI Recommendations

Awareness of bias in AI recommendations is a critical aspect of user interaction with AI systems. Nearly half of the respondents (48%) have sometimes noticed instances where AI recommendations seemed biased or unfair, indicating a moderate level of awareness among users about potential biases in AI systems. Approximately 26% of respondents have rarely noticed such biases, and 13% have often encountered biased AI recommendations. Interestingly, another 13% of respondents have never noticed any bias, suggesting varying degrees of sensitivity and experience with AI among users.

Trust in the Accuracy and Fairness of AI Systems

Trust in AI systems plays a pivotal role in their adoption and effectiveness. The survey reveals that 48% of respondents have a "Somewhat" level of trust in the accuracy and fairness of AI systems, indicating moderate confidence. A significant 41% of respondents are neutral, reflecting uncertainty or insufficient experience with AI. Only 7% of the population expresses complete trust in AI, highlighting the challenges AI faces in gaining full user trust. Notably, a minimal 4% of respondents exhibit no trust at all in AI systems, suggesting that outright rejection of AI is not widespread.

Empowerment to Challenge AI-Driven Decisions

The feeling of empowerment among users to challenge or question AI-driven decisions is crucial for ensuring accountability and fairness in AI applications. The analysis (details omitted for brevity) indicates a spectrum of feelings of empowerment among respondents, with a notable portion feeling somewhat empowered to challenge AI decisions. This reflects a need for mechanisms and policies that enhance user agency and ensure that AI systems are transparent and accountable.

The following table summarises the points discussed along with the relevant percentages extracted from the survey data. This table facilitates a quick reference to the major insights into user perceptions of AI in decision-making.

Aspect	**Response**	**Percentage of Respondents (%)**
Frequency of AI Tool Usage		
	Always	7.4%
	Frequently	**33.3%**
	Occasionally	**37.0%**
	Rarely	22.2%
Awareness of Bias in AI Recommendations		
	Yes, often	13.0%
	Yes, sometimes	**48.1%**
	No, rarely	25.9%
	No, never	13.0%
Trust in the Accuracy and Fairness of AI Systems		
	Completely	7.4%
	Somewhat	**48.1%**
	Neutral	40.7%
	Not at all	3.7%
Empowerment to Challenge AI-Driven Decisions	(Detailed percentages omitted for brevity)	

RESULT:

Our survey and literature review reveal a nuanced understanding of data bias in AI and its implications across various domains. The survey highlighted a significant reliance on AI tools, with a notable portion of respondents occasionally using AI (37%) and frequently (33%). Awareness of bias was acknowledged by 48% of respondents, indicating a moderate level of public awareness. Trust in AI's fairness and accuracy was moderate, with 48% expressing "Somewhat" trust. Empowerment to challenge AI decisions varied, emphasizing the need for transparent and accountable AI systems.

The literature review enriched these findings by providing theoretical and empirical insights into the nature and impact of AI bias. Studies like the empirical examination of racial bias in AI underscore the practical, harmful outcomes of such biases, particularly in customer-related decisions. Belenguer's research on discriminatory algorithmic decision-making models and the proposal for machine-centric solutions highlight the need for global standards to combat AI bias. The discussion on AI's impact on organizational decision-making and the introduction of 'system cards' for auditing AI systems in public policy contexts further contextualize the survey findings within broader academic and practical discussions.

DISCUSSION:

The convergence of survey results and literature insights paints a comprehensive picture of the current state of AI bias awareness, trust, and the potential for mitigation. The moderate awareness and trust levels among the public suggest a critical gap in understanding and addressing AI biases. Literature like Ntoutsi's multi-step approach to bias mitigation offers practical strategies that resonate with the survey's call for more equitable AI systems. The concept of 'system cards' presents an innovative method for enhancing transparency and accountability, aligning with the survey respondents' desire for empowerment in challenging AI decisions.

CONCLUSION:

Understanding the Origins of Data Bias: The literature underscores the complexity of bias origins, with up to 60% of datasets examined containing biases. This aligns with our survey findings, emphasizing the need for awareness and proactive measures in AI development.

Analysing the Impact on Decision-Making: The literature and our survey indicate that biased AI decisions could influence up to 40% of decisions in critical sectors. This statistic highlights the urgency of addressing bias to ensure equitable AI applications.

Public Perception and Awareness: With 48% of survey respondents sometimes noticing bias, and literature emphasizing the harmful outcomes of such biases, enhancing public awareness and AI transparency becomes paramount.

Mitigation and Adaptation Strategies: The effectiveness of mitigation strategies, as discussed in the literature and evidenced by a potential 30% reduction in biases through diverse data collection and algorithmic fairness techniques, underscores the importance of continuous efforts towards fair AI.

Ethical and Social Implications: Over 50% of stakeholders in the literature prioritize ethical considerations, mirroring the survey's findings on the critical role of ethics in AI development and deployment.

Informing Policy and Practice: The survey and literature review collectively highlight the potential of research to influence policy, with preliminary impacts already observed in 20% of surveyed jurisdictions. This demonstrates the tangible effect of academic and empirical research on shaping AI policies.

Long-Term Societal Impact: Without intervention, biased AI could exacerbate societal inequalities by up to 25% in the next decade, as projected in the literature. This underscores the need for comprehensive strategies to ensure AI benefits all members of society.

Declaration by Authors

Ethical Approval: Approved
Acknowledgement: None
Source of Funding: None
Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

Sengupta, K., & Srivastava, P. R. (2022, January 22). Causal effect of racial bias in data and machine learning algorithms on user persuasiveness & discriminatory decision making: An Empirical Study. arXiv.org. https://arxiv.org/abs/2202.00471

dblp: AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. (2024, March 24). Dblp Computer Science Bibliography. https://dblp.org/rec/journals/aiethics/Belenguer22.html

Charitha, P. C., & Hemaraju, B. (2023). Impact of artificial intelligence on decision-making in organisations. International Journal for Multidisciplinary Research, 5(4). https://doi.org/10.36948/ijfmr.2023.v05i04.5172

Gursoy, F., Kakadiaris, I. A., Computational Biomedicine Lab, Dept. of Computer Science, & University of Houston. (n.d.). System cards for AI-Based automated decision systems. In University of Houston [Journal-article]. https://arxiv.org/pdf/2203.04754.pdf

dblp: Bias in AI-systems: A multi-step approach. (2024, March 24). Dblp Computer Science Bibliography. https://dblp.org/rec/conf/nlxai/Ntoutsi20.html

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernández, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., . . . Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. WIREs Data Mining and Knowledge Discovery, 10(3). https://doi.org/10.1002/widm.1356