Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2021, 8(4):23-31

Research Article ISSN: 2394 - 658X

Ethics and Bias in Data Visualization

Paraskumar Patel

CCS Global Tech
San Diego, CA, USA

ABSTRACT

This paper addresses the critical interplay between ethics and bias in data visualization within the information-rich landscape of contemporary society. As data visualization becomes an indispensable tool for elucidating complex datasets through graphical formats like charts, graphs, and maps, it concurrently faces challenges stemming from biases that can obscure the truth and mislead viewers. These biases, originating from varied sources, including data selection, collection methodologies, and the inherent preconceptions of both creators and viewers, necessitate a rigorous ethical examination. The discourse underscores the imperative for conscientious design choices to ensure data is presented accurately and ethically, thereby mitigating the potential for bias and fostering responsible visualization practices. By delving into the sources of bias, proposing strategies for their mitigation, and contemplating the ethical ramifications of visual representation, this review contributes to the broader dialogue on developing fair and transparent data visualization methodologies. Highlighting the moral obligations of data scientists and visualizers, the paper advocates for integrity, transparency, and fairness in creating and disseminating visualizations, aiming to support informed decision-making and promote an equitable society. Through a comprehensive exploration of ethical considerations, challenges, and future directions, the paper positions ethical vigilance and inclusivity as central to advancing the field of data visualization toward more accurate and just practices.

Key words: Cross-Domain Data Engineering, Data Heterogeneity, Interoperability Standards, Cross-domain Data Marketplaces, Cross-Disciplinary Collaboration, Cross-Domain Data Integration Techniques

INTRODUCTION

Data visualization has emerged as a critical tool for understanding and communicating complex datasets in the information age. Through charts, graphs, maps, and other visual formats, data visualization helps to uncover patterns, trends, and insights that might remain hidden in raw data. However, as powerful as these visual representations are, they are not immune to biases that can distort the viewer's understanding and lead to misleading conclusions. The introduction of intentional or unintentional bias raises significant ethical concerns, particularly when visualizations inform decision-making in policy, business, health, and other crucial areas.

Bias in data visualization can stem from various sources, including the choice of data, the methods of data collection and processing, the design of the visualization itself, and even the preconceptions of both the creator and the viewer. These biases can influence the interpretation of the data, potentially leading to decisions that are not based on a fair or accurate representation of the facts. Moreover, because visualizations can be so persuasive, they have the power to reinforce stereotypes, manipulate opinions, and obscure the truth.

Recognizing the potential for bias and ethical pitfalls in data visualization is essential for those who create and consume these visualizations. It prompts a critical examination of how design choices, from data selection and preparation to the final presentation, can introduce or mitigate biases. This paper explores the interplay between ethics and bias in data visualization, highlighting the importance of conscientious design choices in accurately and ethically presenting data. By examining the sources of bias, discussing strategies for mitigation, and considering the ethical implications of visual representation, this review seeks to contribute to developing more responsible and fair data visualization practices.

As we delve into the intricacies of ethics and bias in data visualization, it becomes clear that this is not merely a technical challenge but a moral one. The responsibility lies with data scientists, visualizers, and all who partake in creating and disseminating data visualizations to strive for integrity, transparency, and fairness. Through a critical review of design choices and their impacts, this paper aims to shed light on the path toward more ethical visualization practices that respect the truth and foster a more informed and equitable society.

THEORETICAL FRAMEWORK

A. Definition of Key Terms

Before diving into the complexities of ethics and bias in data visualization, it is essential to define the key terms that form the foundation of our discussion. Bias, in the context of data visualization, refers to any systematic error or unfair influence in data selection, interpretation, representation, and presentation. This can lead to partial or misleading visualizations that do not accurately reflect the underlying data or reality[1]. Ethics in data visualization concerns the moral principles that govern a person's behavior or the conducting of an activity. In this context, it pertains to the responsibility of accurately and fairly representing data, ensuring that visualizations do not mislead or harm the audience or subjects represented by the data. Data visualization itself is the graphical representation of information and data. Using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

B. Conceptualizing Bias in Data Visualization

Bias in data visualization can be categorized into several types, each affecting the integrity and interpretation of the visual representation. Cognitive Bias refers to the various ways human cognition can distort how we perceive and interpret information, including in data visualizations. This includes confirmation bias, where individuals are more likely to favor information that conforms to their pre-existing beliefs. Statistical Bias occurs when there is a systematic error in data collection, analysis, interpretation, or reporting. This type of bias can lead to inaccurate conclusions being drawn from the data. Visual Bias is introduced through the design of the visualization itself, such as the choice of colors, scales, and graphical elements, which can influence how data is perceived and understood.

C. Ethical Considerations in Data Presentation

The ethical considerations in data presentation revolve around the principles of accuracy, transparency, and fairness. Accuracy ensures that data visualizations accurately represent the underlying data and reality without distortion or manipulation. Transparency involves being open about the data sources, the data collection and analysis methodology, and any limitations or uncertainties associated with the data or the visualization. Fairness requires that data visualizations not perpetuate stereotypes, biases, or injustices. This includes being mindful of how data about different groups is represented and ensuring that visualizations do not lead to discrimination or harm.

We can better navigate the challenges and responsibilities of creating and interpreting data visualizations by grounding our understanding of these definitions and conceptual frameworks. The next section will delve deeper into the sources of bias in data visualization, examining how these biases arise and their impact on visual representations' ethical integrity.

SOURCES OF BIAS IN DATA VISUALIZATION

A. Data Collection and Selection Bias

Bias introduced during the data collection phase can significantly affect the reliability and representation of the data. This bias may occur if the methodology used to collect data inherently favors certain groups or if the response rate varies significantly across different population segments. Additionally, selection bias can arise when the data chosen for visualization fails to accurately represent the entire dataset or the phenomenon being studied. This often happens when data that supports a preconceived notion is selected over data that contradicts it, potentially leading to misleading visualizations [1].

B. Preprocessing and Cleaning Bias

The steps taken during data cleaning and preprocessing can also introduce bias. Decisions on handling missing values, for instance, whether by removal or imputation, can significantly affect the outcomes visualized. Furthermore, bias can emerge in the categorization and grouping of data. The choice of categories and the thresholds for these categories can shape the visualization's message, possibly oversimplifying complex data or reinforcing stereotypes [2].

C. Visual Encoding Choices

The choices made in visual encoding can introduce bias through color schemes, scale and axis manipulation, and the selection of graphical elements and shapes. Colors can influence perception, where certain colors might be associated with specific sentiments or ideas, affecting how data is interpreted. Manipulating the scale or axes of a graph can alter the appearance of data trends, and the choice and size of graphical elements can bias interpretation by emphasizing certain elements over others [3].

D. Narrative Framing and Contextual Bias

The narrative or story accompanying a data visualization can guide the viewer towards a particular interpretation, emphasizing certain aspects while downplaying others. Contextual bias arises from the environment in which the visualization is presented, such as in a news article, corporate report, or social media. The context can influence how the same data is interpreted, depending on the viewer's expectations or the content surrounding the visualization [4].

E. Designer and Viewer Bias

The biases of the individuals or teams creating the visualizations can influence every stage of the visualization process, from data selection to design decisions. These biases stem from personal beliefs, experiences, and cultural backgrounds. Additionally, the viewer's biases and preconceptions play a critical role in interpreting data visualizations, with individuals of different backgrounds or beliefs potentially drawing completely different conclusions from the same visualization [5].

F. Algorithm and Model Bias

Algorithm and model bias occurs when the computational methods used to analyze data inherently favor certain outcomes. This can happen, for instance, in a job applicant screening tool that utilizes a machine learning algorithm trained on historical hiring data. Suppose the historical data reflects a bias toward selecting candidates from a particular demographic (e.g., based on gender or ethnicity). In that case, the algorithm may perpetuate this bias by favoring similar candidates in future selections. As a result, visualizations of applicant screening results could misleadingly suggest that one demographic consistently outperforms another, not due to merit but because of the biased training data [6].

G. Social and Cultural Bias

Social and cultural bias in data visualization can lead to the exclusion or misrepresentation of certain groups. An example is a health information dashboard that uses symbols or colors culturally associated with gender (such as pink and blue) to represent data unrelated to gender differences. This not only reinforces gender stereotypes but could also obscure important health trends by oversimplifying the data into binary categories. Furthermore, visualizations that lack diverse cultural perspectives might misinterpret or underrepresent significant social phenomena, leading to skewed public perceptions [7].

H. Accessibility Bias

Accessibility bias in data visualization excludes individuals with disabilities from fully engaging with data. For example, a dynamic chart that relies solely on color differences to distinguish between data series may be completely inaccessible to someone with color vision deficiency. Without alternative text descriptions or patterns that differentiate the series, these individuals miss out on the visualization's insights. An accessible design, in contrast, would include labels, patterns, or an interactive feature that allows users to explore data points through touch or sound, ensuring the information is accessible to a broader audience [8].

I. Temporal Bias

Temporal bias can mislead audiences when visualizations do not adequately account for changes over time. A common example is a graph showing the dramatic growth of a tech company's revenue over five years without indicating that the industry has been growing or adjusting for inflation. Such a visualization might imply that the company's strategies were the sole reason for growth, ignoring broader economic trends or changes in purchasing power that also contributed to the increase [9].

J. Geographical Bias

Geographical bias occurs when certain locations are overemphasized or underrepresented in visualizations. A global study on internet usage primarily features data from North America and Europe, with little to no representation from Africa or Asia, which exemplifies geographical bias. This could lead viewers to assume that internet usage trends in the well-represented regions apply worldwide, ignoring significant regional differences that could affect global internet policy and development strategies.

K. Confirmation Bias in Data Interpretation

Confirmation bias in data visualization leads to the selective presentation and interpretation of data that aligns with pre-existing beliefs. An environmental organization might publish a graph showing a steep decline in polar bear populations over a decade, supporting their campaign against global warming. While the data might be accurate, failing to include other factors affecting polar bear populations, such as hunting or pollution, can lead viewers to attribute the decline solely to climate change, reinforcing the organization's stance without presenting a complete picture.

L. Economic and Political Bias

Economic and political biases are introduced when data visualizations are used to support specific agendas. For instance, a visualization showing the economic growth of a country during a particular administration might omit data from previous years that would reveal a consistent growth trend starting well before the current administration took office. This selective presentation could mislead viewers into attributing economic success solely to the current government's policies, ignoring the contributions of past administrations or global economic trends [10].

Each of these biases illustrates the importance of a critical approach to data visualization, emphasizing the need for transparency, diversity, and inclusivity in every step of the visualization process to ensure that insights derived from data are as unbiased and accurate as possible.

MITIGATING BIAS THROUGH DESIGN CHOICES

Addressing and mitigating bias in data visualization ensures the accuracy and ethical integrity of visual representations and brings a host of benefits that enhance the value and impact of visualizations. By adopting conscientious design choices and adhering to foundational principles, visualizers can significantly reduce the potential for bias, fostering trust, inclusivity, and more informed decision-making. Below, we explore the benefits associated with each mitigation strategy outlined in the principles of ethical visualization and the advantages gained through case studies and tools for bias reduction.

A. Principles for Ethical Visualization

Transparency in data visualization fosters trust among the audience by openly sharing data sources, methodologies, and any adjustments made during the visualization process [11]. This openness allows viewers to understand the context and limitations of the data, encouraging critical engagement and enabling independent verification of findings. Transparency also promotes accountability among designers, compelling them to adhere to high data integrity standards and ethical representation.

Accuracy ensures visualizations truthfully represent the underlying data, fostering credibility and reliability [12]. By choosing appropriate scales, avoiding misleading graphical elements, and accurately representing statistical uncertainties, visualizers can convey complex data in an understandable and trustworthy manner. This accuracy supports informed decision-making by giving stakeholders a clear and correct understanding of the data.

Inclusivity broadens the reach and relevance of visualizations by designing with diverse audiences in mind. This approach ensures that visualizations are accessible to people with different abilities, cultural backgrounds, and perspectives [13]. The benefit of inclusivity is twofold: it enhances the accessibility of information, ensuring that a wider audience can derive value from the data, and it promotes equity by actively working to avoid reinforcing stereotypes or excluding certain groups.

Fair Representation commits to impartially representing all data points, ensuring visualizations do not unduly emphasize certain values or outcomes [14]. This balanced approach prevents the skewing of perceptions and supports a more objective understanding of the data. Fair representation allows viewers to draw conclusions based on a comprehensive and unbiased presentation of facts, leading to more equitable and informed analyses.

B. Case Studies of Ethical Visualization

Case studies, such as Climate Change Data Visualization and Public Health Dashboards during COVID-19, demonstrate the practical application of ethical visualization principles and highlight their benefits. These examples show how accurate and non-sensationalist representations can effectively communicate complex and critical information, fostering a deeper understanding and engagement with pressing issues. They illustrate the power of ethical visualization to inform public discourse, guide policy decisions, and encourage individual action based on a solid foundation of trust and clarity.

C. Tools and Techniques for Bias Reduction

Data Auditing helps identify potential biases in the data collection and processing stages, ensuring that the data used for visualization is as unbiased and representative as possible [15]. This process enhances the integrity of the data and, by extension, the visualizations created from it, leading to more accurate and credible insights.

User Testing with diverse groups uncovers unintended biases in design and allows for their correction, ensuring that the final visualization communicates clearly to a broad audience [16]. This inclusivity not only extends the reach of the visualization but also enriches the perspective from which the data is interpreted, leading to more nuanced and comprehensive understandings.

Standardization of Practices through guidelines and checklists ensures consistent evaluation for bias and adherence to ethical standards across visualizations [17]. This standardization promotes consistency and reliability in data visualization practices, enhancing visualizations' trustworthiness as communication and decision-making tools.

Educational resources raise awareness among designers about the importance of mitigating bias and equipping them with the necessary tools [18]. By expanding the knowledge and skills of visualizers, these resources contribute to improving visualization practices and fostering a culture of ethical visualization that prioritizes accurate, fair, and impactful representations of data.

In sum, the benefits of mitigating bias through conscientious design choices in data visualization extend far beyond individual visualizations' accuracy and ethical integrity. They foster a culture of trust, inclusivity, and informed decision-making that can significantly impact how data is understood and acted upon across various domains.

D. Implementation of Mitigation Strategies

Detailing the specific steps for implementing each mitigation strategy can offer practical guidance for visualizers. For example, for transparency, this could include a checklist of information that should be disclosed (e.g., data sources, collection methodologies, adjustments made during analysis). For inclusivity, guidelines on conducting accessibility audits and incorporating universal design principles would be beneficial. Providing concrete examples of how to apply these strategies in the context of actual visualization projects can help practitioners understand how to operationalize these principles in their work.

E. Continuous Feedback and Adaptation

Emphasizing the importance of iterative design and continuous feedback underscores that mitigating bias is not a one-time task but an ongoing process. Incorporating user feedback mechanisms, such as surveys or user testing sessions with diverse groups, can reveal biases and areas for improvement that were not initially apparent. Discussing methods for incorporating this feedback into the visualization design process can help designers make their visualizations more adaptive and responsive to the needs and perceptions of a broad audience.

F. Role of a Diverse Team

Highlighting the benefits of having a diverse team involved in the visualization process can underscore the value of different perspectives in identifying and mitigating biases. A team with members from various backgrounds (cultural, disciplinary, and experiential) can bring a wider range of insights into potential biases and how they might affect the interpretation of data. Discussing strategies for fostering diversity and inclusion within visualization teams and challenges that might arise and how to address them can provide valuable insights for organizations and individuals committed to ethical visualization practices.

G. Advanced Tools and Technologies

Exploring advanced tools and technologies for bias detection and mitigation can provide visualizers with resources to tackle bias more effectively. This might include machine learning algorithms to identify dataset bias, software tools that automatically check for accessibility issues, or platforms that facilitate collaborative design and feedback. Providing information on these tools and case studies of their successful application can equip visualizers with the latest advancements in the field.

CHALLENGES AND LIMITATIONS

While the strategies for mitigating bias in data visualization are crucial, they are not without challenges and limitations. The process of creating visual representations of data that are both accurate and ethical is complex, and various factors can complicate these efforts. This section outlines some primary challenges and limitations of pursuing unbiased and ethical data visualization.

A. Inherent Limitations of Data Visualization

One of the inherent challenges in data visualization is the balance between complexity and clarity. Simplifying complex data sets for visualization can sometimes lead to losing nuance or important context, potentially introducing bias. Even with the best intentions, the subjectivity inherent in design choices can introduce bias. The designer's perspective influences decisions about color, layout, and graphical elements, potentially affecting the visualization's neutrality. Additionally, data quality and availability can limit the effectiveness of bias mitigation strategies. Incomplete, inaccurate, or inaccessible data sets can skew visualizations, inadvertently introducing bias [19].

B. Balancing Aesthetics and Accuracy

Designers often face the challenge of making visualizations visually appealing to engage their audience while ensuring they accurately represent statistical truths. This balancing act can sometimes lead to choices prioritizing aesthetics over accuracy, potentially misleading viewers. There's a tendency to oversimplify information to make data more accessible. Still, while simplification is necessary for clarity, it can also strip away critical details that affect the overall interpretation of the data [20].

C. The Role of Viewer Interpretation

Viewers bring their own biases to the interpretation of visual data. These pre-existing biases can influence how data is understood, regardless of how ethically it's presented, complicating the goal of impartiality in data communication. The interpretation of visual elements can also vary significantly across different cultures and contexts, leading to misunderstandings.

D. Technological and Resource Constraints

The tools and software available for data visualization come with limitations and default settings, which can introduce bias. For example, default color schemes or graph types may not be suitable for all data types. Small organizations or individuals may lack the resources (time, budget, expertise) to thoroughly audit their data and visualizations for bias, leading to unintended ethical oversights [21].

E. Addressing Challenges with Advanced Mitigation Strategies

Incorporating advanced mitigation strategies, such as leveraging diverse teams for design, iterative user testing, and employing tools for algorithmic transparency, can offer pathways to overcome some of these challenges. A diverse team can provide varied perspectives that help identify and address potential biases early in the design process. Iterative design and user testing with diverse groups can uncover and rectify biases and interpretation issues, enhancing the inclusivity and accessibility of visualizations. Advanced tools and technologies for bias detection and mitigation can assist in navigating the limitations of data and visualization software, facilitating more ethical representations [22].

The challenges and limitations in creating unbiased and ethical data visualizations are significant but not insurmountable. Awareness of these issues is the first step toward addressing them. By acknowledging the complexity of data representation, the subjectivity of design choices, and the diversity of viewer interpretation, designers can approach visualization with a more critical and careful mindset. Furthermore, advancements in visualization tools and methodologies and a commitment to education and best practices can help mitigate these challenges, moving the field toward more accurate, ethical, and effective visual communication.

FUTURE DIRECTIONS

The field of data visualization is dynamic and multifaceted, continuously adapting to meet new challenges, leverage advancing technologies, and fulfill societal needs. Addressing biases and ethical considerations necessitates diligence, sustained effort, and innovative and forward-thinking approaches. This discussion proposes potential future directions to bolster data visualization practices' integrity, inclusivity, and impact.

A. Advancements in Visualization Technologies

Integrating Artificial Intelligence (AI) and Machine Learning into data visualization tools presents advanced capabilities for identifying and reducing bias. These technologies could facilitate the automatic detection of biased data patterns and propose more equitable data representation methods. Additionally, the trend toward more Interactive and Dynamic Visualizations enables users to engage with data in real time, offering a democratized exploration experience that can help users discover and understand biases independently. Moreover, Virtual and Augmented Reality (VR and AR) technologies promise to transform data visualization through immersive experiences, providing novel ways to comprehend complex datasets and introducing unique ethical representation challenges.

B. Emphasis on Education and Training

Developing and integrating curricula focusing on ethical data visualization practices into educational programs can equip future data scientists and visualizers with the tools necessary to prioritize bias mitigation and ethical considerations. Further, ongoing Professional Development opportunities, including workshops, seminars, and online courses, can keep current practitioners up-to-date on the latest tools, techniques, and ethical standards in data visualization.

C. Promotion of Diversity and Inclusion

Fostering Diverse Teams in the design and creation of data visualizations can introduce multiple perspectives into the process, diminishing the influence of unconscious biases on the final products. Promoting Inclusive Design Practices that consider the needs and perspectives of a broad user base ensures that data visualizations effectively and ethically serve diverse audiences.

D. Standardization and Best Practices

Establishing Industry-wide Standards for ethical data visualization can offer practitioners clear guidelines, promoting consistency and integrity in visual representations. Additionally, articulating and disseminating Best Practices for Bias Mitigation provides designers with the necessary knowledge and tools to produce equitable and accurate visualizations.

E. Research and Collaboration

Encouraging Interdisciplinary Research that amalgamates insights from data science, psychology, sociology, and design can deepen understanding of bias in data visualization and strategies for its counteraction. Furthermore, Collaborative Projects that unite organizations, researchers, and practitioners across different fields can drive innovation in ethical visualization practices, leading to the development of novel methods and technologies for bias mitigation [23].

Both challenges and opportunities mark the trajectory of data visualization. As the field progresses, the dedication to tackling bias and ensuring ethical representation becomes more critical. Through technological innovations, education, and training, efforts towards diversity and inclusion, standardization, and collaborative research, the data visualization community is well-positioned to advance, enhancing its capacity to present complex information accurately and equitably. By embracing these future directions, practitioners will contribute to a more informed and just society where data visualizations empower rather than mislead, ensuring a more equitable future for all.

CONCLUSION

In summary, this article has navigated through the multifaceted realm of ethics and bias in data visualization, shedding light on both the perils and potentials inherent in the visual representation of data. Through thoroughly examining the sources of bias—from data collection and preprocessing to visual encoding and narrative framing—this discussion underscores the profound impact that design choices can have on the accuracy, integrity, and ethicality of data visualizations. Moreover, it emphasizes the pivotal role of transparency, accuracy, inclusivity, and fair representation in mitigating these biases and fostering ethical visualization practices.

The exploration of challenges and limitations highlights the inherent complexities and subjective elements in data visualization, acknowledging that while the quest for unbiased and ethical visual representations is daunting, it is not unattainable. The article posits that awareness, critical scrutiny, and a commitment to ethical principles can significantly advance the pursuit of integrity in data visualization. Additionally, it outlines future directions, including advancements in visualization technologies, the importance of education and training, the promotion of diversity and inclusion, the establishment of industry-wide standards, and the value of interdisciplinary research and collaboration.

As we stand on the brink of new technological frontiers and societal demands, the field of data visualization is poised for transformative growth. The ethical challenges it faces are not mere obstacles but opportunities for innovation, learning, and improvement. By embracing a forward-thinking approach that prioritizes the ethical representation of data, the data visualization community can ensure that its practices convey information effectively and contribute to a more informed, just, and equitable society.

The journey towards ethical data visualization is ongoing and demands the collective effort of designers, researchers, practitioners, and consumers alike. As we forge ahead, let us carry the insights and lessons gleaned from this discussion, committed to enhancing our visual representations' clarity, integrity, and impact. In doing so, we uphold the principles of fairness and transparency, ensuring that data visualizations serve as beacons of truth and understanding in an increasingly data-driven world.

REFERENCES

[1]. D. P. Dee, “Bias and data assimilation,” Quarterly Journal of the Royal Meteorological Society, vol. 131, no. 613, pp. 3323–3343, Oct. 2005, doi: 10.1256/QJ.05.137.

[2]. A. Althubaiti, “Information bias in health research: definition, pitfalls, and adjustment methods,” J Multidiscip Healthc, vol. 9, pp. 211–217, May 2016, doi: 10.2147/JMDH.S104807.

[3]. A. Das, S. Anjum, and D. Gurari, “Dataset bias: A case study for visual question answering,” Proceedings of the Association for Information Science and Technology, vol. 56, no. 1, pp. 58–67, Jan. 2019, doi: 10.1002/PRA2.7.

[4]. E. Wall, L. M. Blaha, C. L. Paul, K. Cook, and A. Endert, “Four Perspectives on Human Bias in Visual Analytics,” Cognitive Biases in Visualizations, pp. 29–42, Jan. 2018, doi: 10.1007/978-3-319-95831-6_3/COVER.

[5]. M. Pohl, "Cognitive Biases in Visual Analytics—A Critical Reflection," Cognitive Biases in Visualizations, pp. 177–184, Jan. 2018, doi: 10.1007/978-3-319-95831-6_13/FIGURES/1.

[6]. A. Khosla, T. Zhou, T. Malisiewicz, A. A. Efros, and A. Torralba, "Undoing the damage of dataset bias," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 7572 LNCS, no. PART 1, pp. 158–171, 2012, doi: 10.1007/978-3-642-33718-5_12/COVER.

[7]. T. Tommasi, N. Patricia, B. Caputo, and T. Tuytelaars, "A deeper look at dataset bias," Advances in Computer Vision and Pattern Recognition, no. 9783319583464, pp. 37–55, 2017, doi: 10.1007/978-3-319-58347-1_2/FIGURES/11.

[8]. C. J. Pannucci and E. G. Wilkins, "Identifying and avoiding bias in research," Plast Reconstr Surg, vol. 126, no. 2, pp. 619–625, Aug. 2010, doi: 10.1097/PRS.0B013E3181DE24BC.

[9]. T. L. Lash and T. P. Ahern, "Bias analysis to guide new data collection," International Journal of Biostatistics, vol. 8, no. 2, pp. 1–23, Jan. 2012, doi: 10.2202/1557-4679.1345/MACHINEREADABLECITATION/RIS.

[10]. G. T. Sica, "Bias in Research Studies1," https://doi.org/10.1148/radiol.2383041109, vol. 238, no. 3, pp. 780–789, Mar. 2006, doi: 10.1148/RADIOL.2383041109.

[11]. E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri, "Mitigating the Attraction Effect with Visualizations," IEEE Trans Vis Comput Graph, vol. 25, no. 1, pp. 850–860, Jan. 2019, doi: 10.1109/TVCG.2018.2865233.

[12]. Z. Wang et al., "Towards fairness in visual recognition: Effective strategies for bias mitigation," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 8916–8925, 2020, doi: 10.1109/CVPR42600.2020.00894.

[13]. M. Reynolds, "Bias test to keep algorithms ethical," New Sci (1956), vol. 234, no. 3119, p. 10, Apr. 2017, doi: 10.1016/S0262-4079(17)30603-6.

[14]. K. Hepworth, K. Hepworth, and C. Church, "Racism in the Machine: Visualization Ethics in Digital Humanities Projects," Digital Humanities Quarterly, vol. 012, no. 4, Feb. 2019.

[15]. A. Yapo and J. Weiss, "Ethical Implications of Bias in Machine Learning," Proceedings of the Annual Hawaii International Conference on System Sciences, vol. 2018-January, pp. 5365–5372, Jan. 2018, doi: 10.24251/HICSS.2018.668.

[16]. L. Dixon Jigsaw et al., "Measuring and Mitigating Unintended Bias in Text Classification," AIES 2018 - Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society, vol. 7, no. 18, pp. 67–73, Dec. 2018, doi: 10.1145/3278721.3278729.

[17]. E. M. Bender and B. Friedman, "Data Statements for Natural Language Processing: Toward Mitigating System Bias and Enabling Better Science," Trans Assoc Comput Linguist, vol. 6, pp. 587–604, Jan. 2018, doi: 10.1162/TACL_A_00041.

[18]. S. Barocas and A. D. Selbst, "Big Data's Disparate Impact," SSRN Electronic Journal, Mar. 2016, doi: 10.2139/SSRN.2477899.

[19]. E. Dimara, G. Bailly, A. Bezerianos, and S. Franconeri, "Mitigating the Attraction Effect with Visualizations," IEEE Trans Vis Comput Graph, vol. 25, no. 1, pp. 850–860, Jan. 2019, doi: 10.1109/TVCG.2018.2865233.

[20]. E. Wall, J. Stasko, and A. Endert, "Toward a Design Space for Mitigating Cognitive Bias in Vis," 2019 IEEE Visualization Conference, VIS 2019, pp. 111–115, Oct. 2019, doi: 10.1109/VISUAL.2019.8933611.

[21]. Y. Ahn and Y. R. Lin, "Fairsight: Visual analytics for fairness in decision making," IEEE Trans Vis Comput Graph, vol. 26, no. 1, pp. 1086–1095, Jan. 2020, doi: 10.1109/TVCG.2019.2934262.

[22]. Z. Wang et al., "Towards fairness in visual recognition: Effective strategies for bias mitigation," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 8916–8925, 2020, doi: 10.1109/CVPR42600.2020.00894.

[23]. E. Wall, L. M. Blaha, L. Franklin, and A. Endert, "Warning, Bias May Occur: A Proposed Approach to Detecting Cognitive Bias in Interactive Visual Analytics," 2017 IEEE Conference on Visual Analytics Science and Technology, VAST 2017 - Proceedings, pp. 104–115, Jul. 2017, doi: 10.1109/VAST.2017.8585669.