

Advancing critical data science and statistical literacy for democratic and civic education

Olushina Olawale Awe, Joachim Engel and Florian Weber-Stein
Ludwigsburg University, Germany
olushina.awe@ph-ludwigsburg.de

In an increasingly data-driven world, the integration of critical data literacy and statistical reasoning into civic education has emerged as a vital component of democratic engagement. The ability to interpret, question, and ethically engage with data empowers individuals to participate meaningfully in civic life, challenge misinformation, and advocate for equitable policies. This paper examines the intersection of data literacy, statistical reasoning, and civic education, emphasizing their role in fostering informed and socially responsible citizens. Drawing from global scholarships, pedagogical innovations, and case studies, we argue that a critical approach to data education enhances democratic resilience by promoting analytical thinking, ethical awareness, and social justice. By reimagining curriculum design and instructional strategies, educators can equip learners with the competencies needed to navigate a data-saturated society, promote equity, and contribute to democratic decision-making.

INTRODUCTION

The 21st century has witnessed an unprecedented surge in data generation and utilization, influencing everything from public policy to personal decision-making (Gal et al., 2023). Governments, corporations, and media institutions increasingly rely on data analytics to shape narratives, allocate resources, and justify decisions (Rak & Rezmer, 2022). In this context, traditional civic education, which has historically emphasized constitutional knowledge, voting procedures, and civic duties, must evolve to incorporate critical data literacy (the ability to analyze, interpret, and ethically engage with data) and statistical literacy (the capacity to reason about variability, bias, and uncertainty in quantitative information). The rise of misinformation, algorithmic bias, and data-driven governance underscores the urgency of embedding these literacies into education. Citizens today must be able to evaluate election statistics, public health data, economic indicators, and social media metrics, making data literacy a cornerstone of democratic participation. For instance, during the COVID-19 pandemic, individuals were bombarded with conflicting infection rates, vaccine efficacy studies, and policy justifications, highlighting how data illiteracy can lead to public confusion and mistrust (Krishnannair & Krishnannair, 2022).

The advent of big data has transformed the landscape of democracy and civic engagement. Large quantitative models, ranging from election forecasting algorithms and public opinion analyses to predictive models for resource allocation and social behavior, now significantly influence policy decisions, shape public discourse, and profoundly impact citizens' lives (Alscher et al., 2022). While these models offer unprecedented opportunities for evidence-based governance and participatory decision-making, their inherent complexity often creates a gap between data producers and the broader public. This gap risks undermining transparency, accountability, and citizen trust in democratic institutions.

Addressing this challenge requires a fundamental rethinking of how large-scale quantitative models are taught within civic and educational contexts. Teaching these models not only involves technical instruction in statistical reasoning, data interpretation, and modeling techniques but also demands critical reflection on ethical considerations, power dynamics, and the societal implications of data-driven governance. An interdisciplinary approach, like blending data science, statistics, political science, and ethics, is crucial for equipping citizens with the tools to critically engage with, question, and influence the models that increasingly shape public life. Recent scholarship has increasingly emphasized the intersection of quantitative literacy, civic education, and democratic engagement, reflecting the growing importance of data-driven decision-making in public life. Researchers such as Gal (2023), Engel (2017) and Radermacher (2022) have stressed the central role of statistical and data literacy in preparing citizens for informed participation in modern democracies. Quantitative literacy is now recognized not merely as a technical competence but as a critical civic skill essential for interpreting policies, voting decisions, and media narratives shaped by data. In parallel, civic education theorists

have highlighted the need to integrate contemporary technological and data-centric realities into curricula to ensure that citizens are equipped to navigate the complexities of governance in the digital age (Engel, 2017). Moreover, studies in data science education (Engel & Martignon, 2022; Baumer, 2015; Awe et al., 2024) advocate for pedagogical approaches that move beyond formulaic instruction toward interdisciplinary, project-based, and ethics-centered learning. These approaches recognize that understanding how models are constructed, whose interests they serve, and what limitations they carry is as important as mastering computational techniques. Meanwhile, critical data studies scholars (Weiland, 2017; D'Ignazio & Klein, 2020) have drawn attention to the biases and inequities embedded within big data systems, emphasizing the ethical imperatives in teaching about algorithms and models. Collectively, these literatures underscore the urgent need for educational frameworks that bridge technical knowledge, critical reflection, and democratic engagement, forming the foundation for the analysis and proposals presented in this paper.

The growing influence of large quantitative models in democratic governance marks a significant shift in how public decisions are made and justified. These models, encompassing election forecasting algorithms, predictive analytics for public service delivery, and behavioral models based on social media data, have the potential to improve the efficiency, transparency, and responsiveness of political systems (Engel & Weber-Stein, 2023; Weber-Stein & Engel, 2025). For instance, election forecasting platforms such as FiveThirtyEight have not only transformed how elections are analyzed but also how voters understand probabilities and outcomes (Ghosh et al., 2025). Similarly, predictive models are increasingly used to allocate public resources, optimize urban planning, and anticipate public health and development needs (Awe et al., 2022). By integrating massive datasets and sophisticated computational techniques, such models offer decision-makers new tools for evidence-based policymaking. However, the deployment of these models also introduces critical challenges. Their complexity often renders them opaque to the general public, creating what scholars refer to as "black box" systems, where model inputs, assumptions, and decision rules are not easily interpretable by non-specialists.

Additionally, biases embedded within the data or model design can perpetuate existing inequalities or produce unintended harms. Without mechanisms for transparency, accountability, and public scrutiny, large quantitative models risk distancing citizens from the democratic processes they are intended to enhance. Citizens who cannot scrutinize or question the models that influence public policies are less equipped to hold institutions accountable for their outcomes. Moreover, when models are accepted uncritically, there is a heightened risk of technocratic governance, where decisions appear objective and inevitable but are, in reality, shaped by contestable assumptions and value-laden judgments. Misinformation, mistrust, and alienation can thrive in such environments, weakening the social contract that underpins democratic societies. Addressing these challenges requires not only technical education but also the cultivation of critical data literacy skills that empower citizens to engage thoughtfully and ethically with complex quantitative systems. This paper aims to explore how the integration of critical data science and statistical literacy into civic education can strengthen democratic engagement, promote social justice, and counter data-driven disinformation and exclusion.

CONCEPTUAL FRAMEWORK: OVERVIEW OF CRITICAL DATA AND STATISTICAL LITERACY

What is Critical Data Literacy?

Traditional data literacy focuses on technical skills like collecting, analyzing, and visualizing data. However, critical data literacy expands this definition by incorporating the following:

Questioning data origins: Who collected the data, and for what purpose? - Recognizing bias and power structures: How does data reflect or reinforce societal inequalities?

Ethical engagement: What are the implications of data use on privacy, surveillance, and human rights? This approach aligns with critical pedagogy, encouraging learners to interrogate data as a social construct rather than an objective truth. For example, crime statistics may reflect policing biases rather than actual crime rates, requiring citizens to assess underlying methodologies critically (Gal, 2023).

The Role of Statistical Literacy in Civic Life

Statistical literacy is no longer merely a technical skill confined to classrooms or specialized professions; it is an essential component of active citizenship. It extends beyond basic mathematical

computations to include a good understanding of uncertainty, variability, and methodological rigor. Citizens must not only be able to perform calculations but also critically interpret margins of error, recognize the implications of sample size and sampling techniques, and detect common visual distortions in graphical representations, such as truncated axes or manipulated scales. For example, when interpreting a public opinion poll, it is crucial to consider the margin of error and the representativeness of the sample before concluding. Similarly, graphical data representations can subtly mislead audiences if they are not properly scrutinized. When applied to civic contexts, statistical literacy enables individuals to critically assess the veracity and relevance of policy claims, media reports, and scientific findings. It equips citizens to question not only "what" is being presented but also "how" and "why" particular data points are highlighted or omitted. For instance, a politician might announce a "50% increase in employment," yet without information about the baseline figures, the timeframe over which the increase occurred, or the sectors contributing to this rise, this statistic could create a deceptive impression of economic prosperity. In a democracy increasingly driven by complex data narratives, statistical literacy emerges as a critical defense against manipulation and a powerful tool for informed decision-making.

The intersection with civic education

Civic education traditionally focuses on preparing individuals for responsible participation in democratic processes by imparting knowledge of political systems, rights, and responsibilities. However, the complexity of the modern information landscape demands a broader educational mandate, one that integrates statistical and data literacy into the very fabric of civic education. Through the lens of data, citizens learn to ground their public reasoning in evidence rather than rhetoric, fostering a culture of critical inquiry and debate that is vital to the health of democratic societies.

Integrating data and statistical literacy into civic education promotes evidence-based reasoning in public discourse, enhances participatory democracy by enabling citizens to advocate effectively through data-driven arguments, and combats the growing threat of disinformation by instilling critical thinking and verification habits. This intersection forms the foundation for a data-empowered democracy, where citizens are not passive recipients of information but active interpreters, analysts, and even producers of data narratives. Students trained in this hybrid framework are better equipped to evaluate statistics cited in political debates critically, assess the validity of claims made in media outlets, and advocate for systemic reforms using data-driven arguments, thereby strengthening the democratic fabric from the ground up.

THE DEMOCRATIC IMPERATIVE FOR CRITICAL DATA LITERACY

Data as a Tool for Civic Engagement

In the digital age, data has become a central instrument of governance, activism, and public policy. From election analytics and health statistics to environmental monitoring and urban planning, data informs virtually every aspect of public life. Citizens equipped with critical data skills are better positioned to evaluate the transparency and effectiveness of governmental initiatives, assess how public funds are allocated, and advocate for policy changes based on empirical evidence.

For example, citizens who can analyze municipal budget data may uncover inequities in how resources are distributed across neighborhoods, prompting demands for more equitable urban planning. In another instance, participatory platforms such as open-source environmental monitoring projects empower individuals to track pollution levels and contribute to shaping local environmental policies. Data-driven activism, whether through crowd-sourced mapping of public safety issues, documenting instances of voter suppression, or using social media analytics for community organizing, demonstrates how data can become a formidable tool for fostering accountability, transparency, and social change.

Risks of Data Illiteracy in Democracy

The converse, however, presents a significant threat to democratic integrity. Without critical competencies in data literacy, individuals and communities are vulnerable to sophisticated forms of manipulation and exclusion. Misleading statistics, when wielded uncritically by politicians, corporations, or media outlets, can distort public perception and polarize societies. The politicization of

pandemic data, for instance, illustrates how selective presentation or suppression of data can fuel misinformation and undermine public trust.

Moreover, algorithmic systems increasingly influence access to opportunities, justice, and services. Biased algorithms in hiring practices or predictive policing systems can reinforce systemic discrimination if their outputs go unchallenged by an informed citizenry. A lack of data literacy not only limits participation in debates around these technologies but also perpetuates a cycle of exclusion, where only a privileged few possess the skills necessary to interpret, critique, and influence the data architectures that govern modern life. Case studies from recent elections, public health responses, and tech sector controversies vividly illustrate the real-world consequences of widespread data illiteracy, highlighting an urgent need for critical citizenship and educational interventions (Martínez-Castro et al., 2023; Thelma et al., 2024).

Data Justice and Equity

Addressing the democratic imperative for critical data literacy also requires confronting the structural inequities embedded within data systems themselves. Marginalized communities often suffer from data invisibility, underrepresentation, or exploitation. Incomplete or biased datasets can lead to policies that ignore or harm vulnerable populations, while aggressive surveillance programs often disproportionately target minority communities, exacerbating privacy violations and civil rights abuses.

Furthermore, systemic inequalities in access to digital technologies and education, commonly referred to as the digital divide, prevent many from fully participating in a data-driven society. Ensuring equitable participation requires embedding data justice principles into educational frameworks, including advocating for transparent, inclusive, and community-centered data practices. Critical data literacy education must therefore go beyond technical proficiency to cultivate an ethical and political consciousness, empowering students to recognize, resist, and reform the injustices perpetuated by unequal data systems (Watson, 2003).

PEDAGOGICAL INNOVATIONS FOR DATA-INFUSED CIVIC EDUCATION

This paper proposes the following pedagogical strategy for merging data literacy and civic education.

Project-Based Learning (PBL) with Civic Data

One promising pedagogical strategy for merging data literacy and civic education is project-based learning (PBL) focused on authentic civic issues. For instance, students might analyze local housing data to investigate patterns of gentrification, identifying correlations between rising rent prices and demographic shifts. By developing evidence-based proposals for equitable urban policies, students apply statistical techniques in meaningful contexts, thus linking academic skills to real-world civic challenges. Project-based learning not only enhances statistical reasoning but also fosters critical thinking, collaboration, and a sense of agency (Awe et al., 2024).

Inquiry-Driven Data Exploration

Inquiry-driven approaches further enhance engagement by centering students' questions as the starting point for data analysis. Investigating school funding disparities using publicly available datasets, for example, invites learners to grapple with issues of equity, governance, and public accountability. In posing questions, seeking out relevant data, and drawing conclusions, students develop both their technical and civic capacities, preparing them for active participation in a democratic society.

Interdisciplinary Approaches

True data literacy cannot be siloed within mathematics or statistics classrooms. Interdisciplinary approaches that integrate mathematics, social studies, environmental science, and ethics provide a richer understanding of how data interacts with complex social issues. For example, a project examining climate change could combine the scientific analysis of temperature trends, the ethical implications of environmental justice, and the policy challenges of international cooperation. Such holistic frameworks help students appreciate the multi-dimensional nature of civic problems and the necessity of integrating quantitative and qualitative reasoning (Awe et al., 2022).

Ethics and Critical Media Literacy

Given the prevalence of misinformation and algorithmic bias in today's media ecosystem, it is essential to embed ethics and critical media literacy within data education. Students should be taught not only to analyze datasets but also to interrogate the sources, motivations, and potential biases underlying data production and dissemination. Critical discussions around social media algorithms, government data transparency, digital consent, and privacy rights prepare students to navigate the ethical dilemmas of the information age with discernment and integrity (Rak & Rezmer, 2022).

CONCLUSION

In conclusion, the integration of data literacy into civic education is essential for cultivating an informed and engaged citizenry in the modern democratic landscape. As data increasingly drives public policy, societal narratives, and global discussions, the ability to critically evaluate and engage with data is no longer a luxury but a necessity. Statistical literacy enables individuals to not only understand numerical information but also to recognize its limitations, detect potential biases, and question the validity of the data presented. This skill set empowers citizens to critically assess policy claims, media reports, and scientific studies, ensuring that they are not passive consumers of information but active, informed participants in the democratic process. Moreover, the intersection of data literacy with civic education strengthens public discourse by promoting evidence-based reasoning, fostering participatory democracy, and combating the growing threat of disinformation. By incorporating data literacy into civic education, we enable individuals to engage with complex issues, such as budgeting, urban planning, and public health, using data-driven arguments. This is crucial for holding institutions accountable, ensuring transparency, and advocating for systemic reforms that promote equity and justice. For instance, when citizens are equipped with the ability to analyze municipal budget data or environmental statistics, they can advocate for fairer resource allocation and more effective policies. However, the risks of data illiteracy are profound. Without the necessary competencies, individuals are vulnerable to manipulation through misleading statistics, biased algorithms, and selective data reporting. The politicization of data, particularly in sensitive areas like public health and election campaigns, underscores the importance of critical data literacy in protecting democratic values. Furthermore, as algorithmic systems increasingly govern hiring practices, law enforcement, and social services, the lack of data literacy can perpetuate inequalities and exclusion, leaving marginalized communities without the tools to challenge unfair practices.

To address these challenges, the recommendations for future action are clear. First, educational policies must prioritize the integration of data literacy into curricula from K-12 to higher education, ensuring that all students develop the critical thinking skills necessary to navigate a data-driven world. This involves not only technical proficiency in data analysis but also a justice-oriented approach that considers the ethical implications of data use. Teacher training should be reimagined to equip educators with the knowledge and tools to teach data literacy effectively across subjects. Additionally, community-based initiatives, such as data literacy workshops, seminars in libraries, and civic tech programs, must be expanded to bridge the digital divide and ensure equitable access to data tools and knowledge. As we move forward, it is essential to foster a generation of citizens who understand that data is not just a commodity for analysis but a vital resource for social justice, transparency, and accountability. By embedding data literacy into education and society at large, we can create a more informed, equitable, and participatory democracy for the future. The path forward requires collaboration among educators, policymakers, and communities to ensure that future generations are not only data-literate but also data-empowered, capable of harnessing data to challenge injustice and strengthen democratic ideals.

ACKNOWLEDGMENT

OOA acknowledges the Alexander von Humboldt Foundation for financial support.

REFERENCES

- Alscher, P., Ludewig, U., & McElvany, N. (2022). Civic Education, Teaching Quality and Students' Willingness to Participate in Political and Civic Life: Political Interest and Knowledge as Mediators. *Journal of Youth and Adolescence*, 51(10), 1886–1900. <https://doi.org/10.1007/s10964-022-01639-9>
- Awe, O. O., Love, K., & Vance, E. A. (2022). Promoting Statistical Practice and Collaboration in Developing Countries. Taylor & Francis. <https://doi.org/10.1201/9781003261148>
- Awe, O. O., Love, K., & Vance, E. A. (2024). Teaching Data Science in Africa via Online Team-Based Learning. In O. O. Awe & E. A. Vance (Eds.), *Sustainable Statistical and Data Science Methods and Practices: Reports from LISA 2020 Global Network, Ghana, 2022* (pp. 409–415). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-41352-0_21
- Baumer, B. (2015). A data science course for undergraduates: Thinking with data. *The American Statistician*, 69(4), 334–342. <https://doi.org/10.1080/00031305.2015.1081105>
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. The MIT Press.
- Engel, J. (2017). Statistical literacy for active citizenship: A call for data science education. *Statistics Education Research Journal*, 16(1), 44–49. <https://doi.org/10.52041/serj.v16i1.213>
- Engel, J., & Martignon, L. (2022). Data science for informed citizens: on synergies between data literacy and civic statistics. In S. A. Peters, L. Zapata-Cardona, F. Bonafini, & A. Fan (Eds.), *Bridging the Gap: Empowering & Educating Today's Learners in Statistics. Proceedings of the 11th International Conference on Teaching Statistics (ICOTS11)*. International Association for Statistical Education.
- Engel, J. & Weber-Stein, F. (2023). Teaching social science students to think with data. In EM Jones (Ed.), *Fostering Learning of Statistics and Data Science. Proceedings of the Satellite conference of the International Association for Statistical Education (IASE)*, July 2023, Toronto, Canada.
- Gal, I. (2023). Critical Understanding of Civic Statistics: Engaging with Important Contexts, Texts, and Opinion Questions. In J. Ridgway (Ed.), *Statistics for Empowerment and Social Engagement* (pp. 323–343). Springer. https://doi.org/10.1007/978-3-031-20748-8_13
- Gal, I., Nicholson, J., Ridgway, J. (2023). A Conceptual Framework for Civic Statistics and Its Educational Applications. In J. Ridgway (Ed.), *Statistics for Empowerment and Social Engagement* (pp. 37–66). Springer. https://doi.org/10.1007/978-3-031-20748-8_3
- Ghosh, D., Boettcher, W. A., Johnston, R., & Lahiri, S. (2025). THANOS: A Predictive Model of Electoral Campaigns Using Twitter Data and Opinion Polls. *Data Science in Science*, 4(1). <https://doi.org/10.1080/26941899.2025.2484180>
- Krishnannair, A., & Krishnannair, S. (2022). “Critical Statistical Literacy”, “Social Justice Statistics”, and “Critical Statistical Consciousness” as Higher Education Imperatives, amidst the Covid-19 Pandemic in South Africa. *South African Journal of Higher Education*, 36(1), 154–170. <https://doi.org/10.20853/36-1-4509>
- Martínez-Castro, C.A., Zapata-Cardona, L., Jones, G.L. (2023). Critical Citizenship in Statistics Teacher Education. In G. F. Burrill, L. de Oliveria Souza & E. Reston (Eds.), *Research on Reasoning with Data and Statistical Thinking: International Perspectives* (pp. 213–226). https://doi.org/10.1007/978-3-031-29459-4_17
- Radermacher, W. J. (2022). Statistical awareness promoting a data culture. *Statistical Journal of the IAOS*, 38(2), 453–461. <https://doi.org/10.3233/sji-220956>
- Rak, J., & Rezmer-Płotka, K. (2022). Civic Education for Democracy During Crisis: Measuring State Media Engagement. *The New Educational Review*, 69(3), 82–94. <https://doi.org/10.15804/ner.2022.69.3.06>
- Thelma, C. C., Madoda, D., Sylvester, C., Patrick, M., & Olayinka, Y. (2024). Evaluating the Effectiveness of Civic Education in Enhancing Voter Participation and Political Engagement. *International Journal of Research and Innovation in Social Science*, 8, 4478–4493. <https://doi.org/10.47772/ijriss.2024.8080344>
- Watson, J. M. (2003). Statistical Literacy at the School Level: What Should Students Know and Do? *The Bulletin of The International Statistical Institute*, 54, 1–4. <https://hdl.handle.net/102.100.100/521470>

-
- Weber-Stein, F., & Engel, J. (2025). Implementing Statistical Literacy in Civics Teacher Education. *Journal of Political Science Education*, 1–20. <https://doi.org/10.1080/15512169.2025.2483786>
- Weiland, T. (2017). Problematizing statistical literacy: An intersection of critical and statistical literacies. *Educational Studies in Mathematics*, 96(1), 33–47. <https://doi.org/10.1007/s10649-017-9764-5>