

The Importance of Context in Data Analysis

Bolin Shen

2024-01-21

In the quote, Amia Srinivasan warns us to be careful when we are interested in data that is not very strong but supports ideas that many people have believed in the past. She urges us to be cautious because sometimes this kind of information can be used to prove ideas that have been powerful in societies for a long time. The problem is that these ideas, deeply rooted in our culture, can be used to justify treating large groups of people unfairly. Srinivasan makes us think about how easily we might accept data that agrees with what we already believe, especially if those beliefs have affected society before and might lead to treating some groups unfairly. The quote stresses that we should examine information carefully and be aware of how it might affect fairness and justice. It suggests we should be thoughtful and careful when interpreting information, especially if it supports ideas from the past that have believed to be unfair or discriminatory in society.

D'Ignazio and Klein discuss the concept of “Big Dick Data,” a term coined by them to describe big data projects characterized by patriarchy, aiming to dominate the world through data capture and analysis. They criticize projects such as GDELT (Global Database of Events, Language and Tone) for exaggerating the completeness and accuracy of its data and algorithms. They argue that these projects often ignore context, fetishize size and exaggerate their technical and scientific capabilities. The lack of context in data, especially those obtained from web portals or APIs, is highlighted as an important problem. The Open data movement has also been criticized for prioritizing the “opening up” of data without providing enough context, leading to the phenomenon of “zombie data.”

In addition, D'Ignazio and Klein highlight how power differentials and social dynamics affect the data collection process and resulting data sets, and they emphasize data are not raw inputs but are already “cooked” by social, political, and historical circumstances. They provide examples of power imbalances, as well as issues related to data misinterpretation, such as underreporting on sexual assault cases on some college campuses. They argue for a more critical and context-aware approach to data analysis, urging data scientists to ask about the limitations and validity of the data they use. It also explores emerging practices, tools, and methods for applying context into data analysis and communication, including data user guides, datasheets for datasets, and the role of data intermediaries.

The emergence of a “data creative” class, including data scientists, journalists, artists, and designers, highlights the importance of creatively mining and combining data to produce meaningful insights. However, this creativity should be grounded in an understanding of context to avoid misleading analysis. D’Ignazio and Klein provide examples where creative data synthesis worked well and cases where it failed due to lack of context.

To sum up, the principle of considering context is a key aspect of data feminism, advocating for a deeper understanding of power dynamics, biases, and social inequalities in the data ecosystem. D’Ignazio and Klein emphasize the importance of understanding the context in which data is generated, connecting data to their context, and considering the ethical implications and power dynamics involved in data creation and use.

References: Cowen, Tyler. 2021. “Episode 132: Amia Srinivasan on Utopian Feminism.” Conversations with Tyler, September. <https://conversationswithtyler.com/episodes/amia-srinivasan/>. D’Ignazio, Catherine, and Lauren Klein. 2020. Data Feminism. Massachusetts: The MIT Press. <https://data-feminism.mitpress.mit.edu>. Beame, Katherine. 2021. Thinking like a Data Feminist, <https://medium.com/ihme-tech/thinking-like-a-data-feminist-4dc5e2b58bfc> Dinç, Sibel. 2021. Rules are necessary: Data Feminism and its seven principles. <https://www.internetjustsociety.org/rules-are-necessary-data-feminism-and-its-seven-principles>.