

Ayiana Baynes

Dr. DeSpain

CODES 499

23 July 2025

Katie Rawson and Trevor Muñoz, “Against Cleaning” Annotation

The purpose of this article is to challenge the current perception of “data cleaning” and to explore the various misconceptions that have arisen surrounding this term. The methodology of this article used a mixed-methods approach, combining both qualitative and quantitative methods. The authors examined their own research processes, obtained insights from previous studies that corresponded with their research question, and involved the public to broaden their study. This comprehensive approach allowed for a more detailed understanding of the topic. The key findings from their study indicate that “data cleaning” is a crucial and complex aspect of research. It involves much more than just eliminating duplicate entries; it plays a vital role in preserving diversity within datasets. Therefore, the concept of “data cleaning” demands deeper exploration and greater recognition in the research community.

The research was very detailed as it provided a thorough analysis of the concept of “data cleaning” and how it is extremely oversimplified. It did not present data cleaning as a technical task; instead, it unfolded a narrative that showcased its significance in real-world studies, allowing us to see the complex process that goes along with it. The importance of this article matches the results, especially given the everlasting stigma surrounding “data cleaning.” Initially, I had little experience in this area, but I expected that much of my work would consist of me cleaning the data. Recently, as I researched the next steps for analyzing my survey responses, I

encountered discussions about data cleaning that emphasized the importance of eliminating “mistakes” and how much time it would take. However, when I finally began the cleaning process, I discovered that there was actually no real data I had to remove. Each response I received proved to be important to my research, and the term “cleaning” initially suggested that I would be discarding useless material. This experience, along with the article I read, highlighted how many aspects of data analysis remain a mystery and how much more we have to learn.

How this compares to other research I have encountered is that I have never seen researchers use outside individuals who volunteer to help them “clean data.” It was a concept that is completely new to me and I found it very interesting. This research truly challenges my previous knowledge, as I always assumed that involving outside sources could lead to unpredictability, which was something to avoid. However, by using volunteers, it allowed for a variety of perspectives and different understandings of the task. This diversity not only enhanced the process but also uncovered limitations, creating a unique dynamic that constrained and propelled them forward. This article might inform future research and practice by breaking down the stigma over “data cleaning” and opening our minds to the different possibilities and processes that are much more complex.

Works Cited

Rawson, Katie and Muñoz, Trevor. "Against Cleaning." *Debates in the Digital Humanities*, University of Minnesota Press, Apr. 2019,
<https://dhdebates.gc.cuny.edu/read/untitled-f2acf72c-a469-49d8-be35-67f9ac1e3a60/section/07154de9-4903-428e-9c61-7a92a6f22e51#ch23>. Accessed 8 Aug. 2025