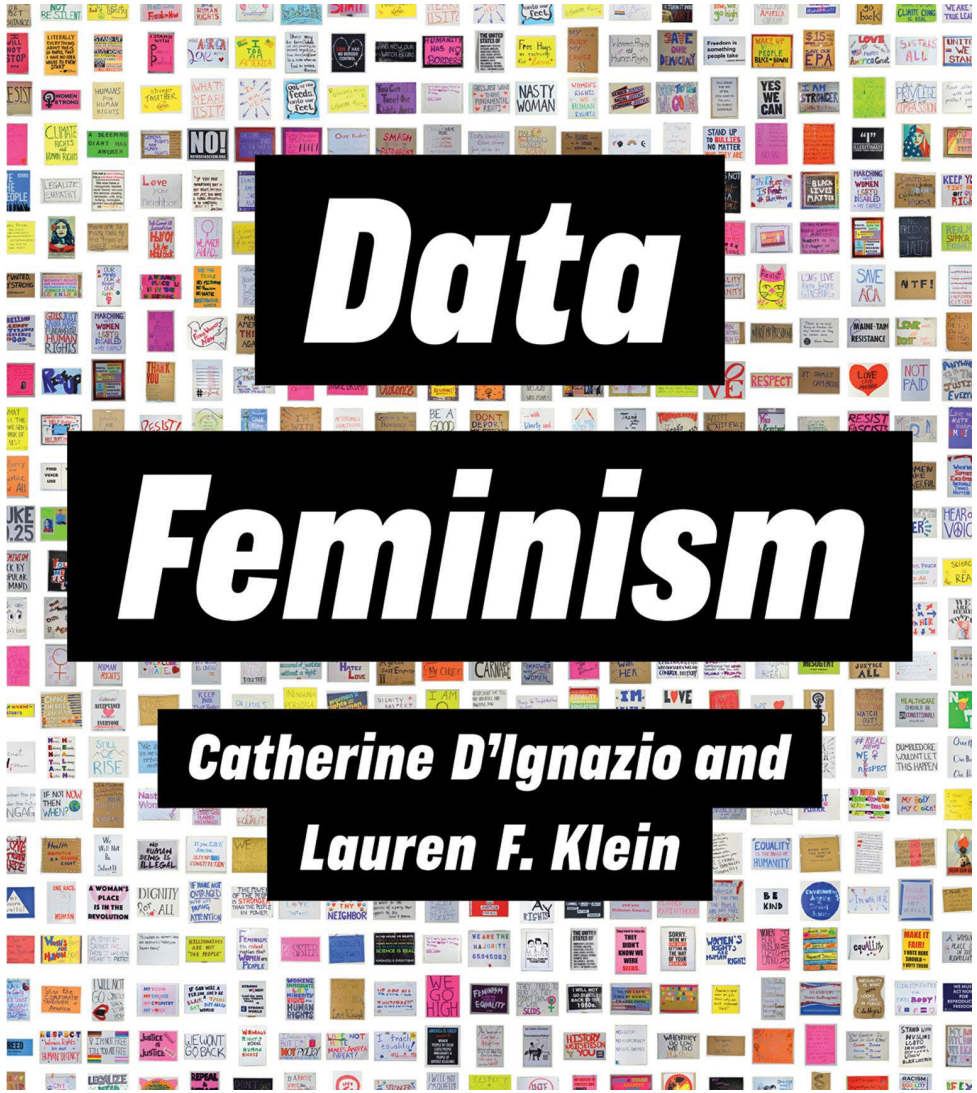
Data Feminism



D'Ignazio and Klein describe **Data Feminism** as:

"A way of thinking about data, both their uses and their limits, that is informed by direct experience, by a commitment to action, and by intersectional feminist thought. The starting point for data feminism is something that goes mostly unacknowledged in data science: power is not distributed equally in the world."

Data Feminism: Four Actions*

DATA IS A STRUCTURE OF POWER

1. Align work with your values to affect positive change in the world.

DATA IS NEVER NEUTRAL OR OBJECTIVE

2. Reflect, align with values, and document all the decisions made when creating data.

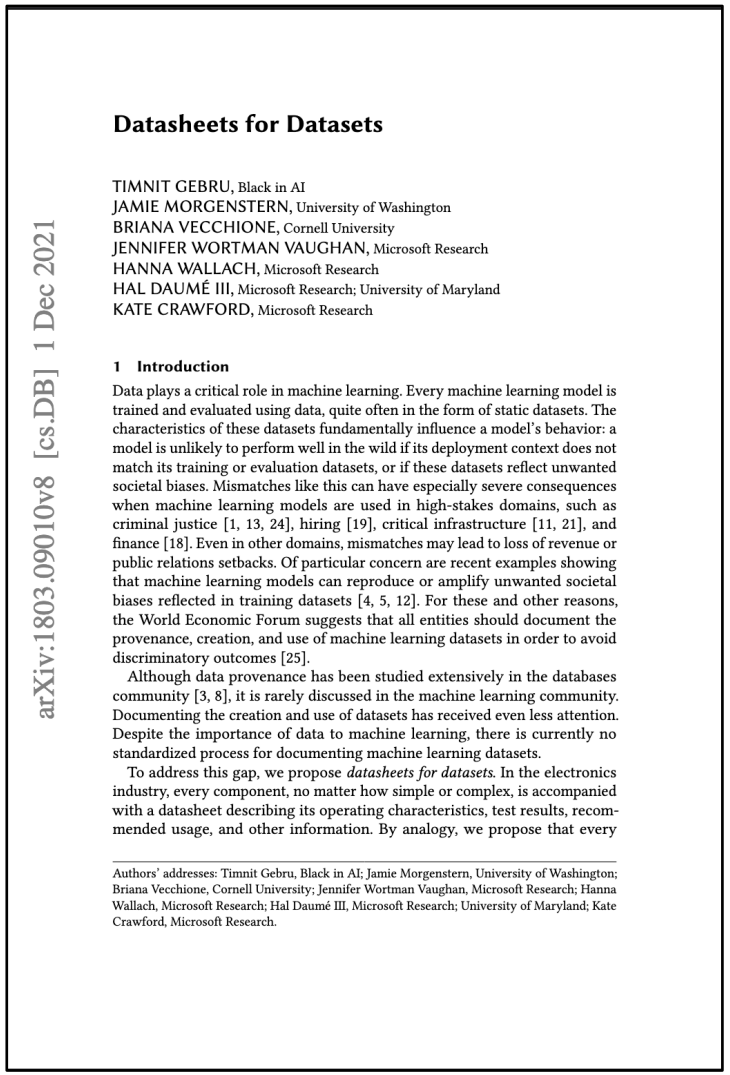
3. Integrate a diverse range of sources; combine qualitative + quantitative analyses.

DATA SCIENCE REQUIRES MANY PARTICIPANTS

4. Surface all the people that contribute to our work as data scientists.

* Note that this is a very opinioned and partial distillation of the original ideas.

Datasheets



arXiv:1803.09010v8 [cs.DB] 1 Dec 2021

Datasheets for Datasets

TIMNIT GEBRU, Black in AI
JAMIE MORGENSTERN, University of Washington
BRIANA VECCHIONE, Cornell University
JENNIFER WORTMAN VAUGHAN, Microsoft Research
HANNA WALLACH, Microsoft Research
HAL DAUMÉ III, Microsoft Research; University of Maryland
KATE CRAWFORD, Microsoft Research

1 Introduction

Data plays a critical role in machine learning. Every machine learning model is trained and evaluated using data, quite often in the form of static datasets. The characteristics of these datasets fundamentally influence a model's behavior: a model is unlikely to perform well in the wild if its deployment context does not match its training or evaluation datasets, or if these datasets reflect unwanted societal biases. Mismatches like this can have especially severe consequences when machine learning models are used in high-stakes domains, such as criminal justice [1, 13, 24], hiring [19], critical infrastructure [11, 21], and finance [18]. Even in other domains, mismatches may lead to loss of revenue or public relations setbacks. Of particular concern are recent examples showing that machine learning models can reproduce or amplify unwanted societal biases reflected in training datasets [4, 5, 12]. For these and other reasons, the World Economic Forum suggests that all entities should document the provenance, creation, and use of machine learning datasets in order to avoid discriminatory outcomes [25].

Although data provenance has been studied extensively in the databases community [3, 8], it is rarely discussed in the machine learning community. Documenting the creation and use of datasets has received even less attention. Despite the importance of data to machine learning, there is currently no standardized process for documenting machine learning datasets.

To address this gap, we propose *datasheets for datasets*. In the electronics industry, every component, no matter how simple or complex, is accompanied with a datasheet describing its operating characteristics, test results, recommended usage, and other information. By analogy, we propose that every

Authors' addresses: Timnit Gebru, Black in AI; Jamie Morgenstern, University of Washington; Briana Vecchione, Cornell University; Jennifer Wortman Vaughan, Microsoft Research; Hanna Wallach, Microsoft Research; Hal Daumé III, Microsoft Research; University of Maryland; Kate Crawford, Microsoft Research.

Datasheets are a proposal to pair datasets with a file that: *"documents its motivation, composition, collection process, recommended uses, and so on. Datasheets for datasets have the potential to increase transparency and accountability within the machine learning community, mitigate unwanted societal biases in machine learning models, facilitate greater reproducibility of machine learning results, and help researchers and practitioners to select more appropriate datasets for their chosen tasks."*

Datasheets: Sections

1. Metadata
2. Motivation
3. Composition
4. Narrative
5. Distribution
6. Attributions
7. References
8. Notes

Two sections removed/renamed because they were specific to machine learning applications.

Three sections were added to better align with ideas from *Data Feminism*.