

Biased Priorities, Biased Outcomes:

Three Recommendations for Ethics-oriented Data Annotation Practices

Gunay Kazimzade*

Weizenbaum Institute for the Networked Society
Technische Universität Berlin
Berlin, Germany
gunay.kazimzade@tu-berlin.de

Milagros Miceli*

Weizenbaum Institute for the Networked Society
Technische Universität Berlin
Berlin, Germany
m.miceli@tu-berlin.de

ABSTRACT

In this paper, we analyze the relation between data-related biases and practices of data annotation, by placing them in the context of market economy. We understand annotation as a praxis related to the sensemaking of data and investigate annotation practices for vision models by focusing on the values that are prioritized by industrial decision-makers and practitioners.

The quality of data is critical for machine learning models as it holds the power to (mis-)represent the population it is intended to analyze. For autonomous systems to be able to make sense of the world, humans first need to make sense of the data these systems will be trained on. This paper addresses this issue, guided by the following research questions: Which goals are prioritized by decision-makers at the data annotation stage? How do these priorities correlate with data-related bias issues? Focusing on work practices and their context, our research goal aims at understanding the logics driving companies and their impact on the performed annotations.

The study follows a qualitative design and is based on 24 interviews with relevant actors and extensive participatory observations, including several weeks of fieldwork at two companies dedicated to data annotation for vision models in Buenos Aires, Argentina and Sofia, Bulgaria.

The prevalence of market-oriented values over socially responsible approaches is argued based on three corporate priorities that inform work practices in this field and directly shape the annotations performed: *profit* (short deadlines connected to the strive for profit are prioritized over alternative approaches that could prevent biased outcomes), *standardization* (the strive for standardized and, in many cases, reductive or biased annotations to make data fit the

products and revenue plans of clients), and *opacity* (related to client's power to impose their criteria on the annotations that are performed. Criteria that most of the times remain opaque due to corporate confidentiality).

Finally, we introduce three elements, aiming at developing ethics-oriented practices of data annotation, that could help prevent biased outcomes: *transparency* (regarding the documentation of data transformations, including information on responsibilities and criteria for decision-making.), *education* (training on the potential harms caused by AI and its ethical implications, that could help data annotators and related roles adopt a more critical approach towards the interpretation and labeling of data), and *regulations* (clear guidelines for ethical AI developed at the governmental level and applied both in private and public organizations).

CCS CONCEPTS

- Computing methodologies ~ Artificial intelligence • Computing methodologies ~ Machine learning • Applied computing ~ Annotation • Social and professional topics ~ Computer supported cooperative work

KEYWORDS

Bias; data; annotation; classification; ethics; power; transparency; priority; profit

ACM Reference format:

Gunay Kazimzade, Milagros Miceli. 2020. Biased Priorities, Biased Outcomes: Three Recommendations for Ethics-oriented Data Annotation Practices. In *AIES '20: AAAI/ACM Conference on AI, Ethics, and Society Proceedings, February 7–8, 2020, New York, NY, USA*. ACM, New York, New York, USA, 1 page. <https://doi.org/10.1145/3375627.3375809>

***Both authors have contributed equally to this work.**

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

AIES '20, February 7–8, 2020, New York, NY, USA

© 2020 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-7110-0/20/02.

<https://doi.org/10.1145/3375627.3375809>