Penultimate draft, final version published in the journal Philosophy & Technology. Available open access at https://link.springer.com/article/10.1007/s13347-022-00549-9

Who Is a Good Data Scientist? A Reply to Curzer and Epstein

Mark Graves[1] and Emanuele Ratti[2,3,4]

A central distinction in Curzer and Epstein (2022) is the one between 'protect the disadvantaged' and 'protect the data'. The distinction opens up discussions about the relationship between ethics and epistemology in the practice of science. Focusing on the disadvantaged to the exclusion of good scientific practices, Curzer and Epstein argue, can harm everyone impacted by medical science, including the disadvantaged. For this reason, they propose that "ethical data scientists should strive for accurate data and scientifically sound data analysis" (2022, p 2) with attention to minimizing data processing errors, bias, and outside influence, and that includes identifying errors caused by tendencies to neglect disadvantaged and historically underrepresented communities and groups. While we agree with several points made by Curzer and Epstein, we also have three main points of concern.

1. Microethics and the 'protect the disadvantaged' approach

We believe that Curzer and Epstein partially mischaracterizes our original paper (Ratti and Graves 2021) as 'protect the disadvantaged'. The author builds upon that characterization to argue for a protect-the-data approach that prioritizes scientific accuracy over extra attention to the disadvantaged. But the "protect the disadvantaged" descriptor misrepresents our project. In the commentary, Curzer and Epstein argue:

"*Ratti and Graves' use of the capabilities approach as well as their examples, leave the impression that ethical data science primarily consists in identifying and avoiding data management decisions that inappropriately negatively impact the disadvantaged" (p 2).*

First, it should be noted that our framework has the explicit goal of shaping all phases of the data science pipeline, and data management phases are only a fraction of it. In fact, we expand our analysis in Graves and Ratti (2021). Next, our focus on the 'disadvantaged' is an artifact

[1] Parexel AI Labs, San Francisco, CA, USA
[2] Institute of Philosophy and Scientific Method, Johannes Kepler University Linz, Austria
[3] Department of Arts and Humanities, Technion Israel Institute of Technology, Haifa, Israel
[4] mnl.ratti@gmail.com

of the proof-of-concept that we develop: it is easy to illustrate the impact on capabilities on those who are blatantly being unequally impacted by mechanisms of social injustice. Moreover, the 'inappropriately' is out of place: when describing moral attention, we explicitly say that it is not a full-blown virtue, as "moral attention may not be effective in choosing and acting, but only in reasoning" (p 10). This means that moral attention is just realizing that there is an impact on capabilities, but whether this impact is positive or negative (appropriate or inappropriate) is pretty much left in the air, and it will depend on either other virtues or the participatory design sketched at the end of the article.

In the same page, Curzer and Epstein adds that "[i]t is also that focusing on a subgroup distorts results of data management and undermines its goals of advancing knowledge" (p 2). We agree that focusing on 'a subgroup' at the expense of others can harm everyone, but we would argue that the process still needs to focus on impinged capabilities. In agreement with the commentary authors, we would argue that data scientists should not focus on the group they initially believe is disadvantaged, but should focus on the impinged capabilities of everyone, recognizing that those with the greatest impinged capabilities would, as a consequence, be considered disadvantaged[5].

2. Ethics and Epistemology, or science and values

In the commentary, emphasis is added on accuracy for the integrity of data science, and how this is important for advancing knowledge (p 2).

However, the author treats 'advancing knowledge' as a pure epistemic goal achievable by pure epistemic means, and we view this as a controversial thesis. Because science always operates in a situation of uncertainty, risks of epistemic errors arise anywhere during scientific practices. There is a rich literature in philosophy of science - known through various labels such as inductive risk, epistemic risk, etc - focusing on connections between scientific choices and ethical considerations (Douglas 2009; Elliott and Richardson 2017). Risks must be managed and balanced in light of values and interests, and this makes value-laden choices in science inevitable: scientists proceed by balancing and managing risks and uncertainty via values (Ward 2021). The case in our original article of choosing which data set to process first, whether EMR or OCR, is a good example. Assuming that resources are limited (which is a defining feature of the context in which medical data scientists operate), you cannot deal with

[5] Partially underlying the agreement is that if data scientists imagine a disadvantaged group, they are likely to imagine one with impinged *functionings*, while the proposed microethical process is guiding them to identify impinged *capabilities*

EMR and OCR data sets with the same level of accuracy – this epistemic desideratum will lean towards a certain direction. You have to make a choice. This can be based on efficacy ('I'll go with EMR because those data sets are bigger and easier to deal with') or on concerns about complete representation, which would include the capabilities of the disadvantaged ('It's likely that OCR data will be about them'). In both cases, we treat data accurately (or to the best of the means available), but the direction is shaped by values.

To make the same point a bit differently, how accuracy is modulated, and in which direction should be pursued, is not a pure epistemic matter. Of course, if we had all the data of the world, all the computational power of the universe, and billions of data scientists, then we could strive for complete accuracy. But that's just not possible. This argument makes 'values' inevitable, given our situation of limited beings with limited information. One can even make a further argument and argue that, beyond the problem of uncertainty, science *per se* has an ethical dimension intertwined with the epistemic one, as it can be value-promoting, in the sense that scientific choices promote certain values while simultaneously obfuscating others (Russo 2021). Philosophy of technology has already previously explored this territory in the context of power relations (Winner 1980).

Even if we want to debate the value-free ideal of science on its own terms, we should consider that the over-reliance on epistemic characteristics, such as accuracy, is problematic. In fact, Curzer and Epstein seem to assume that accuracy has one and only one meaning, that there is one way to measure it, and that data scientists will agree on all these things. But this is a situation that the history of science and technology has shown pretty well is not the case: epistemic desiderata are understood and operationalized in many, and sometimes mutually exclusive, ways (Kuhn 1977). This means that epistemic desiderata such as accuracy require value judgment in order to be operationalized (McMullin 1983): epistemic desiderata are indeed values, and one can even question the distinction between epistemic and non-epistemic values itself (Rooney 1992). But even assuming we do have one notion, consider this. One cannot know to include something like transportation conversion factors in the model unless one realizes they could be a significant factor. There is a debatable point here whether that requires moral attention or just heightened awareness of social factors determining health, but it still requires cultural and/or moral awareness within the data science process that goes beyond pure technical proficiency and epistemic considerations.

Curzer and Epstein seem to imply, even further, that there is a dichotomy between ethics and epistemology. Putting ethics and epistemology in opposition through the 'protect the data' approach has an interesting consequence. Given that ethical considerations must

appear at some point, these will be external to the practice of data science. This is well-formulated at page 4:

"Rather our claim is that the appropriate point at which moral concern about the methodology and application of the study comes into play is not during the study, but rather before or after the study"

We disagree that moral attention should only occur "before" or "after" the data work. We see this as an externality model of science and values that was much criticized by Longino decades ago (1990), when she argued against the assumptions that ethics is completely external to science, and that science within its internal activities is value-free. Moreover, value-free science is not only descriptively false, but also normatively problematic. The case of the diabetes intervention in our article is an example of why moral attention is needed *during* the project. It is clear that an attention to the ethical dimension of data subjects improves the study also from the point of view of epistemic considerations alone, such as brute performance metrics. The same applies to the analysis of missing data.

3. **The social context of data science**

At pages 2 and 3, Curzer and Epstein claim:

"Recognizing that scientific errors can impede the health agency of large numbers of people in sometimes unpredictable ways, good data scientists try to imagine what scientific errors might be introduced by a proposed data management choice, and then take steps to avoid or ameliorate these errors"

This claim is useful to introduce the importance of the social context to which data science is going to make a difference, and why ethics and epistemology are necessarily intertwined because of that context. If 'scientific' is understood as 'technical integrity', which the authors seem to identify with 'data accuracy', then even processing accurate data may lead to negatively affecting health agency of a large number of individuals. If you do this in a society where health injustice is systematic like the USA, then data science tools will simply provide predictions that are informed by the same patterns of systematic injustice, and hence will impact the substantial freedoms of data subjects, as we have documented in our article. However, this is problematic because of the role that medical data scientists play. We claim

that the profession of the medical data scientist inherits the same ethical and epistemic obligations that any member of the medical community has, which is to promote the well-being of patients. But because of the well-documented ripple effects on substantial freedoms that data science tools have, being a medical data scientist implies also a 'protecting human agency' perspective, given that human agency as a substantial freedom is a necessary component of well-being. This can be preserved by our use of the capability approach and the sketch of the participatory design at the end of our article.

In other words, the 'protect the data' position does not sufficiently acknowledge the values influencing all scientific endeavors, the broader social systems in which data science takes place, and the obligations of data scientists *qua* members of the medical community towards those social systems. We might agree with Curzer and Epstein that a scientist should not commit extra resources to one group over another beyond developing a representative model, but we argue that the false separation of scientific and ethical practices fails to acknowledge that the scientific practice occurs within a system that has ingrained biases, and that data scientists have ethical obligations towards those biases, as they impact the well-being of data subjects by being detrimental to their substantial freedoms. In healthcare in particular, a data scientist is tasked with developing models that represent the population and its healthcare needs, and that means *not* incorporating the systemic biases into the modeling framework.

REFERENCES

Curzer, H. J., & Epstein, A. C. (2022). The virtuous data scientist and the ethics of good science. *Philosophy & Technology*, 1–5. https://doi.org/10.1007/s13347-022-00541-3

Douglas, H. (2009). *Science, policy, and the value-free ideal*. University of Pittsburgh Press.

Elliott, K., & Richardson, T. (2017). *Exploring inductive risk: Case studies of values in science*. Oxford University Press.

Graves, M., & Ratti, E. (2021). Microethics for healthcare data science: Attention to capabilities in sociotechnical systems. *The Future of Science and Ethics*, 6, 64–73. https://doi.org/10.53267/20210106

Kelly, T. (2018). *Professional ethics - A trust-based approach*. Lexington Books.

Kuhn, T. (1977). Rationality, value judgment, and theory choice. In *The Essential Tension* (pp 320–339). Chicago University Press

Longino, H. (1990). *Science as social knowledge: Values and objectivity in scientifc inquiry*. Princeton University Press

Mcmullin, E. (1983). Values in Science. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*

Ratti, E., & Graves, M. (2021). Cultivating moral attention: A virtue-oriented approach to responsible data science in healthcare. *Philosophy and Technology*, 34(4), 1819–1846. https://doi.org/10.1007/ s13347-021-00490-3

Rooney, P. (1992). On values in science: Is the epistemic/non-epistemic distinction useful?, *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*

Russo, F. (2021). Value-promoting concepts in the health sciences and public health, *Philosophical News.*

Ward, Z. B. (2021). On value-laden science. Studies in History and Philosophy of Science, 85, 54–62. https://doi.org/10.1016/j.shpsa.2020.09.006

Winner, L. (1980). Do artifacts have politics? Daedalus, 109(1)