The Many Facets of Data Equity

Extended Abstract

H. V. Jagadish
University of Michigan
jag@umich.edu

Julia Stoyanovich
New York University
stoyanovich@nyu.edu

Bill Howe
University of Washington
billhowe@uw.edu

ABSTRACT

Data-driven systems can be unfair, in many different ways. All too often, as data scientists, we focus narrowly on one technical aspect of fairness. In this paper, we attempt to address equity broadly, and identify the many different ways in which it is manifest in data-driven systems.

1 INTRODUCTION

There is concern about fairness today, whenever data-driven systems are used. It is no longer believed that data are impartial and neutral. Nevertheless, the scope of fairness considered is often narrow. Computer scientists are trained to develop algorithms that can solve cleanly stated formal problems. If fairness could be reduced to a mathematical constraint, it would have been addressed by now. The difficulty, of course, is that fairness is more complicated than that. Technical solutions can help, but are not enough in themselves to address the real problems. Our goal is to get past these limitations and address data equity broadly defined.

To reach our goal, we begin with a discussion of data equity in Section 2. Based on this understanding, in Section 3, we examine multiple facets of data equity that must all be addressed.

2 WHAT IS DATA EQUITY

Equity as a social concept promotes fairness by treating people differently depending on their endowments and needs (focused on equality of outcome), whereas equality aims to achieve fairness through equal treatment regardless of need (focused on equality of opportunity) [15, 16, 31, 35, 36]. Equity is not a legal framework per se, yet underpins civil rights laws in the U.S. that restrict preferences based on protected classes, for example in housing or employment [50–52]. It has recently also been operationalized in computer science scholarship, primarily through fairness in machine learning research [5]. However, equity is a much richer concept than a simple mathematical criterion that can be captured in a fairness constraint.

Even in the best of circumstances, underlying structural inequities in access to health care, employment, and housing exhibit themselves in the data record and are propagated through decision systems, automated or otherwise, to become reinforced by policy. A key thing that's missing is a treatment of how decision systems, regardless of consideration of equity, reinforce existing structures. Therefore, any effort to define and improve "Data equity" may consist largely of "happy talk" [6], which involves a willingness to acknowledge and even revel in cultural difference without seriously challenging ongoing structural inequality.

We consider data equity in the context of automated decision systems, while recognizing a broader literature around the role of administrative systems in creating and reinforcing discrimination. Spade argues that administrative systems facilitate state violence encoded in laws, policies, and schemes that arrange and define people by categories of indigeneity, race, gender, ability, and national origin [44]. Hoffman considered how these effects are amplified through data technologies and their purveyors [21]. Decision systems, regardless of consideration of equity, mechanize existing structures, such that any effort to define and address data equity issues are at risk of becoming mere technological "happy talk." To combat these outcomes, we emphasize the need to think about equity broadly, and to own the outcomes realized. Ideally, we have a primacy of equity in the design: the goal is not just to automate and correct for equity, but to design systems that exist to further equity. For example, a machine learning system to help submit insurance claims to maximize payment is designed to counteract the discrimination effected by corporate models to minimize payments. However, we know that no everyone data-driven system can have equity as its purpose. So we also must develop a framework to recognize and remedy the many different ways in which a data-driven system may introduce inequities.

Data and administrative systems construct the very identities and categories presented to us as "natural," both inventing and producing meaning for the categories they administer [45](:pp. 31–32). Administrative systems facilitate state violence encoded in laws, policies, and schemes that arrange and define people by categories of indigeneity, race, gender, ability, and national origin, which Spade calls "administrative violence" [45](:pp. 20–21).

Similarly, transportation apps like Ghettotracker and SafeRoute are designed to help users navigate around "dangerous" or unsafe areas. In practice, they often target neighborhoods populated by people of color by encoding racist articulations of what constitutes danger [19].

That social inequity is reinforced and amplified by data-intensive systems is not new. We know from other domains that advances in data science and AI can be undermined by similar problems: automated decisions based on biased data can operationalize, entrench, and legitimize new forms of discrimination. For example, a defendant's immediate social network may reveal many convictions, but that information must be interpreted through the lens of socioeconomic conditions and prior structural discrimination in the criminal justice system before concluding that an individual is at a higher risk of recidivism or bail violation. Similarly, standardized test scores are sufficiently impacted by preparation courses that the score itself says more about socioeconomic conditions than an individual's academic potential.

In summary, the manner in which data systems are built and used can compound and exacerbate inequities we have in society. It can also introduce inequities where there previously were none. Avoiding these harms results in *data equity*, and is accomplished through constructing socio-technical systems that we call *data equity systems.*

3 FACETS OF DATA EQUITY

We have examined dozens of examples of inequities in data systems, such as those cited in the preceding section. Based on our empirical study, we have identified four distinct facets of data equity [24], which we present here as a rough taxonomy of the issues to be considered in the construction of data equity systems.

3.1 Representation equity

There often are material deviations between the data record and the world the data is meant to represent, often with respect to historically disadvantaged groups [10]. Perhaps the best-known case in this regard has to do with crime records used for predictive policing. Many offenses are recorded only when there is police presence. While citizens may call the police in for some types of crimes, both major (such as a murder) and minor (such as a noisy party), it would be unusual for the police to be called in because of a report of jaywalking or minor drug possession. Rather, these offenses are only entered into the record when police happen to observe them, and choose not to ignore them. Therefore, crimes are more likely to be observed in areas with greater police presence, and among these observed, crimes are more likely to be recorded where the police officer chooses not to give the offender a pass, a choice that has historically been racially biased. In other words, the data record reflects, and can enshrine, historical injustices. The use of this record for future police deployments can lead to a vicious cycle of victimizing communities that have suffered in the past.

Representation issues can arise even when there is no historical record involved. For example, confirmed COVID-19 cases require testing, and there can be racial disparities in both the availability of testing and in the desire of individuals to be tested, leading to systematic biases in collected data. These disparities are found in contemporary data, even if they are rooted in historical discrimination. For example, there may be fewer test sites located in minority neighborhoods, or poor people lacking insurance may worry about the cost of testing and this may reflect in racial statistics. Similarly, a long history of being unfairly treated by the medical profession may make African-Americans naturally wary of such interactions and hence induce reluctance in testing. Whatever be the reasons, the point is that contemporary data may under-represent racial minorities, particularly African-Americans, and hence potentially lead to under-estimating the prevalence of COVID-19 in these communities.

Representation inequities in the data can lead to systemic biases in the decisions based on the data. But it can also lead to greater errors for under-represented groups. Consider facial recognition as an example. It has been extensively documented, across numerous current systems, that these systems are considerably more accurate with white males than with women or people of color. Higher error rates for a community is also a harm, in this case caused by a lack of representation. These error rates may not only be higher, but they could additionally also be biased. For example, Amazon developed software to screen candidates for employment and trained this software on data from the employees it already had. Since its employees were mostly male, women were under-represented in the data record. Worse still, because of historical discrimination, the few women previously in the company had done poorly compared to their potential. A model trained on this data set began classifying most women as unsuitable for hiring, a problem that exacerbated historical difficulties. Amazon had to cancel this project even before it launched.

Representation issues typically, but not exclusively, occur in data about people. But there are many exceptions, which can still have inequitable impacts on people. The city of Boston released an app, called StreetBump, to report potholes in its streets. The app was downloaded and installed by many citizens with smartphones, and reported many potholes to the city. The difficulty was that smartphones were more frequently owned by the better off residents of the city, and these were also more likely to take the effort to install the app because of their history-driven belief in government. The consequence would have been a data record with inadequate representation of streets in poor neighborhoods: a problem that was proactively corrected by the city, through sending out its own pothole recording crews to use the app in poorer neighborhoods. Similarly, richer countries have many more weather stations measuring conditions in the atmosphere and in the ocean. The disparity of representation in the data record can lead to weather predictions being less accurate for poor countries.

Data representation issues, and the harms they cause, may first appear in the input, output, or at any intermediate data processing step, but the majority of research in AI bias and fair ML pertains only to learning. We must develop techniques to introspect and intervene at any stage of the data pipeline. It is not enough to hope that we will mitigate the propagation of data representation issues during a final learning step.

Our solution is to adopt database repair [38] as the guiding principle. We have developed techniques to detect under-representation efficiently for a high number of small-domain discrete-valued attributes, such as those that result from joining multiple tables in a relational schema [4, 25, 28]. Once representation gaps are detected, we consider cases where they can be filled by collecting more data. We have shown how to satisfy multiple gaps at the same time efficiently [4, 25]. We have linked causal models to the conditional independence relationships used in the database repair literature, suggesting a new algorithm for causal database repair such that any reasonable classifier trained on the repaired data set will satisfy *interventional fairness* and empirically perform well on other definitions [37]. We have developed [40], a design and evaluation framework for fairness-enhancing interventions in data-intensive pipelines that treats data equity as a first-class citizen and supports sound experimentation [39, 42].

3.2 Feature equity

All the features required for a particular analysis, or to represent members of some group adequately, may not be available in a dataset. Feature equity refers to the availability of variables needed to represent members of every group in the data, and to perform desired analyses, particularly those to study inequity. For example, if attributes such as race and income are not recorded along with other data, it becomes hard to discover systematic biases that may exist, let alone correct for them.

In the recent COVID-19 pandemic, significant racial disparities have been reported in the United States on both infection rates and mortality rates. Since race is not typically recorded as part of medical care in many jusrisdictions, it has been challenging for policymakers and analysts to explore these racial differences as deeply as they would like, and to devise suitable remedies.

Similarly, eviction data does not typically include race and gender information, and this makes it hard to assess equity.

Intuitively, it is not unreasonable to think about representation equity as being concerned with rows in the data-table and feature equity as being concerned with the columns. However, feature equity includes the full scope of modeling choices made, of which attribute choice is is only one component, albeit a very important one. Another manifestation of feature equity has to do with choice for the domain of attribute values. If a gender attribute is defined to permit exactly two values, male and female, this is a modeling choice that explicitly does not accommodate other, more complex, gender expressions. Similarly, if age has been recorded in age ranges (<20, 20-30, 30-40, 40-50, 50-60, and >60), it is not possible to distinguish between toddlers and teenagers, or between a 61 year old still able to work a full day and a 95 year old no longer able to do so. If these distinctions are not important for the desired analyses, the chosen age range values are reasonable. However, many analyses may care, and may find these value choices very restricting.

When a desired attribute is not recorded at all, or has been recorded in a limited way, we may seek to impute its value. Ideally, we will be able to do this by linkage across datasets. For example, it may be possible to determine race based on census data joined on geography and statistical patterns in first and last names [53].

Where values for missing attributes cannot be determined through direct linkage, they may sometimes still be estimated through the use of auxiliary data sets. Choices among competing sources may introduce other issues; income recorded to determine eligibility for housing services will have different biases than income estimated from buying history. Furthermore, integration among datasets involves schema mapping decisions that can change the result.

Finally, imputation of missing attribute values may involve an algorithm that depends on some model, which may itself be biased. For instance, zip code can be used to "determine" race. Obviously, this cannot work at the individual level, because not everyone in a zip code is of the same race. Furthermore, even in the aggregate, we cannot always assume that the proportion of entries in our data with a particular value for race is equal to the proportion who live in that zip code. For instance, there have been several COVID-19 outbreaks in prisons, where the racial composition of prisoners is likely quite different from that of the surrounding community.

Using a novel concept of EquiTensors, we have demonstrated that pre-integrated, fairness-adjusted features from arbitrary inputs can help avoid propagating discrimination from biased data, while still delivering prediction accuracy comparable to oracle networks trained with hand-selected data sets [54, 55].

3.3 Access equity

So far, we have looked at what is in a data set. Now we look at who has access to it. Typically, data sets are owned by big companies, which spend substantial resources to construct the data set, and want to obtain competitive advantages by keeping it proprietary. On the other hand, customers may not have access to this data, and hence be at a disadvantage in any interaction with the company, even with regard to their own information. Worse still, the company has knowledge of multiple customers, which it can exploit. In contrast, the customer has access to only their own actions with the company. The customer may interact with multiple companies, but the number of companies is usually not very large, and furthermore the customer may not have access to sophisticated tools to predict company actions. In other words, data-driven systems create, and exacerbate, asymmetries, with power going to the entity with more information.

Access equity refers to equitable access to data and models, across domains and levels of expertise, and across roles: data subjects, aggregators, and analysts.

Fundamental asymmetries in information access are difficult to address. Some amelioration is possible through regulation, or voluntary transparency. Privacy policies are a tiny step in this direction, though they are far from enough in themselves, and leave a great deal to be desired the way they are currently implemented in most cases. Right to access information about oneself, as provided through GDPR is Europe, is a more substantial step.

Access to data is a challenge not just for data owned by private companies. We sometimes see similar issues in other domains as well. Researchers may hoard their data for competitive advantage in their research: if they put in the effort to collect the data in the first place, they want to analyze the data and publish their findings before releasing the collected data. Government agencies may also act similarly, driven by parochial thinking, local politics, or other such reasons.

One major impediment to making data public is the need to respect the privacy of the data subjects. A classic example is medical records: there is great potential value in making these available for analysis: surely many new patterns will be found that improve health and save lives. Yet, most people are very sensitive about sharing medical information and it has proved all too easy to re-identify anonymized data, with enough effort and ingenuity And this is even before one considers regulatory constraints on such sharing. Similarly, as citizens, we all desire open government, and would like government agencies to make their data public. But, as subjects, we also may be sensitive about some of our information with the government, and not want it made public. This is a difficult balance, which has to be managed in each instance. Technical solutions can be helpful. For instance, differential privacy may permit privacy preserving release of some information aggregates.

Even when actual access to data is not restricted, the opacity of data systems, as perceived by different groups, can also be an access equity violation. Researchers' reluctance to release data they have invested to collect contributes to the reproducibility crisis. Private companies' tight control of their data impedes external equity audits. Inadequate data release can promote misinterpretation and therefore misinformation and misuse. Data access must be accompanied by sufficient metadata to permit correct interpretation and to determine fitness for use.

A typical data science pipeline will have a sequence of data manipulations, with multiple intermediate data sets created, shared, and manipulated. Often, these data sets will be from disparate sources, and much of the processing may be conducted at remote sites. When using a remote data source, it is important to understand not just what the various fields are, but also how certain values were computed and whether the dataset could be used for the desired purpose. Provenance descriptions can contain all this information, but is usually far too much detailed information for a user to be able to make use of. Additionally, proprietary concerns and privacy limits may limit what can be disclosed. The idea of a nutritional label has been proposed by us, and independently by others, as a way to capture succinctly a small amount of critical information required to determine fitness for use. The challenge is that the information that must be captured depends on the intended use.

We have developed *RankingFacts* [58], the first prototype of an automatically computed nutritional label that helps users interpret algorithmic rankers. The work on a user-facing nutritional label prototype motivated a deeper inquiry into fairness and diversity in set selection and ranking [48, 56, 57], and on designing fair and stable ranking schemes [1, 2]. We also continued this work to compute properties to characterize data sets [49], and to succinctly capture correlation between attributes [32].

Finally, most individuals affected by data-driven systems likely have many other things going on in their lives. So, they may have limited time and attention that they wish to devote to data details. This makes it important that results and data be presented fairly, in a manner that leads to good understanding. Otherwise, inequity in attention availability can lead to errors and misunderstanding. To address such questions, we have initiated a stream of work in cherry-picking [3].

3.4 Outcome equity

Controlling for inequity during processing does not guarantee improvements in quality of life and societal outcomes, either in aggregate or at the individual level, due to, for example, *emergent bias* [18]. It is, therefore, important to monitor and mitigate unintended consequences for any groups affected by a system after deployment, directly or indirectly.

Outcome equity refers to downstream unanticipated consequences outside the direct control of the system — evaluation of these consequences pertains directly to the socio-political notion of equity focusing on equality of outcome. For example, families rejected Boston's optimized bus route system due to disruption of their schedules, despite the system's improvement in both resource management and equity.

It take time, effort, and expense to build a model. In consequence, models developed in one context are often used in another. Such model transfer has to be done with care. We have used 3D CNNs to generalize predictions in the urban domain [54]. We have shown that fairness adjustments applied to integrated representations (via adversarial models that attempt to learn the protected attribute [30]) outperform methods that apply fairness adjustments on the individual data sets [55].

The equity of a data-intensive system can be difficult to maintain over time [27, 34], due to *distribution shifts* [7, 22, 29, 41] that can reduce performance, force periodic retraining, and generally undermine trust. Techniques similar to the transferability methods, described in the preceding paragraph, can help.

To minimize outcome inequity, data-driven systems must be accountable. Accountability requires public disclosure. For example, a job seeker must be informed which qualifications or characteristics were used by the tool, and why these are considered job-relevant [46, 47].

But accountability is not enough in itself: the data subject also should have recourse. We seek contestability by design [33], an active principle that goes beyond explanation and focuses on user engagement, fostering user understanding of models and outputs, and collaboration in systems design [17, 26, 43]. Our goal is to empower users to question algorithmic results, and thereby to correct output inequities where possible.

As a starting point, consider credit scores: a simple tool that has existed for years in the US and in many other countries. A myriad of data sources report on your paying what you owe, and these reports are aggregated into a credit score, which you can see. You have some sense of what goes into building a good score, even though the specific details may not be known. More importantly, you can see what has been reported about you by your creditors and there is a process to challenge errors. The system is far from perfect, but most data-driven systems today are much worse in so many respects, including in particular their mechanisms for providing accountability and recourse.

4 CONCLUSION

Data equity issues are pervasive but subtle, requiring *holistic consideration* of the socio-technical systems that induce them (as opposed to narrowly focusing on the technical components and tasks alone), and of the contexts in which such systems operate. The richness of issues surrounding equity cannot be addressed by framing it as a narrow, situational facet of "final mile" learning systems. We need a *socio-technical framing* that shifts equity considerations upstream to the data infrastructure, combines technical and societal perspectives, and allows us to reason about the proper role for technology in promoting equity while linking to emergent social and legal contexts. This type of approach is rapidly gaining traction in global technology policy [12]. From a technology perspective, we must appreciate that multiple data sets are processed in a complex workflow, with numerous design and deployment choices enroute [23]. Additionally, our socio-technical framing mandates engagement with stakeholders before, during, and after any technology development, affords operationalization of socio-technical equity, as it emphasizing their lived experience as design expertise. It therefore centers *intersectionality*, a framework that focuses on how the interlocking systems of social identity (race, class, gender, sexuality, disability) combine into experiences of privilege and oppression [11, 13, 14, 20]. This framing expands data sciences' existing interpretation of intersectionality from external classification, often a political act [8, 9], to active involvement of those who are classified.

In this extended abstract, we have identified four facets of data equity, each of which must be addressed by data equity systems. For our ongoing work in this direction, please visit our project website at https://midas.umich.edu/FIDES.

ACKNOWLEDGMENT

This work was supported in part by the US National Science Foundation, under grants 1934405, 1934464, and 1934565.

REFERENCES

[1] Abolfazl Asudeh, H.V. Jagadish, Julia Stoyanovich, and Gautam Das. 2019. Designing Fair Ranking Schemes. In *ACM SIGMOD*.

[2] Abolfazl Asudeh, H. V. Jagadish, Gerome Miklau, and Julia Stoyanovich. 2018. On Obtaining Stable Rankings. *PVLDB* 12, 3 (2018), 237–250. http://www.vldb.org/pvldb/vol12/p237-asudeh.pdf

[3] Abolfazl Asudeh, H. V. Jagadish, You Wu, and Cong Yu. 2020. On detecting cherry-picked trendlines. *Proceedings of the VLDB Endowment* 13, 6 (2020), 939–952.

[4] Abolfazl Asudeh, Zhongjun Jin, and H. V. Jagadish. 2019. Assessing and remedying coverage for a given dataset. In *IEEE International Conference on Data Engineering*. IEEE, 554–565.

[5] Solon Barocas and Andrew Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104, 3 (2016), 671–732.

[6] Ruha Benjamin. 2019. Race after technology: Abolitionist tools for the new jim code. *Social Forces* (2019).

[7] Steffen Bickel, Michael Brückner, and Tobias Scheffer. 2009. Discriminative Learning Under Covariate Shift. *J. Mach. Learn. Res.* 10 (2009), 2137–2155. https://dl.acm.org/citation.cfm?id=1755858

[8] Geoffery C. Bowker and Susan Leigh Star. 2000. *Sorting Things out: Classification and Its Consequences.* MIT Press, Cambridge, MA, USA.

[9] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Conference on Fairness, Accountability and Transparency, FAT 2018, 23-24 February 2018, New York, NY, USA*. 77–91. http://proceedings.mlr.press/v81/buolamwini18a.html

[10] Irene Chen, Fredrik D Johansson, and David Sontag. 2018. Why is my classifier discriminatory?. In *Advances in Neural Information Processing Systems*. 3539–3550.

[11] P. H. Collins. 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge, New York, NY.

[12] Council of Europe. 2020. Ad Hoc Committee on Artificial Intelligence (CAHAI), Feasibility Study. https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.

[13] Kimberle Crenshaw. 1989. Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum* 1 (1989), 139–167.

[14] Kimberle Crenshaw. 1991. Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color. *Stanford Law Review* 43, 6 (1991), 1241–1299. http://www.jstor.org/stable/1229039

[15] Ronald Dworkin. 1981. What is Equality? Part 1: Equality of Welfare. *Philosophy and Public Affairs* 10, 4 (1981), 185–246.

[16] Ronald Dworkin. 1981. What is Equality? Part 2: Equality of Resources. *Philosophy and Public Affairs* 10, 4 (1981), 283–345.

[17] Motahhare Eslami. 2017. Understanding and Designing around Users' Interaction with Hidden Algorithms in Sociotechnical Systems. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017, Portland, OR, USA, February 25 - March 1, 2017, Companion Volume*, Charlotte P. Lee, Steven E. Poltrock, Louise Barkhuus, Marcos Borges, and Wendy A. Kellogg (Eds.). ACM, 57–60. http://dl.acm.org/citation.cfm?id=3024947

[18] Batya Friedman and Helen Nissenbaum. 1996. Bias in Computer Systems. *ACM Trans. Inf. Syst.* 14, 3 (1996), 330–347. https://doi.org/10.1145/230538.230561

[19] Dominique DuBois Gilliard. 2018. *Rethinking incarceration: Advocating for justice that restores*. InterVarsity Press.

[20] P.L. Hammack. 2018. *The Oxford Handbook of Social Psychology and Social Justice*. Oxford University Press. https://books.google.com/books?id=ZY9HDwAAQBAJ

[21] Anna Lauren Hoffmann. 0. Terms of inclusion: Data, discourse, violence. *New Media & Society* 0, 0 (0), 1461444820958725. https://doi.org/10.1177/1461444820958725 arXiv:https://doi.org/10.1177/1461444820958725

[22] Jiayuan Huang, Alexander J. Smola, Arthur Gretton, Karsten M. Borgwardt, and Bernhard Schölkopf. 2006. Correcting Sample Selection Bias by Unlabeled Data. In *Advances in Neural Information Processing Systems 19, Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 4-7, 2006*, Bernhard Schölkopf, John C. Platt, and Thomas Hofmann (Eds.). MIT Press, 601–608. http://papers.nips.cc/paper/3075-correcting-sample-selection-bias-by-unlabeled-data

[23] H. V. Jagadish, Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M Patel, Raghu Ramakrishnan, and Cyrus Shahabi. 2014. Big data and its technical challenges. *Commun. ACM* 57, 7 (2014), 86–94.

[24] H. V. Jagadish, Julia Stoyanovich, and Bill Howe. 2021. COVID-19 Brings Data Equity Challenges to the Fore. *ACM Digital Government: Research and Practice* 2, 2 (2021).

[25] Zhongjun Jin, Mengjing Xu, Chenkai Sun, Abolfazl Asudeh, and H. V. Jagadish. 2020. MithraCoverage: A System for Investigating Population Bias for Intersectional Fairness. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*. ACM, 2721–2724.

[26] Daniel Kluttz et al. 2020. Shaping Our Tools: Contestability as a Means to Promote Responsible Algorithmic Decision Making in the Professions. In *After the Digital Tornado: Networks, Algorithms, Humanity*.

[27] Arun Kumar, Robert McCann, Jeffrey F. Naughton, and Jignesh M. Patel. 2015. Model Selection Management Systems: The Next Frontier of Advanced Analytics. *SIGMOD Rec.* 44, 4 (2015), 17–22. https://doi.org/10.1145/2935694.2935698

[28] Yin Lin, Yifan Guan, Abolfazl Asudeh, and H. V. Jagadish. 2020. Identifying insufficient data coverage in databases with multiple relations. *Proceedings of the VLDB Endowment* 13, 12 (2020), 2229–2242.

[29] Zachary C. Lipton, Yu-Xiang Wang, and Alexander J. Smola. 2018. Detecting and Correcting for Label Shift with Black Box Predictors. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018 (Proceedings of Machine Learning Research)*, Jennifer G. Dy and Andreas Krause (Eds.), Vol. 80. PMLR, 3128–3136. http://proceedings.mlr.press/v80/lipton18a.html

[30] David Madras, Elliot Creager, Toniann Pitassi, and Richard Zemel. 2018. Learning adversarially fair and transferable representations. *arXiv preprint arXiv:1802.06309* (2018).

[31] Charles W Mills. 2014. *The racial contract*. Cornell University Press.

[32] Yuval Moskovitch and H. V. Jagadish. 2020. COUNTATA: dataset labeling using pattern counts. *Proceedings of the VLDB Endowment* 13, 12 (2020), 2829–2832.

[33] Deirdre K. Mulligan and Kenneth A. Bamberger. 2019. Procurement as Policy: Administrative Process for Machine Learning. *Berkeley Technology Law Journal* 34, 3 (2019), 773–852.

[34] Neoklis Polyzotis, Sudip Roy, Steven Euijong Whang, and Martin Zinkevich. 2018. Data Lifecycle Challenges in Production Machine Learning: A Survey. *SIGMOD Rec.* 47, 2 (2018), 17–28. https://doi.org/10.1145/3299887.3299891

[35] John Rawls. 1971. *A theory of justice*. Harvard University Press.

[36] John E Roemer and Alain Trannoy. 2015. Equality of opportunity. In *Handbook of income distribution*. Vol. 2. Elsevier, 217–300.

[37] Babak Salimi, Luke Rodriguez, Bill Howe, and Dan Suciu. 2019. Capuchin: Causal Database Repair for Algorithmic Fairness. *CoRR* abs/1902.08283 (2019). arXiv:1902.08283 http://arxiv.org/abs/1902.08283

[38] Babak Salimi, Luke Rodriguez, Bill Howe, and Dan Suciu. 2019. Interventional fairness: Causal database repair for algorithmic fairness. In *Proceedings of the 2019 International Conference on Management of Data*. 793–810.

[39] Sebastian Schelter, Felix Biessmann, Tim Januschowski, David Salinas, Stephan Seufert, Gyuri Szarvas, Manasi Vartak, Samuel Madden, Hui Miao, Amol Deshpande, et al. 2018. On Challenges in Machine Learning Model Management. *IEEE Data Eng. Bull.* 41, 4 (2018), 5–15.

[40] Sebastian Schelter, Yuxuan He, Jatin Khilnani, and Julia Stoyanovich. 2020. FairPrep: Promoting Data to a First-Class Citizen in Studies on Fairness-Enhancing Interventions. In *EDBT*, Angela Bonifati, Yongluan Zhou, Marcos Antonio Vaz Salles, Alexander Böhm, Dan Olteanu, George H. L. Fletcher, Arijit Khan, and Bin Yang (Eds.). OpenProceedings.org, 395–398. https://doi.org/10.5441/002/edbt.2020.41

[41] Sebastian Schelter, Tammo Rukat, and Felix Bießmann. 2020. Learning to Validate the Predictions of Black Box Classifiers on Unseen Data. In *Proceedings of the 2020 International Conference on Management of Data, SIGMOD Conference 2020, online conference [Portland, OR, USA], June 14-19, 2020*, David Maier, Rachel Pottinger, AnHai Doan, Wang-Chiew Tan, Abdussalam Alawini, and Hung Q. Ngo (Eds.). ACM, 1289–1299. https://doi.org/10.1145/3318464.3380604

[42] David Sculley, Gary Holt, Daniel Golovin, Eugene Davydov, Todd Phillips, Dietmar Ebner, Vinay Chaudhary, Michael Young, Jean-Francois Crespo, and Dan Dennison. 2015. Hidden technical debt in machine learning systems. In *Advances in neural information processing systems*. 2503–2511.

[43] Judith Simon. 2015. Distributed Epistemic Responsibility in a Hyperconnected Era. In *The Online Manifesto: Being Human in a Hyperconnected Era*. 145–159. https://doi.org/10.1007/978-3-319-04093-6_17

[44] Dean Spade. 2015. *Normal Life: Administrative Violence, Critical Trans Politics, and the Limits of Law*. Duke University Press. http://www.jstor.org/stable/j.ctv123x7qx

[45] Dean Spade. 2015. *Normal life: Administrative violence, critical trans politics, and the limits of law*. Duke University Press.

[46] Julia Stoyanovich. 2020. Testimony of Julia Stoyanovich before New York City Council Committee on Technology regarding Int 1894-2020, Sale of automated employment decision tools. https://dataresponsibly.github.io/documents/Stoyanovich_Int1894Testimony.pdf.

[47] Julia Stoyanovich, Bill Howe, and H.V. Jagadish. 2020. Responsible Data Management. *PVLDB* 13, 12 (2020), 3474–3489. https://doi.org/10.14778/3415478.3415570

[48] Julia Stoyanovich, Ke Yang, and H. V. Jagadish. 2018. Online Set Selection with Fairness and Diversity Constraints. In *Proceedings of the 21th International Conference on Extending Database Technology, EDBT 2018, Vienna, Austria, March 26-29, 2018*. 241–252. https://doi.org/10.5441/002/edbt.2018.22

[49] Chenkai Sun, Abolfazl Asudeh, HV Jagadish, Bill Howe, and Julia Stoyanovich. 2019. Mithralabel: Flexible dataset nutritional labels for responsible data science. In *Proceedings of the ACM International Conference on Information and Knowledge Management*. ACM, 2893–2896.

[50] Supreme Court of the United States. 1995. ADARAND CONSTRUCTORS, Inc. v. PEÑA, 515 U.S. 200 (1995), No.93-1841. https://supreme.justia.com/cases/federal/us/515/200/#tab-opinion-1959723.

[51] Supreme Court of the United States. 1996. UNITED STATES v. VIRGINIA, 518 U.S. 515 (1996), No.94-1941. https://supreme.justia.com/cases/federal/us/515/200/#tab-opinion-1959723.

[52] Supreme Court of the United States. 2009. Ricci v. DeStefano (Nos. 07-1428 and 08-328), 530 F. 3d 87, reversed and remanded. https://www.law.cornell.edu/supct/html/07-1428.ZO.html.

[53] Ian Kennedy Timothy A. Thomas, Ott Toomet and Alex Ramiller. [n.d.]. The State of Evictions: Results from the University of Washington Evictions Project. https://evictions.study/. accessed 25-Apr-2019.

[54] An Yan and Bill Howe. 2019. FairST: Equitable Spatial and Temporal Demand Prediction for New Mobility Systems. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 552–555.

[55] An Yan and Bill Howe. 2021. EquiTensors: Learning Fair Integrations of Heterogeneous Urban Data. In *ACM SIGMOD*.

[56] Ke Yang, Joshua R. Loftus, and Julia Stoyanovich. 2020. Causal intersectionality for fair ranking. *CoRR* abs/2006.08688 (2020). arXiv:2006.08688 https://arxiv.org/abs/2006.08688

[57] Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In *Proceedings of the 29th International Conference on Scientific and Statistical Database Management, Chicago, IL, USA, June 27-29, 2017*. 22:1–22:6. https://doi.org/10.1145/3085504.3085526

[58] Ke Yang, Julia Stoyanovich, Abolfazl Asudeh, Bill Howe, H. V. Jagadish, and Gerome Miklau. 2018. A Nutritional Label for Rankings. In *Proceedings of the 2018 International Conference on Management of Data, SIGMOD Conference 2018, Houston, TX, USA, June 10-15, 2018*. 1773–1776. https://doi.org/10.1145/3183713.3193568