KRITIKE VOLUME NINETEEN NUMBER ONE (MARCH 2025) 25–49

Article

Ethics and AI Classification: Power, Bias, and the Technical Worldview[1]

Nahum Brown

Abstract: This article offers a critical response to Kate Crawford's discussion of AI classification. In her celebrated book *Atlas of AI* (2021), Crawford examines the ethical consequences of biases, reifications, and oversimplifications that emerge from the decision-making behind AI taxonomies. I argue that there is an ambiguity in Crawford's work about whether she believes that the problems we face with AI are solvable or not, in other words, whether an ethics of AI classification is possible. I investigate the source of this ambiguity by outlining multiple trunks and branches of arguments for and against the claim that, however complicated a solution may be, AI classification distortions are fixable. By uncovering a total of five arguments within her classification chapter, my aim is to clarify and develop Crawford's thesis about the encoding of power while also reflecting on why this question of an ethics is unresolved in her work.

Keywords: AI classification, AI ethics, power, the technical worldview

One of the most celebrated recent books about the rapidly emerging topic of AI ethics is Kate Crawford's *Atlas of AI* (2021).[2] This book is a penetrating and widely-read interdisciplinary piece that combines sociology, philosophy, and environmental studies with a specialist's knowledge of the technical underpinnings of machine learning. While establishing specific theses chapter by chapter, the book effectively engages with a full range of pressing AI issues. These include the physical and material costs of the everyday implementation of AI on Earth (chapter 1); the difficulties that workplaces and labor forces face (chapter 2); conflicts arising from data collection and consent (chapter 3); biases and oversimplifications that come as a consequence of classification processes (chapter 4); the affect

[1] The research for this article is partially supported by a grant from the AI Ethics from the Ground Up Project, National Research Council of Thailand.

[2] Kate Crawford, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (New Haven: Yale University Press, 2021).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

of the human face in terms of face recognition technology (chapter 5); and AI power encoding when viewed from the vantage point of the state (chapter 6). As her title suggests, Crawford describes the book as an atlas, with the chapters performing this "cartographic approach."[3] The book is a "collection of disparate parts, with maps that vary in resolution,"[4] giving the reader the ability to range over and zoom in on a myriad of topics.

This article takes Crawford's cartographic approach seriously by zeroing in on one of the most compelling places in the atlas: classification. In chapter 4 of her book, Crawford analyzes serious ethical problems arising from AI classification from blatant sexist, racist, and ableist biases to the more tenacious systemic prejudices that underlie taxological divisions. Reminiscent of Foucault's theory of power, Crawford argues that the extraction, formatting, and labeling of data are never neutral. Instead, power is encoded in these processes which leads to the oversimplification and reification of reality, often with the unintended consequences of upholding conservative religious, political, and essentialist worldviews.

My primary thesis is to argue that there is an ambiguity in Crawford's work about whether she believes that the problems we face with AI are solvable or not, in other words, whether an ethics of AI classification[5] is possible. This amounts to a question about whether an ethics of AI classification is possible at all, in the sense that whether there are normative claims that would help to resolve the issues of classification, or whether there is really nothing that one can do. As a means to both critically analyze this ambiguity in Crawford's work and also to develop her project further, I outline the most plausible arguments for and against the assertion that one can do an ethics of AI classification: arguments that range from the stance that the biases produced from this process are immanently fixable to the position that it would take resolving systemic injustices to resolve classification biases; from the argument that category divisions are a form of power to the claim that AI classification requires the reification of a technical worldview. In total, I outline five arguments that Crawford either explicitly or implicitly articulates.

Crawford's work focuses primarily on ImageNet's massive taxonomy project as an extended example of AI classification problems. After

[3] Crawford, *Atlas of AI*, 10.

[4] *Ibid.*, 9.

[5] Crawford defines "AI classification" as a "core practice in Artificial Intelligence" (*ibid.*, 127) that contributes to training AI for the implementation of face and object recognition. Data are extracted and labelled for the purpose of developing training models, especially in the form of images under the headings of categories. These taxonomical systems are then used to "predict human identity, commonly using binary gender, essentialized racial categories, and problematic assessments of character and credit worthiness" (*ibid.*, 17).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

an initial conception in 2006, Stanford professor Li Fei-Fei developed ImageNet in 2009 with the aim to "map out the entire world of objects"[6] as a means to create a tool for training data sets. ImageNet is a "large-scale image ontology"[7] that has led to "multiple breakthroughs in object recognition."[8] The ImageNet team made their goals clear in a conference in the same year:

> The digital era has brought with it an enormous explosion of data. The latest estimations put a number of more than 3 billion photos on Flickr, a similar number of video clips on YouTube and an even larger number for images in the Google Image Search database. More sophisticated and robust models and algorithms can be proposed by exploiting these images, resulting in better applications for users to index, retrieve, organize, and interact with these data.[9]

ImageNet at first struggled to find a method for how to extract, format, and label so many images. After realizing that hiring undergraduate students at ten dollars an hour would take nearly a century to complete the project, Li Fei-Fei discovered Amazon Mechanical Turk. This gave her a way to organize a huge low-cost labor force, capable of classifying approximately fifty images per minute per worker.[10] "Suddenly we found a tool that could scale," Li Fei-Fei says, "that we could not possibly dream of by hiring

---

[6] Dave Gershgorn, "Data That Transformed AI Research—and Possibly the World," in *Quartz* (26 July 2017), <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world>.

[7] In philosophy, "ontology" usually refers to the study or science of being, either as its own abstract branch of philosophy or as a subdivision of metaphysics, including ontologies as diverse as Plato's theory of the forms, Hegel's study of being in *The Science of Logic*, and Heidegger's critical analysis of the history of the question of being in *Being and Time*. However, Li Fei-Fei uses the term differently. By "ontology," she refers to the expansive project of cataloging and matching as many nouns as possible with images. This roughly connects to a specific philosophical interpretation of ontology as the study of object existence and recognition but departs from most continental and historical accounts of ontology. Li Fei-Fei uses the term to describe ImageNet's initial aim of being as comprehensive as possible by providing every noun imaginable within the taxonomy. The subsequent 2019 redactions of images based on ethical objections caused limitations for these initially grand claims of doing a total "ontology" of all nouns.

[8] Kaiyu Yang et al., "Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy," in *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020).

[9] Crawford, *Atlas of AI*, 107.

[10] *Ibid.*, 108.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

Princeton undergrads."[11] This enabled ImageNet to create thousands of categories and sort millions of images:

> There were categories for apples and airplanes, scuba divers and sumo wrestlers. But there were cruel, offensive, and racist labels, too: photographs of people were classified into categories like "alcoholic," "ape-man," "crazy," "hooker," and "slant eye." All of these terms were imported from WordNet's lexical database and given to crowdworkers to pair with images. Over the course of a decade, ImageNet grew into a colossus of object recognition for machine learning and a powerfully important benchmark for the field.[12]

It took ten years before enough public criticism—brought on partially by the artist Trevor Paglen's ImageNet Roulette project, which let people "upload images to see how they would be classified"[13]—caused ImageNet to significantly revise its taxonomy. ImageNet subsequently removed approximately 56% of the "Person" category (one of its nine top-level categories) thus attempting to reshape the initially problematic divisions and compartments from its taxonomy:

> The creators of ImageNet published a paper… titled "Towards Fairer Datasets" that sought to "remove unsafe synsets." They asked twelve graduate students to flag any categories that seemed unsafe because they were either "inherently offensive" (for example, containing profanity or "racial or gender slurs") or "sensitive" (not inherently offensive but terms that "may cause offense when applied inappropriately, such as the classification of people based on sexual orientation and religion"). While this project sought to assess the offensiveness of ImageNet's categories by asking graduate students, the authors nonetheless continue to support the automated classification of people based on photographs despite the notable problems.[14]

---

[11] *Ibid.*

[12] *Ibid.*, 108–109.

[13] *Ibid.*, 141. Crawford describes this project as a "multiyear research collaboration" that she did with Paglen. See Crawford, *Atlas of AI*, 257.

[14] *Ibid.*, 141–142.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

The ImageNet team addressed their problematic decisions about AI classification by hiring a small group of graduate students to revise the categories and images within labels that appeared to be biased or offensive.[15] By responding this way, ImageNet reinforced the underlying ethical assumption that such a taxonomy can be made neutral or bias-free and that massive amounts of data can be sorted and classified in a fair and unassuming way that approximately matches the reality of this content. Crawford is critical of this response from ImageNet:

> The focus on the hateful categories is not wrong, but it avoids addressing questions about the workings of the larger system. The entire taxonomy of ImageNet reveals the complexities and dangers of human classification. While terms like "microeconomist" or "basketball player" may initially seem less concerning than the use of labels like "spastic," "unskilled person," "mulatto," or "redneck," when we look at the people who are labeled in these categories, we see many assumptions and stereotypes, including race, gender, age, and ability. In the metaphysics of ImageNet, there are separate image categories for "assistant professor" and "associate professor"—as though once someone gets a promotion, her or his biometric profile would reflect the change in rank.[16]

Why exactly does Crawford think that ImageNet's response is unsatisfactory? Is she claiming that the twelve graduate students hired to revise the categories did a bad job, but that by hiring a full team of experts to redesign and limit the top-level category "Person," it is still possible to correct the biases inherent in ImageNet's initial 2009 taxonomy of all nouns? Does she have an issue with the approaches that the authors of "Towards Fairer Datasets" propose to use to fix these biases—approaches that identify black box problems, ethically questionable data, privacy concerns, imbalanced representations of demographic groups, and algorithmic development responses?[17] Or, is Crawford making the larger claim that it is difficult or even impossible for classification to accurately compartmentalize the complexity of reality?

---

[15] See also Yang et al., "Towards Fairer Datasets," section 4.3, 550.

[16] Crawford, *Atlas of AI*, 142.

[17] See Yang et al., "Towards Fairer Datasets," section 2, 548–549.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

While she clearly rejects ImageNet's 2019 revisions, Crawford articulates an ambiguous range of different answers to this ethical question of whether AI classification could be redesigned in a fair and unbiased way. Sometimes she appears to endorse the claim that there are real solutions, even if these solutions are hard to reach. She suggests this when she emphasizes how daunting the task of revision was for the graduate students (implying that it is nevertheless possible). She also suggests this at the end of the chapter when she discusses a "fluid dynamics" strategy to classification based on Geoffrey C. Bowker and Susan Leigh Star's seminal source text for AI classification *Sorting Things Out: Classification and Its Consequences*.[18] But more often Crawford appears to claim that solutions are impossible to implement, either because it is impossible to make data neutral, because classification is always the encoding of power, or because there is an irretractable technical worldview that comes along with the classification process:

> Images—like all forms of data—are laden with all sorts of potential meanings, irresolvable questions, and contradictions. In trying to resolve these ambiguities, ImageNet's labels compress and simplify complexity. The focus on making training sets "fairer" by deleting offensive terms fails to contend with the power dynamics of classification and precludes a more thorough assessment of the underlying logics. Even if the worst examples are fixed, the approach is still fundamentally built on an extractive relationship with data that is divorced from the people and places from whence it came. Then it is rendered through a technical worldview that seeks to fuse together a form of singular objectivity from what are complex and varied cultural materials.[19]

In this passage, Crawford outlines multiple theses about the nature of classification. Building from her claim that "there are no neutral categories in ImageNet,"[20] she says that images contain contradictory interpretive meanings, multiplicities which are, nevertheless, suppressed under the abstraction process of data collection. She also claims that when ImageNet labels these images, it oversimplifies the meaning of that which it classifies, inferring from this that everyday reality has a great, or even infinite,

[18] Crawford, *Atlas of AI*, 148.
[19] *Ibid.*, 143.
[20] *Ibid.*, 142.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

complexity to it but that the process of AI classification causes a distortion in the form of reduction to simplicity. She then outlines a further thesis about the relationship between classification and power, suggesting that there is an inherent encoding of power embedded within the taxonomy creators' decisions about what categories should exist and what content should be placed within them. She then turns to the issue of extracting content from its context and background, from its communities and places, implying yet another thesis about the impossibility of cataloging things from within their context. She also mentions that there is a technical worldview that comes along with AI classification, drawing on yet another point about the reduction of quality to quantity.

Crawford sometimes even seems to hold the position that the act of classifying things is in itself inherently reductive, as when she writes: "to impose order onto an undifferentiated mass, to ascribe phenomena to a category— that is, to name a thing—is in turn a means of reifying the existence of that category."[21] This makes it sound as if the biases and reification processes that she diagnoses in AI classification are actually more fundamental and are already present in pre-AI classification systems. However, even in statements like these, we can still recognize a form of authority that is distinctively problematic about AI classification in contrast to pre-AI classification systems. AI classification projects the mystique of objectivity that Crawford elsewhere describes as "enchanted determinism."[22] In other words, it projects a false sense of security that the contingent and sometimes arbitrary divisions of AI taxonomy are objectively true. Because AI classification has the anonymous aura of being an expert authority on matters of division and labeling things, it presents a specious objectivity that is difficult for most people to resist. The main reason why AI classification produces this mystique is because the source of the decision-making is hidden behind obscure black box technical processes. The labels and categories of AI classification appear as if they were the object divisions caused by a higher power of nature or objective reality,[23] when in fact they are actually the product of teams of people and companies with political agendas and subjective interests. Technochauvinism also contributes to the mystique of AI classification objectivity. Meredith Broussard defines technochauvinism as "a kind of bias that considers computational solutions

[21] *Ibid.*, 139.

[22] Alexander Campolo and Kate Crawford, "Enchanted Determinism: Power without Responsibility in Artificial Intelligence," in *Engaging Science, Technology, and Society*, 6 (2020), 3.

[23] Cathy O'Neil describes this as the effect of "algorithmic gods." See Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016), 14.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

to be superior to all other solutions."[24] Because of this false sense of objectivity, AI classification exacerbates a reductive process that already appears latently in pre-AI taxonomical systems.

The five arguments that I draw from Crawford's work can be organized under two trunks: normative and non-normative. Proponents of the normative trunk claim that there are practical, reachable solutions and that, if we are more mindful enough, it is possible to redesign categories to accurately account for the myriad distinctions of reality. The three categories that fall under the normative trunk are (1) the argument from immanent resolution, (2) the systemic argument, and (3) the people-are-irreducible-to-things argument. In contrast, proponents of the non-normative trunk claim that it is impossible to match categories accurately with reality. This is, for one reason or another, because the act of classifying things necessarily distorts the things that it classifies. There are two arguments that fall under the non-normative trunk: (4) the power argument and (5) the technical worldview argument.[25]

Since they have significantly different consequences and notably different implications for an ethics of AI classification, it is important that we disambiguate these five arguments in Crawford's work. By doing so, we gain a clearer view of the AI classification normativity debate as a whole—the debate about whether and by what means an ethics is possible—and thus expose the ethical complexity behind the process of dividing and cataloging things with AI. Not all of Crawford's arguments have the same structural implication. Some imply that there are specific strategies that we could employ to resolve biases while others imply that, no matter how much we try to mitigate them, biases are inevitable and unresolvable. The main advantage of disambiguating these five arguments is that we put ourselves in a better position to evaluate the various and sometimes conflicting implications in Crawford's work.

The conclusion I arrive at is twofold. On the one hand, the reason for explicating the arguments is to give evidence for which position Crawford ultimately thinks is best. Although a case can be made for any of these five

[24] Meredith Broussard, *More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech* (Cambridge: MIT Press, 2023), 2.

[25] While there is mutual exclusion between the normative and non-normative trunks, one can find an overlap within the branches of each. The point of these branch distinctions is not to establish the purity of each as categorically separate from the others but to account for the various ways to view classification divisions. I also recognize that because of the vision and subtle insights of Crawford's work, some readers might disagree with me about how the arguments should best be cataloged, claiming either that there are more or less than five arguments or that some arguments should be presented differently to more accurately reflect Crawford's views. This is often how it goes with interpretive readings of texts that are as rich and nuanced as Crawford's.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

arguments at different moments in her text, Crawford seems to offer the most emphatic defense of the fourth argument, i.e., from power. On the other hand, my conclusion is also that part of the value of Crawford's analysis comes down to her ability to outline the range and complexity of these five different arguments. There is at least implicit textual evidence for each argument in Crawford's book. Her apparent indecision about whether AI classification distortions are resolvable or not expresses an ambiguity that is relatable and productive in that it helps to enumerate various philosophical commitments and assumptions embedded within the nature of dividing and naming things. With this in mind, I have two types of readers: readers who are directly engaged with Crawford's work can view this paper as an explicative and critical companion piece to the fourth chapter of *Atlas of AI*. But there are also more general readers who are interested in the relationship between ethics and classification and are concerned about the dangers of abstraction, reification, and the oversimplification of reality.

Before we look at the arguments, let's establish some basic definitions. Crawford's analysis builds from Bowker and Star's definition of the terms "classification" and "classification system" in:

> A classification is a spatial, temporal, or spatia-temporal segmentation of the world. A "classification system" is a set of boxes (metaphorical or literal) into which things can be put to then do some kind of work—bureaucratic or knowledge production.[26]

Bias is the other key term for this study.[27] In a recent article about the technical grammar of AI, Matteo Pasquinelli lays out a helpful set of distinctions about three types of bias: "world bias," "data bias," and "algorithmic bias." He defines "world bias" as the unjust or arbitrary treatment of underrepresented groups of people that is "already apparent in society before technological intervention."[28] World biases include racist, sexist, ableist, classist, and other forms of discrimination based on cultural and historical prejudice. Although this form of bias predates the recent emergence of machine learning technology, the datasets of AI further reify these already ingrained stereotypes. In contrast to world bias, the other two types of biases are directly about AI. "Data bias" refers to a form of

[26] Geoffrey C. Bowker and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge: MIT Press, 1999), 10.

[27] Crawford discusses the historical etymology and usage of the word "bias." See Crawford, *Atlas of AI*, 134.

[28] Matteo Pasquinelli, "How a Machine Learns and Fails – A Grammar of Error for Artificial Intelligence," in *Spheres: Journal for Digital Cultures*, 5 (2019), 9.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

discrimination that comes directly from the "capturing, formatting, and labelling of data from the training dataset."[29] Pasquinelli points to data labelling as that part of the dataset training process that is most susceptible to this type of bias: "Universities, corporations, and military agencies build training datasets with rough and cheap labour. They often make use of old and conservative taxonomies causing a distorted view of world cultures and diversities."[30] There is also a third type of bias, "algorithmic bias," caused by "computational errors, information compression and the approximation of techniques of machine learning algorithms."[31] The opacity in the design and implementation of algorithms within training sets further exacerbates inequalities that are already present in data. As we will see, these three forms of bias are very prevalent in the arguments I will outline; however, some of the arguments refer to one or another of these biases more than others.

The Normative Trunk: The Resolvable, Systemic, and Irreducible-to-Things Arguments

The normative trunk contains a set of arguments drawn from the assertion that AI classification biases are resolvable. These arguments are optimistic about the possibility that AI classifications could eventually achieve an objectivity of object recognition. They are normative in the sense that they make claims about what we should do to fix biases and resolve taxonomical distortions. While normative claims can be applied more broadly to refer to how we should live and what we should do when faced with an ethical dilemma, I use the term "normative" in a more specific way to identify a set of arguments that are built from the assumption that it is both possible and desirable to solve classification problems. These arguments are normative in the sense that, because the biases are resolvable, it is our duty to resolve them. The creators of ImageNet who defend the 2019 revisions of the "Person" category fall into this category. Even though Crawford disagrees with their revisions, we can also interpret three strands of Crawford's own arguments to support the normative approach.

Argument 1: Immanently Resolvable

There are those who think that by simply revising the categories of our classifications, we can set up an adequate taxonomy of the objects of reality that succeeds at fairly representing them. Correcting the taxonomy

[29] *Ibid.*, 10.
[30] *Ibid*.
[31] *Ibid*.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

issues with ImageNet's "Person" top-level category might require a lot of funding to significantly overhaul the initially flawed, sometimes arbitrary, sometimes biased or even racist, typology of things—but this is nevertheless fixable if we pour enough resources into the project. Proponents of the resolvable interpretation support ImageNet's strategy of revising and redacting the categories of their taxonomy but feel that the task needs to be taken more seriously. Pertinent questions for this branch include: How much training data is really enough to develop fair training sets? What measures can we take to uncover and revise categories and images that are deemed to be ethically questionable? Can more effective guidelines be set up to deal with privacy concerns and black box barriers that come from consent and transparency issues? What sorts of expertise and how much funding would it take to more adequately address classification biases? Is there a way to re-configure the demographic distribution to reach underrepresented groups? How much of a factor is the sedimentation of problematic categories over time? Li Fei-Fei and her team at ImageNet have already addressed many of these questions in their 2019 article, but enthusiasts of the resolvable interpretation think that there is more work to be done.

This position has the advantage of being immediate and optimistic. There are real, practical solutions to the problems that arise from AI classification, and it is merely a matter of working out how we should fix these problems. In terms of the "Person" category, the compartments of the taxonomy need to be in touch with the spirit of the times and should be as sensitive and inclusive as possible so as not to further marginalize underrepresented or vulnerable groups of people. It might also be necessary to further revise the categories as time goes on. Annual or semi-regular revisions would help to keep the categories up-to-date and in line with the spirit of the times.

Although she sometimes entertains this position, Crawford for the most part rejects it.[32] While she recognizes that ImageNet has made real improvements and that the categories are relatively fairer than they were, she also argues that there are still inherent problems with the classification system and that there is not an immediate path for further revisions that would clearly work. Part of her argument is that categories associated with a function or social role are obviously problematic because connecting images to them is usually random and subjective:

> People are labeled by careers or hobbies: there are Boy Scouts, cheerleaders, cognitive neuroscientists, hairdressers, intelligence analysts, mythologists,

[32] Crawford, *Atlas of AI*, 140.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

> retailers, retirees, and so on. The existence of these categories suggests that people can be visually ordered according to their profession… ImageNet also contains categories that make no sense whatsoever for image classification such as Debtor, Boss, Acquaintance, Brother, and Color-Blind Person. These are all non-visual concepts that describe a relationship, be it to other people, to a financial system, or to the visual field itself.[33]

But part of her argument comes from a deeper objection to the basic assumption that it is possible to form an adequate objective representation of the categories of reality. Underlying the advantages of the resolvable interpretation is the assumption that fundamental categories unambiguously exist and help to constitute an objective reality that is both discernable and relatively stable, reliably ascertained from a timeless set of essences. Even if the administrators of ImageNet were to hire the most sensitive, tactful, woke experts to overhaul the classification system, decisions would have to be made about gender, race, class, etc., that make essentialist assumptions about the fixity of being.

In a few passages in the text, Crawford seems conflicted about whether it might still be possible to design a fairer, more objective taxonomical network of object recognition. That she views the 2019 revisions to have made at least some progress implies that this normative project could succeed at least partially. Nonetheless, since most of Crawford's analysis is taken up with stronger arguments about power and the consequences of dividing things from things, it is reasonable to conclude that she does not believe that classification biases are immanently resolvable.

Argument 2: Systemic

There is also the more entrenched systemic variation of the normative approach: people who view classification biases as historically embedded in the social fabric of communities. The distortions and biases of classification are fixable, but it would take much more than redactions and policy revisions to redesign the categories in objectively neutral ways. It would take systemic shifts to clear away the latent biases hidden within our everyday presumptions about reality. For example, it would take linguistic alterations of language to clear away sexist, racist, and ableist prejudices that are built into the everyday pronouns and nouns we use to speak. Or another example: it would take a reassessment of inheritance and other legal biases to correct

[33] *Ibid.*, 140–141.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

the systemic imbalance of wealth and historical economic marginalization of certain African American communities in the US.[34]

This variation also includes assumptions about rationality built into the divisions of the categories that privilege identity thinking, organizations based on hierarchies, and formal inference, while suppressing alternative conceptions of rationality, such as dialectic, dialetheist, and rhizomatic models of inference. For example, we can interpret Deleuze and Guattari to outline the rationality version of the systemic argument in the introduction to *A Thousand Plateaus* with their discussion of the arborescent and the rhizomatic.[35] Post-industrialized communities have for the most part prioritized arborescent organizations, privileging linear, rational, hierarchical structures while marginalizing rhizomatic organizations, which contain non-linear, non-rational multiplicities. Rhizomes nevertheless appear prevalently in nature and social life as the patterns of organization in, for example, ant colonies, wolf packs, cancer, capitalism, and the unconscious. Proponents of the systemic argument claim that it would take fundamental changes in the organization biases of contemporary communities to more fairly represent the suppressed non-rational principles that contribute to the constitution of reality.

Arguments from the systemic branch are similar to arguments from the resolvable branch in that both make normative claims about how we should re-design AI classification to resolve biases and match the truth of things. The main difference, however, is that while the resolvable branch aims at the immediate gratification of an imminent resolution, proponents of the systemic variation think that there are much larger historical and cultural barriers. Broussard explains this difference in the two arguments as the difference between conceiving of biases as glitches (resolvable) or as structurally embedded, while arguing throughout her book for the latter:

> Calling something a glitch means that it's a temporary blip, something unexpected but inconsequential. A glitch can be fixed. The biases embedded in technology are more than mere glitches; they're baked in from the beginning. They are structural biases, and they can't be addressed with a quick code update.[36]

---

[34] For discussions about racial bias and AI, see Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code* (Cambridge: Polity Press, 2019) and Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: NYU Press, 2018). See also Broussard, *More than a Glitch*.

[35] See Gilles Deleuze and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. by Brian Massumi (Minneapolis: University of Minnesota Press, 1987), 3–25.

[36] Broussard, *More than a Glitch*, 3.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

Proponents of the systemic argument believe that we are in a long generation-by-generation process of moral evolution. We have stripped away some of our moral prejudice. There are watershed moments in the US, for example, such as the women's suffrage movement and the civil rights movement, which have led to greater gender and racial equality. But there is still a long way to go. Generally, proponents of the systemic branch claim that there is a direct link between our moral evolution and the resolution of the classification biases inherent in AI taxonomy.

Crawford outlines the main ideas of systemic variation in her discussion of 19th-century phrenology where she focuses on Samuel Morton, one of the leading craniologists of the century and his assumptions about racial divisions based on skull measurements.[37] Morton's classification method was viewed at that time to be a scientifically rigorous objective study of race that contributed to a long historical tradition of racial prejudices.[38] In his book *The Mismeasurement of Man*, Stephen Jay Gould helped to debunk the measurement evidence Morton used.[39] When analyzing Gould's objections, Crawford is careful to clarify that the flaws in Morton's arguments are not only based on the applied problems of weak induction, biased sample selection, and the skewed measurements of skull weight data. The concept itself of drawing up a seemingly objective science of racial division based on craniology is deeply flawed. Some readers might view Crawford's discussion of phrenology to indirectly support the systemic variation. Gould's important critical study helped to discredit pseudo-scientific racial classifications. It opened up the possibility for systemic solutions to racial prejudice and thereby to fairer treatments of race. But since Crawford then transitions from Gould's criticisms of Morton to a larger claim about how classification is inherently political and inherently produces power, it is more reasonable to conclude that, for Crawford, AI classification biases are more deeply entrenched than even proponents of the systemic variation would admit. As with the resolvable interpretation, most of the arguments Crawford develops in her work reject these normative solutions.

Argument 3: People are Irreducible to Things

There is also a third normative argument to be found in Crawford's classification chapter: the argument that people are irreducible to things argument. Proponents of this argument believe that because people are moral

[37] Crawford, *Atlas of AI*, 123–126.
[38] *Ibid.*, 125.
[39] Stephen Jay Gould, *The Mismeasure of Man* (New York: W. W. Norton, 1996).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

agents and ends in themselves, they should be treated in fundamentally different ways from how we treat things. If we label people under the "Person" category in the same way that we label things under various "thing" categories, we risk undermining these fundamental differences and, thereby, reducing people to things. Crawford calls ImageNet an "object lesson… in what happens when people are categorized like objects."[40] One of her criticisms of ImageNet is that because they handle the two different categories in the same way, people are misrepresented as something that is primarily of use value.

Proponents of this argument believe that an AI classification system that tracks object recognition *in terms of objects* can be a successful project. However, they also claim that there are significant barriers, biases, and oversimplifications that arise from labeling people as such. Because of this double position, we should think of the people-are-irreducible-to-things argument as a limited and only partial sub-variation of the normative trunk. In contrast to the immanently resolvable and systemic interpretations, which fully embrace the idea that taxonomy biases are solvable, this argument works by separating people from things and recognizing that while we can catalog things successfully, we cannot be as successful with people.

This argument requires a premise about there being a real difference between people and things in terms of treatment and use value. To establish this premise, let's invoke Kant's distinction in the *Groundwork for the Metaphysics of Morals* between people—who are an end in themselves—and things—which are merely a means to some further end:

> The human being, and in general every rational being, exists as end in itself, not merely as means to the discretionary use of this or that will, but in all its actions, those directed toward itself as well as those directed toward other rational beings, it must always at the same time be considered as an end… The beings whose existence rests not on our will but on nature nevertheless have, if they are beings without reason, only a relative worth as means, and are called things; rational beings, by contrast, are called persons, because their nature already marks them out as ends in themselves, i.e., as something that may not be used merely as means.[41]

[40] Kate Crawford and Trevor Paglen, "Excavating AI: The Politics of Images in Machine Learning Training Sets," in *AI & Society*, 36 (2021), 1110.

[41] Immanuel Kant, *Groundwork for the Metaphysics of Morals* (New Haven: Yale University Press, 2002), 45–46.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

For Kant, there are two distinct means–ends relationships, one for people and one for things. While it is fine to treat things in instrumental ways, it is unethical to use people purely in terms of use value. Of course, other people are a means as well—the friend who drives me to work every day has instrumental value—but it would be wrong and reductive to view people only in this way. Proponents of the people-are-irreducible-to-things argument believe that ImageNet's current methods for classifying people endanger this reduction of people by placing them in vulnerable positions that overemphasize their use value.

There are at least two possible normative claims that could address this problem of reducing people to things: (1) we can attempt to refrain from classifying people at all or (2) we can attempt to redesign the person category to better foster the intrinsic value of people:

(1) Whenever possible, we should refrain from classifying people at all. The advantage of this interpretation is that by completely removing the people category altogether, we avoid most of the problems that arise from placing people into categories. The disadvantage, however, is that without this category, there would be real limitations to the range of what AI would be allowed to classify. The end result of developing an object recognition tool for training sets would be less effective and feel less complete. The project really requires that we establish labels for all conceivable nouns, including people, by doing an "ontology." Designers might also object that in practice it is not feasible to fully separate people from the taxonomy, that people are an integral part of reality, and that there is no practical way to limit the labeling process. (2) As an alternative to removing the people category altogether, we can attempt to classify people in more sensitive ways to include their intrinsic value. Are there ways to label people that uphold their intrinsic value as an end in themselves? Crawford discusses the problem of how to allow people to give consent for how they are labeled by AI.[42] We can interpret the fight for consent as a battle over the right to maintain the complexity and freedom of being a person.

It is also worth listing some of the central problems and objections that proponents of this argument would have to address to make this argument more plausible: (1) The argument relies on a hidden premise that the AI classification of people exaggerates use value and marginalizes intrinsic value. Proponents of this argument would need to make this premise more explicit and support it with evidence. Even though there is textual evidence that Crawford views the reduction of people to objects to be a consequence of AI classification, we can also interpret her to reject the

[42] Crawford, *Atlas of AI*, 109–111.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

normative aspect of this argument because it is not clear that AI classification causes the pure instrumental use value of people. (2) The argument also assumes that there is a clear distinction to be made between people and things. What exactly counts as a person? Do animals count as people? Should they also be treated as ends in themselves? What about trees? Where do we draw the line? This is a problem that Kant faced with his claims that only humans are rational beings and have rights, but it is also a problem that resurfaces now with AI.

The Non-normative Trunk: Arguments from Power and the Technical Worldview

The non-normative trunk contains a set of arguments drawn from the thesis that AI classification biases are irresolvable. These arguments are critical of the classification process itself, even of pre-AI classification systems. Some even make strong claims about determinate being. But they are nevertheless appropriate for this discussion because they assert that AI classification crosses a line, exacerbating distortions and biases by upholding a false objective authority of what are actually subjective, political, and arbitrary decisions about the category divisions of reality. Proponents of the non-normative trunk argue that it is impossible to rectify the mislabeling of fundamental reality because there is no way to catalog things fairly and accurately but, nevertheless, there may still be ways to mitigate the damage done by AI.

Argument 4: Power

One branch of the non-normative trunk is based on a claim that there is a necessary correlation between classification and power. Crawford argues for this variation, as when she writes, "artificial intelligence uses classification to encode power."[43] In a 2021 article, she and Paglen frame this argument in terms of politics by claiming that classification is inherently political:

> Despite the common mythos that AI and the data it draws on are objectively and scientifically classifying the world, everywhere there is politics, ideology, prejudices, and all of the subjective stuff of history. When we survey the most widely used training sets, we find that this is the rule rather than the exception.[44]

[43] *Ibid.*, 128.
[44] Crawford and Paglen, "Excavating AI," 1107.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

Arguably her most controversial assertion in the classification chapter is that AI taxonomies are always politically motivated expressions of power. She attacks the common sense position that AI can measure and classify the nouns of the world in politically neutral ways. There is always power and political agenda embedded within the categories and branches of the taxonomies—from the largest details of the top-level categories to the most minute details of an Amazon Mechanical Turk worker's directions for how to label an image.

Crawford claims that even non-AI forms of classification produce political power, that the very act of dividing things into categories is political. But she also claims that because of the authority to project objectivity, AI classifications exponentially magnify the ability to control reality. Classification is the process of dividing things from things, individuating and separating them into common and comparable forms. Crawford claims that AI classification uses the power of division to bolster conservative religious and essentialist values. Let's take the example, once again, of the binary man/woman. A topology of genders that only recognizes this binary exerts power by reducing the complexity of gender possibilities to conform to essentialist preconceptions. And yet, the problem runs deeper than this. Even if designers of AI classification propose a more nuanced topography with multiple genders, sub-categories, and in-between gray categories, even these more subtle forms of division produce power. Proponents of this argument view power to be a necessary and irretractable element of the classification process generally and view the advancements of AI division to cause an exponential increase in the surplus, encoding, and hegemony of this power.

Although Crawford only mentions him in a few passing citations, Foucault is the obvious precursor for this argument.[45] Foucault's influence on Crawford's work can be traced back to Bowker and Star's *Sorting Things Out*, a text that Crawford draws heavily from. Bowker and Star present Foucault in their introduction as a central theoretician of their work:

> Foucault's work comes the closest to a thoroughgoing examination [of the invisible forces of categories] in his arguments that an archaeological dig is necessary to find the origins and consequences of a range of social categories and practices. He focused on the concept of order and its implementation in categorical discourse… [His] practical archaeology is a point of departure for

[45] Crawford only briefly mentions Foucault in endnotes 12 and 66 of chapter 2 (*ibid.*, 250–251).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

> examining several cases of classification, some of which have become formal or standardized, and some of which have not.[46]

According to Foucault, it is not simply individual agents who wield power and make decisions that enforce and proliferate traditions over time. Instead, power emanates from practices, institutions, and systems of knowledge (epistemes). Against common assumptions about the source and stability of power, Foucault claims that power is the product of a series of discourses that cause things to be ordered in specific ways, discourses that de-individualize and obscure the origins and workings of power.[47] We can apply this insight from Foucault to AI classification. Power similarly emanates from the divisions of AI taxonomy, as if from an indeterminate source, rather than from some specific group or sovereign individual. The pseudo-objectivity and black box obscurity of AI cause the structure of the classification process to appear indifferent and faceless—there is no one person or identifiable group behind the divisions and labels. Although the team at ImageNet made initial decisions about the branches of their taxonomy, and while they can still partially guide revisions and attempt to reshape the consequences of their product, AI classification outstrips them and runs its own course.

For Foucault, there is no way out of power relations. Crawford seems to endorse this position, especially when she claims that data could never neutral and that classifying things is always political and always the embodiment of power. However, we can also interpret Crawford to partially diverge from the Foucauldian model when she asks about who gets to choose how AI taxonomy is designed.[48] She complicates the Foucauldian insight that there is no discernable source to power by questioning whether the origins of power can be traced back to decision-makers who design the taxonomy divisions and place images under categories.

In Foucault's version, the reason why there are no normative paths to rectify the encoding of power, other than resistance as a form of mitigation, is because there is no clear source to the emanation of power, since it does not reside in the agency of a sovereign who wields it, since its true source is in the structure of the order of things as a construct. Crawford's version of the argument is different. By attempting to find the source of power in the "who" that is behind the decision-making process, Crawford's argument loses some

---

[46] Bowker and Star, *Sorting Things Out*, 5.

[47] Michel Foucault, *Discipline and Punish: The Birth of the Prison*, trans. by Alan Sheridan (New York: Vintage Books, 1995), 202.

[48] Crawford, *Atlas of AI*, 147. See also Crawford and Paglen, "Excavating AI," 1115.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

of its consistency but gains resources for more robust possibilities of mitigating power. This inconsistency comes about because, on the one hand, like Foucault, she clearly views the power encoding of AI classification to be a necessary consequence of cataloging things. However, on the other hand, she sees ways to mitigate power although not being able to fully rectify it. Crawford seems to endorse the claim that if we could develop ethical guidelines about who gets to choose the foundational designs of the taxonomical divisions, there is a partially normative dimension to the argument that could aid us in developing fairer classifications.

Argument 5: The Technical Worldview

Crawford also frequently cites a second non-normative argument: the technical worldview argument. This argument is based on the claim that there is an insurmountable difference between the real content of things and their external expression when compartmentalized and organized through classification. Crawford can be interpreted to present this line of argumentation when she claims that AI classification causes a distortion of the qualities of things by overemphasizing their measurability and quantity. AI taxonomy nullifies differences so that asymmetrical things can co-exist in the same category. Proponents of this argument think that there is no way for AI classification to fairly approach things because the divisions of the taxonomy process fundamentally alter reality by causing artificial abstractions.

Crawford claims that AI classification requires a technical world and that this worldview both causes a distorted representation of the things that it classifies and causes them to fit into the shape of its conditions. There is a technical worldview embedded in the category divisions of ImageNet that aims to "flatten complex social, cultural, political, and historical relations into quantifiable entities."[49] If we interpret Crawford to mean that the technical worldview is a permanent aspect of AI classification and that it distorts the non-technical things that it classifies by making them fit into its view, then we are committed to the non-normative position that it is impossible to ever reach an objectively fair and neutral system of classification.

Crawford and Paglen lay the groundwork for this argument by claiming that there are underlying theoretical assumptions built into the training sets of AI systems:

> What are the assumptions undergirding visual AI systems? First, the underlying theoretical paradigm of

[49] *Ibid.*, 144.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

> the training sets assumes that concepts—whether "corn," "gender," "emotions," or "losers"—exist in the first place, and that those concepts are fixed, universal, and have some sort of transcendental grounding and internal consistency. Second, it assumes a fixed and universal correspondence between images and concepts, appearances and essences… It assumes that different concepts—whether "corn" or "kleptomaniacs"—have some kind of essence that unites each instance of them, and that that underlying essence expresses itself visually… Images of people dubbed "losers," the theory goes, contain some kind of visual pattern that distinguishes them from, say, "farmers," "assistant professors," or, for that matter, apples. Finally, this approach assumes that all concrete nouns are created equally, and that many abstract nouns also express themselves concretely and visually (i.e., "happiness" or "anti-Semitism").[50]

Proponents of the technical worldview argument claim that matching images to AI classification headings necessarily requires a set of assumptions that irrevocably change the constitution of things. These assumptions include the privileging of static, definable nouns over verbs, actions, relationships, and states of becoming. All things, it is assumed, are measurable. All nouns are created equal. Even the most abstract nouns can be expressed visibly by identifying them with images. These are not assumptions that we can expose and resolve; as the argument goes, these are assumptions that are built into the possibility of classifying things.

The technical worldview presents reality as a mechanical world, dominated by precision, pragmaticism, and abstraction. The main attributes of this worldview are speed, generality, and substitution. By extracting things from their context, the technical worldview produces speed. By sorting things under headline categories, it expresses the most abstract generalizations of things while removing their particularity.[51] By displacing things from their place among others, it enables the substitution of things that are otherwise unique. These characteristics are integral to the concept of data. The technical worldview is built from the assumption that things can easily be repackaged as data. ImageNet's project of an ontology of objects requires a massive

[50] Crawford and Paglen, "Excavating AI," 1113.

[51] Crawford, *Atlas of AI*, 134. Crawford talks about the tendency of AI to cause generality as a kind of induction.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

amount of data. Collecting data is like oil "just waiting to be extracted."[52] "According to this worldview," Crawford writes, "there is always more data to capture from the constantly growing and globally distributed treasure chest of the internet and social media platforms."[53] Two conditions for turning something into data are: (1) things must be taken out of context and (2) quality must be reduced to quantity.

(1) Turning something into data requires that it be displaced from its original context. Text, images, and videos are torn from the context of their environments and given little to no background. This causes an abstraction that is at once effective at aiding the speed and generality of the classification process but is also alarming. If we assume that context and background are part of the constitution of things, then the process of extracting things from their context distorts this constitution, causing an artificial reduction or oversimplification. This is especially problematic for predictive crime AI technology, but it is also generally one of the fundamental characteristics of data extraction to dislocate things from their surroundings.

(2) Crawford also describes data as a reduction of quality to quantity. The technical worldview argument is based on the idea that things are qualitatively distinct from each other, but that AI classification forces them to fit into predominately quantitative measurements. The reduction of quality to quantity is a necessary condition for reaching this worldview's standard of speed, acceleration, and rapidity of movement. This reduction is also an important attribute for substitution. Things that are otherwise incompatibly distinct from each other become exchangeable only if we are able to mute this incompatibility of their qualities and reach a mostly quantitative view of them.[54] The qualitative distinctions between things make exchanges cumbersome and slow to absorb the abstraction process of classification. However, by muting qualitative differences, we become able to sort things in terms of purely numerical differences, which greatly enhances the precision and rapidity of the taxonomy.

Some readers might object to the claim that the technical worldview is a non-normative argument, arguing instead that if we had more data, we could solve these distortion issues. In a provocative article, Chris Anderson

---

[52] *Ibid.*, 113.

[53] *Ibid.*, 94.

[54] For a related book-length study on how the concept of money causes a reduction of the qualities of things to their quantities, see Georg Simmel, *The Philosophy of Money*, trans. by David Frisby (London: Routledge, 1990). See also his essays about the interaction between life and form in Georg Simmel, *On Individuality and Social Forms*, ed. by Donald N. Levine (Chicago: University of Chicago Press, 1971) and Georg Simmel, *The View of Life: Four Metaphysical Essays with Journal Aphorisms*, trans. by John A.Y. Andrews and Donald N. Levine (Chicago: University of Chicago Press, 2010).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

argues that because we have now reached the petabyte age, it is possible to collect massive amounts of data at such a large scale that the old scientific method of hypotheses based on a model becomes obsolete.[55] Some might conclude from this that the reason we face biases and oversimplifications is because we have not yet collected enough data. But proponents of the technical worldview argument reject Anderson's position. They argue that it is not possible to have enough data because no amount of data could ever capture the qualitatively distinct ways that reality exists in our lived experience of it. It is not a question of having enough data because the concept itself of transforming the qualities and relationalities of things through technical conditions that turn it into data is the root cause of the problem. While some might view the normative interpretation of this argument as having traction, there is clearly more evidence for the technical worldview as a non-normative argument. The taxonomy process requires that things be extracted as data; but data extraction is the extraction of things from their context and background. It is not obvious how producing huge amounts of data would resolve this. Even if we were able to collect data at incredibly large scales, the technical worldview would still cause the same issues by forcing things to fit into the shape of their conditions.

Conclusion

There are two ambiguities to sort out in Crawford's work: (1) There is the question of whether there is any form of ethical response to be made that might resolve or, at any rate, mitigate the biases, reductions, and reifications caused by AI classification. For the most part, Crawford's work points to non-normative positions about classification, but sometimes her tone suggests that she is open to normative responses. She seems to think that, on the one hand, the revisions that the ImageNet team have begun to make contribute to a better and fairer system of AI taxonomy and that deeper, systemic responses could contribute further to at least limiting the extent of the distortions. However, on the other hand, readers who weigh the evidence she gives and the ways she emphasizes her positions will no doubt conclude that even the most heroic effort of redesigning the taxonomy will still fail to adequately represent reality. Her strong statements about how data cannot be made neutral and about how AI classifications are always political and power-driven expose her preference for non-normative arguments.

(2) Crawford does not consistently separate the two non-normative arguments, i.e., the power argument and the technical worldview argument.

[55] See Chris Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," in *WIRED* (23 June 2008), <https://www.wired.com/2008/06/pb-theory>.

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

She often lists the insights from the fifth argument alongside her more emphatic defense of the fourth argument, as if the two arguments were similar enough to be the same even when they are not. It is in itself valuable to clearly distinguish the two arguments from each other. Her discussion of power leads to a set of insights into the nature of data, political bias, panopticon structures, and discourses of knowledge. In contrast, her discussion of the technical worldview leads to insights into the oversimplification of complexity, the reification of abstract concepts, and what happens when being is fixed and not allowed to appear beyond caricatures and labels. Crawford clearly favors the power argument as the most important diagnosis of AI classification. Readers who acknowledge this and yet at the same time keep track of her other arguments are in a better position to visualize the various dilemmas of the normativity debate and what sense can be made of AI classification.

Krirk University
Bangkok, Thailand

References

Anderson, Chris, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," in *WIRED* (23 June 2008), <https://www.wired.com/2008/06/pb-theory>.

Benjamin, Ruha, *Race After Technology: Abolitionist Tools for the New Jim Code* (Cambridge: Polity Press, 2019).

Bowker, Geoffrey C. and Susan Leigh Star, *Sorting Things Out: Classification and Its Consequences* (Cambridge: MIT Press, 1999).

Broussard, Meredith, *More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech* (Cambridge: MIT Press, 2023).

Campolo, Alexander and Kate Crawford, "Enchanted Determinism: Power without Responsibility in Artificial Intelligence," in *Engaging Science, Technology, and Society*, 6 (2020).

Crawford, Kate, *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence* (New Haven: Yale University Press, 2021).

Crawford, Kate and Trevor Paglen, "Excavating AI: The Politics of Images in Machine Learning Training Sets," in *AI & Society*, 36 (2021).

Deleuze, Gilles and Félix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. by Brian Massumi (Minneapolis: University of Minnesota Press, 1987).

Foucault, Michel, *Discipline and Punish: The Birth of the Prison*, trans. by Alan Sheridan (New York: Vintage Books, 1995).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330

Gershgorn, Dave, "The Data That Transformed AI Research—and Possibly the World," in *Quartz* (26 July 2017), <https://qz.com/1034972/the-data-that-changed-the-direction-of-ai-research-and-possibly-the-world>.

Gould, Stephen Jay, *The Mismeasure of Man* (New York: W. W. Norton, 1996).

Kant, Immanuel, *Groundwork for the Metaphysics of Morals*, trans. by Allen W. Wood (New Haven: Yale University Press, 2002).

Noble, Safiya Umoja, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: NYU Press, 2018).

O'Neil, Cathy, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Crown, 2016).

Pasquinelli, Matteo, "How a Machine Learns and Fails – A Grammar of Error for Artificial Intelligence," in *Spheres: Journal for Digital Cultures*, 5 (2019).

Simmel, Georg, *On Individuality and Social Forms*, ed. by Donald N. Levine (Chicago: The University of Chicago Press, 1971).

_____, *The Philosophy of Money*, trans. by David Frisby (London: Routledge, 1990).

_____, *The View of Life: Four Metaphysical Essays with Journal Aphorisms*, trans. by John A.Y. Andrews and Donald N. Levine (Chicago: University of Chicago Press, 2010).

Yang, Kaiyu, Klint Qinami, Li Fei-Fei, Jia Deng, and Olga Russakovsky, "Towards Fairer Datasets: Filtering and Balancing the Distribution of the People Subtree in the ImageNet Hierarchy," in *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (2020).

https://doi.org/10.25138/19.1.a1
https://www.kritike.org/journal/issue_36/brown_march2025.pdf
ISSN 1908-7330