HOW COPYRIGHT LAW CREATES BIASED ARTIFICIAL INTELLIGENCE

Amanda Levendowski[*]

ABSTRACT

Breakthroughs in artificial intelligence (AI) threaten to be overshadowed with reports of bias. But why does AI reflect, perpetuate, and amplify human bias rather than eliminating it? Computer science and legal scholars have analyzed many sources of bias, including teaching AI with biased data. This Article is the first to examine the way in which copyright law may be biasing AI.

Artificial intelligence learns by reading, viewing, and listening to copies of human works. This Article examines how copyright law privileges easily available, legally low-risk sources of data for teaching AI, even though those data are demonstrably biased. Teaching AI using biased data as an attempt to avoid legal liability all but ensures that AI will inherit human biases. This Article concludes that teaching AI is not only a fair use under copyright, but one that quite literally promotes fairness.

[*] Technology Law and Policy Clinical Teaching Fellow, New York University School of Law and Research Fellow, NYU Information Law Institute. Thanks to Miles Brundage, Ryan Calo, H. Brian Holland, Craig Konnoth, Tiffany Li, Richard Loren Jolly, Zvi Rosen, Jason Schultz, Scott Skinner-Thompson, Cameron Tepfer, Eva Subotnik, Clinton Wallace, Ari Waldman, and Emily Winston for their thoughtful observations and suggestions. I am also grateful to the organizers and participants of the 2017 Works-in-Progress in Intellectual Property Colloquium, the NYU Lawyering Scholarship Colloquium, and We Robot 2017, where I presented early drafts of this paper.

I. INTRODUCTION

When Joz Wang bought her mom a new Nikon Coolpix S630 camera for Mother's Day, she didn't expect that the camera would turn out to be racist.[2] Wang, a Taiwanese-American blogger, noticed that Nikon's facial recognition algorithm identified her as "blinking" when she took photographs of herself. Except Wang wasn't blinking—she was just an Asian woman taking selfies.[3]

The facial recognition algorithms embedded in cameras like the Coolpix S630 are one example of artificial intelligence (AI).[4] But Wang's experience with the Coolpix S630 revealed a racial bias in the AI. The consequences for Wang were fairly mundane (her camera mistakenly thought she was blinking), but racist, sexist, and other biased predispositions baked into AI can and do have more meaningful consequences. Imagine instead that Wang had been unable to upload a photograph for her passport because the AI did not recognize that her eyes were open,[5] or that law enforcement misidentified her as a suspect because its AI had become a little *too* human and

---

[2] @jozjozjoz, "Racist Camera! No, I did not blink… I'm just Asian!," *Flickr* (May 10, 2009), https://www.flickr.com/photos/jozjozjoz/3529106844.

[3] Joz Wang, "Racist Camera! No, I did not blink… I'm just Asian!," *jozjozjoz* (May 13, 2009), http://www.jozjozjoz.com/2009/05/13/racist-camera-no-i-did-not-blink-im-just-asian. As an analog medium, photography was similarly designed for white faces—Jean-Luc Godard refused to use Kodak film during an assignment in Mozambique, claiming that the film was "racist." David Smith, "'Racism' of Early Color Photogrpahy Explored in Art Exhibition," *The Guardian* ((Jan. 25, 2013), https://www.theguardian.com/artanddesign/2013/jan/25/racism-colour-photography-exhibition. For a thoughtful deconstruction of the racial biases of film and photography, see Syreeta McFadden, "Teaching the Camera to See My Skin," *Buzzfeed* (Apr. 2, 2014), https://www.buzzfeed.com/syreetamcfadden/teaching-the-camera-to-see-my-skin?utm_term=.tkmgevBja#.vuP6ZMVRb.

[4] John McCarhy coined the term "artificial intelligence" in 1955 and described it the science and engineering of making intelligent machines. John McCarthy, *What Is Artificial Intelligence*, STANFORD FORMAL REASONING GROUP (2007), http://www-formal.stanford.edu/jmc/whatisai.pdfwe. As a point of reference, Siri, Apple's artificially intelligent personal assistant, defines AI as "the theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages." *Interview with Siri* (Feb. 26, 2017) (notes on file with author) . Amazon's version, Alexa, defines AI as "the branch of computer science that deals with writing compute programs that can solve problems creatively." (TK).

[5] Indeed, that is exactly happened to Richard Lee, a Taiwanese engineering student from New Zealand whose passport renewal photograph was rejected because the AI identified Lee's eyes as "closed." Hudson Hongo, "Passport Website Rejects Asian Man's Photo for Having 'Closed' Eyes," *Gizmodo* (Dec. 6, 2016), gizmodo.com/passport-site-rejects-asian-mans-photo-for-having-close-1789762968; *see also* Selina Cheng, "An Algorithm Rejected an Asian Man's Passport Photo for Having 'Closed Eyes,'" *Quartz* (Dec. 7, 2016), https://qz.com/857122/an-algorithm-rejected-an-asian-mans-passport-photo-for-having-closed-eyes/. There is something to be said about the fact that both Wang and Lee are of Taiwanese descent and both encountered biased facial recognition, but I will leave that analysis to other scholars.

inherited our cross-racial identification biases.[6] In those instances, the consequences of biased AI can be life-changing—and not for the better.

Biased AI is not limited to Nikon or New Zealand, or even facial recognition technology.[7] AI is embedded in our computers and our smartphones; it tells companies, banks, police officers, and judges who we are and what to think about us.[8] But search results,[9] voice recognition[10] and even sentencing algorithms[11] have all been identified as relying on AI that reflects, amplifies, and perpetuates human bias rather than eliminates it, often due to flawed training data.[12] In 2016, the Obama White House released its whitepaper on the future of AI and acknowledged that "AI needs good data. If the data is [*sic*] incomplete or biased, AI can exacerbate problems of bias."[13]

---

[6] This is a particularly salient question, as law enforcement increasingly relies on facial recognition for the investigation and identification of suspects For a comprehensive and nuanced discussion of law enforcement use of facial recognition technology, see Clare Garvie & Alvaro Bedoya, et al., *The Perpetual Line-Up: Unregulated Facial Recognition in America*, GEORGETOWN LAW CENTER ON PRIVACY & TECHNOLOGY (Oct. 18, 2016), https://www.perpetuallineup.org/ (hereinafter "Garvie & Bedoya, *The Perpetual Line-Up*"). In particular, Garvie, Bedoya, and Frankle unpack the race, gender, and age biases of facial recognition AI used by law enforcement and the consequences of those biases, including overlooking a perpetrator, misidentifying a suspect or mistakenly identifying an innocent person. *See* Garvie & Bedoya, *The Perpetual Line-Up*, at https://www.perpetuallineup.org/findings/racial-bias.

[7] Alice O'Toole, et al., *Face Recognition Algorithms and the "Other Race" Effect*, JOURNAL OF VISION 8(6), 256, http://jov.arvojournals.org/article.aspx?articleid=2136933; *see also* @jackyalcine, Twitter (Jun. 28, 2015) ("Google Photos, y'all fucked up. My friend's not a gorilla."), https://twitter.com/jackyalcine/status/615329515909156865?lang=en (calling out a facial recognition flaw in Google Photos that misidentified multiple photographs of African-Americans as gorillas).

[8] Whether this is an ideal state of affairs is not beside the point (and I would argue that it is not) but we have moved beyond discussing the merits of AI-based decisionmaking but, as a practical matter, it is here now and it appears to be here to stay.

[9] Latanya Sweeny, *Discrimination in Online Advertising*, 56 Comm. of the ACM 5, 44 (2013), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2208240 (Google searches for black-identifying names were 25% more likely to get an ad suggesting an arrest record than for white-identifying names); Amit Datta & Michael Carl Tschantz, et al., *Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination*, arXiv:1408.6491 (2014), https://arxiv.org/abs/1408.6491 (finding that a female-gendered automated tool received fewer instances of ads related to high-paying jobs);

[10] James Rodger & Parag C. Pendharker, *A Field Study of the Impact of Gender and User's Technical Experiene on the Performance of Voice-Activated Medical Tracking Application*,60 INT'L J. OF HUMAN-COMPUTER STUDIES 5-6, 529 (2004), (finding that voice-activated medical technology was more accurate and responsive to men's voices), https://arxiv.org/abs/1408.6491; Daniela Hernandez, "How Voice Recognition Systems Discriminate Against People with Accents," FUSION (Aug. 21, 2015), http://fusion.net/story/181498/speech-recognition-ai-equality/ (discussing recognition biases associated with accented voices observed in Microsoft Cortana and Google Now).

[11] Julia Angwin & Jeff Larson, et al., "Machine Bias: There's Software Used Across the Country to Predict Future Criminals, And It's Biased Against Blacks," *ProPublica* (May 23, 2016).

[12] *Infra* notes 11-13.

[13] EXECUTIVE OFFICE OF THE PRESIDENT, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (Oct. 2016),

Although Nikon and New Zealand's Department of Internal Affairs both declined to explain why their respective AI evinced a racial bias against Asian features,[14] flawed training data is a likely source.[15] Facial recognition AI trained using photographs predominantly of one race or ethnicity, say, Caucasian or white faces, will struggle to accurately identify faces of other races and ethnicities.[16]

Of course, the quandary of biased data producing biased results is not new—it's as old as the first computer. Charles Babbage, the inventor and philosopher credited with creating the first mechanical computer, first addressed the issue back in 1864: "On two occasions I have been asked—'Pray Mr. Babbage, if you put into the machine the wrong figures, will the right answers come out?' . . . I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question."[17] The principle underlying Babbage's reply is so foundational that computer

---

https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

[14] *Supra* notes 3 and 5. For an investigation of the troubling social and financial consequences stemming from the stunning lack of transparency around how data are used, which applies in equal force to AI, see FRANK PASQUALE, BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (2015).

[15] For a taxonomy that isolates and explicates technical issues in data mining and AI that might result in bias, see Solon Barocas & Andrew D. Selbest, *Big Data's Disparate Impact*, 14 CALIF. L. REV. 671, 671-3 (2016). The skewed demographics of engineers responsible for coding and designing AI can also result in biased AI. *See* Kate Crawford, "Artificial Intelligence's White Guy Problem," NEW YORK TIMES (June 25, 2016), https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html. Machine intelligence—perhaps the largest subfield of AI, and the one behind facial recognition, voice recognition, and sentencing technologies mentioned *infra*—is especially gender biased: a recent analysis of 78,768 engineering job listings found that postings for software engineers in the machine intelligence category had a gender-bias score favoring men *more than twice as high* as the next category. Kieran Snyder, "Language in Your Job Post Predicts the Gender of Your Hire," *Textio* (June 21, 2016), https://textio.ai/gendered-language-in-your-job-post-predicts-the-gender-of-the-personyoull-hire-cd150452407d#.rht0s16ov ("In light of the bias distributions above, the apparent scarcity of women in machine intelligence jobs is probably more than anecdotal."). And in 2016, it was estimated that fewer than 14% of the machine learning workforce were women. Cale Guthrie, "The Women Changing the Face of AI," *Fast Company* (Aug. 18, 2016), https://www.fastcompany.com/3062932/mind-and-machine/ai-is-a-male-dominated-field-but-an-important-group-of-women-is-changing-th.

[16] *See* P. Jonathon Phillips et al., *An Other-Race Effect for Face Recognition Algorithms*, 8 ACM TRANSACTIONS ON APPLIED PERCEPTION 14:1, 14:5 (2011) (finding that AI, like humans, evinces a cross-race effect: facial recognition algorithms developed in East Asia performed better on East Asian faces and algorithms developed in the United States and Western Europe performed better on Caucasian faces); *see also* Brendan F. Klare et al., *Face Recognition Performance: Role of Demographic Information*, 7 IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY 1789, 1797 (2012). There are, of course, other sources of bias, including technical biases introduced in classifying datasets and engineering algorithms. *Infra* note 9.

[17] Charles Babbage, *Passages from the Life of a Philosopher*, 66 (1864), https://archive.org/stream/passagesfromlif01babbgoog/passagesfromlif01babbgoog_djvu.txt. I would be amiss to mention Babbage without acknowledging the contributions of Ada Lovelace, who wrote the first algorithm meant to be executed by a machine and is often identified as the first computer programmer. *See*

scientists have reduced it to shorthand: garbage in, garbage out.[18] There is a robust body of scholarship, even entire conferences, dedicated to reducing biases and enhancing fairness of AI.[19] Legal scholars have long examined the complex legal and ethical questions posed by collecting, storing, and processing the quantities of "Big Data" needed to train AI.[20] Absent from the conversation, however, are analyses from copyright scholars about how our legal framework inadvertently biases data selection.

August Ada Lovelace, *Notes*, 3 SCIENTIFIC MEMOIRS 666 (1843) (translating and adding commentary to Luigi Frederico Menabrea's notions on Charles Babbage's Italian lectures on the Analytical Engine). Lovelace's envisioned a machine capable of more than mere calculation, more akin to the social and political computing technologies we use today. For an introduction to Lovelace's contributions to computer science, see Eugene Eric Kim and Betty Alexandra Toole, *Ada and the First Computer*, SCIENTIFIC AM. (May 1, 1999), http://www.cs.virginia.edu/~robins/Ada_and_the_First_Computer.pdf. For a longer examination of Lovelace's role in kickstarting the contemporary computer age, see BETTY A. TOOLE, ADA: THE ENCHANTRESS OF NUMBERS, PROPHET OF THE COMPUTER AGE (1998).

[18] Solon Barocas & Andrew D. Selbest, *Big Data's Disparate Impact*, 14 CALIF. L. REV. 671, 683 (2016)

[19] *See, e.g.*, Nina Gric-Hlaca, et al., *The Case for Process Fairness in Learning: Feature Selection for Fair Decision Making*, Conference on Neural INFO. PROCESSING SYS. (2016) http://www.mlandthelaw.org/papers/grgic.pdf; Ed Felten and Terah Lyons, "Public Input and Next Steps on the Future of Artificial Intelligence," MEDIUM (Sept. 6 2016), https://medium.com/@USCTO/public-inputand-next-steps-on-the-future-of-artificialintelligence-458b82059fc3; Kate Crawford & Ryan Calo, *There is a Blind Spot in AI Research*, NATURE (Oct. 13, 2016). For scholarship examining other sources of bias, as well as how to detect and avoid unfairness in AI, see Faisal Kamiran & Toon Calders, *Data preprocessing techniques for classification without discrimination*, 33 KNOWLEDGE & INFO. SYS. 1, 1 (2012) (suggesting algorithmic solutions—specifically suppression, massaging, and reweighing or resampling of sensitive attributes or discriminatory proxies—for reprocessing data to avoid biased classifications), https://link.springer.com/article/10.1007/s10115-011-0463-8. Since 2014, the FATML workshop, short for Fairness, Accountability, and Transparency in Machine Learning, has brought researchers and scholars into conversation around these questions in AI. *See* FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY IN MACHINE LEARNING (2016), http://www.fatml.org/.

[20] Note that much of the legal scholarship in this area uses Big Data and variants to not only refer to the quantity of data, but also to AI trained using that data and the algorithms that process it. *See, e.g.*, Paul Ohm, *The Underwhelming Benefits of Big Data*, 161 S 339, 340 (2013) (enumerating the "bad outcomes" that will inevitably follow projects reliant on big data); Neil Richards & Johnathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 41-2 (2013) (delving into ways that data collectors are empowered at the expensive of individual privacy and identity, even as collection methodologies and analyses are "shrouded in legal and commercial secrecy"); Danielle Keats Citron & Frank A. Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1 (2014); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 119-120 (2014) (exploring how personal privacy harms, such as revealing sensitive personal information, may stem from use of one's data without knowledge or consent); Ryan Calo, *Digital Market Manipulation*, 82 GEO. WASH L. REV. 995 (2014); Solon Barocas & Andrew D. Selbest, *Big Data's Disparate Impact*, 14 CALIF. L. REV. 671, 671-3 (2016) (identifying how biases in datasets can result in disparate impacts on marginalized populations).

AI is trained on "Big Data," and much of that data is derived from works protectable by copyright law.[21] To avoid legal liability for copying works to use as training data, researchers and companies generally have two options: license an existing database of copyrighted works from a third party or create a database of works they own.[22] I call these data "high-friction data."[23] When designing facial recognition software, for example, it would be prohibitively expensive and time consuming to negotiate licenses with a company like Getty Images, the world's largest repository of photographs, or build a platform like Facebook or Instagram, to which users regularly upload photographs that can, in turn, be used by those companies. It's understandable why many researchers and companies don't bother with high-friction data at all. Instead, researchers and companies rely on data that are easily accessible and *perceived* as legally low-risk. I refer to this type of data as "low-friction data." Common sources of low-friction data, however, are demonstrably biased.[24] In the context of facial recognition, it would be almost impossible to create a diverse dataset by relying exclusively on low-friction data—which may explain why commercial facial recognition struggles so much with diverse faces.

If we hope to reduce bias in AI, researchers and companies ought to be able to incorporate copyrighted works into their datasets to supplement or substitute for the biases of low-friction data. This is not to say that copyrighted works are neutral or bias-free (they aren't) or to suggest that more data equates to better data (it doesn't).[25] But without data derived from copyrighted works data to offset the unattainability of high-friction data and the known biases of low-friction data, AI is condemned to replicate human bias—a result anathema to copyright's fundamental purpose of "promot[ing] the Progress of Science and the Useful Arts."[26]

This Article is the first to address how copyright law channels AI in a fundamentally biased direction by privileging biased data, and I suggest that using copyrighted works to reduce bias in AI is a fair use.

In Part I, I explain the mechanics of teaching AI. In Part II, I explain why reproducing data or circumventing measures meant to limit access to training data for AI both pose risks of copyright liability and suggest that these risks limit the usability of copyrighted training data. Part II also contextualizes the copyright cases of the last three decades to illustrate that copyright owners historically and empirically regard new technologies with skepticism, if not outright hostility, that

---

[21] *See* 17 U.S.C. § 106; *supra* II.B.

[22] There are a number of well-regarded repositories of freely available datasets, but many of these datasets include data derived from copyrighted works.

[23] TK.

[24] *Supra* at III. A. and B.

[25] *Supra* at III. A.

[26] For context on why copyright law has treated robotic readership of copyrighted works as fair use, see James Grimmelmann's excellent discussion of the issue, *Copyright for Literate Robots*, 101 IOWA L. REV. 657, 658 (2016).

inevitably leads to litigation.[27] Part III distinguishes high-friction data, which depends on complex licensing regimes or sizable engineering teams, from low-friction data and unpacks the biases embedded in three common sources of low-friction data: the public domain, Creative Commons-licensed data, and proprietary open data. Finally, in Part IV, I offer a solution to the perverse effects of favoring biased low-friction data to avoid copyright liability by suggesting that teaching AI is a non-expressive, socially beneficial use of copyrighted works and one that is an inherently fair use. I conclude that using copyrighted works to train AI is not only a fair use, but one that can quite literally promote fairness.

[27] *See, e.g.*, Sony Corp. of Am. V. Universal Studios, Inc. 464 U.S. 417 (1984) (copyright owners unsuccessfully suing manufacturer of homevideo recording technology for infringement); Am. Broadcasting Co., Inc. v. Aereo, Inc., 134 S. Ct. 2493 (June 25, 2014) (copyright owners successfully suing provider of Internet-streamed broadcast television programming for infringement); *see also* Lewis Galoob Toys v. Nintendo of Am., 964 F.2d 965 (9th Cir.1992), *cert. denied*, 113 S. Ct. 1582 (1993) (copyright owner unsuccessfully suing manufacturer of video game alteration cartridge for infringement); Sega Ent., Ltd. V. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992) (copyright owners unsuccessfully suing (copyright owner unsuccessfully suing manufacturer of interoperable game cartridges for infringement); A&M Records, Inc. v. Napster, Inc., 239 F.3d 1004, 1013 (9th Cir. 2001) *and* Metro-Goldwn-Mayer Studios, Inc. v. Grokster, Ltd., 259 F. Supp. 2d 1029 (C.D. Cal. 2003), *aff'd*, 380 F.3d 1154 (9th Cir. 2004) (copyright owners successfully suing provider of peer-to-peer file sharing website for infringement); Kelly v. Arriba Soft Corp., 336 F.3d 811 (9th Cir. 2002) *and* Perfect 10, Inc. v. Amazon.com, 508 F. 3d 1146 (9th Cir. 2007) (copyright owners unsuccessfully suing operators of Internet search and retail websites for infringement); Authors Guild v. HathiTrust, 755 F. 3d 87 (2d Cir. 2014) *and* Authors Guild v. Google, Inc., 804 F. 3d 202 (2d Cir. 2015) (copyright owners unsuccessfully suing mass book-digitization projects for infringement); Capitol Records, LLC v. ReDigi Inc., 934 F. Supp. 2d 640 (S.D.N.Y. 2013), *appeal pending* (copyright owners successfully suing operator of secondhand digital music marketplace for infringement).