

UCLA

Recent Work

Title

Epistemic Asymmetries and the Architecture of Exclusion: Decolonizing the "White-as-Default" Paradigm in Dermatological AI

Permalink

<https://escholarship.org/uc/item/65p3f8wh>

Author

Berman, Zach

Publication Date

2026-03-16

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at <https://creativecommons.org/licenses/by-nc/4.0/>

Epistemic Asymmetries and the Architecture of Exclusion: Decolonizing the "White-as-Default" Paradigm in Dermatological AI

Berman, Zach¹

¹Department of Computer Science, University of California, Los Angeles; Los Angeles, California, 90095 USA

*zachberman@g.ucla.edu

Abstract: This paper analyzes how dermatological AI diagnostics reinforce a "white-as-default" standard, and how this marginalizes patients with darker skin. It critiques synthetic data techniques as digital colonialism and proposes community based regulatory solutions by applying frameworks such as Standpoint Theory and the Heteropatriarchal Triad

INTRODUCTION

The introduction of artificial intelligence (AI) into dermatological diagnostics has widely been considered an accomplishment in the medical field, as it promises to democratize specialty medical care and knowledge. However, commonly used AI models are poisoned by discrete sociotechnical realities that prioritize a "white-as-default" medical standpoint. Skin cancer datasets from major providers like ISIC contain over 70% white skin images [4]. This is an imbalance that often causes AI models to perform worse on darker skin. This can perpetuate diagnostic inequalities and lead to worse clinical results for people of color. Some engineers are currently attempting to fix this with novel generative AI techniques to generate images with dark skin. However, this techno-solutionism approach is idealized, and operates somewhat in a vacuum. It extracts and commoditizes the visual qualities and aesthetics of marginalized communities while ignoring their actual lived experiences and standpoints [4]. This paper investigates a few key questions about the current state and future of medical AI in this context:

1. Who, or what organization, is in position to decide what constitutes valid diagnostic knowledge in dermatology data and AI?
2. Can synthetic representations of marginalized bodies ethically substitute for the lived experiences of the real patients?
3. How can marginalized people participate in creating the tools and data used to diagnose their health?

This work is in partial fulfillment of the ENGR184 course using the blueprint curriculum in Ref [1,2] and captured in a collection [3].

METHODS

Feminist Standpoint Theory is applied to challenge traditional scientific "objectivity". This framework posits that knowledge is socially situated, and that marginalized groups possess an epistemic advantage capable of revealing blind spots in hegemonic narratives. Rather than accepting an abstract or generalized "diverse" viewpoint, Standpoint Theory demands adhering to "strong objectivity, centering the concrete lived experiences of marginalized groups, requiring a constant recognition of power dynamics, their roots, and how they impact such marginalization [5].

This study applies Standpoint Theory as a theoretical lens for evaluating medical AI systems. This involves critically examining whose perspectives are embedded in the design, validation, and evaluation processes of diagnostic technologies. By foregrounding the perspectives and lived experiences of historically marginalized patient populations, particularly individuals with darker skin tones, this framework is used to guide the assessment of diagnostic dermatological AI systems beyond purely technical performance metrics and toward the social epistemic conditions under which these metrics are produced.

Additionally, this analysis applies Imani Perry's Heteropatriarchal Triad, which compartmentalizes modern power structures into the interdependent pillars of Sovereignty, Property, and Personhood [6]. The paradox of the triad is that while our system derives stability from these three interconnected pillars, it also inherently is designed to exclude those who do not fit its geometry. The Heteropatriarchal Triad maps to the dermatological AI ecosystem as follows:

1. Property: Tech companies and researchers extract diverse patient data to create profitable, proprietary diagnostic algorithms. In this dynamic, marginalized bodies and their digital footprints are commodified as raw property to enrich corporate datasets [7-12].
2. Personhood: The medical system and AI algorithms diminish marginalized patients to data points. Reducing certain groups to isolated biological aesthetics like skin tone pixels, AI data systems institute non-personhood when they fail to recognize the full human context and situated knowledge of patients [13].
3. Sovereignty: Regulatory bodies like the FDA, alongside biotechnology monopolies, maintain sovereign authority over what is deemed "safe" and "effective". They dictate validation standards of dermatologic diagnostics that can easily overlook minority experiences, exerting a form of algorithmic control and institutional power over marginalized populations [14].

When considered together, Feminist Standpoint Theory and the Heteropatriarchal Triad reveal complementary structures of power within sociotechnical systems. Standpoint Theory interrogates the epistemics of knowledge production, asking whose experiences are recognized as legitimate sources of scientific knowledge. The Heteropatriarchal Triad, by contrast, illuminates the structural and institutional mechanisms through which authority is maintained. When applied simultaneously to dermatological AI, these frameworks allow for analysis across both epistemic and institutional levels. This lets us examine not only how marginalized knowledge is excluded from dataset formation, but also how corporate ownership and authority, and proprietary algorithmic systems reinforce these exclusions.

In this research, these two frameworks are specifically applied to critically examine the recent engineering trend of using Generative AI (such as Stable Diffusion and Low-Rank Adaptation) to generate fake images of dark skin, specifically Fitzpatrick Skin Types V and VI, to fill representation gaps in datasets like the International Skin Imaging Collaboration (ISIC) archive [4]. By analyzing this specific technical case, the methodology exposes how current "techno-solutionism" satisfies the algorithmic requirement for data (Property) while simultaneously erasing the actual marginalized patients (Personhood).

RESULTS AND INTERPRETATION

Applying these frameworks to research unveils a few interesting dynamics. First, is that the "white-as-default" standard in dermatological AI is not just some software error or coincidence. Instead, it's a technical manifestation of a long history of marginalization. Historically, dermatology training relied on older visual photos that were optimized for lighter skin, mostly Fitzpatrick skin types I and II [15]. In the 20th century, photographic tools like Kodachrome film and "Shirley cards" were chemically and mechanically calibrated for the properties of white skin [15]. The actual engineering design of these imaging technologies rendered non-white bodies essentially illegible.

When these archives were transitioned into the machine learning era, these biases were discretely brought in and solidified as "ground truth" for AI. This is very evident in major benchmark datasets like the International Skin Imaging Collaboration (ISIC) archive, where over 70% of the images depict light skin, while darker skin tones (Fitzpatrick V and VI) account for less than 8% of the data [4]. Through the lens of Standpoint Theory, these datasets masquerade as objective, universal science, but actually represent the highly specific, dominant standpoint of a white hegemony. The resulting algorithms learn to prioritize the features prevalent in lighter skin while entirely missing the diagnostically relevant patterns in darker skin tones [4].

To address this data scarcity, modern engineering has increasingly turned to Generative AI, using tools like Stable Diffusion and Low-Rank Adaptation (LoRA) to synthesize artificial images of dark skin [4]. Analyzed through the Heteropatriarchal Triad, this techno-solutionism functions as digital colonialism, more specifically, data colonialism. Data colonialism operates by extracting raw materials, in this case, human behavioral and biological data, from marginalized populations for the profit and optimization of systems owned by tech monopolies in the Global North [9].

Generating synthetic dark skin perfectly illustrates the tension between "Property" and "Personhood." The algorithm requires data (Property) to function and be sold as a viable diagnostic product. By synthesizing melanin pigmentation changes to augment training "accuracy" and validation, developers extract and commodify the biology and aesthetics of marginalized, dark-skinned communities [7]. However, this "technological mimicry" operates in a vacuum, entirely erasing the "Personhood", perhaps confounding lived experiences, of actual patients. A software model of personhood reduces human beings to an information processing system of simply, extractable data points, discarding the embodied, lived experiences that surround health and disease. The deep paradox here is that the AI is trained to successfully "recognize" a synthetic matrix of dark pixels, yet the overarching sociotechnical system continues to refuse recognition of the actual marginalized human being. Even further, the original goal of seeking medicine equally for all is compromised as researchers use this data purely to improve AI scores on test sets and validation.

The final pillar of the triad, "Sovereignty," is exercised by regulatory bodies like the FDA, which maintain authority over what is deemed safe and effective for public use [16]. Currently, the regulatory clearance of medical AI relies heavily on lab based technical sampling and abstract performance metrics, often relying on nondeterministic and non-binding "soft law" guidelines that cater to corporate innovation rather than establishing enforceable standards for true equity [16].

Epistemic injustice is when individuals are systematically stripped of their capacity as credible knowers and interpreters of their own lived experiences [17-18]. These systems were created without requiring any sort of integration of patients' diverse situated knowledge, committing a direct epistemic injustice. When AI models trained on a Western sun-exposure model of disease, where melanoma is expected to appear on sun-exposed skin, are deployed, they often fail to account for how skin cancer actually presents in people of color, such as acral lentiginous melanoma appearing on the palms, soles, or under the nails [18]. Additionally, marginalized patients are excluded from any of the algorithm design and validation processes, so their testimonies and standpoints are similarly excluded. As a result, Black patients face significantly higher mortality rates and are far more likely to be diagnosed at late, incurable stages compared to their white counterparts. So, the failure of AI diagnostic tools cannot simply be attributed to naive oversight. Instead, it's an actively constructed system of marginalization that prioritizes institutional sovereignty over equal clinical results.

CONCLUSIONS

The integration of AI into dermatological diagnostics currently functions as a system of exclusion, reinforcing a "white-as-default" paradigm that harms patients with darker skin tones. By applying Feminist Standpoint Theory and the Heteropatriarchal Triad, this research concludes that the current techno-solutionist approach of synthesizing dark skin data to improve algorithmic accuracy is a form of digital colonialism. This extractivist logic commodifies marginalized bodies as proprietary data (Property) while simultaneously erasing their complex lived realities (Personhood). This leaves the institutional and corporate authority of regulatory bodies (Sovereignty) unquestioned. The failure of these diagnostic tools in this manner is not merely a technical lapse or an accident, as it is a system built on foundations of marginalizing data and algorithms. This is often left unaddressed as doing so prioritizes the efficiency of corporate innovation and institutional control, over equitable clinical outcomes.

To challenge this dominant paradigm, future work in medical AI must go beyond generative dataset augmentation and validation-chasing to actively attack the epistemic asymmetries embedded in AI design. While further technical solutions like adopting the Monk Skin Tone (MST) scale could prove beneficial (MST-AI models have shown to accurately automatically detect skin color in large datasets) the real reform needs to come from fundamentally altering the power dynamics of who is allowed to validate these tools [12]. This includes creating mechanisms that allow affected patient populations to contribute their lived experiences and knowledge to the design and evaluation of medical AI.

This effort will require direct collaboration with the communities and individual people most affected by diagnostic bias. Future governance of dermatological AI should incorporate things like community advisory boards composed of dermatologists of color and patient advocacy organizations who participate in dataset formation and AI validation. This participation and advisorship could take control away from exclusively corporate and regulatory parties and give it to communities and patients that the tools will actually be used on.

REFERENCES

1. Lee, Ethan, et al. "Education for a Future in Crisis: Developing a Humanities-Informed STEM Curriculum." arXiv preprint arXiv:2311.06674 (2023).
2. Carbajo, Sergio. "Nurturing Deeper Ways of Knowing in Science." *Issues in Science & Technology*, 2025, v. 41, n. 2, p. 71. doi:10.58875/jkrw4525.
3. Carbajo, S. (2026). *Queered Science & Technology Center: Volume 4*. Queered Science & Technology Center, 4. Retrieved from <https://escholarship.org/uc/item/0zb930m0>
4. Shabu, Areez Muhammed. "A Generative AI Approach for Reducing Skin Tone Bias in Skin Cancer Classification." arXiv, 2026. Retrieved from <https://arxiv.org/html/2602.14356v1>
5. Harding, Sandra. *Whose Science? Whose Knowledge?: Thinking from Women's Lives*. Cornell University Press, 1991.
6. Perry, Imani. *Vexy Thing: On Gender and Liberation*. Durham: Duke University Press, 2018.
7. "Data Flows and Colonial Regimes in Africa: A Critical Analysis of the Colonial Futurities Embedded in AI Recommendation Algorithms." arXiv, 2025. Retrieved from <https://arxiv.org/pdf/2511.19283>
8. "Making the Case for Africa's Leadership in Responsible AI Development." African Centre for Technology Studies (ACTS). Retrieved from <https://acts-net.org/making-the-case-for-africas-leadership-in-responsible-ai-development/>
9. De Freitas, Marcus Vinícius. *Digital Sovereignty and Data Colonialism: Shaping a Just Digital Order for the Global South*. Policy Center for the New South, 2025. Retrieved from https://www.policycenter.ma/sites/default/files/2025-10/PP_38-25%20%28Marcus%20Vini%CC%81cius%20De%20Freitas%29.pdf

10. "Algorithmic Colonization of Africa." In *Imagining AI: How the World Sees Intelligent Machines*. Oxford Academic. Retrieved from <https://academic.oup.com/book/46567/chapter/408130272>
11. "Re-imagining Algorithmic Fairness in India and Beyond." arXiv, 2021. Retrieved from <https://arxiv.org/pdf/2101.09995>
12. "MST-AI: Skin Color Estimation in Skin Cancer Datasets." *Journal of Imaging*, MDPI. Retrieved from <https://www.mdpi.com/2313-433X/11/7/235>
13. Jones, Mirabelle. *The Data Is of Us But Not Us*. PhD Thesis, University of Copenhagen, 2024. Retrieved from https://di.ku.dk/english/research/phd/phd-theses/2024/Mirabelle_Jones_PhD_thesis.pdf
14. "FDA Guidance on AI-Enabled Devices: Transparency, Bias, & Lifecycle Oversight." CenterWatch: Clinical Trial Listing Database & Insights. Retrieved from <https://www.centerwatch.com/insights/fda-guidance-on-ai-enabled-devices-transparency-bias-lifecycle-oversight/>
15. "Technical Responses to Critique: The Case of Skin Tone." ResearchGate. Retrieved from https://www.researchgate.net/publication/391239665_Technical_Responses_To_Critique_The_Case_Of_Skin_Tone
16. "The Illusion of Safety: A Report to the FDA on AI Healthcare Product Approvals." Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC12140231/>
17. Ajmani, Leah Hope. "Moving Towards Epistemic Autonomy: A Paradigm Shift for Centering Participant Knowledge." arXiv, 2025. Retrieved from <https://arxiv.org/abs/2501.14648>.
18. Jin, Weina, et al. "Ethical Medical Image Synthesis." arXiv:2508.09293 [cs.CY] (2025).
19. Wu, Xinxin, MD, et al. "Disparities in Malignant Melanoma Diagnosis in American Indian/Alaska Native Versus White Patients." Creighton University School of Medicine, Omaha, Nebraska, USA.
20. Gupta, Alpana K. "Skin Cancer Concerns in People of Color: Risk Factors and Prevention." Retrieved from <https://pmc.ncbi.nlm.nih.gov/articles/PMC5454668/>