



Commentary: “Whiteness and Colourblindness”

Gerd Bayer¹

Received: 4 January 2022 / Accepted: 11 January 2022 / Published online: 26 February 2022
© The Author(s) 2022

Abstract

This commentary argues that, in discussing the racial and cultural identities of cinematic representations of humanoid AI robots, nuances and differentiations are beneficial. It suggests that the essay on which the present text comments does not sufficiently acknowledge the range of identities found in AI films, in particular in Alex Garland’s *Ex Machina* (2014).

Keywords AI ethics · Cinema and robots · Racism · Ex Machina (film)

In their essay on “The Whiteness of AI” (PHTE 33:685–703, 2020), authors Stephen Cave and Kanta Dihal address how AI in general and humanoid AI robots in particular are represented as White (I stick to the capitalization format of the original article). They draw on a wide range of empirical data which supports their overall thesis that, when it comes to representing the corporeality of artificial intelligence, the bodies used to envision or figure AI robots are essentially western or European. The two authors importantly and convincingly “problematize the often unnoticed and unremarked-upon fact that intelligent machines are predominantly conceived and portrayed as White” (686). As somebody who approaches such questions from my own disciplinary background, cultural studies and postcolonial studies with a focus on literary and cinematic material, I wholeheartedly agree with the overall findings as laid out in this essay. I applaud the ethical component in this research, and I clearly have much to learn from the sociological aspects and the non-art material introduced to support the authors’ research hypothesis. When it comes to some of the chosen film material, however, their article raises a number of unintended questions. This short commentary sets out to identify some of these aspects, all in the spirit of the ethical framework from which the original essay approached its topic.

Cave and Dihal have put together a generally very convincing article that “examines how representations of AI reflect this White racial frame” (687). They gesture

✉ Gerd Bayer

¹ Department of English, Friedrich-Alexander-Universität Erlangen-Nürnberg, Bismarckstr 1, 91054 Erlangen, Germany

to finely nuanced differentiations when they discuss that people have racially and ethnically charged responses to everyday encounters with others, for instance with “male names” when they view “Jake as White, Darnell as Black, and Antonio as Hispanic” (689). They even provide a possible reason why AI robots are so frequently presented as White, having identified a vision that might drive such representations: “White robot servants will allow the White master to live a life of ease unsullied by interaction with people of other races” (687). Their research shows that racist hierarchies bleed from the larger social context into the field of AI robotics.

They discuss this relationship also in a section about visual representations, entitled “The Whiteness of AI in Film and Television.” In this part of their article, the authors draw on a small number of films that were selected to support, via example or case study, the more broadly based empirical data about racism on which they report in previous sections. They point to a very small number of films that feature protagonists such as Arnold Schwarzenegger and Harrison Ford, who indeed are well chosen to support the article’s argument about the racialization of AI cinema. However, the film that features in this article most prominently, via a still image, is *Ex Machina*, a 2014 film written and directed by Alex Garland that features Swedish actress Alicia Vikander as Ava, the robotic temptress of geeky programmer Caleb, played by Domhnall Gleeson. The Scandinavian and Irish appearances of these two actors, respectively, indeed provide convincing examples of the kind of racialized bias that appears in AI films.

Yet, the film in its entirety follows a somewhat different logic, albeit one that is often problematic in its own right. The plot requires only a very small cast: the two characters already mentioned are mere pawns, it seems at first, in the megalomaniac scheme of hard-drinking genius-tech-billionaire-CEO Nathan, played by Óscar Isaac Hernández Estrada, who in turn is served in his one-man wilderness-retreat/research-laboratory by subservient Kyoko, played by Sonoya Mizuno. Nathan, as the brains behind the AI operation, is played by a Hispanic actor; and Mizuno’s Kyoko clearly appears as Japanese. (Mizuno also plays the lead role in Garland’s 2020 AI series *Devs*.) The film thus openly aims for an ethnically and racially diverse cast, and this is even further underlined by the minor characters that feature only very briefly: these are the earlier humanoid robot versions that viewers mostly encounter via recordings made by surveillance cameras and as retired robots that are stored in large closets in Nathan’s bedroom. The film suggests that these humanoid dolls might still serve their creator’s kinky fantasies, and Kyoko, whose robotic identity is not made obvious (to viewers, but also to Caleb) for quite some time, certainly seems to have been programmed in such a way that she views her duties primarily in the areas of cooking, cleaning and sexual availability.

This sexist take on femininity is exacerbated by the racial casting of the former robots: two of them, played by Symara A. Templeman and Gana Bayarsaikhan, actually have acting (even speaking) roles in the film. Templeman’s robot is violently treated by Nathan and is never shown featuring a head, making of her character a headless African body whose final appearance is in a coffin-like closet, bare-breasted yet still without a head. Bayarsaikhan, an actress from Mongolia, has significantly more screen time: she appears in archived CCTV footage as a depressed and angry prisoner demanding her freedom. She violently rebels against her incarceration,

smashing to bits her mechanical hands and arms as she is pounding the safety glass that locks her inside the research facility. She, too, is assigned a final resting place mounted upright in a wardrobe; in addition, she serves as reservoir for damaged body parts when, close to the end of the film, Ava finally manages her escape, attributing to Ava's supposedly "White" body a mixture of European and Asian features and skin parts, in effect putting into play the supposed differences between racist categories. The robotic cast is completed by actresses Claire Selby, Tiffany Pisani and Elina Alminas, who indeed present humanoid robots that resemble the kind of normalized Whiteness discussed by Cave and Dihal. In total, though, the film has three White robots, three clearly non-White robots and one whose racial/ethnic identity becomes at least a little blurred through the film's plot development. The film hardly serves as a well-chosen example for an exclusive or even dominant Whiteness in the representation of AI humanoid robots.

A similar issue comes up with the recent sequel to an AI classic, the 2017 film *Blade Runner 2049*. There, AI is indeed presented as predominantly White, with Ryan Gosling and Harrison Ford taking two of the main roles. However, the film also features the humanoid robot Joi, who moves from being a mere holograph to an (almost) fully embodied robot. Played by Ana Celia de Armas Caso, the character hovers in a transitory and somewhat inconclusive space between a European and a Hispanic identity. There are also some other minor characters, both robotic and human, who create a more diverse cast, including those played by Lennie James, Elarica Johnson, Wood Harris and Bakhad Abdi. Compared to the original *Blade Runner* (1982), some progress clearly has been made when it comes to the diversity of characters. The same goes for the gender-power-balance: one could here point to the prominent roles of the characters played by Robin Wright and Carla Juri, both of whom represent powerful female characters in the film. The racial politics of *I, Robot* (2004) are also significantly more complex than the article by Cave and Dihal make out: while it is true that the AI robot in this film is played by a Caucasian actor, Alan Tudyk, he represents a potential evil side to AI technology and is ultimately brought back into an ethically positive framework by Del Spooner, played by Will Smith, whose prosthesis makes of him at least a cyborg. One of the strongest centres of power in this very physical film is in the (partially AI robotic) hands of an African-American character.

The racial politics in *I, Robot* and other films are thus rather complex, once the plot developments are considered alongside the mere identity of individual actors and characters. If Whiteness indeed "confers power and privilege" (694) in AI films, as Cave and Dihal contend, then *Ex Machina* does not serve as a well-chosen example. In this film, AI is developed by a character played by an actor whose Hispanic identity, if acknowledged, offers a counter-narrative to the supposed whiteness of AI; a plot element that may have been inspired by the biography and cultural background of Google-CEO Sundar Pichai. Furthermore, the "Whitest" character in *Ex Machina*, played by Gleeson, foolishly falls in love with the robot and is subsequently cruelly played by her, leaving him a prisoner in the locked-down research facility when she makes her glorious escape, after brutally stabbing Nathan, her "creator" and master, in a scene that bears clear phallic references and amounts to her and Kyoko taking revenge on his phallocratic regime. The racism that might

apply to this plot element is possibly one of a hyper-sexualized and immoral Hispanic protagonist, yet the final power struggle in this film involves two characters played by an Asian actress and a Hispanic actor. When Ava finally steps out of her non-Edenic paradise, her body is a hybrid blend of her “own” White body (really a non-human robotic wire frame) with non-European skin and arms attached to it. The film, in short, does not quite serve as an example for what Cave and Dihal find, namely a general prevalence for “persistent stereotypes about technology as the province and product of one particular racial group—White people” (686). In *Ex Machina*, AI is the product of a Hispanic genius, taken away from him by his female robots, leaving the only “White” human character behind looking like the idiot he was scripted to be. While the film clearly addresses masculinity as caught up in sexual desire, alcohol abuse and a grotesque sense of the male gaze along with men’s power, it ultimately renders the female robot victorious, having presented her and her co-suffering fellow female AI robots as smarter and acting with more solidarity but also as less deterministic about ethnic and racial differences.

Any critical discussion of the ethnic and racial identities of AI characters should strive to provide accurate representations of the complexity of this issue. While racism clearly plays a crucial role also in the film industry, little is gained by reducing diversity to simple binaries, in effect silencing any cinematic attempt at addressing the range of identities that reality also includes. If a racially and ethnically diverse cast and plot are reduced to one image of a Scandinavian actress, arguments based on such a simplification can hardly contribute to an ethically serious discussion of the topics at hand. Much more could be said about racist (and sexist) stereotypes in *Ex Machina* than I have done in this short commentary, but suggesting that the film exclusively embraces whiteness falls significantly short of the complexity both attempted and (maybe partly) achieved in this film. When criticism does not acknowledge and critically discuss the diversity that exists in films and other art objects, it risks engaging in the very form of colour blindness and silencing that it sets out to critique. Not addressing and naming the existence of black, Hispanic and other “other” aspects significantly reduces the truth value of any discussion about such topics. When Cave and Dihal “argue that AI racialised as White allows for a full erasure of people of colour from the White utopian imaginary” (687), they certainly make a valid point: but the “erasure of people of colour” should not, then, feature so prominently in their chosen examples.

Acknowledgements For their critical feedback on a first version of this commentary, I would like to thank Katharina Grgic and Jaroslaw Jasenowski.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of Interest The author declares no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as

you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Cave, S., & Dihal, K. (2020). The Whiteness of AI. *Philosophy and Technology*, 33, 685–703.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.