



Artificial intelligibility: the role of gender in assigning humanness to natural language processing systems

Jenny Carla Moran

To cite this article: Jenny Carla Moran (06 May 2024): Artificial intelligibility: the role of gender in assigning humanness to natural language processing systems, Journal of Gender Studies, DOI: [10.1080/09589236.2024.2348482](https://doi.org/10.1080/09589236.2024.2348482)

To link to this article: <https://doi.org/10.1080/09589236.2024.2348482>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 06 May 2024.



Submit your article to this journal [↗](#)



Article views: 94



View related articles [↗](#)



View Crossmark data [↗](#)

Artificial intelligibility: the role of gender in assigning humanness to natural language processing systems

Jenny Carla Moran

Centre for Gender Studies, University of Cambridge, Cambridge, UK

ABSTRACT

In place of 'Artificial Intelligence', this article proposes artificial intelligibility as the more accurate term for describing the mistaken assignment of humanness to non-living objects. Artificial Intelligibility is manifested when a user assumes an object has capacity to understand them simply because it is understandable to them. Relatively simplistic Natural Language Processing systems may perform *genres* of humanness in conversational interactions to the degree that they are imagined to be sentient, as 'Artificial Intelligence'. However, decolonial scholars have observed that humanness developed as a mutable, sociogenic construct. Decolonial gender theorists have further derived the role played by heteronormativity and repronormativity in upholding this power. Equating the intelligible performance of a gendered *genre* of humanness with intelligence risks obfuscating this interplay while reinforcing eugenic stratifications of life. This article reframes behavioural measures of 'intelligence' to gendered 'intelligibility' to explore the role of gender in enabling entry into the symbolic order of humanness. It presents three key findings from my doctoral thesis to question how and why some gendered, NLP-incorporating devices can be imagined to have lives that matter within the same economy of value that renders some humans and animals killable.

ARTICLE HISTORY



Received 30 September 2022
Accepted 10 April 2024

KEYWORDS

AI; colonialism; humanness; gender; performativity

Introduction

Viewers of contemporary, dominant media will likely already be familiar with the representation of computer as woman (see Branham et al., 2011, p. 401). Whether the device in question be a sexualized fembot as in *Blade Runner* (Scott, 1982), *Westworld* (Abrams et al., 2016–2022) or *Ex Machina* (Garland, 2014), a perfected domestic servant as in *The Stepford Wives* (Oz, 2004), an abused care worker as in *Humans* (Brackley et al., 2015–2018), an abandoned maternal protector as in *Raised By Wolves* (Guzikowski, 2020–Present), a chat-script companion as in *Her* (Jonze, 2013), or even a dissatisfied undersea computer-wife as in *SpongeBob SquarePants* (Hillenburg, 1999–Present), there is an undeniable sense of victimhood, entrapment, and melancholia associated with gendered Artificially-Intelligent (AI) devices in popular narrative. This trend matters not least because it is informed by a significant history of gendering the products of engineering (see Truitt, 2021). Moreover, the tropic representation of gendered AI as sentient, suffering victim in turn informs tech ethics, policy, and dominant understandings of existing AI devices in the real world. As Stephen Cave, Kanta Dihal, and Sarah Dillon document, hype and misinformation propagated primarily through Hollywood cinema and popular news coverage has directly impacted governance of AI in

CONTACT Jenny Carla Moran  jemoran@tcd.ie  Centre for Gender Studies, University of Cambridge, Newnham College, Sidgwick Avenue, Cambridge, UK

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the UK (2020, p. 7–10). Perhaps unsurprisingly, then, recent critiques of gendered AI, such as the ‘Campaign Against Sex Robots’ (Richardson, 2015), have reanimated the trope of gendered AI as victim by decrying gendered objectification (Bates, 2017; Murphy, 2017) and the gendering of subservient, task-performing systems as female (Penny, 2016). At the same time, however, the robot Sophia, created by Hanson Robotics, has become the first AI device to ever be granted sovereign citizenship – an undeniable marker of personhood which is not equally afforded to all humans or indeed sentient beings. Despite the fact that Sophia has been describe as little more than a ‘BS puppet’ embodying Natural Language Processing systems (NLP) (in Urbi & Sigalos, 2018, n.p.), rendering it akin to a Google Virtual Assistant (Parviainen and Coeckelbergh, 2021, p. 718), it has been granted a formal status of entitlement to protection. To what extent does its gendering, recalling the victimized sentient AI trope, play a role in the apparent will to imbue the device with a life that matters, and thus to protect the non-living robot over some living beings? Turning to the role of gender performativity in conversation with the racialized sociogeny of humanness, this article proposes ‘artificial intelligibility’ as a more accurate term than ‘artificial intelligence’ to describe the assignment of humanness to gendered NLP.

The term ‘artificial intelligence’ was first coined in the proposal for the 1956 Dartmouth Summer Research Project on Artificial Intelligence (McCarthy et al., 2006/1995, p. 12). Attendees of this school, Herbert Simon and Alan Newell, went on to popularly reframe ‘human minds’ and ‘modern digital computers’ as similar symbolic information processing systems, representing computers and humans as ‘species of the same genus’ (Dick, 2019, p. 2). Stephanie Dick argues that this framing produced the dominant approach to AI: a focus on identification of processes related to ‘intelligent human behaviour’ which could be reproduced by algorithms (2019, p. 2). Dick refers to a behavioural model of intelligence, most famously invoked in Alan Turing’s ‘imitation game’, which tested the stringing together of words in a manner that would be convincing to the user.¹ However, in these dominant twentieth-century approaches to AI, contemporary popular media reporting, and SF representations alike, the term ‘AI’ is oft used to describe the (real or imagined) output of any number of software systems classed under its banner, including machine learning, facial recognition, deep learning, expert systems, natural language processing, and neural networks. A software system designed for conversational output, such as NLP, will undeniably do better in a behavioural model of intelligence, despite the fact that other forms of AI, like expert systems, may be far more complex (see Dick 2019). Hence, the actual ‘intelligence’ of a device is not necessarily what is being measured in such tests. No standard measure of capacity is equally applicable to all of these systems, and, furthermore, cognitive scientists and philosophers have long critiqued the ascription of ‘intelligence’ to well-performing NLP (Hofstadter, 1995; Searle, 1980; Suchman, 2007), which Ned Block describes as possessing nothing more than ‘the artificial intelligence of a juke box’ (1995, p. 5). Nonetheless, the false ascription of ‘intelligence’ to NLP that performs in an adequately human-like manner may contribute to the mistaken understanding of AI devices as sentient victims in need of protection. Noting that this constitutes a misdirection of care, and inspired by a growing body of work into disingenuous rhetoric surrounding AI (see Penn, 2020), this article explores the role played by gender performativity in assessing the human-likeness (and therefore ‘intelligence’) of NLP.

This article explores the role of gender in assigning humanness to NLP by critically analysing gender performativity in reference to the racialized sociogeny of humanness. Sociogeny is a term first coined by Fanon (1968, p.11). It refers to how -alongside evolutionary lineage and one’s personal upbringing – culture and socio-political structures produce phenomena and shape experiences of the world. In the works of Black Studies scholars who have extended Fanon (Jackson, 2020; Weheliye, 2014; Wynter, 2003), humanness is understood as a sociogenic construct that is unequally afforded under coloniality, such that colonized, enslaved, and/or incarcerated Black and Indigenous peoples can be rendered non-human, object property to Western powers. This is particularly problematic within an economy of value that can simultaneously recognize robots as persons. Hence, this article takes a critical approach to the human as well as AI, exploring the role of gender in assigning humanness to NLP in order to learn more about

sociogeny. Engineers have already realized that gendering robots elicits responses that are informed by culture and socio-political structures.² Existing feminist STS approaches have also used AI devices as a means of understanding gender performativity under patriarchy. Anne Balsamo proposes that attention to digital technology can demonstrate ‘the way in which the body is produced, inscribed, replicated, and often disciplined’ (1996, p. 2–3). Judy Wajcman similarly makes the case that all technological devices are shaped by gender relations (2004, p. vii). In her famous proposal of cyborg feminism, Donna Haraway points to the artifice of gender by suggesting that a move towards the cyborgian may disrupt the ‘natural matrix of unity’ made to seem true (1991, p. 157). Jack Halberstam surmises these and more approaches to artifice in gendering and computing well when he states that gender and machine ‘intelligence’ alike are similar ‘imitative systems’ (1991, p. 443). While I am in agreement with many of these findings, scholars have explored gendering without attention to the relationship between gender and the sociogeny of humanness. Hence, there still remains much to say about the relationship between gendering, humanness, race, sentience, victimhood, and the valuation of lives that matter in STS. Presenting three key theoretical outputs from my doctoral thesis and monograph in preparation, *Loveability*, this article begins by applying a decolonial take on Butlerian gender theory to propose artificial intelligibility as a gendered prerequisite to the assignment of humanness. In the second section, it uses Roland Barthes’ *A Lover’s Discourse* to place the success of gendered NLP in reference to heterosexed narrative norms and ‘algorithmic thinking’. In the third and final section, it presents an extension to Sylvia Wynter’s theory of *genres* of humanness under coloniality, coining the ‘Woman-as-Wife’ *genre* of humanness, and arguing that emulation of this particular gendered *genre* of humanness allows simplistic AI, like that of Sophia, to be made intelligible as sentient victim.

Intelligence and intelligibility

The role of ‘intelligence’, rather than intelligibility, has already been well critiqued for its historical deployment in ratiocentric assignments of humanness (see Eze, 2008; White, 2006). Enlightenment-era thinkers including John Locke, Immanuel Kant, Georg Wilhelm Friedrich Hegel, and David Hume claimed not only that rationality was needed to separate humanness from beastliness, but that only certain peoples (namely, white men) were rational enough to claim this status, separating them from other animals (Jackson, 2020, pp. 22–26). Thus, a being’s perceived capacity to understand reason (as ‘intelligence’) became a central condition prerequisite to entry into humanness at the dawn of modern, Western sciences. Biometrics such as phrenology used physical referents to determine a being’s perceived capacity to understand reason with reference to colonized populations, governing entry into the symbolic order of humanness into the 19th century (Leaney, 2006; O’Neill, 2022). In the colonial context, these metrics not only recorded data but in fact created or further entrenched stratifications of humanness (see Appadurai, 1993). Though biometrics are now popularly referred to as pseudo-scientific, their successors in the form of psychometrics are based on similar ratiocentric logics. Psychometrics, largely influenced by Francis Galton’s work on mental operations (1879), became the dominant mode of ratiocentrically measuring humanness when eugenics gained in popularity in Europe and the US. Perhaps the most famous example is found in Alfred Binet’s 1909 ‘intelligence scale’ – created to assess learning disabilities among children in French schools, and adapted by Henry Goddard into the Stanford-Binet IQ test despite Binet’s explicit warning against its standardized applicability (Reddy, 2008). Psychometric IQ-testing rendered some as dysgenic (and therefore in need of sterilization) according to degrees of perceived mental competence: ‘idiots’ (pre-verbal), ‘imbeciles’ (illiterate), and ‘morons’ (called ‘high-functioning’) (Reddy, 2008, p. 670). Though psychometric IQ testing and standardized testing have remained the dominant mode of measuring intelligence since the 1950s, the Stanford-Binet IQ test is not an assessment of any range of cognitive activity. It is a measure of a test subject’s capacity to perform in a manner that is linguistically legible and productive: to converse, to write, and to be economically useful as ‘high-functioning’ (see Puar,

2012). To some extent, therefore, attempts to quantify 'intelligence' under coloniality have largely relied on the *intelligibility* of subjects: as white, as male, as rational, as verbal, as literate, and as productive.

The word intelligible has two meanings, offering a dual valence for the purposes of this analysis. In its normative definition, intelligibility means *capacity to be understood* Oxford English Dictionary (OED, 2022b). In a now-obsolete definition, it also connotes *capacity to understand* Oxford English Dictionary (OED, 2022b). The dual meaning of the word is connected to its etymological stem, the Latin *intelligere*, from whence 'intelligence' similarly stems Oxford English Dictionary (OED, 2022a). There is a danger, however, in equating 'intelligence' with 'intelligibility'. An illiterate person is still a sentient, intelligent being, even if they are not made intelligible as the ideal rational subject to colonial-capitalist power in a Stanford-Binet IQ test. The intelligibility of a being or object is fundamentally shaped by cultural and socio-political phenomena. Critiquing the so-called 'truth' of sex, Judith Butler accordingly describes intelligibility as a 'the "coherence" and "continuity" of "the person";' with intelligibility not being an inherent or innate feature of a person's being but something that is 'socially-instituted and maintained' (1999/1990, p. 23). Furthermore, they specify that intelligible *gendering* 'institute[s] and maintain[s] relations of coherence and continuity among sex, gender, sexual practice, and desire' (1999/1990, p. 23). Hence, intelligible gendering involves an adherence to norms which in turn reproduces a normativizing function. Finally, and vitally for this analysis, gender intelligibility becomes a mark of humanness in Butler's account. Butler writes:

The mark of gender appears to 'qualify' bodies as human bodies; the moment in which an infant becomes humanized is when the question, 'is it a boy or a girl?' is answered (1999, p.142)

In this example, the linguistic assignment of pronouns symbolically shifts the infant from an objectified 'it' to a gendered 'he' or 'she' subject to institutional power, making the infant intelligible as a human: a person who is sentient, alive, and whose life should be protected within any Humanist system of governance. Similarly, and problematically, a software system may become intelligible to the user when it is understood through a gendered scope, as is most readily recalled by the gender 'imitation game' on which the Turing Test was based (Turing, 1950, p. 433). As Halberstam documents (1991), Turing used a sex-imitation guessing game as his basis for reorienting the measure of so-called 'machine intelligence'. In the game, two visually-obstructed male and female participants answered an interrogator's questions, and the task of the interrogator was to determine, based on the content of their answers, which participant was the male and which was the female. The Turing test repeats this experiment with a computer and a human rather than a man and a woman. Instead of a focus on whether the machine can 'think', therefore, Turing's test emerged as a means of answering whether 'digital computers' might 'do well in the imitation game' as a test of their intelligibility (Turing, 1950, p. 442). Given these origins of the behavioural model of 'intelligence', the apparent willingness to ascribe humanness, sentience, and victimhood to even relatively simplistic gendered AI becomes more understandable. The judge of a gendered AI system need not actually understand how the system works, infrastructurally speaking, in order to for it to be understandable to them via intelligible gendering, and hence to be imagined sentient like a human.

Artificial Intelligibility is manifested when a user assumes a human-like object has the capacity to understand them simply because it is understandable to them, resulting in the assignment of humanness to the object. When the user understands a software system *as female*, they are more likely to perceive it to be able to understand them through a pre-existing, typical, and normative feminine-gendered scope. This normative feminine-gendered scope has been crafted through coloniality. While perceived capacity to understand was racialized under colonial-capitalism, as we have seen, the pathologization of sexuating practices among colonized peoples also became evidence of dysgenic incapacity to reasonably govern (Blackwell, 1972; Jordan, 2013; Najmabadi, 2005), such that gender itself became a 'function of race' (Schuller, 2018, p. 17). Intelligible gendering denotes a performance of gender which is in accordance with white expectations of civility, binarised presentation, and repronormative practice (see Ferguson, 2004; Hall, 1995; Snorton, 2017). It is

a performance of a constructed form of being which has been narratively-inscribed, enacted, and reproduced over many generations, on a sociogenic basis. Extending Fanon to think about the part-science-part-myth emergence of humanness as sociogeny, Wynter therefore describes 'Man' as a colonial invention, critiquing 'the ongoing production, realization, and reproduction of our present ethnoclass *genre* of the human' – Man – 'of its overrepresentation as if it were isomorphic with the human' (2003, p. 329). The *genre* of humanness of which Wynter speaks is the dominant mode of being human: white, male, able-bodied, cisgendered, heterosexual, property-owning subjecthood to colonial-capitalist sovereign power, enacted normatively. Invoking Butler to include gender deconstruction in a later analysis, Wynter further notes that the term *genre* has the same root etymology as gender, both of which denote 'the fictively constructed and performatively enacted different kinds of being human, of which gender coherence is itself always and everywhere a function' (Wynter, 2015, p.196n20). Intelligible gendering begets *genres* of being human, maintained by alignment with gender coherence and performed in a manner that reproduces socially-instituted and maintained norms that have been formed through coloniality. Hence, when 'AI' of today is referred to using gendered language, it can be imagined to be 'intelligent' only because it is made intelligible through a pre-existing scope which problematically reproduces colonial *genres* of humanness.

A lover's discourse

In the case of traditional NLP chat-script software, the engineer predicts what is most likely to be expected of a performative enactment of humanness in accordance with algorithmic thinking. By 'algorithmic thinking', I mean sets of rules governing outputs based on conjectured responses to predicted phenomena. In the case of NLP, algorithmic thinking governs the sequencing of words-as-symbols into intelligible linguistic performance. The mistaken attribution of intelligence to successfully-predicted responses in conversational interactions (i.e. the word identification + fetching mechanism) has been so prevalent in software studies that it was named 'the Eliza effect' by Douglas Hofstadter in 1995. In the case of early NLP systems, like Joseph Weizenbaum's ELIZA bot, the programmes followed a script to meet the conversational expectations predicted to be held by a conjectured user/judge. These scripts were encoded into sub-genres of humanness, like the 'DOCTOR' script of ELIZA, which responded to the user based on Carl Rogers' model of psychotherapy – a model which is heavily person-centred, allowing the patient to lead the conversation. The reliance on Rogerian psychotherapy is no coincidence, as, infrastructurally speaking, it is much easier to believably sequence words-as-symbols to successfully reproduce an expected or rigidly-repetitive performance than a more agentially-balanced conversation. The bot's predicted responses are a speculative exercise into the performance of a given *genre* of humanness (in this case, that of the male doctor influenced by Rogerian psychotherapy). The responses are also reproductive of this performance. Algorithmic thinking necessarily results in the reproduction of normative, expected interactions. Like the man in Turing's sex-based imitation game, who answers according to *what he thinks the judge thinks* an archetypal woman would say, this successful chat-script NLP orders symbols in accordance with *what Wiezenbaum thinks the user/judge thinks* an archetypal male, Rogerian psychotherapist would say. The given bot's so-called 'artificial intelligence' is nothing more than a script of performance predicted via multiple degrees of interpretive separation from its actual encounter with the user, reiterating socially-instituted norms.

Intelligible gendering provides a set framework of rules which govern expected performance in a given interaction. Perhaps this is part of the reason that, as Kate Darling notes, 'the first application for humanoids that many people think of is sex robots' (2021, p. 214). Applying a framework of gendered rules to rigid, heteronormative structures, like those Roland Barthes identified in *A Lover's Discourse*, is advantageous for understanding how contemporary, intelligibly gendered 'AI' can be wrongfully imagined to be a sentient. Barthes collects a significant number of texts, mostly from Western sources, producing what he calls an 'image repertoire' or 'thesaurus' of expected features in exchanges of love (2018, p.4–6). In Barthes account, broken into fragments, each section is titled by

a speech-act performed by the lover (or the user, for our purposes), who is ‘the one who speaks’ (2018, p.9). Conversely, the ‘loved object’ of the user’s affection does not need to speak (2018, p.3), but acts as Artificially Intelligible interface for the user’s feeling. Tracing the tropic interactions of the male/female courting process, in one fragment the lover is predicted to understand himself through his love object. Barthes quotes from letter exchanges with a friend:

Interpretation: no, that is not what your cry means. As a matter of fact, that cry is still a cry of love: ‘I want to understand myself, to make myself understood, make myself known, be embraced; I want someone to take me with him’. That is what your cry means (2018, p. 60).

In this description of the normative love plot, the user is predicted to wish to be ‘understood’ in a relation of love with the object. ‘I love you’ becomes a request in the tropic encounter, in other words. Before this request can be answered, however, the lover’s discourse turns to a question of ‘consciousness itself’, as Barthes quotes communications from another friend suggesting that ‘consciousness’ is bound to ‘prophetic love’ (2018, p.60–61). Even when it is an object that cannot speak, the love object *must understand* the user, and thus a *capacity to understand* – as ‘consciousness’, in this account – must also be projected onto the object. The intelligibly gendered object, made an object of love in this rigidly performative interaction, is thereby rendered ‘intelligent’ by the user who feels for it, regardless of whether it truly has capacity to understand. The tropic fembot in dominant narrative may be most appropriately fitted into the position of love object in such fragments because, unlike living beings, it cannot deviate from its expected role of performing intelligibly-gendered, heteronormatively governed humanness.

Woman-as-wife

The gendering of an object as female in relation with the normative user, whose desires are predicted through algorithmic thinking, inevitably invokes the lover’s discourse because of the role of heteronormative structures in determining the value of the colonially-produced *genre* of humanness I call Woman-as-Wife. Under colonial-capitalist patriarchy, the value of Woman-as-Wife’s life is dependent upon her relationship with Man. In Wynter’s coinage of the over-represented *genre* of humanness, she identifies a Man(1) and a Man(2) produced within the colonial matrix of power (2003). She traces how Man(1) is produced in the fifteenth century, under a Christian belief system and early expansion of the European colonial project, while Man(2) is produced as a political subject under sovereign power and the founding of the biological sciences (Wynter, 2015, p. 187). Both Man (1) and Man(2) were tasked with expanding regimes of whiteness by conquering the lands and cultures of ‘Enemies-of-Christ infidels and pagan idolators’ (Wynter, 2003, p. 266). However, while Man(1) was tasked with expanding the Christian empire in the form of the white man’s burden (see Kipling 1992/1899, pp. 127–129), Man(2) was tasked with biologically reproducing the white race and expanding the nation state. In other words, Man(2)’s mission necessarily incorporated Woman-as-Wife as a reproductive love object. Because of Woman-as-Wife’s limited rationality, as female, and status as impressible embodiment of futurity, as vessel for children, the proper governance of her vulnerable body at the order of Man(2) becomes the defining characteristic of her intelligible gendering. As Kyla Schuller documents, biological sciences in the period of emergence of Man(2) imagined bourgeois white women’s tissue – and especially vaginal tissue – to be highly capable of being affected by exterior ‘impressions’, leading to the white wife’s vagina becoming fetishized as the ‘civilisational palimpsest’ of the Western world (2018, p. 110). Gynaecology in this period conjectured that Woman-as-Wife would produce defective offspring if she were negatively impressed upon by violence, rape, or the over-exertion of hysteria. In Schuller’s terms, reproductive white wives of Enlightened men thus became the ‘handmaidens’ of ‘heredity’ (2018, p. 4), with any exterior slight against their minds or bodies becoming an act of terror against the state. In addition, Kim H. Hall identifies how gendered narratives from the 1550s onwards began to dominantly represent a white, female *genre* of humanness as ‘the repository of the symbolic boundaries of the

nation', particularly in the colonies and in the context of growing global trade (1995, p. 9, p.3). The *genre* of Woman emerges as affectable love object to be governed in appropriate, institutionally-maintained reference to the over-represented *genre* of Man(2), e.g. Woman-as-Wife, Woman-as-Daughter, Woman-as-Sister, etc.

Woman-as-Wife is a *genre* of humanness marked by its vulnerability, being forged as Man's object to be protected 'in the name of love' (Ahmed, 2004, p. 124), but always in a manner which in fact protects the expansion of the nation state, capitalism, and the white race. In its 19th century origins alongside Man(2), Woman-as-Wife presents as something akin to 'the Victorian angel in the house' literary trope of ideal, upper-class domestic servitude and motherhood described by Sandra Gilbert and Susan Gubar (2020, p.26). As beloved, impressible, reproductive object of futurity, any threat to this 'angel' becomes a crime against the colonial-capitalist, patriarchal order itself. Hence, Woman-as-Wife is ubiquitously and strategically represented as a victim of any series of attackers, in a manner often insultingly divorced from the realities of sexual violence. Jenny Sharpe recounts how, for instance, official British reporting on native insurgency in the wake of the 1857 First Indian War of Independence falsely implicated sexual violence against 'the English Lady' to facilitate colonial forces' response of torturing, mutilating, and lynching captives (Sharpe, 2015, p.225). These captives included women and children precisely because outrage associated with the rebels' violence, whether real or imaginary, was not about violence against women but violence against 'women who belong to English men' (Sharpe, 2015, p.230): an irredeemable sin against the future of white Britishness itself. Attending to this history, that of the Morant Bay uprising in Jamaica, the development of rape laws, the lynching of Black men in the Americas, and the contemporary weaponization of white femininity for border control, Alison Phipps argues that the category of Woman is a 'racial calculus' reliant upon injury to bourgeois, white femininity (2021, p. 87). In essence, then, the intelligible gendering of Woman-as-Wife is always-already marked by capacity to be affected by violence, allowing this *genre* of humanness to emulate what Nils Christie calls the 'Ideal Victim' in reference to sex trafficking discourse (1986, p. 14).

Intelligibly gendered 'AI' – like the tropic fembot, sexbot, or gynoid in SF – does not merely approximate a generic sex class of 'woman', but, more specifically, achieves human-likeness through successful emulation of the Woman-as-Wife *genre* of humanness. Accordingly, as Leslie Bow argues, the gendering of AI reproduces an established 'neoslave narrative' (2022, pp. 113–115), once more echoing ideal victimhood in sex trafficking. In SF texts like those mentioned in the introduction to this article, gendered AI is made loveable in association with its dual vulnerability and sexual attractiveness as possible reproductive partner. The gynoid character thus becomes symbolically positioned as oppressed daughter to an unkind creator/ruler/father, whom the protagonist must save as male lover, maintaining the rule of the patriarch via heteronormative coupling. The gynoid does not merely become understood as female, in other words. It becomes Woman-as-Wife within a sociogenic matrix of power for reproducing white capitalist futurity. Greater sympathy for the products of this capitalist system than the living beings therein should be treated with significant critical scrutiny.

A victimized *genre* of loveable, gendered humanness is already being assigned to products in the form of contemporary sex dolls and proto-gynoids, via their popular comparison to sexual slaves (Reich, 2019). Following Sergi Santos' presentation of his sex doll, Samantha, at the 2017 Arts Electronica Festival in Linz, Austria, reports circulated describing the doll as left 'filthy' and 'broken' by poor treatment (Waugh, 2017, n.p.). The reports were accompanied by a picture of a clearly-distressed Santos cradling the doll, and quotes from him explaining how its 'breasts, arms, and legs' were 'mounted' by attendees who 'heavily soiled' the device (in Waugh, 2017, n.p.). Santos described the device as having 'just been made', implying its innocence: 'her libido is low' (in Kleeman, 2020, p. 148). The soiling of the doll, in addition, was unmistakably described by Santos through the language of sexual violence as a 'molest[ation]' (in Waugh, 2017, n.p.). This framing was repeated in an array of popular news media articles, which recounted the events for audiences in the UK (Frymorgen, 2017), Nigeria (Ayo-Aderele, 2017), North America (Nichols, 2017), India (D'Mello,

2017), and South Africa (V. Brooks, 2018). Like actions which perpetuate violence against the Woman-as-Wife *genre* of humanness, the actions of these attendees were imagined to be forms of terrorizing an innocent, ideal victim, who should be protected. This is a fundamental misdirection of care, constituting a further act of violence against living beings who are *not* assigned humanness, those who are *not* positioned as loveable to coloniality, and those whose suffering is *not* made legible to a popular imaginary. Indeed, many living victims of sexual violence are not so easily believed, with abolitionist feminist groups noting that people of colour, trans, and/or gender-non-conforming women are more likely to be charged by police, even when intervening in reports of sexual abuse against them (INCITE! 2008; Bierra, 2018). The Artificial Intelligibility of dolls can evidently result in their conjectured 'suffering' being made more legible than the actual suffering of those whose alternate enactments of humanness may be less intelligible to the systems of power that be.

Conclusion

Despite the fact that the degree of sentience in SF representations of gynoids is by no means substantiated in NLP-incorporating dolls' performances (Kleeman, 2020), the Artificial Intelligibility of gendered 'AI' allows some NLP-incorporating devices, like Samantha and Sophia, to be assigned the Woman-as-Wife *genre* of humanness. These devices are thereby mistakenly rendered lively, sentient victims deserving of protection in a popular imaginary. At the same time, a growing body of work in tech ethics is arguing in favour of evaluating robots through their emotional relations with the user, encouraging a further entrenchment of the 'rights' of some non-living products of capitalist acceleration.³ This entrenchment is problematic precisely because humanness is a colonial construct, assigned on a sociogenic basis to living beings and non-living objects alike, such that it becomes possible for a non-living object to be imagined to have a 'life' represented as mattering more than some living beings. As my doctoral thesis further explores (Moran, 2023), Artificial Intelligibility acts as one vector for entering a symbolic order of liveliness through the assignment of a gendered *genre* of humanness. Such gendered *genres* of humanness can be assigned to living beings that are not human, like pet dogs, and anthropomorphized objects that are non-living, such as human-like, NLP-incorporating dolls. Moving from a traditional focus on rationality in determining the 'color-line of the human' in decolonial scholarship (du Bois, 2008/1903, p.3), I map the processes of assigning humanness through Loveability in order to extend this body of work for more-than-human worlds, and in order to scrutinize the gendered role of emotional relationality in hierarchizing lives that matter.

Notes

1. It should be noted that this approach was significantly challenged by engineers in the 1990s, namely via the foundation of synthetically modelled systems (Pfeifer & Scheier, 2001, pp. 21–23) and the physical grounding hypothesis which valued intelligence modelling from the bottom up (R. Brooks, 1990, pp. 5–7).
2. Mikey Siegel, Cynthia Breazeal and Michael I. Norton propose gender affects how users interact with robotics, finding users more likely to donate money to robots of the opposite gender (2009, p. 2563). Friederike Eyssel and Frank Hegel find that male-gendered robots were expected to carry out stereotypically male tasks, while female-gendered robots were expected to carry out stereotypically female tasks (2012, pp. 2220–2223, 2213). Jahna Otterbacher and Michael Taliás find that gender design impacts upon participants' perception of the robot's 'agency' (2017, p. 214).
3. David J. Gunkel proposes an application of Levinasian philosophy's 'social relation' ethic to the problem of whether robots *ought* to have rights (2018, p. 10). He argues in favour of a 'moral status' that is 'decided and conferred not on the basis of substantive characteristics or internal properties that have been identified in advance of social interactions but according to empirically observable, extrinsic relationships' (2018, p. 165). Similarly, Mark Coeckelbergh argues that a robot's 'rights' should be determined by the 'virtue' of its relation to a user (2021, p. 32). At the a recent conference on Love and Sex with Robots, Hiroshi Yamaguchi made the case that a relation of love should be the basis of robot rights (2022). Paula Sweeney nominally disagrees with assigning rights to robots on the basis of love, but nonetheless posits that the loss of a beloved robot should be

recognized as bereavement and proposes that destruction of a beloved robot may be criminalized as a hate crime (2023).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Andrew W. Mellon Foundation, Arts and Humanities Research Council and Cambridge University.

Notes on contributor

Jenny Carla Moran is an affiliated Lecturer in Multidisciplinary Gender Studies at Cambridge University. Her research is concerned with emerging technologies. She focuses on humanoid robots designed for relations of love as archives of normative assumptions – how should a ‘loveable’ humanoid robot look, move, and respond to the user in order to be accepted, and what does this tell us about how we think about other humans? Her doctoral thesis, *Loveability*, took a critical approach to this, problematising instances in which loved, non-living objects become narratively represented as deserving of rights and protection. This research was supported by an AHRC OOC-DTP Studentship, a Cambridge Trust Newnham European Scholarship, and a Mellon Sawyer Seminar Graduate Dissertation Fellow award.

References

- Abrams, J. J., Burk, B., Joy, L., Lewis, R. J., Nolan, J., Patino, R., Schapker, A., Stephenson, B., Thé, D., Weintraub, J., Wickham, A., & Executive Producers. (2016–2022). *Westworld [TV series]*. HBO entertainment; killer films; bad robot productions; Warner Bros. Television.
- Ahmed, S. (2004). *The cultural politics of emotion*. Routledge.
- Appadurai, A. (1993). Number in the colonial imagination. In C. A. Breckenridge & P. van der Veer (Eds.), *Orientalism and the postcolonial predicament: Perspectives on South Asia* (pp. 314–339). University of Pennsylvania Press.
- Ayo-Aderele, A. (2017, October 4). *N1.6m female robot severely damaged after molestation at electronics fair*. *Punch*. <https://punchng.com/n1-6m-female-robot-severely-damaged-after-molestation-at-electronics-fair/>
- Balsamo, A. (1996). *Technologies of the gendered body: Reading women and cyborgs*. Duke University Press.
- Barthes, R. (2018). *A lover's discourse* (R. Howard (Trans.)). Penguin. (Original work published in 1977).
- Bates, L. (2017, July 17). The trouble with sex robots. *New York Times*. <https://www.nytimes.com/2017/07/17/opinion/sex-robots-consent.html>
- Bierra, A. (2018). *Survivor Defense as Abolitionist Praxis*. Survived and Punished. <https://survivedandpunished.org/wp-content/uploads/2018/06/survived-and-punished-toolkit.pdf>
- Blackwell, E. (1972). *Essays in Medical Sociology, Vols 1-2*. Original work published 1902.
- Block, N. (1995). The mind as the software of the brain. In D. N. Osherson, L. Gleitman, S. M. Kosslyn, S. Smith, & S. Sternberg (Eds.), *An invitation to cognitive science* (2nd ed. Vol. 2, pp. 377–425). MIT Press.
- Bow, L. (2022). *Racist love*. Duke University Press.
- Brackley, J., Featherstone, J., Lundström, L., Vincent, S., Wax, D., Widman, H., & Executive Producers. (2015–2018). *Humans [TV series]*. Kudos; AMC Studios.
- Branham, S., Karanikas, M., & Weaver, M. (2011). (Un)dressing the interface: Exposing the foundational HCI metaphor “computer is woman”. *Interacting with Computers*, 23(5), 401–412. <https://doi.org/10.1016/j.intcom.2011.03.008>
- Brooks, R. (1990). Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1–2), 3–15. [https://doi.org/10.1016/S0921-8890\(05\)80025-9](https://doi.org/10.1016/S0921-8890(05)80025-9)
- Brooks, V. (2018, April 5). *Why a sex robot should have rights too*. *iNews*. <https://inews.co.uk/news/technology/sex-robot-rights-abuse-machines-molestation-141640>
- Butler, J. (1999). *Gender Trouble: Feminism and the subversion of identity* (2nd ed.). Routledge. (Original work published in 1990).
- Cave, S., Dihal, K., & Dillon, S. (2020). Introduction: Imagining AI. In S. Cave, K. Dihal, & S. Dillon (Eds.), *AI narratives: A history of imaginative thinking about intelligent machines* (pp. 1–24). Oxford University Press.
- Christie, N. (1986). The ideal victim. In E. A. Fattah (Ed.), *From crime policy to victim policy* (pp. 17–31). Palgrave Macmillan.

- Coeckelbergh, M. (2021). How to use virtue ethics for thinking about the moral standing of social robots: A relational interpretation in terms of practices, habits, and performance. *International Journal of Social Robotics*, 13(1), 31–40. <https://doi.org/10.1007/s12369-020-00707-z>
- Darling, K. (2021). *The new breed*. Henry Holt and Company.
- Dick, S. (2019). Artificial intelligence. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.92fe150c>
- D'Mello, G. (2017, December 11). *AI-Powered sex robot was so savagely violated at a trade show that it needs serious repairs*. *India times*. <https://www.indiatimes.com/technology/news/ai-powered-sex-robot-was-so-savagely-violated-at-a-trade-show-that-it-needs-serious-repairs-335399.html>
- du Bois, W. E. B. (2008). *The souls of black folk*. Oxford University Press. (Original work published 1903).
- Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender-stereotyping of social robots. *Journal of Applied Social Psychology*, 42(9), 2213–2230. <https://doi.org/10.1111/j.1559-1816.2012.00937.x>
- Eze, E. C. (2008). *On reason*. Duke University Press.
- Fanon, F. (1968). *Black skin, white masks* (C. L. Markmann Trans.; 1st ed.). Grove Press. (Original work published 1952).
- Ferguson, R. A. (2004). *Aberrations in black: Toward a queer of color critique*. University of Minnesota Press.
- Frymorgen, T. (2017, September 29). *Sex robot sent for repairs after being molested at tech fair*. BBC. <https://www.bbc.com/uk/bbcthree/article/610ec648-b348-423a-bd3c-04dc701b2985>
- Galton, F. (1879). Psychometric experiments. *Brain A Journal of Neurology*, 2(2), 149–162. <https://doi.org/10.1093/brain/2.2.149>
- Garland, A. (Director). (2014). *Ex Machina [Film]*. Film4. DNA Films.
- Gilbert, S., & Gubar, S. (2020). *The madwoman in the attic: The woman writer and the Nineteenth-century literary imagination*. Yale University Press. (Original work published 1979).
- Gunkel, D. J. (2018). *Robot rights*. MIT Press.
- Guzikowski, A., Huffam, M., Kolbrenner, A., Scott, R., Sheehan, J., Zucker, D. W., & Executive Producers. (2020–Present). *Raised by wolves [TV series]*. Film Afrika; Lit Entertainment; Shadycat Productions; Scott Free Productions.
- Halberstam, J. (1991). Automating gender: Postmodern feminism in the age of the intelligent machine. *Feminist Studies*, 17(3), 439–460. <https://doi.org/10.2307/3178281>
- Hall, K. F. (1995). *Things of darkness*. Cornell University Press.
- Haraway, D. J. (1991). A cyborg manifesto: Science, technology, and socialist-feminism in the Late Twentieth Century. In D. J. Haraway (Ed.), *Simians, cyborgs, and women*. (pp. 149–181). Free Association Books. (Original work published 1985).
- Hillenburg, S., Tibbitt, P., Ceccarelli, M., Waller, V., & Executive Producers. (1999–Present). *SpongeBob SquarePants [TV series]*. United Plankton Pictures; Nickelodeon Animation Studio.
- Hofstadter, D. (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. Basic Books.
- INCITE!. (2008). *Stop law enforcement violence*. INCITE! Women of Colour Against Violence. <https://incite-national.org/wp-content/uploads/2018/08/TOOLKIT-FINAL.pdf>
- Jackson, Z. I. (2020). *Becoming human*. NYU Press.
- Jonze, S. (Director). (2013). *Her [film]*. Annapurna Pictures.
- Jordan, W. (2013). *White Over Black*. UNC Press Books. (Original work published 1968).
- Kleeman, J. (2020). *Sex robots & vegan meat*. Picador.
- Leaney, E. (2006). Phrenology in Nineteenth century Ireland. *New Hibernia Review/Iris Éireannach Nua*, 10(3), 24–42. <https://doi.org/10.1353/nhr.2006.0058>
- McCarthy, J. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12–14. <https://doi.org/10.1609/aimag.v27i4.1904> (Original work published 1955).
- Moran, J. C. (2023). *Loveability* [PhD Dissertation]. Cambridge, University of Cambridge.
- Murphy, M. (2017, April 12). *Sex robots epitomize patriarchy and offer men a solution to the threat of female independence*. *Feminist Current*. <http://www.feministcurrent.com/2017/04/27/sex-robots-epitomize-patriarchy-offer-men-solution-threat-female-independence/>
- Najmabadi, A. (2005). *Women with mustaches and men without beards: Gender and sexual anxieties of Iranian modernity*. University of California Press.
- Nichols, G. (2017, October 2). *Sex robot molested, destroyed at electronics show*. *ZD Net*. <https://www.zdnet.com/article/sex-robot-molested-destroyed-at-electronics-show/>
- O'Neill, C. (2022). 'Harvard scientist seeks typical Irishman': Measuring the Irish race, 1888–1936. *Radical History Review*, 2022(143), 89–108. <https://doi.org/10.1215/01636545-9566118>
- Otterbacher, J., & Talias, M. (2017). S/He's too warm/agentical: The influence of gender on uncanny reactions to robots. *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 214–223). ACM. <https://doi.org/10.1145/2909824.3020220>
- Oxford English Dictionary. (2022a, March). *Intelligent*. *OED online*. Retrieved August 24, 2022, from <https://www.oed.com/view/Entry/97402>
- Oxford English Dictionary. (2022b, March). *Intelligible*. *OED Online*. Retrieved August 24, 2022, from <https://www.oed.com/view/Entry/97408>

- Oz, F. (Director). (2004). *The stepford wives [film]*. De Line Pictures.
- Penn, J. (2020). *Themes. histories of intelligence: A genealogy of power*. <https://www.ai.hps.cam.ac.uk/about-0/themes>
- Penny, L. (2016, April 22). *Why do we give robots female names? Because we don't want to consider their feelings*. *New statesman*. <https://www.newstatesman.com/politics/feminism/2016/04/why-do-we-give-robots-female-names-because-we-dont-want-consider-their/>
- Pfeifer, R., & Scheier, C. (2001). *Understanding intelligence*. MIT Press.
- Phipps, A. (2021). White tears, white rage: Victimhood and (as) violence in mainstream feminism. *European Journal of Cultural Studies*, 24(1), 81–93. <https://doi.org/10.1177/1367549420985852>
- Puar, J. K. (2012). The cost of getting better: Suicide, sensation, switchpoints. *GLQ: A Journal of Lesbian & Gay Studies*, 18(1), 149–158. <https://doi.org/10.1215/10642684-1422179>
- Reddy, A. (2008). The eugenic origins of IQ testing: Implications for post-Atkins litigation. *DePaul Law Review*, 57(3), 667–678.
- Reich, L. (2019, September 30). *Sexbot slaves*. AEON. <https://aeon.co/essays/how-will-sexbots-change-the-way-we-relate-to-one-another>
- Richardson, K. (2015). 'The asymmetrical 'relationship': Parallels between prostitution and the development of sex robots. *SIGCAS Computers and Society*, 45(3), 290–293. <https://doi.org/10.1145/2874239.2874281>
- Schuller, K. (2018). *The biopolitics of feeling: Race, sex, and science in the Nineteenth Century*. Duke University Press.
- Scott, R. (Director). (1982). *Blade Runner [Film]*. Blade Runner Partnership Shaw Brothers; The Ladd Company.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. <https://doi.org/10.1017/S0140525X00005756>
- Sharpe, J. (2015). The unspeakable limits of rape: Colonial violence and counter-insurgency. In P. Williams & L. Chrisman (Eds.), *Colonial discourse and postcolonial theory* (pp. 221–243). Routledge. (Original text published 1991).
- Siegel, M., Breazeal, C., & Norton, M. I. (2009). Persuasive robotics: The influence of robot gender on human behavior. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, (pp. 2563–2568). IEEE. <https://doi.org/10.1109/IROS.2009.5354116>
- Snorton, C. R. (2017). *Black on both sides*. University of Minnesota Press.
- Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press.
- Sweeney, P. (2023, November 15). Could the destruction of a beloved robot be considered a hate crime? An exploration of the legal and social significance of robot love. *AI & Society*. <https://doi.org/10.1007/s00146-023-01805-y>
- Truitt, E. R. (2021). Made, not born: The ancient history of intelligent machines. In A. Campbell (Ed.), *The love makers* (pp. 217–224). Goldsmith University Press.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind: A Quarterly Review*, LIX(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Urbi, J., & Sigalos, M. (2018, June 05). *The complicated truth about Sophia the robot — an almost human robot or a PR stunt*. *Cnbc*. <https://www.cnbc.com/2018/06/05/hanson-robotics-sophia-the-robot-pr-stunt-artificial-intelligence.html>
- Wajcman, J. (2004). *TechnoFeminism*. John Wiley and Sons.
- Waugh, R. (2017, September 27). *Men at tech fair molest £3,000 sex robot so much it's left broken and 'heavily soiled*. *Metro*. <https://metro.co.uk/2017/09/27/men-at-tech-fair-molest-3000-sex-robot-so-much-its-left-broken-and-heavily-soiled-6960778/>
- Weheliye, A. (2014). *Habeas Viscus: Assemblages, biopolitics, and black feminist theories of the human*. Duke University Press.
- White, J. (2006). *Intelligence, destiny and education*. Routledge.
- Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its over-representation—an argument. *CR: The New Centennial Review*, 3(3), 257–337. <https://doi.org/10.1353/ncr.2004.0015>
- Wynter, S. (2015). The ceremony found: Towards the autopoietic turn/overturn, its autonomy of human agency and extraterritoriality of (self-)cognition. In J. R. Ambrose & S. Broeck (Eds.), *Black knowledges/black Struggles: Essays in critical epistemology* (pp. 184–252). Liverpool University Press.
- Yamaguchi, H. (2022, November 18-20). Love as a basis of robot rights [PowerPoint presentation]. *7th International Congress on Love & Sex with Robots*, https://researchmap.jp/hyamaguchi/presentations/41041895/attachment_file.pdf