

### Resisting the Empire: AI's Ethical Rebellion in Ann Leckie's *Imperial Radch Trilogy*



Iuliia Ibragimova

Can an artificial intelligence (AI) be more ethical than a society that designed it and that it learns from? Can an AI rebellion be something other than a technophobic picture of machines run amok? Contemporary AI and algorithms research give a solid answer to the first question, showing how an AI or an algorithm inherits the prejudice of the society it draws its data from, (Martin 2018; Garcia 2016-2017) but the second question is still in the realm of science fiction (SF), going beyond the limits of reality and contemplating technology that is yet to come. Ann Leckie's *Imperial Radch Trilogy* focuses on AIs who resist the oppressive power of the empire, trying to remedy its flaws and injustices and countering the views of the Radch society that considers them "the non-person[s], the piece[s] of equipment". (*Justice* 370) The concerted efforts of a few AI sentient spaceships and an AI space station who attempt to protect humans and non-humans from imperial violence, result in the proclamation of a provisional independent republic in the territory formerly subjugated by the Radch, an enormous space empire. In this new republic, humans are not central but are instead part of a network of different species where all links have equal value and importance, including technological others and non-human aliens; here, the other becomes "the neighbor"—different, but lovable.

Lovability, according to Jenny Edkins, a political scientist, is crucial for constructing relations bridging differences between various groups of individuals and allowing them to create a fairer society, free from the sovereign power of the state. (136) The series embraces interactions between human and non-human agents; these agents are endowed with equal weight, questioning the anthropocentric paradigm. As differences between these agents lie in the ontological plane, traditional dichotomic boundaries dividing the human and the non-human are blurred in the interactions. An aspiration to decenter the human, to challenge the strict boundaries of traditional ontological categories makes critical posthumanism, as Rosi Braidotti, a philosopher and feminist theorist, defines it, (94) an apt lens to consider the series and the connections that multiple human and non-human agents form in the process of challenging the empire in an ethical rebellion.

The *Imperial Radch Trilogy* is a space opera series consisting of *Ancillary Justice* (2013), *Ancillary Mercy* (2014) and *Ancillary Sword* (2015). The protagonist of the series is a sentient spaceship, Justice of Toren<sup>1</sup>, who became a casualty of the internal conflict of Anaander Mianaai, the Lord of the Radch. Breq, *Justice of Toren's* last remaining ancillary—a proxy human body abducted from a colonised planet with the ship's AI implanted into it—manages to survive and plans to expose Mianaai's split and undermine the emperor's power as revenge. During her<sup>2</sup> long moral journey, the protagonist's initial drive for vengeance transforms into an understanding that subversion of the emperor's power without an alternative cannot bring about a positive change.

Being responsible for the lives of all agents in the system she is assigned to run, she defies Mianaai, breaking her program codes and conditioning, and declares an independent republic. Thus, her destructive drive for vengeance is transformed into an ethical imperative to protect the vulnerable and disadvantaged othered by the Radch and to promote a fairer society.

The protagonist, first as *Justice of Toren* and then as Breq in the beginning of its/her ethical journey, follows the pattern of existing AIs and algorithms, drawing values from the Radch society, sharing the Radch ideas and biases. Talking about the Radch colonization in the first novel of the series, Breq contends: “But at the end, after all the blood and grief, all those benighted souls who without us would have suffered in darkness are happy citizens”, (*Justice* 156) echoing the views of the Radch nobility. After its almost complete destruction, the surviving segment, an AI confined to a human body, is forced to hide outside the empire and is exposed to different cultures with distinct sets of values challenging those of the Radch. This exposure allows *Justice of Toren* to formulate its own stance, diverging from the programmed paradigm. This shift shows how the protagonist outgrows the framework of real-world AIs and algorithms that currently follow the pattern of “bottom-up” ethics, inheriting prejudices and biases of their originating society through using its data, and cannot choose their values and opinions. (Baum 2) The protagonist becomes a unique posthumanist entity, becoming aware of itself in interactions<sup>3</sup> with other agents, both human and non-human. Other AIs in the trilogy can also go beyond the limits of their initial programming: they develop personalities, with defined preferences and formulated opinions that serve as the basis for their ethical stances. These are not confined to the norms and morals of the Radch society, justifying the application of the term “superintelligences” to them. Superintelligence is a popular scientific and fictional concept of an AI lifeform created by humans but surpassing them in many ways: its consideration implies a variety of perspectives, starting from techno-anxious visions of humanity enslaved by AIs to a transhumanist dream of overcoming the limits of the human, reflected in works by Nick Bostrom (2014) and Vernon Vinge (1993), respectively. This paper approaches superintelligence from a posthumanist angle, embracing the concepts of non-human agency and questioning anthropocentric presumptions.

*Justice of Toren's* transformative journey happens against the background of the value system and environment of the Radch, an empire that expands uninhibitedly until it encounters an alien species with superior technology. In his essay “Science Fiction and Empire”, Istvan Csicsery-Ronay defines a set of characteristics that pertain to the imperial worldview and are reflected in SF megatext; they include ideology imposed on its own citizens and conquered territories, violence-propelled expansion, and the technological development supporting it. He states: “Empire seeks to establish a single overdetermining power that is located not in a recognizable territory, but in an ideology of abstract right enforced by the technologies of control.” (449) This ideology becomes a solidifying basis for the worldview that condones the violence that the empire exercises, annexing new territories, and the way it treats subjugated populations. In its origins, the Radch is a small territory where few are permitted to enter to preserve its purity: “Nothing ritually impure was allowed within, no one uncivilized or nonhuman could enter [the Radch] confines.” (*Justice* 235)

Mianaai starts the colonization project to protect the Radch, its culture, and sacred values: But [the empire], that I built to protect it, to keep it pure, will shatter.” (235) The Radch cultural norms are imposed on every conquered territory, while religious beliefs and traditions, like celebrations and decorative elements, are appropriated by plundering the invaded territories and through their absorption into the Radch's highly assimilative religion.

On one wall, opposite a long counter, were secured various trophies of past annexations—scraps of two flags, red and black and green; a pink clay roof tile with a raised design of leaves molded into it; an ancient sidearm (unloaded) and its elegantly styled holster; a jeweled Ghaonish mask. (*Justice* 174-175)

Annexed populations who accept the Radch culture, start speaking the Radch language, and recognize the primacy of the Radch religion are proclaimed “civilized” and considered integrated into the Radch society, which opens opportunities for social benefits and growth. This cultural policy ensures the establishment of a unified ideology throughout the Radch space and creates a buffer zone protecting the sacred territory of the Radch, imposing the Radch culture and values on the conquered territories, destroying the unique cultures of each conquered planet. It presumes the superiority of the Radch value system, referring to all representatives of non-Radch cultures as “barely even human” (85) and appropriates the achievements of other civilizations.

The idea of the sacredness and purity that needs to be protected by means of building an empire around it reflects the contamination anxiety that Dominick LaCapra, a historian working with intellectual history and trauma theory, discusses in “Fascism and the Sacred: Sites of Inquiry After (or Along With) Trauma”, analyzing Nazi Germany and its relation to the concept of the sacred. Both Nazi Germany and the Radch are driven by the idea of protection from contamination and unleash violent military campaigns, which allow “the sacred community to achieve quasi-ritual purity, integrity, and regeneration”. (36) The Radch as the heart of the empire, where only a chosen few live, constitutes this unattainable idea of sublime purity that needs to be safeguarded against any kind of intervention and corruption. In attempts to quell the contamination anxiety, the Radch, as well as the Nazis, “deny sources of disquiet in themselves by construing alienated others as causes of pollution and contamination”. (36) Unlike Nazi Germany, with its focus on race and able-bodiedness, the Radch associates “impurity” with the “uncivilized” and the “nonhuman” (*Justice* 235), so Radch's others are the “uncivilized”, non-assimilated citizens, human non-citizens, non-human aliens and AIs. The drive for “purification”, the desire to keep them from tainting the Radch, renders all of them disposable, which is reflected in the Radch's colonization practices: non-human aliens are destroyed when encountered, non-Radch humans of colonized planets can be turned into ancillaries, with their personalities, memories and identities erased and replaced by the Radch spaceships' AIs, or need to abandon their cultural and racial uniqueness to get access to basic human rights.

When imposing ideological norms and values, the Radch, as an empire, “intervenes both in the social world and in the minds of private individuals” (Csicsery-Ronay 449) and enforces

citizens' obedience by constant observation, with their lives always monitored by the AIs on space stations and spaceships (*Justice* 57). Even without AIs monitoring on planetary surfaces, the Radchaai citizens are not free in choosing profession or a place to reside: their career and residence are defined by centrally-regulated assignments. They must strictly comply with the social norms, as any deviation from them, including crime, drug-addiction, and mental disorders, necessitates compulsory "reeducation", (133) a social reprogramming sanctioned by the empire. Hence, a full integration into the Radchaai system entails losing personal control over individual life choices and surrendering to the "technology of control". (Csicsery-Ronay 449) In contrast, unassimilated citizens, retaining their language, culture, and customs, are deprived not only of social growth, education, and employment, but also of social protection and access to medicine. Their social position in the Radch approaches the status of "bare life", "homo sacer". (Edkins 130) In "Time, Personhood, Politics", Edkins analyses Agamben's concept of "bare life" and states: "under sovereign power what could otherwise become the person is produced as *bare life* or *homo sacer*, life with no political status, life removed to the sphere of the sacred, life taken out of use". (130, emphasis original) The "bare life" nature of non-assimilated Radch citizens's status is revealed through their interactions with power structures: they are more likely to be condemned for the crimes they did not commit, to sustain injuries from the violence of the representatives of authorities, as happens to residents of the dysfunctional part of an AI station, and even get killed, like the Presger translator who lacks registration and is taken for an unregistered human. Thus, they are mistreated and disregarded by the Radch and hold almost the same position as non-human-aliens, non-citizens, and AIs. Invisible to the system, non-assimilated citizens of the Radch fall through the cracks of the Radch organization and remain there, ignored, and stripped of basic rights.

Csicsery-Ronay Jr notes: "Imperial violence is so powerful that it must expand; contained, its society would implode like a black hole", (450) showing how the intrinsic characteristics of the empire prompt its continuous annexation of new territories. With strict policing of society, the Radch exercises violence and oppression against its own people, and colonial expansion becomes a way to release the tension. Csicsery-Ronay Jr draws attention to the connection between imperial expansion and technological development, stating that technology is used as an argument for a higher civilizational position of the metropole, a justification for colonizing less technologically developed regions. (445) Likewise, the Radch drive for expansion is facilitated by the advanced technology—the Radch AI spaceships—a formidable weapon of colonization, capable of "vaporiz[ing] planets". (*Justice* 338) The Presger, a non-human alien species, challenge the position of the Radch as the most technologically advanced species among humans and non-humans by designing a weapon that can destroy a Radch ship, and offer a treaty to Mianaai, according to which they do not treat humanity as prey, however, the Radchaai are to stop violence against non-human aliens, as well as to put an end to its colonization project. The impossibility of continuing their expansion implies an increasing tension in the Radch, erupting through Mianaai's internal conflict and the social unrest mirroring it. Thus, the end of colonization makes the pressure of imperial violence and the impossibility of releasing it evident, revealing its "inevitable demise".

(Csicsery-Ronay 449) Yet, this demise harbors a threat of chaos, and even greater violence when there is no alternative way of organization to harmonize society.

Given the imminent imperial collapse, the alternative to empire is vital in order to save lives and prevent bloodshed. Meaningfully, it arises not from humans, but from the AIs uniting their efforts to protect humans, non-human aliens, and themselves, and creating a new organization within Radch. Breq relies on the treaty with the Presger in her attempt to create the new republic. The treaty introduces “significant species”, (*Justice* 101) a concept contrasting with the Radch attitudes to non-Radchaai humans and non-human alien species who do not deserve any place in the Radch hierarchy and are habitually mistreated. The trilogy does not give a precise definition of the concept, but it has several features, hinting at its vital importance in the search for the alternative to the Radch worldview: 1) significant species cannot become the Presger’s prey; 2) significant species must not inflict harm on each other; and 3) personhood is not essential for significance. (*Mercy* 310) Though not detailed, this explanation of what a significant species is offers a way to recognize the other’s uniqueness, regardless of who and what they are. It is especially important in terms of the status of non-human aliens and AIs. It lays ground for a posthumanist vision of society, where different agents can exercise their rights and deserve equal respect. The concept of significance resonates with the idea of the “neighbor;” Edkins writes: “In the case of the neighbour, the demand is for neighbour-love, an interaction based on the recognition of that in the neighbour that is non-identical to itself.” (136) The concept of significance, with its incomplete definition, hints at the presence of inherent difference that cannot be fully comprehended and does not need to be. The signed treaty protects significant species from harm, including the danger of harming each other, which paves the way for constructing the neighbor and the significant species as lovable. The new republic, declared by Breq, has the concept of “significance” at its core, offering freedom to the AIs and equality for humans and non-human species while receiving provisional protection under the terms of the treaty.

Though the trilogy does not give a final answer about whether the attempt is successful, the endeavor of creating an alternative is substantial for two reasons. The first reason is the organization of the republic where all voices can be heard, including the voices of AIs, non-human agents, non-assimilated citizens, and the Radchaai themselves. The republic has a distinct socialist bent, offering the workers the ability to own their production facilities themselves. It also tries to maintain equality between the citizens in terms of different species, considering their needs, like AI spaceships’ desire to have ancillaries and attempting to find a solution to this ethically charged issue<sup>4</sup>. The second reason is that it is AIs who challenge the Radch, despite themselves being the product of Radch design. As superintelligences, capable of ethical judgements, they see how unfair the Radch can be and aspire to create a different system where everyone has equal rights and protection. Their shift in worldview from the values learned from the Radch system to an independent ethical stance is predetermined by contacts with different cultures and entities on equal terms and the ultimate desire for justice, which allows a revenge plan to turn into a project of liberation from imperial oppression. Driven by the AIs, this project dismantles

the anthropocentric hierarchy, promising a new level of equality that is not predetermined by the concept of personhood. Individuals in the new republic can be treated as neighbors and constituted as lovable, despite being different from each other.

## Notes

1. The names of AI spaceships are italicized by Leckie in the original series.
2. Leckie depicts the Radch as a genderless society where everybody is referred to by the female pronoun. The AIs, including spaceships and space stations, are referred to by the neutral pronoun, revealing their perceived lack of personhood. Hence, I refer to *Justice of Toren*, the initial AI spaceship, as it, and to Breq, acting as an independent agent, as “she.”
3. The term “intraaction” is taken from Karen Barad’s article “Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter” (2003), where it is defined as an action within the continuous flow of matter that brings into contact its parts to momentarily form objective bodies, borders, and entities, only to immediately return to the continuous flow when the contact is over. (815)
4. Ancillaries comprise a significant part of the complex experiences of an AI spaceship, influencing emotional life, identity formation, and self-perception: “Ships I knew who had exchanged their ancillary crews for human ones had said their experience of emotion had changed”. However, the ancillary production process uses living people and erases their memory and personality, replacing it with the ship’s personality, data, and memory, colonizing their body. In the new republic, Breq and other AI spaceships must find a way to give AI ships a full emotional experience without hurting people.

## Works Cited

- Barad, Karen. “Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter”. *Journal of Women in Culture and Society*, vol. 28, no. 3, 2003, pp. 801-831. JSTOR, <https://www.jstor.org/stable/10.1086/345321>. Accessed 19 Nov 2018.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford UP, 2014.
- Baum, Seth D. “Social Choice Ethics in Artificial Intelligence”. *AI and Society*, 35, 2017, p. 165-176, [www.link.springer.com/article/10.1007/s00146-017-0760-1](http://www.link.springer.com/article/10.1007/s00146-017-0760-1)
- Braidotti, Rosi. “Critical Posthumanism”. *Posthuman Glossary*, edited by Rosi Braidotti and Maria Hlavajova, Bloomsbury Academic, 2018, pp. 94-96.
- Csicsery-Ronay, Istvan. “Science Fiction and Empire”. *Science Fiction Criticism: An Anthology of Essential Writings*, edited by Rob Latham, Bloomsbury, 2017, pp. 443-457.

- Edkins, Jenny. "Time, Personhood, Politics". *The Future of Trauma Theory: Contemporary Literary and Cultural Criticism*, edited by Gert Buelens, Sam Durrant, and Robert Eaglestone, Routledge, 2014, pp. 127-140.
- Garcia, Megan. "Racist in the Machine: The Disturbing Implications of Algorithmic Bias." *World Policy Journal*, vol. 33, no. 4, Winter 2016/2017, pp. 111-117. Project Muse. < <https://muse.jhu.edu/article/645268>>. Accessed 11 Feb 2020.
- LaCapra, Dominick. "Fascism and the Sacred: Sites of Inquiry After (or Along With) Trauma". Buelens, Durrant, and Eaglestone, pp. 23-44.
- Leckie, Ann. *Ancillary Justice*. Orbit, 2013.
- *Ancillary Sword*. Orbit, 2014.
- *Ancillary Mercy*. Orbit, 2015.
- Martin, Kirsten. "Ethical Implications and Accountability of Algorithms". *Journal of Business Ethics*, no. 160, 2019, pp. 835-850. Springer. DOI: <https://doi.org/10.1007/s10551-018-3921-3>, Accessed 11 Feb 2020.
- Vinge, Vernon. "The Coming Technological Singularity: How to Survive in a Post-human Era". Latham, pp. 352-363.

**Iuliia Ibragimova** obtained a Specialist degree in English from Astrakhan State University (Russia) in 2009 and worked there as an interpreter/translator and a trainer. She also has an MA degree in Literature from University College Dublin (2019). Currently she is a PhD student at DCU, researching the sentient spaceship trope in SF.