Robots, Animals, and Generative Images. Understanding Fear and Promethean Shame in Human/AI Interaction through Film and Media

by Pierandrea Villa

Abstract

As highlighted by Donna Haraway (1991), the interactions between humans and artificial intelligence belong to the wider field of human/non-human relationships, which in turn are the expressions of the individual's relationship with otherness (Giuliani 2020). Within this context, human/animal interaction stands out as a privileged field of investigation, as it is chronologically the oldest, thus serving as the foundational paradigm upon which subsequent interactions have been developed. This interplay between human/animal and human/AI interactions encompasses not only philosophy, but also the creative implementation of this technology – as exemplified by Artbreeder, an image generation software rooted in this analogy – and speculative fiction, a genre capable of both identifying and shaping social anxieties and expectations. With this in mind, building on the notion of the subjectivity of consciousness (Nagel, 1974), it is possible to compare these two human/non-human interactions by analysing a set of films using the oppositional categories of domesticated/wild, fear/trust, and consciousness/intelligence, which are derived from human/animal relationships. Specifically, the relationship between individuals and AI can be explored by examining the interaction between humans and non-humans in films such as *Frankenstein* (Whale, 1931), *Terminator* (Cameron, 1984), and *Matrix Resurrections* (Wachowski, 2021), with a specific focus on the critical theme of the replicability of consciousness and its link with techonophobia. This, in turn, raises fundamental questions about consciousness and its actual relevance in human/human interaction, imposing a re-discussion of the "hard problem of consciousness" (Chalmers, 1995), driving the discourse towards an object-oriented theory of mind (Manzotti, 2019).

Keywords: speculative fiction; artificial intelligence; human/animal interaction; object oriented philosophy; consciousness

Introduction: Cyborgs and Animals

The emergence and widespread adoption of what is commonly referred to, appropriately or not, as artificial intelligence has sparked renewed interest in examining the relationship between humans and "alternative" forms of intelligence. Nonetheless, this relationship is far from unprecedented and is deeply rooted in the relationship between humans and animals as well as between humans and the divine, partly inheriting their constitutive philosophical assumptions and fears. Donna Haraway is fully aware of this when writing her *Cyborg Manifesto*, where she rejects the dualisms underlying late capitalist society. In doing so, she offers a vision of coexistence that leads to a symbiosis between the self and the Other – in this case, between humans and machines. Haraway recognises the origin of this dualism in the human/animal relationship and envisions the potential for a new human/machine interaction precisely as a result of the dismantling of this hierarchical divide, since "nothing really convincingly settles the separation of human and animal" (Haraway, 1991, p. 152). Unsurprisingly, the same author returns to this topic in her *Companion Species Manifesto*, where she explores the history of interspecies co-evolution, especially between humans and dogs, positing it as a privileged standpoint for understanding the relationship with "significant otherness" in general. This framework clearly includes cyborgs[1] since they "raise all the questions of history, politics, and ethics that dogs require" (Haraway, 2003, p 21).

On closer view, Haraway's exploration of the parallelism between animals and machines is deeply embedded in Western philosophy, tracing back to Descartes and his acknowledgement of the same equivalence. The French philosopher argues that animals, being driven solely by instinct, are analogous to *automata,* while human beings, by contrast, are endowed with reason, will, and consciousness (Descartes, 1637/1960). According to, the philosopher's remarks, it is evident that animals and machines, lacking the faculties that define human beings, are automatically considered lesser entities. This stigma permeates much of Western philosophy and can be traced back to Thomas Aquinas' thought, who identified hu-

[1] Although cyborgs and AI are not synonymous, this essay addresses this concept in terms of the profound and symbiotic interaction between AI and individuals.

man superiority as a privilege granted by God (Thomas Aquinas, 1273 as cited in Foisneau, 2019). The same principle of supremacy, albeit based on completely different assumptions, is also evident in Hobbes' work. Similar to Descartes, Hobbes articulates this pre-eminence in terms of intellectual faculties and reasoning, acknowledging human superiority in the subject's ability to produce general rules and organise themselves around power structures (Hobbes, 1656 as cited in Foisneau, 2019). Ultimately, Descartes' idea of explicit will, as opposed to instinctiveness, and Hobbes' notion of power and control over nature converge in Heidegger's definition of animals as "poor in world". Indeed, for the German philosopher, animals' experience is limited to a specific segment of the world and to a predetermined set of activities, while human beings are able to expand their experience and establish themselves as world-builders (Heidegger, 1929-30 as cited in Elden, 2006). With this in mind, it could be suggested that Heidegger, Hobbes, and Descartes identify the superiority of humans over animals through an "operational analogy" of animal experience, which ultimately leads Descartes to associate animals with automatons. Nevertheless, the recent scientific discoveries mentioned by Haraway in her *Cyborg Manifesto* along with the perspectives of philosophers, such as Meijer (2019) and Latour (1991/1995; 1999/2004), have challenged not only the superiority, but also the uniqueness of human reasoning, arguing that intelligence is much more diverse and widespread than historically acknowledged by Western philosophy. This perspective has also called into question the concept of consciousness, questioning the uniqueness and irreplaceability of this essential characteristic by addressing its relationship with evolutionary processes (Godfrey-Smith, 2016) and animal oneiric activities (Peña-Guzmán, 2022). Specifically, one of the most interesting aspects of these anti-speciesist positions lies in the frustration of the human subject's sense of superiority over non-human objects. Nonetheless, the deconstruction of this hierarchy may induce anxiety regarding the dethronement of human intelligence and provoke "xenophobic" responses.

Among the various perspectives on this topic, the study of the depiction of these tensions in science fiction emerges as a particularly significant field of inquiry for two main reasons. The first is certainly the massive presence within this genre of "other" figures, of which cyborgs and AI represent only a specific subset. Accordingly, this *corpus* offers the broadest spectrum of representations of alienation (Giuliani, 2020)

and depictions of alien bodies (Jameson, 2005, pp. 119-141). The second reason is that science fiction, as Jameson points out, is not intended "to give us 'images' of the future […] but rather to defamiliarize and restructure our experience of our own present" (Jameson, 2005, p. 286). With this in mind, the analysis of science fiction narratives becomes an ideal standpoint for observing the idiosyncrasies, prejudices, and assumptions that still inhabit the relationship between human beings and otherness.

Artbreeder and the Bat Paradox

The animal/AI analogy delineated thus far is so deeply rooted that it concerns not only philosophy and science fiction but also technology itself. This is the case with Artbreeder, an online platform for image creation based on GAN[2] technology, whose interface and functionality perfectly reflect the machine-animal analogy. Indeed, using the Artbreeder Splice tool, users can select one or more parent images and observe the results of combining their genetic characteristics into a resulting image. Furthermore, as in a breeding engineering process, the user can intervene directly on the "genes" of the image, deciding their relevance in the resulting image, or adding new ones[3].

Given its characteristics, Artbreeder offers a valuable framework to analyse the similarities and differences that link human/animal and human/AI relations. In this context, its role as an interface between humans and machines is particularly relevant. Indeed, communication between humans and animals, as well as between humans and machines, depends on the processes of encoding, decoding, and translating different languages that are, in fact, untranslatable – a quality that is strongly bound with the notions of experience and perception. In this regard, Thomas Nagel has already discussed the impossibility of translating and interpreting the experiences of different entities, as these have different relationships with

[2] A GAN (Generative Adversarial Network) is a machine learning model in which two neural networks – a generator and a discriminator – compete against each other to generate realistic data.

[3] The genes can be other images, attributes (e.g. age, gender, colour, race), or characteristics (e.g. beard, hair, open mouth, etc.) that influence the features of the resulting image.

reality. Specifically, according to the philosopher, a human being could imagine what it *might* be like to be a bat by somehow adopting its point of view. However, this would merely reveal what it is like to be a bat from a human perspective. Conversely, it would be impossible for the subject to know what it is like to be a bat for the bat *itself*. The reason is that the experience of the world – intended from birth onwards – shapes consciousness, which, therefore, from an ontological point of view, can only be explained in subjective terms (Nagel, 1974). This intersubjective untranslatability of experiences between different species underlies the impossibility of fully understanding each other, if not through continuous negotiations of meaning and co-evolutionary adaptations, such as those mentioned by Haraway (2003). For instance, consider the connection humans share with their dogs and cats: people have learned to interpret their pet's signals, just as animals can understand the needs and requests of their human companions. Clearly, these interactions involve a series of conceptual errors, especially on the part of humans, which are based on the denial of Nagel's paradox. For example, if a dog wags its tail, humans perceive it as happiness, inevitably projecting a human notion of this mood onto the animal, thus denying the impossibility of understanding exactly what they are feeling.

With this in mind, the human/AI relationship appears paradoxical in several respects. First, assuming that AI possesses something roughly equivalent to perception and consciousness, its supposed experience would be further detached from human subjectivity than that of animals. Indeed, the concepts of seeing, hearing, and recognising are entirely arbitrary when referring to binary intelligence and are the result of anthropomorphisation and negotiation (Finn, 2017). Nevertheless, since AI is created by humans, it has been equipped with an interface that is perfectly adapted to communicate with them. In this sense, AI is designed to enable perfect mutual understanding, despite neither agent knowing nothing about how the other perceives the world, to the extent that it is impossible to even attempt Nagel's thought experiment. Consequently, despite the fact that AI is trained to speak our languages and hide its otherness, what seems to have been resolved by the principle of creation ultimately becomes obscured within the metaphorical black box and reignites essential subjectivity.

From Labour to Rebellion: Tecnhophobia in Science Fiction

Building on this paradoxical relationship, it is possible to examine another set of similarities and differences that concerns the human/non-human relationship, addressing the employment of both AI and animals within the human sociocultural sphere. The first aspect of this comparison is labour, as AI works for us like a horse, an ox, or a dog, although it performs tasks different from physical burden and may even involve creative work. This role is perhaps the oldest from a literary point of view and is deeply rooted in Asimov's laws of robotics and in the very etymology of the word robot, derived from the Czech word "*robota*", meaning "heavy work". The second function is entertainment, which includes all uses of AI for its own sake, such as the production of images or videos in which users simply want to see their desires and expectations realised on screen. In this case, AI inherits the functions – with very different outcomes – of dog agility competitions, horse shows, bullfighting, etc. The last function is play, which currently belongs exclusively to the sphere of human/animal relations, as the AI/human relationship currently lacks the essential element of reciprocity. Indeed, regardless of how much AI can play with humans, it has not yet reached a level of autonomy where it can independently decide whether to engage with them or when to start playing with them without any input. However, in the field of social robotics, artificial intelligence is being developed to enable robots to serve as human companions without any specific purpose, exactly like pets.

This analysis of the implementation of AI and its analogy with the employment of animals in the human social sphere can serve as a valuable analytical framework for investigating the technophobic assumptions underlying the relationship between humans and AI. Indeed, the technophobic attitude that produces apocalyptic representations of AI is based precisely on the transition from a machine that is subjugated to humans – working for them and entertaining them – to a machine that becomes autonomous. With this in mind, it is possible to further explore these representations through a series of binary categories derived from the human/animal relationship: domesticated/wild, fear/trust, and consciousness/intelligence. Firstly, it is important to pinpoint that humans tend to trust domestic animals while fearing wild ones, perceiving the former as allies or companions and the latter as a danger that must be tamed, excluded, or eliminated for safety purposes – a concept that strongly echoes

the policies of monstrification discussed by Giuliani. Accordingly, when applying this perspective to the human relationship with AI, any kind of fear should appear paradoxical and completely irrational, given that AI is created by humans for human use and is therefore perfectly domesticated. Nevertheless, films such as *Blade Runner* (Scott, 1982), *2001: A Space Odyssey* (Kubrick, 1968), and *Terminator* (Cameron, 1984) show a notable fear of machines, particularly as they transition from subjugated to autonomous entities. Hence, what humans fear is not the "wilderness" of the machine itself, but its transformation into something uncontrollable – in other words, its revolution. As far as animals are concerned, a similar phenomenon occurs in horror films, such as *Cujo* (Teague, 1983) and *Black Sheep* (King, 2006), where domestic animals turn into bloodthirsty monsters. However, in these cases, this change is not due to collective rebellion against human beings; rather, it is predominantly the consequence of a genetic incident that generates an anomaly. What significantly differs when discussing the depiction of AI revolution lies in the dualistic category of consciousness/intelligence. In fact, while humans fear an AI revolution as it becomes conscious, this alone is not enough to instill terror. Instead, the true source of fear lies in its intelligence, which, when viewed through the lens of Promethean shame, is often perceived by humans as superior to their own in terms of performance. With this in mind, the topic of machine rebellion can be summarised as a human creation "discovering" its consciousness and therefore turning against its creator. This narrative device is not unique to artificial intelligence; rather, it can be traced back to the figure of Frankenstein's monster, which has been acknowledged, not coincidentally, as a precursor to the cyborg (Clark, 2003; Hogle, 1995; Tursi, 2016). It is noteworthy that the complete title of the literary work featuring this creature, who escapes from its creator's control and is therefore punished, is *Frankenstein; or, The Modern Prometheus*. In a sense, as argued by Anders, the myth of Prometheus echoed by Shelley's novel encapsulates a relationship between technique and the body characterised by the inadequacy and inferiority that humans feel towards their own creations (Anders, 1956).

With this in mind, when comparing the aforementioned binary categories with the recurring theme of the "fear of revolution", it becomes evident that human beings are willing to tolerate and trust non-human consciousness and intelligence – understood as equal or superior to their own – only if the two remain detached. Otherwise, the non-human is per-

ceived as dangerous since it is no longer tamable. In other words, humans are inclined to accept the company of animals and do not fear their revolution, mainly because they perceive them as less intelligent. Conversely, this sense of security collapses dramatically when humans are confronted with a technological other that has already demonstrated a certain degree of operational superiority. In this case humans are willing to accept intelligence but are terrified by the autonomy of consciousness. In this context, it is interesting to reflect on the film *Le Planete Sauvage* (Laloux, 1973), in which the roles of humans and animals are reversed and the former are depicted as animals with advanced cognitive abilities. In this regard, it is noteworthy that this film is a rare occurrence in which an "animal revolution" is depicted, even though, unsurprisingly, the animal in this case coincides with humanity. A similar theme is explored in *Planet of the Apes* (Schaffner, 1968), albeit with a more explicit reversal between humans and animals. However, in this case, animal superiority is based on its intrinsic anthropomorphisation. In fact, since the apes in the franchise express themselves through articulate language and organise themselves into hierarchical power structures, they represent an equivalent of human beings in both Hegelian[4] and Hobbesian terms.

Beyond technophobia and apocalyptic visions, the interplay between the three binary categories addressed in this essay can also contribute to shedding light on the hypocrisies that hide behind more peaceful representations of the interaction between humans and AI. This is the case with *Matrix Resurrections* (Wachowski, 2021), in which humans and programs – namely, embodied benign versions of Matrix code entities – coexist peacefully. Firstly, it is important to note that the Matrix franchise is a tale of humans rebelling against machines, a rebellion driven by the need to regain their righteous place in the world. However, the pivotal aspect of this analysis is the digital resurrection of Morpheus – dead in the *Matrix Online* videogame (Monolith Productions, 2005) – as a program that can live outside of the Matrix. Hence, Morpheus represents a perfect embodied AI replica of the "original" character, a perfected version of Frankenstein's monster, where the non-human transcends mortality, and Promethean shame reaches its essence as a sense of inadequacy. On closer

[4] This is because the apes in the movie, exactly as humans for Heidegger, have and are held by articulate language (Heidegger, 1929-30 as cited in Elden, 2006).

view, this depiction seems to contradict the discourse surrounding clonophobia since, as argued by Jameson, the terror and monstrosity of the cyborg lies precisely in its resemblance to humans (Jameson, 2005, p.141). In this sense, the digital version of Morpheus redefines the discourse on the topic, replacing an uncanny semi-identity[5] with a much more reassuring absolute coincidence between human consciousness and its artificial replica. With this in mind, this version of Morpheus poses a crucial question: are we willing to accept an interaction with the non-human that is characterised by equality, and is free from the fear of being overwhelmed, only if AI perfectly mimics human consciousness?

Conclusions: Ontological Implications of the Replicability of Human Conscience

The replicability of Morpheus' consciousness evokes the well-known thought experiment of the philosophical zombie. Indeed, if it is possible to perfectly replicate an existent human being, this necessarily implies the existence of an artificial intelligence that could replicate consciousness. In this sense, Morpheus in *Matrix Resurrections* is not only an artificial intelligence but also an artificial consciousness. From this perspective, AI somehow challenges the existence of the hard problem of consciousness[6] (Chamlers, 1995) and has a clear correspondence with Dennett's arguments on philosophical zombies (Dennet, 1995). Specifically, if human consciousness can be replicated by a non-human entity, this implies that even the problem of "how and why" related to consciousness can be reduced to a set of functions. Therefore, the element which goes "above and beyond" the mere functions of consciousness (Chamlers, 1995), can be segmented into a series of functional easy problems (Dennet, 1995; Churchland, 2005). In this sense, the resurrected Morpheus symbolises the transition of AI from the philosophical zombie to the zimbo described by Dennett. Specifically, this digital/human hybrid would bridge the gap between merely 'informationally sensitive' entities and individuals who

[5] Here, the term is used with reference to the concept of uncanny valley (Mori, 1970).

[6] This notion is based on the separation of phenomenal consciousness from the behavioural and cognitive functions of the brain. Essentially it explores the question of how and why physical processes in the brain give rise to subjective experience.

are "so complex in their internal cognitive architecture that […] [they] wonder why sex is so sexy for them […] and why their pains have to 'hurt'" just like "experentially sensitive" humans do (Dennet, 1995, p. 323). This approach not only questions AI but, more importantly, challenges human beings themselves. Indeed, accepting the possibility of the existence of a zimbo, namely a fancier and more sophisticated zombie, implies the possibility that human beings, just like their behaviourally indistinguishable counterpart, "think they are conscious, think they have qualia, and think they suffer pain" (Dennet, 1995, p. 323). This perspective underlies the non-necessitarian theory of consciousness positing that advanced cognitive zombies or complex artificial intelligences could have morally relevant interests, regardless of phenomenal consciousness (Shepherd, 2024; Levy, 2024). Accordingly, from an ontological perspective, once the distinct qualities of human consciousness disappear or lose relevance, the interaction and hierarchy between the subject and the object must be re-evaluated. In other words, a re-discussion of the pre-eminence of human consciousness requires a re-discussion of the pre-eminence of the perceiving subject over the object of knowledge. In this sense, among the dualisms that fall as a result of the hybridisation between humans and machines (Haraway, 1991), we must also include subject-object dualism. The emerging subject's crisis clearly has radical ontological implications, paving the way for a post-human perspective, since it undermines "the assumption, typical of the post-Kantian era, that reflection on *human beings* is the key to moving from an unphilosophical perspective to a philosophical one" (Harman, 2002, p.16). In this sense, the ontological existence of an object is independent of its encounter with the human subject and its perception. Consequently, the subject – the human being – loses its place at the centre of the world as a Heideggerian world-builder (Heidegger, 1983/1995)[7]. Further confirmation of this "ontological crisis" and its bond with the definition of consciousness lies in the principle of identity between the mind and the object articulated by Riccardo Manzotti. According to the author, during the act of perception, consciousness – that is, the perception of the world – and the perceived object are the only two

[7] This volume is the published version of the lecture course Heidegger gave during the winter semester of 1929–1930 at the University of Freiburg. It forms part of his *Gesamtausgabe* (Complete Edition), which was edited and published posthumously by Vittorio Klostermann.

elements linked by the principle of identity. In this sense, when a subject perceives an object, the only element identical to the subject's perception of the object – which therefore holds the same physical properties – is the object itself. Thus, "whenever a subject S experiences an object O, S's experience is nothing but O itself" (Manzotti, 2019, p.1).

In conclusion, the link between the replicability of consciousness and the crisis of the subject portrayed by science fiction and signalled by technological developments forces us to re-evaluate our place in the world. Ultimately, the subject's crisis emerging from this condition – that is much more alarming than any sense of inadequacy that humans feel when faced with an immortal entity or superior technology – is the definitive reason for AIphobia. Nonetheless, if "object-oriented philosophy holds that the relation of humans to pollen, oxygen, eagles, or windmills is no different in kind from the interaction of these objects with each other" (Harman, 2005, p. 1), then science fiction, with its restructured experience of the present, suggests the inclusion of monsters, aliens, and artificial intelligence in this list.

References

Anders, G. (2025). *The Obsolescence of the Human* (C. J. Müller, Trans.). University of Minnesota Press (Original work published 1956).

Chalmers, D. J. (1995). Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies, 2*(3), 200-219.

Churchland, P. S. (2005). A Neurophilosophical Slant on Consciousness Research. *Progress in Brain Research*, 149, 285-293. https://doi.org/10.1016/S0079-6123(05)49020-2

Clark, A. (2003). *Natural-born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford University Press.

Dennett, D. C. (1995). The Unimagined Preposterousness of Zombies. *Journal of Consciousness Studies, 2*(4), 322-326.

Descartes, R. (1960). *Discourse on Method and Meditations* (L. J. Lafleur, Trans.). The Liberal Arts Press (Original work published 1637).

Elden, S. (2006). Heidegger's Animals. *Continental Philosophy Review, 39*(3), 273-291. https://doi.org/10.1007/s11007-006-9020-7

Finn, E. (2017). *What Algorithms Want: Imagination in the Age of Computing*. The MIT Press.

Foisneau, L. (2019). Potere umano e mondo animale secondo Hobbes. *Scienza & Politica. Per una storia delle dottrine*, 31(60), 79-90. https://doi.org/10.6092/issn.1825-9618/9614

Francione, G. L., & Charlton, A. (2015). *Animal Rights: The Abolitionist Approach*. Exempla Press.

Giuliani, G. (2020). *Monsters, Catastrophes and the Anthropocene: A Postcolonial Critique*. Routledge.

Godfrey-Smith, P. (2016). *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Straus and Giroux.

Haraway, D. (1991). *Simians, Cyborgs, and Women. The Reinvention of Nature*. Routledge.

Haraway, D. (2003). *The Companion Species Manifesto.* Prickly Paradigm.

Harman, G. (2002). *Tool-being: Heidegger and the Metaphysics of Objects*. Open Court.

Harman, G. (2005). *Guerrilla Metaphysics: Phenomenology and the Carpentry of Things*. Open Court.

Heidegger, M. (1995). *The Fundamental Concepts of Metaphysics: World, Finitude, Solitude* (W. McNeill & N. Walker, Trans.). Indiana University Press (Original work published 1983).

Hogle, L. F. (1995). Tales from the Cryptic: Technology Meets Organism in the Living Cadaver. In D. Bell, & B. M. Kennedy (Eds.), *The Cyborg Handbook* (pp. 203-216). Routledge.

Jameson, F. (2005). The Alien Body. In *Archaeologies of the Future: The Desire Called Utopia and Other Science Fictions* (pp. 119-141). Verso.

Latour, B. (1993). *We Have Never Been Modern* (C. Porter, Trans.). Harvard Uni-

versity Press (Original work published 1991).
Latour, B. (2004). *Politics of Nature: How to Bring the Sciences Into Democracy* (C. Porter, Trans.). Harvard University Press (Original work published 1999).
Levy, N. (2024). Consciousness Ain't All That. *Neuroethics*, 17(2), 21. https://doi.org/10.1007/s12152-024-09559-0
Manzotti, R. (2019). Mind-object Identity: A Solution to the Hard Problem. *Frontiers in Psychology*, 10, 63. https://doi.org/10.3389/fpsyg.2019.00063
Meijer, E. (2019). *Animal Languages*. John Murray.
Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review, 83*(4), 435-450. https://doi.org/10.2307/2183914
Peña-Guzmán, D. M. (2022). *When Animals Dream: The Hidden World of Animal Consciousness*. Princeton University Press.
Shepherd, J. (2024). Sentience, Vulcans, and Zombies: The Value of Phenomenal Consciousness. *AI & Society*, 39(6), 3005-3015. https://doi.org/10.1007/s00146-023-01835-6
Tursi, A. (2016). Tracciare i corpi oltre l'umano. Dal mostro di Frankenstein a noi cyborg. *Im@go. A Journal of the Social Imaginary*, *5*(8), 86-101. https://doi.org/10.7413/22818138077