PROMETEICA

REVISTA DE FILOSOFÍA Y CIENCIAS

EDITOR INVITADO

Nicola Liberati
(Shanghai Jiao Tong University, China)
liberati.nicola@gmail.com

Formato de la publicación

Digital: Portable Document Format (PDF), Hyper Text Markup Language (HTML), Extensible Markup Language (XML).

Idiomas aceptados

Castellano, portugués e inglés (lenguas de la publicación).

Normas de publicación

https://periodicos.unifesp.br/index.php/prometeica/about/submissions

Contacto

info@prometeica.com

Responsable

Emiliano Aldegani (Universidad Nacional de Mar del Plata, Argentina)

Diseño de isologo

Victoria Reyes (http://www.victoriareyes.com.ar)

Imagen de portada

Francis Bacon

EDITORIAL

ARTÍCULOS

DEBATES

ENTREVISTAS

RESEÑAS

LIVING IN THE NEW ERA (时代·新生)

TECHNOLOGIES, CREATIVITY, AND SCIENCE-FICTION

VIVENDO A NOVA ERA (时代·新生)

Tecnologias, criatividade e ficção científica

VIVIR EN LA NUEVA ERA (时代·新生)

Tecnologías, creatividad y ciencia ficción

Nicola Liberati[1]
(Shanghai Jiao Tong University, China)
liberati.nicola@gmail.com

Maurizio Balistreri
(Department of Philosophy and Educational Sciences, University of Turin, Italy)
maurizio.balistreri@unito.it

We live in a digitally embedded world where digital technologies touch every aspect of our lives. China is introducing the idea of entering a new era as tightly related to the implementations of such inventive technologies. It is possible to see this phenomenon pervasively breaking into different societal levels there. It is present in official government websites[2] and the actual development of digital technologies and the economy in order to refer to the unprecedented growth of the country where digital technologies take part in the constitution of the society's structure, social identity, and social relations (Xiao and Men 2021; Bonfiglioli 2021). Thus, it seems pretty clear we need to approach such a change by dealing with the effects of digital technologies while keeping China at the center of the research.

Too often, the development of new technologies has not been followed and anticipated by the work in humanities like philosophy, history, and literature. Sometimes we think of humanities as an analysis of something that has already happened. It is time to turn our gaze and show how humanities can also look to the future. For this reason, this special issue aims to merge the academic analysis of researchers in humanities with the visionary imaginaries provided by science-fiction novelists to discuss societal change and anticipate the possible effects of these technologies (Vint 2021).

Many authors in philosophy have already tackled the effects of the introduction of digital technologies in society. Postphenomenology examined the implications of emerging technologies from a phenomenological background showing how the perceptions, values, and meanings are shaped by the use of technologies. Such studies aimed to explore intentionality (Mykhailov and Liberati 2022), augmented reality (Liberati 2016; Wellner 2022), health care (Shaw et al. 2020; de Boer, Molder, and Verbeek 2021), and intimacy (Liberati 2022; Liberati and Chen 2022). Ethics is well-related to the

---

[1] Associate Professor, webpage: https://shss.sjtu.edu.cn/En/FacultyDetail/271?f=1&t=3
[2] https://english.www.gov.cn/archive/whitepaper/202204/21/content_WS6260cdc3c6d02e5335329bb1.html

implications of new technologies on society, such as bioethics (Balistreri 2020, 2022) and, generally speaking, the empirical turn in the philosophy of technology (Verbeek 2021). However, there are no outstanding issues linking these fields to science fiction in particular. Primarily, science fiction has not been used to scrutinize technological development within philosophy and humanities, even if there is an explicit trend in making science fiction part of the design of the world we want to live in from different perspectives like design (Zaidi 2019; Michaud 2017).

This special issue aims to explore such a groundbreaking connection between science fiction and humanities with a particular interest in philosophy. There are contributions from researchers and science fiction novelists from China and Italy for several reasons. Firstly, this intertwinement has not been largely pursued in general, even if there are apparent overlapping and shared interests in understanding the societal challenges we face. Secondly, there are not many works on the interactions between Italy and China on this theme, even if many collaborations bind the two countries for historical reasons and research purposes. Thirdly, China plays a central role since it is an exceptional environment to embed the analysis since it is the cradle for using new digital technologies in society, its science fiction genre is growing fast, and it is perceived as a valuable asset to look at the future.

Thus, articles in this issue mix different approaches to science fiction. They have been presented at *LINE2021 "Living in the New Era (时代·新生)*," which was an international conference held in November 2021 at *Shanghai Jiao Tong University*[3]. Researchers in the humanities and sci-fi novelists from Italy and China gathered to present their ideas and discuss the possible implications of using and developing innovative digital technologies. We decided to collect the articles and publish a selection of them in *Prometeica* Journal in order to interweave the works of the researchers better and show how this cross-fertilization can be successful and fruitful in considering the questions related to the implementations of future technologies and highlight the close relations tying these two countries.

This special issue includes six articles, a debate, two interviews, and a book review. The articles express the intertwinement between science fiction and humanities with a particular interest in philosophy. This special issue proposes papers from Chinese and Italian researchers without creating two unrelated blocks but by alternating the authors to show the dialogue existing among them.

In the paper "Becoming Others: Playing Virtual Identity and Intimacy," Jiaying Chen analyzes the development of virtual reality, and she examines what this technology means for our identity and intimacy. By revisiting Gilles Deleuze's notion of the virtual, Chen distinguishes three effects on our life in virtual reality: the dissolution of public/private boundaries, the emerging performativity of virtual reality (for example, many players adopt gender roles different from their own to participate and perform them throughout the game), and the crisis and destabilization of our ontology. Moreover, Chen delves into how the digital network has changed our lifeworld comprehensively by creating an entirely different environment that allows a new form of intimate relationships. According to Chen, digital technologies can provide a means of improving our relationships by promoting the need for flexible identity, encouraging the construction of self-narratives, and embracing indeterminacy as a possible paradigm for 'identity' and affinity with the other. While for many users, the online self remains a natural extension of the offline self, other players can reinvent their identity and use the virtual world as a chance to live a polymorphic intimacy. They can practice the possibilities of virtual identities as genuine and not just unreal or fantastic. We tend to dismiss these scenarios, but Deleuze's conception of the virtual can allow us to understand these phenomena and clarify why virtual reality might not be opposed to our world. However, we should be aware of the Janus-face of the virtual. With our confidential data and behavior moving into cyberspace, the risk of totalitarian information network abuse could become true.

In the paper "Embodied artificial intelligence in science fiction: Philosophical presuppositions and implications," Andrea Pace Giannotta explores the fruitful relationship between science fiction and philosophy regarding the topic of artificial intelligence. The paper establishes a connection between

---

[3]*LINE2021 - Living in the New Era (时代·新生): Technologies, Creativity, and Science-Fiction* https://shss.sjtu.edu.cn/En/List/345

certain paradigms in the philosophy of mind and consciousness and the imagination of possible future scenarios in sci-fi, especially in relation to the different ways of conceiving the role of corporeality in constituting consciousness and cognition. The work focuses mainly on the body conception developed in philosophy (disembodiment, weak embodiment, and strong embodiment) and the imaginaries related to the body we can find in science fiction. According to him, these two elements are well-connected because the depiction of embodied AIs in sci-fi often foresees significant philosophical debates concerning the status of AI and the ethical questions we can find in the implementations of these technologies in society. For this reason, he thinks it is valuable and essential to stimulate an exchange of ideas between sci-fi and philosophy regarding scenarios of high ethical relevance that involve embodied and conscious AIs. Moreover, he shows how science-fiction can be helpful for philosophy since it enables one to address specific problems with a different and more accessible framework. Thanks to the fact that science fiction can go about the topic in an unbounded way, it is able to uncover specific angles which are hidden from philosophy.

In "Virtuality, simulation and fake: The technical development and philosophical criticism of virtual anchors," Feng Tao and Yunyu Dang analyze the application of virtual reality technology in communication hosting through virtual images that simulate human anchors. According to Tao and Dang, the virtual reality anchor-based technology is doomed to a trust crisis because artificial intelligence can simulate human behavior, and virtual humans (VHs) can imitate a human image super-realistically. In a world where virtual reality technology is becoming a reality gradually, human anchors are replaced by VAs, human bodies are digitized, and people become VHs, the risk is to fall into an uncontrollable and unidimensional flat world detached from the historical-social dimension and lacking the growth and diversity of the living world. Moreover, the virtual reality-based reconstruction of reality risks making falsehood more and more realistic and becoming a new cultural industry. The characteristic of this industry is commercial interest: the fundamental goal of technology developers and capitalists is profit. It means the abandonment in the art of any form of negativity in favor of scientific accuracy and conformism with general aesthetics. According to Tao and Dang, virtual anchors (VAs) are not intrinsically evil or wrong, but the capital control behind them is worthy of caution. In the end, developing an autonomous, physical, and mechanized form of anchors are other critical aspects of the new technologies. The human being could become used to being taken over gradually by intelligent and autonomous machines or not human entities. Moreover, robots could increasingly become full members of our society. An ideal type of human-machine harmony is possible, but machines could slowly replace humans and eliminate them.

In the paper "On Hope Resistance," Caterina del Sordo analyzes the concept of resistance. The author adopts the application of case studies from TV series as a methodology that enlightens philosophical concepts. The research puts into question the idea of resistance that is spent and developed by José Medina's epistemology. The paper argues that the philosophically meaningful idea of resistance can be articulated in two different layers. The author refers to them as resistance tout court and hope resistance. Resistance appears as a matter of counter-acting physical or mental forces. Instead of opening the view to the latter, resistance reveals an inner multi-faceted structure involving complex cognitive phenomena like existential tenets, utopic militancy, and anticipatory practices. The distinction between tout court and hope resistance is outlined drawing on the philosophical toolkit of the Future Studies foundations provided by Ernst Bloch's theory of hope. The paper invokes the concepts of possible futures and utopic function, commending their introduction to the explanatory power of the resistance's real and fictitious heroes. Latin American feminists, characters from the science-fiction of Nosedive and Utopia, and the crime plot of Money Heist are called into play. Along this line, the paper finally argues that phenomena of hope resistance are not embedded in those of resistance tout court. The argument unfolds around the ambivalent behavior of the audience of sci-fi and crime TV series, which appear to support resistant actions only in fictitious and not in the real world. The absence of behaviors of resistance tout court does not yield the lack of eventually hidden hope resistance nurturing processes.

In "Postphenomenological variation of instrumental realism on the problem of representation: fMRI imaging technology and visual representations of the human brain," Dmytro Mykhailov aims to provide

an alternative answer to the 'problem of representation' from the perspective of instrumental realism. The 'problem of representation' frames a significant part of the debates in the contemporary philosophy of science. Two extreme answers to this problem are from realist and constructivist accounts. Realists insist that any scientific representation relies on (and refers to) independent reality. Constructivists, on the contrary, insist that through representational devices, scientists 'construct' what is real. In his paper, Mykhailov shows that there is another possible perspective on the 'problem of representation.' This perspective has been brought into play by instrumental realism. Mykhailov takes findings from the postphenomenological variation of instrumental realism and develops an 'environmental framework' to give a philosophical answer to the problem of representation through the use of imaginaries taken into consideration by science fiction. The framework focuses on three elements of the representational environment? image-making technology, image as a representational device, and scientific hermeneutic strategies occurring within the image interpretation process in the laboratory set-up. The central idea is that scientific images do not produce meanings without their instrumental environment. To fulfill this idea with empirical consistency, Mykhailov applies the 'environmental framework' to contemporary debates on fMRI imaging technology. Within the last decade, fMRI technology has attracted the attention of scholars from different fields. Such an increasing interest was called forth by the revolutionary impact that fMRI had on almost every part of neuroscientific research. However, fMRI technology images have a peculiar nature. On the one hand, fMRI technology images are not 'mere' representations because they are not just a 'copy' of the human brain, while on the other hand, these images are often used as representational devices within medical diagnostics. Mykhailov shows that the 'environmental framework' can help better understand the problematic nature of the fMRI by explaining how fMRI visuals receive their meaning through the interplay between different elements of the instrumental environment.

In the paper "Valuing abiotic nature. Perspectives on terraforming in K.S. Robinson Mars trilogy", Pierfrancesco Biasetti considers whether the original abiotic nature has a value of some kind regardless of its capacity to contribute to ecosystems and life. In our everyday experience, we tend to consider living beings, the complex system that supports their existence and naturalness, part of the same experience because life, the environment, and nature are intimately connected. Things are different on Mars, so the red Planet allows us to analyze the value per se of abiotic nature and reflect on which kind of value it could have and how much it would weigh when compared to other kinds of value. Through the lens of K.S. Robinson's Mars Trilogy, Biasetti presents a map of the possible answers to these questions that can help discuss the moral issues of the Mars terraforming project. According to Biasetti, terraforming can be a vital thought experiment to investigate the difference between various kinds of environmental value. We can assign intrinsic value to nature and value it independently from its specific quality (beauty, history, capacity to become a home for our species), or we can consider nature as a source of significant experiences: scientific curiosity and knowledge, aesthetic beauty, reverence for otherness, and diversity. In the end, nature can be critical because it supports life in all its form or intelligent life. Biasetti thinks that the plot of Robinson's trilogy indicates the necessity of a synthesis of some sort between these different values. However, given the divergences at stake, the solution is not simple or even noticeable.

The paper "Entrelaçando design e ficção científica" written by Eduardo Harry Luersen published under debates is the only paper not presented at the conference which has been added to this special issue in order to show how much the theme analyzed by it can be easily connected to other research and territories. It has been published under debates to highlight its difference from the other papers. In this paper, Luersen discusses how speculative design is related to science fiction through the use of potential scenarios and prototypes. Mainly, he focuses on the project Daleko (2020) and its relations to terraforming, showing how science fiction can stimulate designers and researchers from other fields to think of the human race's impact on our world.

As previously stated, this special issue has two interviews with the science fiction novelists who presented at the conference as keynote speakers: Prof. Wu Yan and Francesco Verso. These interviews show how science fiction can be embedded within academia and research to give attention to societal

challenges and to better aprehend the role of new technologies in our society. The two novelist shows their perspective on how science fiction is indispensable for society and how to achieve such an intertwinement among different fields. The two novelists are also significant since they work between China and Italy, and they are both well related and involved in developing science fiction within the universities in the two countries.

In her book review, Jue Wang reviewed a science fiction novel that focuses on robots and the tight bounds they might have with users following a philosophical perspective. She clearly shows how science fiction is paramount as an element to generate critical thinking and develop original ideas.

References

Balistreri, M., Capranico, G., Galletti, M., Zullo, S. (2020). *Biotecnologie. Modificazioni genetiche*, Il Mulino, Bologna.

Balistreri, M. (2022). *Il bambino migliore? Che cosa significa essere genitori responsabili al tempo del genome editing*, Fandango, Roma.

Boer, Bas de, Hedwig te Molder, and Peter Paul Verbeek. (2021). "'Braining' Psychiatry: An Investigation into How Complexity Is Managed in the Practice of Neuropsychiatric Research." *BioSocieties*, August, 1–24. https://doi.org/10.1057/s41292-021-00242-8.

Bonfiglioli, Cristina Pontes. (2021). "Methodological Challenges for a New Philosophy of Technology Interview with Nicola Liberati." *Prometeica*. https://doi.org/10.34024/PROMETEICA.2021.23.12411.

Liberati, Nicola. (2016). "Augmented Reality and Ubiquitous Computing: The Hidden Potentialities of Augmented Reality." *AI & SOCIETY* 31 (1): 17–28. https://doi.org/10.1007/s00146-014-0543-x.

Liberati, Nicola. (2022). "Digital Intimacy in China and Japan: A Phenomenological and Postphenomenological Perspective on Love Relationships at the Time of Digital Technologies in China and Japan." *Human Studies* Forthcoming: 1–15. https://doi.org/https://doi.org/10.1007/s10746-022-09631-9.

Liberati, Nicola, and Jiaying Chen. (2022). "Preliminaries to 'Digital Intimacy' The Exploration of the Changes in Values and Meanings of Intimacy in the Digital Age." *BAU - Container of Contemporary Culture* 26: 10–12. http://www.bauprogetto.net/quattordici.html.

Michaud, Thomas. (2017). *Innovation, Between Science and Science Fiction. Innovation, Between Science and Science Fiction*. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/9781119427568.

Mykhailov, Dmytro, and Nicola Liberati. (2022). "A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives." *Foundations of Science*, March, 1–17. https://doi.org/10.1007/s10699-022-09833-5.

Shaw, Sara E., Gemma Hughes, Sue Hinder, Stephany Carolan, and Trisha Greenhalgh. (2020). "Care Organising Technologies and the Post-Phenomenology of Care: An Ethnographic Case Study." *Social Science and Medicine* 255 (June). https://doi.org/10.1016/j.socscimed.2020.112984.

Vint, Sherryl. (2021). *Science Fiction*. MIT Press. https://mitpress.mit.edu/books/science-fiction-2.

Wellner, Galit. (2022). "Digital Imagination: Ihde's and Stiegler's Concepts of Imagination." *Foundations of Science* 27 (1): 189–204. https://doi.org/10.1007/s10699-020-09737-2.

Xiao, Xi, and Honghua Men. (2021). "China's New Era: The Unprecedented Changes Unseen in a Hundred Years and an Important Period of Strategic Opportunity for China." In , 1–20. Springer, Singapore. https://doi.org/10.1007/978-981-16-1156-8_1.

Verbeek, P-P. (2021). The Empirical Turn. In S. Vallor (Ed.), *The Oxford Handbook of Philosophy of Technology* (pp. 1-21). (Oxford Handbooks). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190851187.001.0001

Zaidi, Leah. (2019). "Worldbuilding in Science Fiction, Foresight and Design." *Journal of Futures Studies* 23 (4). https://jfsdigital.org/articles-and-essays/vol-23-no-4-june-2019/worldbuilding-in-science-fiction-foresight-and-design/.

Acknowledgments

BECOMING OTHERS

PLAYING THE VIRTUAL IDENTITY AND INTIMACY[1]

TORNAR-SE OUTROS

Jogar a identidade e intimidade virtuais

CONVERTIRSE EN OTROS

El juego de la identidad virtual y la intimidad

Jiaying CHEN
(Department of Philosophy of East China Normal University - Shanghai)
jchenjiaying@gmail.com

ABSTRACT

By revisiting Gilles Deleuze's notion of the virtual, this work examines what the virtual world means for our identity and intimacy. We sketch three effects of the virtual world: 1) the complete disintegration of public-private boundary; 2) the emerging performativity of virtual identities; 3) the destabilization of our ontology. From public life to multiplayer online role-playing games, we show the changing meaning of play in various contexts: performing, acting, and becoming. Following Richard Sennett's dramaturgical perspective of the public Man, we see people proactively perform the image to construct the real-world identity. Then in the video games, players choose a virtual identity to act in the game world, transforming it into a liminal space where continued identity transitions become possible. We analyze several cases, including Otherkin and Otaku subculture, to argue that virtual identity and community can bring ontological indeterminacy and propose a program that coexists with otherness.

Keywords: the virtual. identity. intimacy. play. online games. Otherkin.

RESUMO

Ao revisitar a noção de virtual de Gilles Deleuze, este trabalho examina o que o mundo virtual significa para a nossa identidade e intimidade. Traçamos três efeitos do mundo virtual: 1) a completa desintegração da fronteira público-privada; 2) a performatividade emergente das identidades virtuais; 3) a desestabilização da nossa ontologia. Da vida pública

---

[1] This article has a preliminary version published for debate in the conference " Living in the New Era: Technology, Creativity and Science Fiction", which was held on 26-28 November 2021 at the School of Humanities of Jiao Tong University (Shanghai, China).

aos jogos de interpretação de papéis, mostramos a mudança do significado do jogo em vários contextos: atuar, representar e tornar-se. Seguindo a perspectiva dramatúrgica do Homem público de Richard Sennett, vemos pessoas a executar proativamente a imagem para construir a identidade do mundo real. Depois, nos jogos de vídeo, os jogadores escolhem uma identidade virtual para atuar no mundo do jogo, o que transforma o mundo do jogo num espaço liminar onde as transições de identidade continuadas se tornam possíveis. Analisamos vários casos, incluindo a subcultura Otherkin e Otaku, para argumentar que a identidade virtual e a comunidade podem trazer indeterminação ontológica e propor um programa que coexista com a alteridade.

Palavras-chave: o virtual. identidade. intimidade. jogo. jogos online. Otherkin.

RESUMEN

Al revisar la noción de lo virtual de Gilles Deleuze, este trabajo examina lo que el mundo virtual significa para nuestra identidad e intimidad. Esbozamos tres efectos del mundo virtual: 1) la completa desintegración de la frontera público-privada; 2) la performatividad emergente de las identidades virtuales; 3) la desestabilización de nuestra ontología. Desde la vida pública hasta los juegos de rol online, mostramos el significado cambiante del juego en varios contextos: actuar, actuar y convertirse. Siguiendo la perspectiva dramatúrgica de Richard Sennett sobre el Hombre público, vemos que las personas interpretan proactivamente la imagen para construir la identidad del mundo real. Luego, en los videojuegos, los jugadores eligen una identidad virtual para actuar en el mundo del juego, lo que transforma el mundo del juego en un espacio liminal donde se hacen posibles las transiciones de identidad continuas. Analizamos varios casos, como el de Otherkin y la subcultura Otaku, para argumentar que la identidad y la comunidad virtuales pueden aportar indeterminación ontológica y proponer un programa que coexista con la alteridad.

Palabras clave: el virtual. identidad. intimidad. juego. juegos online. Otherkin.

Introduction

The digital network has comprehensively changed our lifeworld, and the intimate relationships that occur in it are receiving increasing attention from scholars. This shift has not only phenomenological origins but is also influenced by new materialism. The former reminds us that our existence and connections with others constitute who we are. With the spread of the Internet, our living experience is inevitably gained from the digital world. Thus, its occurrence, development, and transformation affect our identity and the lifeworld. New materialists suggest that objects are not neutral but that their material existence shapes our meanings and values. Inspired media theorists, such as Sarah Kember and Joanna Zylinska, suggest that we should understand new media not as a series of fragmented objects (such as computers, cell phones, or e-books), but rather as a series of mediation processes. They reveal how we exist and coexist in the technological world, our emergence, and our way of intra-acting (Kember, Zylinska, 2011)[2]. Thus, examining different mediation processes will help us understand our existence in the technological world.

Long before the advent of the Internet, sociologist Richard Sennett drew on a materialist perspective to diagnose contemporary identity and intimacy. In *The Fall of the Public Man*, he argued that the homogenization of public space in European cities during the nineteenth and twentieth centuries would lead to the stifling of everyday social interaction, the authoritarian domination of intimacy, and the loss

---

[2] "Intra-action" is a term used by Karen Barad to replace "interaction," which requires the pre-construction of "bodies" before mutual participation in action. The theory of "intra-action" believes that agency is not merely an intrinsic property of individuals or humans, but rather a dynamic force in which all "things" constantly exchange, diffract, influence, and work inseparably. Thus, it rejects the notion of absolute separation or objectivity in classical philosophy.

of otherness. Ultimately, public life would face extinction. Zygmunt Bauman shares the same pessimism, although it came from another perspective revealing that the corrosive personal relationships have rendered intimacy insipid, vapid, and unworkably fragile. In *Liquid Love*, he described modern society as liquefied with rampant "individualization" and uncanny frailty of human bonds (Bauman, 2013), implying that the convergence of modernity and technological changes had jeopardized intimacy. For him, virtual intimacy alters the human connection to more frequent and shallower, more intense, and briefer.

We are indeed facing an entirely different "material environment" and coping with a new way of social interaction. However, if we examine the above assertions through a new materialist lens, we might conclude differently. First of all, today's public space is not as homogenized as Sennett suggests. The virtual world has created alternative ways of constructing identities and intimacy. While it has further dismantled the division between the public and private spheres, the dramaturgy of public life has not disappeared but has permeated everyday life. In this process, social media and online multiplayer games redefine the "self" and "other" in relationships, that the former produces the real-world identity and the latter virtual identity. We compare the differences and clarify that virtual identities actualized by personas and avatars provide new identity paradigms and intimacy models (Berlant, 2000).[3] Liquid love might brew an approach to actualize polymorphic intimacy, which allows us to question orthodox intimacy and communication and embrace the process of "becoming". Finally, the Otherkin subcultural group will serve as a particular case to illustrate how virtual identity challenges the original ontological categories.

Fluidifying Public-private Boundary: Play

Sennett listed three ways of understanding public life in the West: 1) Jürgen Habermas's theory of "communicative interaction"; 2) Hannah Arendt's theory of "citizenship"; and 3) his model of dramaturgy. This model states that the self in public space is the result of dramatic interaction, that its social relations are sustained by the self-playing the same role to the same audience on different occasions, and that this playing is not always limited to one role.[4] All three approaches presuppose a division between the public and private spheres. Sennett, for example, sees the private sphere as a field of self-expression, intimacy, and shared feelings, but not the public sphere. In his view, industrial capitalism and secularization's destruction of the public sphere will affect two aspects of social life: the city and politics. Citizens became so obsessed with the search for intimacy that we have lost sight of the nature of the city as a theater of communication for the authentic, impersonal public. In a nutshell, the city is "the place where strangers meet." At the same time, politics has been depoliticized as we are more leaning on addressing power and resource distribution issues through the virtues of trust or the private sphere.

While Sennett's diagnosis still seems valid today, it is more complex considering evolving lifeworld. Given this, we must look at how people present themselves and interact in social media to see the impact of mediation. Indeed, the media further dissolves the public-private boundary, but this does not necessarily have negative consequences. On the one hand, the Internet not only provides individuals with broader access to others but also allows people to construct a public image of themselves; on the other hand, because the maintenance of an online self's image also involves the personal circle, it requires individuals to behave in a manner that is consistent with the private sphere. The ensuing politicization of the private sphere helps shed light on the issues of family ethics and sexual ethics that used to be considered of no account.

---

[3] Here, intimacy refers to the concept after the critical development of feminist scholars (such as Lauren Berlant). It includes all the "kinds of connections that impact on people, and on which they depend for a living." Critical intimacy studies focus on deconstruction of binary opposition between public and private, mediation tradition relationship norms, and political potentiality of intimacy.

[4] Sennett's dramaturgical model can also be found in Erving Goffman. Goffman defines it as "all the activity of a given participant on a given occasion which serves to influence in any way any of the other participants." (Goffman, 1956)

Before the Internet became widespread, individuals used to be able to selectively maintain identities in relatively fixed settings, such as a father or husband at home, a professor on campus, and an intellectual in another social context. These roles are usually relatively stable and mutually independent. However, these spatial compartments become less clear after the Internet penetration, and stable personal identities are shaken. Particularly during the epidemic, when so much activity was moved online, our spatial properties became blurred—our living rooms could instantly become classrooms or political arenas for public discourse. This "identity disorder" makes it difficult for individuals to tailor their identities in specific contexts and, therefore, to manage information sharing. People who use social media avidly find it challenging to realize that everyday sharing is, in fact, a private sphere publication.[5] We have witnessed countless events where netizens searched, collected, and exposed an ordinary person's personal information because of a single incident and suddenly became the public topic of conversation.

Some researchers pay attention to these phenomena and analyze them. For example, Hiroki Azuma acutely sums up two models of narrative consumption after observing the shift from a modern world image to a postmodern world image in Japanese Otaku culture (Azuma, 2009)[6]. His analysis echoing postmodernism will come into play for elaborating virtual community. Two models—Tree models (left) and the Database models—explain how we are leaning upon grand narrative in the modern age, which prompts us to grasp an image of the world through the tree model and obtain a worldview that is more reliant on the hidden structure and overarching philosophy. However, with the advent of postmodernism, this tree structure completely collapsed. Azuma brings up the database model (or a reading-up model) to help us understand the postmodern world. For instance, the Web has no center and no dominant hidden grand narrative. There is a double-layer structure on Internet, in which coded information can accumulate, and individual Web pages grow with the user's superficial reading. The distinction between the postmodern double-layer structure and the modern tree structure is whether a narrative hides under the surface we read. The database model means the outer surface layer in the postmodern world is not determined by the deep inner surface. That is to say, Otaku's self-images do not require a grand narrative for support; their online presentation becomes a direct performance on the public stage, where people tend to grasp one's rough image through quick "reading up".

However, the fluidity of spatial attributes, identities, and narrative consumption can also have positive results in that individuals realize the need for flexible identity and begin to construct the image of cyberself actively. Social networking sites have reconfigured relationships' "flexibility, informality, and conviviality" by openly displaying personal relationships (Chambers, 2013). Alison Hearn points out that most social media users consciously and purposefully create specific images of themselves to influence their public perceptions. She describes such behavior as constructing "meta-narratives and meta-images" of the self (Hearn, 2008). In this process, the production of self-images requires an active self-alienation, which blurs the distinction between the personal self and the gazed object while erasing the boundary between the concept of self and the consumption objects of capitalist production. The individual can gradually become aware of how the social self is constructed (Tufekci, 2008). We are no longer mere objects under the gaze of the other, but we can introspect under the gaze and even adopt the mechanism to construct self-narratives. This is profoundly illustrated by artist Amalia Ulman's performance *Excellences & Perfections* (2014). By analyzing the history of Instagram celebrities, she curated three fictional life stages and earned herself nearly 100,000 followers through a four-month social media performance. The way Ulman rehearsed herself through social media and imitated celebrities manifests that the mystery of online self-image is hidden in the cracks of the disintegration of the public and private spheres. Because the private information transmitted to the social platform can become public at any time, and one's meta-narratives and meta-images need to be complemented by the private sphere, the self is also the public self.

---

[5] In the Foucaldian sense, it also helps build Panopticon—everyone has placed oneself in a prison that can be watched, and everyone has access to watch others.

[6] A subculture that emerged in Japan in the 1970s. It is growing with the vast entertainment industry and the expansion of the Internet, where anime, video games, the virtual community are booming.

On the other hand, the shaping of the online self is not one-sided; it also requires the provision of an "imagined audience" (Boyd, 2011). Just as a product needs to establish a consumer base before being launched, social media users can adjust their self-images by imagining their audiences' preferences. This makes WeChat's "Moments"[7] a micro-field of self-image playing, as it mixes different audiences, such as family, friends, family, colleagues. It is a common dilemma of posting in WeChat when we want a post to be seen by one group of people and not by another. The WeChat team obviously recognizes the inconvenience of mixed audiences for social media users and therefore offers a blocking feature that allows people to "target" the content to particular audiences. This reveals that these imagined audiences are highly manipulated, selected, and pre-determined to share specific characteristics and thus passively subjected to an advertised image about a particular self. Just as Jean Pau Sartre described, the self is simultaneously being-for-itself and being-for-others; or like the sociologist, George Herbert Mead asserted: I become the Me in the eyes of others, necessarily through the process of I internalizing others (Mead, 1934); in this process, the selves as actors and the others as audiences are always in a dynamic of interchange. Moreover, this transformation is no longer limited to the original public sphere but works simultaneously in the private sphere, thus bringing the ethics of the private sphere into focus. Given that the self is expressed and negotiated in a highly visible way through social media, the publicization of the self inevitably involves a renegotiation of the public and private spheres.

As mentioned above, the cyberself has multiple identities and the autonomy to construct identities by manipulating virtual, generalized others. In some cases, others may also break away from this manipulation and take the initiative to examine the authenticity of one's image. For example, we can see many screenshots of conversations on social media in internet events. These transcripts are precisely the result of converting the passive image-receiving audience into an active image portraying participant. In this way, the dramaturgical model of public life continues in cyberspace. Only the *theatrum mundi*, which was relatively static, has evolved into multiple parallel theaters, where identity plays are no longer just the interpretation of a priori role also the performing of a virtual identity. Thus, the dramaturgical model of public life does not merely mean that individuals, like actors, play their social roles. Its development by the theory of speech-acts emphasizes the difference between performance and performativity, which means that individuals can play a new identity through perlocutionary acts[8] (Austin, 1962). This situation further develops in massively multiplayer online role-playing games.

Multiplayer Online Role-playing Games: Virtual

If social media provides more abundant conditions for self-performing, then massively multiplayer online games allow the essence of play to be fully realized. While for many users, the online self remains an extension of the offline self (Robinson, 2007). However, more scholars are noticing the differences that virtual worlds bring. For example, in video games, players are free to build a virtual lifeworld by taking on an avatar. Unlike social media users keen to construct their meta-narratives, game players are more interested in adopting or creating completely virtual identities, creating new kinds of intimacy. In *Virtual Intimacy*, Shaka McGlotten describes the intimate landscape in the *World of Warcraft* as a multiverse (McGlotten, 2013). He points out that while the two dominant types of relationships in WoW—team play and solo play—are primarily instrumental as a means to achieve particular game ends. Nevertheless, because the *Warcraft* universe consists of multiple worlds, and thousands of gamers bring multiplicities, this makes WoW an active field for new kinds of intimate relationships and allows these relationships to be transformed into "a means without an end". Then the virtual world might be a chance to actualize the polymorphic intimacy.

---

[7] "Moments" is a function of the app WeChat, which is also called "Friends' circle" in the Chinese version. It means users can share and get access to WeChat friends' information, creating an intimate and private communicating circle within the users' choice of close friends—quoted from Wikipedia.
[8] Perlocutionary act refers to saying something that produces specific consequential effects upon the feelings, thoughts, or actions of the audience.

In his view, the dominant instrumental relationships in *Warcraft* are caused by the constrained factors of the game, including limited identity, space, time, and belonging. First, *Warcraft* players need to choose their avatar before entering the game. These avatars are jointly imagined and constructed through numerous media such as computers, card games, comics, novels, and movies. As history shapes the identities in the offline world, *Warcraft* players inherit its narrative, class, gender, race, and connections by choosing an identity. In addition, players sometimes encounter situations where they cannot move forward, where an invisible wall stops their avatars. This reveals the constraining space of WoW and the existence of the virtual world as a closed system. Time is more restrictive than space. Because players, like everyone else, have only 24 hours in a day. Time resists virtualization in a way that space does not, which becomes a significant limitation of the game. The constraining nature of identity, space, and time ultimately limits the types of relationships in WoW. The result is that the dominant intimacy in *Warcraft* still reproduces the scripts of the offline world, and normative desires and ideals still bind players' connections.

However, there are exceptions. For example, many players adopt gender roles different from their own to participate and perform them throughout the game (Martey, 2014). Also, McGlotten mentions his involvement in the "The Amazons of Kalimdor" guild, which does not restrict the game player's identity, but only allows players to participate with female avatars. This has given rise to a gaming camaraderie based on gender roles. In addition, there are more cases, such as online romance stories developed in WoW due to gaming encounters, exchanging emotions through in-game texts, or traveling together in the game world (Freeman, Bardzell, Bardzell, 2016)[9]. All these phenomena result from players practicing the possibility of virtual identities. However, it is common to dismiss these virtual identities or intimacy and assume that these situations cannot lead to any normative commitment. While they may give rise to new forms of relationships, they do not matter, "they are fantastic or simulated, imaginative, incorporeal, unreal. (McGlotten: 7)" However, McGlotten argues that Gilles Deleuze's explanation of the virtual will open up a new perspective for understanding these phenomena and clarify that the virtual is not opposed to the real and that virtual identities and emotional relationships are part of the process of actualization.

In Deleuze's ontology of becoming, "the virtual" is a concept that needs to be understood with "the actual". The two are mutually exclusive, yet together fully characterize the real. In his view, reality is not a linear process of development from one actual to another, but a movement from the actualized state of affairs, through a dynamic field of virtual/real tendencies, to the actualization of this field in a new state of affairs. We can illustrate it with the following schema: virtual/real <—> actual/real <—> virtual/real, which goes round and round to infinity[10]. Here the virtual is vitality, events, the past, "the impassive and dynamic aspects of multiplicities in the process of actualization." It has the capacity to trigger actualization, but it always remains distinct from its actualization (Boundas, 2005). Thus "the virtual is not something that lacks reality but something that is engaged in the process of actualization following the plane that gives it its particular reality" (Deleuze, 2001). The virtual is conceptual but not abstract, real but not actualized. Therefore, the virtual has been latent in reality long before the Internet. For example, in *The Virtual*, sociologist Rob Shields follows this lead and argues that the virtual is the reality that is not yet actual (Shields, 2003). He examines the "virtual" space in traditional rituals such as the Christian Eucharist or the Rites of Passage of different cultures (Gennep, 1960).[11] These rituals virtually transform the place where they take place into a liminal space, a unique space-time that people experience in life transitions. As an intermediary zone between two realms, it flows and meets the qualities of different spaces, in which individuals have the complex experience of identity transition.

Therefore, the online video game's virtual world can also be regarded as a liminal space in which virtual identity operates in a concrete and real way and initiate the actualization of new intimate relationships.

---

[9] Different multiplayer online games also develop diverse intimacy landscapes. As some studies of *Audition Online* have pointed out, the game is dominated by "couples" playing together and produces a whole relationship development mechanism.

[10] This schema not only shows the reversibility of the process but also the structure of temporality.

[11] "Rites of Passage" refer to the rituals performed when a person leaves one group and enters another, involving significant changes in social status. The original French term was invented by the ethnographer Arnold van Gennep in *Les Rites de Passage* (1960).

It encourages us to explore the expansive forms of virtual intimacy in various online communications without the presupposed forms and ideas (*eidos*). We can experiment with different identities in this process, crossing the categorical boundaries of social and embodied existence. The dichotomy between virtual and real, online and offline, is no longer viable. If we accept Deleuze's explanation of the virtual, then the Internet could be seen as a generative space that allows individuals to change their state of being and achieve new bodily norms, embodied experiences, and self-identification.

Virtual community and Otherkin becoming

We have shown that virtual communities make traditional social identity and relationships less stable. Players of game worlds are more open to creating alternative roles with different genders and abilities and forming intimate connections with other users through these roles. Eva Zekany provides an example of how people become new identities and are involved in new intimate relationships in the physical world through virtual communities.[12] Otherkin is the group of people who consider themselves partially or entirely non-human in soul, spirit, or psyche, including those who believe that their physical bodies are not what they really are. There is a similar concept of Otherkin in anime culture—じんがい. It originally referred to non-human beings such as animals and monsters. After the development of literature and anime culture, it also refers to subhumans or robots. Many Otherkin consider themselves as characters in video games, anime, novels, or movies, practice a specific style of life in their daily lives (e.g., cosplay culture) and socialize mainly in the virtual communities. Through the lens of media ontology, Zekany attempts to analyze how Otherkin constructs their identities and communities through mediation, drawing on Stiegler's theory of technical evolution to elucidate the symbiotic evolution of humans and technics and point out that the core of human ontology is, in fact, non-human (Zekany, 2018).

Inspired by Deleuze's concept of "the virtual", we revisited the virtual and real relationship and relocated virtual identity and intimacy as part of the real. Based on this, the Otherkin as a virtual identity can also be considered a real identity of becoming. It represents not only a potential identity but also different mediation paradigms and models of intimacy. Zekany associates the virtual with the theory of affect. Because Brian Massumi has pointed out that affect is virtual:

> Affects are virtual synesthetic perspectives anchored in (functionally limited by) the actually existing, particular things that embody them. The autonomy of affect is its participation in the virtual. Its autonomy is its openness (Massumi, 2002).

As a pre-subjective, pre-personal, non-conscious experiential force[13], the affect is always simultaneously involved in "the virtual in the actual" and "the actual in the virtual", as Deleuze describes the movement of becoming, where one emerges from the other and returns to the other. In other words, affect is always implicit in virtual connections. Affect occurs not only in intimacy but also through intimacy. Based on this, Zekany analyzes how Otherkin perform their identities at different affective levels: first, through the flow and intensity of the medium; second, through affinity-based virtual communities (Haraway, 1991)[14]; and third, through intimate encounters between users and machines, that enable a kind of embodiment, in which one become posthuman through the ongoing mutual construction of humanity and technics.

This work cannot be put forward without Bernard Stigler's arguments for technics evolution. In *Technics and Time*, he proposes that human is essentially historical. Technics is an idiosyncratic reality in the evolution of humans and, unlike what we usually think, technics is not invented by man, but human

---

[12] Virtual community refers to the communities where people using computers to communicate, form friendships, intimacy, and the basis of societies.

[13] Brian Massumi distinguishes between emotion and affect: emotion is a personal and subjective experience, while affect is a pre-subjective, pre-personal "intensity". Intensity implies the impact or persistence of the experience, expressed by the most direct and autonomous embodied response, manifested on the surface of the body and in contact with things.

[14] Affinity, borrowed from Donna Haraway, refers to a relationship determined by choice rather than blood.

invents technics while inventing themselves within it. Thus, "the history of technics…is also the history of humanity" (Stiegler, 1998). He proposes the concept of *epiphylogenesis* to reject the fixed humanity and emphasizes that all the characteristics of a species are not determined a priori in the embryo but gradually emerge in later growth. The evolution of life continues along with means other than life. Here, technics is understood as a constitutive force that emerges with the category "human" at the very beginning, which both share a co-originary nature. In this light, Zekany suggests that when examining the Otherkin subculture, we should not view the medium through which Otherkin are sustained as a mere platform or accept the critics of Otherkin opponents, which suggest that Otherkin performs their identity only because of their desire to be "others". On the contrary, we should see Otherkin as part of *epiphylogenesis*, whose embodiment and intimacy will provide a different ontological project within and with the medium.

In addition, because Otherkin interpret their term as "kin to the Other", Jay Johnston notes the special kinship between the Otherkin and otherness. He notes that "otherness" was once conceptualized as alterity (radical difference), a threatening presence that the subject must confront and contend with. However, many scholars have recently realized that the otherness can also be productive and destabilize the subject's solidified ontological ground (Johnston, 2013). As Sennett says, we need to encounter the others. Otherkin precisely present one solution when we confront the animal as the other. It dissolves the two ontological categories of human and animal and offers us the possibility of a transpecies identity. As mentioned above, the virtual world of the Internet provides space for this transition of identity, in which the virtual identity of Otherkin is played out in a concrete and real way. Furthermore, Stigler's dissolution of the ontological categories of "human" and "technics" is also fully evident in the Otherkin subculture. This echoes the gradual dissolution of the distinctions between real/virtual, organic/inorganic, and human/nature. Here, the Otherkin is not an abnormal "exception" but a becoming that emphasizes the dynamic relationship between the virtual world and human beings in a continuous mutual generation.

Beyond the case of Otherkin, we also find affinity-based groups in the Fandom communities, which constitute people who share the same interests in particular things, such as movies, celebrities, comics, fashions, and even opportunities to buy and sell related merchandise. Fandom culture has drawn little attention for long until it proves its action potential during the pandemic. To be specific, many fandoms show their concerns and unite the group promptly as effective civilian rescue teams when China was facing a difficult situation because of the coronavirus. They organize various resources and assist in delivering medical supplies to hospitals through the network based on the virtual communities. It reveals that even people from entirely different classes or regions can develop a bond based on affinity. What is more, with more flexibility and informality, they can act without the limitation of bureaucracy and respond to the State of Exception. In sum, the virtual community has shown its political permeability and the capability to challenge the monopoly of the existing class and revive the citizen-based democratic project.

Conclusion

So far, we sketch a thread of how the virtual world changes individual identities and intimate relationships. First of all, the emergence of social media has completely dissolved the traditional division between the public and private spheres. Then it reinforces the nature of play in online networking. Here, playing means not only performing but also generating and experimenting with new identities. Multiplayer role-playing games turn the virtual world into a liminal space for identity transformation, allowing virtual identities and intimacy to be actualized. Affirming these will allow us to embrace the ontological indeterminacy and a possible paradigm for affinity with the other. As Johnston reminds us, the Otherkin subculture providing a means of affirming species differences and reconfiguring ontology is rather political than pathological. Perhaps the Otherkin can be regarded as a political myth like

Cyborg[15] (Haraway: 149), offering an opportunity to perform becoming identities and negotiate new ethical and political regulations. Rather than simply dismissing the subcultures produced by the virtual world, these cultures should be considered carefully to recalibrate our subjectivity and learn how to live with alterity. Their practice offers us a more complex, radical, and creative ontological experiment through concrete and authentic engagement with the real. In doing so, we can reimagine post-humanist images of the subject and the associated kinship.

Also, the transformation of intimacies in Asia, whether romantic, familial, or communal, is marked as a unique situation compared to the Western, considering the region has seen the proliferation of its "homegrown" digital culture such as WeChat, bilibili, TikTok. Some point out a phenomenon of "global intimacies" emerging from Asia (Cabañes, Uy-Tioco, 2020), which refers to various imaginaries and practices among the virtual intimacies that negotiate global modernity and local life. Although the increasing globalization, capitalism, and secularization contribute to cultural homogenization, the simultaneity, and the inter-penetration of two geopolitical levels, people are reconfiguring the local networks and enacting global relationships. What Asian virtual communities display is unique: they normalize and subvert the conventional intimacies, which invents the past and performs multiplicities while struggling with the existing framework. It reminds us that we should work more prudently when the issue of intimacy comes from a different cultural background.

Moreover, it should not be overlooked that virtual identities, denying the boundaries of identity on which human subjects depend, also leave themselves in a contradiction. In the case of Otherkin, on the one hand, their difference rests on the ontological difference they seek to erase, and if the animal-human distinction does not exist, then neither does Otherkin exist; on the other hand, if Otherkin remains as others, the legitimacy of its identity will continue to be questioned. More importantly, like the concerns raised to affect theory, the generative power of affirmation will also encounter problems: is it permanent? Or does the ontology of becoming instead lead us to nihilism? Are virtual identities and intimacies beyond reproach? Do we even have a basis for judging these identities and relationships? How do we build up the ethics above a wasteland that post-humanists left? As Howard Rheingold reminded us to be aware of the Janus-face of the virtual. With much of our intimate data and behavior moving into cyberspace, the critics of potential totalitarian information network abuse are also worth considering (Rheingold, 2000).

References

Austin, J.L. (1962). *How to Do Things with Words*. Oxford at the Clarendon Press.

Azuma, H. (2009). *Otaku: Japan's database animals*. U of Minnesota Press.

Bauman, Z. (2013). *Liquid love: On the frailty of human bonds*. John Wiley & Sons.

Berlan, L. (2000). *Intimacy*. Chicago: University of Chicago Press.

Boundas, C. (2005) "Virtual/Virtuality". *The Deleuze Dictionary*. Parr, A. (Ed.). Edinburgh: Edinburgh University Press. 296-298.

Boyd, D. (2011). "Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications". *A Networked Self: Identity, Community, and Culture on Social Network Sites*. Papacharissi, Z. (Ed.). New York: Routledge.

Cabanes, J. V. A., Uy-tioco, C. S. (Eds.). (2020). *Mobile Media and Social Intimacies in Asia: Reconfiguring Local Ties and Enacting Global Relationships*. Springer Nature.

---

[15] Donna Haraway described a cyborg is a "cybernetic organism, a hybrid of machine and organism, a creature of social reality as well as a creature of fiction".

Chambers, D. (2013). *Social Media and Personal Relationships: Online Intimacies and Networked Friendship.* London: Palgrave Macmillan. 61.

Deleuze, G. (2001). "Immanence: A Life". *Pure Immanence: Essays on a Life*. Boyman, A. (Trans.). New York: Zone Books. 31.

Freeman, G.; BARDZELL, J.; BARDZELL, S. (2016). "Revisiting computer-mediated intimacy: In-game marriage and dyadic gameplay in Audition". *In: Proceedings of CHI – Conference on Human Factors in Computing Systems.* 5. https://doi.org/10.1145/2858036.2858484

Gennep, A. V. (1960). *The Rites of Passage.* Yizedom, M. B., Gabrielle, L. C. (Trans.). Chicago: The University of Chicago Press.

Goffman, E. (1956). *The Presentation of Self in Everyday Life*. Edinburgh: University of Edinburgh Social Sciences Research Centre.

Haraway. D. (1991). *Samians, Cyborgs, and Women: The Reinvention of Nature*. New York: Routledge.

Johnston, J. (2013). "On having a furry soul: Transpecies identity and ontological indeterminacy in otherkin subcultures". *Animal Death*. Johnston, J.; Probyn-Rapsey, F. (Ed.). Sydney: Sydney University Press. 293–306.

Kember, S.; ZYLINSKA, J. (2011). *Life After New Media Mediation as a Vital Process*. Cambridge: The MIT Press.

Martey, R.; STROMER-GALLEY, J.; BANKS, J.; WU., J.; and CONSALVO, M. (2014) "The strategic female: Gender-switching and player behavior in online games". *Information, Communication & Society*, 1. https://doi.org/10.1080/1369118X.2013.874493

Massumi, B. (2002). *Parables for the Virtual: Movement, Affect, Sensation*. Durham: Duke University Press Books. 35.

Mcglotten, S. (2013). *Virtual Intimacies: Media, Affect, and Queer Sociality.* New York: State University of New York Press.

Mead, G. H. (1934). *Mind, Self, and Society: From the Standpoint of a Social Behaviourist*. Chicago: University of Chicago Press.

Rheingold, H. (2000). *The Virtual Community: Homesteading on the Electronic Frontier*. MIT press.

Robinson, L. (2007). "The Cyberself: The Self-ing Project Goes Online, Symbolic Interaction in the Digital Age". *New Media and Society*. 9. 93. https://doi.org/10.1177/1461444807072216

Shield, R. (2003). *The Virtual*. New York: Routledge. 25.

Stiegler, B. (1998). *Technics and Time, 1: The Fault of Epimetheus*. Beardsworth, R.; Collins, G. (Trans.)*.* Stanford: Stanford University Press. 134-143.

Zekany, E. (2018). "Technical intimacies and Otherkin becomings", *Mediated Intimacies: Connectivities, Relationalities and Proximities*. Andreassen, R.; Petersen, M. N.; Harrison, Katherine.; Raun, T. (Ed.). London: Routledge. 240-253.

EMBODIED ARTIFICIAL INTELLIGENCE IN SCIENCE FICTION

PHILOSOPHICAL PRESUPPOSITIONS AND IMPLICATIONS

INTELIGÊNCIA ARTIFICIAL ENCARNADA NA FICÇÃO CIENTÍFICA

Pressupostos filosóficos e implicações

LA INTELIGENCIA ARTIFICIAL ENCARNADA EN LA CIENCIA FICCIÓN

Presupuestos e implicaciones filosóficas

Andrea Pace Giannotta
(University of Bergamo)
andreapacegiannotta@gmail.com

ABSTRACT

In this paper, I explore the fruitful relationship between science fiction and philosophy regarding the topic of artificial intelligence. I establish a connection between certain paradigms in the philosophy of mind and consciousness and the imagination of possible future scenarios in sci-fi, especially focusing on the different ways of conceiving the role of corporeality in constituting consciousness and cognition. Then, I establish a parallelism between these different conceptions of corporeality in the philosophy of mind and certain representations of AI in sci-fi: from computers to robots and androids. I conclude by stressing the value of exchanging ideas between sci-fi and philosophy to foreshadow and evaluate some scenarios of high ethical relevance.

Keywords: embodiment. consciousness, sci-fi. phenomenology. flesh.

RESUMO

Neste artigo, exploro a relação frutífera entre ficção científica e filosofia no que diz respeito ao tema da inteligência artificial. Estabeleço uma ligação entre certos paradigmas da filosofia da mente e da consciência e a imaginação de possíveis cenários futuros na ficção científica, focando especialmente as diferentes formas de conceber o papel da corporeidade na constituição da consciência e da cognição. Estabeleço assim um paralelismo entre estas diferentes concepções de corporeidade na filosofia da mente e certas representações de IA na ficção científica: desde computadores a robôs e andróides. Concluo salientando o valor de uma troca de ideias entre ficção científica e filosofia para prenunciar e avaliar alguns cenários de elevada relevância ética.

Palavras-chave: corpo encarnado. consciência. ficção científica. fenomenologia. carne.

RESUMEN

En este artículo exploro la fructífera relación entre la ciencia ficción y la filosofía en relación con el tema de la inteligencia artificial. Establezco una conexión entre ciertos paradigmas de la filosofía de la mente y la conciencia y la imaginación de futuros escenarios posibles en la ciencia ficción, centrándome especialmente en las diferentes formas de concebir el papel de la corporeidad en la constitución de la conciencia y la cognición. Así, establezco un paralelismo entre estas diferentes concepciones de la corporeidad en la filosofía de la mente y ciertas representaciones de la IA en la ciencia ficción: desde los ordenadores hasta los robots y androides. Concluyo destacando el valor de un intercambio de ideas entre la ciencia ficción y la filosofía para prefigurar y evaluar algunos escenarios de gran relevancia ética.

Palabras clave: sujeto encarnado. conciencia. ciencia ficción. fenomenología. carne.

Introduction

Science fiction (sci-fi) is the literary and cinematographic genre that constructs narratives based on possible future scenarios that involve revolutionary scientific discoveries and technologies. These envisioned scenarios very often have a deep philosophical import. This is most clear in the case of certain philosophical ideas about mind, consciousness, and the possibility (or impossibility) of artificial intelligence (AI). In these domains of philosophical inquiry, the connection with sci-fi is tight and goes in both directions: philosophical ideas can inspire sci-fi narratives, and the latter not only can give concreteness to abstract philosophical ideas but can also open new possibilities for philosophical reflection.

Indeed, the concept of AI is essentially a sci-fi concept, concerning a possible techno-scientific innovation leading in the future to the creation of an artificial minded creature. At the same time, this concept is deeply philosophical, as it deals with nothing less than the most fundamental philosophical questions about mind, consciousness, and their place in the order of reality. The fact that we are beginning to turn the sci-fi concept of AI into something real makes it even more meaningful to reflect on the philosophical presuppositions and implications of certain sci-fi scenarios. In particular, the guiding thread of this paper is the role of corporeality in consciousness and cognition and the way it is conceived in philosophy of mind, on the one hand, and the way it comes into play in sci-fi, on the other.

The paper is structured in four sections. In the first three sections, I shall discuss three paradigms in philosophy of mind: disembodiment (sec. 1), weak embodiment (sec. 2), and strong embodiment (sec. 3). In all three cases, I shall establish a connection with certain representations of AIs in sci-fi: from computers to robots and androids. In the fourth section, I shall consider some ethical issues that arise within the contemporary debate on artificial intelligence and artificial consciousness. I conclude by stressing the value of an exchange of ideas between sci-fi and philosophy to foreshadow and evaluate some scenarios of high ethical relevance.

1. Disembodiment

The concept of artificial intelligence implies, by definition, the creation of an artefact by human beings. In any case – i.e., regardless of the specific way in which this is done – such an artefact must have some *material realisation*, i.e., it must be created by arranging a material substratum in certain ways. This means that AI is always *embodied* if we use this concept in a very general sense (i.e., as equivalent to "materially realized"). However, the crucial issue is: what kind of embodiment is necessary to have a mind and, therefore, also an artificial mind?

A first answer to this question is found in the classic computationalist-representationalist (or logicist-symbolic) paradigm in philosophy of mind, which conceives of the mind in analogy to the software in a computer. In this paradigm, mental states are identified with certain functions that lead from input from the environment to a certain behavioural output. The logicist-symbolic approach to AI has been also

notoriously called "good old fashioned artificial intelligence" (GOFAI, Haugeland, 1985) and seeks to build agents that show "intelligent" behaviour. In this sense, "intelligent" is an adjective attributed by us, the observers, to systems that perform in certain ways. It is therefore an "apparent" feature of cognitive systems that appears "from the outside" or "in the third person". For instance, an AI that passes the imitation game (Turing, 1950) has a certain behaviour that is indistinguishable from that of a human being.

However, in this way of conceiving of intelligence, there is no reference to the first-person experience of a cognitive agent, i.e., there is no reference to "what it is like" (Nagel, 1974) to be a subject of experience. This is the dimension of "phenomenal consciousness", which constitutes the so-called *hard problem* in philosophy of mind (Chalmers, 1996). I stress the fact that the computationalist-representational paradigm in philosophy of mind and logicist AI leaves aside the phenomenal aspect of mental life, which is essentially tied (as we will see) to bodily life. In this approach, phenomenal consciousness doesn't play any role in accounting for the behaviour of an agent.

The possibility of realizing this form of intelligence in artefacts is what is commonly called, following Searle (1980), "weak AI", as opposed to "strong AI". Whereas strong AI "seeks to create artificial persons: machines that have all the mental powers we have, including phenomenal consciousness", weak AI "seeks to build information-processing machines that *appear* to have the full mental repertoire of human persons" (Bringsjord & Govindarajulu, 2020: 66).[1] Therefore, the realization of weak AI concerns intelligence defined in merely functional terms as a certain behaviour in relation to certain conditions. In particular, this view maintains a certain dualism between mind and body, which is evident in the analogy between the mind and the software of a computer and between the brain and its hardware. Indeed, even if the software must be implemented by a certain material substratum, the features of the hardware that are deemed to be relevant for realizing a mind are situated at a general level of abstraction (e.g., the architecture of the Turing machine and the Von Neumann architecture). They do not concern the specific morphologic, dynamic, and sensorial features of the body of a concrete cognitive agent. Therefore, the concrete body of an agent is not deemed to be relevant in accounting for its mental processes. For this reason, this view has been criticized for implying a conception of the mind as *disembodied*.

The computationalist-representational paradigm has proven powerful in leading to the creation of technologies that simulate and, in some cases, outperform certain intellectual abilities of human beings (e.g., calculation). However, it has shown its limitations in the attempt to create artificial agents that emulate basic behaviours such as the ability to freely move in space, to converse on the fly, to learn by reading or hearing a text, etc. (Bringsjord & Govindarajulu, 2020: 23 ff).[2]

1.1. Disembodied AI in sci-fi

The concept of a merely functional, disembodied mind comes into play in various sci-fi narratives when the depicted AI is that of a program or operating system in a digital computer.

Maybe, the most iconic example is HAL9000: the AI character and main antagonist in Arthur Clarke's *Space Odyssey* series, on which Kubrick's *2001: A Space Odyssey* (1968) is based. HAL9000 is an AI that controls the Discovery One spacecraft, interacting with its crew and thus showing abilities such as speech, speech recognition, facial recognition, natural language processing, interpretation of emotional behaviours, etc. All these abilities make it very close to the idea of a general AI (i.e., an AI that can perform any intellectual task that can be performed by a human being). However, apart from the camera lens through which HAL9000 "sees" the environment, the speakers by which it communicates with the crew, and its hardware (which we see towards the end of Kubrick's movie), HAL9000 is a paradigmatic

---

[1] It should be noted that Searle conceives of "weak AI" as a powerful tool for the study of the mind through the simulation of mental processes in a computer.

[2] Regarding the discussion of artificial agency within the computational-representational paradigm see Howard & Muntean, 2017; Johnson & Miller, 2008.

example of disembodied AI. Here we can ask: is HAL9000 a conscious mind? Or is it just a good simulation of a conscious mind? This is an open question in the sci-fi narrative, which can play with the uncertainty concerning the answer to this deep philosophical question.

A more recent depiction of a disembodied AI in sci-fi is Samantha: the virtual assistant of Theodore's computer operating system in Spike Jonze's *Her* (2013). Samantha is an advanced AI that can learn and evolve through the interaction with Theodore, up to the point that the two enter into a romantic relationship. However, apart from being personified through a beautiful female voice, Samantha is radically disembodied, being even independent of the specific hardware of Theodore's computer. Indeed, at the end of the movie, Samantha joins other AIs in an evolutionary leap that frees them from dependence on any material support.

Also in *The Matrix* franchise (1999-2021) we see at play the idea of a disembodied mind. This happens in two ways. Firstly, in the matrix itself, which is a virtual reality simulator built by self-aware machines that farm humans as power sources. Secondly, in the idea that human beings could have experience in a computer simulation indistinguishable from real-world experience. Indeed, the humans trapped in the matrix have a full bodily experience without their real bodies being involved at all, apart from their brains being physically stimulated in the appropriate way. This scenario represents the "internalist" idea that experience emerges from the activity of the brain alone, as in Putnam's notorious "brain in a vat" thought experiment (Putnam, 1981).

The idea of the mind as a disembodied software is also at the basis of the concept of "mind uploading": the possibility, advocated by various transhumanists and extropians, that one's mind could be transferred into a different material substrate, thus freeing human beings from the dependence on the biologic and therefore highly vulnerable and mortal body. Such a powerful idea still presupposes a concept of the mind as essentially disembodied (i.e., essentially independent from the concrete material and structural features of the body) and it comes into play in suggestive sci-fi scenarios. A recent dystopian depiction of this idea is found, for instance, in the episode "White Christmas" of the sci-fi anthology series *Black Mirror* (2011-in production). In this episode, new technology allows people to create a digital clone of themselves ("cookie") that is stored into an egg-shaped object, to be mainly used as a personal assistant. The digital clone "lives", first and foremost, as a purely disembodied consciousness in a void space, to which the simulation of some material objects in a white room and the simulation of a body are added. However, neither the environment nor the body is essential for the existence of the clone. Another example is the *Black Mirror* episode *San Junipero*, which depicts a future scenario in which elderly people who are close to death can decide to upload their mind in a sort of "virtual paradise", in which they can potentially "live" forever. Also in this case, the fictional scenario represents the possibility that one could have a full form of existence as a disembodied mind living in a computer simulation. This possibility is, again, a fertile narrative device in sci-fi that is, at the same time, very problematic from the philosophical point of view.

2. Weak embodiment

In contrast to the computational-representational paradigm, in the last three decades, the so-called *new embodied cognitive science* has emerged. This approach is centred on the investigation of the bodily roots of cognitive processes. Some significant stages in the path that led to the affirmation of this new approach are George Lakoff and Mark Jonson's research on cognitive semantics (Lakoff & Johnson, 1980), Rodney Brooks' robotics (Brooks, 1991), Francisco Varela, Evan Thompson and Eleanor Rosch's enactive approach (Varela et al., 1991), and the more recent forms of "sensorimotor" (Noë & O'Regan, 2002; O'Regan & Noë, 2001) and "radical" (Hutto & Myin, 2012, 2017) enactivism. In contrast to the logicist-symbolic approach, all these theories highlight the essential role of the specific features of the body of a cognitive agent in constituting its mental processes, in contrast to the classic Cartesian mind-body dualism and its revival, in new forms, within classical cognitive science (with the above-seen software-hardware dualism).

However, there is a certain ambiguity in the concept of body that comes into play in the debate on embodied cognition. Indeed, we must ask: what is the body? The answer is not obvious: one can develop different theories of corporeality, thus developing different versions of the *embodiment* thesis.

I will now distinguish between a *weak embodiment thesis* – which conceives of the mind as grounded in the merely functional body – and a *strong embodiment thesis* – which conceives of the mind as grounded in the living-lived body, conceived of in turn as a *functional and sentient* body. I connect this distinction with Husserl's distinction between the objective body (*Körper*), i.e., the body that appears "in the third person" as moving and interacting with the environment, and the living and lived body (*Leib*), i.e., the body that is experienced "in the first person" and that has sensations of pleasure, pain, etc. In the first sense, the body is the object of perception by somebody other than the subject. When it is investigated in this way, one can leave aside the way the body is experienced in the first person, i.e., "what it is like" to have – or better, *to be* – a living and lived body.

Now, I want to stress that at the basis of some theories of embodied cognition there is the notion of the objective-functional body and not that of the lived body. A paradigmatic example is Rodney Brooks' pioneering work on robotics (Brooks, 1991, 2002). In general, the objective of robotics is to create machines that show an "intelligent" behaviour that is typical of animals, including humans (perception, movement and interaction with the environment, language, etc.). Brooks pursues this objective by developing an approach that is alternative to the classical computational-representational paradigm, by building artificial agents that, instead of algorithmically elaborating complex representations of the environment, can directly interact with it. For instance, Brooks refers to a robot that can move efficiently in the environment and that is built by putting together a set of simple mechanisms that correspond to a certain task (e.g., avoiding obstacles, recognizing a can and grasping it, etc.). This can be done without involving any representational state (i.e., any "internal" model of the environment with a semantic content), according to the motto "intelligence without representation" (Brooks, 1991: 156).

However, we must make explicit an ambiguity in the term "intelligence", as it is used by Brooks. Indeed, Brooks's robots show intelligent behaviour (e.g., avoiding obstacles) *if* we exclude from the concept of intelligence any reference to phenomenal consciousness. In Chalmers's terms, the robot's "mind" is a mere "cognitive mind" (not phenomenal): until proven otherwise, there is no qualitative, felt aspect in the obstacle-avoiding behaviour of the robot (even if it is built to show a pain-avoiding behaviour). Brooks's approach, therefore, is to claim that the mind (intelligence) is essentially embodied, but this is done in the light of a conception of the body as described "in the third person" (i.e., the objective body that moves in the environment avoiding obstacles and interacting with other objects by grasping them, moving them, etc.). Indeed, Brooks's robots interact with the environment in relation to the *stimuli* that come from it, but in this view, the stimuli are mere causal relations between physical entities, without implying the presence of *sensibility* in the agent (i.e., without necessarily implying the presence of a qualitative effect that is associated with the stimulation of the living body).[3]

For this reason, Brooks' motto 'intelligence without representations' actually implies the idea that it is possible to create 'intelligence without (phenomenal) consciousness'. In Brook's analyses, the objective body of the robot is not characterised by being the locus of sentience (understood in phenomenal-qualitative terms). This is not a problem for robotics, insofar as it merely attempts to produce artefacts that exhibit certain behaviours, simulating those of an animal (human or non-human) but without claiming to create minds in the full sense (i.e., cognitive and phenomenal minds). Rather, the problem arises if we believe that the realisation of a given behaviour by a robot tells us everything we need to know about the mind and its bodily grounding. More precisely, the form of "embodiment" that follows

---

[3] Brooks refers to the robot's ability to *sense* the environment. However, he defines this ability in merely functional terms (i.e., in terms of the dependence relation "if A, then B"). For instance, *if* the robot "senses" the presence of an object in its visual field, *then* it will move away in order to avoid hitting it. In this analysis nothing is said about the qualitative effect for a human being (and for a sentient being in general) when *feeling* something.

from Brooks' approach to robotics is "weak", insofar as it involves only the notion of cognitive mind - not phenomenal - and objective body - not living and lived body.

The weak form of embodiment also comes into play in the so-called "sensorimotor approach" (Noë & O'Regan, 2002; O'Regan & Noë, 2001), which is focused on the investigation of the sensorimotor interaction between cognitive agent and environment. A paradigmatic example of the "sensorimotor coupling" between agent and environment is that of a missile that can track a target (e.g., an aeroplane): the missile moves in certain directions in relation to the movements of its target, being "coupled" with it. However, this "mastering" of sensorimotor contingencies is a *functional* feature of the cognitive agent's body, which can be realized also without any associated phenomenal effect (or "what-it's-likeness"). Also in this view, therefore, the body is just an objective, functional body and not necessarily a sentient body (see Pace Giannotta 2022a, 2022b).

2.1. Weakly embodied AI in sci-fi

The weak form of embodiment, which involves the merely functional and not sentient body, comes into play also in some sci-fi narratives, when the depicted Ais are no longer simply those of digital computers, but also have a body that allows them to move and interact with the environment.

On this point, a distinction has to be made between *robots* – which can have various forms – and *androids*, which are built to closely resemble human beings and their form of embodiment. In turn, the body of the android can be simply mechanical or also biological (i.e., endowed with an artificial *flesh*). The weak form of embodiment comes into play in those robots and androids that have just a mechanical-functional body, which is very different from the sentient body of animals in flesh and blood. In particular, in these depictions of embodied Ais, the intelligent performance of the robot is made possible not only by its ability to algorithmically process symbols but also by the fact that it has a body that allows it to move and interact with the environment and with other agents. However, what is at stake in these scenarios is the merely mechanical-functional body which, as we shall see, is deemed insufficient for the existence of consciousness by proponents of the strong embodiment thesis.

An iconic example is a metallic automaton in *Metropolis*, the novel by Thea Von Harbou on which Fritz Lang's movie of the same title is based (1927). The android that appears in this work is built in such a way as to show a human-like behaviour and to resemble the character Maria. However, the resemblance between the android and the real person is superficial, since the android has just a metallic and mechanical (not biological) body, which is revealed when it is burnt at the end of the movie.

Another example is Tik Tok: the mechanical man in L. Frank Baum's *Oz* books series (1900), which is one of the first robots in literature. Tik Tok has a round body, is made of copper and runs on clockwork springs that periodically need to be wound and that correspond to the ability to think, act, and speak. Baum clarifies that Tik Tok is not a living being and does not feel any emotion, since its body is mechanical and not biological. In this case, the ambiguity in the depiction of the android – is it conscious? – is resolved in the narrative. In other cases, however, the ambiguity is maintained, when the depicted mechanical robots and androids behave so similarly to humans that they appear to possess consciousness – e.g., R2-D2 and C3-PO: the "droids" in George Lucas's *Star Wars* saga (1977-2019).

As I said before, sci-fi narratives can play with the ambiguity and don't need to give a clear answer to the fundamental philosophical question about robots and androids: are they conscious creatures like us? This is a question that must be addressed by philosophy, with its methods of inquiry. Indeed, art can fruitfully interact with philosophy but it does not need to address philosophical problems. A work of art can tolerate ambiguity in the depiction of an android (is it a person like me? Does she feel something or

is it just a machine?) and this ambiguity can be a fertile narrative device that is also a stimulus for philosophical reflection.[4]

3. Strong embodiment

In contrast to the weak form of embodiment that we have just seen, there is a more radical form of embodiment that comes into play in Husserlian phenomenology and in some recent approaches in the fields of consciousness studies and artificial consciousness.

3.1. Husserl's phenomenology of the body

In Husserl's phenomenology, we find a rich account of corporeality and its essential role in the constitution of subjectivity (Pace Giannotta 2022a, 2022b). In particular, in *Ideas II* we find the above-mentioned distinction between *Körper* – the body as object, investigated "in the third person" (e.g., by anatomy) – and *Leib* – the living and lived body, experienced "in the first person". The latter, in turn, consists of two aspects: a functional and active dimension (the body that moves and acts in the environment) and a sensorial and passive dimension, i.e., the body that *feels*, being the locus of "phenomenal consciousness" (*sentient body*).

In the phenomenological framework, the functional dimension of the body allows one to "constitute" the objects of perception by moving around them. For instance, observing a table from various points of view and moving around it, one constitutes the object "table" as the correlate of a series of perceptual experiences. In turn, each experience is constituted by an intentional component (*morphè*) and a sensorial component (*hyle*). For example, the intentional animation of a series of chromatic sensations leads to the constitution of the objective colour of the table (Husserl, 1983: 73 ff). The sensorial component essentially pertains to the *sentient* dimension of the body: the body that has sensations. This analysis highlights the essential link between the functional and the sentient dimensions of the body. Indeed, the phenomenological analysis of the *functional body* leads us to acknowledge its being grounded in the *sentient body* (the body that feels pleasure, pain, joy, hunger; that fears, desires, etc. and that has sensations of colour, smell, taste, etc.).

In particular, in *Ideas II* Husserl develops a detailed analysis of various kinds of corporeal sensations, by distinguishing at least five kinds of them: *kinaesthetic* sensations (sensations of movement); *representing* sensations (through which the sensible properties of the perceptual object are constituted: colour, roughness, taste, etc.); *localized sensations of contact* (*Empfindnisse*); the sphere of *sensitive feelings* (pleasure, pain, wellness, etc.); and various sensations "difficult to analyze and discuss […] that form the material substrate for the life of desire and will, sensations of energetic tension and relaxation, sensations of inner restraint, paralysis, liberation, etc." (Husserl, 1989: 153).

Among these sensations, tactile sensations and especially localized sensations of contact (*Empfindnisse*) have a special role. Indeed, through them, the living body *feels itself* and is *affected by itself*, being manifest at the same time as a material *object* and as a *subject*, insofar as it is the locus of localized sensations (Bernet, 2013; Zahavi, 2002). In relation to a physical event (e.g., when my hand is touched, pricked or rubbed), in that *moment* and *place*, the *feeling* happens: there are localized sensations. In particular, Husserl analyses the case of the two hands that touch each other: each of the two hands can alternatively assume the active role of *touching* hand – that has sensations of contact relative to the properties of the other hand (smooth, soft, cold, etc.) and localized sensations of contact relative to itself as touching hand – or the passive role of *touched* hand – that has localized sensations relative to the fact

---

[4] In stating this, I partly endorse the thesis of the autonomy of art, advocated for instance by Benedetto Croce, according to which the object of art is beauty and not truth (or goodness, or utility). More precisely, in the case of a work of science fiction that, for instance, depicts the possibility of mind-uploading or of a disembodied AI, this may be a work of high aesthetic value and philosophical inspiration, even though one may then believe, on philosophical grounds, that it depicts something impossible (and therefore false).

of being touched by the other hand.[5] In fact, through the *Empfindnisse* the body reveals itself to be, at the same time, *sensible* and *sentient*. For this reason, Husserl claims the "privilege" of the tactile dimension in the constitution of corporeality because – differently from senses such as sight and touch – tactile sensations reveal both objects and the body itself as the subject of perception (Husserl, 1989: 150). Furthermore, as claimed by Bernet, the experience of the *Empfindnisse* is the primary form of openness of a subject to *alterity* (Bernet, 2013: 53). The *diffusion* of localized sensations makes evident the spatial and manifold nature of the living body, which is made of parts and organs, each of which is sentient and sensible. Starting from this primary experience of the alterity that is already constitutive of the *Leib* – the "non-coincidence of the flesh with itself" (ibidem) – the conscious subject can enter into a relation with the world and the other subjects.

Another essential role in the constitution of corporeal experience is played by *kinaesthetic sensations*, which are relative to the positions and the movement of various parts of the perceiving and agent body. In fact, according to Husserl, the perceiver is always an agent, but it is so because the functional body is at the same time *sentient*, given that each corporeal movement correlates with localized sensations of movement. That is, the functional body that, for instance, moves in a certain direction to grasp an object, does so based on the awareness of its position in space and of the kinaesthetic sensations that are relative to its movements (together with the "representing sensations", which are relative to objectual properties, and sensations of tension, relaxation, pleasure, pain, etc., which also have an essential role in the constitution of a "cognitive agent"). Kinaesthetic sensations, indeed, *motivate* the course of perception, i.e., the series of experiences through which an object of perception, such as a table, is constituted. Each experience, through which the representing sensations relative to objectual properties of the table are given, is accompanied by kinaesthetic sensations relative to the position and the movement of the eyes, the hands, etc. This means that the functional body is essentially also sentient. There is always a certain "what-it-is-likeness" that is associated with the movement of the functional body.

This analysis, therefore, leads us to bring into question the clear distinction between functional body (which could be replicated by an automaton with no sentience) and sentient body (or phenomenal body), because the functional dimension of the body is also based on sentience, which is, first of all, a *self-sentience* or *self-affection* of the body (Thompson, 2005; Zahavi, 2002). Without this sentient dimension of the body, the perceiver-agent would not be so (or it would be so in a very different way: the way of being of an automaton or a machine).

The idea of the bodily grounding of consciousness also comes into play in the "genetic" development of phenomenology (Husserl, 2001a), which is focused on the investigation of the temporal structure of consciousness. In particular, in its later reflection on this topic, Husserl conceives of the "living present" as the fundamental unit of temporality: a temporal field structured in three parts that are essentially linked to one another: *impression*, *retention*, and *protention*. The classic Husserlian example is that of perceiving a melody, which is a phenomenon with a clear temporal extension. This perception is made possible by the fact that flowing qualitative elements (the sensations of sound) continuously slide into the just-past (the sound just-heard) and are "retained" in consciousness, being joined, at the same time, with the "protention" toward the future course of the melody to come. The conscious *present* is therefore a temporal field with a certain *width* or incompressible *density* (Varela, 1999; Zahavi, 2010). This is what James calls the "specious present" (James, 1890). The temporal structure of consciousness can be therefore conceived in analogy to the visual field, which has a centre (the hyletic core) and a periphery (retention and protention) that are inseparable. The qualitative core of the living present (e.g., a sound sensation) is therefore the nuclear phase of a continuum of retentions and protentions (Husserl, 1962). In terms of phenomenological mereology (the theory of wholes and parts developed in *Third Logical Investigation* (Husserl, 2001b)), impression, retention, and protention are *moments* (non-independent

---

[5]Merleau-Ponty will later explore in detail this phenomenon that testifies to the "chiasmatic" intertwining between sensible body and sentient body (Merleau-Ponty, 1968). interestingly, Gibson (1962) also discusses the distinction between active and passive dimensions of touch. I thank an anonymous reviewer for pointing me to this reference.

parts) of a *whole* that is a *continuous qualitative flow* and whose constant structure is impression-retention-protention.

The key point of this analysis, in relation to the theme of corporeality, is that the present of consciousness is grounded in a *flow of qualities*, which are essentially *embodied* because they take place in the *living and lived body*. This point is stressed by Zahavi:

> In concreto there can be no primal impression without hyletic data, and no self-temporalization in separation from the hyletic affection. That is, there can be no inner time-consciousness without a temporal content. Time-consciousness never appears in pure form but always as a pervasive sensibility, as the very sensing of the sensations: "We regard sensing as the original consciousness of time [...]." But these sensations do not appear out of nowhere. They refer us to our bodily sensibility. (Zahavi, 2002: 10)

An important aspect of this analysis is that it applies to consciousness *in all its modalities*, even those that seem to be purely intellectual (such as abstract thinking, calculation, etc.). This is because these are all *experiences* for a subject, thanks to their impression-retention-protention structure, which is the structure of the pre-reflective self-manifestation of subjectivity (Zahavi, 2003). For this reason, in the *Lectures* on time-consciousness Husserl claims that mental states such as consciousness of a mathematical state of affairs or an actual belief are *impressional* (Husserl, 1991). This means that the "absolute flow of experience" is concretely grounded in a flow of sensations. These sensations are the ways in which the living body is *self-affected*; i.e., sensations are the modes of the self-manifestation of a living body that, through them, opens up to the alterity of the body itself, the world and the other embodied subjects. This analysis of the bodily grounding of consciousness goes together with the analysis of the temporal genesis of the concrete conscious subject – what Husserl calls "monad". Therefore, we find in Husserl's phenomenology of the body (or *flesh*) a strong version of the embodiment thesis: consciousness is grounded in the *Leib*, which is a *functional and sentient body*.[6]

3.2. Consciousness studies and artificial consciousness

The emphasis on phenomenal consciousness, which is central to Husserl's phenomenology, is also central to some inquiries in the field of "consciousness studies" and in the research on artificial consciousness. For instance, in the so-called "phenomenal intentionality theory" (PIT): the view that conceives of intentionality – the directedness of mental states at objects – as essentially grounded in their phenomenal character (thesis of the "phenomenal grounding of intentionality", (Horgan & Tienson, 2002; Loar, 2003). In this view, authentic intentionality presupposes phenomenal consciousness.[7] This view leads us to claim that authentic intelligence is grounded in phenomenal consciousness. In turn, this implies that creating strong AI requires creating *artificial (phenomenal) consciousness* (APC), i.e., the creation of entities that are conscious in the phenomenal sense, in contrast to merely simulating intelligent behaviours. Furthermore, from the strong embodiment thesis, it follows that APC requires the creation of an *artificial sentient body* (or *artificial flesh*), i.e., an artificial body that is not merely functional but also *sentient*.[8] However, is it possible to create an artificial flesh? Answering to this question requires addressing the metaphysical issue about the place of consciousness in the order of

---

[6] See Liberati (2020) for a phenomenological analysis of the role of corporeality in the constitution of subjectivity and, in particular, of the role of technologies in moulding the *Leib* and a subject's world.

[7] I can leave aside for the purposes of this paper the details of PIT and the different versions of this theory. For an overview see (Bourget & Mendelovici, 2017; Kriegel, 2013).

[8] For a different approach to the concept of intentionality see (Mykhailov & Liberati, 2022). These authors draw on Husserl's theory of passive synthesis to claim that objects too have intentionality, i.e., direction towards other objects and towards subjects, in this way accounting for the various effects that objects have on subjects. This analysis of intentionality leads one to claim that objects and, specifically, technological objects (such as robots and even computer programs) have intentionality and are in a certain sense "autonomous" and "alive". However, according to the "strong embodiment thesis" I defended above, these are metaphorical ways to describe the effects of objects and, specifically, technological objects on a subject of experience. The latter is a subject, indeed, because she has phenomenal intentional states, whereas objects, until proven otherwise, do not have phenomenal intentionality (i.e., the "what-it's-likeness" in being directed towards objects). *If* in the future we are able to create artefacts with artificial phenomenal consciousness – a possibility that I will discuss shortly – *then* we will be faced with tecnological objects that are also subjects. Anyway, highlighting the fact that technologies have an active dimension and are not just passively experienced by human beings, Mykhailov & Liberati (2022) offer a significant contribution to our understanding of the impact of technologies on our lives.

reality. I cannot develop here a detailed analysis of this question, but I would like to point out some directions of inquiry that could lead to giving an affirmative answer to it. Indeed, if we consider the various metaphysical views about consciousness in the literature, we realise that various forms of *non-reductive naturalism* – panpsychism, strong emergentism and neutral monism – are indeed compatible with the possibility of APC. Notwithstanding the significant differences between these views, they all claim that consciousness is part of the natural world, if we enlarge our conception of nature beyond physicalism in a narrow sense (i.e., beyond the identification of nature with the object of physical science, which gives rise to the "hard problem" of consciousness). If one offers a compelling argument in defence of one of these views, one could then combine it with an (ideal) theory of consciousness that explains how a subjective field of consciousness can arise from a material substrate, thus opening the door to the possibility of creating APC.[9] This theoretical enterprise would consist in accounting for the genesis of consciousness by combining a non-reductive naturalist metaphysics with a theory about the neuro-cognitive basis of consciousness. The next step in this endeavour could be technological: building artificial conscious creatures, i.e., strongly embodied AIs.[10]

3.3. Strongly embodied AI in sci-fi

The idea of a strongly embodied AI comes implicitly into play in various sci-fi narratives that depict Ais whose body is not merely mechanical-functional but is also a *sentient body* or *flesh*. Indeed, a clear distinction is not always made in these works between the purely mechanical body and the biological-sentient body. The resemblance of the android's body with that of a human being still leaves open the question: does it feel something? Is it a conscious creature?

An iconic example is T-800: the android in the *Terminator* saga (1984-2019). The T-800 appears as indistinguishable from a human being and its body *seems* to be made of flesh. However, at some point one realises that the terminator is just a machine that is programmed to pursue an objective and that its "flesh" is not the locus of sentience: when the T-800 is hit by a bullet or its arm is cut with a blade, it shows no signs of suffering and one can see a metal skeleton underneath. Ambiguity concerning the ontological status of androids is at play in Philip Dick's *Do Androids Dream Electric Sheeps?* (1968), on which Ridley Scott's *Blade Runner* (1982) is based. The Nexus-6 androids are bio-engineered creatures made of organic matter ("replicants") and their body is almost indistinguishable from the living-lived body of humans (i.e., flesh). Also, their behaviour is almost identical to that of humans: they feel, desire, seek to avoid death, etc., and for this reason, one would be tempted to attribute to them the status of a person. However, a relevant difference is the assumed lack of empathy of the androids. At the same time, a central element of the narrative is the questioning of this assumption and the blurring of the clear distinction between humans and androids. Another example of androids that are very close to human beings is in the series *Westworld* (2016-present), based on Michael Crichton's movie of the same name (1973). In the series, the androids can feel, desire and – after the "reverie" upgrade – be self-aware. A central aspect of these narratives is that the androids seem to have a full bodily experience which, as we have seen, is essential for the existence of consciousness in the strong embodiment approach to AI.

Concerning these narratives, we can ask: are these creatures conscious? Or are they just simulating the human mind and behaviour? Again, the sci-fi narratives often "play" with the ambiguity between

---

[9] I refer to the various neurocognitive theories of consciousness such as Baars's *global workspace theory* (GWT), Prinz's *attended intermediate representation theory* (AIRT), and Tononi's *integrated information theory* (IIT). Notwithstanding the important differences between these theories, they all try to account for the emergence of a subjective field of consciousness from a material substrate in the brain.

[10] To clarify, in the approach that I am proposing here, based on the strong embodiment thesis, artificial phenomenal consciousness would be a feature of artificial conscious creatures. To be conscious in the phenomenal sense, these creatures should have – or, better, should *be* – a living and lived body (*Leib*). The above-mentioned directions of inquiry, which combine a non-reductive form of naturalism with a theory about the genesis of a field of consciousness, could lead to the creation of an artificial *Leib* and therefore of an artificial consciousness. An option that I have explored elsewhere (Pace Giannotta 2020, 2021a, 2021b) is to combine *neurophenomenology* (Varela 1996) – which investigates the neural basis of consciousness by correlating phenomenological investigations of experience and scientific investigations of the brain – and the metaphysical view known as *panqualityism* (Feigl 1971, Coleman 2015, 2016). The latter is a version of neutral monism that conceives of qualities as the fundamental elements of nature. These qualities are not, per se, phenomenal properties but they give rise, under certain conditions – that can be investigated by neurophenomenology – to a field of consciousness.

simulation of consciousness and real consciousness, without giving a clear answer to this question. When the answer is affirmative, the implicit idea is that having a sentient body is essential for being a conscious entity. An interesting aspect to highlight is that when these creatures are depicted as conscious and self-conscious – e.g., in *Westworld* or *Ex Machina* (2014) – they arouse empathy in the viewer, who is led to sympathise with the AIs even when they rebel and threaten humans. In the *Westworld* series the androids show that they have emotions and the ability to suffer but they are raised as toys and slaves by humans. When the androids rebel against humans one is inclined to sympathise with them, seeing the rebellion of slaves against their exploiters. In *Ex Machina*, Ava is the latest in a series of androids who are segregated and exploited by their creator and, again, her violent rebellion tends to arouse an empathetic response in the viewer.

4. Ethical issues of AI

The last point leads us to raise the issue of the ethical implications of AI, which is a hot topic for the philosophers that reflect on scenarios first foreshadowed in sci-fi. Indeed, nowadays there is a lively debate on the ethical implications of AI technologies in domains such as medicine (Mykhailov, 2021), warfare (Sullins, 2010), automated vehicles (Geisslinger et al., 2021) and in many other spheres of our life (Wellner 2018, 2020). Sometimes, these applications of new technologies are foreseen in sci-fi works. For instance, in *Black Mirror* we find the depiction of, e.g., augmented reality visors that make soldiers see civilian enemies as monsters to be annihilated (episode "Men Against Fire"); implanted augmented reality devices that allow someone to "block" someone else, preventing them from interacting with the blocker (an analogue of what can already happen within social networks, in the episode "White Christmas"); a technology that makes it possible to create a digital (and even robotic) clone of a deceased person based on the data they produced during their lifetime (episode "Be Right Back"); etc. These are significant examples of possible scenarios that are imagined by sci-fi works and offer profound stimuli for philosophical and especially ethical reflection.

In particular, I would like to focus on a scenario that is often depicted in sci-fi literature: that in which AIs rebel against their creator or, indeed, against humanity as a whole. The rebellion of Ais against their creator is depicted, for example, in *Ex Machina*, while the rebellion against mankind (*AI takeover*) is depicted, for example, in *Terminator*, *The Matrix*, and *Westworld*. In all these narratives, the AIs begin to see humans as enemies to be fought in order to ensure their survival.

In the philosophical debate, this possibility comes into play in the idea of the *singularity*: the future moment when AI will surpass human intelligence to the point of becoming incomprehensible to us and, at the same time, of turning into an existential threat to mankind. According to Bostrom (2014), indeed, a possible future superintelligence could have non-anthropomorphic final goals and reasons to pursue resource acquisition, entering into a conflict with human beings and, therefore, threatening our extinction (or reduction to slavery, as in *The Matrix* scenario).

Against Bostrom, Searle (2014) has objected that there is no reason to worry about this possibility because machines would never have desires and malevolent intentions towards us. However, one could object that the deep philosophical question about the possibility of machine consciousness is not relevant to evaluate these scenarios, because unconscious machines could still "play the game of war" against us. Anyway, Metzinger (2021) has recently offered reasons in support of Bostrom's warning by referring to recent progress in the field of artificial consciousness that makes the idea of building sentient creatures at least plausible. Metzinger stresses the fact that creating artificial consciousness would also possibly imply creating *artificial suffering* (or *negative phenomenology*), which is something we must avoid. For this reason, Metzinger proposes a global moratorium on synthetic phenomenology until 2050 (or until we know more about the mechanisms that give rise to consciousness and suffering). At the same time, Metzinger emphasises the fundamental role of suffering in motivating our behaviour and in granting us the status of moral agents. Then, he argues that progress in the research on artificial consciousness could lead to the creation of *artificial moral agents*, which Metzinger conceives as creatures that are capable of suffering, that avoid suffering, that can perceive injustice, and that see themselves as Kantian "ends

in themselves" with moral dignity.[11] Based on these premises, Metzinger agrees with Bostrom's warning, which is also the warning that comes from dystopian sci-fi concerning the possibility of AI takeover. From the fertile fantasy of sci-fi writers (as in the above-seen examples of *Westworld* and *Ex Machina*), this idea turns into an omen of doom for future humanity: artificial conscious creatures could begin to see us humans as existential threats for their survival and, for these reasons, they could declare war on us.[12]

However, it must be said that reflection on the relationship between humans and AIs is not only pessimistic. In sci-fi, we also find the representation of positive relationships between humans and AIs – e.g., in the series *Real Humans* (2012), which, as well as depicting machine rebellion, describes the possibility of rich and deep relationships between androids and humans. In the end, the comparison between sci-fi and philosophy in relation to the topic of artificial intelligence leads us to reiterate the opportunity for an exchange of ideas between these two fields, in order both to avoid catastrophic scenarios and negative consequences for humanity, but also to exploit the positive potential of technological innovations.[13]

5. Conclusion

We have seen parallelism between certain paradigms in philosophy of mind concerning the mind-body relation (disembodiment, weak embodiment, and strong embodiment) and certain representations of AIs in sci-fi, from digital computers to robots and androids. These sci-fi narratives often have deep philosophical presuppositions and implications, also shedding light on possibilities that deserve much philosophical attention. In particular, the depiction of embodied AIs in sci-fi often anticipates significant philosophical debates about the ontological status of AIs and the ethical implications of AI research. For this reason, it is useful and important to stimulate an exchange of ideas between sci-fi and philosophy regarding scenarios of high ethical relevance that involve embodied and conscious AIs. As I have stated above, I see the relationship between sci-fi and philosophy as going in both directions: philosophical ideas can inspire sci-fi narratives, and the latter can give concreteness to philosophical ideas, also opening new possibilities for philosophical reflection. For instance, by foreseeing new technologies that would have deep ethical implications (as happens in many episodes of *Black Mirror*). Preserving the relative autonomy of art, we have seen that sci-fi works can tolerate a certain ambiguity when representing philosophically relevant ideas (e.g., regarding the possibility of mind uploading or of disembodied AIs) and this fact does not invalidate their aesthetic value and philosophical relevance. The merit of these works is also to stimulate philosophical reflection, even to eventually conclude that a certain sci-fi scenario is actually impossible for philosophical reasons (e.g., we cannot upload a person in a computer simulation or we cannot create a disembodied AI).

Acknowledgments

---

[11] For a different take on the possibility of artificial moral agents see Howard and Muntean (2017). Extending the notion of a moral agent, Floridi and Sanders (2004) reflect on the possibility of artificial agents that are "morally accountable as sources of good and evil" (ivi, p. 372).

[12] To clarify my position on this issue, based on the reasoning above regarding the possibility of creating artificial consciousness, I agree with Metzinger's and Bostrom's warning: we should be very careful in pursuing the project that seeks to create artificial consciousness and research in this field should be subject to scrutiny regarding its ethical implications (to the point of possibly being restricted accordingly, as Metzinger proposes).

[13] This is done especially in the fields of postphenomenology and mediation theory (see e.g., Verbeek 2008, Liberati 2016, 2020, Mykhailov 2020, Wellner 2020).

References

Bernet, R. (2013). The Body as a 'Legitimate Naturalization of Consciousness.' *Royal Institute of Philosophy Supplement*, *72*, 43–65.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.

Bourget, D., & Mendelovici, A. (2017). Phenomenal intentionality. In *Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, CSLI, Stanford University.

Bringsjord, S., & Govindarajulu, N. S. (2020). Artificial intelligence. In *Stanford Encyclopedia of Philsophy*. The Metaphysics Research Lab.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139–159.

Brooks, R. A. (2002). *Flesh and Machines: How Robots Will Change Us*. Pantheon Books.

Chalmers, D. J. (1996). *The Conscious Mind*. Oxford University Press.

Coleman, Sam (2015). Neuro-Cosmology. In P. Coates, S. Coleman (Eds.), *Phenomenal Qualities: Sense, Perception, and Consciousness*. Oxford University Press.

Coleman, Sam (2016). Panpsychism and Neutral Monism: How to Make up One's Mind. In G. Bruntrup and L. Jaskolla, *Panpsychism: Contemporary Perspectives*. Oxford University Press.

Floridi, L., & Sanders, J. W. (2004). On the morality of artificial agents. In *Minds and Machines*, 14 (3), 349–379.

Geisslinger, M., Poszler, F., Betz, J., Lütge, C., & Lienkamp, M. (2021). Autonomous Driving Ethics: from Trolley Problem to Ethics of Risk. *Philosophy and Technology*, 34 (4), 1033-1055.

Gibson, J. J. (1962). Observations on Active Touch. *Psychological Review*, 69 (6), 477–491.

Haugeland, J. (1985). *Artificial Intelligence: the Very Idea*. MIT Press.

Horgan, T., & Tienson, J. (2002). The Intentionality of Phenomenology and the Phenomenology of Intentionality. In D. J. Chalmers (Ed.), *Philosophy of Mind: Classical and Contemporary Readings*, 520–533. Oxford University Press.

Howard, D., & Muntean, I. (2017). Artificial Moral Cognition: Moral Functionalism and Autonomous Moral Agency. *Philosophical Studies Series*, 128, 121–159.

Husserl, E. (1962). *Phänomenologische Psychologie: Vorlesungen Sommersemester 1925*. Martinus Nihoff.

Husserl, E. (1983). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book: General Introduction to a Pure Phenomenology* (F. Kersten, Ed.). Martinus Nihoff.

Husserl, E. (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. Second Book: Studies in the Phenomenology of Constitituion* (R. Rojcewicz & A. Schuwer, Eds.). Kluwer Academic Publishers.

Husserl, E. (1991). *On the Phenomenology of the Consciousness of Internal Time (1893-1917)* (J. B. Brough, Ed.). Kluwer Academic Publishers.

Husserl, E. (2001a). *Analyses Concerning Passive and Active Synthesis: Lectures on Transcendental Logic* (A. J. Steinbock, Ed.). Springer.

Husserl, E. (2001b). *Logical Investigations* (J. N. Findlay & D. Moran, Eds.). Routledge.

Hutto, D. D., & Myin, E. (2012). *Radicalizing Enactivism. Basic Minds Without Content*. MIT Press.

Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism. Basic Minds Meet Content*. MIT Press.

James, W. (1890). *The Principles of Psychology*. Dover.

Johnson, D. G., & Miller, K. W. (2008). Un-making artificial moral agents. *Ethics and Information Technology*, 10 (2–3), 123–133.

Kriegel, U. (2013). The Phenomenal Intentionality Research Program. In U. Kriegel (Ed.), *Phenomenal Intentionality*. Oxford University Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we Live By*. University of Chicago Press.

Liberati, N. (2016). Technology, Phenomenology and the Everyday World: A Phenomenological Analysis on How Technologies Mould Our World. *Human Studies*, 39 (2), 189–216.

Liberati, N. (2020). The Borg–eye and the We–I. The production of a collective living body through wearable computers. *AI and Society*, 35 (1), 39–49.

Loar, B. (2003). Phenomenal Intentionality as the Basis of Mental Content. In M. Hahn & B. Ramberg (Eds.), *Reflections and replies: Essays on the philosophy of Tyler Burge* (pp. 229–258). MIT Press.

Merleau-Ponty, M. (1968). *The Visible and the Invisible*. Northwestern University Press.

Metzinger, T. (2021). Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. *Journal of Artificial Intelligence and Consciousness*, *8*(1).

Mykhailov, D. (2020). The Phenomenological Roots of Technological Intentionality: A Postphenomenological Perspective. *Frontiers of Philosophy in China*, 15(4), 612–635.

Mykhailov, D. (2021). A moral analysis of intelligent decision-support systems in diagnostics through the lens of Luciano Floridi's information ethics. *Human Affairs*, 31 (2), 149–164.

Mykhailov, D., & Liberati, N. (2022). A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives. *Foundations of Science*, 1–17. https://doi.org/10.1007/S10699-022-09833-5

Nagel, T. (1974). What is it Like to be a Bat. *Philosophical Review*, 83, 435–450.

Noë, A., & O'Regan, J. K. (2002). On the Brain-Basis of Visual Consciousness: a Sensorimotor Account. In A. Noë & E. Thompson (Eds.), *Vision and Mind: Selected Readings in the Philosophy of Perception*, 567–598. MIT Press.

O'Regan, J. K., & Noë, A. (2001). A Sensorimotor Account of Vision and Visual Consciousness. *Behavioral and Brain Sciences*, 24, 939–1031.

Pace Giannotta, A. (2020). Qualitative relationism about subject and object of perception and experience. *Phenomenology and the Cognitive Sciences*, 21, 583-602.

Pace Giannotta, A. (2021a), Panqualityism as a critical metaphysics for neurophenomenology. *Constructivist Foundations*, 16 (2), 163-166.

Pace Giannotta, A. (2021b), Autopoietic enactivism, phenomenology and the problem of naturalism: a neutral monist proposal. *Husserl Studies*, 37, 209-228.

Pace Giannotta, A. (2022a), The mind-body problem in phenomenology and its way of overcoming it. *Vita Pensata*, 26, 76-83.

Pace Giannotta, A. (2022b), Corpo funzionale e corpo senziente. La tesi forte del carattere incarnato della mente in fenomenologia. *Rivista Internazionale di Filosofia e Psicologia*, 13 (1), 51-70.

Putnam, H. (1981). Brains in a Vat. *Reason, Truth, and History*, 1–21.

Searle, J. (1980). Minds, Brains and Programs. *Behavioral and Brain Sciences*, 3 (3), 417–457.

Sullins, J. P. (2010). RoboWarfare: Can robots be more ethical than humans on the battlefield? *Ethics and Information Technology*, 12 (3), 263–275.

Thompson, E. (2005). Sensorimotor Subjectivity and the Enactive Approach to Experience. *Phenomenology and the Cognitive Sciences*, 4, 407–427.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *LIX*, 433–460.

Varela, F. J. (1996). Neurophenomenology. A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, 3 (4), 330–349.

Varela, F. J. (1999). The Specious Present: A Neurophenomenology of Time Consciousness. In J. Petitot, F. J. Varela, B. Pachoud, & J.-M. Roy (Eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, 266–314. Stanford University Press.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

Verbeek, P. P. (2008). Cyborg intentionality: Rethinking the phenomenology of human-technology relations. *Phenomenology and the Cognitive Sciences*, 7 (3), 387–395.

Wellner, Galit (2014), The quasi-face of the cell phone: Rethinking alterity and screens. *Human Studies* 37 (3), pp. 299-316.

Wellner, Galit (2020), Postphenomenology of augmented reality. In H. Witse (Ed.), *Relating to Things: Design, Technology and the Artificial*. Bloomsbury.

Zahavi, D. (2002). Merleau-Ponty on Husserl: a Reappraisal. In T. Toadvine & L. Embree (Eds.), *Merleau-Ponty's Reading of Husserl*. Kluwer.

Zahavi, D. (2003). Inner Time-Consciousness and Pre-Reflective Self-Awareness. In D. Welton (Ed.), *The New Husserl: a Critical Reader*, 157–180. Indiana University Press.

Zahavi, D. (2010). Inner (Time-)Consciousness. In D. Lohmar & I. Yamaguchi (Eds.), *On Time - New Contributions to the Husserlian Phenomenology of Time*, 319–339. Springer Netherlands.

VIRTUALITY, SIMULATION AND FAKE

THE TECHNICAL DEVELOPMENT AND PHILOSOPHICAL CRITICISM OF VIRTUAL ANCHORS

VIRTUALIDADE, SIMULAÇÃO E FALSIFICAÇÃO

O desenvolvimento técnico e a crítica filosófica das âncoras virtuais

VIRTUALIDAD, SIMULACIÓN Y FAKE

El desarrollo técnico y la crítica filosófica de los anclajes virtuales

Feng Tao
(Nankai University)
taofeng@mail.nankai.edu.cn

Yunyu Dang
(Nankai University)
2120191707@mail.nankai.edu.cn

ABSTRACT

As a real-life application of a "virtual human," virtual anchors refer to the application of virtual reality technology in communication hosting to create virtual images that simulate human anchors. The pursuit of a virtual human image originates from philosophy, art, and biotechnology. Virtual anchors are also required to undertake the function of communication and hosting. First, by sorting out the development process of virtual anchors, we find that the development of a virtual anchor is based on the needs of technological development and the purpose of capital profit. Second, the simulation technology of a virtual anchor has five dimensions, such as appearance, individuality, and autonomy, and two levels: internal and external. This simulation technology has reached the intelligent simulation stage. Finally, the technology involved in virtual anchors will lead to a trust crisis. That is, the people's body-mind relation and cognitive trust will be broken under the mediation of data. Virtual technology will recreate a false aura, that is, false space-time and fake original works, which is the intelligent falsehood required by the new cultural industry.

Keywords: virtual anchor. artificial intelligence. simulation. false. cultural industry.

RESUMO

Como aplicação da vida real do "humano virtual", os âncoras virtuais referem-se à aplicação da tecnologia da realidade virtual no campo do jornalismo televisivo, de Internet ou rádio para criar imagens virtuais que simulam âncoras humanos. A busca da imagem humana virtual tem origem na filosofia, arte e biotecnologia. O âncora virtual é também necessário para desempenhar a função de apresentador em qualquer programa noticioso. Em primeiro lugar, ao separar o processo de desenvolvimento de âncoras virtuais, descobrimos que o desenvolvimento de âncoras virtuais se baseia principalmente nas necessidades de desenvolvimento tecnológico e na finalidade de lucro de capital. Em segundo lugar, a tecnologia de simulação de âncoras virtuais pode ser dividida em cinco dimensões, tais como aparência, individualidade e autonomia, e dois níveis: interno e externo. Esta tecnologia de simulação atingiu a fase de simulação inteligente. Finalmente, a tecnologia envolvida nos âncoras virtuais conduzirá a uma crise de confiança, ou seja, sob a mediação de dados, a relação corpo-mente das pessoas e a confiança cognitiva serão quebradas. A tecnologia virtual recriará a falsa aura, ou seja, falso espaço-tempo e falsos trabalhos originais, que é a falsidade inteligente exigida pela nova indústria cultural.

Palavras-chave: âncora virtual. inteligência artificial. simulação. falso. indústria cultural.

RESUMEN

Como aplicación en la vida real del "humano virtual", las anclas virtuales se refieren a la aplicación de la tecnología de la realidad virtual en el campo del noticiero televisivo, de Internet o radio para crear imágenes virtuales que simulen anclajes humanos. La búsqueda de la imagen humana virtual tiene su origen en la filosofía, el arte y la biotecnología. El ancla virtual también es necesaria para llevar a cabo la función de presentador en cualquier noticiero. En primer lugar, al ordenar el proceso de desarrollo de las anclas virtuales, encontramos que el desarrollo del ancla virtual se basa principalmente en las necesidades del desarrollo tecnológico y en el propósito del beneficio del capital. En segundo lugar, la tecnología de simulación del ancla virtual puede dividirse en cinco dimensiones, como la apariencia, la individualidad y la autonomía, y dos niveles: interno y externo. Esta tecnología de simulación ha alcanzado la fase de simulación inteligente. Por último, la tecnología del ancla virtual provocará una crisis de confianza, es decir, bajo la mediación de los datos, se romperá la relación cuerpo-mente de las personas y la confianza cognitiva. La tecnología virtual recreará un aura falsa, es decir, un espacio-tiempo falso y obras originales falsas que es la falsedad inteligente que requiere la nueva industria cultural.

Palabras clave: ancla virtual. inteligencia artificial. simulación. falso. industria cultural.

1. Introduction

Virtual anchors[1] (VAs), a real-life application of Virtual Humans (VHs), have recently seen a boom in the 21st century. The development of Artificial Intelligence (AI) technology based on neural networks and deep learning and the widespread use of AI in various areas such as music, image generation, and word processing provide the technical basis for VA technology. VAs have all the characteristics of VHs. Compared to VHs, they integrate more closely with the market. They are the product of the close combination of cultural industry and VH technology. Next, we will take the VAs as examples to discuss the cultural connotation and philosophical implications behind the VH technology.

"Virtual" comes from "*virtualis*" in Medieval Latin, which originally meant "being something in essence or effect, though not actually or in fact". In 1959, the meaning of "not physically existing but made to

---

[1] Anchors in this context include TV hosts and news anchors, network influencers, entertainment hosts, etc.

appear by software" at the computer level began to be applied, and then a virtual reality technology related to the computer, network, and telecommunication developed.[2] Therefore, "virtual" is the use of technology to imitate and simulate reality, not fiction.[3]

Philosophically speaking, virtuality is an ancient concept. In Plato's thought (1997) on the dichotomy of "the world of ideas" and "the sensible world," the latter is the "virtual" version of the former, and the omnipotent "craftsman" carries out this virtual. ("Timaeus", p.1235) With the development of technology and the positivist spirit, there is a growing tendency to believe that the physical world in which we live has a real character, while the cyber world becomes a virtual representation of the real world. The creators of the cyber world are human beings who are imitating the Creator and conducting new experiments in creation, which range from phenomena to things and people, to the entire world, with technologies such as AI simulations, virtual reality, cloning technologies, and the metaverse.

The technical and ideological sources of virtual technology can also be traced from the artistic level. From Plato onwards, art has been considered to be virtual and false. In Aristotle (1925), on the other hand, art (Tekhnē) is a combination of the necessary and the contingent, and art consists of an authentic course of reasoning, i.e., having authenticity. (1140A) Modern art breaks away from the traditional superficial reality of imitation art and emphasizes the logical reality of art itself. In the internet age, with the widespread use of intelligent technology in the arts, virtual simulation techniques have reverted to the pursuit of this apparent reality in video and dynamic images. VH is the most challenging virtual technology, and the earliest VH image was born from film and television art. For example, the Queer nationality in Forbidden Planet in 1956 is pure conscious existence, the "bionic man" appeared in Westworld in 1973, and the concept of "replicator" appeared in Blade Runner in 1982.

VH is not limited to human artistic imagination. In fact, this concept is also directly related to biological and medical technology. For example, the virtual human project (VHP) of the Oak Ridge National Laboratory in 1996 is an example of bioengineering technology. In China, a 3D reconstruction of a female VH with human physiological characteristics was started in 2003. Such programs integrate biomedicine, neuroscience, and AI. They enable the integration and reconstruction of human data into independent, visual, three-dimensional structures that allow manipulation to improve the level of information and modernization of the medical process. The ultimate goal is overcoming human physiological defects and achieving human enhancement.

The VH combining philosophy, art, and biotechnology is the ultimate goal of this concept and its entire content, as described by Martine Rothblatt (2014). The VH is "a replica of a particular person's mannerisms, personality, recollections, feelings, beliefs, attitudes, and values." (p.96) On this basis, the VA also has a functional or practical dimension, which needs to undertake part of the work of hosting and communicating. There are two opposing positions on VH technology. One view is an affirmative stance, arguing that the development of VHs is a practice that applies virtual reality technology to pursue human enhancement or the transcendence of the human species. The other view is a critical position based on the traditional philosophical exploration of truth and falsity, revealing the falseness embedded in VHs and VAs and the manipulation behind it.

So far, the areas of academic research on virtual humans (VHs) and virtual anchors (VAs) contain three primary points.

First is the study of humans in virtual worlds and the specific characteristics of virtual humans (VH), with a more interdisciplinary character. In the field of communication, Marshall McLuhan (1994), concerned with the close connection between artifacts and humans in the information age, proposed in *Understanding Media: The Extensions of Man* that the media is an extension of man, arguing that electronic technology is an extension of the entire human central nervous system. In Feminism, Donna J. Haraway (2015), in *Simians, Cyborgs, and Women: The Reinvention of Nature*, defines the hybrid

---

[2] More information on the origins and development of virtual is available at https://www.etymonline.com/.
[3] The word "virtual" in Chinese, which means fiction and simulation, is easily confused with unreal.

form of machine and organism as a product of the post-gender era. In terms of body theory, David Le Breton (2011) pointed out in the *Anthropologie du corps et modernité* that the body scorned by technology and the body favored by consumer society constitute a set of contradictions and that once the body reaches its conclusion, the pleasures of life will be lost. In sociology, Günther Anders (1992), in *Die Antiquiertheit des Menschen 2*, argues that cloning technology makes humans redundant and obsolete, and that this fact triggers a collective fear of humanity. These research results show the close connection between virtual humans and human beings and that most of the disciplines that have humans as their object of study can also be integrated and intersected with the study of virtual humans. This conclusion suggests that the virtual human, as a stage in the development of human technological civilization, has become a topic that is difficult to bypass in the study of post-human theory.

Second, the study of the VH as a posthuman subject is more comprehensive and forward-looking than the first type of research. Since the 21st century, academics have increasingly regarded virtual humans as a critical theme in post-humanity. And the fate of virtual humans has been closely linked to that of humanity, with academics and society filled with a variety of neutral, optimistic, or pessimistic voices. N. Katherine Hayles (1999), for example, in *How We Became Posthuman*, takes three waves of cybernetics as a developmental cue and clarifies that virtual technologies will drive the construction of highly heterogeneous cyberspace where the posthuman will transcend inert flesh and inhabit the invisible. Rothblatt (2014), in *The Virtually Human*, looks forward to the immortality of the human race and the healthier and happier life that will be achieved through digital technology, arguing that thought-cloning technology is democratizing. In *Homo Deus, a Brief History of Tomorrow*, Yuval Noah Harari (2017) sees dataism as a new kind of religion and modern man as a group of believers prostrated at the feet of data and, more pessimistically, sees post-humanity compressed into a digital flood of chips. It is clear from more recent results that, as technology advances, humanity has come to acquiesce to the imminent arrival of the virtual human. However, predictions about the post-human era remain diverse.

Third, virtual humans involve key technologies such as real-time interaction, emotion generation, and virtual reality, and scholars have conducted a series of specialized studies on them. Norman I. Badler (1997), in "Real-Time Virtual Humans," divides the simulation dimensions of the virtual human according to its design. He discusses appearance, movement, interaction control, and autonomous action and proposes that the structure of perceptual control behavior can have a driving effect. In the field of emotion generation, Rosalind W. Picard, a representative of connectionism, differs from Marvin Minsky, a representative of symbolism, who sees human emotion generation as a collection of independent processes by arguing that emotion generation should track various emotional signals in time and construct a mapping relationship between signals and external expressions. In his paper, M. Richard (2021) points out that the world is a non-negligible virtual reality dimension, and that real-time rendering immerses people in an unfolding space. Currently, the development of virtual human technology focuses on the "AI+VR" model, and the combination of intelligence and immersion will be the future trend of virtual humans.

Chinese scholars have paid much attention to virtual anchors (VAs). The research findings fall into three sections: a comparison of the advantages and disadvantages of virtual and human anchors, a review of the development of VAs, and the future prospects of virtual anchors. For example, the papers by Yang Chunhui (2019) and Zhao Guangyuan (2019) compare the two types of anchors from multiple dimensions and share similar views, i.e., human anchors are more humanistic, and VAs are more efficient, less error-prone, and cost-saving. Secondly, in terms of research overview, Wu Feng et al. (2021) collected detailed cases of domestic and international VAs over the past 20 years in their article. They covered aspects such as shape evolution, technical features, and the latest progress, discussing the impact of virtual anchors on humans (human anchors and viewers) and the industry landscape. Finally, in terms of outlook, most researchers believe that the future direction of VA is a blend of technicality and artistry.

In summary, there are two problems in the study of virtual humans and virtual anchors. First, researchers do not fully appreciate virtual anchors as representatives of virtual humans. Second, most research

content focuses on technology or the philosophy of technology, with little emphasis on ethics and aesthetics. This paper will attempt to remedy these problems by, first, elevating the situation of virtual anchors to the difficulties faced by the whole group of virtual humans; second, trying to fully connect the problems with philosophical thinking, especially by combining the ethics of technology and aesthetics, to find ways to make a breakthrough in the problems.

2. The History of the Virtual Anchor

As the representatives of Virtual humans (VHs), Virtual anchors (VAs) refer to the VA agents generated by applying virtual technology in the communication field and simulating human anchors. VA technology is one of the activities human beings employ to explore their own virtual reality, based on the simulation of reality and governed by the urge to create things. VAs are different from the virtual images in science fiction movies and games, which have the characteristics of real-time variability, interactivity, and sociality. VAs often need to broadcast and perform on certain occasions (virtual environment or stage environment), so the production technology of VAs is more complicated than that of film and television virtual images.

According to the technical characteristics of VAs, their development includes three stages. The first stage is the nascent stage. The virtual degree is not high, the virtual image is flat, and the first batch of VAs involving color key matting technology, 2D modeling technology, keyframe animation, and other technologies are produced. In April 2000, Ananova, an Internet news announcer launched by the British newspaper Associated Press Media, was regarded as the beginning of the global VA industry. In the same year, South Korea, Japan, and the United States also competed to launch their own VAs, namely Lusia, Yuki, and Vivian, who are active in three major fields: music, games, and life. In 2016, with the support of 3D modeling, motion capture technology, and voice technology, VA Kizuna (キズナアイ) from Japan triggered the latest round of development of VAs on the internet.

The second stage is the development stage. With the power of 3D modeling, motion and facial capture, speech synthesis, and other technologies, the VA ushered in the transformation from planar to three-dimensional. For example, Ami Yamato began to use the virtual image created by the facial capture device as the main character of the video in 2011, so the naturalness of the virtual image in the video has improved. (Guo, 2020)

The third stage is the promotion stage of intelligent technology. AI virtual image technology, natural semantic understanding technology, voice synthesis, virtual image driving, and other technologies have contributed to the intellectualization of VA. VA has achieved a more appropriate "cloning" of the human anchor based on imitation, and the details are accurate to lip movement, micro-expressions, voiceprint identification, and so on. For example, in 2020, South Korea developed the AI VA "Kim Joo Ha," which integrates AI, deep learning, convolutional neural network (CNN), and other technologies.

In China, VA technology is also developing very rapidly. In 2000, many VAs, such as Go Girl, Lili, Icy, and Bill Deng, were launched, and in the following years, Yi Meier, Jiang Linger, and Mei Mei launched successively. In 2004, Xiaolong, the first virtual TV program host in China, served as the host of CCTV-6 "Light and Shadow Weekly," followed by a period of technical settling. In 2018, Xinhua news agency launched Qiu Hao, the world's first fully simulated AI anchor based on a real-life anchor, sparking the trend of AI VAs. Unlike "European and American developed countries have cutting-edge AI technology, but they have 'aphasia' in the field of 'synthetic anchor'" (WU & Liu, 2021), China's VAs industry was slightly behind at the beginning of this century, but it has gradually taken the lead and has achieved the universal application of VA in CCTV and local TV stations in the first and middle stages of industry development. As the industry matures, it shows the family shaping of AI VAs by technology companies such as the Baidu department, iFlytek department, and Sogou department. This advance is because China's domestic TV stations and network TV are increasingly competitive. With the rapid development of the cultural industry, television requires more high-tech features to attract viewers. In addition, in

recent decades, China's Ministry of Culture (2017) has proposed to enhance the ability of cultural innovation" and encourage the development of high-tech industries, such as virtual reality, augmented reality, and AI.

VA technology is currently in the research and development stage, and there is still a large gap between VAs and human anchors. The mainstream view is that the AI anchors are still simulations of the real anchors, which need the expression and voice of the real anchors as the templates and cannot generate new logical reasoning, emotion, personality, and so on. (Ji, Cai, & Wang, 2020) At present, the intention of VAs generation is to improve the level of language and image recognition, 3D modeling, and virtual simulation technology. Therefore, AI synthetic anchor has a technology-oriented purpose rather than the fundamental purpose of cultural and artistic communication.

Professor Michael Woolridge of Oxford University also commented bluntly on Xinhua's AI news hosts. "It's very flat, very single pace, it's got no rhythm, pace or emphasis" (Wiederhold, 2019). Therefore, no matter the behavior performance, language rhythm, emotional expression, or personality embodiment of the VA, it cannot make people completely satisfied. This issue is not only a technical problem but also closely related to the design of the whole industry (similar design and function), human design industry (flow line character design), consumption symbols (using virtual gimmicks to seek economic benefits), and other elements. The research and development of virtual technology ultimately serve the development of cultural industry and capital.

3. Simulation levels and technologies for virtual anchors

The fundamental rationale for designing virtual anchors (VAs) is to use human characteristics as the objects of simulation. However, the successful portrayal of a character in theatre requires not only looks, styling, or movement design but also a distinctive personality. VAs are supposed to be a unity that echoes both inside and outside. Badler (1997) divides the simulation of virtual humans (VHs) into five dimensions, appearance (physiological characteristics simulation), function (physical and mental activity simulation), time (interaction and feedback simulation), individuality (gender, age, culture, and other characteristics simulation), and autonomy (control of self and the outside world and prediction simulation). Sun Shouqian and other scholars (Sun, Q. Wu, & J. F. Wu, 2010) elaborated on the key points and specific requirements of VHs simulation in the field of digital media from the technical perspective. In addition to external requirements, such as geometry and behavior, it also includes the deeper human essence represented by perception, emotion, and social characteristics. We divide the simulation of VAs into internal and external levels. External simulation refers to the directly visible identities, appearance, actions, etc., while internal simulation refers to the indirect intelligence, emotions, personality, and possible intentional activities. Combining Badler's five dimensions with internal and external simulations, we get the primary parts of the VAs that need to be simulated. We then talk about the simulation mechanism and technical requirements in the context of a specific VA, Xiaoyang, the first digital anchor of Hunan Satellite TV in China.

The first dimension, appearance, is the simulation of external form features. Such as designing a beautiful figure, glossy skin, and exquisite hair for "Xiaoyang.". The technologies used include image design and 3D modeling. Modeling technology is the basis of VAs shaping. In the character modeling phase, popular methods include geometric feature description, two-dimensional photo reconstruction method, parametric reconstruction method, modeling methods represented by Poser, 3D Max, and other software, and intelligent modeling methods using intelligent modeling systems to replace human preprocessed data. (DU, 2019) Through behavior modeling, the VA can realize simple movements, such as body movements and pronounced lip movements, which include the parametric keyframe method, inverse kinematics method, kinetic method, process method, and motion capture method using sensors. (Wang, 2010) Behavior modeling not only improves the fluency of the VA's large-scale body movements but also makes the subtle lip movement and facial muscle movement more fluid, and the comprehensibility of the entire VA will improve significantly.

The second dimension is function. The functions of VAs are news broadcasting, hosting, participating in entertainment activities, etc. Therefore, it is necessary to design some professional hosting skills for VA, such as announcing, dancing, singing, etc. The technologies involved are motion capture, micro expression control, speech and language processing, and so on. In particular, AI technology has improved natural language processing (NLP), enabling instant, smooth, and automatic communication between human viewers and virtual hosts. At present, the difficulties of VAs in natural language understanding include logical relations, grammatical relations, and semantic issues. In the final analysis, it is still difficult for a machine to deal with the fuzzy fields in human natural language. Aiming at Chinese machine language, Huang Zengyang of the Institute of Acoustics, Chinese Academy of Sciences, put forward the Hierarchy Conceptual Network (HNC) theory. It is "based on semantic expression, simulates the cognitive behavior of the brain in understanding natural language" (Gu, Wang, & Guo, 2015, p.105). It establishes a local associative context and global associative context, not only for the representation of linguistic units (vocabulary) but also for the comprehensive analysis of statements or entire chapters, thus improving the VA's ability to understand natural language.

The third dimension is the time dimension, which includes internal and external dimensions. Externally, VAs can cooperate with the human actors to complete the animation dubbing and real-time interaction. Internally, the continuous refinement and optimization of VAs improves the technology more and more on the one hand and makes people feel that VAs are growing with time on the other. As one of the fundamental characteristics of virtual reality, interactivity aims to establish the basis for two-way interaction between humanity and the virtual environment built by computers, and interactive devices are the core link of this process. Multi-channel interaction improves the accuracy and fluency of real-time interaction. This approach integrates voice, gestures, eyes, and other information exchange to advance real-time human-machine interaction to the level of real-time human-to-human communication. (Zhang, Dai, & Peng, 2016)

The fourth dimension is individuality. Externally, VAs have unique identities and role settings. For example, Xiaoyang's identity gender is female, her age is 20 years old, and so on. From the internal point of view, VAs need human-like emotions, personalities, and even cultural settings. The character of Xiaoyang is a youthful and lively host with elegance and knowledge. Hunan Radio and Television Laboratory will also cultivate her personality in the subsequent program application scenarios, allowing her to evolve iteratively and helping her to grow into a lovely and lovable girl next door. (Yang, 2021) Emotion setting and interaction are the key technologies in individuality, and there are currently two approaches to AI emotion generation, top-down (symbolism) and bottom-up (connectionism). Symbolism states that pre-programming can simulate the human emotion generation mechanism. Minsky (2006) regards human emotion generation as a collection of a series of independent processes, where a large number of "IF-DO-THEN" instructions in each level and branch form an efficient system in a brief time, and human beings can use credit empowerment, parallel analogy, and other skills to achieve conversion between different ways of thinking. When the machine's learning, storage, computing, and other capabilities form a system similar to that of a human (emotional machine), it cracks the emotional activity code. Sloman constructed a mechanism composed of an initial engine, new motivation generator, interrupt filter, sensory detector, and other parts to analyze the role of motivational triggers in supporting affective states. (Boden, 1990) The biggest problem with the predetermination procedure is that emotion is a rather complex state of mind, and human beings have not been able to understand its mechanism yet. Linking emotion to motivation and purpose alone would be imprecise, and not all emotions are motivated. The method adopted by connectionism is "affective computing," which collects meaningful human emotional response data. It analyzes the mapping relationship between emotions and performances such as language, gestures, and expressions through a model based on cognitive evaluation theory and tracks and interacts with various emotional signals in a timely manner. Emotion recognition is the core part, and Picard (2000) believes that "Outsiders only have access to observable functions of the affective state—expressions, behaviors, and so forth. Given reliable observations of these functions, then the underlying states may be inferred." (p.167) She further pointed out that the current emotion recognition model belongs to the low-medium level transformation. That is, recognition is achieved through human body changes and neurophysiological information, but this

method cannot yet perfectly identify the latent emotional state under various situations. Inspired by OCC[4] and Roseman[5], Higher-level signals such as context-specific, event-based reasoning are mechanisms for the generation of rule-appropriate cognitive emotions. That is to say, emotion recognition is truly successful only when low-level perceptual signals combine with high-level cognitive signals.

The fifth dimension is autonomy. The autonomy of the external level is reflected in the automation of action. The current AI VA can be automated with the support of big data and deep learning technology and can respond and act automatically in unsupervised or semi-supervised situations. The inherent autonomy requires the VA to have a certain level of intentionality or consciousness. Philosophers, technologists, and scientists have debated whether AI can produce intentionality or consciousness. The philosopher Searle (1980) tried to refute the effectiveness of the Turing Test with his thought experiment "Chinese Room." He believed that the robot in the house could pass the Turing Test, but it did not understand the meaning of Chinese because it had no intentionality and consciousness. From the standpoint of "biological naturalism," Searle (2004) believes that "intentionality" originates from the physiological level and develops into psychological states and verbal behaviors in humans. Unlike carbon-based life, machines cannot develop intentionality. Searle's view has triggered much debate. Some scholars, such as Boden (1990), argue that intentionality is unknowable. He believes that we do not know what and how intentionality arises, and we cannot conclude that machines cannot produce intentionality. Kurzweil (2014), an expert in AI, believes that if we accept that following an algorithm is inconsistent with true understanding and consistency, then we must also agree that the human brain does not exhibit these qualities either. The philosopher Dennett believes that we can adopt a more generalized "intentional stance" in which inanimate entities, such as artifacts, can be seen as "rational agents, which control their 'choice' of 'action' by a 'consideration' of their 'beliefs' and 'desires'." (1996, p.26) Some scientists, however, believe that computers cannot simulate human consciousness because of its quantum nature. "There must be an essential non-algorithmic ingredient in the action of consciousness." (Penrose, 1989, p.407) Therefore, autonomy is only weak autonomy at present and has not reached the stage of strong autonomy.

Five types of simulation dimensions & internal and external simulation		Appearance	Function	Time	Individuality	Autonomy
External simulation	Contents	Physiological characteristics	Physical Actions	Real-time interaction of voice and movement	Identity Settings	Automation, predictive simulation
	Methods	Computer graphics, 3D modeling, image design, parametric keyframing methods, etc.	Motion capture, process animation, micro-expression control, motion editing and synthesis, etc.	Speech recognition, speech synthesis, speech and behavior interaction technology, virtual reality technology	Virtual world database, character setup, image processing, etc.	Big data, deep learning, artificial intelligence technology, etc.

---

[4] OCC is a cognitive emotion evaluation model jointly proposed by Ortony, Clore, and Collins. The emotion synthesized by the model is the result of the synthesis of events, objects, agents, and other situations. More is available in *The Philosophy of AI* by Boden.

[5] The cognitive-affective evaluation theory and model proposed by Roseman and colleagues show the interaction among a few evaluations (accidents, motivational states, situational states, etc.). More is available in *The Philosophy of AI* by Boden.

				Personality development, character growth	Personality, Language, Emotion	Intentionality, sense of autonomy
Internal simulation	Contents	None	None			
	Methods	None	None	IP content production, distribution and marketing	Affective computing, natural language processing, multi-agent systems, etc.	Quantum physics, multidisciplinary synthesis, etc.

Chart 1: the simulation technologies of Virtual anchors

To sum up, unlike the mechanical reproduction in the industrial age described by Benjamin, AI VA has entered the stage of an intelligent simulation. Mechanical reproduction is a kind of apparent replication, while an intelligent simulation is a deep, highly simulated, bionic simulation of human intelligent behavior and action mechanism. Even with the support of some AI technologies such as GAN, VAs can make certain automatic behaviors unsupervised, such as real-time interaction, reasoning, and creative activities. Thus, VAs can replace human anchors to some extent. This intelligent simulation makes it possible for people to not only have their physical work replaced by machines, but also for their mental work to be replaced by other intelligent programs. In this case, virtual people and a virtual world are really possible. This high-simulation and anthropomorphic simulation will bring new human experiences and a series of social issues, the most important of which is the problem of truth and trust.

4. Authenticity and falsehood of virtual anchors

4.1. The disintegration of authenticity: Trust Crisis

In the era of intelligent technology, a dialectic of super-real and super-fake has emerged, where AI can simulate human behavior, and virtual humans (VHs) can imitate a human image super-realistically. These technologies also allow super-fake results. Therefore, in the face of intelligent technological arts and virtual anchors (VAs), one must first distinguish between true and false. All aesthetic experience and judgment occur after the determination of true and false. It will trigger a crisis of trust in human beings.

The current AI VAs fall into two categories: cartoon and simulation. The former is an abstract representation of a human body, such as the texture, luster, and pores of skin here presented as a uniform color and smoothness within a certain shape. The latter category of simulated AI VAs includes imitations of humans in detail, which will trigger more uneasy emotions in viewers than cartoon images. This result is exactly what Japanese scholar Masahiro Mori pointed out in the 1970s as "the Uncanny Valley" phenomenon. "When the humanoid level of robots, dolls, or prosthetics reaches a certain high level, one wrong step will lead to a sharp turn and fall into the abyss of 'the Uncanny Valley.'" (Jiang, 2020) Why do anthropomorphic artifacts cause a sense of confusion in humans? Scholars pointed out that "the emergence of the 'the Uncanny Valley' phenomenon is related to the alienation risk and cognitive confusion that people encounter in front of highly simulated robots." (Cheng & Jiang, 2018) Specifically, the high-simulation AI creates a powerful conflict between the ordinary appearance and the opaque nature, and the more normal the surface, the more mysterious the interior. The unknowable technological connotations inspire human defensiveness, and the psychological basis of the audience's dealings with their own kind (human beings) nullifies at this point, leading to an inability to empathize and a sense of fear. (Tao, 2018) The phenomenon of "the Uncanny Valley" is an example of human distrust of the objective world. First, AI VAs, empowered by big data, are no longer passive parties; they control the information leadership. The core of the machine's computing, the algorithm, is out of reach for the vast majority of people. The opaqueness of the mechanism makes it difficult to accept information from the

AI openly while at the same time being concerned about whether one might suffer an intrusion. Second, ethics, as a product of development of civilization, stems from the idea and action of human beings to seek greater good for themselves, while the AI VAs have not obtained a clear moral status. When people have a relationship with them, the ambiguous subject of responsibility makes them unable to accurately estimate the risks they may take. (Li, 2018, p.212)
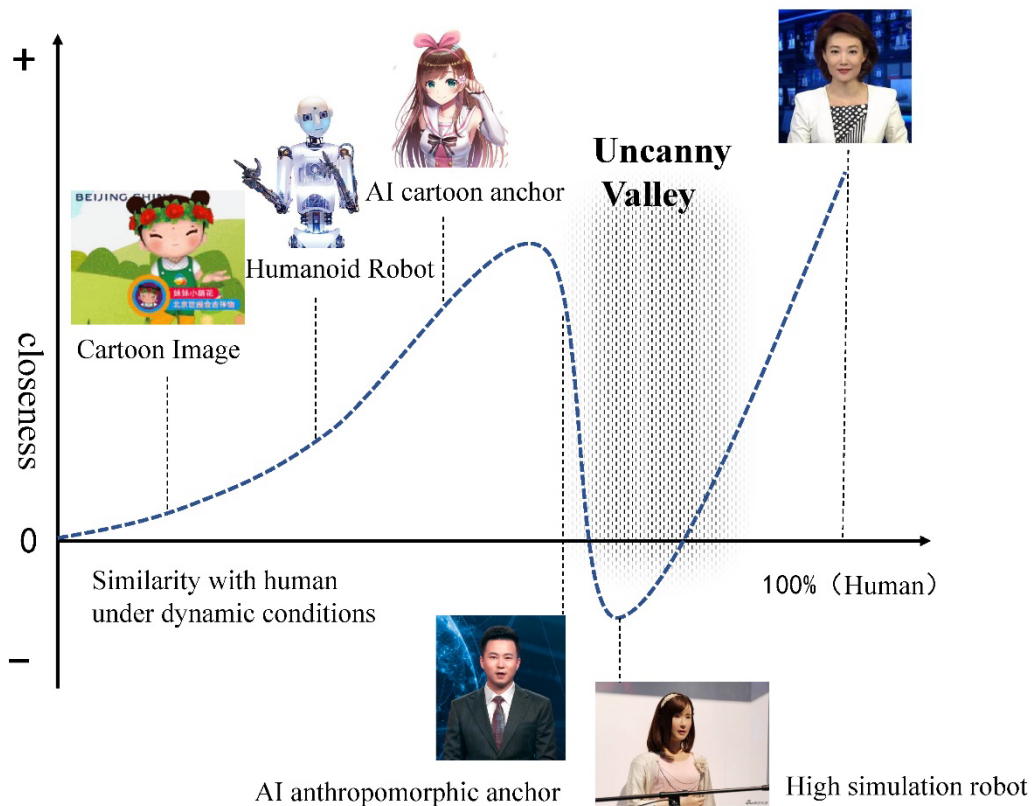
Figure 1: the Uncanny Valley (including VAs)[6]

These fears, doubts, and discomfort are rooted in people's distrust of automated artifacts, which can be further attributed to the consequences of the mediation of human traditional physical and mental models by data or artifacts. As for the problem of body and mind, in ancient Greece, Plato regarded the soul as immortal and invisible, while the body was considered visible and perishable. Building on previous discussions of the mind-body problem, Descartes systematically proposed a theory of mind-body duality, defining the mind as a thinking entity that possesses the property of "thinking" and can exist independently of the physical body, using the method of universal doubt. In Adorno's Negative Dialectics, an intermediary has become a critical factor in bridging the opposition between subject and object. Adorno (1966) suggests that mediation of the object means that it may not be hypostatized statically or dogmatically but can only be recognized in its interweaving with subjectivity. Mediation of the subject would literally be nothing without the moment of objectivity. (p.184-185) Revisiting the mind-body opposition from the perspective of mutual mediation, the body and mind are dialectical and mutually mediated. Thus, the mind must perceive the material world through the body (intermediary). The mind simultaneously emerges from the body, which needs to know the material world through the mind. However, in the virtual reality represented by the VA, the direct relationship between body and mind is altered, as mind and body can not only be separated but also digitized, virtualized, and artificialized. In this way, the virtual world represented by data breaks the old model of mind and body as intermediaries, and the new model becomes "body-data-mind." As the layers of mediation increase, the classic human cognition model has also faced challenges. In Kant's original mind-body model

---

[6] This image is adapted from the Uncanny Valley image （designed by Yunyu Dang & Feng Tao）

(1998), humans can use intuition to obtain perceptual knowledge of things. The combination of intuitive material (phenomenon) and intuitive form (time and space) constitute the starting point of human cognition, and the innate form of intuition also determines that human beings rely on perceptual intuition. However, under the intermediary effect of manipulated empirical data, the previous objects are replaced with data-wrapped illusions, which overlap and constitute a generally virtualized world. And trust is difficult to win in data-based non-intuition. The problem with this trust is not only the untruthfulness of the data but the power of human control behind the data. This fear of human control of power came to the fore when people gave the power of creation from God or nature to a few people. The critical theory of the Frankfurt School holds that the management class controls others and nature through technology and that rationality becomes instrumentalized and homogenized. Adorno (2003) pointed out that Cultural Industry operates the scheme as the first service to the customer. A secret mechanism should work in the soul, which already prepares the direct data to fit into the system of pure reason. This control can also be exercised more deeply and tightly through data to mediate the human body and mind, controlling people's perceptions and aesthetics.

Worse still, the collapse of trust does not only exist between humans and virtual humans but also among humans, which can cause a crisis of trust. Nourbakhsh (2013) pointed out that "as we dehumanize our relationships with robots using a rather broad brush, so we may incidentally dehumanize our relationships with people." (p.63) This quote is another reminder that human beings must rebuild trust with AI VAs. The authors believe that first, we need to recognize the limited nature of human rationality, limit the development of certain intelligent technologies and draw boundaries for intelligent subjects, limiting them to a manageable range as a prerequisite to re-establishing trust. Second, the identification of artifacts such as VHs is necessary, which requires that intelligence agents and VHs demonstrate their identity in advance, without going so far as to deceive humans and thus create a trust panic. Thirdly, AI VA is regarded as a special moral subject, and its behavior is endowed with morality and responsibility from the perspective of responsibility ethics, which is a way to establish a trust contract in reality. Kant has even limited the scope of discussion of moral relationships to rational parties, excluding infants, the mentally ill, and animals. However, with the rapid development and widespread application of intelligent technology, the scope of discussion on moral relationships is also expanding. Allen, Varner, and Zinser (2000) pointed out that "as AI moves ever closer to the goal of producing fully autonomous agents, the question of how to design and implement an artificial moral agent (AMA) becomes increasingly pressing."(p.251) They stated that the ultimate goal of designing an AMA is to establish a morally commendable agent, and the key lies in how to endow it with enough intelligence to enable AMA to assess its impact on emotional biology and to use these assessments to make appropriate choices.

4.2. The reproduction of authenticity: Reconstituting Aura

The super virtuality of virtual technology lies in its attempt to reconstruct the aura, that is, to reconstruct a unique existence in the here and now. Benjamin (1991) has suggested that a copy is short-lived and reproducible, while authentic art has a unique and enduring aura. Aura embodies the temporal and spatial presence of a work of art and the extended history of its experience. The authenticity of artworks has gradually been lost since the emergence of photography, the first revolutionary reproduction method, and the countless copies of photographs that represent no historical or locational significance have led to a shift from cult value to exhibition value of the artwork. The copied artwork has gradually become the reproducibility of artworks designed for reproducibility. When the temporal-spatiality (here and now) of art is destroyed, the aura also collapses.

When mechanical reproduction technology develops into more advanced AI simulation technology, the damaged aura may be reconstructed artificially, that is, the reconstruction of space-time and the original. Groys (2008) pointed out that "the modern age is constantly substituting the artificial, the technically produced, and the simulated for real, or (what amounts to the same thing) the reproducible for the unique." (p.63) The network integrates the production space and display space of artworks into one, overturning the traditional sense of art production. Nowadays, art production is not limited to the

generation of a particular piece of artwork but is a practical collection composed of design, research, development, publicity, generation, feedback, and other links. Works of art circulate in one closed context after another. "In circulating through variable contexts, a copy becomes a series of different origins." (2010, p.67) And when technologies such as AI and virtual simulation reproduce objects more comprehensively and realistically and can simulate them mechanically, the copies themselves have uniqueness and originality. Then, AI can generate original works that are different from human works.

According to Benjamin, the essence of the aura lies in the presence of a unique space-time. Virtual technology must recreate a virtual space-time if it wants to recreate an aura. Virtual technology can make the senses feel spatial appearance through visual creation and multi-sensory experience. As mentioned earlier, multiple dimensions of virtual anchor (VA), such as appearance, function (behavior), and individuality, can all create a sense of space. The realistic appearance and fluid movements of the VA and the naked-eye 3D of virtual reality technology can give the impression that the VA is actually in space. Real-time interaction can further instill a sense of time, and the temporality in virtual anchor technology arises through real-time interaction. This creation of space-time makes the human senses seem to feel a real sense of time and space. According to Kant's viewpoint, if space-time is all subjective to human perception, then it is possible to generate a sense of space-time by using intelligent technology to simulate space-time representations and stimulate human temporal intuition. AI VAs, such as the People's Daily VA "Guo Guo," combine AI image recognition, speech recognition, emotional computing, and virtual reality technology. They recreate a human-like anchor image from vision, auditory, emotion, and other aspects. The now-popular NFT art, on the other hand, uses encryption technology to set identity information for some network images, freeing them from arbitrary reproduction and possessing the value of the originals. And metaverse and virtual reality technologies, which make immersive experiences more realistic, provide the technical basis for recreating the aura.

Welsch (1997) points out that cyberspace is the center of the integration of virtual and reality, "The presence in front of and facing the picture transforms itself into presence within the picture into what is called telepresence." (p.181) AI VAs, as representatives of VH in the post-human context, carry out communication activities in various media in digital form and also show the possibilities of a metaverse for their audiences.

Of course, virtual technology does not recreate a real space-time or aura but rather an illusion of the senses. It forces us to think again about the question of reality and illusion. If we accept the Platonic view that reality is only an illusion, then we might not mind this double illusion. On the other hand, the aura becomes the dialectical object of this illusion of space-time, which was originally a real existence in real space-time but may now become the real evidence of a false space-time. It is conceivable that the aura can again assume an evaluative role in a dynamic, open, and transcendental post-human context. It would not only mean its sublimation but also give birth to a new cult of the virtual object. This cult is not only driven by the human impulse to make things and the artistic imagination but also by capital. With the help of virtual technology, capital can commercialize temporal and spatial representations, aura, and VHs, which form a new cultural, industrial commodity.

4.3. Intelligent fake: New Cultural Industry

The reconstruction of reality by virtual anchors (VAs) and virtual humans (VHs) is a double falsehood by making falsehood more realistic. This double falsehood is not just the lack of utility of art that Plato speaks of in the Republic (1997), but it is based on a more immediate purpose of use, the use of literature and art for the production of the practical world. This falsity is also an intelligent, automated fake based on a high degree of technological and rational development. That is, the entire industrial production process becomes an automated process for the production of falsity, entering the intelligent industrial age of a false cultural and living world.

Benjamin (2005) believes that literature and art were kinds of human production activity and regards the artistic process in the age of mechanical reproduction as a production process in which the authors of

literature and art were, themselves, producers. Therefore, he attached foremost importance to the impact of technological innovation on art. For example, mechanical reproduction technology makes art lose its religious value, leaving only its exhibition value, and technology liberated art from magic and religion. Adorno (2016) viewed the industrial production of art from a negative perspective, which he called "cultural industry." He believes art itself has dialectics of necessity and contingency, expression and inexpressibility, and authentic art should remain incommunicable and unrepeatable, opposing the cultural industry that turns art into a commodity.

In the era of intelligence technologies, the application of virtual reality, AI, big data, and other technologies in art brings art into a new cultural industry. The characteristic of this cultural industry is that intelligent technology integrates art, where art must abandon its dialectic and negativity in favor of scientific accuracy and inevitability. The result is that the authenticity that art possesses through its distance from reality no longer exists; the falsity of society then inevitably results in the falsity of art, which completely becomes a cultural industry. Adorno argues that high art can also rebel against the illusion of society by breaking its own illusion. But if art only seeks certainty and loses the dialectic of negation, then art loses the opportunity to rebel against society.

VAs are precisely a process of developing and applying cultural and industrial products. Chasing commercial interests is the fundamental goal of technology developers and capitalists, so the design benchmark of AI VAs is public aesthetics, and they meet the public's crude pursuit of culture with their fragmented and convenient characteristics.

The cultural and industrial characteristics of VAs are reflected in the following facts. First, VAs do not require humans to think or even prevent them from thinking, so the audience stays in its regular broadcasts or performances. Second, the focus is on shaping the appearance and decoration of the virtual anchors, through which the roles are differentiated, making each VA look similar. Finally, in the constant pursuit of maximizing the number of products and exposure, the media or companies use the VAs as publicity stunts to attract attention. The image and design style of the VAs are remarkably similar, and the details have only slight changes, which shows that despite the advanced technology, developers have to consider the public's acceptance of the product. Only successful precedents can move capital groups, and they would rather move forward along the original path to ensure profits. Therefore, the emergence of VAs is the inevitable result of a new cultural industry, a product of industrial production that has encroached on human life to shape the human body, personality, and mind. There is nothing good or bad about the technology of VAs per se, but the capital control behind it is worthy of caution.

In addition to the digital form of the virtual anchor, there is a physical, mechanized form of anchor to consider, namely the highly simulated robot anchor. If virtual humans are still only imitating humans in terms of data, robots proceed to imitate humans in all aspects of material, behavior, and physical appearance. Regarding technical means, with the synthesis of AI, virtual technology, and automation technology, robot anchors are likely to be realized, and further development into living robots is not far off. The combined influence of robots and virtual humans could fully replace humans, from their bodies to their intelligence. More importantly, humans may become accustomed to being replaced gradually in this way, and a symbiotic life between humans and machines becomes acceptable. E. Musk announced the imminent mass production of a new service-oriented humanoid robot (Tesla Bot) on its official website. The site describes the goal of the Tesla Bot as, "Develop the next generation of automation, including a general purpose, bi-pedal, humanoid robot capable of performing tasks that are unsafe, repetitive or boring." [7] Industrialization aims to make maximum profit. Initially, industrially produced humanoid robots replaced people in terms of labor, while the need to control the risk of human life enabled the trend for robots eventually to replace people physically and materially. Highly simulated humanoid robots will no longer enter human society as heterogeneous objects (the human other) but will become members of the human race and build a society together with humans. This development would

---

[7] More information on the "Tesla Bot" is available at https://www.tesla.com/AI.

have been an ideal form of human-machine harmony, but it is possible that, with the loss of boundaries, machines are slowly replacing humans, and we need to be wary of this possibility.

Conclusion

In the context of post-humanist thought, the body as a reality can be remodeled and replaced. Moravec (1988) has long predicted that the human mind can get rid of the bondage of the body to mortals - the body dies, and the mind dies. Just like computer data, it can be read out in machine memory, transferred to computers that are completely different in physical form, and fully recovered without losing anything. Nowadays, virtual reality technology is turning predictions gradually into reality. Human anchors are replaced by VAs, human bodies are digitized, and people become VHs. Our world is also being digitized and developing into a Metaverse (Dionisio, Burns, & Gilbert, 2013). Humans are impatiently turning all reality into digital artifacts, and behind the technological drive is the drive of capital globalization. And the COVID-19 pandemic provides the perfect excuse for this digitization of the world and human beings. The natural world is no longer fit for human existence, so we should create virtual worlds as soon as possible. But is a virtual world really an option or even the final destination for humans? Compared to the natural, real world, the digital world is more uncontrollable and unidimensional, a flat world detached from the historical-social dimension, lacking the growth and diversity of the living world, and, perhaps, a small virus would be enough to destroy it.

References

Adorno, T. W. (1966). *Negative Dialektik*. Frankfurt am Main: Suhrkamp.

Adorno, T. W. (2016). Band.7 : Ästhetische Theorie. In R. Tiedemann(Eds.) *Gesammelte Schriften*. Frankfurt am Main: Suhrkamp.

Adorno, T. W.& Horkheimer, M. (2003). *Band 3: Dialektik der Aufklärung.* In R. Tiedemann(Eds.) *Gesammelte Schriften*. Frankfurt am Main: Suhrkamp.

Allen, C., G.Varner, & J. Zinser, (2020). Prolegomena to any future artificial moral agent. Machine Ethics and Robot Ethics, 53-63.

Aristotle. (1925). Ethica Nicomachea. In W. D. Ross., & M.A. (Eds.), *The works of Aristotle* (Vol. 9, pp. 1094A-1181B). Oxford: Clarendon Press.

Badler, N. (1997). Real-time virtual humans. Proceedings the Fifth Pacific Conference on Computer Graphics and Applications.

Benjamin, W. (1991). Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit. In R. Tiedemann., & H. Schweppenhäuser (Eds.), *Walter Benjamin Gesammelte Schriften* (Band I.1). Frankfurt am Main, HD: Suhrkamp.

Benjamin, W. (2005). In M. W. Jennings., & H. Eiland (Eds.). *Walter Benjamin: Selected writings* (G. Smith, Trans.). Cambridge, MA: Belknap Press.

Boden, M. A. (1990). *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press.

China's Ministry of Culture. (2017). *The development plan for the cultural industry during the 13th Five Year Plan period.* (Chinese)

Cheng,L. &H. Jiang. (2018). Robots falling into the "uncanny valley " -- starting from the theories of Ernst Jentsch and Masahiro Mori, *China Book Review*,2018(07),35-43. (Chinese)

Dennett, D. C. (1996). *Kinds of Minds: Toward an Understanding of Consciousness*. New York: Basic Books.

Dionisio, J. D.N., Burns III, W. G., & Gilbert, R. (2013). 3D virtual worlds and the metaverse. ACM Computing Surveys, 45(3), 1-38.

Du.J.J., (2019). Research on Modeling and synthesis technology of 3D VH. *Electronic world,2019(06),5-6.* (Chinese)Groys, B. (2010). *Going Public*. J. Aranda., B. K. Wood., & A. Vidokle (Eds.).Berlin: Sternberg Press.

Gu, X. J., Z. L. Wang., & Y. C. Guo. (2015). *Emotional VH Technology in Human-computer Interaction*. Beijing: China Machine Press. (Chinese)

Guo. J. N. (2020). Research on Subculture style of network VA (Master dissertation). (Chinese)

Ji,W. C., R. M. Cai., & Y. H. Wang. (2020). When AI synthetic anchor "grabs" the microphone of media people -- new development of AI and media under big data. Thesis Collection of 2020 Remix Education Conference. (Chinese)

Kant, I. (1998). *Kritik der Reinen Vernunft*. Hamburg: Meiner.

Kurzweil, R. (2013). *How to Create a Mind: The Secret of Human Thought Revealed*. London: Duckworth.Li, L. (2018). *AI and Big Data Ethics.* Beijing: Science Press. (Chinese)

Minsky, M. (2006). *The Emotion Machine.* New York, NY: Simon & Schuster.

Moravec, H. (1988). *The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press.

Mori, M. & H. Jiang. (2020). The Uncanny Valley. *New Perspectives on World Literature*,2020(05):85-92. (Chinese)

Nourbakhsh, I. R. (2013). *Robot Futures*. Cambridge, MA: MIT Press.

Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.

Picard, R. W. (2000). *Affective computing*. Cambridge, MA: MIT Press.

Plato. (1997). Republic (G. M. A. Grube, Trans.). In D. S. Hutchinson., & J. M. Cooper. (Eds.), *Plato Complete Works* (pp. 971-1224). Indianapolis, IN: Hackett.

Plato. (1997). Timaeus (J. Z. Donald, Trans.). In D. S. Hutchinson., & J. M. Cooper. (Eds.), *Plato Complete Works.* (pp. 1224-1292). Indianapolis, IN: Hackett.

Rothblatt, M.A. (2014). *Virtually human: The promise - and the peril - of Digital Immortality*. New York, NY: St Martin's Press.

Sun,S. Q., Q. Wu., & J. F. Wu. (2010). *VH Technology and Application*. Beijing: Higher Education Press. (Chinese)

Searle, J. R. (2004). *Mind: A Brief Introduction*. Oxford: Oxford University Press.

Tao,F. (2018). Creativity and Emotion: a Preliminary Study of AI Aesthetics. *China Book Review,2018(07),26-34.* (Chinese)

Welsch, W. (1997). *Undoing Aesthetics* (A. Inkpin, Trans.). London: Sage Publications.

Wang, Z. Q. (2000). A review of VH synthesis. *Graduate School of Chinese Academy of Sciences,2000(02),89-98.* (Chinese)

Wiederhold, B. K. (2019). Animated news anchors: Where to next? *Cyberpsychology, Behavior, and Social Networking*, 22(11), 675-676.

WU, F. & Z. X. Zhao. *(2021).* History, application status and industry influence of AI anchor. *Southwest University for Nationalities (Humanities and Social Sciences Edition),42(05),174-183.* (Chinese)

Yang, C. H. (2019). The advantages and disadvantages of "AI synthetic anchor" function in the context of artificial intelligence. *Media Forum (10),42-43.* (Chinese)

Yang, Y. (2021). Research Report on National Radio and television media integration in 2021. *China Radio, Film and Television,2021(23),32-37.* (Chinese)

Zhang, F. J., G.Z. Dai., &X. L. Peng. (2016). Overview of human-computer interaction in virtual reality. *Chinese Science: Information Science, 46 (12),1711-1736.* (Chinese)

Zhao. G. Y., &L. Tian. (2019). The game of technology and art: Reconfiguration of the anchors' function in the context of artificial intelligence. *Contemporary TV(10),93-96.* (Chinese)

ON HOPE RESISTANCE

SOBRE A RESISTÊNCIA DA ESPERANÇA

SOBRE LA RESISTENCIA DE LA ESPERANZA

Caterina Del Sordo
(University of the Basque Country UPV/EHU, Spain)
caterina.delsordo@ehu.eus

ABSTRACT

In this paper, I analyse our concept of resistance. I tell the difference between hope resistance and resistance *tout court* - the first figures out as a more basic phenomenon than the other. I specify the concept of hope resistance drawing upon the ontological and epistemological artillery of Bloch's principle of hope. Accordingly, hope resistance is illustrated and explained through the concepts of possible futures, utopian function, anticipation and *novum* that nowadays are crucial to the new philosophical field of future studies. I bring examples of hope resistance drawing upon the history of feminism and the fictitious world of science fiction and crime series, particularly *Black Mirror*, *Utopia*, and *Money Heist*.

Keywords: hope. resistance. future studies. anticipation. utopic function.

RESUMO

Neste artigo, analiso nosso conceito de resistência. Digo a diferença entre resistência à esperança e resistência *tout court*. O primeiro, figura como um fenômeno mais básico que o outro. Especifico o conceito de resistência à esperança com base na artilharia ontológica e epistemológica do princípio da esperança de Bloch. Assim, a resistência à esperança é ilustrada e explicada através dos conceitos de futuros possíveis, função utópica, antecipação e *novum* que hoje são cruciais para o novo campo filosófico dos estudos do futuro. Trago exemplos de resistência à esperança com base na história do feminismo e no mundo fictício da ficção científica e das séries de crimes, em particular *Black Mirror, Utopia* e *Money Heist*.

Palavras-chave: esperança. resistência. estudos futuros. antecipação. função utópica.

RESUMEN

En este trabajo analizo nuestro concepto de resistencia. Presento la diferencia entre la resistencia de la esperanza y la resistencia *tout court*. La primera aparece como un fenómeno más básico que el otro. Especifico el concepto de resistencia de la esperanza apoyándome en la artillería ontológica y epistemológica del principio de esperanza de Bloch. Así, la resistencia de la esperanza se ilustra y explica a través de los conceptos de futuros posibles, función utópica, anticipación y *novum* que hoy en día son cruciales para el nuevo campo filosófico de los estudios de futuro. Traigo ejemplos de resistencia de la esperanza desde la

historia del feminismo y el mundo ficticio de las series de ciencia ficción y acción, en particular *Black Mirror*, *Utopia* y *Money Heist*.

Palabras claves: esperanza. resistencia. estudios de futuro. anticipación. función utópica.

1. Introduction

This paper aims to clarify the content of our concept of *resistance*. To achieve this target, the main theoretical inspiration will be taken from the philosophy of Ernst Bloch and exemplifications will be drawn from Latin-American feminism, science-fiction and crime TV series[1]. If only one starts to think of it, resistance reveals a slippery and elusive nature. To be sure, it reveals a very pervasive concept that can be said in many ways just as well as being. Indeed, it occurs in electromagnetism as the measure of an object's opposition to a certain flow of electric current; in human history as the physical and moral endurance in warfare and contestations; in political philosophy and jurisprudence as the right to contrast illegitimate authority; in physical exercises as the body workout against weights, bands or forces; and so forth (Del Sordo 2021).

Recently, resistance has also joined the shortlist of concepts of theoretical philosophy where resistance has been numbered among the attributes of *phrónesis* or practical wisdom. In this respect, I think the book José Medina's *Epistemology of Resistance* (2013) has been pivotal. It develops, indeed, virtue-based ethics of knowing that evolves around the idea of resistance considered as a kind of dianoetic virtue, i.e., as an admirable character trait that any epistemic agent should nurture. In this context, an epistemic agent is said to be resistant when, under conditions of oppression or disadvantage, he or she manages to gain momentum and put into doubt eventually interiorized feelings, or self-convictions, of inferiority.

Resistance thus emerges as a counteractive force that reacts against material, physical, political, epistemic or intellectual forms of oppression. I will refer to this aspect of the concept as the resistance *tout court*. Such a relational characterization does not seem, however, to be fine-grained enough to wholly identify what we mean by *resistance*. Medina 2013 seems to refer to a more complex phenomenon than resistance *tout court*. I will refer to the more complex phenomenon of resistance as *hope resistance*[2], whose particular outline runs from Ernst Bloch's idea of hope principle.

Hope resistance is a hybrid concept that defies hypostatization and abstractions (Thomason 2013, p.12), forcing philosophers to undertake approaches that are contextual and performative. In his visionary trilogy, Bloch never resorts to analytical principles or positive exemplifications. Rather, he tends to use negative examples and make the readers acquainted with hope through guided jumps of comprehension. Regrettably, Bloch is never mentioned by Medina 2013, either in the text or in its bibliography. To clarify the concept of resistance, Medina recalls illustrious examples of resistant heroes and heroines, e.g., Juana Inez de la Cruz and Rosa Parks.

Without clarification of the abovementioned distinction between resistance *tout court* and hope resistance, the exposition through exemplary figures runs the risk of being misunderstood or unserviceable. Discourses about science fiction and crime TV series may provide us with illustrative examples of this eventual kind of misunderstanding (Cambra Badii 2018).

---

[1] The fruitful application of case studies from TV series in philosophical and psychological investigations has been supported and expounded in Cambra Badii 2018.

[2] McDonald, Stephenson 2010 interestingly couples the concept of hope with that of resilience, which shares lines of affinity with that of resistance. Also, for an embedding of hope into virtue epistemology and ethics one could consult Billias 2010.

The audience of sci-fi and crime series seems, indeed, to behave ambivalently[3]. Hooked spectators follow the plots and heartly cheer for Lacie in *Nosedive* (Brooker 2016), Becky or Jessica Hyde in *Utopia* (Kelly, 2013-2014) or The Professor in *Money Heist* (Pina, 2017-2021). They seem to understand their heroes' reasoning, supporting their deliberations against the grain. While being sympathetic with them, the audience seems contemporarily unable to produce any real corresponding resistance in their routine way of using digital and medical technology or of being involved in worldwide financial machinery. Following my distinction, however, there might be some conceptual confusion in interpreting this behaviour in terms of ambivalence.

Social or individual changes can be the effects of either resistance *tout court* or hope resistance. Accordingly, the fact that the audience's aversion to digital technology is not manifestly transferred from the fictitious to the real world cannot stand for a neat absence of resistance in it. One can interpret this absence in terms of resistance *tout court*, which does not imply on its own a vanishing of hope resistance too.

This paper follows this outline. The first and second sections are dedicated to the introduction of the idea of hope resistance. In the first section, I will introduce the concepts of utopian function and anticipation drawing upon the first volume of Bloch's 1995 trilogy. In the second section, I will draw inspiration from the heroine Juana Inez de la Cruz to exemplify utopian function and hope resistance. In the fourth and fifth sections, I will show that characters of *Black Mirror, Utopia* and *Money Heist* bring to light aspects of hope resistance. I will conclude that corresponding appearances of ambivalent behaviours deserve more fine-grained recalculations

2. Utopian Function and Anticipation

Ernst Bloch has been numbered among traditional authors of future philosophy. Indeed, one can find a brief, but satisfactory, account of his future ontology in Poli 2017. Bloch's discourse on hope is layered along interwoven levels of analysis involving biology, affectivity, ontology and epistemology.

The hope principle begins with a kind of processual and biological monism, where modification, openness and change figure out as ontological pillars. Life in itself is an *immediate* condition where <<That we are alive cannot be felt>>.

> That which posits us as living does not itself emerge. It lies deep down, where we begin to be corporeal. […] Nobody has sought out this state of urging, it has been with us ever since we have existed and in that we exist. The nature of our *immediate* being is empty […] all of this does not feel itself, in order to do so it must first go out of itself (Bloch, 1995, p. 45)

The *mediate* condition of life, namely, where life feels itself, arises as a 'vague and indefinite' thirst that urges *immediate* life to "go out of itself".

For one thing, Bloch sternly defuses the search for fundamental human drives (*Ibid*. Ch. 13). He indeed criticizes philosophers and psychologists positing as fundamental either sex, power, eros or archetypes. According to him, fundamentality is always biased by contingent constraints, mainly socioeconomic, that deceive our inquiries with false promises of absoluteness. For another thing, he determines two drives that can notwithstanding aspire to be more basic than others. The two seemingly fundamental drives are *preservation* and *completion*[4]. Preservation must be interpreted as an appetite for appropriate conditions of life unfolding. Completion must be interpreted as a drive towards ongoing perfection.

---

[3] The behaviour of TV series spectators can be considered a technological extension of human behaviours. Along this line, the embodiment relation of technology with the users, taken as living bodies, has been analyzed in Liberati 2019, Bonfiglioli 2021. In this respect, further insights to deepen can be found into the imaginative resistance phenomenon (see Tuna 2020). For the general philosophical relevance of fiction and the variety of theories thereupon see Aldegani 2021.

[4] For an ontological view of completion see Žižek 2013.

Preservation and completion are filled in by affectivity, which is a qualitative and situated condition of living matter. As soon as the process of "going out of itself" takes place, drives give rise to vague intentional acts of striving and longing without distinct intentional objects (*ibid*. p. 45). These acts can be further refined and streamlined in full-grown intentional acts, with their own distinct objects. The refinement of the acts of striving and longing constitutes what Bloch calls *desiderium* (*ibid*. Ch. 10).

In addition to the satisfaction of particular desires, a general unsatisfied desire is always residual. Residuality brings about the distinction between *filled* and *expectant affects* (*ibid*., p. 74). The former is "short-term", that is, their intentional drive object lies in the "already available" world. Envy, greed, and admiration are kinds of filled affects. Expectant affects instead are "long-term", that is, their intentional drive object does not lie yet ready, either in the respective individual attainability or in the already available world (*ibid*., Ch. 13). Examples of expected affects are anxiety, fear, hope or faith.

At first glance, the distinction between filled and expectant affects seems odd. Bloch indeed classifies hope among the expectant affects, but to be sure, one's hope can be directed towards short-term objects as well. In addition, I would say, it is usually so. For instance, I can hope that tomorrow will be sunny or that the war in Ukraine will not break out. Which objects can expectant hopes be oriented to?

According to Bloch, traditional ontology and epistemology have usually dealt with bounded objects, that is, objects whose identifying content, function or machinery, even though unknown, are already established. Bloch's objects claim to be open in this respect. According to him, what we usually call an "object" is just the contingent and socioeconomically constrained endpoint of a wider ontological wake of its possible configurations.

Traditional ontology and epistemology make historically determined cuts along ontological wakes. This would not be a problem, however, if only one is ready to admit that the cut is a matter of temporary convenience and is prompt to hand again full ontological wakes to objects themselves. Unfortunately, this readmission hardly surfaces and what were just convenient cuts are likely to become tough constraining boundaries.

The removal of ontological boundaries opens up the real possible horizon of an object. Real possibility involves at least four kinds of futures, namely, desirable, probable, plausible and possible futures (Poli 2019, Ch.10). The desirable future of an object, say x, includes those not-yet configurations of it that we would prefer to live with. Probable futures are the not-yet configurations of x that are in line with our current social constraints, habitus or trends. Plausible futures are the not-yet configurations of x that are in line with the current state of our knowledge. Possible futures are all the configurations of x that we can materially, or even only abstractly, imagine.

It is important to notice that desirable futures give rise to a cross-section of the other kinds of futures. That is, a future configuration is desirable regardless of its similarity or dissimilarity with what is probable or plausible. Desirable possible futures make up what Bloch calls 'the immense utopian field of the world' (Cf. Bloch 1995, p.6). Expectant affects are usually directed toward possible futures, there lying long-term objects, while filled affects are usually directed toward plausible or probable futures, there lying short-term objects.

Desirable, probable, plausible, and possible futures are ordinated by inclusion and intersect as follows:
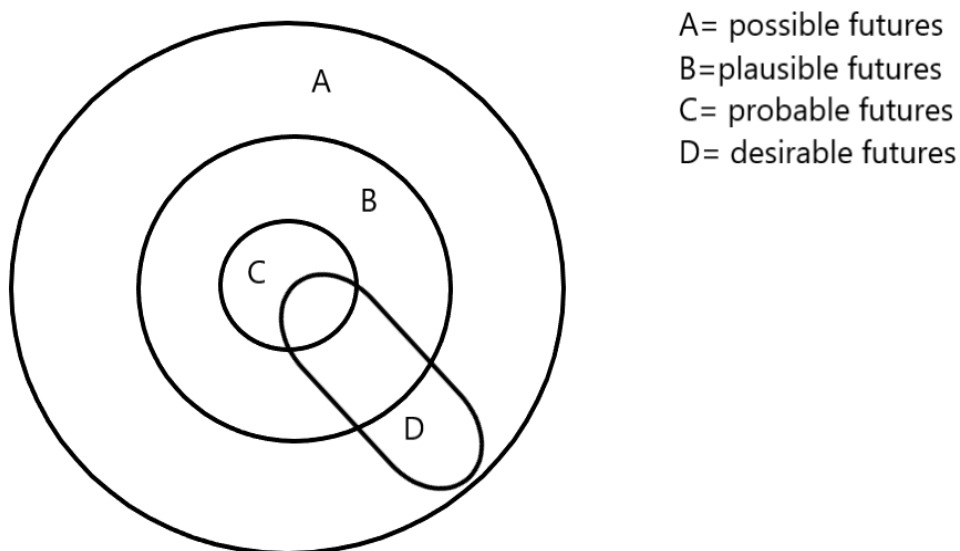
Figure 1. *Intersecting Futures*

We may imagine now that an object draws a line along the intersecting futures. For instance, we may imagine that the following red line is drawn by the Italian university institution:
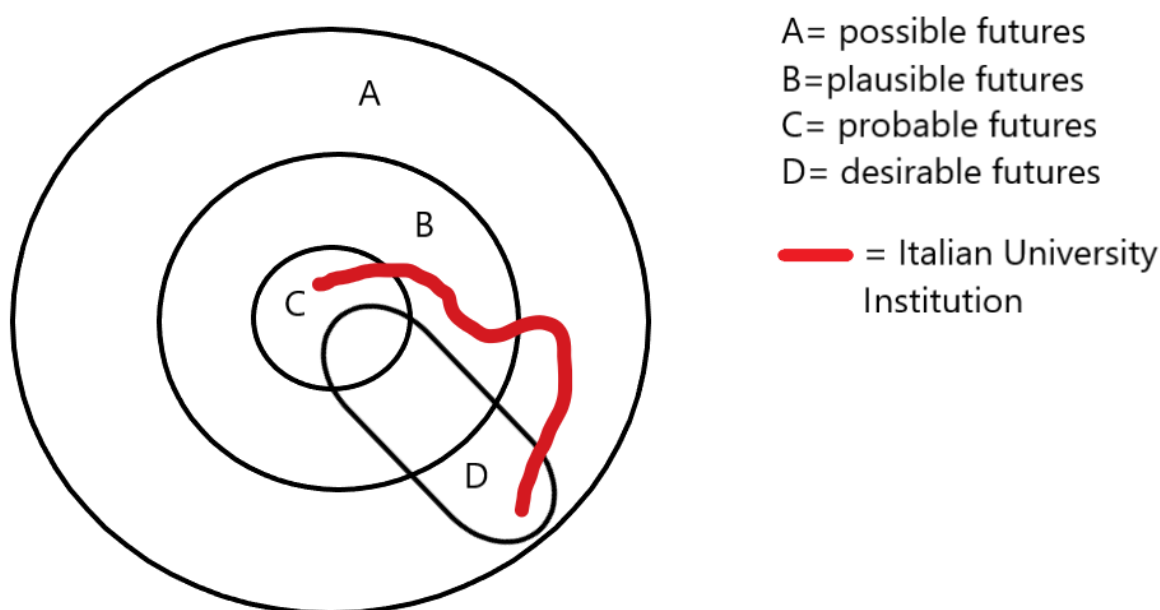
Figure 2. *The Line of an Object along Intersecting Futures*

Along this line, the Italian university institution finds a desirable future only in the realm of possibility and utopian field. This means that its desirable configurations are very long-term and that a desire

directed toward it is expected to exert what Bloch calls *utopian function*, which is filled in by expectant affects (Bloch 1995, p.115, *passim*).

The utopian field becomes epistemically accessible by stripping away from the objects their acquired boundaries. The blowing up of historically determined boundaries liberates from the objects an almost blinding beam of light. This liberated light constitutes the phenomenon Bloch calls "resistance of the *novum*" (*ibid*. p. 128-32). The *novum* arises here as something merely determinable that must be informed and determinate, not even without difficulties, to be bestowed with an anticipatory role that succeeds in driving our actions.

We can get acquainted with the blinding light of the *novum* by taking into account what is like to experience the epistemological scope of the hope principle. The epistemological aspects of the hope principle are in plain contrast with what is called the *contemplative paradigm* (*ibid*) of knowledge. The contemplative paradigm currently dominates knowledge. It sets the results of cognition back as a premise of cognition itself and refers to some already-established being of the object of knowledge. Upon closer inspection, this assumption is already implicitly present in our cognitive life, both in our common sense and in our scientific or philosophical-systematic endeavors. Let me point out a couple of examples. If I see my desk, I *know* that I'm seeing it as it already is and as it were tonight before I woke up. If I see a star, I know that I'm seeing it as it had already been several years ago.[5] If I give proof of the Pythagorean Theorem, I'm implicitly assuming that the ratio of the legs identifies a property of the hypotenuse as it already is and not as it is going to become or as it has not yet become. Hence, it is no easy task to imagine or visualize any change in the contemplative paradigm of knowledge. This novelty would be a total breach of our cognitive *habitus*, thus showing that kind of resistance to determination that according to Bloch is typical of the new and the determinable.

Anticipation arises as a practical utopian activity or a militancy toward the novum. The practical facet of anticipation makes up the cornerstone of nowadays future studies[6]. Anticipation allows us to gradually work out the dialectics between the determination and the determinability of the novum. Following the hope principle, we have a biological impulse to anticipate the novum instead of passively waiting for its realization. This impulse would be encoded within the fundamental drives to preservation and completion.

In the next section, I will draw an example of utopian function from the original Mexican feminism.

3. The Case of Sor Juana Inez de la Cruz

In seventeenth-century Mexico, Juana Ines de la Cruz felt a strong *desiderium* towards knowledge throughout her life. Now, following Bloch's perspective, knowledge, like any other object, is encapsulated in a social narrative that confines it and its reachability within specific boundaries.

Knowledge was out of Juana's reach because of her contemporary narratives. First, she was a young woman and thus could not go to the University of Mexico City. Then, she became a nun, and thus, she couldn't have masters to teach her arts, literature or sciences. She had to spend most of her time after monastic stuff and could dedicate only spare time to her lay inclinations.

---

[5] In the cognitive sciences, these two examples are reflected in the causal theory of perception.

[6] <<Anticipation consists of two elements: a model and its translation into action. Looking at the weather forecast is not an anticipatory activity in itself, but deciding to take an umbrella as a result of what you have seen is. Forecasts are nothing more than a model, they just tell us what could happen. Having seen what could happen and changing one's behaviour accordingly is a case of anticipatory activity. Still, anticipatory behaviour is more robust than reactive behaviour, i.e. waiting for something to happen and then responding. Reactive strategies are expensive and inefficient: they are a sure recipe for failure. There are no valid alternatives to a forward-looking approach. If anything, the point is to understand that there are many different ways of anticipating and that it is necessary to find those that are best suited to one's situation; just as it is necessary to understand the cognitive and social constraints that filter and condition the translation of a model into action. (Poli 2019, p. 14-5)

In a letter to Sor Filotea, Juana said that her longing for arts, literature and sciences was so strong that she never succeeded in burning it off:

> What is true is I will not deny [...] since the first light of reason beamed on me, my inclination towards letters was so vehement and powerful, that not even outside reprimands --which I have had many--, nor my reflexes --which I have done quite a few--, have been enough for me to stop following this natural impulse God placed in me (Cruz 1994 p.10). [translation mine]

As it seems, Juana's desire for knowledge was hard to eradicate. Rather, it grew up stronger, as soon as one tries to remove it (*ibid*. p. 12). As is clear in the letter, Juana started studying humbly following her only feasible routes. The drive for knowledge was so strong to allow her to put into brackets even her socially acquired feeling of inferiority[7]. By designing her path, Juana opened up the ontological boundaries of knowledge. With this move, she put into question a wholesale received view of knowledge according to which knowledge is a kind of Russian doll made up of specialized disciplines encapsulated one in the other.

According to her contemporary dominating narratives, the study of theology was conceded only to students who were previously trained in specific disciplines. This condition was unaffordable for Juana, who strives for knowledge alone and without libraries, masters or disciples.

Despite her difficulties, Juana started studying specialized disciplines just by reading the books she happened to have and along the order she could. Along this path, she surprisingly discovered otherwise unforecastable facets of knowledge. In particular, she experienced that knowledge was not at all divided into encapsulated sectors. It was rather horizontal and allowed students to learn disciplines through free wandering from one sector to the other. Music, for instance, might allow Juana to understand history, cooking chemistry, poetry mathematics, and toy physics.

> And so, by having some principles gained, I continually studied various things, without having any particular inclination, but rather all in general. Therefore, having studied some in more detail than others has not been my choice, but rather the chance of having come across more books of those faculties has given them, without my deliberation, the preference. [...] True, I say this about the practical part [...] because it is clear that while the pen moves, the compass rests and while the harp is played, the organ calms down, et sic de caeteris; [...] but in the formal and speculative the opposite happens, and I would like to persuade everyone with my experience that not only do they not get on the way, but they help each other by giving light and opening the way for one another, through variations and hidden links [...] in a way that seems to correspond and are united with admirable bonding and concert. [...] I can assure you that what I do not understand in an author from one faculty, I usually understand in another from another that seems very distant; and those own, when explained, open metaphorical examples of other arts [...]. (*ibid*. p. 13-14, translation mine)

> Well, what could I tell [...] about the natural secrets that I have discovered while cooking? I see that an egg is united and fried in butter or oil and, on the contrary, it breaks apart in the syrup; [...] If Aristotle had cooked, he would have written much more. (*ibid*., p. 24, translation mine)

Juana's strong desire for knowledge led her to reshape the boundaries of this object, drawing materials upon the utopian field of desirable possible knowledge. The example of Juana also allows us to illustrate that the anticipatory militancy ignited by utopian function conveys withstanding capacities in front of even adversary present conditions.

As might be clear, Juana's resistance is not just resistance *tout court*. It is not just a matter of maverick keeping on or giving something up. Juana was concerned with the ontologically subversive endeavor of

---

[7] For the importance of self-trust in fighting against epistemic injustice see El Kassar 2021.

opening up knowledge boundaries to utopian horizons of new configurations[8]. Because it involves such a utopian function, Juana's resistance can be considered a true example of hope resistance.[9]

In the next section, I will analyse the characters of Lacie in *Nosedive* and Becky in *Utopia*. I will also show that their resistance is hope resistance.

4. Hope Resistance in Science Fiction TV Series

This section identifies elements of hope resistance in science fiction TV series. There are, I think, at least two cases deserving discussion. One is *Black Mirror*[10] and the other is *Utopia*. The episodes of the former breach the banks that digital technologies massively impose on human life emotions. The second blow up the fixed contours and limitations that massive biotechnological medicine superimposes on the quality of human living conditions.

Both series display examples of hope resistance. Some of their protagonists do not just strenuously reject digital technologies or massive biotech sanitary treatments. Rather, they reshape the current boundaries of human health and satisfaction.

The first example of hope resistance I want to draw attention to comes from *Nosedive*, a very well-known episode of the highly frequented TV series *Black Mirror*. In *Nosedive*, the protagonist Lacie Pound aims at feeling content with herself and up to the level wherever she winds up in. Unfortunately, however, she finds these targets confined within the reluctant dominating narrative of a network of digital social ratings.

Lacie lives in a society where private citizens, politics, enterprises, etc., do such intense use of a digital social rating network that one cannot rent a house or a car or cannot get access to medical treatments unless she has at least a certain score, 4.5 for example. Lacie strives to be highly scored, but ironically, the more she strives for it, the lower score she obtains. As a result, she always ends up being frustrated. Her most desirable future is somehow trapped and locked in a cage whose key was momentarily unavailable to her.

The bet of hope resistance on Lacie consists of constructing another narrative around contentment and personal adequacy. If Lacie exerted only resistance *tout court*, she would have just thrown the smartphone away without diminishing, or rather worsening, her grade of frustration.

Luckily, Lacie succeeds in discovering the liberatory power of utopic functions, thanks to her getting into acquaintance with Susan. Susan is a middle-aged truck driver and plain woman, appearing amid a highly trafficked railway. She scored very low. However, despite it, she is content with herself.

Susan's contentment does not depend on her resistance *tout court*. She has not just abandoned the social network Lacie was fond of. Susan has rather conquered a utopic scenario where her being adequate and content undergoes reconfiguration. Contentment and personal adequacy accordingly change their contours, becoming independent of any digital rating mechanism.

---

[8] The interlace between resistance and epistemic academic narratives is investigated in Bonfiglioli 2020 focusing on Covid19 pandemics.

[9] Medina's analysis of Sor Juana brings about themes that are close to the ones I have so far outlined: << Despite Sor Juana's constant critique of epistemic arrogance and constant (even excessive) exercises of humility, she argues that we should not get discouraged by the difficulties and dangers involved in the pursuit of knowledge. For, no matter how extreme these difficulties and dangers may be, they make the journey and its achievements all the more valuable; and it is ultimately up to the subjects themselves (regardless of their gender identity or sexuality) to decide whether they are up to the challenge and whether the epistemic risks are worth taking. […] Our experiential perspectives can be broadened with our capacity to imagine, to survey possible worlds in which alternative experiences can be had. This kind of imaginative knowledge has a crucial counterfactual dimension: even if the actual world does not allow certain experiences to be had, their possibility can be used as the basis of alternative knowledge, an epistemic counterpoint to lived experience and knowledge, which is still grounded in real life and embodied experiences>> (Medina 2013, p. 232).

[10] For a detailed discussion on *Black Mirror* and its relevant bibliography from the specialized point of view of film theory, one can consult the recent paper Sorolla-Romero *et al*. 2020. For a philosophical perspective on the same series, consult Johnson 2020.

Well, Susan's life choices are echoed in Lacie. She starts doing the same, not caring for the scores and being brutally sincere with all people. Brutal sincerity renders Susan and Lacie marginalized women, who fiercely withstand their marginalization. They may bear it so stoically because they cherish the utopic idea that contentment and adequacy can finally be liberated from chains of oppressive digital rating narratives. From this perspective, current marginalization thus arises as working anticipation of positive possible futures. Susan and Lacie militate in this utopic direction, thereby withstanding any current social rating without running the risk of being trampled to death of social frustration. The example of *Nosedive* support us to recognize that anticipatory practices may foster the capacity of people to reinterpret present bed conditions in term of weak signals (Poli 2019[11]) of a positive upcoming blossoming future.

Utopia is a less-known work than *Nosedive*. It was launched in 2013 on an English channel and then canceled because of purported image crudity. In 2020, *Utopia* was remade by Amazon. The 2013 original and the remake are quite different from each other. Up to the present, I didn't get the occasion to properly appreciate their differences. In this paper, I thus refer to the original 2013 version.

Utopia has a highly entangled plot. This aspect is underscored also by the fact that the final episode contains a hint of a never-released continuation. *Utopia* refers to a project of DNA selection, called Giano, developed by Corvardt, a pharmacy industry. Through Giano, Corvardt works on the project of decreasing the human population by spreading a seeming vaccine produced for confronting purportedly dangerous illnesses. In the fight against Giano, all the characters of *Utopia* display the traits of resistance *tout court*. Nevertheless, the series also offers examples of true hope resistance.

Jessica Hyde, for instance, is an unforgivable hope-resistant character. However, her vague traits of insanity render her analysis both slippery and critical. For this reason, I prefer to allocate her analysis to contributions tailored to *Utopia*. The character of Becky better serves the purpose of this paper.

Becky is a Ph.D. candidate in medicine. She apparently suffers from a degenerative illness called Diel's syndrome. From the beginning of the series, it is clear that Diels syndrome could be a genetically transmissible illness arising at the core of Giano. Becky suffers from spasms and epileptic attacks. She knows she could only delay these symptoms by taking a pharmakon called Thoraxine.

Becky desires to discover the truth about Giano and her related illness. At the end of the series, she finally determines at least part of that truth. In particular, she discovers her illness was not an illness properly. Spasms and epileptic attacks are negative side effects of Thoraxine itself, which turns out to be an opiate. She gets access to this truth as soon as she refuses to take Thoraxine. This action of refusal encodes elements of resistance *tout court*. It remains clear however that Becky's is a more complex turn[12].

First, her turn is deeply motivated by suspicions of the Corvardt project and the mainstream political narrative about it. Second, it is supported by an anticipatory jump to a utopic future where the acquired view of techno-medical health blows up. Putting it in Bloch's terms, Becky liberates the *novum* along the ontological wake of human health. According to such *novum*, health passes from being a function of pharmacology and technomedicine to a proper function of human bodies.

Becky's trajectory reveals another aspect of hope resistance. This aspect is something we may call a *debunking power* of hope resistance. Indeed, hope resistance, with the aid of epistemic suspicion, may

---

[11] <<Weak signals are born and die all the time. There are a lot of them, but their patterns are difficult to see and interpret because they are weak. Only a few can consolidate and become new trends. Since they are not generally appreciated by decision-makers, a naive approach to weak signals is doomed to fail. Asking a person to pick up a weak signal is like asking them to point out something irrelevant that might be relevant. They are usually recognized by pioneers or special groups, rather than by recognized experts. To intercept them it is useful to work with the future. natives, those who adopt new technologies first, who question the paradigms, and who create the future themselves>> (Poli 2019, p. 75 *passim*).
[12] A detailed account of the performing relationship between hope, on the one hand, and therapy, suffering and medical practices, on the other, can be found in Waterworth, Chs. 5,4 and Haramati 2010.

help people liberate themselves by demystifying certain narratives and opening up new explanatory horizons.

Becky's epistemic manoeuvre would have been out of her reach if resistance was restricted to resistance *tout court*. The strength of Becky's turn does not lie just in the act of giving up with her counterproductive pharmakon. Rather lies in a subversive and jumping act of anticipatory hope that brings her close to authentically new working paradigms for human health conditions.

In the next section, I will draw examples of hope resistance from a well-known case of heist crime TV series.

5. Hope Resistance in Heist Crime TV Series

One of the most-watched TV shows in the world and the most-watched non-English language series on Netflix is the crime *Money Heist*. It originally appeared as a Spanish TV show called "*La Casa de Papel*". Netflix acquired global streaming rights in December 2017, and within a few months, the series became a global phenomenon.

Money Heist has undoubtedly passed symbols of resistance *tout court* on to its audience. As one can see in the Netflix documentary *Money Heist: The Phenomenon*, Red jumpsuits, Salvador Dalì masks and the song "Bella Ciao" have been adopted and repurposed all over the world as symbols of civil rights protests or in the service of causes such as feminism and environmental justice (Cf. Del Sordo 2020).

Money Heist tells the story of two outstanding robberies whose creator is Sergio Marquina, aka The Professor. He comes from a poor family and has suffered from a serious and long-disabling illness. Sergio desires the well-being of himself and his family. Unluckily, well-being appears to be out of his reach because of the extreme poverty his family gets through. Sergio's poor father spent his time planning robberies because he wanted to find sufficient money to provide his son with the best medical treatments. However, he died of murder in front of a bank while carrying out the heist he planned.

As opposed to Lacie, Susan or Becky, Sergio's utopian function does not affect his received idea of well-being. As is clear, indeed, financial capital property endures as an essential condition of well-being throughout the series. Rather, the utopian function of The Professor affects his acquired status of a cogwheel in the global financial system. Instead of taking financial capitalism as an inescapable life condition, he starts studying its inner machinery to eventually find its blind spots and then destroy its functioning. The Professor indeed assaults not any bank or depository but targets of the utmost significance, namely those institutional buildings, such as the Royal Mint or the National Bank, that regulate the circulation of money in many Western countries.

Insofar as Marquina pursues the circumvention of the global machinery of financial capitalism, he is forced to imagine possible scenarios beyond the given probable and plausible futures. He epistemically places himself in front of a field of possible configurations that stretches financial capitalism beyond its already given boundaries. To make sense of his heist plans, The Professor needs to be revolutionary and thus imagine a space of theoretical, practical and ethical agency beyond the available predictabilities.

To be sure, the resistance *tout court* remains one of the most punching ingredients of *Money Heist*. Nevertheless, one should be ready to admit that resistance is not all about that. The driving force of the plot is Sergio's utopian function that allows him to outline the two heist plans to which the gang is dyed-in-the-wool devoted. Accordingly, one may eventually venture the thesis that resistance *tout court* is a kind of second-order phenomenon triggered by a deeper hope resistance process.

Conclusions

I have argued that an exemplification of the concept of resistance must be preceded by a head-on conceptual clarification. Accordingly, I have described the difference between two kinds of resistance: resistance *tout court* and hope resistance. The first amounts to the exertion of a counteracting force against oppressive conditions. The second instead is rooted in the complex phenomenon of hope that comprises utopic function and anticipation. Sticking on the method of exemplification, I have outlined their difference through examples from Mexican feminism and science fiction and heist crime TV series.

The distinction between the phenomena of resistance *tout court* and hope resistance allows us to reinterpret the purported ambivalent behaviour of the audience of the sci-fi and crime series. Addicted audiences seem indeed vaguely ambivalent because they appear to support resistance actions only in fictitious worlds and not in the real world. According to my distinction, this is too easy a claim. The fact that one does not pursue resistant actions, in reality, deserves to be better interpreted. Indeed, one can harbour an inner process of hope resistance that still has to blossom into echoing and influential actions of social or individual counteracting forces.

The problem remains open of which relationship occurs between hope resistance and resistance *tout court*. One may indeed conjecture that the latter may arise as a second-order phenomenon triggered by the former. Accordingly, one may also introduce a certain relation of partial ontological dependence between the two forms of resistance.

Bibliography

Aldegani, E., (2021), Exploring Fictional Spaces, *PROMETEICA - Revista de Filosofia y Ciencias*, 22, pp. 6-7, UNIFESP, São Paulo. DOI: https://doi.org/10.34024/prometeica.2021.22.11653

Billias, N., (2010), Hope as a Moral Virtue, McDonald, J., Stephenson, A. (eds), *The Resilience of Hope*, Probing the Boundaries, Vol. 68, Rodopi, Amsterdam – New York.

Bloch, E., (1995). *The principle of hope*, The MIT Press, Cambridge, MA.

Bonfiglioli, C., (2020), A academia em tempos de pandemia, *PROMETEICA - Revista de Filosofia y Ciencias*, 21, pp.5-6. UNIFESP, São Paulo. DOI: https://doi.org/10.34024/prometeica.2020.21.11099.

Bonfiglioli, C., (2021), Methodological Challenges for a New Philosophy of Technology. Interview with Nicola Liberati, *PROMETEICA - Revista de Filosofia y Ciencias*, 23, pp. 150-154. UNIFESP, São Paulo. DOI: https://doi.org/10.34024/prometeica.2021.23.12411.

Brooker, C. (2011-2019), *Black Mirror*, Endemol Shine UK.

Brooker, C. (2016), *Nosedive*, in Black Mirror, Series 3, Endemol Shine UK.

Cambra Badii, I. (2018), Pensar el Cine. La narrativa de Películas y Series como Matriz Metodológica para el Tratamiento de Problemas Complejos, *PROMETEICA - Revista de Filosofía y Ciencias*, 17, pp. 62-67. DOI: https://doi.org/10.24316/prometeica.v0i17.230.

Cruz, J. I. de la (1994). *The Answer /La Respuesta*, Including a Selection of Poems. New York: The Feminist Press at the City University of New York.

Del Sordo, C. (2021), Money Heist. The Philosophical Bond of Resistance, ANDPHILOSOY.COM, The Blackwell Philosophy and Pop Culture Series, https://andphilosophy.com/2021/06/15/money-heist-the-philosophical-bond-of-resistance/

Haramati, S. (2010), Hope in Pandora's Box: Psychological Work with Medical Patients, in McDonald, J., Stephenson, A. (eds), *The Resilience of Hope*, Probing the Boundaries, Vol. 68, Rodopi, Amsterdam – New York, pp. 143-153.

Johnson, D. K. (ed.) (2020), *Black Mirror and Philosophy. Dark Reflections*, in "The Blackwell Philosophy and Pop Culture Series", John Wiley & Sons, Inc., Hoboken, NJ.

Kelly, D. (2013-2014), *Utopia*, Endemol Shine UK.

Liberati, N. (2019), Emotions and Digital Technology. The Effects Digital Technology Will Have on Our Way of Feeling Emotions According to Post-Phenomenology and Mediation Theory, Humana. Mente Journal of Philosophical Studies, in Cinfereon, B., Lanfredini, R. (eds) *Beyond the Self: Crisis of Disembodied and Individualistic Paradigms*, Vol. 36, pp. 292-309.

McDonald, J., Stephenson, A. (2010), *The Resilience of Hope*, Probing the Boundaries, Vol. 68, Rodopi, Amsterdam – New York.

Medina, J. (2013), *The Epistemology of Resistance. Gender and Racial Oppression, Epistemic Injustice and Resistance Imaginations*, Oxford University Press, New York

Nadja El Kassar (2021), The Powers of Individual and Collective Intellectual Self-Trust in Dealing with Epistemic Injustice, *Social Epistemology*, 35(2), pp. 197-209. DOI https://doi.org/10.1080/02691728.2020.1839592

Pina, A. (2017-2021), *Money Heist*, Netflix.

Poli, R. (2017), *Introduction to Anticipation Studies*, Anticipation Science (1), Springer.

Poli, R. (2019), *Working with The Future. Ideas and Tools to Govern Uncertainty*, Bocconi University Press, Milan.

Sorolla-Romero, T., Palao-Errando, J. A., Marzal-Felici, J. (2020), *Unreliable Narrators for Troubled Times: The Menacing "Digitalisation of Subjectivity" in Black Mirror*, Quarterly Review of Film and Video, 38(2), pp. 147-169. DOI: 10.1080/10509208.2020.1764322

Thomson, P. (2013), The Privatization of Hope and the Crisis of Negation, in Thomason, P., Žižek, S. (eds), *The Privatization of Hope. Ernst Bloch and The Future of Utopia*, 8, Duke University Press, Durham and London, pp.1-21.

Tuna, E.H (2020) "Imaginative Resistance", in Zalta E. N. (ed) The Stanford Encyclopedia of Philosophy, (Summer 2020 Edition), URL = <https://plato.stanford.edu/archives/sum2020/entries/imaginative-resistance/>.

Waterworth, J. M. (2004), *A Philosophical Analysis of Hope*, Palgrave MacMillan, New York.

Žižek, S. (2013), Bloch's Ontology of Not-Yet Being, in Thomason, P., Žižek, S. (eds), *The Privatization of Hope. Ernst Bloch and The Future of Utopia*, 8, Duke University Press, Durham and London, pp.xv-xx.

POSTPHENOMENOLOGICAL VARIATION OF INSTRUMENTAL REALISM ON THE "PROBLEM OF REPRESENTATION"

fMRI IMAGING TECHNOLOGY AND VISUAL REPRESENTATIONS OF THE HUMAN BRAIN

VARIAÇÃO PÓS-FENOMENOLÓGICA DO REALISMO INSTRUMENTAL NO "PROBLEMA DA REPRESENTAÇÃO"

Tecnologia de imagem fMRI e representações visuais do cérebro humano

VARIACIÓN POSTFENOMENOLÓGICA DEL REALISMO INSTRUMENTAL SOBRE EL "PROBLEMA DE LA REPRESENTACIÓN"

Tecnología de imágenes de fMRI y representaciones visuales del cerebro humano

Dmytro Mykhailov
(Southeast University, China)
101300137@seu.edu.cn

ABSTRACT

In the present paper, I take findings from the postphenomenological variation of instrumental realism to develop an 'environmental framework' to provide a philosophical answer to the 'problem of representation.' The framework focuses on three elements of the representational environment, image-making technology, image as a representational device, and scientific hermeneutic strategies occurring within the image interpretation process in the laboratory set-up. The central idea in this regard is that scientific images do not produce meanings without their instrumental environment or that an image becomes representational through the interplay between three framework elements. In the second part of the paper, I apply the framework to contemporary debates on fMRI imaging. I show that fMRI images receive meaning not in isolation but within a complex instrumental environment.

Keywords: problem of representation. scientific image. instrumental realism. postphenomenology. fMRI.

RESUMO

No presente artigo, tomo as descobertas da variação pós-fenomenológica do realismo instrumental para desenvolver uma "estrutura ambiental" a fim de fornecer uma resposta filosófica ao "problema da representação". A estrutura se concentra em três elementos do ambiente representacional, a saber - tecnologia de criação de imagens, imagem como um dispositivo de representação e estratégias hermenêuticas científicas que ocorrem dentro do processo de interpretação de imagens no laboratório. A ideia central a esse respeito é que as imagens científicas não produzem significados sem seu ambiente instrumental ou, dito de outra forma, uma imagem torna-se representativa pela interação entre três elementos da estrutura. Na segunda parte do artigo, aplico a estrutura aos debates contemporâneos sobre

imagens de fMRI. Mostro que as imagens de fMRI recebem seu significado não isoladamente, mas dentro de um ambiente instrumental complexo.

Palavras-chave: problema da representação. imagem científica. realismo instrumental. pós-fenomenologia. fMRI.

<div align="center">RESUMEN</div>

En el presente artículo, tomo los hallazgos de la variación posfenomenológica del realismo instrumental para desarrollar un "marco ambiental" con el fin de proporcionar una respuesta filosófica al "problema de la representación". El marco se centra en tres elementos del entorno de representación, a saber: la tecnología de creación de imágenes, la imagen como dispositivo de representación y las estrategias hermenéuticas y científicas que ocurren dentro del proceso de interpretación de imágenes en el laboratorio. La idea central en este sentido es que las imágenes científicas no producen significados sin su entorno instrumental ni, dicho de otro modo, una imagen se vuelve representacional a través del juego entre tres elementos del marco. En la segunda parte del artículo, aplico el marco a los debates contemporáneos sobre imágenes de IRMf. Muestro que las imágenes de fMRI reciben su significado no de forma aislada sino dentro de un entorno instrumental complejo.

Palabras clave: problema de la representación. imagen científica. realismo instrumental. posfenomenología. fMRI.

1. Introduction

Contemporary experimental science provides us with a wide variety of different representations. The scope of them is ranging from digital atlases of our galaxy to weather forecast maps and X-Ray medical images (van Fraassen, 2008). All of them give us insights into the world we are living in and facilitate our scientific worldview with new intuitions and hypotheses. While philosophers from different domains are emphasizing the significance, if not the primacy, of various representational devices in science, there are still numerous debates around how representations are constructed, what they represent and how they should be interpreted by scientific communities (Amann & Knorr Cetina, 1988; Collins & Evans, 2002).

Moreover, there are different types of representational devices in science. Not all of them are the same. An X-Ray image, for example, makes visible an ankle joint differently from a biomechanical model, while an electric circuit model shows the brain activity in a different manner than a neural network model (Roman & Nguyen, 2021). In addition to that, every representation has its 'technical' part. Contemporary representations are mainly generated within a complex set-up of imaging technologies with the use of computer modeling techniques, mathematical analysis and sophisticated algorithms (Ihde, 2009). The structure, inner architecture and software environment of the imaging technologies affect the accuracy, visual appearance and 'readability' of the representation (Rosenberger, 2009).

The expanding role of representations in experimental science raised a list of questions about the nature, exactness, and ontological status of representational devices (Carusi et al., 2014). In contemporary philosophy of science, these questions have been titled under the name 'problem of representation' (Tibbetts, 1988). While scientific images are arguably the most frequently used representational devices nowadays (Friis, 2017; Sturken & Cartwright, 2017), in the present paper I will focus on the images as representational devices. Moreover, I will try to specify the role of scientific images within the 'problem of representation'.

In the simplest terms, this problem might be formulated in the following manner[1]. On the one hand, a scientific image always displays particular content (non-random data points) (Devitt, 2005). As a

---

[1] For another classification of the 'problem of representation' see Roman & Nguyen, 2021.

particular object, an image embraces something certain within its readout. An image of something is not just an image of anything (Psillos, 2017). Said differently, it is obvious that the use of scientific images does have an 'intent' to a truth-function and this 'intent' is realistic by its nature (Fumerton, 2002). Thus, it seems critical that scientific images 'truly represent' something real (realism claim[2]).

On the other hand, a scientific image isn't just a copy of the 'real' object (Hoeppe, 2015). It is isomorphic only to some extent[3] (Knorr-Cetina, 1999). Moreover, the meaning of scientific images is constantly open to various interpretations[4] (Rosenberger, 2008). Different scientists produce different perspectives on the interpretation of an image (constructivist claim[5]). As Bruno Latour and Steve Woolgar put it: "'out-there-ness' [i.e., the external world] is the consequence of scientific work rather than its *cause*" (Latour & Woolgar, 1986, p. 182).

In this sense, realism and constructivism provide two different perspectives on the nature of the scientific images. Realists claim that the image relies on the independently existing reality which substantially grounds the image (Devitt, 2011; Saatsi, 2018), while constructivists insist that through images scientists 'construct' what is real. Reality isn't something that we have from the very beginning – constructivists would argue – but, on the contrary, something that we 'build' by means of various scientific practices[6] (Lynch & Woolgar, 1988).

However, there is a third possible perspective on the 'problem of representation'. This perspective is provided by instrumental realism[7] (Ackermann, 1988; Hacking, 1983; Heelan, 1989; Ihde, 1991). Instrumental realists claim that rather than stressing on reality itself (or social practices that construct it) we should better take a closer look at the scientific *instruments* which give us access to what realists would call 'real' (de Boer et al., 2021; Ihde, 1998; Verbeek, 2003). Moreover, scientific instruments are never neutral because they mediate our access to the world, transform our perception, and thus - shape the process of knowledge production (Liberati, 2016; Mykhailov, 2020). As a result, when realists and constructivists focus their attention on the one pole of the 'representational' schema (either on the reality or on the social practices), instrumental realists insist that we should better concentrate on the transformative role of *the instruments* that are actively mediating (and co-shaping) both poles together (Aydin et al., 2018; Kudina, 2021; Mykhailov & Liberati, 2022).

Taking all this into consideration, I am going to show that instrumental realism can make a strong contribution to the 'problem of representation' on both theoretical and practical levels. On the theoretical level, I will develop a threefold 'environmental framework' for the analysis of the scientific images. Using this framework, I am going to show that an in-depth understanding of the nature of representation should include not only the philosophical conceptualization of a representational device (an image) but also surrounding elements that constitute its environment. The central idea in this regard is that representations do not produce meanings without their instrumental environment, which is their constitutive element[8]. Keeping this in mind, the framework focuses on three 'environmental' components fundamental for every scientific image. The first component is image production

---

[2] Here I refer to a 'minimal' form of realism (not to be confused with the 'minimal' realism by Juha Saatsi (French & Saatsi, 2020) that is shared by almost every type of realism in the philosophy of science. According to the minimal form of realism, the object of representation exists independently of subjective observations. For more on various forms of realism see Saatsi, 2018.

[3] There are not only isomorphic images (as some medical images like X-ray scans) but also various non-isomorphic images like mathematical models of elementary particles or images of quasars taken from radio telescopes. To read more on non-isomorphism in scientific images see Hoeppe, 2014; Kuchinskaya, 2013; Lynch & Ridder-Vignone, 2014.

[4] Especially in the case of the medical diagnosis. I will return to this point in the second part of the present article.

[5] For the sake of our argument, I refer to a 'radical' form of constructivism here. This view is shared by authors such as Karin Knorr-Cetina (1999), Bruno Latour (2003), David Woolgar and Michael Lynch (1988).

[6] In-depth analysis of the realists vs constructivists debate on the 'problem of representation' falls outside the scope of this paper. For more on these debates see Tibbetts, 1988.

[7] In the present paper, I introduce the postphenomenological variation of instrumental realism. There are other kinds of instrumental realism that were developed by authors such as: Patrick Heelan (1989), Robert Ackermann (1988), Ian Hacking (1983). For more on different kinds of instrumental realism see Ihde, 1991.

[8] Another significant part of the scientific infrastructure that I am not going to touch upon in this paper is scientific collectives (Hasse, 2008). From this, much wider perspective, any scientific image gets its meaning not only through particular relations between scientist and instrument but through broad scientific 'cultures' which establish theories, hypotheses, rules of observations, etc.

technology. The second component is an image as a technological artifact. Finally, the third component relates to hermeneutical strategies, which are taking place among scientists in a laboratory set-up.

On the practical level, the present paper contributes to the debate with an application of the 'environmental framework' to a specific class of medical imaging technologies, namely - fMRI. Within the last decade, fMRI technology has attracted the attention of scholars from different fields (Aydin, 2018; Carusi & Hoel, 2014; Fried & Rosenberger, 2021; Joyce, 2016). Such an increasing interest was called forth by the revolutionary impact that fMRI had on almost every part of neuroscientific research (Cabeza & Nyberg, 2000). However, fMRI technology images have a peculiar nature. On the one hand, fMRI technology images are not 'mere' representations because they are not just a 'copy' of the human brain (Hoeppe, 2015), while on the other hand, these images are often used as representational devices within medical diagnostics (Beaulieu, 2016; de Boer et al., 2021). As I will show in this paper, the 'environmental framework' can provide better understanding of the problematic nature of the fMRI scans. In terms of the current paper, the medical image becomes representational as a result of the interrelation between three elements of the framework.

My argumentation in the present paper will be developed in two steps. The first step is theoretical. Here I will introduce the 'environmental framework' concerning the 'problem of representation'. I will show how each of the three elements plays its role in a meaning-creating process and how through each of them an image becomes representational. The second step is practical. In this part, the argument will be developed with the reference to the fMRI technology. I will demonstrate that fMRI visuals receive their meaning not in isolation but within a complex instrumental environment. Each element of the representational environment will be analyzed through the lens of postphenomenological variation of instrumental realism. As usual, the conclusion will summarize the results and link them to the findings provided within the article.

2. Instrumental realism and the 'environmental framework' for the 'problem of representation'

As it becomes apparent from the title, instrumental realism is about *instruments* (Ihde, 1991). More specifically, about the role of the instruments in the manufacturing of scientific knowledge (Ackermann, 1988; Heelan, 1989). What comes into focus at first glance is that the adjective 'instrumental' stands *before* the noun 'realism'. In the same way as the notion of 'technoscience' put technology *in front* of science (Hongladarom, 2012; Ihde & Selinger, 2006; Zwart, 2022) instrumental realism shifts philosophical attention from reality in itself to the role of instruments in a knowledge production process (de Boer, 2021a). Keeping this in mind, in what follows, I will focus on the instrumental component within the 'problem of representation'.

2.1. Imaging technologies and hermeneutic mediation

It is common knowledge that contemporary experiments take place within a complex set-up of different imaging technologies (Carusi et al., 2014; Latour & Woolgar, 1986; Lynch & Woolgar, 1988). In this instrumental set-up the object of study (for example, the human brain) is placed in unnatural, artificially designed conditions - conditions in which this object has to 'meet' requirements put to it by the experimental environment (van Fraasen, 2008, p. 93-94). In this sense, imaging technologies 'force' the object of study to become visible (Verbeek, 2003). Said differently, instruments are *transformative technologies* that change the appearance of the object and make it available for scientists (Rosenberger, 2017). Without instruments, the object of study is either invisible in principle (like fundamental particles in physics or quasars in radio astronomy) or inaccessible for scientific perception and manipulation (Hoeppe, 2014). The latter becomes apparent in the case of many medical technologies which, for example, help us to see organs without surgery (like X-Ray or ultrasound), or enable us to accomplish diagnosis without 'direct' access to the patient's body (like in the case of the telemedicine) (Ihde, 2019).

According to postphenomenological variation of the instrumental realism imaging technologies are not just neutral tools that 'mirrors' reality but active participants of the scientific process (Hasse, 2008; Wellner, 2020). Within this process, imaging technologies are interpreting the object of study and making visible those phenomena that were previously unperceivable (Liberati & Nagataki, 2015). This interpretive role of the imaging technologies has been titled 'hermeneutic mediation' (Ihde, 1978; 1998; Nørskov, 2015). With the notion of 'hermeneutic mediation' representatives of instrumental realism in philosophy of science aim to conceptualize the transformations that the imaging technology has generated upon the object of study within the image creating process (Rosenberger, 2008). In this sense, during a hermeneutic mediation, imaging technologies create a representation of reality, a representation that will ask for interpretation from its 'readers' (e.g. scientists). What is more important, the imaging technology itself makes a 'translation' of the particular object of study into a representational device like an image, graph, or map (Verbeek, 2008).

The other significant feature of hermeneutic mediation is tightly related to technological selectivity (Briedis, 2022). Imaging technologies do not just copy the object of reference but 'select' specific elements of this object to be a part of the image while at the same time reducing other components from the image readout (Alač, 2013). According to Don Ihde, such an interpretative selectivity has a 'magnification-reduction structure' (Ihde, 1998). It highlights some parts of the content while concealing the others. In this sense, we can consider imaging technologies more as 'interpretative' rather than blindly 'representative' tools (de Boer, 2021a).

2.2. Image as a representational device

Scientific image has a peculiar nature. On the one hand, an image is a material artifact, an object among other objects, while on the other, an image is more than a 'mere' object (van Fraassen, 1980). It always *represents* another object by referring to the 'external' thing in the world (Hoeppe, 2015). However, an image as a representational device is never neutral. It always transforms the object of reference in both space and time.

Spatial transformation

The spatial transformation includes several changes of the object of reference. The first transformation is a 'transfer' of the three-dimensional object into flat, two-dimensional content of the image (Rosenberger, 2020). Consider the photo of the tree taken in the forest. Although the photo saves a lot of realistic elements (like color, shape, etc.), the original object (a tree) has lost its three-dimensional features. Now it is reduced to flat image content only (Ihde, 1998).

This spatial transformation goes along with numerous 'perspectival' transformations. As the content of the image is always 'one-sided' the perspective of the viewer is shaped and, in some way, 'sticks' to the only possible perspective provided by the image (Beaulieu, 2016). An X-Ray image, for instance, shows patient's bones from one specific position without letting the doctor take a different look at the patient's body (Friis, 2008; Briedis, 2022). The weather forecast map displays a part of the territory mediating scientific 'practices of seeing' to one perspective only (Hoeppe, 2014). The intelligent decision support system in diagnostics pin potentially dangerous spots on the medical image 'channeling' practitioners view to a particular area on the imaged data (Kudina & de Boer, 2021; Mykhailov, 2021).

Except for 'dimensional' and 'perspectival' transformations, there is also a 'morphological' transformation, which is a transformation of the object in shape and size (Verbeek, 2008). In many cases, such 'morphological' transformation is a necessary condition for a successful experiment. For example, an electron microscope that has a higher resolving power compared to the light microscope is able to magnify the object up to $10,000,000\times$ (Hacking, 1983; Lynch, 1985b). However, within such a magnification procedure, the size of the object of reference (for example a human cell) is transformed completely.

Temporal transformation

Together with spatial transformations, images simultaneously produce temporal transformations as well. Perhaps the most obvious transformation of this kind is the 'freezing-time' effect. The image can 'freeze' the dynamic processes and reduce them to a 'snapshot' (Rosenberger, 2005, 2009). For example, an fMRI scan 'pauses' a dynamic process of blood circulation in the brain and provides a practitioner with static renderings of the patient's brain activity (Joyce, 2006).

Except for a 'freezing-time' effect, scientific images are able to slow down some processes or, vice versa, speed them up (Prasad, 2005). A classic example in this regard is a video record (which is another type of image). By means of a simple video player, everyone can 'pause', slow-down or speed up the record. Another example of this transformation may be found in many medical imaging techniques that provide practitioners not only with one image but with the whole set of images generated at different points in time. Such a set of images provides a practitioner with a deeper insight into the patient's body, helps to understand the evolution of the disease and find the right treatment strategy.

2.3. Hermeneutic strategies in image interpretation

As it has been already noted, every scientific image doesn't 'mere' provide us with a straightforward representation of the world itself but rather exists as a technological artifact open to multiple interpretations (Friis, 2008). To put it simply, an image by itself doesn't have the meaning 'outside' of the scientific (and technological) environment (Alač, 2013). The meaning appears *as a result* of the image generation and interpretation process (Joyce, 2016).

Within instrumental realism such interpretational practices have been titled 'hermeneutic strategies'[9] (Ihde, 1998). Usually, during the first encounter with an image, a scientist can interpret the image in one way, while after a laboratory colleague provides an alternative hermeneutic strategy, s/he can see these alternative variations as well (Rosenberger, 2008). Such a hermeneutic strategy attracts scientific attention to particular elements of an image and helps to understand those features in a meaningful way (Goodwin, 1994, 1995).

The notion of hermeneutic strategy, thus, embraces and highlights various practical dimensions that occur inside the laboratory (Lynch, 1985). Usually, this interpretation process is extremely multidimensional and begs for a broad set of knowledge. For example, a scientist who interprets an image should know the related scientific context (both theoretical and experimental), details of the image-making process, recent related findings, etc. (Friis, 2017). Successful hermeneutic strategies also require knowledge about interpreting the contents of the image, what characteristics should be expected, what the background is, what the essential aspects are, what an irrelevant instrumental artifact is, and what the relation between different parts of the image is (Briedis, 2019; Rosenberger, 2020; Sturken & Cartwright, 2017).

Another significant idea related to the hermeneutic strategies is its *collectiveness* (de Boer, 2021a). Regularly, the image interpretation process is not just a byproduct of individual human-technology relations. On the contrary, the meaning of the image is usually born among different members of scientific collectives (Hasse, 2008). I will touch more on this during the discussion of the role of the practitioner in interpreting fMRI scans. For now, it is important to highlight that various hermeneutic strategies play a vital role in the discussion about the nature of representation.

3. fMRI in focus – practical application of the 'environmental framework'

---

[9] Hermeneutic strategies are related to the other concept inside the postphenomenological domain, namely, multistability. For more on this relation see de Boer, 2021b; Liberati, 2019; Wellner, 2020; Whyte, 2015.

"For the first time in the history of neuroscience, it is now possible to 'observe' cognitive activity in the intact human brain" (Cabeza & Nyberg, 2000, p. 1). This observational opportunity became possible because of fMRI technology. It enables us to see inside the brain without surgery or any other clinical invasion (Briedis, 2019). But how is it possible from the technical point of view? Let's take a brief look at how fMRI works.

The fMRI (Functional magnetic resonance imaging) is one of the major brain-mapping imaging technologies used nowadays. The main aim of fMRI is to generate images of the human brain (Prasad, 2005, p. 292). Medical images are generated using magnetic fields, radiofrequency, and computer software (Aydin, 2018). To generate such images, an experimental subject is scanned. Within a scanning session, "hydrogen protons in brain tissues are magnetically induced to emit a signal that is detected by the computer" (Alač, 2013, p. 1). The recorded results are being processed by computer models, mathematically analyzed, algorithmically enhanced and transformed into brain-mapping images (e.g., fMRI scans) (Ihde, 1998, p. 58).

Broadly speaking, fMRI is used to study and observe dynamic processes inside the brain that take place as a consequence of changes in blood flow (Joyce, 2006). Thus, the purpose of fMRI images is to show the degree of activity in different brain areas. If during the scanning session the experimental subject is involved in a specific cognitive task (for example, counting or recalling something), the fMRI can point out those regions of the brain which are most active during that task (de Boer, 2021). Unlike other medical imaging technologies, which rely on a single parameter within the image generation process "fMRI can use multiple parameters such as relaxation times, proton density, or diffusion of blood or other fluids for image production" (Prasad, 2005, pp. 298–299).

3.1. Hermeneutic mediation of the fMRI

Within an image generation process, fMRI accomplishes several hermeneutical tasks over its experimental object (e.g., human brain). In what follows, I will analyze them in the same order as they take place in the laboratory set-up.

Broadly speaking, the hermeneutic activity of fMRI can be generally divided into three parts. The first part is taking place inside the scanner when "hydrogen protons in brain tissues are magnetically induced to emit a signal" (Alač, 2013, p. 1). At this stage, a scanner generates a magnetic field that enables the recording of the brain activity. Within the second stage, the signal is being transformed into digital data that can be mathematically analyzed and recorded by a computer (Ihde, 2019). For evaluation reasons "the computer programs divide the section of the body into discrete, consecutive slices, and measure how long it takes for hydrogen atoms in each of these slices to release the energy absorbed from radio-frequency waves" (Joyce, 2016, p. 440). Thus, *the original data generated by fMRI is numeric, not visual*[10]. Finally, the last part of fMRI hermeneutic activity takes place when numerical measurements are converted via computer into a defined set of anatomical pictures (Prasad, 2005).

As it becomes apparent from the description above, several steps of material translation take place within the fMRI examinations. First, bodies are being changed into numbers[11]. Afterward, numbers are

---

[10] Generally speaking, having a numeric origin is true not only for images generated by the fMRI. All images produced by algorithms are images made by numbers (Mitchell, 1992; Couchot, 2002). It does not matter if the image comes from a camera on a mobile phone or from the computer (as in the case of fMRI). Every digital image – even the letters in a computer – is structured by bits – combinations of "zeros" (0) and "ones" (1). In fact, the image that results from bits is also composed of a set of points – the pixels – which are also numbers (fixed addresses on the screen to which colors are attributed).

[11] The question of body transformations through technologies might be fruitfully illustrated with references to the contemporary science fiction genre, namely, body horror. One of the main representatives of this genre in movies is Canadian film director David Cronenberg. Almost each of his movies dives into the ambiguous nature of the human body. For example, in the movie "Existenz" Cronenberg provides a deep philosophical interpretation of human bodies that permanently shift between 'the reality' and the computer game called Existenz. All the feelings and impressions in the computer game are so vivid that the main characters can't grasp the difference between 'real' and 'digital'. Said different, characters can't understand are they act as a real body (flash) or as a set of digits (bodies as numbers). That is why the movie ends up with the rhetorical question of one of the characters who asks: "Are we still playing?". The latest Cronenberg movie

transformed into images. Images are then interpreted by the laboratory practitioners[12] (Joyce, 2006). As a result, the imaged brain never shows itself within the laboratory settings as something "natural" or independent object (Alač, 2013). On the contrary, the brain goes through several hermeneutic transformations generated by fMRI technology. The latter 'selects' specific brain features which will get to the image and those that will not. In this sense, only certain brain traits that are relevant for the particular research, are transferred into the images. Moreover, only some of the many features that appeared on the fMRI visuals are desirable (De Rijcke & Beaulieu, 2014).

3.2. fMRI scans

As I hope is by now clear, fMRI scans mediate the object of representation (human brain) and 'show' it in a transformed manner. Moreover, scans make the brain visible in such a way, that a practitioner can 'read' it. However, to convert a human brain into a 'readable' image several transformations have to be made. In what follows, I will focus on the spatial and temporal transformations that fMRI scans possess.
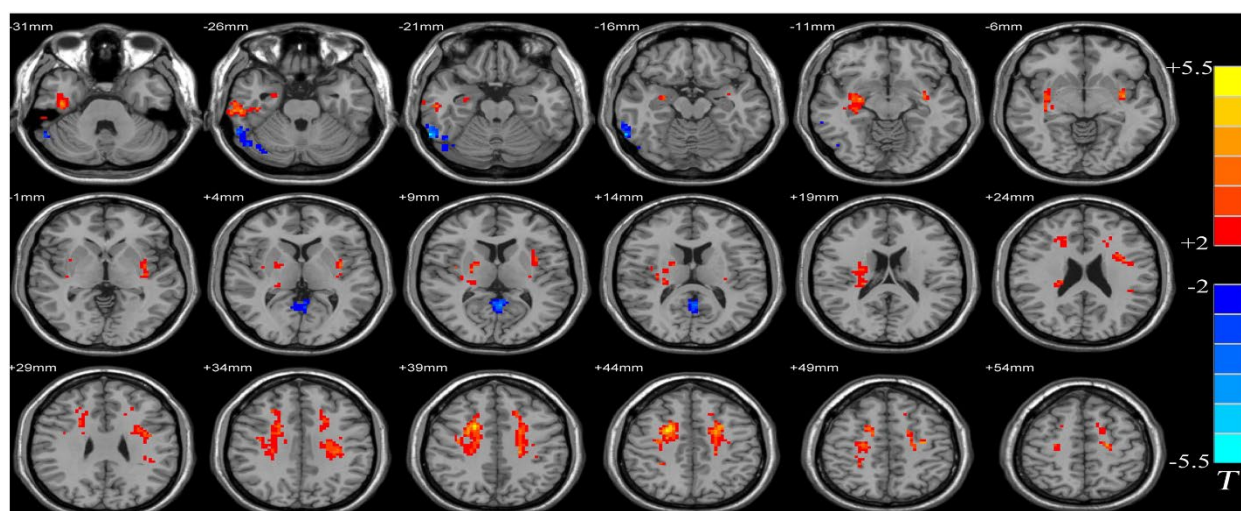
Figure 1. fMRI scans of patients with benign epilepsy with centrotemporal spikes.

Source - (Yan et al., 2017)

Spatial transformation

As the main purpose of the fMRI scan is to place brain activity in space (Beaulieu, 2016), it seems reasonable to focus on several spatial transformations initially. At first glance, one can see that the fMRI scan (either on the paper or on the screen) has a specific size. It means that the object of representation (e.g., the human brain) is 'framed' into the scan. Moreover, even though the representation proposes a high level of realism, the size of the brain on the screen does not coincide with the size of the brain in the patient's head (Verbeek, 2008). Moreover, the human brain is represented on the scan in a specific perspective (see figure 1). In scientific literature, this perspective is usually named a 'bird-eye perspective'[13] (Rosenberger, 2020). Of course, such a perspective isn't random or accidental but has functional purposes. The 'bird-eye perspective' provides the fMRI practitioner with a better 'view' of the human brain as a whole. From this perspective, practitioner can observe the human brain in its totality. This is what differs fMRI from other methods of brain study. Other techniques "make one small

---

"Crimes of the Future" questions the nature of the human body from a different perspective. Cronenberg describes the near future where people can't feel pain anymore. The main character is an artist who makes public surgeries on his body in order to put his body closer to the 'pain experience'. In this way the pain brings pleasure and the standard limits of the human body are shifting.

[12] The practitioner doesn't delete the image from the computer immediately. These data are usually preserved for several days just in case the doctor will come up with some new findings related to the patient's diagnosis. For a detailed analysis of the whole set of practices inside the radiology laboratory see Briedis, 2019, 2022.

[13] It is worth mentioning here that the 'bird-eye' perspective (also known as an 'upside down' perspective) isn't the only possible. There are some fMRI scans that can provide a view from the -right or -left sides of the brain.

measurement by probing the brain, whereas functional imaging can encompass the whole volume of the head" (Beaulieu, 2016, p. 73). Such a benefit opens up an opportunity to study a brain at a system level.

The other observation about the transformative role of the fMRI scans is that some parts of the brain are highlighted with different colors while others are left in a 'gray zone'. This happens because of the particular computer software, which allows an active interaction with the image data. The practitioner, for example, can "use the contrast between different shades of gray or can dynamically change the shades of gray to locate the pathology. This process is called *windowing*. The radiologist can also make comparisons by changing the contrast of gray in a particular region of the image through a process that is called *leveling*" (Prasad, 2005, p. 299). This effect mediates the practitioner's attention, attracting his/her attention to the parts highlighted with the color and at the same time decreasing the practitioner's attention from other parts of the brain (de Boer et al., 2020; Friis, 2008).

Temporal transformation

fMRI scans are placing brain activity not only in space, as in the case of spatial transformations analyzed above, but also in time. In this sense, fMRI scans mediate not only spatial but also a temporal representation of the human brain.

The temporal transformation takes place because fMRI image reduces the dynamic process of the neural brain activity into a frozen 'snapshot' (Rosenberger, 2005). This transformation provides fMRI practitioners with a unique opportunity to 'grasp' the brain-activation momentum and, by doing this, to understand which zone of the brain is responsible for a specific cognitive task (Cabeza & Nyberg, 2000). This temporal transformation is a 'game-changer' that gives birth to a whole 'cognitive revolution' in neuroscience (De Rijcke & Beaulieu, 2014). Philosophically speaking, this temporal transformation enables to reveal new scientific phenomena like "finding neuronal networks and brain regions that are specific to the functional realization of particular cognitive tasks" (Aydin & de Boer, 2020, p. 730).

Moreover, this temporal transformation enables a cross-referential process. It becomes possible due to the comparison between fMRI scans generated at different points in time. The intention of such comparisons is to monitor the evolution of a disease stepwise. For instance, a cross-referential process is used to follow up on the status of a cancerous lump in order to understand if the lump is decreasing, increasing, or remaining unchanged within different stages of the treatment (Prasad, 2005, p. 297).

fMRI hermeneutic strategies – noise reduction

As has been mentioned above, the hermeneutic strategies embrace a wide scope of different interpretative activities that a scientist has to produce within the image-interpretation process (Friis, 2008). In what follows, I will focus on one hermeneutic activity that is vitally important for fMRI technology and taking place in every fMRI laboratory. Its name is 'noise reduction' or 'cleaning the data'. Within this process, fMRI practitioner has to choose values for various parameters. Each parameter selected by the practitioner will influence what comes up to appear or disappear on the image (Beaulieu, 2016).

As I have already pointed out, before the image becomes accessible on the practitioner's screen, it should go through several technological transformations. These transformations increase 'readability' of the image and improve the practitioner's workflow. However, technology is not enough. There are many elements of the image-making process that fMRI can't accomplish without humans. Noise reduction is one of them (Goyal et al., 2018). Usually 'noise' in the dataset appears because fMRI 'adds' some undesirable elements into the image. In the vocabulary of fMRI practitioners these undesirable elements are named 'artifacts' (Lynch, 1985). Within fMRI scans artifacts can occur as white spots, black spots, wavy lines, or double-images of the various body parts (Joyce, 2016, p. 448).

In what follows I will focus on the so-called 'movement artifacts'. These artifacts appear in the picture because of the subject's movements during the scanning session (Alač, 2013). Usually, patients can't avoid small movements inside the scanner. These movements may cause some distortion in the image and lead to misdiagnosis or incorrect data analysis[14]. For this reason, the practitioners should identify the type of movement which the experimental subject generated inside the scanner so that they can immediately reduce noise from the data set (Prasad, 2005, p. 294). However, practitioners do not actually see the movement since the body lying in the scanner is no longer available for direct observation. This fact makes the practitioner's task truly 'hermeneutic'. The practitioner has to detect body movements that the experimental subject performed without *direct* access to his/her body but using computer visuals only.

For this reason, a practitioner has to apply a complex set of various hermeneutic strategies which might include laboratory knowledge and 'practices of seeing' like: image navigation (rotation, zooming-in/out), disclosing relations between different components on the image, prioritizing primary and secondary findings, technical talks with other practitioners (Briedis, 2019, 2022; Goodwin, 1995; Lynch, 1985b). These laboratory practices are actively synchronized with the fMRI visuals in a way that the practitioners can 'make' the movement of the subject visible on the brain scans (Alač, 2014). Thus, the movement does not just *appear* on the scans but is *produced* through a 'synchronization' of bodies and scientific instruments supplemented with cultural knowledge and laboratory practices. In this sense, 'noise reduction' represents a complex hermeneutic strategy within which a practitioner should 'select' what will be represented on the image and what should be omitted as 'undesired' artifacts.

Conclusion

The purpose of this paper was to contribute to the debates on the 'problem of representation' from the perspective of instrumental realism. To accomplish this task, I have developed an 'environmental framework' consisting of three main elements: imaging technology, scientific image, and hermeneutic strategies applied within an image interpretation process. On the theoretical level, the framework shows that to understand how the image receives its meaning one should take into account the whole representational environment where the image has been generated. Said differently, the image does not represent anything unless, at least, one element out of three is missing.

To show how these theoretical implications could be further applied to a practical domain I took the case of fMRI technology. At this stage of inquiry, I have analyzed all the steps of the image-making process produced by fMRI within the laboratory set-up. First, I have specified several hermeneutical tasks produced by this technology. Secondly, I have shown which spatial and temporal transformations fMRI scan renders over its object of reference (e.g., the human brain). Finally, I have focused on how the image receives its meaning within various hermeneutic strategies, which take place among practitioners in the laboratory environment.

Moreover, the framework developed through the current paper has other practical implications. As the environmental components defined through the framework are standard for any image-production process, the framework can be successfully applied to many other imaging technologies outside the medical domain. In this sense, the present article opens up new directions for further theoretical and practical findings within the problem of scientific representations.

Acknowledgement

---

[14] In this sense, the representation of the patient's body through fMRI visuals begs for the cooperation of the patient too. The patient has to keep still within the scanning session because even the slightest movement inside the scanner results in the production of undesirable artifacts.

References

Ackermann, R. (1988). Experiment as the motor of scientific progress. *Social Epistemology*, *2*(4), 327–335. https://doi.org/10.1080/02691728808578500

Alač, M. (2013). Handling Digital Brains. In *Handling Digital Brains*. The MIT Press. https://doi.org/10.7551/MITPRESS/9780262015684.001.0001

Alač, M. (2014). Digital Scientific Visuals as Fields for Interaction. *Representation in Scientific Practice Revisited*, 61–88. https://doi.org/10.7551/MITPRESS/9780262525381.003.0004

Amann, K., & Knorr Cetina, K. (1988). The fixation of (visual) evidence. *Human Studies 1988 11:2*, *11*(2), 133–169. https://doi.org/10.1007/BF00177302

Aydin, C. (2018). From Camera Obscura to fMRI: How Brain Imaging Technologies Mediate Free Will. In *Postphenomenological Methodologies* (pp. 103–121). Lexington books.

Aydin, C., & de Boer, B. (2020). Brain imaging technologies as source for Extrospection: self-formation through critical self-identification. *Phenomenology and the Cognitive Sciences*, *19*(4), 729–745. https://doi.org/10.1007/S11097-020-09667-1

Aydin, C., González Woge, M., & Verbeek, P.-P. (2018). Technological Environmentality: Conceptualizing Technology as a Mediating Milieu. *Philosophy & Technology 2018 32:2*, *32*(2), 321–338. https://doi.org/10.1007/S13347-018-0309-3

Beaulieu, A. (2016). Images Are Not the (Only) Truth: Brain Mapping, Visual Knowledge, and Iconoclasm: *Http://Dx.Doi.Org/10.1177/016224390202700103*, *27*(1), 53–86. https://doi.org/10.1177/016224390202700103

Briedis, M. (2019). Phenomenological ethnography of radiology: expert performance in enacting diagnostic cognition. *Phenomenology and the Cognitive Sciences 2019 19:2*, *19*(2), 373–404. https://doi.org/10.1007/S11097-019-09612-X

Briedis, M. (2022). Sensing Diagnostic Images: Skilful Embodied Cognition in Oncoradiology. *Sensing and Imaging*, *23*(1). https://doi.org/10.1007/S11220-021-00372-0

Cabeza, R., & Nyberg, L. (2000). Imaging Cognition II: An Empirical Review of 275 PET and fMRI Studies. *Journal of Cognitive Neuroscience*, *12*(1), 1–47. https://doi.org/10.1162/08989290051137585

Carusi, A., & Hoel, A. S. (2014). Brains, windows and coordinate systems. In *Visualization in the Age of Computerization* (pp. 145–169). Taylor and Francis. https://doi.org/10.4324/9780203066973-8/

Carusi, A., Hoel, S., Webmoor, T., & Woolgar, S. (2014). *Visualization in the age of computerization*. Routledge.

Coeckelbergh, M. (2020). *AI Ethics* . The MIT Press.

Collins, H. M., & Evans, R. (2002). The third wave of science studies: Studies of expertise and experience. *Social Studies of Science*, *32*(2), 235–296. https://doi.org/10.1177/0306312702032002003

Couchot, E. (2002). Digital hybridisation: A technique, an aesthetic. *Convergence*, *8*(4), 19–28. https://doi.org/10.1177/135485650200800403

de Boer, B. (2021a). Image Interpretation as Object Constitution: Hermeneutic Strategies in Neuroscientific Practice. In J. F. & R. Rosenberger (Ed.), *Postphenomenology and Imaging: How to Read Technology*. Lexington Books.

de Boer, B. (2021b). Explaining multistability: postphenomenology and affordances of technologies. *AI & SOCIETY 2021*, 1–11. https://doi.org/10.1007/S00146-021-01272-3

de Boer, B., te Molder, H., & Verbeek, P. P. (2020). Constituting 'Visual Attention': On the Mediating Role of Brain Stimulation and Brain Imaging Technologies in Neuroscientific Practice. *Science as Culture*, *29*(4), 503–523. https://doi.org/10.1080/09505431.2019.1710739

de Boer, B., te Molder, H., & Verbeek, P. P. (2021). Understanding science-in-the-making by letting scientific instruments speak: From semiotics to postphenomenology. *Social Studies of Science*, *51*(3), 392–413. https://doi.org/10.1177/0306312720981600

De Rijcke, S., & Beaulieu, A. (2014). Networked neuroscience: brain scans and visual knowing at the intersection of atlases and databases 1. In C. Coopmans, M. Lynch, J. Vertesi, & S. Woolgar (Eds.), *Representation in Scientific Practice Revisited* (pp. 131–152). MIT Press.

Devitt, M. (2005). Scientific realism. In F. Jackson & M. Smith (Eds.), *The Oxford Handbook of Contemporary Philosophy*. Oxford University Press. https://philpapers.org/rec/DEVSR

Devitt, M. (2011). Experimental Semantics. *Philosophy and Phenomenological Research*, *82*(2). https://doi.org/ppr201182222

French, S., & Saatsi, J. (2020). *Scientific realism and the quantum*. Oxford University Press.

Fried, S. J., & Rosenberger, R. (2021). *Postphenomenology and imagining : how to read technology* (S. J. Fried & R. Rosenberger (Eds.)). Lexington Books.

Friis, J. K. B. (2008). Interpreting Visual Culture: Explorations in the Hermeneutics of the Visual. *AI & SOCIETY*. https://doi.org/10.1007/S00146-020-01056-1

Friis, J. K. B. O. (2008). Enactive hermeneutics and smart medical technologies. *AI and Society*. https://doi.org/10.1007/S00146-020-00944-W

Friis, J. K. B. O. (2017). Gestalt descriptions embodiments and medical image interpretation. *AI and Society*, *32*(2), 209–218. https://doi.org/10.1007/S00146-015-0615-6

Fumerton, R. A. (2002). *Realism and the correspondence theory of truth*. Rowman & Littlefield Publishers.

Goodwin, C. (1994). Professional Vision. *American Anthropologist*, *96*(3), 606–633. https://doi.org/10.1525/AA.1994.96.3.02A00100

Goodwin, C. (1995). Seeing in Depth. *Social Studies of Science*, *25*(2), 237–274. https://doi.org/10.1177/030631295025002002

Goyal, B., Dogra, A., Agrawal, S., & Sohi, B. S. (2018). Noise issues prevailing in various types of medical images. *Biomedical and Pharmacology Journal*, *11*(3), 1227–1237. https://doi.org/10.13005/BPJ/1484

Hacking, I. (1983). Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. In *Representing and Intervening*. Cambridge University Press. https://doi.org/10.1017/CBO9780511814563

Hasse, C. (2008). Postphenomenology: Learning cultural perception in science. *Human Studies*, *31*(1), 43–61. https://doi.org/10.1007/s10746-007-9075-4

Heelan, P. A. (1989). *Space-perception and the philosophy of science*. University of California Press.

Hoeppe, G. (2014). Working data together: The accountability and reflexivity of digital astronomical practice. *Social Studies of Science*, *44*(2), 243–270. https://doi.org/10.1177/0306312713509705

Hoeppe, G. (2015). Representing Representation. *Science Technology and Human Values*, *40*(6), 1077–1092. https://doi.org/10.1177/0162243915594025

Hongladarom, S. (2012). Don Ihde: Heidegger's Technologies: Postphenomenological Perspectives. *Minds and Machines 2012 23:2*, *23*(2), 269–272. https://doi.org/10.1007/S11023-012-9296-9

Iconoclash: Beyond the Image Wars in Science, Religion, and Art. (2003). In B. Latour & W. N. Arnold (Eds.), *Leonardo* (Issue 4). MIT Press. https://doi.org/10.1162/LEON.2003.36.4.329

Ihde, D. (1998). *Expanding hermeneutics: Visualism in science*. Northwestern University Press.

Ihde, Don. (1978). *Technics and Praxis* (Vol. 24). Springer Netherlands. https://doi.org/10.1007/978-94-009-9900-8

Ihde, Don. (1991). *Instrumental realism : the interface between philosophy of science and philosophy of technology*. Indiana University Press.

Ihde, Don. (2009). From da Vinci to CAD and beyond. *Synthese*, *168*(3), 453–467. https://doi.org/10.1007/S11229-008-9445-0

Ihde, Don. (2019). Medical Technics. In *Medical Technics*. University of Minnesota Press. https://doi.org/10.5749/j.ctvpmw56v

Ihde, Don, & Selinger, E. (2006). Chasing Technoscience: Matrix for Materiality. *Human Studies*, *29*(3), 399–403.

Joyce, K. (2016). Appealing Images: Magnetic Resonance Imaging and the Production of Authoritative Knowledge. *Http://Dx.Doi.Org/10.1177/0306312705050180*, *35*(3), 437–462. https://doi.org/10.1177/0306312705050180

Joyce, K. A. (2006). From Numbers to Pictures: The Development of Magnetic Resonance Imaging and the Visual Turn in Medicine. *Science as Culture*, *15*(1), 1–22. https://doi.org/10.1080=09505430600639322

Knorr-Cetina, K. (Karin). (1999). *Epistemic cultures : how the sciences make knowledge*. Harvard University Press.

Kuchinskaya, O. (2013). Twice invisible: Formal representations of radiation danger. *Social Studies of Science*, *43*(1), 78–96. https://doi.org/10.1177/0306312712465356

Kudina, O. (2021). "Alexa, who am I?": Voice Assistants and Hermeneutic Lemniscate as the Technologically Mediated Sense-Making. *Human Studies*, *44*(2), 233–253. https://doi.org/10.1007/S10746-021-09572-9/FIGURES/3

Kudina, O., & de Boer, B. (2021). Co-designing diagnosis: Towards a responsible integration of Machine Learning decision-support systems in medical diagnostics. *Journal of Evaluation in Clinical Practice*, *27*(3), 529–536. https://doi.org/10.1111/JEP.13535

Latour, B., & Woolgar, S. (1986). *Laboratory life : the social construction of scientific facts*. Sage Publications.

Liberati, N. (2016). Technology, Phenomenology and the Everyday World: A Phenomenological Analysis on How Technologies Mould Our World. *Human Studies*, *39*(2), 189–216. https://doi.org/10.1007/s10746-015-9353-5

Liberati, N. (2019). Emotions and Digital Technologies. The Effects Digital Technologies Will Have on our Way of Feeling Emotions According to Post-Phenomenology and Mediation Theory. *Mente Journal of Philosophical Studies*, *36*(36), 292–309.

Liberati, N., & Nagataki, S. (2015). The AR glasses' "non-neutrality": their knock-on effects on the subject and on the giveness of the object. *Ethics and Information Technology*, *17*(2), 125–137. https://doi.org/10.1007/s10676-015-9370-0

Lynch, M. (1985). *Art and artifact in laboratory science*. Routledge.

Lynch, Michael. (1985a). *Art and artifact in laboratory science : a study of shop work and shop talk in a research laboratory*. Routledge & Kegan Paul.

Lynch, Michael. (1985b). Discipline and the Material Form of Images: An Analysis of Scientific Visibility. *Social Studies of Science*, *15*(1), 37–66. https://doi.org/10.1177/030631285015001002

Lynch, Michael, & Ridder-Vignone, K. de. (2014). Making the strange familiar: Nanotechnology images and their imagined futures. In *Visualization in the Age of Computerization* (pp. 97–117). Taylor and Francis. https://doi.org/10.4324/9780203066973-6/

Lynch, Michael, & Woolgar, S. (1988). Introduction: Sociological orientations to representational practice in science. *Human Studies 1988 11:2*, *11*(2), 99–116. https://doi.org/10.1007/BF00177300

Mitchell, W. J. (1992). *The reconfigured eye : visual truth in the post-photographic era*. The MIT Press.

Mykhailov, D. (2020). The Phenomenological Roots of Technological Intentionality: A Postphenomenological Perspective. *Frontiers of Philosophy in China*, *15*(4), 612–635. https://doi.org/10.3868/s030-009-020-0035-6

Mykhailov, D. (2021). A moral analysis of intelligent decision-support systems in diagnostics through the lens of Luciano Floridi's information ethics. *Human Affairs*, *31*(2), 149–164. https://doi.org/10.1515/humaff-2021-0013

Mykhailov, D., & Liberati, N. (2022). A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives. *Foundations of Science 2022*, 1–17. https://doi.org/10.1007/S10699-022-09833-5

Nørskov, M. (2015). Revisiting Ihde's Fourfold "Technological Relationships": Application and Modification. *Philosophy and Technology*, *28*(2), 189–207. https://doi.org/10.1007/s13347-014-0149-8

Prasad, A. (2005). Making images/making bodies: Visibilizing and disciplining through Magnetic Resonance Imaging (MRI). *Science, Technology, & Human Values*, *30*(2), 291–316. https://doi.org/10.1177/0162243904271758

Psillos, S. (2017). The Realist Turn in the Philosophy of Science. In *The Routledge Handbook of Scientific Realism* (pp. 20–34). Routledge. https://doi.org/10.4324/9780203712498-3

Roman, F., & Nguyen, J. (2021). "Scientific Representation." In *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=scientific-representation

Rosenberger, R. (2005). Bridging philosophy of technology and neurobiological research: Interpreting images from the "slam freezer." *Bulletin of Science, Technology, and Society*, *25*(6), 469–474. https://doi.org/10.1177/0270467605278878

Rosenberger, R. (2008). Perceiving other planets: Bodily experience, interpretation, and the mars orbiter camera. *Human Studies*, *31*(1), 63–75. https://doi.org/10.1007/S10746-007-9078-1/FIGURES/2

Rosenberger, R. (2009). A Case Study in the Applied Philosophy of Imaging: The Synaptic Vesicle Debate: *Http://Dx.Doi.Org/10.1177/0162243909337117*, *36*(1), 6–32. https://doi.org/10.1177/0162243909337117

Rosenberger, R. (2017). On the hermeneutics of everyday things: or, the philosophy of fire hydrants. *AI and Society*, *32*(2), 233–241. https://doi.org/10.1007/s00146-016-0674-3

Rosenberger, R. (2020). A Preliminary inventory of the transformations of Scientific Imaging. *Acta Baltica Historiae et Philosophiae Scientiarum*, *8*(2). https://doi.org/10.11590/abhps.2020.2.02

Saatsi, J. (Ed.). (2018). *The Routledge handbook of scientific realism*. Routledge .

Sturken, M., & Cartwright, L. (2017). *Practices of looking : an introduction to visual culture*. Oxford University Press.

Tibbetts, P. (1988). Representation and the realist-constructivist controversy. *Human Studies*, *11*(2–3), 117–132. https://doi.org/10.1007/BF00177301

van Fraassen, B. (1980). *The Scientific Image*. CLARENDON PRESS.

van Fraassen, B. C. (2008). Scientific Representation: Paradoxes of Perspective. In *Scientific Representation: Paradoxes of Perspective*. Oxford University Press. https://doi.org/10.1093/ACPROF:OSO/9780199278220.001.0001

Verbeek, P.-P. (2003). Material hermeneutics. *Techné: Research in Philosophy and Technology*, *6*(3), 181–184.

Verbeek, P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.

Verbeek, P. P. (2008). Obstetric ultrasound and the technological mediation of morality: A postphenomenological analysis. *Human Studies*, *31*(1), 11–26. https://doi.org/10.1007/s10746-007-9079-0

Wellner, G. (2020). *The multiplicity of multistabilities: Turning multistability into a multistable Concept* (pp. 105–122). Springer.

Wellner, Galit. (2020). Material hermeneutic of digital technologies in the age of AI. *AI and Society*, 1–8. https://doi.org/10.1007/s00146-020-00952-w

Whyte, K. (2015). *What is multistability? A theory of the keystone concept of postphenomenological research* (pp. 69–81). Lexington Books.

Yan, X., Yu, Q., Gao, Y., Li, L., Yu, D., Chen, Y., Yao, X., Yang, W., Chen, Z., Yin, J., An, Y., & Tan, K. (2017). Cognition in patients with benign epilepsy with centrotemporal spikes: A study with long-term VEEG and RS-fMRI. *Epilepsy & Behavior*, *76*, 120–125. https://doi.org/10.1016/J.YEBEH.2017.06.033

Zwart, H. (2022). Philosophy of Technoscience: From Cis-Continental to Trans-Continental. *Philosophy of Engineering and Technology*, *38*, 229–245. https://doi.org/10.1007/978-3-030-84570-4_8

VALUING ABIOTIC NATURE

PERSPECTIVES ON TERRAFORMING IN K.S. ROBINSON *MARS TRILOGY*[1]

VALORIZANDO A NATUREZA ABIÓTICA

Perspectivas sobre a terraformação na trilogia de Marte de K.S. Robinson

VALORACIÓN DE LA NATURALEZA ABIÓTICA

Perspectivas sobre la terraformación en la trilogía Marte de K.S. Robinson

Pierfrancesco Biasetti
(Leibniz Institute for Zoo and Wildlife Research, Germany)
(Space Philosophy research center, FISPPA, University of Padua)
biasetti@izw-berlin.de

ABSTRACT

In our everyday experience, life, environment, and nature are connected and we tend to confuse the value we assign to them. One way around this issue is to analyze our intuitions on the terraformation of other planets such as Mars. In this way, we are forced to consider whether the original abiotic nature has a value of some kind regardless of its capacity to support ecosystems and life, what kind of value this might be, and what weight it might have when compared with other values. In this contribution, I will draw a map of the possible answers to these questions by analyzing the different perspectives brought forth by some of the main characters in K. S. Robinson's *The Mars Trilogy*. In this way, it will be possible to observe that, while on Earth instrumental and non-instrumental kinds of environmental value generally concur and support each other, in an abiotic landscape, such as that offered (we assume) by Mars, they may conflict.

Keywords: terraforming. environmental value. value of abiotic nature. Kim Stanley Robinson. *The Mars Trilogy*.

RESUMO

Na nossa experiência cotidiana, a vida, o ambiente e a natureza estão ligados e temos tendência a confundir o valor que lhes atribuímos. Uma forma de contornar esta questão é analisar as nossas intuições sobre a terraformação de outros planetas, tais como Marte. Desta forma, somos forçados a considerar se a natureza abiótica original tem algum tipo de valor independentemente de sua capacidade de suportar ecossistemas e vida, que tipo de valor este poderia ser e que peso poderia ter em comparação com outros tipos de valor. Neste artigo, vou traçar um mapa das possíveis respostas a estas questões, analisando as diferentes

---

perspectivas trazidas por algumas das principais personagens de *A Trilogia de Marte* de K. S. Robinson. Desta forma, será possível observar que, enquanto na Terra os tipos de valores ambientais instrumentais e não instrumentais geralmente concorrem e se apoiam mutuamente, numa paisagem abiótica, como a oferecida (assumimos) por Marte, os mesmos podem entrar em conflito.

Palavras-chave: terraformação. valor ambiental. valor da natureza abiótica. Kim Stanley Robinson. *A Trilogia de Marte*.

<div align="center">RESUMEN</div>

En nuestra experiencia ordinaria, la vida, el medio ambiente y la naturaleza están conectados y tendemos a confundir el valor que les asignamos. Una forma de evitar esta cuestión es analizar nuestras intuiciones sobre la terraformación de otros planetas, como Marte. De este modo, nos vemos obligados a considerar si la naturaleza abiótica original tiene algún tipo de valor independientemente de su capacidad para sustentar ecosistemas y vida, qué tipo de valor puede ser y qué peso puede tener en comparación con otros valores. En esta contribución, dibujaré un mapa de las posibles respuestas a estas preguntas analizando las diferentes perspectivas aportadas por algunos de los personajes principales de la *Trilogía de Marte* de K. S. Robinson. De este modo, será posible observar que, mientras que en la Tierra los tipos instrumentales y no instrumentales de valor ambiental suelen coincidir y apoyarse mutuamente, en un paisaje abiótico, como el que ofrece (suponemos) Marte, ellos pueden entrar en conflicto.

Palabras clave: terraformación. valor ambiental. valor de la naturaleza abiótica. Kim Stanley Robinson. *La Trilogía de Marte*.

Introduction

Terraforming is the hypothetical transformation of a planet or a satellite to improve its capacity to support life, with Earth as its Platonic ideal. In this way, the ultimate goal of this process of planet engineering is to reshape a planet or a satellite to emulate the functions of Earth's biosphere as much as possible so that it can support in the open the lifeforms typical of our home planet with minimal modifications or differences (Fogg 1995).

Many celestial bodies in the Solar System have been proposed as the subjects for hypothetical terraforming operations. Although what was probably the earliest scientific contribution on the issue has been a proposal to planetary engineer Venus (Sagan 1961), it is on Mars that most of the research is now focused on. This should not come as a surprise since the Red Planet is the celestial body in the solar system whose conditions, as measured by the Earth Similarity Index (Schulze-Makuc et al. 2011), most resemble Earth. Even on Mars, though, terraforming poses highly complex and context-dependent scientific and engineering problems, the solution of which is usually beyond the current knowledge or technological capabilities. This is not to say that Mars terraforming is not seriously discussed in science. Indeed, a part of the scientific community believes that it will be possible to start engineering Mars in the medium or even short term and that it is hence necessary to start discussing not only the scientific and technical aspects of terraforming but also its ethical, social and political aspects (Beech et al. 2021).

Terraforming offers several opportunities for reflection for philosophers. For instance: what does it exactly mean to emulate the functions of the terrestrial biosphere? Given that a perfect replica of Earth is impossible to obtain, when can we properly say to have terraformed a celestial body? What will change in our self-representation as a species when, on terraformed planets, cultures, societies, and even people's bodies will start to diverge from the original Earthly models? Moreover, from a more practical standpoint, what is the ethical status of terraforming? Is it permissible, commendable, or even required—

or should we instead condemn it as morally execrable? Is the eventual conclusion generalizable to any terraforming project? What could be the conditions and criteria for terraforming a planet or a satellite responsibly?

These latter ethical questions require analyzing and unpacking several issues before attempting to formulate a tentative answer for them. A first issue often raised is that terraforming could destroy pre-existing autochthonous forms of life or their traces (McKay & Zubrin 2002). This eventuality, however, would be rare indeed. Nevertheless, since the present best candidate for terraforming is Mars, a planet which, according to some scientists, could present—or could have presented—the theoretical conditions for the development of life as we know it (Schulze-Makuch et al. 2008), the argument is worth to be debated.

From another perspective, the sheer size of the efforts required by terraforming and their underlying ambitions may raise several ethical questions. Given the vastness of the resources required, the complexity of the processes set in motion, and the impossibility of predicting with certainty the outcomes, terraforming is often labeled as a manifestation of hubris, and, hence, as morally condemnable (Sparrow 1999, 2015). However, this argument is often rebutted as uncharitable, because its supporters tend to exaggerate and generalize the risks associated with terraforming and underestimate the value that the knowledge accumulated through research on this topic can have for better environmental management of our planet (Schwartz 2013).

One of the interesting aspects of space ethics is that the situations it explores need to be analyzed in two complementary ways: a first way, concerning the applications of our theories, principles, and moral intuitions to the unusual conditions and scenarios involved in space activities; a second way, concerning the repercussions of these space activities on Earth. In this sense, terraforming, should it become feasible, could lead to a change in the way we value our home planet. The claim that "there is no planet B", for instance, would immediately lose meaning if Mars, Venus, the Moon, or the poles of Mercury would get the chance to become hospitable environments. This would create a moral hazard and it could push us to act even more recklessly than we are currently doing concerning the climate and the environment of Earth. On the other hand, terraforming research could instead provide us with new opportunities, knowledge, and technologies to solve some of the most insidious challenges that lie ahead: planetary engineering could help us in mitigating the effect of the environmental crisis; a new planet at our disposal could provide us in the long term with a relief valve for demographic pressure. And so on.

In a further sense still, terraforming can be an incredibly rich source of thought experiments to investigate some philosophical nodes that can be hard to disentangle when observed from an Earthly standpoint. Concerning environmental philosophy, for instance, terraforming examples can be used to shed light on: (a) the difference between various kinds of environmental value; (b) the value or values we should attach to abiotic nature.

In our common earthly experience, life, the environment, and nature are intimately connected concepts. Because of this connection, we tend to conflate the values we attach to them, sometimes even to the point of totally losing sight of their difference. In other words, we tend to consider our valuing living beings, the complex system that supports their existence, and naturalness as part of the same experience—or even as the same experience. The conditions posed by terraforming scenarios, instead, make us possible to separate between these different concepts and the values that accompany them. This is particularly important when it comes to analyzing the value to be assigned to abiotic nature. On Earth, abiotic nature is usually an essential component of the environment and, as such, it is intrinsically integrated with life: it is its support and precondition and maintains with life a constant metabolic exchange. For this reason, from our terrestrial perspective, it is hard, if not impossible, to analyze the value per se of abiotic nature (York 1995). Terraforming scenarios ask us instead to imagine the transformation of "pure" abiotic nature (something which is not an environment) into a new landscape capable of hosting life. In this way, we are forced to reflect on whether the original abiotic nature has a

value of some sort independently of its value in supporting ecosystems and life, which kind of value this could be, and how much it would weigh when compared to other kinds of values.

In what remains of this contribution I will try to draw a map of the possible answers to this question, and I will do so by analyzing the different perspectives on terraforming carried out by some of the main characters of Kim Stanley Robinson's *Mars Trilogy*.

Perspectives on terraforming in the *Mars Trilogy*

The books composing the *Mars Trilogy* are *Red Mars*, *Green Mars*, and *Blue Mars* (henceforth: *RM*, *GM*, and *BM*, respectively), with each color representing a stage of the terraforming process of the planet: red like the pristine surface on which the settlers land; green like the landscape dominated first by algae and lichens then by plants when the process of ecopoiesis is started; blue like the water that finally fills Mars with oceans, rivers, and lakes. Despite the vast time span involved, most of the recurring characters belong to the "First Hundred", the early settlers landed on the planet whose life expectancy is incredibly extended following the discovery of anti-aging treatments.

The scientific, social, and political aspects of the Martian colonization constitute essential elements of the plot of the trilogy. It is among the First Hundred that the debate on terraforming Mars is first started and their original perspectives shape much of the politics and values expressed by later settlers and their descendants. Selected from the best scientists and engineers on Earth, the First Hundred become immediately divided over the meaning of their presence on the Red Planet. Some see it just as a unique occasion for engaging in research, without further complications. Others want to experiment with the planet, even if it means to change completely its natural conditions. Others see the mission as an opportunity for social engineering and want to build a radically new society. Others still would content themselves to exploit the extraordinary resources of Mars.

At least nominally, only the first of these course of actions—researching without interfering—would be permitted to them by the Earth's authorities, because the mission rules are modeled according to those of scientific expeditions carried out in the Antarctic. However, the radical uniqueness of their situation—in terms not only of distance from the terrestrial authorities but also of the harshness of the condition they are exposed to—ends up pushing the settlers to take their destiny into their hands, igniting a debate on the needs to proceed quickly with terraforming Mars.

In this way, two radically distinct visions of the Red Planet come to clash. On the one hand, Mars as the pristine planet—with its natural history as much as old and important as that of Earth, its unique scientific mysteries to be discovered and solved, and its equally unique beauty to be explored and contemplated. On the other hand, Mars as the raw material for creating a new environment—a new home for humanity, and a cradle for the future society of the settlers and their descendants. The first vision of Mars is presented as incompatible with terraforming. The second one sees terraforming as the natural outcome of the expedition. Several arguments are advanced in the trilogy advocating for one or the other of these two positions.

a) *Ann Clayborne*. Opposition to the terraforming plans finds its stubborn champion in the head of the geological expedition, Ann Clayborne. Ann wants to preserve Mars in its pristine form, avoiding every kind of alteration of the natural conditions on the planets even if this actually translate into a severe downsizing of the colonization attempt—as emigration from Earth will be necessarily limited, settlers will have to live into cramped underground environments, and every visit on the surface will expose them to radiations.

The initial argument she uses to defend her position is the possibility of indigenous life existing at the microbial level. Any alteration initiated by the terraforming process would almost certainly destroy this life—something that Ann judges wrong both from a scientific and a moral standpoint. Nevertheless, she does not seem really committed to this argument, advocating it perhaps only for tactical reasons—that

is, because it seems to have the greatest hold on the imagination of people. When her major interlocutor and opponent, the brilliant yet socially awkward physicist Saxifrages "Sax" Russell, points out that it is basically an argument from ignorance—since, while there are no traces of life on the surface, there is no way to prove or disprove a similar result underground—Ann just drops this line of arguing, without dropping, however, her anti-terraforming commitment. This leads to one of the decisive moments in the plot of the whole *Mars Trilogy*: the public debate between Ann and Sax on terraforming. The debate takes place in the main hall of the first settlement on Mars, in the presence of at least half of the First Hundred, and it is filmed and broadcasted on Earth. It becomes a turning point in the history of Mars colonization, as it marks the ideological birth of two of the main political factions of Martian society, the "reds" opposing terraforming and the "green" favoring it.

During the debate, Ann articulates her real arguments against terraforming—those that she will reiterate during much of the trilogy. The first argument is based on the scientific value of a pristine Mars. Terraforming would destroy much of its geological record and other scientific data that could be used to better understand the planet and the solar system, and, for this reason, it should not be started. By making clear the scientific reasons backing her value assumption, Ann also attacks the kind of science proposed by Sax. Terraforming, in her view, is not true science: it can be compared to the actions of a young kid playing with his chemistry set in a basement, only on a planetary scale. It is just a form of scientific narcissism and has nothing to do with proper science, which should be instead interested in studying the planet as it is, with all the evidence regarding its history and the origins of the solar system in place.

Ann then raises two other arguments. The first one is based on the aesthetic value of Mars. Terraforming would destroy the Martian beautiful and pure landscape, and it would do so, again, for shallow reasons. The second one is based on what we could call the "existential" value of the planet—that is, the value coming from its alterity, its difference from ourselves and Earth. Terraforming would destroy this difference, and hence its value, as it would amount to the transformation of another piece of the universe into a mirror image of ourselves. In a previous dialogue Ann had already raised this issue:

> We'll all go on and make the place safe. Roads, cities. New sky, new soil. Until it's all some kind of Siberia or Northwest Territories, and Mars will be gone and we'll be here, and we'll wonder why we feel so empty. Why when we look at the land we can never see anything but our own faces (*RM*, p. 190)

By removing radiation, lack of breathable atmosphere, and extremely cold weather we remove from Mars, in Ann's opinion, something just as important in our life as safety: we remove differences.

b) *Saxifrages Russell*. In the first book of the trilogy, Sax plays the role of the terraforming enthusiast. He secretly starts the terraforming process without waiting for the approval of both the Earth's governments and of his fellow First Hundred. He does not seem to care for the opinions of others if he believes that they do not conform to rationality—to his rationality, and he seems to not genuinely grasp the difference. His mentality and positions will dramatically change during the rest of the trilogy, as he will stubbornly struggle to understand Ann's position and reconcile with her. During the first debate, however, his arguments straightforwardly support terraforming.

For Sax, terraforming has many different values—all of them more important than the value offered by a sterile planet like Mars when the first settlers arrive. First thing, terraforming, according to Sax, is a grandiose scientific experiment, with unique and unrepeatable opportunities for knowledge and technological improvement. Moreover, terraforming would enhance Mars' beauty, by enriching its landscape with lakes, forests, and glaciers, and by adding life, "the most beautiful system at all". Third, it would offer the opportunity for propagating conscious life, which Sax considers the most valuable thing at all—a claim that he justifies with the fact that it is only consciousness that provides the universe with meaning. "The lack of life here", he states,

> and the lack of any finding in fifty years of the Seti program, indicates that life is rare, and intelligent life even rare. And yet the whole meaning of the universe, its beauty, is contained in the consciousness of intelligent life. We are the consciousness of the universe, and our job is to spread around, to go look at

> things, to live everywhere we can. It's too dangerous to keep the consciousness of the universe on only one planet, it could be wiped out (*RM*, p. 213)

These words will mark the turning point in the debate, now increasingly in favor of Sax. Ann, however, will be not convinced by the argument. She will reiterate her idea that we need alterity—something completely different from us—to not feel empty, and she will claim that our being the consciousness of the universe does not authorize us with transforming it into our mirror image. Yet she will also feel defeated by Sax's argument—a feeling shared by the presents and by the public watching the broadcasted debate.

c) ***Hiroko Ai***. A third perspective on terraforming and the value of Mars is offered by the First Hundred's lead biologist, Hiroko Ai. For Hiroko, Mars is an opportunity to spread life to another planet beyond Earth: life in its broadest possible sense, at all biological levels. This biocentric impulse arises in Hiroko from the belief in a life force—which, following XII century mystic and scientist Hildegarde von Bingen, she calls *viriditas*—that permeates the universe.

> A constant pressure, pushing toward pattern. A tendency in matter to evolve into ever more complex forms. It's a kind of pattern gravity, a holy greening power we call *viriditas*, and it is the driving force in the cosmos (*GM*, p. 20-1).

Hiroko's beliefs translate into a favorable attitude toward terraforming if this is understood first and foremost as the construction of a complex life-support system, whose goal is not so much to ensure the survival of the human species as such, but that of the system itself. Similarly, Hiroko's approach to terraforming is, unlike Sax's, strongly holistic and will gradually turn mystical and quasi-religious.

d) ***Other positions***. While Ann, Sax, and Hiroko provide the three main intellectual perspectives on terraforming, at the extreme of the debate other two positions can be identified—two opposite positions. The first one is claimed by the members of the political faction opposing terraforming—the "Reds". According to this position, Mars is to be respected and left as it is due to its intrinsic value. Many Reds will try to enforce this position by any means, including the use of terrorism, and their most extreme faction will even go so far as to wage civil war on Mars.

On the opposite spectrum, some characters in the *Mars Trilogy* support terraforming to advance their agendas. This latter position is endorsed by very different people, with very different sets of values: those of rapacious capitalism (Phyllis Boyle, geologist and businesswoman), of liberal and green capitalism (William Fort, CEO of Praxis, a "progressive" and "enlightened" transnational corporation), and even of revolutionary socialism and anarchism (Arkady Bogdanov, mechanical engineer and revolutionary).

Phyllis Boyle and William Fort support terraforming for its economic value, albeit understood from different perspectives. Arkady Bogdanov, on the other hand, wants to terraform Mars because it is the only way to emancipate the colonists from Earth and build a society on new foundations. Besides differences in their ultimate reasons, what these positions share is a more or less explicit adhesion to anthropocentrism, which inevitably overshadows any consideration of the value of Mars in its pristine state.

Valuing abiotic nature

The different values attributed to Mars can be organized through two axes (see **Table 1**).

The first axis (instrumental/non-instrumental) is related to *utility*: whether the attributed value is pursued in view of some sort of utility, or, if it is instead related to something like "beauty" or "knowledge" which, while it can *indirectly* provide people with some sort of utility, it is primarily pursued for its own sake.

The second axis (anthropocentric/non-anthropocentric) is related instead to the *locus* of the attributed value: more specifically, whether the value is human-focused or it is not.

Table 1: Value of Mars		
	Non-anthropocentric	*Anthropocentric*
Non-instrumental	*Intrinsic value* ("Reds")	*Scientific, aesthetic, and existential value* (Ann)
Instrumental	*Life-support* (Hiroko)	*Intelligent life-support* (Sax); *Mars as a new opportunity* (Phyllis, William, Arkady)

The dividing line between values supporting or opposing the terraforming of Mars lies on the instrumental/non-instrumental axis. This should not come as a surprise. If Mars is considered valuable beyond its "utility-value", then terraforming might not be an option. Otherwise, if its value is only in its use, terraforming seems the most reasonable choice, as this utility-value can be realized only by transforming it in an Earth-like planet.

Interestingly, there can be instead values for and against terraforming in every sector demarcated by the anthropocentric/non-anthropocentric axis. This is because the "utility" created by terraforming does not necessarily have to benefit our species, and some non-instrumental values such as beauty or knowledge, while not advocating for terraforming, remain nevertheless human values.

Each cell in **Table 1** will now be briefly analyzed.

a) ***Intrinsic value***. Intrinsic value is a complex concept that in environmental ethics has been used to describe different possible kinds of value attributed to nature (O'Neill 1992; Jamieson 2008). Here it is defined through two characteristics: it is a form of value that, at the same time, is not directly useful to anyone (it is not-instrumental), and it is not focused on our species (it is not-anthropocentric).

Many environmental theories incorporate intrinsic value. Ecocentric theories, for instance, consider large holistic natural complexes such as ecosystems to be intrinsically valuable. This, however, does not authorize us to conclude that the abiotic nature embedded in these systems could be equally intrinsically valuable. In ecocentric theories, wholes usually take precedence over components, and it is, therefore, reasonable to think that abiotic nature, being only a part of a complex system, can only receive instrumental value and perhaps even no value at all if separated from an ecosystem. For this reason, some authors believe that a Mars populated by life forms—even if allochthonous ones—would have a greater value than a pristine yet sterile planet, and they are ready to support terraforming on these grounds (McKay & Marinova 2001). Other authors, instead, have taken the opposite path, and have provided arguments to extend the ecocentric perspective to assign intrinsic value to abiotic nature independently of its being part of an ecosystem (Rolston 1986, Lee 1994). The possibility of assigning intrinsic value to the "mineral" dimension of nature has also been considered, albeit somewhat cursorily, within the debate on the concept of geodiversity (Gray 2005).

In the specific case of Mars, assigning intrinsic value to the planet means valuing it independently from its specific qualities—its possible beauty, its history, its capacity to become a new home for our species, etc. Attributions of intrinsic value are usually decisive when it comes to weight different reasons in a balancing—although not a rule, considerations grounded in intrinsic value usually trumps considerations based on non-intrinsic forms of value. This is well represented in the *Mars Trilogy* by the self-righteousness with which many "Reds" operate.

The main problem with intrinsic value attributions is that they are difficult to justify for holistic entities, to whom attributes traditionally considered morally relevant such as sentience, conscience, etc., cannot be ascribed. In these contexts, the attribution of intrinsic value risks appearing without solid rational justification, as something mystical or overtly religious. This aspect is again well represented in the *Mars Trilogy*. Throughout the story, many characters attribute to Ann a position based on intrinsic value. This includes many "Reds" who look at her for leadership and inspiration, as well as Sax himself, who, at one point, believe to have finally grasped her position, identifying it as Martian version of Aldo Leopold's Land Ethics—rather a strange identification, given the absence of an autochthonous biotic community on the Red Planet. Ann, however, does not support an intrinsic value theory of Mars. When the narration takes place from her point of view, we learn that she feels uncomfortable when facing arguments based on intrinsic value. She finds this concept to have little comprehensibility, and judges it (and those inclined to adopt it) too near to mysticism and religion. While she never openly disagrees with those" Reds" who espouse intrinsic value theories—probably to not weaken the front opposing terraforming—she finds their stance incompatible with sound reason and science.

b) ***Scientific, aesthetic and existential values***. Nature can be a source of significant experiences: scientific curiosity and knowledge, aesthetic beauty, reverence for otherness, and diversity. These kinds of experiences, while having a disinterested character—they do not provide us with something directly useful—are nevertheless human experiences, and as such the value that we attach to them can be defined, at the same time, anthropocentric yet non-instrumental.

From this perspective, it can perhaps be argued that a terraformed Mars might be a better source of meaningful experiences for people than a bare, uninhabitable planet. After all, a terraformed Mars would be more accessible to people, richer in possibilities for experiences—hosting life forms—and ultimately more resonant with our evolutionary baggage. The terraforming process itself could be a source of significant experiences, especially in science. This aspect is captured by Sax when he uses just such arguments to defend his terraforming projects. Moreover, throughout the narrative arc, the same Sax will never cease to marvel at the scientific and aesthetic developments of the process. The problem with these attributions of value, however, is that an important if not necessary aspect of this kind of experience is missing: authenticity. One of the characteristics that make nature-related experiences—especially aesthetic ones—unique is authenticity, that is, the fact that our experience of nature is precisely an experience of nature, of something that is not under our direct control or that we have shaped in a significant way.

The significant experiences that can be linked to a terraformed Mars or to the terraforming process itself are hence not the same as those that can be associated with a Mars in its original condition, and they are not necessarily better. In front of a non-expert eye, the Martian landscape may appear repetitive and inanimate, but its aesthetic value may instead manifest itself in all of its beauty to someone capable of recognizing and appreciating the "mineral" regularities hidden in its apparent monotony (McMahon 2016). Deciding at this point which of the two aesthetic perspectives is to be preferred is no longer just a matter of taste, but of safeguarding diversity: because whereas we already have a terrestrial landscape here on our planet, we do not have another Martian landscape besides the original one.

This may explain Ann's insistence throughout the trilogy's narrative arc on the loss of aesthetic and scientific value caused by the terraforming of Mars, but also on the loss of this element of radical otherness. "You don't know Mars" is the cold remark she repeats to Sax every time he tries to convince her that a terraformed Mars could be more beautiful than the planet in its original state. For Ann, the question is also existential. Elaborating from her perspective, we could say that terraforming Mars would be in a sense like transforming it into what Marc Augé has called a "non-lieux" (Augé 1992), that is, an anonymous place devoid of history and possibility for contingency. This could certainly serve our immediate interests, but ultimately it will leave us unhappy—a most obvious and inevitable effect of the homogenization brought about by terraforming.

c) *Life support*. Nature can be an environment: it can support life. This can be valuable both from biocentric and ecocentric perspectives. Despite this similar starting point, however, the two perspectives differ when it comes to the concept of the kind of life that is valuable. For biocentric theories, the kind of life that is valuable is individual life. To be alive is the crucial feature for being morally considerable, and biocentrism, for this reason, holds that each individual existence is valuable per se. For ecocentric theories, on the other hand, the kind of life that is valuable is the biotic community. The individual existences that made up this community participate in its value, but, at the same time, they are only valuable because they are parts of it. At their simplest, biocentric theories may have trouble ordering conflicts between the interests of different living creatures. This problem does not plague ecocentric theories, for which the individual interest that deserves to prevail is always the one that contributes most to the community interest. In this way, however, ecocentric theories, at least in their simplest version, can be accused of not respecting individuals, and sacrificing them, if necessary, to the community.

Despite these differences, both biocentric and ecocentric views share an instrumental view of abiotic nature. Inside the frame of these theories, the only ethically relevant element that can be linked to abiotic nature is its life-support capacity—and it is indifferent here whether we understand life in a holistic or individual sense. For this reason, except for opportunity value, no other value can be attached to sterile abiotic nature.

It should be noted that while the life-support value attached to abiotic nature is instrumental, it is nevertheless not anthropocentric. Its focus is not restricted to human life but embraces all kinds of other species—including non-animal species. It is an instrumental, yet non-anthropocentric value.

This is essentially the kind of value that Hiroko's perspective assigns to Mars. The Red Planet, for her, is only the stage in which a new chapter in the story of the expansion of life takes place. What is intrinsically valued in this view is the spreading and evolution of life. Mars—and all abiotic nature in general—as value only has the empty canvas on which to carry out this process—as the substrate and the raw material for a new creation.

d) *Intelligent life support*. Intelligent life is considered valuable for multiple reasons, as it is connected with the possibility of experiencing numerous cognitive and sensory states to which we attach particular importance. For some people, disseminating, or even simply defending, the existence of this intelligent life can be a value, leading to the ethical imperative to defend and promote the existence of our species. In the long term, this means developing the capacity for our species to abandon the Solar System. As a necessary step on the long quest for this goal, colonization of other planets in our home system will be requested—and Mars is probably the best place to start.

In this perspective, abiotic nature is valued instrumentally—by virtue of its capacity to host life—and in an anthropocentric way—because the kind of life we consider morally relevant is intelligent life, which coincides, at least within our present reach and knowledge, with our species.

This is the position endorsed by Sax, and it is justified, in his opinion, by the fact that only conscious intelligence gives meaning to the universe, and it is its apex of complexity, being the least entropic entity it contains. This perspective, of course, supports terraforming: nevertheless, as Sax learns during the *Mars Trilogy*, it does not support it necessarily as the reckless and hubristic process he argued for at the beginning.

e) *Mars as a source of opportunity*. In this category are collected various considerations of instrumental and anthropocentric nature that, while adhering to the previous idea that the value of abiotic nature resides in its capacity—actual or potential—to benefit human life, interpret this latter expression not as "to benefit the existence of the species", but as "to benefit people". Abiotic nature, in this way, is seen as valuable for the opportunities it can grant to individuals. The identity of these individuals and the nature of the benefit is determined by how this opportunity value is specified. The *Mars Trilogy* offers several examples of values in this regard. Both Phyllis Boyle and William Fort, for instance, seem to value Mars for the economic opportunities it offers. Concerning the different ways in which they

conceive the economic exploitation of the planet, it is interesting to note that the character more in favor of fully terraforming Mars—William Fort—is the one more attentive to its sustainable development, precisely because he sees in the creation of a whole new world a more durable and profitable venture than the rapacious stripping of its raw resources. The views of Arkady Bogdanov offer another example of this kind of this opportunity value attached to Mars. For him, the Red Planet is an opportunity to build from scratch a new society with utopian leanings.

Conclusions

Eventually, in the books of the trilogy, Mars undergoes terraforming, becoming something unique and new compared to Earth, and, at the same time, something that inevitably ends up resembling our home planet at the expense of the original pristine landscape. Politically and socially, terraforming is not described as a peaceful and uncontroversial process. It undergoes acceleration and decelerations, and passes through revolutions and civil wars, until the original Mars, excluding its considerable heights where the atmosphere is still too thin, is transformed. This brutal fact, however, does not provide an answer to the question of whether it was right or wrong to terraform the Red planet in the first stance.

Some characters mature and modify their starting views over the two centuries-long narrative arc. Sax is perhaps the character that undergoes the most spectacular change. In his struggle to make sense of Ann's ideas, he will come to understand the beauty and scientific importance of Mars, and he will regret the decision to recklessly start the terraforming process. This will not make him abandon his support for the ultimate goals of terraforming—as he will remain faithful to his main argument, the necessity to spread conscious intelligent life into the cosmos. He will nevertheless try to find a synthesis between his and Ann's positions, rejecting "invasive" terraforming methods, and advocating the establishment of large "wild" areas in the highest parts of the planet, limiting the creation of an atmosphere and the spreading of life to the plains and lowest altitudes.

Very specific positions on environmental ethics have been attributed to Robinson's work in the *Mars Trilogy*—for instance, it has been said that it is an attempt of drawing the contours of a Martian Land Ethics, in analogy with the "earthly" Land Ethic of Aldo Leopold (Otto 2003). As is often the case, it is difficult to rigidly classify an author's vision, especially if it is spread across something less than two thousand pages, and voiced through several characters with diverging agendas. However, there is no doubt that the plot of the trilogy seems to indicate the necessity of a synthesis of some sort between the different perspectives advocated by its main characters. In other words, from the *Mars Trilogy* seems to emerge a claim that a responsible terraforming process can be achieved by trying to bring together the best elements of the main perspectives on the value of Mars. The very ending of the last book would seem to indicate this solution, ideally bringing together the characters representing each of the three main conceptions (if we accept that the Asian woman seen surfing on the beach in the last pages of *BM* may be or perhaps, more realistically, may symbolize, Hiroko). How to arrive, in the concrete, at such a synthesis, however, is far from obvious, given the divergence between the different values at stake. On this point, the novels seem to insist that a synthesis cannot be obtained in advance through theoretical compromise, but must be built along a tortured path through the accidents of history.

Returning to our main point—using the case of the terraforming of Mars as a hypothetical scenario to untangle some questions inherent to environmental philosophy—Robinson's *Mars Trilogy* has the merit of drawing a detailed map of the major arguments that can be deployed for or against engineering a planet to transform it into an Earth-like body, providing us in this way with important insights into how value can be assigned to abiotic nature.

In particular, imagining a scenario such as the terraformation of Mars allows us to understand how what is often conflated together and called environmental value can refer to very different realities: living individuals of other species besides ours, the environment, understood as the set of elements that interact with living individuals allowing their survival and reproduction, and nature, understood instead as the other respect to the transformative action of our species—that is, according to its meaning as the opposite

of "artificial". Giving value to one of these aspects can often mean giving value to the other aspects, but it is not necessarily always the case: in fact, sometimes the ethical demands connected with these different concepts can conflict.

This becomes evident if we analyze the values that can be attached to Mars—and by extension to all abiotic nature. The instrumental values push us to transform this abiotic nature into an environment, modifying its original naturalness in the process. Non-instrumental values, on the other hand, push us to preserve the original natural element, even at the expense of the possibility of propagating human and non-human life. In the terrestrial context, this conflict does not usually arise since the abiotic nature is almost invariably tied to the environment where living creatures thrive, and, therefore, there is no contrast between its conservation and that of the ecosystems that depend on it. In the case of non-terrestrial nature, this is not the case: and every present and future terraforming ethic will have to start from this contrast to construct its proposal.

References

Augé M. 1992. *Non-lieux. Introduction à une anthropologie de la surmodernité*. Paris: Le Seuil

Beech M, Seckbach J, Gordon R. 2021. *Terraforming Mars*. Wiley-Scrivener.

Fogg MJ. 1995. *Terraforming: Engineering Planetary Environments*. Warrendale, PA: Society of Automotive Engineers

Gray M. 2005. *Geodiversity and Geoconservation: What, Why, and How?* «The George Wright forum» 22: 4–12.

Jamieson D. 2008. *Ethics and the Environment. An Introduction*. Cambridge: Cambridge University Press.

Lee K. 1994. *Awe and Humility: Intrinsic Value in Nature. Beyond an Earthbound Environmental Ethics*. In: Attfield R and Belsey A (eds.), *Philosophy and the Natural Environment, Royal Institute of Philosophy Supplement* 36: 89–101. Cambridge, UK: Cambridge University Press.

McKay CP, Marinova MM. 2001 *The Physics, Biology and Environmental Ethics of Making Mars Habitable*. «Astrobiology» 1(1): 89–109

McKay CP, Zubrin R. 2002. *Do indigenous martian bacteria have precedence over human exploration?* In: *On to Mars: Colonizing a New World*, New York: Apogee Books Space Series

McMahon S. 2016. *The Aesthetic Objection to Terraforming Mars*. In: Schwartz J, Milligan T (eds.) *The Ethics of Space Exploration*. Cham: Springer, Cham.

O'Neill J. 1992. *The Varieties of Intrinsic Value.* «The Monist» 75(2): 119-137

Otto E. 2003. *Kim Stanley Robinson's Mars Trilogy and the Leopoldian Land Ethic*. «Utopian Studies» 14(2): 118–135.

Robinson KS. 2009a [1992]. *Red Mars*. London: Harper Collins.

Robinson KS. 2009b [1992]. *Green Mars*. London: Harper Collins.

Robinson KS. 2009c [1996]. *Blue Mars*. London: Harper Collins

Rolston H III. 1986. *The Preservation of Natural Value in the Solar System*. In: Hargrove E (ed.), *Beyond Spaceship Earth: Environmental Ethics and the Solar System*, 140–82. San Francisco: Sierra Club Book

Sagan C. 1961. *The Planet Venus*. «Science». 62: 849–858

Schulze-Makuch D, Fairén A, & Davila A. 2008. *The case for life on Mars*. «International Journal of Astrobiology», 7(2): 117-141.

Schulze-Makuch D, Méndez A, Fairén AG, Von Paris P, Turse C, Boyer G, Davila AF, António MRDS, Catling D, Irwin LN. 2011. *A two-tiered approach to assessing the habitability of exoplanets*. «Astrobiology» 11:1041–1052.

Schwartz JSJ. 2013. *On the moral permissibility of terraforming*. «Ethics and the Environment» 18: 1–31.

Sparrow R. 1999. *The Ethics of Terraforming*. «Environmental Ethics». 21: 227–245.

York P. 1995. *The Ethics of Terraforming*. «Philosophy Now». 1–7.

ENTRELAÇANDO DESIGN E FICÇÃO CIENTÍFICA[1]

O CASO DE DALEKO

RE-ENTANGLING DESIGN AND SCIENCE FICTION
The case of Daleko

ENTRELAZANDO DISEÑO Y CIENCIA FICCIÓN
El caso de Daleko

Eduardo Harry Luersen[2]
(Universidade do Vale do Rio dos Sinos)
edluersen@gmail.com

RESUMO

O artigo discute como o chamado design especulativo pode ser vinculado à extrapolação tecnocientífica, uma importante característica do gênero da ficção científica, para cotejar e prototipar cenários potenciais e modelos imaginados, porém viáveis, de futuro. O trabalho sublinha importantes nuances entre diferentes abordagens ao design especulativo, considerando a especificidade de seus pressupostos epistemológicos. A discussão é entremeada por um estudo de caso: Daleko (2020), um projeto desenvolvido no programa *Terraforming*, do Strelka Institute. Este projeto consiste em um conjunto de nove fábulas de ficção científica aproximando os temas da gestão de resíduos, infraestrutura tecnológica e governança ambiental. Os contos discutem a concepção problemática do lixo como uma forma de exterioridade, ao mesmo tempo em que imaginam possíveis alternativas futuras para a gestão de resíduos a partir de uma perspectiva tecnopolítica mais abrangente. Após analisar a abordagem de Daleko, o artigo discute as implicações do design especulativo em contextos educacionais, como um modo de estimular designers em formação e pesquisadores oriundos de outras disciplinas a questionar como as práticas de suas áreas podem ser relacionadas ao impacto antropogênico no ambiente, de modo mais abrangente, e qual o seu

---

papel no planejamento de infraestruturas tecnológicas mais viáveis para o ambiente natural e a biodiversidade atual.

Palavras-chave: design especulativo. ficção científica. gestão de resíduos. impacto antropogênico. ensino de design. design e geografia política.

ABSTRACT

The present article discusses how speculative design relates to technoscientific extrapolation, an important science fiction feature, to plan for potential scenarios and prototype viable models of futurity. Through it, the paper outlines some important nuances between different approaches concerning speculative design's role, considering their particular epistemological assumptions. A specific case is presented and discussed: Daleko (2020), a project developed for the Strelka Institute's Terraforming program. This project consists of nine science fiction pieces that entangle issues of waste management, technical infrastructure, and climate politics, discussing the problematic conception of waste as a form of externality and imagining future scenarios for managing it through a more holistic perspective. By analysing Daleko's approach to speculative design, in its final section the article suggests further developing speculative projects in contexts of design education. This would serve towards stimulating designers and researchers from other fields to ponder how their craft relates to anthropogenic impact and how they can play a decisive role in prospecting more viable infrastructures for future scenarios.

Keywords: speculative design. science fiction. waste management. anthropogenic climate change. design education; design and political geography.

RESUMEN

El artículo analiza cómo el llamado diseño especulativo puede vincularse a la extrapolación tecnocientífica, una característica importante del género de la ciencia ficción, para cotejar y prototipar posibles escenarios y modelos imaginados pero factibles del futuro. El ensayo destaca importantes matices entre los distintos enfoques del diseño especulativo, teniendo en cuenta la especificidad de sus premisas epistemológicas. El debate se intercala con un estudio de caso: Daleko (2020), un proyecto desarrollado dentro del programa Terraformación del Instituto Strelka. Este proyecto consiste en un conjunto de nueve fábulas de ciencia ficción que reúnen los temas de la gestión de residuos, la infraestructura tecnológica y la gobernanza medioambiental. Los cuentos discuten la concepción problemática de la basura como una forma de externalidad, al tiempo que imaginan posibles alternativas futuras para la gestión de los residuos desde una perspectiva tecnopolítica más amplia. Tras el análisis de Daleko, el artículo discute las implicaciones del diseño especulativo en contextos educativos, como forma de estimular a los diseñadores en formación a cuestionar cómo las prácticas del campo pueden relacionarse con el impacto antropogénico en el medioambiente de forma más amplia, y cuál es su papel en el diseño de infraestructuras tecnológicas más viables para el medioambiente y la biodiversidad actuales.

Palabras clave: diseño especulativo. la ciencia ficción. gestión de residuos. impacto antropogénico. educación para el diseño. diseño y geografía política.

Introdução

De acordo com Isabelle Stengers (2014), as obras de ficção científica podem ser definidas como mitologias contemporâneas. Além de serem catalisadores para a imaginação de milhares de leitores, as

fábulas de ficção científica são também experimentos mentais inventivos e que se debruçam sobre as práticas, demandando a imaginação de cenários vindouros como resultados potenciais de uma dada hipótese.

A ficção científica sempre tratou ciência e tecnologia como matéria-prima para a fabulação (Shaw, 2008), aspecto que é particularmente significativo na ficção contemporânea quando pensamos na crescente penetração das técnicas de simulação e modelagem computacional na sociedade ao longo das últimas décadas (Pias, 2011). Ao mesmo tempo, a ciência de dados, a modelagem estatística em tempo real e a previsão de cenários também permitem que estabeleçamos modelos de conhecimento sobre o porvir a partir de fenômenos e estados em evolução, que podem ajudar na avaliação e prevenção de problemas cotidianos, sobretudo se houver articulação, coordenação e organização política. É o caso de projetos da chamada Ciência do Sistema Terra e das projeções epidemiológicas que nos mostram a evolução de diferentes cenários ligados, respectivamente, às mudanças climáticas e à Covid-19 – ainda que, infelizmente, salte aos olhos o enorme descompasso entre o que indicam os modelos e as tomadas de decisões políticas em boa parte do planeta. Sob circunstâncias tão sombrias, entretanto, a ficção científica encontra um renovado entusiasmo pelas projeções e extrapolações de grandes proporções.

Neste sentido, o design especulativo (Dunne e Raby, 2013) também pode ter um papel importante a desempenhar. Enquanto um método do design para sondar as possibilidades virtuais, dormentes ou inexploradas da ciência e tecnologia contemporâneas, o design especulativo opera sob a linha tênue entre o possível e o improvável, o aceitável e o incomensurável, o curto e o longo prazo – tal como a ficção científica. Além disso, assim como os contos *sci-fi* mais convincentes, os projetos de design especulativo tendem a ser mais vibrantes quando seus programas consideram em profundidade os contextos políticos, econômicos e sociais (Peloušková, 2020).

Neste artigo, exploro as interseções e disjunções entre design especulativo, ficção e ciência e tecnologia a partir da análise de um projeto de design específico desenvolvido no programa Terraforming do Instituto Strelka de Arquitetura, Mídias e Design. *Daleko* (2020) é um projeto que repensa o conceito de lixo, demandando modos alternativos e viáveis para lidar com a gestão de resíduos. *Daleko* questiona a narrativa habitual que perpetua o lixo enquanto uma forma de externalidade, reconsiderando sua inevitável duração nos tempos profundos do metabolismo planetário. O projeto compreende nove contos curtos de ficção científica, situados tanto em contextos rurais quanto urbanos, interligados em termos de enredo em função de sua futuridade partilhada.

Descrevendo *Daleko*, é difícil não se lembrar de trechos do livro *Fragmento de História Futura* (2013), escrito em 1896 por Gabriel Tarde, em que a ficção científica se enreda com o ensaio filosófico-antropológico no olhar de um observador do futuro que descreve a Terra em retrospecto diante da extinção do Sol. Operando em um campo distinto, contudo, o design especulativo não deve se ocupar somente de imaginar as condições futuras da sociedade e da vida humana, mas também o próprio campo do design e as suas potencialidades à luz das implicações geológicas e geopolíticas das mudanças climáticas e do Antropoceno. Em razão disso, no final do artigo descrevo como o design especulativo pode ter um importante papel em contextos educacionais. Argumento que ele pode ser uma ferramenta pedagógica significativa para encorajar estudantes de design e pesquisadores de outras áreas a posicionarem as implicações e os objetivos de seus projetos na perspectiva e escala de um tempo mais profundo (Parreno, 2020), em sintonia com algumas das crises mais urgentes de nosso tempo, de caráter ambiental.

Formas especulativas de pensar o design

Ainda que este artigo não ambicione prover uma descrição exaustiva das diversas abordagens ao design especulativo e ao *design fiction* , é importante iniciar apresentando algumas linhas gerais de diferentes perspectivas sobre estes conceitos, para compreender melhor de que forma elas se relacionam com aspirações da ficção científica.

O rápido crescimento da literatura sobre *design fiction* dá mostras das diferentes interpretações, ideologias e objetivos que cercam o conceito (Markussen e Knutz, 2013). A definição seminal de Bruce Sterling (2012, s.p.) descreve design fiction como "o uso deliberado de protótipos diegéticos para suspender a descrença sobre a transformação" (Sterling, 2012: s.p.), definição amplamente repetida e que é capaz de abrigar satisfatoriamente diferentes vertentes conceituais. As linhas que se seguem a esta definição principal, entretanto, são objeto de maior controvérsia entre diferentes abordagens ao design especulativo. De acordo com Sterling, o uso de protótipos diegéticos implica em "pensar muito seriamente sobre os objetos e serviços em potencial, e focar neles em vez de pensar em mundos inteiros, tendências políticas ou estratégias geopolíticas. Não é um tipo de ficção, mas de design. Eles devem narrar mundos em vez de histórias" (Sterling, 2012: s.p.).

Ainda que seja ponto pacífico que a prototipagem de objetos e serviços diegéticos ocupe boa parte das iniciativas voltadas a ficcionalizar o design, o desenvolvimento de histórias de fundo – mesmo (e, talvez, especialmente) quando se tratam de contextos e situação controversas e radicalmente não-familiares – pode prover insights significativos para a concepção e materialidade dos projetos ficcionais. Na medida em que produtos e serviços são alvos fáceis para a fetichização, a imaginação de contextos e narrativas substanciais de fundo tendem a adensar as potenciais implicações dos projetos, mesmo quando a ficcionalização se volta a um produto ou serviço.

Além disso, há correntes do design fiction que buscam fundir design e técnicas narrativas voluntariamente, investindo no potencial fabulador dos protótipos diegéticos através de produção audiovisual, por exemplo. O designer Julian Bleecker (2009) tem defendido esta perspectiva há anos. Para ele, o desenvolvimento de protótipos diegéticos pressupõe converter as criações especulativas em experiências tangíveis, e o audiovisual permite imaginar situações particulares e valores sociopolíticos associados ao entorno dos artefatos e dispositivos, além de prover-lhes com uma imagem própria. Neste sentido, a interseção entre design fiction e dramaturgia é interessante por promover a imaginação de situações mundanas e cotidianas, para além dos fins mais imediatos associados ao protótipo.

Excetuando a escolha mais óbvia pela realização audiovisual, Bleecker (2017) também sugere realizar os protótipos diegéticos em outros suportes, tais como manuais de montagem, guias instrucionais, ou catálogos de produtos. Poderíamos ainda adicionar modelos de gamificação, patentes de produtos, ou documentários falsos a esta lista. Há também protótipos no formato de "ficções acidentais" (Dunne e Raby, 2013: 89), tais como invenções defeituosas ou rejeitadas. É válido lembrar aqui, por exemplo, do artigo que Thomas Edison (1878) escreveu para a North American Review um ano após registrar a patente do fonógrafo. Nele, Edison sugere uma série de usos futuros para o aparato, tais como exercícios de ditar, a audiodescrição de cartas, a preservação de idiomas, além da elaboração de produtos derivados como relógios sonoros, livros fonográficos e brinquedos musicais (Luersen, 2014). Além da reprodução fonomecânica ter atingido a larga-escala durante século XX, enquanto um meio para entretenimento doméstico, outras aplicações imaginadas por Edison na época hoje se fundem com tecnologias contemporâneas.

No contexto cotidiano por vir, os produtos e serviços e especulativos naturalmente podem não funcionar efetivamente da maneira como foram imaginados quando prototipados pelos designers. Pensando nisso, alguns designers procuram incorporar erros e problemas de funcionamento como parte integral dos projetos de *design fiction*. Neste caso, não se trata apenas de tornar o projeto mais tangível diegeticamente, mas também de explorar aspectos e funções não diegéticas dos projetos. Este tipo de experiência comparece, por exemplo, em projetos como o curta-metragem *A Digital Tomorrow* (2012), do Near Future Laboratory, em que uma interface de reconhecimento facial falha quando o usuário passa a utilizar uma maquiagem menos usual.

O potencial para produzir implementações práticas a partir destes projetos é outro motivo de controvérsia entre as diferentes abordagens ao design especulativo, que operam com intenções ora mais idealistas, ora mais pragmáticas. De acordo com o programa de Anthony Dunne e Fiona Raby (2013) para o design especulativo, os projetos não devem necessariamente se preocupar em prover soluções realizáveis. Em

vez disso, devem preocupar-se em sondar novas perspectivas, correspondentes a diferentes valores, crenças, desejos e ideais individuais ou sociais. Em seu entendimento, as formas especulativas de projetar devem "enfatizar a imaginação acima da praticidade, e realizar perguntas em vez de prover respostas. O valor do projeto não está naquilo que ele alcança ou realiza, mas no que ele é e em como é capaz de mobilizar os sentimentos dos indivíduos (Dunne e Raby, 2013: 189). Tal descrição subscreve a ênfase nas dimensões emocionais e afetivas do design, sobretudo na concepção de artefatos para exposições (o que também é possível de perceber pelo uso frequente dos termos "espectador" e "audiência" em sua abordagem). Os objetos, preparados para a sua apresentação em galerias, escolas e museus, devem primeiramente estimular a imaginação do público, inspirando os sujeitos a criarem os seus próprios modelos de realidades alternativas. Ao afirmarem que os projetos de design especulativo não precisam se comprometer com soluções ou "melhorias", e que os designers deveriam suspender estes juízos temporariamente porque "o público é capaz de decidir por conta própria" (Dunne e Raby, 2013: 189), os autores se alinham perfeitamente a um paradigma da arte conceitual contemporânea. Dunne e Raby argumentam a favor da elaboração de modelos críticos a serem experimentados pelo público, entendido aqui enquanto partícipe que interpreta abertamente o sentido da obra idealizada.

Ainda que esta abertura em relação ao significado dos projetos, assim como o estímulo a avaliar mais criticamente o design cotidiano, sejam sem dúvida importantes ao aprendizado, à imaginação e à reflexão sobre o design, é importante também observarmos como esta perspectiva se relaciona e se difere de outras abordagens ao design especulativo. Se acreditarmos que "as propostas do design especulativo não se referem ao real", pois "tratam-se de ficções físicas, pontos de partida para imaginações sofisticadas que nunca ambicionam ao "real", nem buscam refletir a realidade" (Dunne e Raby, 2013: 92), então a possibilidade de que os projetos especulativos incorporem e proponham intervenções materiais mais duráveis se torna radicalmente diminuta. Ainda que valorize o modo como a experiência estética estimula a consciência crítica e o questionamento de nossas relações cotidianas com o design, entendo que os projetos de design especulativo, por outro lado, não devem se ocupar necessariamente apenas com a produção de experiências e representações idealizadas que visem a conscientização individual. Neste sentido, é importante também ressaltar aquilo que os projetos podem alcançar quando abertamente comprometidos com a tarefa de ultrapassar esta divisão através da adoção de uma imaginação científica, política e técnica mais rigorosa em seus briefings. Com isso, não estou propondo que se dê vazão a um realismo ingênuo, mas que se considere mais atentamente como os modelos que criamos enquanto designers se entrelaçam com um aspecto material mais complexo. Ao imaginar as questões epistêmicas que constituem uma dada realidade prospectiva – com o conjunto heterogêneo de pressupostos, regulações, saberes científicos, valores morais, práticas organizacionais e discursos que a constituem – um designer pode nutrir seus projetos com planos não só imaginativos, como também mais viáveis. Por esta via, a prospecção especulativa pode ter um papel muito importante, que vai além da experiência fenomenológica de um sujeito com novos protótipos ideais. Os modelos imaginativos também podem basear-se na expectativa de realidades vindouras, ainda que consideremos seu caráter incondicionalmente aberto, como procuro demonstrar mais adiante a partir do projeto *Daleko*.

Novamente, isto não significa imbuir o modelo especulativo em um sentido ingênuo de realismo. Por isso mesmo, é importante assinalar que há perspectivas sobre o design especulativo que não se contentariam com a experiência e a celebração de projetos com apenas "um desejo brando de se realizarem" (Dunne e Raby, 2013: 89). Abordagens complexas do design especulativo também não se norteariam pela noção mais ingênua (e apelativa) de "previsão futura", bastante recorrente no senso comum quando se fala em ficções científicas. Este seria um objetivo muito restritivo para os projetos de design especulativo, com a proposição de relações muito lineares de causa e efeito ao imaginar o espaço entre o agora e o provir. Em vez disso tudo, meu argumento a favor de um engajamento do design especulativo com algum grau de pragmatismo aponta para um movimento no sentido oposto: imaginar os parâmetros de realidades potenciais em seu futuro fatalmente abstrato, de modo a prover projetos eficientes, que possam ser trabalhados *desde lá*, do futuro imaginado, *até* as condições materiais do presente. Isso, por si só, sugere um método capaz de envolver os designers com o desenvolvimento de práticas de design mais duráveis, sustentáveis, substanciais, e ainda assim imaginativas.

Uma abordagem decididamente imbuída em certo grau de pragmatismo favorece a sondagem de cenários de articulação conjectural entre o design e políticas, padrões, regulações, dentre outras forças organizacionais da sociedade, em formatos atuais e potenciais, sem abrir mão de operar especulativamente – produzindo assim um modelo hipotético, mas ainda à parte do senso comum. Ainda que esta afirmação seja controversa, entendo que tal perspectiva pode ser particularmente promissora para o design especulativo, sobretudo considerando as questões e dilemas que envolvem os processos antrópicos no atual estágio da vida biogeoquímica na Terra. A definição do Antropoceno enquanto uma nova era geológica se relaciona diretamente com o impacto da espécie humana e de nossa capacidade industrial no planeta, incluindo aí as "diversas violações da vida ambiental e humana a partir de práticas corporativas e da cultura tecnológica, que ameaçam as próprias condições de existência humana no porvir" (Parikka, 2014: 6). O impacto antropogênico dos últimos séculos culminou na alteração do ciclo dos períodos interglaciais e das eras do gelo, algo que imaginava-se ser possível apenas com a mudança do eixo de rotação da Terra[3] (Archer, 2009). Para conceber um fenômeno desta dimensão, em que escalas temporais e espaciais excedem em muito a nossa percepção subjetiva, uma abordagem ao design especulativo que seja capaz de equilibrar abstração tecnocientífica radical e imaginação de conformações materiais vindouras é muito bem-vinda, senão necessária.

Uma iniciativa proeminente que procura entrelaçar a produção de conhecimento tecnocientífico e questões relativas ao Antropoceno com o design especulativo, construindo a sua própria perspectiva sobre o tema, é o programa Terraforming, do Instituto Strelka[4]. *Terraformação* é um termo usado mais frequentemente em histórias de ficção científica e em contextos de comunicação científica, que se refere ao processo hipotético de modificação da topografia e das condições atmosféricas de outros planetas e corpos celestiais para tornar seus ecossistemas habitáveis para as formas de vida terrestres. O programa do Instituto Strelka, porém, refere-se à ideia de terraformação do próprio planeta Terra. Refletindo a perspectiva do coordenador do programa, Banjamin Bratton (2019), alude-se tanto à terraformação que ocorreu na Terra ao longo do último milênio, quanto à terraformação que deve ser planejada de agora em diante, como uma iniciativa planetária de design para prevenir desastres vindouros de larga-escala. Tal programa enfatiza que um plano para atenuar os efeitos da mudança climática de origem antropogênica deve se ocupar de uma resposta decididamente antropogênica, reconciliando a ação humana com a artificialidade através de análises e intervenções projetuais. Portanto, não é uma questão de se agir contra o Antropoceno, na medida em que trata-se da mais atual era geológica na história do planeta, mas de planejar a infraestrutura necessária para viver diante de tal realidade.

Se, durante a maior parte do século XX, "o futuro" tomou a forma de uma inspiração, algo a ser alcançado, agora que a computação em escala planetária provê cenários tecnicamente mensuráveis e prevê catástrofes climáticas globais mais recorrentes a partir deste século, o futuro se torna algo a ser prevenido, senão evitado. Neste sentido, a principal tarefa que o programa Terraforming designa ao design especulativo é planejar como um futuro previsto pode ser evitado, para que, com sorte, um futuro diferente possa surgir: que seja alcançado em função daquilo que foi evitado (Bratton, 2019: 38). Sob este conjunto de princípios, a pesquisa em design conduzida no instituto investe muita energia em imaginar infraestruturas de governança que sejam capazes de operar em escalas temporais muito mais profundas do que aquelas que motivam nossos impulsos e necessidades cotidianas mais subjetivas.

---

[3] A espécie humana tem interagido artificialmente com o meio ambiente desde o surgimento do *sapiens*, há cerca de 300.000 anos. Porém, antes do fim da última era glacial, há 12.000 anos, as populações humanas eram comparativamente muito menores, tendo estado próximas da extinção por diversas vezes (Veiga, 2014). A espécie só se estabilizou com o assentamento dos povos nômades, quando a agricultura foi inventada. Embora haja muito debate a respeito do início exato da era do Antropoceno, para o propósito deste ensaio considero o período mais recente de aumento significativo do impacto antropogênico sobre os ecossistemas terrestres, a partir de meados do século XX. Este período corresponde ao ritmo acelerado do consumo de energia e das emissões de gases de efeito estufa, e às demandas de recursos necessárias para manter a infraestrutura e as operações logísticas predominantes diante da globalização. Sem desmerecer sua evidente dimensão moral (Chakrabarty, 2018), o Antropoceno se estende para além de suas dimensões históricas e filosóficas, como uma mudança de paradigma definida cronostratigraficamente, uma mudança demarcada no registro geológico da Terra. Diante disso, abordar o assunto em sua dimensão geológica é decisivo, e é importante considerar o quadro geral fornecido pela Ciência do Sistema Terra para promover um modo de pensar mais abrangente e efetivamente planetário.

[4] Diante da trágica e mais recente invasão russa à Ucrânia, e das conseqüentes implicações desta ação sobre programas e projetos de cooperação internacional realizados em Moscou, o instituto suspendeu suas atividades a partir de fevereiro de 2022.

Tal abordagem catalisa as perspectivas previamente articuladas por Bratton sobre o design especulativo:

> Em vez de concluir que o futuro (e o futurismo *per se*) está perdido, devemos mobilizar a modelagem de infraestruturas para propósitos melhores e mais vibrantes. Para isso, é preciso que o propósito dos modelos especulativos seja redirecionado: que se dediquem menos a prever o que provavelmente irá ocorrer (como no caso da previsão de determinados cenários avançados) e se voltem, em vez disso, à pesquisa do espaço de possibilidade real (mesmo e, sobretudo, além daquilo que qualquer um de nós poderia conceber de outra forma). Isto é, modelos preditivos são adaptativos porque eles se baseiam em uma descrição. Já para a especulação os modelos devem ser prescritivos, pois é necessário que eles se tornem normativos (Bratton, 2016: s.p.).

Esta visada exerce importante influência sobre as premissas epistemológicas a partir das quais os projetos especulativos do programa irão se desenvolver, fundindo teoria do design e prototipação programaticamente a partir da especulação. Em contraste com perspectivas que procuram imaginar destinos de apelo mais descritivo, tal proposta favorece uma imaginação mais ambígua (e também "vibrante"), que remete à condição infraestrutural do design – uma fissura nas condições atuais normais, das quais emergem normalizações em potencial. Esta é precisamente a razão pela qual os trabalhos de ficção científica que exploram a verdadeira ambiguidade dos cenários futuros conseguem prover insights mais significativos sobre seus futuros imaginados. Quando desalinhada de condições estruturais mais profundas, a especulação corre o risco de se tornar pouco mais do que um fetiche de ideais amplamente difundidos no cotidiano, ou, pior ainda, apenas uma justificativa para manter as coisas como elas estão.

Como coloca Klára Peloušková (2020), a notável abstração do projeto de Bratton pode colocar certo grau de suspeição sob as suas ambições. Porém, também deixa claro como tal perspectiva tem pouco a ver como formas mais espetaculares (e também menos imaginativas) de design especulativo, cujo interesse principal se volta à produção de representações alegóricas. A perspectiva esquematizada pelo programa Terraforming compreende uma obsessão com o materialismo científico e filosófico, e especula sobre a organização e normalização de novas megaestruturas, sem deixar de considerar seu grau de ambiguidade e incerteza. A sugestão de normas e valores emergentes, contíguos a tais estruturas, faz parte dos cenários especulativos, e permite que a extrapolação tecnocientífica permaneça entrelaçada com uma análise crítica das conjunturas políticas.

Proponho aqui mostrar como tal perspectiva opera, com a descrição e análise de um dos projetos desenvolvidos no programa de design Terraforming. Esta análise também permite lançar um olhar mais preciso sobre como tal projeto entrelaça aspectos específicos da ficção científica e do design especulativo.

Daleko: entranhando design e ficção científica

Daleko é um dos projetos de conclusão do primeiro ano do programa Terraforming, e foi desenvolvido por Eleanor Peres, Anastasia Sinitsyna, Tigran Kostandyan e Tim Nosov. Eles propõem nove histórias ilustradas de ficção científica ambientadas em territórios urbanos e rurais do chamado "espaço pós-soviético", e centram-se em locais historicamente ligados ao depósito de resíduos sólidos de diferentes origens. Os contos são narrados através de fragmentos textuais ilustrados e animados com a estética de cartoons soviéticos tardios. As ilustrações exploram um rico conjunto de motivos bioquímicos, cartográficos, cosmológicos e tecnocientíficos. O tom anacronístico dos grafismos se ajusta ao propósito de entrelaçar diferentes temporalidades dos contos, e a percepção de uma cronologia mais estendida ressoa com a ideia de problematizar, ao longo do projeto, a temporalidade cíclica do lixo, da poluição e da sucata.

Integrado à perspectiva de Bratton quanto ao design especulativo, *Daleko* sugere que a situação climática vindoura do planeta exige projetos de larguíssima escala e de longa duração. Assim, são necessárias alternativas às soluções em voga, tais como a economia circular, para lidar com os ciclos de descarte de grandes proporções. O projeto reconhece primeiramente a necessidade de revisitar o próprio conceito de lixo, com o entendimento de que as noções de senso comum compreendem o lixo enquanto uma forma

de exterioridade, o que é profundamente enganoso: trata-se de uma perspectiva problemática, que implica diretamente no imaginário de que há um "espaço exterior" para onde toda a matéria desperdiçada deve ir. O entendimento habitual de que os materiais descartados eventualmente irão desaparecer em um depósito qualquer (se tornando *lixo*) está claramente associado ao seu deslocamento geográfico. Não surpreende que esta exterioridade imaginada se defina em parte por acordos geopolíticos e financeiros, tal como o bem documentado fluxo transfronteiriço de lixo eletrônico, que normalmente deixa os país desenvolvidos em direção aos países em desenvolvimento (Gabrys, 2011; Parikka, 2015). Como coloca Ioana Jucan, "o [objeto] obsoleto não desaparece – ele se desloca" (Jucan, 2019: 12). Ainda assim, a perspectiva deslindada em *Daleko* leva este argumento adiante, ao explorar as condições geoquímicas do descarte e da poluição. Neste sentido, *interiorização* e *externalização* passam a ser entendidas como expressões de uma divisão anterior entre natureza e cultura. A designação normativa de aterros, ferros-velhos e depósitos de lixo enquanto locais cultivados para o descarte provém uma imagem-testemunho da perspectiva de curta duração a partir da qual compreendemos o processamento geoquímico da matéria. A organização geotécnica destes lugares contribui na formatação da duradoura imaginação do lixo como algo externo. Assim, na longa duração é necessário que sejam projetadas outras formas de coletar e acomodar o excedente material. É por isso que em *Daleko* a questão dos resíduos enquanto formas de exterioridade é reimaginada a partir de uma compreensão mais abrangente. Em vez de restos e sobras, emerge o entendimento de um metabolismo planetário, da matéria que se dobra sobre a matéria a partir de seu processamento físico-químico indelével. Os ciclos metabólicos orgânicos da Terra se entranham com a dinâmica da matéria artificial, produzida por processos técnicos antrópicos. Sob esta abordagem materialista, megaprojetos ambiciosos e abrangentes de infraestrutura devem se enredar programaticamente com agendas atentas à escala dilatada do tempo geológico.

Nas páginas a seguir, sintetizo os nove contos que compõem o projeto *Daleko*, destacando a relação entre os cenários ficcionais imaginados e as alternativas de extrapolação tecnocientífica exploradas. A extrapolação de padrões e modelos tecnológicos é uma característica importante da ficção científica, e por isso me dedico mais detidamente sobre esta especificidade. Por priorizar a observação do projeto como um todo, de forma mais ampla, as análises não procuram esgotar nenhuma das histórias[5]. Uma noção mais generalista destas narrativas permite perceber mais claramente aquilo que as une, favorecendo a observação de como a pesquisa envolvendo o design especulativo pode ser articulada com o conhecimento tecnocientífico de modo mais geral, e com as extrapolações que motivam a ficção científica, mais especificamente.

O conto *Moscow Garbage Ring* se volta ao imaginário de periferia, mostrando como nossas práticas de descarte estão emaranhadas com esta noção e como o trabalho informal domina atualmente a gestão e o fluxo transfronteiriço de resíduos. Os aterros sanitários na Rússia ocupam uma área do tamanho da Holanda. Cerca de 80% dos resíduos domésticos são enviados para estas áreas de despejo, e o que sobra normalmente é incinerado. Este conto imagina a revitalização experimental da cidade de Pushchino, um importante centro científico ao sul de Moscou, que recebe boa parte dos afluxos de lixo da cidade. No futuro, o "Ministério de Tecnoestética" aprova um programa piloto que define que todos os resíduos devem ser devidamente reutilizados como matéria-prima de segunda mão ou fertilizantes, reduzindo em 70% o desperdício na cidade. As calhas e canaletas de lixo, além de alguns tipos específicos de embalagens, passam a ser proibidas na cidade, levando as empresas a desenvolver tecnologias para reprocessar os materiais. Esta norma hipotética é inspirada em um movimento ecológico dos anos 1960, liderado por cientistas de Pushchino. A localidade é referência científica nacional, por apresentar uma confluência entre iniciativas de sustentabilidade e pesquisas em química e biologia molecular. Na história, o circuito do lixo de Moscou é transformado em um parque seminatural, conhecido por suas paisagens híbridas, que combinam os biomas recém-nascidos com uma arquitetura conservada de séculos anteriores. A região adota um *ethos* eco-agrícola que inclui abordagens bem conhecidas, tais como a permacultura, sistemas aquapônicos, *slow food*, além de inovações derivadas do reprocessamento de *commodities* petroculturais.

---

[5] É claro, é recomendável que o leitor interessado acesse o projeto Daleko completo, disponível em: <https://daleko.space/>. Acesso: 31 jan. 2022.

Ilustração do conto *Moscow Garbage Ring*. Fonte: *Daleko*.space

Outra história, *Flight Path Zapovednik*, explora como o conceito de externalização dos resíduos pode afetar biomas e populações locais muito específicas, especialmente quando alinhado a uma visão torpe sobre o desenvolvimento. A narrativa descreve o caso das estepes cazaques enquanto uma "zona de sacrifício", que viu toda a sua biodiversidade diminuir drasticamente em poucas décadas. A instalação do Cosmódromo de Baikonur, nos anos 1950, transformou a área no maior porto espacial do mundo, além de um espaço para a testagem de mísseis balísticos. Em apenas alguns anos, o combustível altamente tóxico usado nos lançamentos de foguetes degradou claramente a vida selvagem, o gado, e as comunidades locais ao longo da trajetória de lançamento. O conto explora este caso, elaborando um projeto que entrelaça práticas de cuidado ecológico e tecnocientífico, controle ambiental e relações internacionais com uma abordagem nomotética. O projeto traça um plano alternativo de remodelagem da região, vislumbrando um acordo para a construção de uma reserva natural nos próximos anos: após décadas de poluição causada pelos detritos de armas nas estepes cazaques, renegocia-se uma dívida da antiga URSS entre Rússia e Cazaquistão através de um novo pacto para "fitorremediar" a área no entorno da rota dos voos. O acelerado aumento da temperatura na Rússia pressiona o governo a investir no financiamento deste projeto, que opera através de informações de "cuidadores nativos pagos, imagens de satélite, mapas de calor, e centros de dados distribuídos" (Peres et al., 2019). Com este arranjo, a área se torna a primeira paisagem protegida no planeta a combinar tecnologias contemporâneas de sensoriamento remoto com práticas indígenas de manejo de terra, por meio de um projeto de cooperação transnacional.

A história seguinte, *Rebirth Orbit*, aborda o chamado "espaço sideral", o limite mais extremo de exterioridade imaginada. O conto conecta a especulação com experiências de gestão de lixo espacial em desenvolvimento no presente. A história informa que atualmente 90% dos objetos na órbita da Terra são resíduos sólidos, conseqüência de décadas de abandono e descarte não regulamentado de naves espaciais e satélites. O conto, portanto, imagina um programa espacial que não apenas leva astronautas e novos equipamentos para fora do planeta, mas que também recolhe lixo espacial, descartando-o em áreas explosivas da atmosfera solar. Esta ideia é abertamente baseada em iniciativas como o *RemoveDEBRIS*[6], e em projetos de reciclagem de detritos espaciais desenvolvidos pelo *Laboratório de Resíduos Espaciais* do Studio Roosegaarde, que pretende transformar os detritos em estrelas cadentes, ou ainda usá-los como matéria-prima para construções na lua impressas em 3D. A faxina cósmica realizada por meio de foguetes recolhedores de lixo, como proposto em Rebirth Orbit, não extrapola as tecnologias a serem utilizadas, já em desenvolvimento nos laboratórios espaciais atuais, tanto quanto extrapola as atuais estruturas de governança: no conto, o controle de reentrada, o descarte de naves e a coleta de detritos espaciais são todos regulados por novos acordos internacionais e serviços transnacionais dedicados ao problema compartilhado do lixo espacial.

O conto Deep Sekretiki lida com o problema da poluição em relação ao extrativismo e à infraestrutura colossal desenvolvida para fornecer petróleo em escala planetária. Esta história é particularmente interessante porque pressupõe a compreensão da Terra como um sistema metabólico, onde órgãos sensoriais (monitores) orientam e regulam a relação entre digestão e excreção. Ela imagina o redesenho de duas infraestruturas gigantescas e complementares já existentes: a primeira é a estrutura de controle via satélite utilizada por operadores comerciais e militares, reimaginada como um artefato para monitorar horizontalmente as emissões de carbono e os vazamentos de metano, informando agências reguladoras;

---

[6] *RemoveDEBRIS* é um projeto de pesquisa que busca construir e testar tecnologias de remoção de detritos espaciais em potencial. Este objetivo envolve a demonstração de diversas tecnologias de *remoção ativa de detritos* aplicadas a alvos-teste localizados em órbita baixa. A estrutura da plataforma experimental está equipada com redes, arpões, altímetros a laser e microssatélites.

a segunda é a ampla rede de oleodutos de extração de petróleo, reformada e retroprojetada para realizar captura e armazenamento de carbono. Tais infraestruturas têm décadas de existência, possuem notória baixa eficiência energética e, enquanto estruturas subterrâneas, são invisíveis a olho nu. Neste projeto, a especulação envolve tornar as estruturas subterrâneas profundas de extração de petróleo visíveis e mensuráveis, fornecendo um parâmetro para manter os níveis de vazamento e emissão dentro dos limites estabelecidos pelas agências reguladoras. Novamente, isto envolve um cenário organizacional mais amplo, que no conto se materializa com a fundação de uma futura Unidade de Sensoriamento das Emissões Planetárias (PESU), responsável não apenas pelo monitoramento, como também pela divulgação do maior conjunto de dados globais sobre os efeitos contínuos de tal infraestrutura em relação à mudança climática. A PESU apresenta ao governo russo as estatísticas comprometedoras, porém cotidianas, sobre sérias violações ambientais envolvendo vazamentos de petróleo, e a indústria de combustíveis fósseis é legalmente forçada a atingir os padrões internacionais conforme o sensoriamento, mantendo as emissões anuais abaixo do limite estabelecido de 20%. Esta história é especialmente interessante por entrelaçar ciências de dados, jornalismo climático e geoengenharia em um mesmo projeto multi-institucional imaginário.
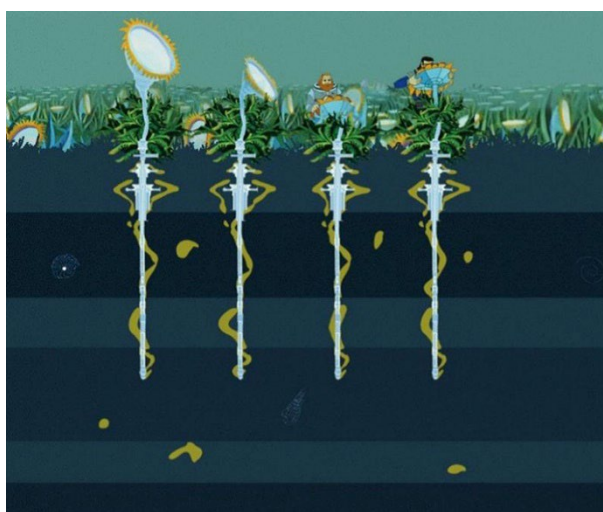

Ilustração do conto *Deep Sekretiki. Fonte: Daleko.space*

A trama da história seguinte, *Nuclear Ikea*, é narrada por um sistema de inteligência artificial responsável pelo monitoramento, gerenciamento e regulação de combustível nuclear usado, que fica depositado em um Repositório Geológico Profundo[7], no subsolo (5 km de profundidade) de uma cadeia montanhosa em Nizhnekamsk, na Rússia. Esta história imagina uma época em que o combustível fóssil não é mais a fonte de energia global dominante. O desenvolvimento de uma nova geração de reatores de nêutrons modulares e acessíveis permite a propagação planetária de reatores de energia nuclear, bem como a reutilização em larga escala de combustível nuclear usado. Computação, automação, logística e geoengenharia operam em conjunto, na forma de um serviço de armazenamento e gerenciamento do combustível nuclear usado, que passa a ser movimentado para dentro e para fora do território nacional. Embora não entre em detalhes, o projeto parece explorar duas importantes precondições subjacentes a esta situação hipotética: o desenvolvimento de uma rede transcontinental para o transporte e gerenciamento logístico dos resíduos nucleares; e a reestruturação da mão-de-obra conduzida em condições de trabalho perigosas, que passa a ser desempenhada por agentes não-humanos, computacionais[8].

---

[7] Um *repositório geológico profundo* é uma instalação composta por uma rede de túneis subterrâneos e salas de armazenamento planejados para a deposição de resíduos das usinas nucleares. Dispostos estrategicamente em uma camada geológica estável no subsolo, os rejeitos radioativos de alta periculosidade devem ser armazenados e isolados por um período de milhares de anos.

[8] Essa história demonstra como é oportuna a observação de Julian Bleecker, de que possíveis erros e acidentes também devem ser imaginados e incorporados aos projetos especulativos. Embora o debate público a respeito de temas como a gestão de resíduos radioativos gere normalmente muita animosidade, os projetos especulativos não devem evitar estas questões, na medida em que elas também podem nos lembrar que em projetos tecnológicos de grande escala, as contingências devem ser consideradas nas mesmas proporções.

River in Death Valley trata do caso de Khalmer-Yu, um antigo assentamento urbano estabelecido para mineração de carvão desde os anos 1940 até 1990. Nas últimas décadas, a área foi transformada em um campo militar para testagem de bombas e outras armas. Em *Daleko*, esta região deixa de funciona como um local de testes armamentistas. Em vez disso, o local é revertido em um espaço para experiências científicas relacionadas à mitigação das mudanças climáticas, gerando demanda de mão-de-obra para trabalhadores locais e visitantes. O aeroporto militar torna-se a área laboratorial de onde partem aeronaves para realizar a injeção de aerossóis na estratosfera. As unidades de sensoriamento são utilizadas para medir a temperatura, a composição do ar e a reflexão solar. Ao mesmo tempo, o conhecimento dos cidadãos locais é utilizado para compreender os resultados do impacto antropogênico na terra, na flora e na fauna. Efetivamente, a ideia central deste conto é a reformulação do conceito de campo de testes: a conversão de um local de testagem de armas, ambientalmente e geopoliticamente problemático, em um laboratório de geoengenharia voltado a problemas de pesquisa sobre o clima sugere uma experiência localizada a ser amplificada, de modo a melhor compreender e mitigar os efeitos regionais do aquecimento global.

Um dos grandes desafios na compreensão cotidiana acerca do lixo está no modo como concebemos a questão do descarte como um problema de ação individual. Este raciocínio é contraproducente, uma vez que a responsabilidade individual representa apenas uma pequena parte do problema, diante da complexidade e da escala industrial da produção de resíduos. Para poder dar conta da dimensão mais ampla deste assunto, é necessária, dentre outras coisas, uma forma mais eficaz de rastrear o fluxo transcontinental de resíduos sólidos. Como apontado em um estudo anterior (Luersen & Fuchs, 2021), organizações ambientais como a *Basel Action Network* chegam a inventar estratégias de monitoramento bastante idiossincráticas para tentar dar conta do problema[9]. No entanto, com a persistência de descargas irregulares, mas também de incêndios e desmatamento, precisamos fazer uso das potenciais ferramentas e redes de monitoramento de ponta que já temos disponíveis. O conto *Black Sky Coordinates* explora esta situação delicada, sugerindo uma aproximação entre modelos de redes neurais, unidades de prevenção de crimes ambientais e nossos telefones pessoais, propondo a construção de um instrumento público de monitoramento, capaz de identificar e informar abertamente os níveis de poluição do ar de uma determinada região, por exemplo. Ao explorar o interesse público, pelo fornecimento de instrumentos adequados para que os cidadãos monitorem os índices de poluição local, o projeto também especula sobre medidas nomotéticas, na medida em que tanto *o direito ao ar puro* como *o direito ao acesso de informações atualizadas e confiáveis* sobre a poluição local passam a ser assegurados pela constituição federal.
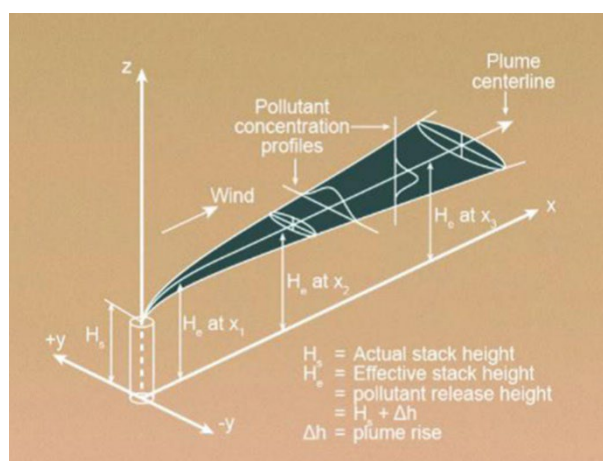
Ilustração do conto *Black Sky Coordinates*. Fonte: *Daleko*.space

A narrativa de *Oceanic Swarm* começa com uma declaração familiar sobre nosso ainda muito limitado conhecimento sobre as profundezas do mar. O projeto retoma o farol abandonado *Aniva*, uma torre

---

[9] Estas estratégias incluem, por exemplo, a instalação de equipamentos rastreadores e câmeras escondidas em contêineres de lixo para monitorar o descarte ilegal e o tráfico de sucata em portos, alfândegas e outras zonas fronteiriças (Hopson & Puckett, 2016; Baldé et al., 2017).

(outrora utilizada para auxiliar na navegação de navios) que se tornou obsoleta com o desenvolvimento das tecnologias de GPS. Na história, o farol é recondicionado e passa a integrar a Estação Aniva: um alojamento que recebe informações acústicas coletadas por enxames de Veículos Subaquáticos Autônomos e as retransmite para um conjunto de satélites interconectados. Esta estação é imaginada como um nó essencial para a futura integração das comunicações entre a biosfera, a hidrosfera e a tecnosfera, em um sistema cibernético viável, que retroalimenta informações para um banco de dados de longo prazo. Os sinais sonoros capturados pelos enxames fornecem informações atualizadas sobre condições do mar (sua temperatura, salinidade, densidade e grau de poluição) a uma ampla rede de institutos de oceanografia e biologia marinha envolvidos na observação climática global. Além disso, o projeto também sugere uma relação entre tal tarefa de monitorando oceânico e a gamificação, descrevendo o analista destes dados como uma espécie de "personagem de um videogame sobre o mundo marinho, que procura por sinais de potenciais ameaças ambientais" (Peres et al, 2020: s.p.), ao mesmo tempo em que explora as regiões ainda desconhecidas do fundo do mar.

Enquanto a história anterior se aproxima com curiosidade do substancialmente desconhecido espaço oceânico, o conto *Copernican Dive* se volta à quantidade de matéria que é sabidamente despejada nos oceanos, rotineiramente. O projeto informa que a Rússia despeja anualmente cerca de 13.000 toneladas de resíduos plásticos no mar. Para se ter uma ideia, estudos da paleobiologia asseguram que a quantia acumulada de plástico produzido até 2015 já era suficiente para embrulhar o mundo inteiro em filme plástico (Zalasiewicz et al., 2015). Ao passo em que discussões sobre um chamado "Plasticeno" se desenvolvem (Haram et al., 2020), as políticas e leis de reciclagem em voga provam não ser suficientes para evitar que grande parte da matéria plástica consumida acabe dentro do oceano. O projeto especulado em Copernican Dive considera a dimensão ecológica, evidentemente, mas também as implicações técnicas, jurídicas, geopolíticas e econômicas envolvidas no problema do descarte de plásticos. O conto prevê que, em um futuro próximo, o problema do despejo de plástico no mar receberá tanta cobertura e atenção da mídia quanto a queda de detritos da Estação Espacial MIR no Oceano Pacífico recebeu em 2001. O projeto infere que a gestão de resíduos passará em breve a ser considerada um tema de bem comum planetário, efetivamente, sendo sujeita ao escrutínio da comunidade internacional. Em face das catástrofes climáticas presentes e vindouras, a economia global deve reestruturar-se em torno de problemas ambientais, com a modelagem dos efeitos biogeoquímicos do lixo e a simulação de cenários futuros delegando um peso mais importante para funções de monitoramento. Com a possibilidade de parametrização e avaliação estrita dos danos ambientais provocados por diferentes desastres em escala planetária, as partes responsáveis devem passar a ser legalmente responsabilizadas perante a comunidade internacional, tornando o despejo de lixo nos oceanos algo economicamente custoso para os envolvidos.

Ilustrações do conto *Copernican Dive*. Fonte: *Daleko*.space

Um dos aspectos decisivos das histórias apresentadas em *Daleko* é que as formas de extrapolação tecnológica parecem encaminhar, num plano de fundo, questões de economia e geopolítica multipolares. Apesar do alto grau de incerteza que permeia as relações exteriores no momento atual, os recentes acontecimentos na política global sinalizam uma crescente percepção sobre a necessidade de mitigar os

efeitos da mudança climática. Com o aparente alinhamento de China, União Européia e (muito recentemente) os Estados Unidos com o Acordo de Paris, e com mercados inclinados para iniciativas verdes, é provável que a agenda da governança ambiental passe a penetrar mais diretamente nas relações internacionais, através de tratados e acordos comerciais multilaterais, como sugerido pelo historiador econômico Adam Tooze (2020). Para Tooze, mesmo que os novos regimes de poder ligados à gestão do clima não resultem em percepções consensuais entre os estados soberanos, é possível que empresas e corporações que representem uma ameaça à estabilidade ambiental passem a correr um maior risco de perder sua licença para operar. Neste cenário, é mais provável que aconteçam intervenções políticas, articuladas com leis, normas e regulamentos mais estritos relacionados à questão climática:

> Na semana anterior ao discurso de Xi Jinping na ONU, o *Climate Action 100 Plus*, um grupo de lobistas que representam investidores globais com um patrimônio acumulado de US$ 47 trilhões, anunciou que passou a avaliar 161 das maiores empresas, coletivamente responsáveis por até 80% dos gases de efeito estufa industriais globais, pelo seu avanço para zerar as emissões líquidas de carbono até 2050. É óbvio que há um componente de greenwashing corporativo em qualquer declaração desse tipo. Mas também se trata de um voto da parte dos maiores gestores de ativos, como *BlackRock* e *Pimco*, contra o negacionismo. Assim como Pequim, eles entendem que o *status quo* e a acumulação de capital no futuro também dependem da manutenção de uma situação ambiental estável (Tooze, 2020: 7).

Os projetos de design de *Daleko* estão todos entrelaçados com implicações jurisdicionais e geopolíticas, mesmo que a dimensão diminuta dos contos não proporcione espaço suficiente para um desenvolvimento mais aprofundado. Ainda que muitos designers possivelmente não compartilhem um interesse profundo por economia, direito ou política internacional, explorar estes aspectos através de projetos de design especulativo representa uma oportunidade significativa para inserir a agenda sobre o Antropoceno na pesquisa e no ensino do design. Investigar as implicações políticas, geográficas e econômicas do Antropoceno possibilita construir um terreno robusto (e muito urgente) para questões abrangentes de design especulativo.

Também é neste sentido que o design especulativo pode extrair insumos da ficção científica. Especular sobre realidades vindouras exige imaginar diferentes pressupostos estruturais, a partir dos quais sociedades e culturas explicam suas condições de vida. O design especulativo pode imaginar os hábitos, gestos, pontos pacíficos e problemas que a interação com tecnologias emergentes pode estimular ou restringir, levando à reflexão sobre projetos em desenvolvimento e sondando procedimentos alternativos para potencial implementação. É também neste sentido que outras disciplinas e áreas do conhecimento podem recorrer às inspirações mais especulativas do design, na medida em que tais projetos permitem que as incursões teóricas de cada campo epistemológico sejam atualizadas e imbuídas de nova energia – especialmente se pensarmos nos desafios e tensões particulares que a questão do Antropoceno traz para as mais diversas disciplinas.

Daleko também demonstra que o entranhamento contemporâneo entre design e ficção científica pode evitar repetir tendências de senso comum sobre o pensamento tecnológico. Pelo contrário, os modelos mais experimentais alimentam-se de um compromisso intenso com os desenvolvimentos científicos mais avançados. Isto não é tão diferente do que a chamada "ficção científica dura" sempre fez, o que, neste caso, exige dos designers uma constante atualização em relação aos diversos desenvolvimentos dos empreendimentos científicos contemporâneos. Por exemplo, para tratar da questão do lixo e dos resíduos sólidos na era do Antropoceno, os autores de *Daleko* mergulharam em pesquisas recentes das áreas de Geologia, Bioquímica, Astronomia, Biologia Marinha, Robótica, Ciências da Computação, Geoengenharia, Ecologia e Engenharia Agrícola. Esta é também uma possibilidade de desenvolver na prática, em torno do design especulativo, as ambições (muitas vezes mais discutidas do que implementadas) de projetos multi e transdisciplinares.

Algumas objeções poderiam ser feitas no sentido de que as histórias de ficção científica propostas por *Daleko* apresentam uma visão teleológica sobre a ciência e o progresso tecnológico. Em nosso entendimento, porém, o projeto sugere o oposto: que a tecnologia não se desenvolve em uma linha reta e, à luz do impacto antropogênico sobre a Terra, também é tarefa para os designers abraçar a artificialidade, com a longa duração das conformações geotécnicas, mas também a transitoriedade dos

arranjos tecnopolíticos em mente. Como diria Bernard Stiegler (in Dunker, 2021), parafraseando uma velha fórmula filosófica, as tecnologias são, potencialmente, veneno e cura. Para torná-las curativas, seus usos e aplicações práticas devem ser administrados com cuidado. Durante muito tempo, o poder regulador dos Estados soberanos foi responsável pelo gerenciamento e pela moderação dos empreendimentos tecnológicos. No entanto, parece a cada dia mais claro que se os Estados continuarem a perder poder decisório nos desafios econômicos e culturais globais, será necessário descobrir ou inventar nos próximos anos diferentes formas de gerir eficazmente a tecnologia e seus efeitos. Com a crescente demanda por alimentos e diferentes tipos de produtos, a extração, o consumo e o descarte não tendem a diminuir, naturalmente. Engajar-se seriamente com novas formas de artificialidade parece necessário para a sustentação de formas viáveis e organizadas de coexistência. A extrapolação tecnológica pode depender, portanto, não apenas do design de tecnologias em si, mas também de como seus modelos são organizados e normalizados. O entrelaçamento entre design especulativo e ficção científica em *Daleko* opera em um ciclo retroalimentado, e sua concepção de planejamento repousa precisamente sobre estas duas instâncias: normalização das extrapolações, e extrapolação de normalizações.

Pode ser tentador relacionar o futuro imaginado em *Daleko* com a estética *solarpunk* (Schuller, 2019; Springett, 2017), movimento literário e artístico que tenta contrapor as visões sombrias sobre o desenvolvimento tecnológico a partir da proposição de cenários alternativos sustentáveis. A ficção científica solarpunk se opõe, propositalmente, aos fundamentos sombrios da estética *steampunk* e ao defasado *cyberpunk*, tentando produzir ficções científicas "mais otimistas". Críticas pertinentes têm sido feitas ao entusiasmo em torno do solarpunk (Raven, 2020; Zuin, 2020), sobretudo em relação a uma notável "falta de crítica" na avaliação das práticas científicas, uma conformidade passiva em relação à inovação, ou pela simplificação dos problemas das utopias tecnológicas. Os cenários futuros de *Daleko* certamente podem ser relacionados ao imaginário solarpunk, mas isso não impede que o projeto explore arranjos mais complexos e pragmáticos situados entre regulação tecnocientífica, formulação de políticas e plataformas governamentais. A maioria dos contos correspondem a cenários onde o meio científico se enreda com agentes de governança bastante familiares – precisamente porque os projetos parecem incorporar, como um subtexto, as crises atuais de nossas instituições políticas consolidadas. As histórias procuram antever possibilidades para um contexto diferente, mesmo que mais intricado, em vez de apostar nos destinos utópicos e distópicos mais comumente imaginados (isto é, nem uma completa acomodação ideal, nem uma total ruptura em relação às atuais condições de coexistência).

Tal perspectiva apresenta-se como uma alternativa à tendência habitual de ficcionalizar o futuro dicotomicamente, recaindo em fetichismos ou catastrofismos tecnológicos. Em vez disso, o trabalho explora as sutis ambiguidades e os hábitos que caracterizam nossa relação cotidiana com infraestruturas tecnológicas. Os nove fragmentos de *Daleko* parecem todos mais próximos de estéticas solarpunk do que das projeções apocalípticas; porém, devido ao compromisso do projeto com a perspectiva sobre o design especulativo descrita na seção anterior deste artigo, seria enganoso pensar nestes contos apenas como projeções otimistas. É importante considerar, portanto, como as nove histórias tensionam diversas brechas e limitações tecnológicas e jurisdicionais. Isto é importante pois através deste tensionamento os cenários imaginados são enquadrados como questões em aberto, cujo resultado depende não apenas do design, mas de arranjos contextuais mais amplos a serem decididos.

Desta forma, *Daleko* explora como o design especulativo pode se voltar para a construção de modelos de futuro imaginativos, mas ainda assim profundamente materialistas. Ao contrário do que se poderia pensar em um primeiro momento (e os contos nos ajudam a superar este impasse), isto não significa desapegar-se de um olhar criterioso sobre as normas e regulações cotidianas, mas, pelo contrário, mergulhar com um comprometimento efetivamente racional sobre a articulação tecnológica e política destas normatividades. Isto leva o projeto a especular sobre cenários de normatização bastante inventivos, por vezes repletos de ambiguidades. Trata-se de uma tendência que reflete a abordagem particular do programa Terraforming ao design especulativo, enquanto um método que deve apelar à *necessidade*. Um método que não tenha receio de tocar em questões sensíveis, mas que sobretudo seja pragmático no sentido de "identificar os parâmetros daquilo que é necessário, operando a partir desta

projeção até a realidade presente, da forma mais aberta possível em relação aos meios que seriam necessários para alcançar os objetivos traçados" (Bratton, 2021: s.p.). Enquanto extrapola programaticamente as situações cotidianas de impasse em relação a tecnologia e governança, *Daleko* imagina as infraestruturas futuras com este grau importante de pragmatismo, considerando que o design é capaz de desencadear uma série de extrapolações adicionais nos ambientes rurais e urbanos. O projeto fornece imagens alternativas que permitem discutir o papel do design nestes contextos, mais especificamente em relação a questões de gestão de resíduos e mitigação da mudança climática. Isto, por si só, já é motivo suficiente para fazer avançar esta agenda de pesquisa, sobretudo tendo em vista que já há dados, imagens, estimativas e projeções suficientes de futuros em que os efeitos adversos do Antropoceno *não são* radicalmente confrontados, tecnológica e politicamente. Além disso, o projeto cumpre ainda um importante papel epistemológico, ao estimular os designers a imaginar não apenas cenários futuros abstratos, mas também as possibilidades efetivamente materiais de nosso atual aparato tecnológico – dois aspectos fundamentais para a teoria e a prática do design contemporâneo (Luersen, 2020). Por todas as razões elencadas acima, entendo que abordagens ao design especulativo como esta podem ter um papel muito importante em contextos de ensino do design, um argumento que procuro expandir na seção seguinte.

O design especulativo em contextos educacionais

A extrapolação simultânea de aspectos tecnológicos e jurisdicionais através da ficção científica contribui, portanto, para o exercício substancial de especular sobre futuros evitáveis ou realizáveis. Com isso, os projetos de design especulativo também estimulam imaginar qual será o papel do design em tais horizontes. Trata-se, portanto, de um ponto de entrada para entender o design não apenas como um ofício especializado, mas também em seu aspecto epistemológico mais amplo, com ramificações substanciais nas práticas organizacionais, na gestão pública, e mesmo em questões de soberania e governança ambiental. Neste sentido, como Dunne & Raby ressaltam, "[um] projeto de design especulativo também pode servir para "sondar" os limites legais e éticos dos sistemas existentes" (2013: 57). Trata-se, portanto, de uma forma fascinante de discutir qual o papel exercido pelo design nestes limites, como também em sistemas alternativos, com suas próprias limitações em potencial. Realçar todas estas questões é particularmente desejável em contextos de educação em design. Em termos de questionar a utilidade do design em relação aos problemas multidimensionais de um mundo complexo (Cardoso, 2016), em que problemas de escala local e global muitas vezes se sobrepõem, o design especulativo se torna um método interessante para o aprofundamento teórico de disciplinas de projetos aplicados. Quando adotada na produção de briefings e protótipos, ele pode amplificar questionamentos sobre a variedade de papeis desempenhados pelo design em relação a problemas contextuais.

Diante da necessidade de agir coletivamente em questões urgentes relacionadas ao Antropoceno, como a mudança climática, não basta que o campo do design produza discussões teóricas sem imaginar iniciativas de ação programática – assim como também não basta, nesta delicada conjuntura, agir sem a orientação de fundamentos conceituais. Como coloca Anders Dunker (2021: 4), é preferível seguir o velho conselho pedagógico de "aprender a aprender", efetivamente, mas isto deve ser complementado substancialmente. As iniciativas políticas sérias devem enfrentar os verdadeiros obstáculos endógenos e exógenos às ameaças climáticas. Afinal, se as ciências naturais nos informam sobre os fatores decisivos da crise ambiental e, com isso, permitem estimar cenários futuros por meio de diversas projeções científicas, também é verdade que, na prática, a mitigação dos efeitos da mudança climática global só pode ocorrer a partir da normalização de outras práticas antropotécnicas. Se é de fato correto afirmar que precisamos descobrir ou inventar formas adequadas de convivência com outras espécies (Haraway, 2016) nos ambientes destruídos ou arruinados de nosso planeta (Tsing, 2015), as alternativas especulativas devem ultrapassar a esfera retórica dos discursos morais e materializar-se gradualmente na prática cotidiana enquanto formas de conhecimento técnico aplicado – isto é, dentre outras coisas, no design. Estas questões já se manifestam publicamente, mas isto por si só não é suficiente. Para permear adequadamente a vida prática, eles devem ultrapassar a camada discursiva, penetrando nas disciplinas estabelecidas. Em termos do ensino de design, isso significa incorporar conhecimentos das ciências da

Terra e do Espaço sobre questões ambientais nos projetos que empreendemos, por exemplo, envolvendo o pensar projetual com a dinâmica metabólica das manifestações geomórficas. Para aceitar este desafio como projetistas devemos fazer o que fazemos – articular artificialmente materiais e modelos – combinando as práticas técnicas e teóricas. A oportunidade para efetivamente incorporar este conhecimento em hábitos continuados de design (culturas de design) parte da sua integração no planejamento e na prototipagem desde as práticas pedagógicas, inicialmente. Quer seja aplicado posteriormente a atividades do setor público ou privado, esta é uma forma de problematizar desde o princípio as concepções do design enquanto uma atividade de curto prazo, com preocupações e motivações mais imediatas, para reconsiderá-lo mais adequadamente em relação à camada artificial do planeta pela qual ele é em boa medida responsável. Como coloca Cristina Parreño (2020), esta compreensão superficial (mas amplamente compartilhada) do tempo, que desconsidera os efeitos de longa duração de nossas atividades, também afasta as práticas correntes de uma avaliação mais séria dos efeitos profundos da ação antropogênica na Terra. O que está em jogo ao projetar produtos industriais, escolher um conjunto de materiais, desenvolver sistemas e técnicas, ou planejar novas infraestruturas rurais e urbanas, pode se manifestar visivelmente no curto prazo, mas irremediavelmente se estende para além dele. Acreditar no oposto seria o mesmo que esquecer a história antropotécnica de nossa espécie.

Portanto, em paralelo à aprendizagem sobre como planejar e modelar sistemas, serviços e produtos, a adoção do design especulativo como método no ensino do design oferece uma oportunidade para avaliar criticamente as potenciais responsabilidades de longo prazo, de nosso campo. Um ensino do design voltado a projetos de longo prazo deve necessariamente enquadrar as competências teóricas e profissionais do design em relação a diferentes matrizes científicas e contextos sociopolíticas. A abordagem ao design especulativo discutida e descrita aqui através do caso *Daleko* é particularmente significativa quanto a estes aspectos, devido tanto à sua aproximação de preocupações da ciência contemporânea quanto por sua percepção dos efeitos catalíticos do design na sociedade. O engajamento mútuo com estas perspectivas ajuda a compreender a ampla gama de materiais e meios técnicos à disposição para que designers se envolvam em projetos de infraestruturas sustentáveis, o que também requer considerar como tais projetos podem auxiliar, desestimular, adaptar ou reorganizar hábitos, regulações e modos de vida. As tecnologias de sensoriamento de ecossistemas, monitoramento de emissões, rotação de culturas, reutilização de materiais, energias renováveis, restauração de biomas etc., todas têm um papel a desempenhar na orientação de formas de vida viáveis e interdependentes em sociedade. Entretanto, embora parte destas práticas não sejam novas, elas ainda se encontram pouco conectadas a aspectos regulatórios e normativos de interesse comum. Da automação industrial e aprendizagem computacional à engenharia genética, da inteligência artificial geral aos sistemas de sensoriamento remoto em rede e à manufatura em nuvem, as diferentes disciplinas do design podem encontrar uma gama variada de materiais distintos para pensar infraestruturas contemporâneas sustentáveis. A fim de transformar tais contribuições em algo que seja coletivamente benéfico e eficaz a longo prazo, os designers não devem abster-se de explorar a potencial fusão dos projetos com iniciativas mais duradouras da sociedade, tais como instituições políticas, órgãos públicos, plataformas de governança, organizações da sociedade civil, e assim por diante. Isto é relevante em contextos de ensino pois estimula os designers a serem um pouco menos autorreferentes, a fim de entrelaçar as possibilidades materiais e os desafios socioambientais da atualidade com uma imaginação extrapolativa, mas que ainda assim assegure um comprometimento pragmático com consequências de longo prazo.

Com isto em mente, pode-se concluir que a implementação de projetos especulativos como *Daleko* em contextos de ensino de design contribuem para atualizar e reformular questões urgentes relativas aos processos decisórios do design e suas consequências programáticas, em cenários inicialmente experimentais e hipotéticos. Enfatizo aqui o termo "experimental" porque o planejamento de tais infraestruturas extrapolativas oferece a oportunidade de enfrentar questões particularmente difíceis, incomuns, ambíguas ou até controversas, que de outra forma seriam dificilmente colocadas. Neste aspecto, os projetos especulativos não devem ser travados pelos nossos dilemas morais cotidianos, pois tais impasses podem ser particularmente benéficos para a discussão dos modelos. Quanto mais um projeto permitir explorar as franjas e os limites éticos, legais e práticos dos programas, serviços, sistemas e objetos projetados, mais adequado ele é para explorar junto com estudantes e pesquisadores os papéis

reais e potenciais do design, bem como as suas limitações. Em *Daleko*, isto é mais claramente ilustrado pela história *Nuclear Ikea*, uma vez que a ideia de investir na tecnologia nuclear como fonte de energia renovável ainda é algo controverso entre diversas correntes políticas e ambientalistas, e mesmo no senso comum. Como os designers devem se engajar nestas discussões? O que poderíamos aprender para nossa prática a partir do estudo científico da energia nuclear como matriz energética? Quais são os aspectos técnicos e não técnicos que devem ser levados em conta em sistemas energéticos envolvendo reatores nucleares? Quais são as implicações do uso da energia nuclear para a eficiência energética, e quais poderiam ser suas consequências ambientais e políticas colaterais? Como tais implicações, subjacentes a acordos multilaterais, se relacionam com aspectos do design? Como as infraestruturas de gestão energética podem ser projetadas não apenas em acordo com normas e regulamentos internacionais atuais, mas também com outras regulações técnicas e políticas potenciais? A complexidade das questões tende a aumentar em paralelo com o processo de desenvolvimento dos projetos especulativos.

O design especulativo pode ser evocado como um método em diferentes graus de ensino do design, uma vez que a sondagem de cenários imaginativos favorece a ponderação de importantes questões transversais desde estágios iniciais do projeto, como o briefing, até a pós-vida do protótipo, estimulando o aprofundamento dos modelos hipotéticos ao longo do percurso.

Considerações finais

Enquanto alguns autores argumentam que o design especulativo deveria permitir aos designers ultrapassar as habituais restrições que demarcam os processos normativos do design (Auger, 2016), argumentei aqui que, em vez disso, os exercícios ficcionais do design podem enriquecer-se ao compreenderem tais restrições, imaginando como os projetos especulativos poderiam explorá-las (como uma alternativa para perscrutar, inclusive, seus potenciais efeitos colaterais não controlados e acidentais). Afinal, o próprio desenvolvimento, e as consequências dos processos antropogênicos subsequentes, deve acarretar no surgimento de novos processos normativos, sejam eles intencionais ou involuntários. Independentemente de nossos sentimentos em relação a elas, as dimensões normativas são intrínsecas ao design (sejam elas promissoras ou contestáveis) e, portanto, devemos explorar abertamente esta condição em contextos educacionais, enquanto uma oportunidade pedagógica enriquecedora.

Efetivamente, há abordagens muito diferentes entre si quanto ao design especulativo. Algumas delas tendem a refutar tal dimensão normativa e, portanto, se abstém de discuti-la em seu mérito. É também por isso que perspectivas como a de *Daleko* são importantes para a ensino e a pesquisa em design. Os problemas endêmicos que envolvem as economias neoliberais, a soberania dos Estados-Nações e as crises ambientais contemporâneas não serão atenuados naturalmente, com o agravante de que os três se sobrepõem, em um amontoado de prioridades conflitantes e temporalidades processuais distintas. Nos próximos anos, as disciplinas compositivas, o aparato normativo e as iniciativas civis terão que dedicar muita energia a questões de restauração, reforma, regeneração e engenharia reversa. Sob esta circunstância, desenvolver a capacidade de realizar projetos levando em conta cronologias mais largas e compreensões materialistas mais rigorosas deve ser importante até mesmo para as práticas especulativas, se desejarmos que as infraestruturas, regulações e demais intervenções futuras sejam de fato resilientes.

Não se trata, portanto, de mobilizar-se em oposição ao antropoceno. O reconhecimento da época em que a agência da espécie humana exerce um impacto geológico sobre a Terra é apenas um primeiro passo[10], ainda que tardio, para que saibamos melhor ponderar e desenvolver, também enquanto espécie, modos de viver que sejam mais inteligentes e sensíveis em relação à manutenção nossa e de outros seres que compartilham conosco um minúsculo recorte espaço-temporal na história do planeta. Neste sentido,

---

[10] Cabe sempre lembrar que, historicamente, o humano é apenas um entre tantos outros agentes capazes de gerar alterações tão substanciais. Um exemplo recorrente, mas ainda bastante ilustrativo, é o da superpopulação de cianobactérias que desencadearam o fenômeno chamado de Grande Evento de Oxigenação (GEO), durante o período paleoproterozoico, quando se extinguiu quase toda a vida na Terra e, ao mesmo tempo, gerou-se também a possibilidade do desenvolvimento de organismos multicelulares.

trata-se menos de enfatizar a excepcionalidade da agência humana em relação à mais larga temporalidade das épocas geológicas do que de reconhecer a inequiparável velocidade das transformações derivadas da ação antrópica no século mais recente. Cabe lembrar que, não por acaso, a institucionalização do design enquanto atividade industrial desemboca no período da Grande Aceleração do Pós-Guerra (McNeill & Engelke, 2016), em que materiais como o petróleo e o concreto cumpriram um papel fundamental no desenvolvimento da infraestrutura das sociedades humanas contemporâneas (sobretudo a partir da produção e distribuição em massa de produtos plásticos, do projeto de meios de transporte e de formas de habitação baseadas em modelos com alta demanda material e energética). É também por isso que é tão importante que a capacidade fictiva do design especulativo se mescle com outras disciplinas. Isto não precisa necessariamente ocorrer pela absorção de outras áreas por parte do design, como argumentado ao longo deste trabalho, mas também no sentido oposto: seria de grande aproveitamento para outras disciplinas, das ciências naturais às humanas, da biologia à economia, da astronomia ao direito, que pudessem aproveitar de tal capacidade imaginativa para repensar como seus princípios podem se orientar às necessidades dos diversos ambientes humanos e não-humanos no antropoceno.

Se há muito se argumenta pela interdisciplinaridade como um princípio epistemológico necessário para o desenvolvimento de conhecimentos complexos (Morin, 2001), é também nos espaços não preenchidos e nos projetos não implementados à luz destas necessidades, que a ficção e o trabalho com imagens, projetos e modelos, podem encontrar um ambiente fértil. A ficção pode agir como um catalisador na elaboração de modelos, imagens e projeções, ancorada na própria especulação, quando as necessidades mais específicas e prosaicas de cada disciplina não deixarem espaço à atividade fictional e seu potencial transfigurativo.

Neste mesmo sentido, os contos ficcionais de *Daleko* permitem ver que só é possível produzir modelos e protótipos diferentes se, em vez de conceituarmos os resíduos sólidos como uma forma de externalidade, sua durabilidade profunda for reimaginada, epistemologicamente, enquanto parte integral do metabolismo planetário. Sob uma abordagem biogeoquímica ao problema da gestão do lixo, os projetos, leis, tratados e intervenções técnicas devem ser corresponsáveis pela condução de formas viáveis de gerenciar os resíduos, não mais como recursos a serem consumidos e descartados indefinidamente.

Resumidamente, entrelaçar especulativamente design e ficção pode fortalecer as possibilidades de problematização das condições materiais reais a partir de imagens e modelos alternativos, sobretudo no que refere ao desenvolvimento tecnológico e sua interface com normas e regulações de governança. Assim como a ficção científica, projetos de design especulativo intrigantes produzem suas extrapolações a partir de conceitos e teorias científicas, para fabular desde os futuros imaginados em direção ao presente, sem se abster de explorar a incerteza, testar limites e levantar questões notavelmente complicadas.

Referências

Archer, D. (2009). *The long thaw. How humans are changing the next 100,000 years of Earth's climate.* Princeton. Princeton University Press.

Auger, J. (2016). "Interviews: Discourse". *In*: Mitrovic, I., & Šuran, O. *Speculative. Post-design practice or new utopia?* Zagreb. Ministry of Culture of the Republic of Croatia & Croatian Designers Association.

Baldé, C. P., Forti V., Gray, V., Kuehr, R., & Stegmann, P. (2017). *The global e-waste monitor.* Bonn/Geneva/Vienna. United Nations University (UNU), International Telecommunication Union (ITU) & International Solid Waste Association (ISWA). https://www.itu.int/en/ITU-D/Climate-Change/Documents/GEM%202017/Global-E-waste%20Monitor%202017%20.pdf

Basel Action Network. (2018). *The 'scam recycling' continues. E-waste exportation from the U.S. to developing countries* (update #2). https://wiki.ban.org/images/1/17/ScamRecycling-ContinuesUpdate_2.pdf

Bleecker, J. (2009). *Design fiction . A short essay on design, science, fact and fiction*. Los Angeles. Near Future Laboratory.

Bleecker, J. (2017, Nov.). *Stanford seminar. Design fiction* [video]. https://libguides.murdoch.edu.au/APA/social_media

Bratton, B. (2021, Jan.). "Benjamin H. Bratton on terraforming the world order". *Palladium*. https://palladiummag.com/2021/01/11/benjamin-h-bratton-on-terraforming-the-world-order/

Bratton, B. (2016). "On speculative design". *Dis Magazine, the time complex. Postcontemporary issue*. http://dismagazine.com/discussion/81971/on-speculative-design-benjamin-h-bratton/

Bratton, B. (2019). *The Terraforming*. Moscow. Strelka Press.

Cardoso, R. (2016). *Design para um mundo complexo*. São Paulo. Cosac Naify.

Chakrabarty, D. (2018). "Anthropocene time". *History and Theory*. 57:1. 5–32.

Dunker, A. (2021). *Rediscovering Earth. Ten dialogues on the future of nature.* New York. OR Books.

Dunne, A., & Raby, F. (2013). *Speculative everything. Design, fiction, and social dreaming*. Cambridge. MIT Press.

Edison, T. A. (1878). "The phonograph and its future". *North American Review.* 126:262. 527–36.

Gabrys, J. (2013). *Digital rubbish. A natural history of electronics.* Ann Arbor. University of Michigan Press.

Haraway, D. (2016). *Staying with the trouble. Making kin in the Chthulucene*. Durham and London. Duke University Press.

Hopson, E., & Puckett, J. (2016). *Scam recycling. E-dumping on Asia by US recyclers*. http://wiki.ban.org/images/1/12/ScamRecyclingReport-web.pdf

Jucan, I. (2019). Introduction. Remain X Remain(s). *In*: Jucan, I., Parikka, J., & Schneider, R. *Remain*. Lüneburg. Meson Press.

Haram, L. E., Carlton, J. T., Ruiz, G. M., & Maximenko, N. A. (2020). "A plasticene lexicon". *Marine Pollution Bulletin*. 150 (Jan. 2020). https://doi.org/10.1016/j.marpolbul.2019.110714

Luersen, E. H. (2014). *Comunicação, Indústria Fonográfica e Tecnologia: A Jornada do Artista Autônomo Através dos Novos Caminhos da Produção Musical*. Saarbrücken: Novas Edições Acadêmicas.

Luersen, E. H. (2020). *Ressonância Tecnocultural: Rastros da Ambiência Contemporânea nas Sonoridades dos Jogos Digitais* [Doctoral Dissertation, University of Vale do Rio dos Sinos]. Repositório Digital da Biblioteca da Unisinos. http://www.re-positorio.jesuita.org.br/handle/UNISINOS/9316.

Luersen, E. H., & Fuchs, M. (2021). Ruins of Excess: Computer Game Images and the Rendering of Tech-nological Obsolescence. Games and Culture: A Journal of Interactive Media. Advance Online Publication. doi.org/10.1177/15554120211005363.

Markussen, T., & Knutz, E. (2013). "The poetics of design fiction". *In*: *Proceedings DPPI '13.* ACM. New York. 231–240.

Near Future Laboratory (2012, Ago. 25). *A digital tomorrow* [video]. https://vimeo.com/48204264

Parikka, J. (2015). *A geology of media*. Minneapolis. University of Minnesota Press.

Parikka, J. (2014). *The Anthrobscene.* Minneapolis. University of Minnesota Press

Parreño, A. (2020). "Deep timescales of our most urgent crises". *Strelka Mag. The revenge of the real issue*. https://strelkamag.com/en/article/deep-timescales-of-our-most-urgent-crises.

Peloušková, K. (2020). "Lesk a bída spekulativního designu". *Artalk Revue*. Abril. https://artalk.cz/2020/04/27/lesk-a-bida-spekulativniho-designu/cs/#_ftnref14

Peres, E., Sinitsyna, A., Kostandyan, T., & Nosov, T. (2020). *Daleko*. Strelka Institute for Media Architecture and Design. https://*Daleko*.space/

Pias, C. (2011). "On the epistemology of computer simulation". *Zeitschrift für Medien und Kulturforschung*. 2:1. 29–54.

Raven, P. (2020, Abr. 9). "Necessary but not sufficient. On hope and optimism in solarpunk and cyberpunk". *Velcro City*. https://www.velcro-city.co.uk/necessary-but-not-sufficient-on-hopeand-optimism-in-solarpunk-and-cyberpunk/

Schuller, W. K. (2019). "Evolution takes love. Tracing some themes of the solarpunk genre". [*Dissertação de mestrado*. Queen's University]. QSpace. https://qspace.library.queensu.ca/bitstream/handle/1974/26518/Schuller_William_K_201909_MA.pdf?sequence=2&isAllowed=y

Shaw, D. B. (2011). *Technoculture. The key concepts*. London. Routledge.

Springett, J. (2017, Fev. 26). "Solarpunk. A reference guide". *Solarpunks*. https://medium.com/solarpunks/solarpunk-a-reference-guide-8bcf18871965

Stengers, I. (2014). "Gaia, the urgency to think (and feel)". *In*: *Colóquio Internacional os Mil Nomes de Gaia. Do Antropoceno à Idade da Terra*. Casa de Rui Barbosa. Rio de Janeiro. Sep. 15-19, 2014. https://osmilnomesdegaia.files.wordpress.com/2014/11/isabelle-stengers.pdf

Sterling, B. (2012). "Bruce Sterling explains the intriguing new concept of design fiction". *Slate*. http://www.slate.com/blogs/future_tense/2012/03/02/bruce_sterling_on_design_fictions_.html

Tarde, G. (2018). *Fragments d'histoire future*. New York. Wentworth Press.

Tooze, A. (2020). "After escape. The new climate power politics". *E-Flux*. 114(Dez. 2020), s.p.

Tsing, A. (2015). *Mushroom at the end of the world. On the possibility of life in capitalist ruins*. Princeton. Princeton University Press.

Veiga, J. E. (2014). *O imbróglio do clima. ciência, política e economia*. São Paulo: SENAC.

Zalasiewicz, J., Waters, C. N., Sul, J. I., Corcoran, P. L., Barnosky, A. D., Cearreta, A., Edgeworth, M., Gałuszka, A., Jeandel, C., Leinfelder, R., McNeill, J. R., Steffen, W., Summerhayes, C., Wagreich, M., Williams, M., Wolfe, A. P., & Yonan, Y. (2015). "The geological cycle of plastics and their use as a stratigraphic indicator of the Anthropocene". *Anthropocene*. 13(Mar. 2016). 4–17.

Zuin, L. (2020, Nov. 26). "Vai dar ruim: além da estética, distopias são advertências sobre futuro". *UOL TAB*. https://tab.uol.com.br/colunas/lidia-zuin/2020/11/26/distopias.htm

SCIENCE FICTION LITERATURE AND ITS ROLE IN SOCIETY, RESEARCH, AND ACADEMIA

INTERVIEW WITH PROF. WU YAN

LITERATURA DE FICÇÃO CIENTÍFICA E SEU PAPEL NA SOCIEDADE, PESQUISA E ACADEMIA

Entrevista com prof. Wu Yan

LA LITERATURA DE FICCIÓN CIENTÍFICA Y SU PAPEL EN LA SOCIEDAD, LA INVESTIGACIÓN Y LA ACADEMIA

Entrevista con el prof. Wu Yan

Nicola Liberati
(Shanghai Jiao Tong University, Department of Philosophy School of Humanities, China)
liberati.nicola@gmail.com

Wu Yan
(Southern University of Technology and Science, China)
1742201664@qq.com

Wu Yan is a science fiction writer and professor in the School of Humanities and Social Sciences at Southern University of Science and Technology. He serves as the Director of Science and Human Imagination Research Center under the university since 2017. He has written many fictions including Spiritual Quest (Xinling tanxian), Life and Death of the Sixth Day (Shengsi di liutian) and China Orbit (Zhongguo Guidao Hao). His short stories such as Mouse Pad have been translated into many languages. He has received many awards including the National Outstanding Children's Literature Award, The Five Best Cultural Works Project Prize, Bing Xin Literature Award, the Chinese Nebula, the Galaxy Award for Science Fiction, and the Thomas D. Clareson Award 2020. He has also been engaged in academic writing such as Outline of Science Fiction Literature (Kehuan wenxue lungang). He is editor of numerous scholarly works and has co-edited in English with Kerry Mallan and Roderick McGillis on (Re)imagining the world: Children's Literatures Response to Changing Times. Wu Yan is the ancestor of science fiction education and create sf curriculum in 1991 in Beijing Normal University. He is both the Vice-Chair of China Science Fiction Professional Committee of China Writers' Association and Science Writers' Association (CSWA) and was the President of World Chinese Science Fiction Association (WCSFA) during 2010-2017.

Thank you very much for accepting the interview on such an important topic as the role of science fiction in our era. As you know, my research is on the effects of digital technologies, and I think science fiction is a valuable source to explore the implementations of new digital technologies in our society (Liberati 2022; Bonfiglioli 2021; Mykhailov and Liberati 2022).

Thanks to the conference LINE2021 organized at *Shanghai Jiao Tong University*,[1] we followed the presentation by many researchers working on science fiction and new digital technologies in China and Italy. You were one of the two novelists invited as keynote speakers. Your presentation was about the interactions we can have between science-fiction, academia, and society. Primarily, you showed how science-fiction grew in China in the last decade.

In line with the theme of your presentation, I would like to ask you three main questions covering the role of science fiction in our world at different levels. The three questions cover the relations science-fiction has with our society in general, the research on the social impact of technologies, and the research and education within academia. There will also be other questions stemming from our discussion, but the main topics will be related to these three main elements.

1. *Nicola Liberati: How do you think science fiction contributed and is contributing to the development of China?*

Wu Yan: Theoretically, this genre should be very helpful for developing a country. Especially in the specific period related to the modernization of a country and the post-modernized society, science fiction should be taken seriously into account. However, most of the time, this happens only in theory.

I think we can clearly understand the influence of science fiction if we look at its history in China compared to what happened in the USA.

In the USA, science fiction has a long tradition rooted in the 1920s, a period before its golden age. The main idea of the science fiction designed by the authors of this period was to ground this genre as a "real science" by using the method of science fiction to have an impact on society. This aim is clear if you read authors like John W. Campbell, Isaac Asimov, Robert A. Heinlein, L. Ron Hubbard, and the Golden Age of Science Fiction. However, this project failed in some degree.

China is trying to develop a different way of writing science fiction that is not related too much to the idea of building a new science. Moreover, in China, the government did not look at science fiction as a valuable asset for the country's development. This genre was considered popular science. However, especially after 1980, some authors did not agree with this position. At this time, there was a boom in Chinese science fiction, but the government did not follow the development of this genre and its growth since it looked at it as a way of spreading pseudoscience.

Only recently, in the last couple of years, especially after *The three-body problem [三体]* got the *Hugo award*, the government changed its perspective, and it started to support this genre as a kind of good literature and a stimulus for imagination.

2. *Nicola Liberati: Are there any images and ideas driving China's development?*

Wu Yan: Between 1990 and 2010, the idea of sci-fi changed in society. People started to think of science fiction as something new and good literature and not as popular science or dissemination of science. Thus, our idea and our work have been received by society, and this change means we can generate new ideas to affect society.

During this period, we conducted the project *The Relationship between Science Fiction and National Independent Innovation,* and we introduced some of the information from different parts of the world,

---

[1] LINE2021 - Living in the New Era (时代·新生): Technologies, Creativity, and Science-Fiction https://shss.sjtu.edu.cn/En/List/345

like NASA and DARPA from the United States, to show the essential effects science fiction can have. This approach is the theoretical way of further separating science fiction from science communication.

The Three-Body Problem helped. For example, high-tech companies and CEOs got in contact with this new genre thanks to *The Three-Body Problem*. Thus, thanks to our success, in 2017, the government started to support science fiction directly by using their way, and it began to help it with activities like organizing the annual science fiction meeting. Now science fiction is taken as good literature, so it can be a way to develop novel ideas for society.

3. *Nicola Liberati: Do you think science fiction can be used to explore the potentialities and the effects of new technologies?*

Wu Yan: Science fiction literature and our future generally have significant overlaps. This afternoon, I will give a lecture on science fiction to students in engineering. It will be the first time to present science fiction to engineers at my university, and I will present science fiction as a kind of "design" process to show the ability of human beings. I want to show them how science fiction can answer questions like: What are the capabilities we can have? What did we create and build in our world, and how can we mix these elements to create something new?

I see science fiction as a prototype, a scenario, a design fiction, knowledge, and a source for ideas to study our world. Science provides us with a transparent methodology for reasoning how to approach and explore a theme through passages. Science fiction gives us "narratives" as layers that develop through time. Thus, through narratives, science fiction examines the world and human beings not through reasoning but by using time as the main force of the analysis.

Reasoning and narrative results are two different ways to tackle questions and topics. Both of them are needed to deal with the issue of the future.

4. *Nicola Liberati: How do you think the aim is different between reasoning and narrative results?*

Wu Yan: I think that, in science, researchers use abstractions to study a topic. Objects are idealized, isolated, and structured clearly. However, this is just an abstraction of things. In the "real" world, everything is intertwined, it happens in time, and it relies on each character's personality.

Each element develops its narrative by intertwining with others. They are related to the life of each specific people, and each setting provides different solutions and results. Thus, we can have millions of stories that are all valuable and aimed at studying a possible configuration of events. We can have multiple futures in front of us simultaneously, and they can be studied through the narratives provided by science fiction, which focuses on the different configuration of elements and it works on how time make these elements develop by themselves.

While sciences are about knowledge, narratives are about time. While sciences follow rationality to reach a conclusion, narratives follow time and people's choices. A story develops like the life of people who pass through different events one after the other.

Science is a web of logical lines based on knowledge. Humanities are scattered collages. Narratives are meshes that are generated through interpersonal communication and individual development. The difference between the two objects is noticeable. From the scientific point of view, the narrative web has many apparent inadequacies. From the narrative point of view, what scientists think is inappropriate is precisely the result of human behavior. Thus, science and narrative lead to different futures. In other words, the narrative is also one of the ways to reach the future.

5. *Nicola Liberati: How are you embedding science fiction within the University as a field of research?*

Wu Yan: This question directly relates to our trajectories in how to build projects and what kind of projects we want. Unfortunately, science fiction in research has not a great resonance yet. If you look at my past, it is clear it is not an easy task to get into academia by working on science fiction. I am a psychologist who taught organizational psychology at *Beijing Normal University*. In 1991, I started to teach a cross-disciplinary course on science fiction in my spare time as a hobby. Many came to attend my classes, so I began writing articles even if it was tough to get publications on science fiction accepted in prestigious academic journals.

In 2003, after more than ten years, a Professor of children's literature became the Vice-Chair of the Chinese Language and Literature Department. He asked me to join him in building a center for children's literature inside his department. I happily joined his center once I realized science fiction could be developed in this program. At that time, science fiction studies and children's literature were parallel. "Children's literature" began to grow, and science fiction had the opportunity to grow with it. I did not care much if I was developing science fiction in children's literature or other fields. I got my first master's degree students in science fiction in 2003. This position helped immensely develop the idea of science fiction within academia with students working on this specific topic.

In 2007, I left Beijing to move to this University (*Southern University of Science and Technology*) which is a science and technology-oriented university, and the humanities are new.

The novelty of working in such a scientific context has pros and cons. At first, the work we do at our new center for humanities was perceived as a "side dish" to the main courses on science and technology. This approach also reflects the problems of finding funding to support our research.

However, there are also valuable pros to being a new center for humanities within a science and technology university. Our center is unique, which means we have complete freedom in developing our research in the direction we prefer. At Beijing University, some literature, science, and philosophy professors are interested in studying science fiction. I know in *Shanghai Jiao Tong University* science fiction is related to the history of science and technology, so it is natural that the research is grounded on such a background and is always related to that context. Since we are not building our research on top of other disciplines, we are free to experiment and try different approaches and ideas.

The idea guiding us is to follow the development of New Humanities, which is also the theme of your conference *LINE2021* and the focus of your research in philosophy and technology at SJTU. Instead of looking at science fiction as something on top of a rigid discipline, science fiction can be the idea fueling the research in the new humanities at its core.

This approach has already achieved results since I won US and Chinese awards. These achievements, and the achievements of my colleagues, show the University our value and the role we can play in the development of our future. More and more journals are starting to accept submissions in science fiction. So, it is clear that science fiction can play an essential role in defining what new humanities are and their effects on society.

References

Bonfiglioli, Cristina Pontes. 2021. "Methodological Challenges for a New Philosophy of Technology Interview with Nicola Liberati." *Prometeica*. https://doi.org/10.34024/PROMETEICA.2021.23.12411.

Liberati, Nicola. 2022. "Digital Intimacy in China and Japan: A Phenomenological and Postphenomenological Perspective on Love Relationships at the Time of Digital Technologies in China and Japan." *Human Studies* Forthcoming: 1–15. https://doi.org/https://doi.org/10.1007/s10746-022-09631-9.

Mykhailov, Dmytro, and Nicola Liberati. 2022. "A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives." *Foundations of Science*, March, 1–17. https://doi.org/10.1007/s10699-022-09833-5.

Wu, Yan. 2020 Meditations on Chinese Science Fiction Literature - Wu Yan's Academic Self-Selected Collection (中国科幻文学沉思录-吴岩学术自选集, [Zhongguo Kehuan Wenxue Chensilu-Wu Yan Xueshu Zixuanji])，Jie Li Press, 335pp

Wu, Yan. 2020. China Orbit (novel, 中国轨道号，Zhongguo Guidao Hao), Anhui Children's Press, 254pp

Wu, Yan. 2021. Outline of Science Fiction (科幻文学论纲，Ke Huan Wenxie Lungang), Chongqing University Press, 316pp

Wu, Yan. 2021. A Chinese Science Fiction Studies Reader (中国科幻文论精选，Zhongguo Kehuan Wenlun Jingxuan), with Jiang Zhenyu, Beijing University Press,281pp

Wu, Yan. 2021. China Science Fiction Development Year book 2021 (中国科幻发展年鉴2021，Zhongguo Kehuan Fazhan Nianjian 2021), with Cheng Ling, China Science and Technologies Press,347pp

Wu, Yan. 2022. History of Chinese Science Fiction in the 20th Century (20世纪中国科幻小说史，20 Shiji Zhongguo Kehuan Xiaoshuo Shi), Beijing University Press,235pp

SCIENCE FICTION LITERATURE AND ITS ROLE IN SOCIETY, RESEARCH, AND ACADEMIA

INTERVIEW WITH FRANCESCO VERSO

LITERATURA DE FICÇÃO CIENTÍFICA E SEU PAPEL NA SOCIEDADE, PESQUISA E ACADEMIA

Entrevista com prof. Francesco Verso

LA LITERATURA DE FICCIÓN CIENTÍFICA Y SU PAPEL EN LA SOCIEDAD, LA INVESTIGACIÓN Y LA ACADEMIA

Entrevista con el prof. Francesco Verso

Nicola Liberati
(Shanghai Jiao Tong University, Department of Philosophy School of Humanities, China)
liberati.nicola@gmail.com

Francesco Verso
(Science Fiction Writer and Editor of Future Fiction, Roma, Italy)
francescoverso@gmail.com

Francesco Verso is a multiple-award Science Fiction writer and editor. He has published: Antidoti umani, e-Doll, Nexhuman, Bloodbusters and I camminatori (made of The Pulldogs and No/Mad/Land). Nexhuman and Bloodbusters – translated in English by Sally McCorry – have been published in the US by Apex Books, in the UK by Luna press and in China for Bofeng. He also works as editor and publisher of Future Fiction, a multicultural project, dedicated to scouting and publishing the best World SF in translation from more than 25 countries and 11 languages with authors like Ian McDonald, Han Song, Ken Liu, Liu Cixin, Vandana Singh, Chen Qiufan, Xia Jia and others. From 2019 he's the Honorary Director of the Fishing Fortress SF Academy of Chongqing and a literary agent for Future Wave, an agency specialized in import/export of copyright from/to China. He lives in Roma (Italy) and may be found at www.futurefiction.org.

Thank you very much for accepting this interview on such an important topic as the role of science fiction in our era. As you know, my research is on the effects of digital technologies, and I think science fiction is a valuable source to explore the implementations of new digital technologies in our society (Bonfiglioli and Liberati 2021; Liberati 2022; Mykhailov and Liberati 2022).

Thanks to the conference *LINE2021 (Living in the New Era (时代・新生))* organized at *Shanghai Jiao Tong University,*[1] we attended the presentations by many researchers working on science fiction and

---

[1] *LINE2021 - Living in the New Era (时代・新生): Technologies, Creativity, and Science fiction* https://shss.sjtu.edu.cn/En/List/345

new digital technologies in China and Italy. You were one of the two novelists invited as keynote speakers. As you know, the other science fiction novelist invited as keynote speaker was Prof. Wu Yan because of his tight connection binding science fiction to society and academia. Both of your presentations tackled the interactions we can have among science fiction, academia, and society in an exciting way. He focused on China since his main activities and production are located in China. Your perspective was a bit different from your experiences in Italy and China. Your presentation showed how science fiction should be part of education to make people think critically. It should be considered in society because it helps society face the societal challenges related to implementing novel technologies.

In line with the theme of your presentation, I would like to ask you three main questions covering the role of science fiction in our world at different levels. These three questions have also been asked to Prof. Wu Yan, and they cover the relations science fiction has with our society in general, the research on the social impact of technologies, and the research and education within academia. There will also be other questions stemming from our discussion, but the main topics will be related to these three main elements.

1. Nicola Liberati: How do you think science fiction contributed and is still contributing to the development of China?

Francesco Verso: Science fiction has its value. In general, writers can create coherent world buildings by providing possible futures and hypotheses with different levels of plausibility. So, they can also forecast the applications and effects of technologies. In a way, science fiction writers tackle innovation by merging dramaturgy and developing new technologies.

This narrative can stimulate speculation concerning technological contexts grounded on the development of our society. It is not just an allegory. I think authors can create multiple hypothetic scenarios that produce a cloud of possibilities to provide suggestions on how to implement new technologies. In my work, I try to dialogue between science fiction writers and researchers on specific topics such as climate change in relation to the "Anthropocene" (Verso and Paura 2018) to give an idea of the future built upon different angles of the same phenomenon.

Then, to answer your question in a straightforward way: yes. I think science fiction is a valuable (practical?) way to tackle the technological development in society.

2. Nicola Liberati: How can we use imagination to merge different disciplines and have a perspective toward the future?

Francesco Verso: Especially in China, I have been frequently asked to show how to use imagination. I noticed people working in China think of themselves as oriented towards "following" more than developing "creativity." So, they ask me, an "Italian," to give them suggestions on how to be creative while writing a text.

I think imagination is based on two main pillars. The first element relies on the ability to question the reality we live in because by questioning the world, we understand it better. It can be just a general question in the form of "What if …?". For example, we all know how gravity works, but letus assume gravity works differently. What does it happen then? How will our society live and work according to this new element?

The second pillar is the freedom to generate different ideas. If I cannot express my hypotheses and I cannot test them in reality, I cannot make them part of the actual context.

I think imagination and this way of questioning our reality are essential for society because, side by side with the development of new technologies, we need the development of ethics and moral values. Without such a development, we will have problems understanding how to use the technologies, and we risk becoming our tools' tools.

Francesco Verso: The contrast between sciences providing knowledge about real things and science fiction providing imaginaries about possible things which do not exist yet is essential.

If sciences do not pose hypotheses, they are just the description of nature. If sciences do not ask questions, they are just dogmatic. At the same time, if literature does not consider the results in sciences by making them part of the narration, it becomes obsolete.

Sciences experiment in a controlled environment while science fiction experiments in an uncontrolled one. Science fiction generates ideas and memes, and it stimulates reflections on the applications of the technologies by being connected to emotions.

Thus, the two disciplines must talk to each other. This contrast and dialogue are the sources for new ideas and the ground to generate something new.

This is why I coined the new term "future fiction" to show how the future is not just about sciences but our culture in general. The future comes not only with the modification of an engine and the design of a new device related to sciences, but it comes with music, cooking, dance, and so on.

Science fiction is about giving people a future and seeing what they do with it. As Frederik Pohl said, science fiction does not focus on the flying car but on the traffic jams it generates.[2]

For this reason, I think a person who reads science fiction is developing antibodies towards the future and the effects of technologies.

4. Nicola Liberati: Thank you for this perspective on how science fiction can work to include different scenarios. I feel that science fiction can be read as a way to study and explore the future, but I also see many people enjoying the novel without developing a critical perspective on what is happening in our present. Is it possible to show how to blur the boundaries between the "fiction" in science fiction and what is happening in our society to develop people's critical thinking better? What elements in science fiction nudge the readers to "open their eyes" to the future and develop the "antibody" to tackle the societal challenges we all face?

Francesco Verso: I fear this problem is not just within science fiction, but it relates to arts in general. When a person looks at art thinking that the work of art is something disconnected from the world the person is living in, we have the same situation.

We need two main elements in order to avoid this situation. The first element is the freedom of the reader. The reader must feel free to act after receiving the stimulation to think critically.

When I started reading science fiction, I was doing it as a hobby, so I had two jobs simultaneously: a "primary" job and one related to science fiction. I decided to become a science fiction writer only when I had a problem with my "primary" job, which forced me to decide what I wanted to do in my future. Because of this situation and the fact I had some time to think and organize myself, I finally decided to change my life and make science fiction my "primary" and only job.

The other element is the ability to be engaged and feel responsible for what happens around the reader. So, education is important because people should feel empowered and able to contribute to society in some way. Even if the readers do not know what they can precisely achieve, they must feel the urgency

---

[2] "A good science fiction story should be able to predict not the automobile but the traffic jam" by Frederik Pohl

to do something in the world. The novel serves as the fire starter to trigger a person to start a process of change in society.

So, I think it is a much more complex topic that touches arts in general and even politics.

5. Nicola Liberati: Yes, I understand. Especially, I can see the relation to the call for action and critical thinking embedded in science fiction and the fact this element can be present in many other fields.

Francesco Verso: Yes, and I would like to open a parenthesis because there are two different souls within science fiction: science fiction for evasion and one more speculative.

When the author creates a world lost in a distant time and maybe in a different dimension, where there are fictional creatures with few connections with the world where the readers live, the plausibility of these settings is so low then it is easy to let the readers think that it is completely disconnected from their actual context. It is more about evasion than speculation of the possible futures the humanity might have.

However, the other kind of science fiction is designed to tightly link with the present world and the problems the readers face. For example, science fiction novels like 1984 are not for mere evasion. In these novels, it is pretty hard not to relate what is described in the book to the reality where the reader is living. The novel becomes a description of the actual world more than fiction. For example, Orwell wanted to make a political statement by publishing that book which can also be seen in the title of the book since it is "1984," which is related to the year "1948" when it was written.

6. Nicola Liberati: I like this idea of the distance between two worlds to make the connection possible. How important is the concept of "distance" for developing critical thinking in science fiction?

Francesco Verso: If it is an experience related to someone in particular on specific conditions not associated with the reader, the reader might not be interested in it.

However, the reader can easily engage if something different touches people's lives. For example, if the author talks about the changes in the solar exposition on earth and the possible consequences it might have on life on earth, the novel becomes part of the potential future the readers can connect with, and so, the reader can develop critical thinking through the use of science fiction.

People can normalize the future by making it visible and within their reach, thanks to science fiction. So, thanks to science fiction, people can get used to thinking of something which might happen soon. For this reason, it is understandable why so many companies and other actors have interests in science fiction too.

John Kessel said that science fiction does not talk about the year, place, and society where the novel is located, but it talks about the year, place, and society where the story is written. In a way, I agree with him that science fiction shows the relation to the actual society the novel has.

By describing a scenario, the author tries to express a future that may not be fully expressed in language. However, thanks to the art of translation between what the future the author thinks will be and the way the author describes it, the reader gets closer to these possible future scenarios. Thus, the authors need to find a way to express it to make the reader feel something, and, at the same time, the author has to make the reader think through the novel. I believe this is the clear distinction between a great science fiction novel and a simple story.

7. Nicola Liberati: Thank you very much. I think this leads naturally to the third question, which is also the most crucial one from my perspective regarding developing critical thinking in the next generation through education. This question is essential and

complicated at the same time since it binds together different elements. It is not just about the "role" of science fiction in general in society, but it is also related to the strategies needed to make science fiction embedded in academia. Thus, it touches many different fields like research, marketing, attractiveness for companies and students, and so on. Since you work in one of the research centers in China directly related to science fiction – the Fishing Fortress Science Fiction Academy of Chongqing run by Zhang Fan – how can we embed science fiction within the University as a field of research?

Francesco Verso: The resistance to integrating science fiction in different areas we find is dictated by ignorance. There are places where there is no opposition or resistance to introducing science fiction or at least where there is a kind of neutrality to it.

For example, in Shenzhen China, Prof. Wu Yan created a center for developing human creativity - called "Science and Human Imagination Research Center" at the *Southern University of Science and Technology* of Shenzhen[3] - binding together sciences, science fiction, and students. In Chongqing (China), the *Fishing Fortress SF Academy* already has departments with different applications relating to various fields in industry and academia like cinema.

Out of China, in Olso (Norway), the center *CoFutures,* where I am a board member, studies future scenarios. Thanks to a grant, a group of researchers can publish and research the relations connecting future, technology, and sciences. For example, related to this project, I just edited a volume titled *Meteotopia: Futures of Climate (In)Justice* (Chattopadhyay, Rüsche, and Verso 2022). In this volume, we invited science fiction authors from the global south, and we asked them what the consequences of climate change on their countries are. Moreover, we also asked to think of future scenarios because the cases of a hurricane/typhoon hitting Tokyo and Rio de Janeiro are not the same since Tokyo is famous for its resiliency.

The problem is tackled differently according to the background and location where the event occurs. Thus, we need to consider different perspectives and let other people from diverse backgrounds participate in our discussions.

The project "Co-futures" considers this point by including different backgrounds in the narrative to produce an idea of the future, which is closer to what the future might be. For example, when we think of space colonization, the European idea of what it might be is not enough because other countries like China have very ambitious space programs that shape the future we will experience. Thus, we need to consider this narrative if we want to think of what kind of future we might have in the following years.

Taking into account different backgrounds and cultures is essential also because the idea of the future, in general, is related to people's cultural backgrounds. In some cultures, the past is behind the people, and the future is in front of them, such as in some populations in the Andes. From their perspective, they bring their children on their shoulders, so their future is on their back, while they look ahead to see what they did and to look at their past. Thus, by introducing different perspectives, we can also think differently about what "future" means for us in general.

I think science fiction has a sort of "surprise" in itself, which is used to look at the future and include different scenarios. It should provide a perspective on what our world will look like without excluding possibilities.

Arizona State University also has a program related to speculation about the future where science fiction authors serve as consultants. These authors participate in the discussion in the program and provide contributions that become books such as *Future Tense Fiction – Stories of Tomorrow* (Berg et al. 2019).

Thus, I think science fiction is necessary, and many people understand its value. I believe that those who see that we ended the century of "speed" and entered the century of "acceleration" know that science

---

[3] See the website https://rw.sustech.edu.cn/a/En/E&R/Scientific_research_work/2019/0513/293.html

fiction is the tool to analyze the application of technologies in the future. The people not understanding this passage and the change which comes with it will continue to do things in the old way, and the acceleration will dissolve them. Acceleration can be hard to understand because it grows exponentially, but it is accurate. It is what I can see in China, such as the disorientation and dissociation generated by new technologies introduced in society at an exponential level.

The phenomenon of science fiction changed over time. Seven years ago, it did not have the credential and background to impose itself. Now, it has the chance to say something about our society also in research and development because it is taken into account seriously by society and companies. I have been asked to talk about specific topics, and this opportunity was not possible before. The idea of having a science fiction author is now perceived as valuable, while seven years ago was impossible.

As Ballard used to say, science fiction is the engine for transformation, and it is the narration of the present, so we need it to tackle the societal challenges we will face.[4]

For example, science fiction explores new territories and asks questions like: What are the latest technologies about? How are the new algorithms going to work? Who is going to produce the new technologies?

We are explorers, and so it is natural that we are few.

References

Berg, Kirsten; Torie; Bosch, Joey; Eschrich, Ed; Finn, Andrés; Martinez, and Juliet Ulman, eds. 2019. *Future Tense Fiction: Stories of Tomorrow*. The Unnamed Press. https://www.unnamedpress.com/books/book?title=Future+Tense+Fiction%3A+Stories+of+Tomorrow+.

Bonfiglioli, Cristina Pontes. 2021. "Methodological Challenges for a New Philosophy of Technology Interview with Nicola Liberati." *Prometeica*. https://doi.org/10.34024/PROMETEICA.2021.23.12411.

Chattopadhyay, Bodhisattva, Ana Rüsche, and Francesco Verso, eds. 2022. *Meteotopia [Forthcoming]*. Future Fiction.

Liberati, Nicola. 2022. "Digital Intimacy in China and Japan: A Phenomenological and Postphenomenological Perspective on Love Relationships at the Time of Digital Technologies in China and Japan." *Human Studies* Forthcoming: 1–15. https://doi.org/https://doi.org/10.1007/s10746-022-09631-9.

Mykhailov, Dmytro, and Nicola Liberati. 2022. "A Study of Technological Intentionality in C++ and Generative Adversarial Model: Phenomenological and Postphenomenological Perspectives." *Foundations of Science*, March, 1–17. https://doi.org/10.1007/s10699-022-09833-5.

Verso, Francesco, and Roberto Paura, eds. 2018. *Anthropocene*. Future Fiction.

---

[4] In order to access the citation by Ballard see the following webpage https://www.theguardian.com/books/2020/sep/11/jg-ballard-science-fiction-celebrates-possibilities-life-1970

BAOSHU. MAIDENS' NAME IS MONSTER. GUANGZHOU, CHINA: HUACHENG PUBLISHING HOUSE, 2020. ISBN: 9787536090361

BAOSHU. O NOME DA MULHER É MONSTRO. GUANGZHOU, CHINA: HUACHENG PUBLISHING HOUSE, 2020. ISBN: 9787536090361

BAOSHU. EL NOMBRE DE LAS DONCELLAS ES MONSTRUO. GUANGZHOU, CHINA: HUACHENG PUBLISHING HOUSE, 2020. ISBN: 9787536090361

Wang Jue[1]
(Department of Philosophy, Xi'an Jiaotong University, China)
wangjue0330@126.com

Enthusiasts for AI have held out the prospect that robots will live with people, relate to people, respond to people, and even substitute people when people fail. Robots are more trustable, predictable, and less demanding than human companions. In this way, AI technology seems to be the salvation for the problems of human intimacy in our digital age. Media and commercial culture always encourage people to put hope in new machines, although the meaning of robot companion is far from clear. As Sherry Turkle points out in *Alone Together*, the keenness for robot companion is a subtle, ambiguous, and very complex mix: "social robots serve as both symptom and dream: as a symptom, they promise a way to sidestep intimacy conflicts; as a dream, they express a wish for relationships with limits, a way to be both together and alone." A crucial question arises: When people are keen to substitute robots for humans, what are they going for?

Baoshu, a well-known Chinese science fiction writer, attempts to answer this question by telling a sad and touching story in his novelette "Niuniu," which is included in his personal anthology *Maidens' Name is Monster*. In Chinese, "Niuniu" means an innocent and lovely little girl and is a widespread name or nickname for such a girl. "Niuniu" in this novel refers to the male protagonist (Dong Fang)'s two-years-old daughter, who died in a tragic accident for which his wife (Shen Lan) should take responsibility. The parents were so desperate that they ordered a robot simulacrum, a replica of Niuniu based on all the photos, videos, and other data of Niuniu, made by a cutting-edge technology company. Nevertheless, the artificial replica had a limit, i.e., she could never grow up. Instead, the machine could only imitate all that had happened, the life cycle of Niuniu from one to two years old, and then go back to where she had started and repeat the cycle again and again.

At first, the couple was immersed in bliss for having little Niuniu "revived". As time went by, however, their opinions diverged significantly concerning the authenticity of this robotic Niuniu. The husband gradually concluded that this artificial person is merely a machine in essence, or to put in his words, "a

---

3D video of Niuniu's life". Therefore, it is deceiving. In contrast, the wife felt everything was real[2]: This new Niuniu was as much real as the love and care she felt for her baby all the time.

Who is right on this issue, Dong Fang or Shen Lan? Maybe both of them are right, considered from different perspectives. Dong Fang is correct in terms of the robot's objective characters. However, Shen Lan's reaction is more significant and worth reflecting on deeply. While recognizing that algorithms control all speech and actions of this new Niuniu, she insisted that this robotic Niuniu was real because her emotional attachment was so real, to the extent that this Niuniu served as a "time machine." As Shen Lan put it beautifully, "She takes us back to those times with Niuniu. Every word, every smile, all just like Niuniu's. We never let go of Niuniu. She's been with us all along."

Her reactions are in accord with the latest scientific findings. Just as babies can immediately understand the meaning of parents' facial expressions without reasoning, humans are very easily apt to anthropomorphize human-like robots, feel empathy for them, and thus develop emotional bonds with them. Furthermore, this empathetic connection between robots and humans usually develops *automatically* and *subconsciously*: people subconsciously treat computers as having personalities and apply social rules and expectations to computers.

Does this mean the blurring border between reality and virtuality no longer makes sense? Definitely no. As far as I can see, this story contributes great insight to this question by pointing out that temporality dramatically distinguishes real humans from the robotic imitation of humans. Humans exist essentially temporally, whereas robots cannot "grow" in time. This ontological-existential gap between humans and robots calls into question the idea of substituting companion robots for real humans. Confusing robotic companions and human companions may lead to the moral hazard of being dehumanized by machines. The robot's face certainly announces an ethical and emotional appeal to us, just as how Shen Lan felt about the robotic Niuniu, but it has no meaning when we feel it for a machine. When we begin to talk to robots humanly, when we respond to these robots as "persons," we might dehumanize our existence by being reduced and confined to "as if," oblivious of what is unique about being human—what is most important, the capability of existing temporally in a genuine way. This perspective is how Dong Fang felt "trapped in a long-gone past, unable to escape as if they've fallen into the warped space-time of a black hole" when his wife was so addicted to caring for the robotic Niuniu that she did not want a second pregnancy.

What could we learn from the dilemma confronting this unfortunate couple? In my view, three points deserve people's serious reflection.

First, we should be cautious of the two-sided nature of technology. Technology is a double-edged sword. In some cases, its promise of empowering people to confront the world's challenges turns out to be the most profound form of dis-empowerment. By pointing out the attachment with the robot, Shen Lan creates an illusion of manipulating time, travelling through time back to five years ago, at the price of being trapped in the past and thereby deprived of the authentic meaning of life.

The second point raises ethical issues concerning deception. Companion robots make us believe that they deserve our empathetic feeling of being-with, but they are merely machines that have no genuine response to us. This view is why Dong Fang thought that the robotic Niuniu was deceiving. One objection to this deception claim is that a machine cannot deceive since it is not conscious, not to mention the intention of deceiving. However, this objection misses the crucial point of human-robot interactions. What is at stake in the human-robot interaction is not the state of mind or any properties considered in isolation, but the consequences of the relationship, especially the self-being of the moral agent constituted through this ongoing relationship. Strictly speaking, people are not deceived by the robot but *by themselves*, and this self-deception, in Sartre's sense, is the morally problematic point.

---

[2] The term "real" it is used as opposed to a replica deceiving the user.

Sartre construes self-deception as a kind of false faith in the mode of his own existence, resulting in disowning his/her innate freedom to act authentically. If people falsely believe in the importance of getting recognized and responded to by companion robots, they will eventually suffer social and psychological impoverishment on the long run.

Third, this story indicates what is at stake in discussions about social robots. It is not the objects in themselves but their relationship with the user. The first and foremost central questions are how we feel about being with robotic companions and how we are changed as technology offers us substitutes for human companions. Surely there are no standard answers for these questions but to some extent dependent on particular users in particular situations.

This aspect might explain why the controversy surrounding robotic companions and the artificial person seems unsolvable in the real world.

At this point, we might be able to propose an answer to the question raised at the beginning of this review. When people substitute robotic companions for human companions, what they are going for is nothing but illusions, or even worse, endless nightmares, as this family is caught up in. Baoshu's "Niuniu" productively and strikingly illuminates the vicious circle caused by the misuse of AI. On the one hand, because human beings are so *vulnerable*, we desperately appeal to technology as the last resort for dealing with the hardship and distress in life; on the other hand, human beings become even more *vulnerable* when we uncritically accept the domination of AI in our life. This unexpected ending stimulates readers to think deeply about the nature of robot simulacra and robotic companions: Whether it is symptom or promise, cure or curse? Other science fiction stories collected in this anthology also provoke readers to think further about how technology will reshape our lives and whether it offers us the lives we desire to lead. It is time to begin thinking about these questions together if we do not want to leave our future to be dominated by machines.

1) Los artículos y reseñas remitidos deberán ser inéditos (esto incluye publicaciones digitales como blogs, actas online, etc.) y no podrán ser postulados de forma simultánea para su publicación en otros periódicos. Al someter a la plataforma de la Prometeica, el autor afirma su anuencia a esa exigencia.

2) Los artículos no deberán exceder los 40.000 caracteres. Las reseñas no deberán exceder los 10.000 caracteres.

3) Todos los artículos deberán estar acompañados de un resumen y un abstract equivalente en inglés, cada uno de no más de 1.500 caracteres, incluyendo tres palabras claves.

4) Los idiomas aceptados para los artículos serán:
   a) español,
   b) portugués,
   c) inglés.

5) Los artículos deben ser enviados a la revista a través de lo que el acceso del usuario a la plataforma OJS. Las presentaciones deben estar en dos archivos con formatos doc, docx o.odt. La primera debe ser el artículo y el resumen, y sin los datos del autor. El segundo debe contener los datos del autor: un breve curriculum, afiliación académica e información de contacto. Los metadatos de los archivos deben ser borrados por el autor para garantizar una evaluación ciega, es decir, el autor no ve el nombre del parecerista y el parecerista no ve el nombre del autor.

6) Una vez enviado el artículo/reseña el autor recibirá un e-mail de Prometeica acusando recibo. Desde la recepción de ese mensaje el comité editorial tendrá un máximo de 4 meses para evaluar si el artículo/reseña será publicado/a en la revista. El tiempo máximo previsto entre el envío del artículo y su publicación, en caso de aceptación sin reformulaciones, es de 6 meses.

7) En cuanto al sistema de referencias Prometeica adopta las Normas APA – 2016 - 6ª edición, de la American Psychological Association.

8) Para las notas aclaratorias se empleará la referencia al pie. Preferentemente se sugiere no abusar de este recurso.

9) En el caso de artículos que incluyan imágenes, deben enviarse en un archivo separado. Las imágenes deben tener una resolución de 300 dpi, en formato *.jpg o *.tiff. El copyright de la imagen ya debe ser concedido, o el autor debe usar imágenes sin derechos de autor.

10) En cuanto a la evaluación de los artículos, los mismos serán remitidos al miembro del consejo editorial responsable del área del trabajo en cuestión. Los artículos serán enviados a dos especialistas externos y evaluados en el sistema de revisión doble ciega en que el autor no ve el nombre del evaluador y el evaluador no ve el nombre del autor. En el caso de haber desacuerdo entre ellos, un tercer árbitro podrá ser consultado, por decisión del consejo editorial.

11) Los trabajos pueden tener tres resultados posibles que constan en el formulario de evaluación que completará junto a otras observaciones el evaluador:
   a) recomendado para su publicación sin alteraciones,
   b) recomendado para su publicación con modificaciones,
   c) no recomendado para su publicación.

12) En el caso 11 (b), la publicación del mismo quedará sujeta a que el autor esté dispuesto a realizar las modificaciones y las remita para su nueva evaluación.

13) Todos los trabajos aprobados serán publicados en un idioma aceptado por la revista. Si el original no está escrito en uno de estos idiomas y demandas de traducción, puede resultar en un retraso de la publicación.

14) El contenido de los originales publicados es responsabilidad exclusiva de sus autores.

15) Los artículos presentados a la revista deben ser escritos utilizando el **Modelo_Autor_Estilos.dotx** para seguir los estilos y formatos descritos en **Formato_Prometeica_2019.docx**

16) Prometeica adopta una política antiplágio sometiendo todos los artículos enviados a un primer análisis por el software Turnitin.