

**ARGUMENTOS
DE
RAZÓN TÉCNICA**

ARGUMENTOS DE RAZÓN TÉCNICA

Número 26, 2023

Este número ha contado con el apoyo de las siguientes instituciones y empresas:
Colegio Oficial de Ingenieros Industriales de Andalucía Occidental
Departamento de Metafísica, Corrientes Actuales de la Filosofía, Ética y Filosofía Política
(Universidad de Sevilla)
Asociación para la Investigación y la Cooperación Industrial en Andalucía (AICIA)
ISOTROL. Ingeniería, Software y Control
SADIEL. Tecnologías de la Información e Ingeniería



SEVILLA, 2023

Redacción, administración, secretaría, distribución e intercambio:
Departamento de Metafísica, Corrientes Actuales de la Filosofía, Ética y
Filosofía Política (Facultad de Filosofía)
C/Camilo José Cela, s/n 41018 SEVILLA (España)
Tlfnos. 95 455 77 59 – 95 455 77 77 / Fax 95 455 97 25
Email: art@us.es

Gestión de la edición electrónica:
Secretariado de Publicaciones de la Universidad de Sevilla
C/Porvenir nº 27 – Edificio Corominas – 41013 SEVILLA (España)
Tlfno. 95 448 74 46 / Fax 95 448 74 43
Correo electrónico: secpub2@us.es, secpub4@us.es

Edición electrónica:
<http://editorial.us.es/es/argumentos-de-razon-tecnica>

Diseño de Cubierta:
Sara Mariscal

© José Antonio Marín Casanova (Ed.), 2023



Edición bajo licencia de *Creative Commons*

BY Reconocimiento (Attribution): El material creado por un artista (o autor) puede ser distribuido, copiado y exhibido por terceras personas si se muestra en los créditos.

NC No Comercial (Non-Commercial): El material original y los trabajos derivados pueden ser distribuidos, copiados y exhibidos mientras su uso no sea comercial.

ND Sin Obra Derivada (No Derivate Works): El material creado por un artista (o autor) puede ser distribuido, copiado y exhibido pero no se puede utilizar para crear un trabajo derivado del original.

Argumentos de Razón Técnica se encuentra evaluada o indexada en PUBLONS (ESCI), ERIH Plus, CAPES, CARHUS Plus+ 2014, CIRC 2012, MIAR live, The Philosopher's Index-EBSCO, LATINDEX (Directorio & Catálogo), Conicet (nivel 1), DICE, RESH-CINDOC-CSIC, ISOC, Norwegian Register of Scientific Journals, REBIUN, SUNCAT, DIALNET Plus, Dulcinea, SHERPA/RoMEO, REDIB, Electr@, Livre y en otros índices de calidad o catálogos internacionales de revistas académicas.

ISSN: 1139-3327

Depósito Legal: SE-893-1998

Producción: Gráficas Ulzama

Impreso en España. Printed in Spain

Director/Editor:

José Antonio Marín Casanova (Universidad de Sevilla, España)

Director Adjunto/Assistant Editor:

José Barrientos Rastrojo (Universidad de Sevilla, España)

Director Honorario/Honorary director:

Joaquín Luque (Universidad de Sevilla, España)

Consejo de Redacción/Editorial Staff:

Ramón Queralto Moreno (†) (Primer Director)

Manuel Pavón Rodríguez (†) (Primer Secretario)

Sara Mariscal Vega (Secretario, Universidad de Cádiz, España)

Pascual Martínez Freire (Universidad de Málaga, España)

Antonio Diéguez Lucena (Universidad de Málaga, España)

Julio Gallego (Universidad de Huelva, España)

César Moreno Márquez (Universidad de Sevilla, España)

Paula Cristina Pereira (Universidad de Oporto, Portugal)

Consejo Asesor/Board of Consulting Editors:

Prof. Dr. M. Losada (Premio Príncipe de Asturias de Investigación Científica y Técnica)

Prof. Dr. E. Agazzi (Presidente de la Academia Internacional de Filosofía de la Ciencia, Bruselas, Bélgica)

Prof. Dr. J. Aracil (Catedrático de Automática de la Universidad de Sevilla, España)

Prof. Dr. J. Echeverría (Instituto de Filosofía, C.S.I.C. – Fundación Ikerbasque, España)

Ing. M. A. Luque (Director General del Instituto Andaluz de Tecnología)

Ing. J. Iglesias (Presidente del Comité Europeo sobre Gestión de la Innovación)

Prof. Dr. C. Mitcham (Colorado Mines School, U.S.A.)

Dr. J. A. Pérez Mercader (Los Álamos National Laboratory, U.S.A., Director del Centro Nacional de Exobiología, Madrid, España)

Prof. Dr. B. Kanitscheider (Miembro de la Academia Internacional de Filosofía de la Ciencia, Bruselas, Bélgica)

Prof. Dr. A. Pérez Luño (Catedrático de Filosofía del Derecho de la Universidad de Sevilla, España)

Prof. Dr. R. Infante Macías (Catedrático de Investigación Operativa de la Universidad de Sevilla, España Director del Centro de Informática Científica de Andalucía, España)

Prof. Dr. J. Rodríguez Sacristán (Presidente de la Real Academia de Medicina, Sevilla, España)

Prof. Dr. J. Serna Arango (Universidad Tecnológica de Pereira, Colombia)

SUMARIO / CONTENTS

© = © Todas las contribuciones de este número están incluidas bajo la licencia CC, no derivación, no uso comercial y cita obligada/*All contributions are subjected to Creative Commons License, no derivation, no commercial use and mandatory quotation*

ESTUDIOS / ARTICLES

Artur Dydrov/ Anastasia Gulevtataya, The conflicting discourse of modern media microformats / <i>El discurso conflictivo de los microformatos mediáticos modernos</i>	13
Roberto Pizarro Contreras, El dilema de la alteridad (o alternancia) tecnológica / <i>The dilemma of technological alterity (or alternation)</i>	35
Tobias Endres, On the reality and value of technology / <i>Sobre la realidad y el valor de la tecnología</i>	51
Regina Penner, The evolution of media through the theoretical optics by M. McLuhan / <i>La evolución de los medios a través de la óptica teórica de M. McLuhan</i>	85
José Luis Delgado Rojo, El saber en la época de su reproductibilidad técnica. Sobre la relación entre la inteligencia artificial y la humana / <i>The knowledge in the age of mechanical reproduction. On the relationship between artificial and human intelligence</i>	107
Erick Rubio - Leandro Giri - Andrés Ilcic, Desafíos epistemológicos en la era de las redes neuronales artificiales: abordando sistemas complejos desde una perspectiva computacional / <i>Epistemological challenges in the era of artificial neural networks: addressing complex systems from a computational perspective</i>	145
Elvira Valeeva – Ekaterina Milyaeva – Regina Penner – Elena Sosnovskih, Digital ascetism for everyday life / <i>Ascetismo digital para la vida cotidiana</i>	179

Gonzalo Manzullo, **Oswald Spengler: visiones sobre la tecnificación** / *Oswald Spengler: views on technification* 207

Maximiliano Bozzoli - Mariano J. Domínguez, **Aspectos epistemológicos sobre la validación de modelos computacionales en cosmología** / *Epistemological aspects on the validation of computational models in cosmology* 249

Ricardo Andrade, **Problemas filosóficos de la inteligencia artificial general: ontología, conflictos ético-políticos y astrobiología** / *Philosophical problems of artificial general intelligence: ontology, ethical-political conflicts and astrobiology* 275

SECCIÓN BIBLIOGRÁFICA / REVIEWS

GONZÁLEZ REYES, L. - ALMAZÁN A., *Decrecimiento: del qué al cómo. Propuestas para el Estado español* (ANDONI ALONSO) 303

ESTUDIOS /ARTICLES

**THE CONFLICTING DISCOURSE OF MODERN MEDIA
MICROFORMATS**

***EL DISCURSO CONFLICTIVO DE LOS MICROFORMATOS
MEDIÁTICOS MODERNOS***

ARTUR DYDROV

South Ural State University, Russia
emdydrovaa@susu.ru

ANASTASIA GULETVATAYA

South Ural State University, Russia
gulevataiaan@susu.ru

RECIBIDO: 03/03/2023

ACEPTADO: 25/04/2023

Abstract: This article conceptualizes the Internet microformats (memes, short stories, gifs, etc.) as reagents of events and, in particular, conflict situations. The authors argue that microforms also perform equally important functions of event interpretation and faking (actually creating an event). As an empirical base, the case with the popular Internet blogger Anatoly Shariy was analyzed. It is known that the blogger was arrested due to inciting international conflicts. Analytics of the time of publication of Internet content (primarily memes) allows us to formulate some theoretical provisions about microformats. Significant characteristics of microformats are a high speed of response to events and processes, as well as a short information life cycle. Summarizing the results of the analysis we can identify three main functions of microformats: microformats as reagents of events; as interpreters of events; as creators of events. The first function is destructive with the speed of microformat occurrence. The second function has a hermeneutic specialization. detection of microformats of external (secondary) meanings that can hide or transform factology. Finally, the third function is creative. Microformats act as carriers of creative intentions that arise from the creation of imaginary events. These characteristics significantly distinguish microformats from monumental cultural forms that have referential foundations, a historical and cultural background. With the semantic diffraction,

conceptualized in the philosophy of structuralism and poststructuralism, memes and other microformats should be not only pedagogical tools to create interest in learning, but also a special, independent object for analysis and study in an educational setting.

Keywords: Internet; microformats; memes; reagent; interpretation; faking; education

Resumen: Este artículo conceptualiza los microformatos de Internet (memes, cuentos, gifs, etc.) como reactivos de eventos y, en particular, de situaciones de conflicto. Los autores argumentan que las microformas también realizan funciones igualmente importantes de interpretación y falsificación de eventos (realmente creando un evento). Como base empírica se analizó el caso del popular bloguero de Internet Anatoly Shariy. Se sabe que el bloguero fue arrestado por incitar a conflictos internacionales. La analítica del tiempo de publicación de contenidos en Internet (principalmente memes) nos permite formular algunas previsiones teóricas sobre los microformatos. Las características significativas de los microformatos son una alta velocidad de respuesta a eventos y procesos, así como un breve ciclo de vida de la información. Resumiendo los resultados del análisis podemos identificar tres funciones principales de los microformatos: los microformatos como reactivos de eventos; como intérpretes de los acontecimientos; como creadores de eventos. La primera función de destrucción con la velocidad de ocurrencia del microformato. La segunda función tiene una especialización hermenéutica. detección de microformatos de significados externos (secundarios) que pueden ocultar o transformar la factología. Finalmente, la tercera función es creativa. Los microformatos actúan como portadores de intenciones creativas que surgen de la creación de eventos imaginarios. Estas características distinguen significativamente a los microformatos de las formas culturales monumentales que tienen un fundamento referencial, un trasfondo histórico y cultural. Con la difracción semántica, conceptualizada en la filosofía del estructuralismo y postestructuralismo, los memes y otros microformatos deberían ser no solo herramientas pedagógicas para crear interés en el aprendizaje, sino también un objeto especial e independiente para el análisis y el estudio en un entorno educativo.

Palabras clave: Internet; microformatos; memes; reactivo; interpretación; falsificación; educación

Introduction

The ontological and the related epistemological problematics of the one and the multiple is obviously not exclusively connected with the postmodernist mainstream, and had been tackled on a certain level long before the concepts of nomadology, intertextuality and diffraction emerged. In particular, the Italian philosopher P. Virno in «The Grammar of the Multitude» refers to the contrasting positions of Spinoza and Hobbes on socio-political problems and the corresponding demarcation of the concepts of people and multitude (Virno, 2013, p. 10-11). «People» is known to be a concept of national, territorial and ideological structure, providing a synthesizing unity: «Russian», «people of their country», «people and party are one», etc.

P. Virno's concept «multitudo», on the contrary, does not imply such unity and at the same time is not an intermediate form, it is literally and tautologically the existence of multiple as multiple. Demarcations in the cartography of multiplicity are conditioned by a pool of factors: confessional, political, economic, ideological, cultural, etc., within each of which, in turn, there are innumerable and weakly registerable segmentations. Multiplicity at the categorical level fixes potentially infinite differences of any order in all conventionally allocated spheres of social life. Of course, this state of affairs is also characteristic of culture with its inherent multiplicity of interpretations, intertextuality, and permanent growth of the world of ideas and things. Internet technologies and, in fact, the network itself have catalyzed processes of information explosion and diffraction (spontaneous, chaotic dissemination) that do not lend themselves to traditional growth («exponential», «logistical»). In the second half of the last century informational expansion was grasped by individual authors, Heidegger in the 50s («Gelassenheit»), Toffler and McLuhan in the 60s, etc. However, they were focused on

«historical» media channels that lost monopoly. The web had the advantage of chaotic, virtually uncontrolled distribution, the limits of which were by no means constituted (in the twenty-first century this is obvious) by the globe. Recalling Heidegger's assessment («today cognition of everything is available so quickly and so cheaply...» (Heidegger, 1991, pp. 102-111), it is possible to describe the cost of modern network information as dumping, or «super-cheap». It is no coincidence that the massovization and domestication of the Internet has aroused the enthusiasm of many optimistic intellectuals: some believed that the network is the result of radical free-thinking and secession from the world of «weary giants of flesh and steel» (Barlow, 2004, p. 349), others saw in the network the conditions for a technological leap from capitalism to cybernetic communism (Barbrook, 2015). The network is a special topos, accommodating an actual and potential statistically uncovered multitude of data, connections, and relationships. The avant-garde of the information revolution, the marker of our time, is not only the permanent growth of information (not fixed in traditional terms), but also the intensification of the elementary binary connection of «request – response», compressed to a few seconds.

Exceptional emphasis on this temporally insignificant connection can lead to an erroneous assessment of the technology impact on human cognitive functions (easy access to information causes dysfunction) and to a concomitant erroneous assessment of technology itself («degrading», «primitivizing» innovations). Having made the most elementary request, the user receives on the output thousands (often millions) content variations. An undefined part of the resulting information is duplicated, but this does not cancel the fact of encountering multiple. Of course, user reactions are well known (it is not our task to search for psychological determinants of these reactions), which influence the following behavioral strategies: to turn to the first page of a browser and copy

the information from the title site (a problem of pedagogy, psychology, teaching and research) or to analyze the query results, sifting out irrelevant information. The first strategy does not require any significant cognitive effort, the second, on the contrary, requires a pool of skills: the user, speaking hermeneutically, must have «precognition» to perform at least elementary navigation and mapping of electronic resources. Returning to the Heideggerian evaluation, marked by the combination «fast and cheap», we note that it does not imply «easy». The second strategy (in fact, the only one claiming to be a strategy without quotation marks) really confronts the individual with multiplicity. The result of the first behavioral response omit multiple.

Theoretical Framework and Methodology

Theoretical framework and methodology of microformats research are inextricably linked to the discourses of structuralism and poststructuralism, which worked directly with language and various modes of expression (including performative). Internet microformats are characterized by semantic diffraction (R. Barthes) and an orientation to instant effects (affectation, emotionality), miniaturization (perception requires comparatively little effort, grasping at once). Conceptually, the trend was expressed by futurologists (Bestuzhev-Lada, 1990): in particular, M. Penn and K. Zalesne argued that the future is largely determined by «micro-trends» (Penn, Zalesne, 2009). The fact is that once marginalized and sporadically emerging micro-formats over time have formed the mainstream, that is, they have actually turned into a cultural trend. The miniaturization and «spontaneity» of their emergence and the short life of many microformats is inextricably linked to their potential: orientation on instant impression, low barrier to entry,

often functioning in the context of fake industry (M. Foucault, J. Baudrillard), functioning as an interpretative and openly simulative propaganda micro-mechanism. Of course, this characteristic is not attributive and appears in specific microforms interpreting political, international relations, etc.

Content analysis, search query research, and semiotics methodology are adequate as tools for microformats in the context of the Internet environment. Research problems arise already at the stage of assembling and codifying material that is scattered on the web and occurs sporadically on a variety of platforms and sites. Working with the web is difficult without big data analytics, and without specific tools to collect and analyze user search queries. The most common and simplest of the tools are, for example, Google Trends and Yandex WordStat. The latter provides chronologically strictly limited data (from 2019-2020). In turn, Google Trends captures «full» statistics and ignores small values of user queries. Working with visual and textual material requires contextual knowledge, since microformat generally does not represent the context of the event.

Results

Strategies (psychological, cognitive, social, cultural) of thinking and behavior under conditions of permanent and hyperactive information growth have yet to be studied. Using P. Virno's conceptual framework, it is obvious that knowledge of the «grammar» of multiplicity is on the agenda today among educational trends. In research discourse, «grammar» can carry a range of names, such as the popular «Internet skills» (Van Deursen, 2011; Hargittai, 2019; Miliou, 2021). In essence, research tactics are often reduced to the codification of a certain number of skills and is accompanied by a

declaration of the importance and even necessity of putting this spectrum into practice in the education system. To illustrate the point, let us quote O. Miliou and C. Angeli: «... the use of these technologies [Internet technologies – hereinafter the author's note] is limited and may not cover sufficiently the knowledge and skills necessary for effective performance of Internet activities in higher education. In addition, [students'] exposure to digital devices and tools varies qualitatively. Consequently, higher education institutions need to close the skills gap» (Miliou, 2021, 1359).

The latter point, despite its frequent reproduction in research texts, is not self-evident and needs to be argued for. An ontology of multiplicity is inextricably linked to a pluralism of forms and contents, formats and ideas. Again, in the discourse of network libertarians (Barlow) and even «cyber-communists» (Barbrook), cyberspace was positioned as a frontier domain, free from the hierarchies and structures of the «flesh-and-steel» world. In general terms, this meant that the network could encompass the entire possible spectrum of propositions, without any of the elements of the spectrum being forcibly annihilated. Today, at the end of the first quarter of the twenty-first century, we know that cyberspace is not an anarcho-libertarian topos, but is in fact a territory segmented into zones of discursive influence. Like a cybernetic «black box», cyberspace receives an event, indefinitely processes it, and produces an interpretive pool in which it is difficult (or even impossible) to see any regularities. Of course, the position of the person acting as a user is also in question. The uncertainty of the position is fundamentally due to the complexity of the relationship and link between the user and the network space: does the user connect to the processes taking place in the «black box» and, therefore, retain relative autonomy, or is he a «black box» element? There should not be a strict disjunction here, at least if we examine the problematics from different angles. In the old conceptual-anthropological optics,

it is as if man remains multidimensional, eccentric, a special being, not reduced to any single dimension in principle (figuratively speaking, he is always «irrelevant»). In the context of the contemporary cultural situation, the network has become an attractor, absorbing the potencies and impulses of the individual. It is not only a matter of psychological «addiction» to the machine: from an objectivist perspective, the user is (by analogy with the Deleuze-Gwattari concept) a machine of opinions. The latter has been elevated to a cult by almost all democratic constitutions, declarations and legal instances. It is important to understand that opinion, an ambiguous phenomenon, for example, in the optics of Platonism, is positioned as one of the highest values in the semantic register of the individual. As a consequence the Web cannot be called a space of dialogue. This does not mean that dialogue is fundamentally impossible in a network: the very architecture of the network, with its endless junctions, intersections and transformations of texts, turns dialogue into a random, «spontaneous» event. The label «polylogue space» would also be questionable; «polyphony» seems appropriate, in this terminology structuralists and semiologists recorded textual multiplicity.

In the context of conflict situations with cross-acting agents (political structures, social institutions, social movements, individuals) the trend of plurality of opinions has a twofold meaning. On the one hand, a wide range of ideas emerges, preventing the transition from totalization of any particular opinion (expansion of the sphere of influence) to totalitarianism (monopoly of ideas and maintenance by punitive structures); on the other, «noise», i.e. the simultaneous functioning of various agents, significantly impedes any attempts to constructive dialogue. One cannot help thinking that silence (even forced silence) remains the only and, at the same time, a dubious way out. We assume that many members of the younger generation pass over conversations about large-scale conflicts not

because they are not «interested» in the topic or because they do not accept political ideas or any state ideologies on a «genetic» level, but because the conditions of permanent socialization today are significantly different from those of the past. Despite the persuasiveness of the conceptual framework of organoprojection (as well as intellectual projection (Gubanova, 2018, p. 254), we should remember the historical uniqueness of net status: it is possible to talk about it in terms of topology. This means that the net accommodates. According to common sense, any topology is inherently bounded and, therefore, the architecture of the net also has limits. The question of limits has been known to arise acutely throughout virtually all of human history and, of course, has rarely been quickly resolved. It is risky to assume that the net has no boundaries at all (cyberlibertarians), but just as unreasonable would be the view that the topos is quite limited. In the latter case, net boundaries are reduced to phenomena outside the net (e.g. silicon, electricity, etc.) that is, to the material substratum. The polar and extreme points of view on the boundaries of the net, of course, concern not only and not so much its “physical” limits, as the behavioral and thinking practices of man.

The ultimate libertarian position in anthropological optics is characterized by the following attitudes (we will formulate them in a primary form that requires further elaboration).

1. Unlike disciplinary units that are strictly limited, integrated into other structures, and have certain legalized limits, the Internet is limitless and, therefore, can accommodate anything. By «everything» is meant not only content-volume, but also semantics. The libertarian point of view crystallized at the turn of the web development in the 90's. The flip side of free expression of any, even the most provocative ideas, is anonymity. It could be said that it was positioned as a kind of protective mechanism that ensured impunity. According to E. O. Trufanova, «anonymity implies complete

protection of the private, since the real person and his private life remain beyond the reach of the Other's gaze. Anonymity, the use of fictitious role images, personas, were characteristic of the early period of Internet development, but since the Internet has been mediating a large number of socially significant activities – trade, bureaucratic procedures, negotiations – the degree of users' anonymization is decreasing, masks are increasingly being replaced by real people» (Trufanova, 2021, p. 20).

The final part of the quoted proposition, in our opinion, is far from being unambiguous: where and how is it possible to find the criteria that allow claiming that some «real people» are replacing the simulacra of the 1990s? The author seems to be right when she speaks about the illusion of anonymity directly influencing the behavioral practices of the network. This illusion is a kind of legacy of the libertarian ideology of the last century. The Internet in libertarian optics is a unique (paradoxical) topos without borders. The only thing with which it can be compared is the universe. Dozens of «theories» of the structure, origin and evolution of the Universe create a poly-discursive noise and turn it into a topos-enigma. With the advent of the World Wide Web, the Internet has «squeezed» the Universe into its direct competitor.

2. Another libertarian proposition is that «everything can be found on the Internet». The guide to the total repository of information, or rather to the world of universal trans-epochal collective knowledge, is, of course, the search query. It is hardly possible to say exactly when the infocommunication network created by military structures became enriched with social mythology and «transformed» into an information universe – most likely, this process of «transformation» took years, although it was largely accelerated by libertarian ideas about the boundlessness of the network.

3. The searching-and-consumeristic orientation, according to

which searchers must «find everything», is supplemented by a primitivized dialectical opposite – the productive orientation. Infoproducts are massively «thrown» into the Internet, generating disputes about quality and consumer properties. The illusion of an absolutely free commodity turnover, not mediated by monetary equivalent, conditions the practice of copying and «uncontrolled» duplication of information. In its turn, the freedom of the market as freedom of production has as its flip side the mass dumping of information that has no social value (flooding) and is harmful in the light of jurisprudence and legislation (advertising of legally forbidden goods and services, propaganda of nationalistic, racial and other similar ideas). The examples outlined are obvious. However, there are immeasurable amounts of microformat content – in particular user generated content – that have an ambiguous impact on the overall state of the Internet environment. A specific feature of such content is its viral or quasi-viral nature, sporadic emergence and uncontrolled spread. Using Plato's conceptual apparatus, we can say that any spectrum of commentary de facto turns into a conflict of multiple opinions and positions, many of which openly claim to be true. At special moments in the development of society and the state, polydiscursive «noise», of course, plays a contradictory role and is almost always both a cause and a consequence of conflicts. Commentaries are specific content, subject to analytics. A fertile environment for the emergence and spread of this content is social networks, which technically support this type of user activity (or destructiveness). Obviously, moderation filters do not always effectively cope with massive commentary activity. As a rule, clusters of comments lack structure because they do not fit the format of a dialogue or conversation. They are capable of creating pseudo-tree-like structures that rapidly disintegrate and are replaced by a new fragile architecture. The example of one of the microformats – user comments – shows that traditional research methods (above all,

observation) can produce unjustifiably limited and even distorted results in the output. The methodological problem stimulated the search for alternative solutions and the entry into the research vanguard of special methods and techniques.

With the development and domestication of artificial intelligence systems, microformats have advanced to a new stage. “Domestication” literally means the social absorption of the technology. The technology becomes easily accessible to any user with basic personal computer or mobile phone skills. Artificial intelligence has made it possible to algorithmize certain processes not previously available to algorithmization. This applies in particular to contentmaking and, among others, memes.

Some researchers correctly and reasonably argue that artificial intelligence, at least at the current stage, is incapable of “understanding” memes. The growing capabilities of artificial intelligence have catalyzed the view that machines will exceed humans and acquire an intelligence equal to or greater than that of humans. This circumstance is often viewed in alarmist optics as exceptionally negative. According to Ishaani Priyadarshini & Chase Cotton, artificial intelligence has become a “pariah,” a subject that generates a storm of negative emotions and gloomy predictions. The Technological Singularity, that is, the hypothetical development of computing power to the level of autonomy and self-governance of intelligent systems, is considered a serious problem in both academia and business. In reality, according to the assessment of research teams, artificial intelligence is not omnipotent and cannot definitively outperform humans. The scientists focused their research strategy on identifying the limitations of machines and analyzing their “incompetence,” which can draw a line between humans and so-called machine “intelligence”. The authors did this on the basis of Internet meme cases. Internet memes are a mixture of images, videos, underlying messages, ideas, feelings, humor and

experiences. An Internet meme is perceived by humans in a way that is not exactly how a machine “understands” it. Results have shown that “understanding” memes is indeed a difficult task and thus a serious limitation for the functioning of artificial intelligence and the creation of simulation algorithms (Cotton & Priyadarshini, 2021, p. 781).

The Internet meme industry is expanding rapidly, driven by the intensive exploitation of social networks and special platforms. Memes generate humor that travels the web and connects people through popular images and humor. Humor, according to the authors, is the main obstacle to imitation by machine algorithms. Hundreds of thousands of memes are generated every day, some refined from previous patterns. Many patterns are actually being created for the first time. Meme humor expresses not only mainstream trends, but also specific, local topics. These topics are difficult to reproduce with an algorithm. Meme generation data is constantly updated, and such data is difficult to manage. Combinatorial creativity has made it possible to present the same meme template in millions of different variations. Behind the millions of variations are millions of meanings and implications. This is a difficult task for AI to understand because AI learns from patterns and specificity (hence repeatability) of data. These patterns and features are very different when it comes to Internet memes. In the future, the number of memes and patterns for them will only grow, and understanding Internet memes will become a much more difficult problem for AI, as it must learn what is already there and make decisions for those it has never seen. Consequently, total understanding of Internet memes may always remain an intractable problem for AI (Cotton & Priyadarshini, 2021, p. 796). It should be noted that a total understanding of memes and the fixation of their circulation is also not available to the user, even those with research skills. However, if we consider humanity rather than an individual

as the subject of meme production, it demonstrates much greater semantic flexibility than algorithms of artificial intelligence systems.

The average user, who has no special qualifications, can now make content of a practically “industrial” volume. Internet resources that generate memes have emerged. These resources work on relatively simple AI algorithms. The general principle of the resources is as follows: the user selects any available meme template and activates the “generate meme” option. The meme template is constituted by the original image and/or phrase. For example, the famous character of the Lord of the Rings, the warlord of the southern state of Middle-earth Boromir, became a meme hero relatively long ago thanks to the phrase “One does not simply walk into Mordor”. The resource offers a template and the first part of the famous phrase. After activating “create meme” a picture appears with the addition of the second part. The user can download the meme and upload it to any other available resources. An analysis of the capabilities of the AI-powered meme generator suggests that its combinatorial resource is actually not exhaustible. The user is offered 48 templates, each of which can be filled with tens of thousands of variations. Our work with the pattern “One does not simply...” was suspended at the 500th variation. We recorded the absence of lexical repetitions. However, the first words of the second part of the meme were of course repeated. Frequent markers were as follows: “create”, “meme”, “meme war”. The semantics of a number of propositions is unreadable and cannot be qualified as humor. Here are some examples of the results of the generation of memes belonging to the group of unqualified semantic correlation (probably, we can talk about the absence of correlation as well): "One does not simply / Pass on a Second to be a free wall"; "One does not simply / Create a car without a party"; "One does not simply / Stop talking about the Master"; "One does not simply / Stop the Senate of the free state", etc. The general meaning of these

propositions is not clear and probably has nothing to do with humor. However, in the list of generated variations there are propositions that contain explicit humorous correlates: “One does not simply / Stop smoking in Math Class”; “One does not simply / Remove the cat period”.

The intensity of the memes circulation, as well as the possibility of rapid generation of memes using platforms with artificial intelligence, has ambiguous social consequences. On the one hand, meme culture is rapidly evolving, and there is a build-up of information content. The dynamics of meme content has increased. In particular, the growth of dynamics is due to the automation and algorithmization of contentmaking. Artificial intelligence working with databases is trying to generate humor and irony. Researchers analyze various aspects of such generation. In particular, in Beth Singler’s article “The AI Creation Meme: A Case Study of the New Visibility of Religion in Artificial Intelligence Discourse” the author argues that the creation of memes by AI (actually by artificial intelligence) will generate many notable cultural artifacts, arising in a variety of places, to be analyzed and discussed for religiosity, apocalyptic and posthumanist narratives (Singler, 2020).

Just as a specialist philologist is able to distinguish between a poem written by a human being and a generated rhyming product, so AI-generated memes are significantly different from those made by humans. It does not follow from this statement that AI is completely safe and that its “products” are easy to recognize. We agree with Cotton & Priyadarshini's position on the learnability of AI. It is unable to “understand” memes, because the meme industry shows a tremendous increase in information, a permanent change of meaning and variations in representation. However, there are propositions in memes, including those created by AI or literally generated by an “intelligent” system, that are directly borrowed from the cultural mainstream. These include racist, chauvinistic, nationalistic and

anti-religious ideas that insult the honor and dignity of a person belonging to a particular nation, ethnicity or denomination. Researchers pay attention to this problem. In particular, employees of Innopolis University in Russia drew attention to the undefined and dynamic group of so-called “hateful memes” in the article “Hateful Memes Classification using Machine Learning”. In the article, the authors focus on the virulent nature of memes. Users often limit themselves to memes instead of reading any article about an event that contains extended interpretation and extended data. They express fears, supported by physiologists and psychologists, of reduced attention to content. Not surprisingly, Internet content is “miniaturizing,” with popular bloggers often creating 5-15 minute videos or small podcasts and targeting large numbers of views and positive reactions. It would take millions of content analysts to track and qualify malicious memes containing blatant propaganda, calls for violence, nationalism, and other similar ideas (Badour & Brown, 2021). Even this hypothetical circumstance is unlikely to save the situation. The hybrid structure (picture and text) makes it difficult to analyze memes and crystallize malicious content. Already from these theses it is clear that machine learning could replace millions of content analysts, but its application is also difficult. Qualifying a picture requires an advanced image recognition and labeling mechanism, qualifying text requires text recognition, respectively. Added to this is the problem of comparing the two layers of the meme, because separately the picture and the text can be qualified as “permissible”, but in synthesis contain information that is degrading. In terms of elementary linguistics, the analysis of individual semes, lexis and icons does not equal the analysis of the totality of these signs. Nevertheless, researchers have created and described several models for segmenting images and text to identify hate labels.

Probably, the activity of science centers is not the only way to combat miniaturized viral Internet content. But it is unequivocally

the most effective way today because it directly engages artificial intelligence, Big Data and Machine Learning. Private agents create apps to suit the market and the needs of the public. In particular, the market for meme-based content retrieval applications is expanding, as the latter is rapidly becoming obsolete. It seems possible to use these applications in addition to serious technological solutions from institutions and professional communities. We believe that entertainment content as well as special research content (software, mathematical models, AI models, etc.) can be used for socially acceptable purposes - regulating the behavior of social agents, tracking malicious messages and images, revealing the attitudes of social groups. Such research is already being conducted on cases of online games and game chats. However, this work has yet to be put into practice.

Conclusion

Summarizing the results of the analysis we can identify three main qualities of microformats. 1. Microformats as reagents of events; 2. Microformats as interpreters of events; 3. Microformats as creators of events. The first quality is directly related to the speed and intensity of microformats emergence. It fixes a relatively high speed of their appearance and circulation. It should be stipulated that the life cycle of a microform can be very short, although this, of course, is not the general rule. When we refer to a comparative degree, it means that microforms circulate intensively in the space of media unlike «monumental» cultural practices (films, educational lectures, etc.). The latter accumulate social and cultural experience, refer to an array of sources, are subjected to tests of validity, are mediated by a set of complex technical operations and procedures, etc. Microformat is often technically simple and can be implemented at

home. The comparative ease of implementation of microformats is responsible for the hard-to-measure, unprecedented frequency of their appearance. There is a feeling that virtually any event (whether with or without the participation of media persons) is instantly represented. On May 5, 2022 (12:28), news about the arrest of Ukrainian politician and blogger Anatoliy Shariy in Spain on charges of treason and incitement of hatred was published. On May 5 on lux.fm there was an article «Memes about the Anatoliy Shariy's arrest that warm the soul» (16:17). At 17:21 there was a rewrite about the arrest of the blogger and the fact that Ukrainians mocked the event with memes.

Acting as interpreters of the event, microformats record the assessment of the event and contain connotative superstructures focused on the refraction of the consumer's view of Internet content. The meme industry has accompanied the blogger's arrest event with numerous connotations. In particular, a meme appeared with a scene from the movie «Ivan Vasilievich Changes His Profession», depicting a phone call from A. Shpak: «Hello, is that the SSU? (Security Service of Ukraine). My neighbor is giving likes to Shariy». The same message is conveyed by memes with a shocked Dwayne Johnson turning to a car passenger («Did you really like Shariy?») and a puzzled Chuck Norris («How can you like Shariy without the IP noticing?»).

Microformats also act as fakemakers, «completing» what happened «after the fact» or literally inventing the event. In particular, the arrest of the blogger was accompanied by a cluster of memes directly pointing to pedophilia and imprisonment. This special cluster of memes is characterized by an abundance of obscenities, scenes of violence, etc., so we cannot refer to specific examples. Fakemaking is one of the most powerful and controversial functions of microformats that determine the general orientation of the Internet mainstream. In essence, it is the flip side of the rapid response to an

event and the intensification of interpretive mechanisms. The skills of decoding the products of the fake industry are on the agenda today. It should not so much be a special knowledge, preceded by a special initiation, but an educational trend, pragmatically oriented to each individual.

The intensity of spread, the high rate of occurrence, and the shortness of the text make microformats a convenient and effective tool for demythologizing and critical thinking. For example, a recent study of Internet memes contained the following argumentative conclusion: memes act as tools for decoding manipulative media techniques. Memes reveal the selectivity of the media in presenting information (part of the information is deliberately hidden), word games (playing with polysemy, abstractness, etc.), the skillful use of the audience's emotions, etc. (Troshchenkova, 2022). But it is also clear that memes can be a tool of visual and verbal manipulation. As an example, let us refer to the study of political content. In particular, an analysis of the British National Party on the Facebook page showed that the production of fabricated images (including photos and memes) is a practice that is also typical for power structures (Klein, 2020, p. 154).

The general strategy of responding to microformats under conditions of meaning diffraction can work effectively within the framework of education. In pedagogy, for example, there is already considerable experience in the use of Internet memes for learning. In the recent past, there has even been research on memes for young children. Researchers argue that meme-culture helps children who leave home and move into institutions to adapt to the changed conditions of life (Westbrook, 2021). Memes are used as stimulating and motivational tools for learning. However, it makes sense to use memes as the main subject for the deciphering of cultural, political, ideological and other codes. The practice of demythologization is well known to linguists, sociologists, and social philosophers. At the moment,

unfortunately, civilized societies have not yet developed the practice of integrating memes (and other microforms) into education as basic objects for analysis.⁷

References

- Badour, J. & Brown, J. (2021). Hateful Memes Classification using Machine Learning. 2021 *IEEE Symposium Series on Computational Intelligence* (SSCI). DOI:10.1109/SSCI50451.2021.9659896
- Barbrook, R. (2015). *The Internet Revolution: From Dot-com Capitalism to Cybernetic Communism*. Ad Marginem Press.
- Barlow, J. (2004). A Declaration of the Independence of Cyberspace. *Information Society*. AST.
- Bestuzhev-Lada, I. V. (1990). «Future phobia Effect» in the Ordinary and Bureaucratic Consciousness. *Sociological Researches*, 11, 72-81.
- Cotton, C. & Priyadarshini, I. (2021). AI Cannot Understand Memes: Experiments with OCR and Facial Emotions. *Computers, Materials & Continua*, 70(1), 780-800. DOI:10.32604/cmc.2022.019284.
- Gubanova, A. S. (2018). Internet technologies as a continuation of the human body. *Bulletin of the Saratov University. New Series. Series: Philosophy. Psychology. Pedagogy*, 18(3), 252-255. DOI: 10.18500/1819-7671-2018-18-3-252-255.
- Hargittai, E. & Micheli, M. (2019). *Internet Skills and Why They Matter. Society and the Internet: How Networks of Information and Communication are Changing Our Lives*. Oxford University Press. DOI: 10.1093/oso/9780198843498.003.0007.
- Heidegger, M. (1991). *Detachment. Conversation on a country road*. High School.
- Klein, O. (2020). Misleading memes: The effects of deceptive

visuals of the british national party. *Partecipazione e Conflitto*, 13(1), 154-179. DOI: 10.1285/i20356609v13i1p154.

Miliou, O. & Angeli, C. (2021). Measuring the Internet Skills of Gen Z Students in Higher Education: Validation of the Internet Skills Scale in University Settings. *Seventh International Conference on Higher Education Advances*, 1359-1368. DOI: 10.4995/HEAd21.2021.13070.

Penn, M. & Zalesne, K. (2009). *Microtrends: The Small Forces Behind Tomorrow's Big Changes*. Twelve.

Singler, B. (2020). The AI Creation Meme: A Case Study of the New Visibility of Religion in Artificial Intelligence Discourse. *Religions*, 11, 253. DOI:10.3390/rel11050253

Troshchenkova, E.V. (2022). Memes as tool for deconstructing media manipulation. *Voprosy Kognitivnoy Lingvistiki*, 2, 74-86.

Trufanova, E. O. (2021). Private and public in the digital space: blurring the boundaries. *Galactica Media: Journal of Media Studies*, 3(1), 14-38. DOI: 10.46539/gmd.v3i1.130.

Van Deursen, A. & Van Dijk, J. (2011). Internet skills and the digital divide. *New Media & Society*, 13(6), 893-911. DOI: 10.1177/1461444810386774.

Virno, P. (2013). *A Grammar of the Multitude. For an Analysis of Contemporary Forms of Life*. Ad Marginem Press.

Westbrook, F. & Hunkin, E. & White, J. (2021). Lost in translation: An experiment with memes for research translation in Australian early childhood education and care (ECEC) contexts: Researchtranslation. *Video Journal of Education and Pedagogy*, 65(3). DOI: 10.1163/23644583-bja10019.

EL DILEMA DE LA ALTERIDAD (O ALTERNANCIA) TECNOLÓGICA

THE DILEMMA OF TECHNOLOGICAL ALTERITY (OR ALTERNATION)

ROBERTO PIZARRO CONTRERAS

University of Science and Technology of China, China

Universidad de Chile, Chile

roberto.pizarro.c@outlook.com

RECIBIDO: 12/03/2023

ACEPTADO: 24/04/2023

Resumen: Este artículo pretende formalizar un dilema moral relativo a la frecuencia, medida y condiciones que han de determinar la decisión de refaccionar o sustituir la base tecnológica sobre la que desplegamos una actividad o conjunto de acciones. Se le denominará *dilema de la alteridad tecnológica*, por cuanto implica un *ethos* que llama a la apertura a otro conjunto posible de tecnologías. Se trata de un dilema, en primer lugar, debido a que el despliegue potencial de una acción confrontaría siempre dos realidades: la acción a través de la base tecnológica actual, o bien, de una nueva que puede ser pensada e implementada. Lo es también, porque su solución responde a condiciones específicas del contexto en que el dilema aparece (una empresa en riesgo de quiebra enfrenta el dilema de forma distinta a una corporación académica que desea mejorar la difusión de su saber entre los ciudadanos). Se intentará fundamentar el dilema recurriendo a: 1) la hipótesis del determinismo tecnológico, 2) la teoría de sesgos cognitivos y 3) el concepto de alteridad (otredad). El sesgo cognitivo, recogido de la psicología, permitirá sostener desde un punto de vista empírico un determinismo tecnológico “suave”, según el cual la tecnología presupone una determinación en la forma de sesgo, pero que es posible contrarrestar elaborando estrategias de evasión, con lo cual la responsabilidad moral del sujeto frente a un acervo o conjunto de tecnologías resultaría en cualquier caso indefectible.

Palabras clave: determinismo tecnológico; sesgo cognitivo; fijación funcional; alteridad (otredad); tecnología.

Abstract: This article aims to formalize a moral dilemma regarding the frequency, measurement, and conditions that must determine the decision to refurbish or

replace the technological base on which we deploy an activity or set of actions. It will be called *the dilemma of technological otherness*, since it involves an *ethos* that calls for openness to another possible set of technologies. It is a dilemma, firstly, because the potential deployment of an action would always confront two realities: the action through the current technological base, or a new one that can be thought out and implemented. It is also a dilemma because its solution responds to specific conditions of the context in which the dilemma appears (a company facing bankruptcy would face the dilemma differently than an academic corporation that wants to improve the dissemination of its knowledge among citizens). The article will attempt to substantiate the dilemma by resorting to: 1) the hypothesis of technological determinism, 2) the theory of cognitive biases, and 3) the concept of alterity (otherness). The cognitive bias, taken from psychology, will allow us to support, from an empirical point of view, a “soft” technological determinism, according to which technology presupposes a determination in the form of a bias, but which can be countered by developing evasion strategies, so that the moral responsibility of the subject facing a set or collection of technologies would be indefectible in any case.

Keywords: technological determinism; cognitive bias; functional fixedness; alterity (otherness); technology

Introducción a los conceptos fundamentales del dilema

Determinismo tecnológico

El *determinismo tecnológico* es una teoría que sostiene que la tecnología es un factor determinante que moldea la sociedad y la cultura. Esta perspectiva argumenta que los avances tecnológicos impulsan el cambio social, al punto de que, por su magnitud, pueden tornar despreciables las decisiones humanas, así como los factores políticos, económicos o culturales. En otras palabras, el determinismo tecnológico sostiene que la tecnología es una fuerza motora –si no la principal– detrás de la evolución social y cultural de la humanidad. (Héder, 2021)

Uno de los ejemplos más comunes de determinismo tecnológico, en la esfera de las ciencias sociales, se aprecia en las denominadas tesis “ferrocarrilistas”. En ellas se argumenta que la construcción de

ferrocarriles en el siglo XIX fue uno de los principales agentes que impulsaron la Revolución Industrial y transformaron la sociedad en el mundo occidental. Según esta forma de determinismo tecnológico, la tecnología del ferrocarril indujo una profunda transformación en los modos de producción, el comercio y el transporte, lo que a su vez provocó cambios sociales significativos como la urbanización y la creación de nuevas clases sociales¹.

Otro ejemplo de determinismo tecnológica es la perspectiva contenida en las tesis “ciborgistas”. Ellas postulan que las tecnologías de la informática y la robótica están cambiando la naturaleza de lo que significa ser humano. Los partidarios de esta teoría sostienen que los avances en la tecnología están llevando a la fusión de lo humano y lo tecnológico, lo que a su vez está transformando nuestra identidad y nuestra relación con el mundo que nos rodea².

La teoría del determinismo tecnológico ha sido trabajada por varios autores. Ya el materialismo dialéctico de Karl Marx y Engels había señalado cómo la tecnología determinaba el modo de producción e incidía directamente en la estructura social (Bimber, 1990). Con todo, uno de los primeros y más influyentes intelectuales que trabajaron sobre esta hipótesis de forma intensiva fue el historiador y filósofo Lewis Mumford, quien argumentó que la tecnología estaba moldeando el desarrollo humano y que los avances tecnológicos estaban impulsando cambios sociales y culturales, aunque no de forma absoluta ni sin concebir ello problemáticamente:

Todo el aparato de la vida se ha vuelto tan complejo y los procesos de producción, distribución y consumo se han vuelto tan especializados y subdivididos, que la persona individual pierde confianza en sus propias

¹ Ver, por ejemplo, el trabajo del historiador económico y empresarial Simon Ville (2004).

² Ver, por ejemplo, el trabajo del investigador en tecnologías digitales Garfield Benjamin (2016).

capacidades: está cada vez más sujeta a órdenes que no cumple, que no comprende; a merced de fuerzas sobre las que no ejerce un control efectivo, moviéndose hacia un destino que no ha elegido. (Mumford, 1970, pp. 358-359).

Otro pionero en el tratamiento explícito del determinismo tecnológico fue el sociólogo y economista estadounidense Thorstein Veblen, para quien la tecnología era un factor decisivo en la evolución económica y social, mostrándose también crítico de su influjo:

El proceso de la máquina impregna la vida moderna y la domina en un sentido mecánico. Su dominio se ve en la aplicación de medidas y ajustes mecánicos precisos y la reducción de todo tipo de cosas, propósitos, actos, necesidades y comodidades de la vida, a unidades estándar. (Veblen, 1904, p. 306).

Otros autores como Marshall McLuhan (1964), Jacques Ellul (1964) y Neil Postman (1992) han tratado la cuestión en sus propias obras.

Ellul, en particular, reconoció en la tecnología una fuerza capaz de afectar la esencia misma del ser humano:

La técnica ha calado en lo más hondo del ser humano. La máquina tiende no sólo a crear un nuevo entorno humano, sino también a modificar la esencia misma del hombre. El medio en el que vive ya no es suyo. Debe adaptarse a un universo para el que no fue creado. [...] Fue creado con cierta unidad esencial y, sin embargo, está fragmentado por todas las fuerzas del mundo moderno. (p. 325)

Más recientemente, el autodenominado “fenomenólogo de la tecnología” Éric Sadin –quien se reconoce además como un gran deudor de la obra de Ellul –aborda el determinismo tecnológico a través de su concepto de *antrobología*, con el que pretende señalar la necesidad de una dialéctica entre el influjo inevitable de la tecnología sobre la acción humana, por un lado, y una acción humana

decidida a contrapesar ese influjo, por otro (Sadin, 2018, pp. 149-155), aunque sin detallar la medida de ese esfuerzo ni proponer acciones específicas distintas a las que se han propuesto hasta acá desde el criticismo tecnológico (como “desconectarnos” de la tecnología, pensar en sus peligros más que en sus bondades, “reconectarnos” con nuestra historia, etc.).

Así, en el tratamiento del determinismo tecnológico es posible distinguir dos posturas:

- **Determinismo tecnológico fuerte:** Sostiene que la tecnología es la fuerza principal que determina el curso de la historia humana. Según esta visión, la tecnología no solo influye en la forma en que vivimos, sino que también determina nuestras decisiones y acciones. En otras palabras, la tecnología es vista como un agente autónomo e independiente que domina sobre la sociedad y la cultura.
- **Determinismo tecnológico suave:** Reconoce que la tecnología puede influir en la sociedad de forma significativa, pero no la determina completamente. En esta visión, la tecnología es vista como un factor importante, pero no único en la evolución de la cultura y la historia.

Finalmente, el determinismo tecnológico ha sido objeto de controversia desde que se popularizó en la década de 1960. Como se dijo, algunos sostienen que la tecnología no es el único factor que influye en la sociedad y que hay otros, como la política, la economía y la cultura, que juegan un papel tanto o más importante en la configuración de la sociedad. Otros, que están más bien en las antípodas, proponen que la tecnología no es una fuerza autónoma que moldea la sociedad y la cultura, sino que más bien es al revés, es decir, la tecnología no es una causa, sino un efecto del cambio social y cultural³.

³ Ver, por ejemplo, Castells (2010, p. 13).

Para poder formular el dilema de que consta este trabajo, se adhiere acá a la hipótesis de un determinismo tecnológico suave, el cual hallará sustento empírico en lo sucesivo en la teoría de sesgos cognitivos.

Sesgo cognitivo

En psicología, el *sesgo cognitivo* se refiere a la tendencia de las personas a procesar la información de una manera que está influenciada por sus prejuicios, creencias, experiencias y emociones. Esto puede conducir a una interpretación distorsionada de la realidad y a tomar decisiones que no son completamente racionales.

Los precedentes del sesgo cognitivo se remontan a la década de 1950, cuando los psicólogos sociales Fritz Heider (1958) y Harold Kelley (1950) comenzaron a investigar cómo las personas atribuyen causas a los eventos en sus vidas.

En la década de 1970, el psicólogo Daniel Kahneman y su colaborador Amos Tversky realizaron una serie de estudios que exploraron cómo las personas toman decisiones en situaciones de incertidumbre, observando cómo ciertas distorsiones –los sesgos cognitivos– las afectaban (Kahneman y Tversky, 1974; 2000). Sus investigaciones ayudaron a establecer el campo de la economía conductual y le valió a Kahneman el premio Nobel de Economía en 2002.

Desde entonces y a lo largo de muchas décadas, otros investigadores han realizado importantes contribuciones a la comprensión de los sesgos cognitivos como patrones de pensamiento comunes en la población humana.

Existen muchos tipos de sesgos cognitivos, algunos de los más famosos son: sesgo de confirmación, sesgo de género, sesgo de autoridad, sesgo de urgencia, sesgo de disponibilidad, sesgo del

costo hundido, entre otros. En particular, se recogerá aquí la variante denominada “fijación funcional”.

El concepto de *fijación funcional* fue propuesto por el psicólogo alemán Karl Duncker en su famoso artículo *On the solving problem* (Duncker, 1945) y se refiere a la tendencia a limitar nuestra comprensión de un artefacto –o conjunto de ellos– que ha tenido éxito previamente a su uso habitual. En otras palabras, la fijación funcional nos lleva a ver las cosas solo en términos de lo que se supone que deben hacer o cómo se supone que deben ser utilizadas, y no en términos de otras posibles aplicaciones o usos, lo que impacta también en la forma de considerar nuevas soluciones o maneras de resolver un problema.

Por ejemplo, si tenemos una taza, es posible que la veamos solo como un objeto para beber líquidos. Sin embargo, si somos capaces de superar la fijación funcional, podríamos ver la taza como un objeto que también puede ser usado para guardar lápices o como elemento decorativo.

Otro ejemplo, en el orden de los sistemas tecnológicos complejos, nos lo provee Andy Zynga, experto en estrategia y transformación de negocios, analizando el problema de la innovación en las corporaciones capitalistas. En su artículo *El sesgo que nos impide innovar*, Zynga (2013) describe el problema que tiene el equipo de desarrollo de productos de PepsiCo, el cual, como consecuencia de la fijación funcional, es incapaz de pensar modos diferentes en que ciertas tecnologías pueden ser aplicadas para alcanzar un determinado nivel de salinidad en sus patatas fritas.

En definitiva, la fijación funcional puede ser perjudicial, porque nos limita en la forma en que pensamos y nos impide ver nuevas soluciones o posibilidades creativas. Para superar la fijación funcional, es necesario ser consciente de la existencia del sesgo y estar dispuesto a considerar nuevas perspectivas y formas de uso de

los artefactos tangibles e intangibles que tienen a su disposición las personas.

Alteridad

El concepto de *alteridad* (a veces referido como “otredad”) es un término utilizado en filosofía e implica la comprensión de que existen otras formas de ser diferentes a las propias, y que estas diferencias deben ser consideradas, respetadas y valoradas, ya que nuestras acciones o inacciones afectan (alteran) lo otro irremediablemente.

Ha sido trabajado por diversos autores y se rescata aquí la acepción formalizada por el filósofo lituano Emmanuel Lévinas, quien argumenta que la alteridad es la base de la ética y que la relación con lo otro es fundamental para la comprensión de la realidad y de nosotros mismos, y para establecer nuestra responsabilidad en consecuencia. Según Lévinas, con la alteridad, lo otro se erige como una condición previa y fundamental para el pensamiento, ya que nos obliga a reconocer su existencia y asumir nuestra responsabilidad hacia él.

La alteridad, es decir, la heterogenidad radical de lo Otro, únicamente es posible si lo Otro es tal con relación a un término cuya esencia es permanecer en el punto de partida, servir de entrada a la relación, ser Otro absolutamente [...], fuera de toda posibilidad de individuación a partir de un sistema de referencias [que un sujeto intentara imponerle para aprehenderlo]. (Lévinas, 2002, p. 60).

Un ejemplo típico de alteridad comprende el encuentro entre dos personas de diferentes culturas, que tienen diferentes formas de pensar y comportarse. En principio, cada persona se ve a sí misma como un sujeto individual y al otro como un objeto. Sin embargo, al mismo tiempo, producto de la alteridad, cada persona reconoce la

humanidad del otro y su capacidad para tener experiencias y perspectivas diferentes. Y de esta forma, en lugar de simplemente juzgar o ignorar al otro, cada parte intenta comprender y respetar su perspectiva y encontrar un terreno común para alcanzar una convivencia pacífica y armoniosa.

En lo que concierne a la aplicación de la alteridad a la tecnología, puede decirse, por un lado, que esta puede ser vista como una herramienta que facilita la comunicación y la interacción entre personas de diferentes culturas y orígenes. Por ejemplo, las redes sociales y las plataformas de mensajería instantánea pueden ayudar a las personas a conectarse y compartir información más allá de las barreras geográficas y culturales.

No obstante, por otro lado, también es posible que la tecnología perpetúe la desigualdad y la exclusión social, al limitar la capacidad de algunas personas para acceder a ella o utilizarla plenamente. Por ejemplo, las personas con discapacidades pueden tener dificultades para acceder a ciertas tecnologías (uno de los casos más sencillos involucra a los lectores digitales de huella dactilar, que son inútiles en personas con sus dedos amputados o en aquellas que tienen su huella corroída) y aquellas que no tienen acceso a internet pueden estar en desventaja en un mundo cada vez más conectado.

Así, la tecnología puede ser vista como un medio para promover la alteridad, si al permitir que las personas se conecten y comprendan mejor las perspectivas y experiencias de los demás, también se la considera problemáticamente como un medio potencial para perpetuar la exclusión y la marginación.

En definitiva, es importante reflexionar sobre cómo el potencial tecnológico puede ser utilizado de una manera ética. Y es en este sentido y no otro, en efecto, en el que la alteridad es aquí considerada.

La construcción del dilema de la alteridad tecnológica

Introducidos sus elementos fundamentales, el *dilema de la alteridad tecnológica* puede ser formulado de la siguiente manera:

PRIMERO. De acuerdo a la hipótesis del determinismo tecnológico suave, estamos determinados por la tecnología de forma inevitable, pero no absoluta. Esto quiere decir que como sujetos contamos con un margen de acción que nos permite subvertir o compensar el influjo tecnológico a través de una agencia activa y deliberada.

SEGUNDO. El determinismo tecnológico suave puede hallar sustento empírico o una forma posible de expresión científica recurriendo al sesgo cognitivo denominado “fijación funcional”, el cual expresa nuestra tendencia a limitar la interpretación de la tecnología a su uso ordinario o convencional. Y aunque esto nos hace perder de vista la posibilidad de usarla en otro sentido, restringiendo a su vez nuestra capacidad de innovar para resolver los problemas que se nos presentan, en cuanto sesgo podemos combatirlo mediante la construcción de estrategias concretas que nos permitan considerar nuevos modos de usar la tecnología, evitando así una determinación tecnológica absoluta.

TERCERO. La consideración anterior no es sino una forma de alteridad tecnológica, por cuanto existe una apertura a nuevas y diferentes performances o bases tecnológicas sobre las cuales catapultar nuestros pensamientos y acciones. Sin embargo, este proceso nos lleva a enfrentar un dilema moral: ¿hasta qué punto debemos considerar la alteridad tecnológica como alternancia y no como estabilidad de la tecnología, sabiendo que la alternancia conlleva renunciar a una base tecnológica estable? Dicho de otro modo, al tiempo que emerge un universo de nuevas e inusitadas posibilidades con la reforma virtual de la tecnología, ellas deben ser contrapesadas éticamente con los riesgos inherentes que conlleva el proceso, es decir, con la potencialidad de problemas del recambio que son también nuevos e inusitados.

En fin, el problema que se expresa no dice relación con el hecho de que la tecnología nos determina o atrapa de alguna manera (este es el presupuesto ontológico del dilema), sino con cómo podemos encontrar un equilibrio entre la ruptura de ese atrapamiento en la alternancia tecnológica y, por otro lado, una base tecnológica

estable, que sea funcional a un despliegue ético de nuestras actividades y existencias.

Conclusión

Una vez que se cobra conciencia del dilema de la alteridad tecnológica –o al menos de su posibilidad–, la operación acrítica de los artefactos y sistemas tecnológicos, así como su conservación sin más en cualquier ámbito, se vuelve negligente, si no tiene un severo costo moral. El dilema no implica necesariamente la acción transformadora, pero sí al menos la reflexión permanente de la base tecnológica sobre la que proyectamos y desplegamos nuestras acciones.

Aunque, como miembros de la civilización o de una sociedad particular, podríamos contar con la mejor de las performances tecnológicas posibles, esto es imposible saberlo. Y al revés, también es posible que la migración a nuevos modos tecnológicos nos arroje al peor de los mundos, si bien no hay manera de estar seguros. La cantidad de variables susceptibles de consideración para determinar la bondad o nocividad del cambio o mantención de la tecnología es inconmensurable, aunque puede delimitarse un conjunto de ellas para aplicarlas a un contexto determinado.

El dilema se vuelve más imperioso teniendo en cuenta la finitud a la que estamos afectos como individuos –y que muy posiblemente nos afecte también en cuanto especie–, sin contar que toca un amplio abanico de problemas y preguntas consabidas que se dan en la filosofía y las humanidades acerca de la tecnología, tales como:

- ¿La tecnología está transformando nuestra comprensión de la realidad? ¿Cómo se relaciona la tecnología con nuestra percepción del mundo?

- ¿La tecnología está mejorando o empeorando nuestra calidad de vida?
- ¿La tecnología nos hace más felices o más infelices?
- ¿La tecnología está reduciendo nuestra capacidad de reflexionar y pensar críticamente?
- ¿La tecnología nos está haciendo perder autonomía?
- ¿Estamos perdiendo habilidades cognitivas esenciales debido a nuestra dependencia de la tecnología?
- ¿La tecnología está deshumanizando nuestras interacciones? ¿Estamos perdiendo la habilidad de conectarnos y comprender a los demás debido al uso excesivo de la tecnología?
- ¿La tecnología está creando una brecha entre los que tienen acceso a ella y los que no?

Para concluir, se presentará un ejemplo concreto del dilema de la alteridad tecnológica, aplicándolo a la tecnología RPA (*Robotic Process Automation*, por sus siglas en inglés).

RPA es una tecnología que comprende robots de software (digitales) capaces de imitar la actividad humana para realizar una amplia gama de tareas dentro de los procesos empresariales, como entrada de datos, cálculos, procesamiento de transacciones, conciliaciones de registros bancarios, contables y financieros, y más (Moffitt et al., 2018). Los robots del RPA interactúan con las aplicaciones de la misma manera que lo hacen los usuarios humanos, pero sin la necesidad de la intervención de estos, lo que permite una mayor eficiencia y precisión. Además, estas entidades no se enferman, pueden trabajar las veinticuatro horas del día, combinarse con soluciones de inteligencia artificial y ser reprogramadas en función de la contingencia, todo lo cual supone en muchos casos incrementos extraordinarios en los indicadores de productividad.

Sin embargo, hay un problema. La programación de estos artefactos inteligentes conlleva un efecto análogo a la fijación

funcional, dada la primacía que su programa concede a las habilidades y funciones técnicas para la maximización del rendimiento de los procesos, sin tener en cuenta la dimensión ética y moral de sus acciones. Aunque esta tarea bien podría ser desempeñada por un equipo de profesionales de la filosofía que fuera contratado para complementar la performance del RPA en la empresa, supondremos que no es así en la mayoría de los casos —si no en ninguno— y que el foco está siempre en la rentabilización o la productividad de acuerdo a los cánones tradicionales de la economía capitalista.

Es en este punto donde entra en juego el dilema: ¿cómo se relacionarán estos robots con los seres humanos y otros seres vivos que cohabitan en el mismo mundo, conociendo nosotros sus enormes beneficios, pero también sus riesgos (por ejemplo, el desplazamiento de trabajadores que no podrán ser reconvertidos)? ¿Considerará su programación alguna vez la perspectiva y los intereses de otros seres o actuarán puramente en función de su itinerario técnico? ¿Cómo se definirá la relación entre humanos y robots en términos de derechos, responsabilidades y moralidad? ¿Qué clase de ajustes podrían hacerse sobre esta tecnología y cuáles son los más apropiados para asegurar su efecto benigno sobre los seres y cosas del mundo?

En suma, esta forma del dilema expresa un conflicto hoy latente entre la creciente automatización de las corporaciones —y, con ellas, de las sociedades capitalistas occidentales— y la necesidad de preservar, en principio, la alteridad entre robots y humanos. Pues el RPA puede ofrecer grandes beneficios en términos económicos, pero también puede tener consecuencias impredecibles y perjudiciales si se pierde de vista el efecto final —y no parcial (asociado, por ejemplo, a la privacidad de los datos o a la mejora de los sistemas de salud)— que produce sobre la existencia humana, la diversidad y la diferencia en el mundo.

Referencias

- Benjamin, G. (2016). *The Cyborg Subject: Reality, Consciousness, Parallax*. Palgrave Macmillan.
- Bimber, B. (1990). Karl Marx and the Three Faces of Technological Determinism. *Social Studies of Science*, 20(2): 333–351.
- Castells, M. (2010). *The Rise of the Network Society: The Information Age: Economy, Society, and Culture Volume I*. Wiley-Blackwell.
- Duncker, K. (1945). On the solving problem. *Psychological Monographs: General and Applied*, 58(6): i-113.
- Ellul, J. (1964). *The Technological Society*. Vintage Books.
- Halpern, O. (2015). *Beautiful Data: A History of Vision and Reason since 1945*. Duke University Press.
- Héder, M. (2021). AI and the resurrection of Technological Determinism. *Információs Társadalom (Information Society)*, 21(2), 119–130.
- Heider, F. (1958). *The psychology of interpersonal relations*. Wiley.
- Kahneman, D., y Tversky, A. (1974). *Judgment under uncertainty: Heuristics and biases*. *Science*, 185(4157), 1124-1131.
- Kahneman, D., y Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge University Press.
- Kelley, H. H. (1950). The warm-cold variable in first impressions of persons. *Journal of Personality*, 18(4), 431-439.
- Lévinas, E. (2002). *Totalidad e infinito. Ensayo sobre la exterioridad*. Ediciones Sígueme.
- McLuhan, M. (1964). *Understanding Media: The Extensions of Man*. McGraw Hill.
- Moffitt, K., Rozario, A., Vasarhelyi, M. (2018). Robotic Process Automation for Auditing. *Journal of Emerging Technologies in Accounting*, 15(1): 1-10.

Mumford, L. (1970). *The myth of the machine: The pentagon of power*. Brace and Company.

Postman, N. (1992). *Technopoly: The Surrender of Culture to Technology*. Vintage Books.

Sadin, E. (2018). *La humanidad aumentada: La administración digital del mundo*. Caja Negra.

Thorstein, V. (1904). *The Theory of Business Enterprise*. Charles Scribner's Sons.

Ville, S. (2004). Transport. En R. Floud & P. Johnson (Eds.), *The Cambridge Economic History of Modern Britain*. Cambridge University Press.

Zynga, A. (13 de junio de 2013). The cognitive bias keeping us from innovating. *Harvard Business Review*.

ON THE REALITY AND VALUE OF TECHNOLOGY

SOBRE LA REALIDAD Y EL VALOR DE LA TECNOLOGÍA

TOBIAS ENDRES

Technische Universität Braunschweig, Alemania
t.endres@tu-braunschweig.de

RECIBIDO: 26/05/2023

ACEPTADO: 25/07/2023

Abstract: In this paper I present an interpretation of Ernst Cassirer's diagnosis that modern culture is in need of an 'ethicization' (*Ethisierung*) of technology. The conclusion of the 1930 essay *Form and Technology* has not merely met with incomprehension among researchers. From the side of the Frankfurt School, Cassirer, presumably because of this choice of term, even earned ridicule. In its peculiarity, the concept of ethicization was not even attempted to be translated in the latest English translation (2013) of Cassirer's essay on technology, which is why it can be assumed that it is still not understood, even among experts. I thus present a close reading of *Form and Technology* that is in line with Cassirer's system of symbolic forms from his main work and also with relevant posthumous writings. Cassirer's position turns out as a defense of an autonomy principle that is at the base of any expression of the human mind, including technology. The downside of technology for modern culture lies in its entanglement with science and economy in the historical shape of capitalism. For none of those symbolic forms is essentially normative, philosophy's task is to reflect on and state those cultural resources that can help to break up this ligation.

Keywords: Continental Philosophy of Technology; Form and Technology; Ernst Cassirer; technology as symbolic form; Symbolic Forms; *Lebensphilosophie*; Transcendental Philosophy

Resumen: En este artículo presento una interpretación del diagnóstico de Ernst Cassirer según el cual la cultura moderna necesita una 'eticidad' (*Ethisierung*) de la tecnología. La conclusión del ensayo de 1930 *Forma y tecnología* no sólo ha suscitado incompreensión entre los investigadores. Por parte de la Escuela de Frankfurt, Cassirer, presumiblemente a causa de esta elección del término, se ganó

incluso el ridículo. En su peculiaridad, el concepto de eticidad ni siquiera se intentó traducir en la última traducción al inglés (2013) del ensayo de Cassirer sobre la tecnología, por lo que cabe suponer que sigue sin comprenderse, incluso entre los expertos. Así pues, presento una lectura atenta de *Forma y tecnología* que se ajusta al sistema de formas simbólicas de Cassirer de su obra principal y también a escritos póstumos relevantes. La postura de Cassirer resulta ser una defensa de un principio de autonomía que está en la base de cualquier expresión de la mente humana, incluida la tecnología. El inconveniente de la tecnología para la cultura moderna reside en su imbricación con la ciencia y la economía en la forma histórica del capitalismo. Puesto que ninguna de esas formas simbólicas es esencialmente normativa, la tarea de la filosofía consiste en reflexionar y enunciar los recursos culturales que pueden contribuir a romper esa ligadura.

Palabras clave: Filosofía Continental de la Tecnología; Forma y Tecnología; Ernst Cassirer; tecnología como forma simbólica; formas simbólicas; *Lebensphilosophie*; Filosofía trascendental

Introduction

This essay asks about the reality and actuality of technology. This does not formulate an essentialist question that asks about the substance of technology, but rather reflects on the function of technology in the construction of culture, that is, on the mental and material resources in discovering and inventing our world. Consequently, technology is not seen primarily as an instrument for the domination of nature, but above all as a meaning-generating phenomenon of human life. Such a perspective is to be gained following Ernst Cassirer, who understands technology as a symbolic form.¹ Starting from the three-volume *The Philosophy of Symbolic*

¹ The Hamburg edition of Cassirer's collected works is referred to below with the established abbreviation ECW and the posthumous writings (*Nachlass*) with the abbreviation ECN.

Forms (1923-1929)² and the article *Form and Technology* (1930)³ the hypothesis shall be substantiated that the reality of technology can only be grasped if it is understood as a mental principle of form and not as a mere tool. Only from this point of view it is possible to decide about the value and unvalue of technology as well as the problem connection between freedom and alienation through technology.

Technology as a symbolic form

For Cassirer, the “primacy of technology” (Cassirer, FT, p. 272/ECW 17, p. 139) of his time is in a sense a fact, insofar as technology in a very real way extends its impact into almost all areas of culture. In 1930, inventions such as television are already on their way to becoming mass media and the first analog computers have been invented; newspapers are at their historical peak and report on current events in Berlin up to four times a day. The so-called Philosophy of Life (*Lebensphilosophie*), for example in the writings of Ludwig Klages (1929) or Georg Simmel (1911), had already expressed its unease about the pervasion of all areas of life by culture and technology at the time when Cassirer was trying to understand the problem of an antagonism of life and spirit (and thus technology) in a deeper way. For Cassirer, the problem is formed in such a way that the cultural “*counterforces* to technology” (Cassirer, FT, p. 272/ECW 17, p. 139), one would probably think - Cassirer does not specify it - of religion, myth, art and language, enter into an alliance with technology in order to be able to subsist, which ultimately leads

² The English translation of Cassirer’s main work is additionally referred to by the abbreviation PSF.

³ The English translation of this writing of Cassirer, which is central for the context of technology, is abbreviated FT.

to a subjugation of their own meaning under that of technology. At the same time, however, he also reminds us of the autonomy of any “energy of spirit” (Cassirer, 1923/ECW 16, p. 79 [76]⁴), as which he understands symbolic forms such as myth, language, science, but also technology. Contrary to the Philosophy of Life, which sees the fate of the modern, technically saturated culture as a tragedy, Cassirer is concerned to create hope for freedom through the further development of the entire culture, because the “essence and basic determination of spirit does not tolerate any *external* determination” (Cassirer, FT, p. 273/ECW 17, p. 139). As a mental (*geistig*) principle, according to Cassirer’s thesis, technology cannot in the long run suppress its own extra-technical formation conditions. Technology originates in the freedom of the spirit and will find its way back to it in an act of self-knowledge. This designates the task of a philosophy of technology, which is to perform this act of an “ideal demand” (Cassirer, FT, p. 273/ECW 17, p. 140).

In fact, this formulates a task rather than a solution to the problem of technology. Cassirer diagnoses that the development of modern technology has found an echo in philosophy since Fichte’s *Wissenschaftslehre*, via the Value theory of the Baden School of Neo-Kantianism up to Pragmatism and, of course, to the Philosophy of Life, but technology and philosophy continue to stand in a “disparity” (Cassirer, FT, p. 275/ECW 17, p. 141) to each other. According to Cassirer, the philosophy of technology must step out of its niche position and shed its “peripheral character” (Cassirer, FT, p. 274/ECW 17, p. 141) by being incorporated into Kant’s critical, transcendental program and “brought . . . before its forum” (Cassirer, FT, p. 274/ECW 17, p. 141) to ask about the validity, the *quid juris*, of its “meaning and right” (Cassirer, FT, p. 274/ECW 17, p. 141). But what does it mean that the unity of technology and philosophy can only be gained through their critique, only through

⁴ The number in brackets here and in the following refers to the English translation.

“the insight and clear and frank acknowledgment that this particular case involves more than a mere difference” (Cassirer, FT, p. 275/ECW 17, p. 142)? In Cassirer’s further development of Kant’s transcendental philosophy, the critique of reason famously becomes the “critique of culture” (Cassirer, PSF 1, p. 9/ECW 11, p. 9). From this point of view, the critique of technology and the recognition of its independence and autonomy can only mean wanting to integrate it into the system of symbolic forms. And this is precisely what Cassirer wants to do with *Form and Technology*.⁵ Such an undertaking has consequences for Cassirer’s system, which hitherto has consisted of the symbolic forms of myth, custom, religion, art, language, and science,⁶ because although Cassirer had conceived his system as open from the ground up (Kreis, 2010, pp. 388-401; Endres, 2020, pp. 124-129), the addition of new elements changes the configuration from which the philosopher’s point of view first emerges. Cassirer writes:

This fact determines the task that philosophy has to fulfill with respect to the current development of technology. This task cannot be limited to assigning technology a predetermined “place” in the whole of culture and, therefore, in the whole of a systematic philosophy that aims to be the

⁵ It would have to be examined elsewhere whether this is an attempt to assimilate the discipline of philosophy of technology to transcendental philosophy or a proposal to conceive philosophy of technology as a special case of philosophy of culture (Favuzzi, 2017, p. 15). Against this would be the fact that Cassirer explicitly acknowledges the emergence of modern philosophy of technology with Ernst Kapp’s main work of 1877 as well as many contemporary contributions originating from technicians in their independence. Undoubtedly, he calls for a “critical’ consideration and justification” (Cassirer, FT, p. 274/ECW 17, p. 141) of the philosophy of technology, but whether this must necessarily come from the side of the philosophy of culture seems doubtful to me in view of Cassirer’s conciliatory philosophical attitude.

⁶ In addition, for the first time in *Form and Technology*, as the quotation below shows, economy, state and law also appear as symbolic forms alongside technology.

intellectual expression of culture. Technology cannot simply be placed next to other areas and formations [Gebilden], such as “economics” and “the state,” “morality” and “law,” “art” and “religion.” For in the realm of spirit, separate domains never stand simply together or next to one another. Here, the community is never spatially static but possesses a dynamic character. One element is found “with” the other only to the extent that both assert themselves in opposition to each other, thereby mutually “setting each other into opposition” [auseinandersetzen]. Thus, every introduction of a new element [Element] not only widens the scope of the spiritual horizon in which this confrontation [Auseinandersetzung] takes place but also alters the very mode of seeing. This process of configuration not only expands outwardly; it also experiences in itself an intensification and heightening so that a simultaneous qualitative transformation, a specific metamorphosis, occurs (Cassirer, FT, p. 275/ECW 17, p. 142).

At no other place in his work Cassirer makes the consequences of the appearance and dominance of a new intellectual force like technology clearer to his reader: The critique of technology has as a consequence the critique of all symbolic forms, since critique is first of all an incomplete process (Cassirer, FT, p. 274/ECW 17, p. 141) and since secondly the meaning of all other symbolic forms is changed by the appearance of technology.⁷ Only insofar as the question of meaning of all symbolic forms is continuously posed, the “questions of being and validity” (Cassirer, FT, p. 275/ECW 17, p. 142) of each can be decided and philosophy can fulfill its task as the “logical *conscience* of culture” (Cassirer, FT, p. 275/ECW 17, p. 142).

With the question of meaning, Cassirer’s philosophical idealism now also enters the scene in the question of technology, an idealism, admittedly, that does not simply follow on from the traditional

⁷ It would have to be examined to what extent this statement leads to problems for the horizontal (Moss, 2015, pp. 4, 11f.) and the complementary (Luft, 2015, pp. 14, 166-168, 178, 210) reading of the system of symbolic forms.

Platonism, but starts from a correlativity of ideality and materiality⁸ - from a facticity-dependent Platonism (Hogrebe, 2006, p. 235), one could say. He justifies this by pointing out that the progress and formation of the philosophy of technology has already overcome the materialism of the 19th century. More recent authors such as Friedrich Dessauer, Eberhard Zschimmer, and Max Eyth, who are not primarily philosophically trained but are first and foremost natural scientists, engineers, and pioneers of technology, ask about the ideational content of technology, implicitly going back to Plato, for whom the relationship between idea and appearance grounds “not in the figures of nature but in the works and formations [Gebilde] of τέχνη [techne]” (Cassirer, FT, p. 277/ECW 17, p. 143). Technology, this is Cassirer’s point, is not a skill for reproducing natural entities, but rather shaping the material world through a mental vision, which can operate in a nature-mimetic way, but at the same time also proceeds analogically, i.e. according to the model of already created mental entities, but is ultimately essentially symbolic in nature (Cassirer, PSF 1, pp. 133-146/ECW 11, pp. 133-146; Cassirer, FT, p. 303/ECW 17, p. 170). Through the fact that humans are symbolically active, they are essentially free in whether they form their environment in close reference to nature or close to the abstract. This freedom, considered as a whole, is the basis of Cassirer’s idealism, according to which all symbolic forms, including technology, are free expressions of the human mind. Against this background, philosophical idealism means to take a step back from the final products of technology and to ask for the principles of form and construction of technology, i.e., to “focus on the *concept of form* rather than on the *concept of being* of natural

⁸ This point seems to me to be lost in Rohbeck’s (1997, p. 202) analysis, when he understands Cassirer’s Platonism of Technique in such a way that one can also make the beginning with the general principle of form and does not have to start with an original correlation.

science” (Cassirer, FT, p. 278/ECW 17, p. 145). Asked about the conditions of possibility of, for example, the Eiffel Tower, one can consequently answer in a technical-scientific way and explain the chemical and physical nature of the iron used and the rivets, which were produced according to the puddling process, or the principle of a caisson for the construction of the foundations. Or one answers in view of, which mental (*geistig*) configuration must be given, in order to want to establish such a building. For this, however, the question of meaning must first be posed correctly, or spoken with Cassirer:

If, instead of beginning from the existence of technological works, we were to begin from the form of the effective action of technology and shift our gaze from the mere product to the mode and type of production—and to the lawfulness revealed in it—then technology would lose the narrow, limited, and fragmentary character that otherwise seems to adhere to it. Technology adapts itself—not directly in its end result, but with a view to its task and problematic—into a comprehensive sphere of inquiry within which its specific meaning and original spiritual tendency can be determined (Cassirer, FT, p. 278/ECW 17, p. 145).

As already in the context of the structural analysis of language, myth and science, Cassirer wants to direct the view, as he says in connection with Wilhelm von Humboldt (Cassirer, 1930b/ECW 17, p. 205 [879]; Cassirer, FT, p. 281/ECW 17, p. 148), from the *ergon* to the *energeia* of technology. When Cassirer speaks of the lawfulness of the ‘spiritual energy’ technology, he is concerned with the autonomy of this form. However, this must first be uncovered, for it is an essential finding of Cassirer’s that this question is obscured by the question of the value of technology. With regard to culture as a whole, according to Cassirer, German idealism, in distinction from Rousseau and the Age of Enlightenment, in which cultivation was tied to notions of happiness and moral perfection, brought about a turning point in time. It was not until Kant’s *Critique of Judgment* (1790) that the independence and autonomy of the

beautiful, as distinct from the feelings of pleasure and displeasure and the demands of ethics, was discovered. And it is from here that Cassirer looks to technology when he states that contemporary philosophy ascribes to culture purposes “that are foreign to the pure creative will [*Gestaltungswillen*] and pure creative power [*Gestaltungskraft*] of technology” (Cassirer, FT, p. 280/ECW 17, p. 147). The hypothesis that Cassirer thus advocates is that the reality of technology is grasped only when we have grasped the laws of its mode of generation peculiar to it as compared to other forms of intellectual expression, when we understand the mode that is peculiar to it alone in the construction of culture. We therefore do not ask about the value, unvalue, benefit, disadvantage or tragedy of technology, but about its “authentically objective ‘form’” (Cassirer, FT, p. 281/ECW 17, p. 148).

Language and technology

Cassirer begins the ‘objective analysis’ of technology with an analogy to language, in which he draws on two lectures by the “technician and . . . thinker of technology” (Cassirer, FT, p. 282/ECW 17, p. 149) May Eyth. Eyth argues that the cultural history of man is determined by an interrelation between (linguistically constituted) knowing-that and (action-oriented) knowing-how, and that the development of civilization is characterized by a dominance of knowing-that over skillful action from which knowledge derives. In this he sees the philosophical problem that man forgets about “the tool of the spirit . . . the spirit of the tool” (Cassirer, FT, p. 282/ECW 17, p. 149). It is not surprising that for Cassirer this is a “real *philosophical* problem” (Cassirer, FT, p. 282/ECW 17, p. 149), since *The Philosophy of Symbolic Forms* began in its first volume with an

analysis of language.⁹ If technology can be understood as a symbolic form, then it is already clear at this point that the mental principle from which it grows must be the symbol and that technology, according to its function, must be classified in the spectrum of the known symbolic functions of expression, presentation and pure signification and that its real development follows the scheme of a mimetic, analogical and symbolic phase, which is also known from *The Philosophy of Symbolic Forms* (Endres, 2020, pp. 103-133). Cassirer consequently objects to Eyth that knowing-that and knowing-how, i.e. language and action, do not form an antagonism so much as they are “originally united” (Cassirer, FT, p. 284/ECW 17, p. 150), since in both a principle of form becomes apparent and language *acts* exactly like technology. No linguistic act, according to Cassirer, who in this follows Humboldt, is a depiction of the world, but language, just like any practical skill, is an action in the world, a “real act of world-creation” (Cassirer, FT, p. 284/ECW 17, p. 150) by “raising up of the world to form” (Cassirer, FT, p. 284/ECW 17, p. 150). If this is so, if the analogy of language and technology thus bears, then by analyzing the form of technology it must also be possible to trace a change of form, an inner transformation of its meaning. The meaning of technology cannot lie in a linear development from the body as a kind of first technical means to an ever more far-reaching mastery of space and time, because in this one would remain in a purely instrumental view of technology and we had already shown at the beginning that value, means, purpose are subordinate categories to the question of meaning. The yield of the analysis of form is to show that in the

⁹ In Cassirer research it is discussed whether this is not a tension to the central position of myth within the philosophy of symbols, which Cassirer refers to in various places as the “mother soil” (Cassirer, 1925/ECW 16, p. 266 [168]) or “mother earth” (PSF 2, p. 1/ECW 12, p. 1) of culture. Important contributions to the resolution of this tension can be found in Rudolph (1992, pp. 79 ff.; 2003, pp. 10, 78-80, 222-226).

extension of action through technology lies a qualitative change, whereby the creation of a “new world-matter” (Cassirer, FT, p. 284/ECW 17, p. 151) becomes possible.

Homo divinans and homo faber

For Cassirer, an analysis based on Kant’s critical method always means ‘reconstructive analysis,’ a method that comes from Kant’s *Prolegomena* (1783)¹⁰ and that Cassirer adapted and further developed from his neo-Kantian teachers Hermann Cohen and Paul Natorp (Endres, 2020, pp. 47-52). It is reconstructive on the one hand because in it one goes back from the *forma formata* to the *forma formans*, i.e., to the functioning and making; and it is reconstructive on the other hand because one goes back to the mimetic, i.e., ultimately mythic, phase of the corresponding symbolic form in order to understand its inherent lawfulness and the principles at work in it. In Cassirer’s work, myth in comparison to more developed symbolic forms often serves as a contrasting foil to make such principles and peculiarities visible. Now in the will to master nature Cassirer sees a crossroads that separates humans in early history from humans in later epochs of cultural development. Ethnology as well as comparative linguistics and religious studies at the time provide Cassirer with the material to establish such a crossroads in the light of a comparison between “cold cultures” (Lévi-Strauss) and the industrialized contemporary culture. Against this background, Cassirer distinguishes between *homo divinans*, the human being who wants to be effective magically, and *homo faber*, who knows how to be effective through technology. In his own words:

¹⁰ Its definition reads: “[A]nalytic method... signifies... that one proceeds from that which is sought as if it were given, and ascends to the conditions under which alone it is possible” (Kant, 2004, p. 4:277, § 5 note *).

Humans from an earlier stage are distinguished from those of a later stage, just as magic is distinguished from technology. The former may be designated as *homo divinans* and the latter as *homo faber*. The whole development of humanity presents itself, then, as a completed process, containing innumerable intermediary forms through which the human being moves from the initial stage of *homo divinans* to the stage of *homo faber* (Cassirer, FT, p. 285/ECW 17, p. 151).

The technical human is in this optics therefore not a late product of the modern civilization¹¹, but a human ideal type, which becomes already in the transition from the prehistory to the early history style and form defining for mankind. Cassirer now reproaches contemporary ethnology for not describing the relationship of *homo divinans* to *homo faber* without prejudice, if the latter is characterized by the fact that he projects his subjective drives and volitions into the outside world, and the latter by the fact that he knows the objective causal connections and limits of his volitional influence. This explanation of the specific difference of both types is namely circular, insofar as it is presupposed that subject and object are already always fixed categories through which humans experience themselves and their world. But this is not so. According to Cassirer's idealism

these borders are not 'in themselves' objectively before us; rather, they must first be set down and secured, they must first be erected by the labor of spirit. The manner of setting these borders takes place differently according to the overall attitude in which spirit exists and according to the direction in which it moves. Each transition from one comportment and direction into another always ends in a new 'orientation,' a new relationship between the 'I' and 'reality.' (Cassirer, FT, p. 286/ECW 17, p. 153).

¹¹ This could be thought, because the basic idea of *homo faber* goes back to Anaxagoras, but in the modern age it is only through Henri Bergson and Max Scheler that it becomes widespread in philosophy.

One could say with Cassirer that reality is formed by a specific understanding of reality and that the degree of reality of a certain world view cannot be judged externally, e.g., from the scientific point of view. Rather, what is considered real to a person is what follows the internal standards and criteria of the dominant intellectual orientation. Thus, one misunderstands the magical human if one applies the category of causality, which presupposes an already established separation of subject and object, to his actions. Herein lies, as it were, a fallacy and a loss of object, for the specific difference between magical and technical action stands and falls with the recognition of the peculiar form of magical thought. Cassirer literally: “If we assume that the principle of ‘causality’ [Kausalität] and the question concerning the ‘reasons’ of being and the ‘causes’ [Ursachen] of events already prevail in the magical apprehension of nature, then the barrier between magic and science falls away” (Cassirer, FT, p. 287/ECW 17, p. 153). Magical action, then, is not a kind of imperfect experimental physics, as James George Frazer influentially postulated, but the real consciousness of the omnipotence of the I: cause and effect demand a form of congruence that is incomprehensible to causal thinking, for an action has taken place in real terms not only when the result has turned out as desired, but also when this has been anticipated ‘correctly,’ that is, according to magical ideas. Cassirer:

All “real” actions, if they are to be successful, need such magical preparation and anticipation. Warring or raiding, fishing or hunting can succeed only if every individual phase is magically anticipated and at the same time “rehearsed” in the right way. Already in the magical view of the world, the human being tears himself away from the immediate presence of things and builds his own empire, with which he reaches out into the future. However, if, in a certain sense, he is freed from the power of immediate sensation, he has only exchanged it for the immediacy of desire. In this immediacy, he believes he is able to seize reality directly

and to conquer it (Cassirer, FT, pp. 288 f./ECW 17, p. 155).

Thus already the magical human is no longer a plaything between the forces of nature and his own drives and sensations, but already exists temporally and in certain respects freely. But still the desire and the will to be effective is an immediacy, is not yet postponed and seized planning after causal understanding. In the two “originary-forms of magic” (Cassirer, FT, p. 289/ECW 17, p. 156), the word-magic and the image-magic, in which temporal and spatial ideas are generated by repetition, lies accordingly an “accomplishment of ‘subjectivity’” (Cassirer, FT, p. 289/ECW 17, p. 156), which first lead to the separation of I and world, of being and deed. Man has no longer succumbed to his impressions, but has taken the path of the “first active direction” (Cassirer, FT, p. 289/ECW 17, p. 156), which leads to technical behavior. This in turn is characterized by a double process in which the will can experience itself as directly acting, but must also be capable of postponement and renunciation. Only in this way can the technical object come into being by relegating its manipulation as a goal to the distance and by recognizing its dependence on the being of nature, not on that of desire. This process precedes the planning grasp and makes it possible only through the growing recognition of an independent being of things and of the laws that are operative in them.

The thesis of the mediateness of the tool and the mind

The central theses of Cassirer’s philosophy of technology have been substantiated up to this point to the extent that it has been possible to prove (1) that technology can be understood as a symbolic form, as an autonomous power of the mind, and (2) that in the mythical worldview an anthropological turning point from *homo divinans* to *homo faber* is taking place, which in a sense allows subject, object

and mediality to come into being in the first place. Perhaps the most central thesis in *Form and Technology*, however, Cassirer advances when he asserts that from a philosophical point of view there is no difference between high technology and primitive tools. The guiding idea here is that it is through *homo faber* that mediality in the true sense is first gained. While *homo divinans* remained trapped in the immediacy of desire, the insight that a mediation between the realm of needs and that of nature is required was revealed to the technically competent human being. The tool is thus mediating medium between these two spheres and thus essentially mediateness. Every simple tool and every marvel of technology have as their principle that they are constructed as media mediating desire and reality. This hypothesis, in turn, supports the thesis of technology as a symbolic form, for mediateness and mediality are specific to all symbols. And this observation, in turn, supports the thesis that technology is a 'spiritual energy' and that there is a 'spirit' of technology, which Max Eyth already wanted to place alongside the techniques of the mind. In Cassirer's words:

It would not be an exaggeration to say that the transition to the first tool not only contains the seeds of a new *mastery of the world* but also marks a turning point in *knowledge*. The mode of action established here grounds and steadies, for the first time, a type of mediacy that belongs to the essence of thought. All thought in its pure logical form is mediated. It is directed to the discovery and extraction of a mediating structure that joins the opening sentence and the ending sentence of a communicative chain. The tool fulfills the same function, presented here in the logical sphere, in the *objective* sphere. It is grasped, as it were, in objective intuition; it is not merely the *terminus medius* of thinking. It sets itself between the first position taken by the will and its goal. Only in this in-between position is it permitted to separate them and set them at a proper distance. So long as the human being makes use only of his limbs, his bodily 'organs,' in order to achieve his goals, such distancing is not yet reached. Admittedly, he *effectively acts* on his environment—however, there is a great distance between this effective activity and the *knowledge*

[Wissen] *of this effective activity*. Whereas all human doing is absorbed in apprehending the world, human beings cannot yet *comprehend* [*ergreifen*] it, because they do not yet conceive [*begreifen*] of it as an objective figure, as a world of objects. The elementary taking-possession-of, immediate physical grasping [*Fassen*], is not a constructive ‘comprehending’ [*Erfassen*]. It does not lead to a construction in the region of pure looking or in the region of thinking. In the tool and its use, however, the goal sought after is, for the first time, moved off into the distance (Cassirer, FT, pp. 291 f./ECW 17, p. 158f.).

Thinking and acting thus become two sides of the same coin, which also confirms the simultaneous originality of technology and language claimed at the beginning: both participate in the construction of an object world opposed to us, which we humans strive to control mentally and technically. This gain, however, is at the same time a loss, because the fact that the immediacy of desire and its magical fulfillment are successively pushed back by a new understanding of reality opens up the path of quasi-unlimited possibilities, which is not hedged in by modern technology, but consistently pursued and expanded. The construction of the Eiffel Tower does not simply fulfill a predetermined goal, for example, to be the entrance portal and observation tower of the 1889 World’s Fair, but generates meaning and significance for countless other possibilities of thought and action. Not only was it for a time the tallest building in the world and thus simply a demonstration of what is technically possible, it was also used for military communications and is still used today for radio and television reception, and at the same time it is the national symbol of France and an icon of modernity. The meaning of technical doing, Cassirer concludes, is consequently transformed in modernity into the pure form of action. According to this thought, the meaning of technology does not lie in the achievement of predictable ends, rather “it is the pure form of doing, the type and direction of the constitutive force as such, that determines this meaning” (Cassirer, FT, p. 297/ECW 17, p. 164).

Thus, the meaning of technology is essentially determined in terms of freedom: Modern technology allows human beings almost unlimited dominion over nature, but without being able to provide an answer to the goal of this dominion. Technology becomes autonomous, and through it human beings come to the self-knowledge that the meaning potentials of technology are in their own hands, which in the end simply goes hand in hand with responsibility for technology and the reality it creates for us.

Does technology alienate?

Technology as symbolic form, as we have seen, participates in the demarcation between I-consciousness and the external world and thus in the construction of objective reality. In this it stands in analogy to language, which in its development also marks out the boundaries of subject and object more and more clearly. Therein lies the function of technology (and language). The meaning of technology thus lies on the one hand in the construction of a counter-world, which for one thing follows its own laws, but is also increasingly mastered by technology itself, and on the other hand in the absence of an absolute purpose. In its very nature, technology is objectifying and autonomous. Against the background of this (preliminary) result, Cassirer now dares to include the question of value in the reflection and to express the suspicion, familiar from Value theory and the Philosophy of Life, that technology in its execution essentially alienates human beings from themselves, from their subjectivity. In any case, the idea that the cultural world, especially in technology, has the potential to crush rather than liberate the individual cannot be dismissed out of hand. Cassirer puts it this way:

Is not what was regarded here as the authentic achievement of technology

nothing other than the basic evil from which it suffers? Does not this exploitation of the world of objects [Objekte] at the same time necessarily result in the estrangement of human beings from their own essence, from what they originally are and feel? With the first step into the world of facts that technological labor secures and constructs for him, the human being also appears to be subjected to the law, to the brute force of factual matters. And is this brutality not the strongest enemy of the inner life enclosed in his I, in the being of his soul? All technology is a creation of spirit; spirit can only ground its own mastery in this way because it conquers all the forces that find themselves enclosed within it, despotically holding them down. To become master, it must not only restrict the free realm of the soul but also deny and destroy it. No compromise is possible in this conflict. Spirit, whose goal and power emerge in technology, is the irreconcilable opponent of the soul. And as it progressively estranges the human being from his own center of life, the same thing occurs concerning the human relationship to the whole of nature, insofar as this is not taken in one of the senses already distorted by technology, insofar as it is not thought of as a mere mechanism obeying general laws, but is felt in its organic peculiarity and fullness of life. The more the power of technology grew within the spheres of modern culture, the more passionately and inexorably did philosophy levy this complaint and accusation against it (Cassirer, FT, pp. 279 f./ECW 17, p. 164).

Cassirer identifies Klages as the most exposed representative of the idea that the spirit and thus the entire culture, above all technology, is in fundamental conflict with the soul of the human being. In this view, the human becomes a deficient being, not in morphological terms, but in psychological terms: the animal has ahead of him the ability to live in harmony with the cosmos. Cassirer takes this idea seriously and points out that such an anthropological view does not simply refer to the problematic instrumental rationality, which is especially expressed in technology, so that its excesses can, as it were, be outweighed by more ‘reasonable’ products of culture in the emphatic sense. Because the raised question “is directed not to the consequences but to the ground” (Cassirer, FT, p. 298/ECW 17, p.

165), the attempt to resolve it cannot consist in “to compare the pernicious effects of the rational-technical spirit . . . with other pleasant and beneficial consequences, drawing an acceptable or favorable balance out of this comparison by a ‘hedonistic calculus’” (Cassirer, FT, p. 298/ECW 17, p. 165).

The question of the ‘ground’ for technology is now an anthropological one: technology as a symbolic form, as an inherent factor in the construction of culture, must be understood as a “necessary path toward becoming human and as a particular phase along this path” (Cassirer, FT, pp. 298 f./ECW 17, p. 165). In this perspective, not only is culture the medium of expression of the human being, but the human being itself becomes the medium of culture. Every symbolic form enables and constitutes a certain mode of being of the human. Thus, according to the “*functional considerations and analysis*” (Cassirer, FT, p. 298/ECW 17, p. 165), in a certain sense technology is what makes human beings human. Through humans, one could say, technology realizes their freedom by snatching them from the passivity of impressions and opening up a mode of freely shaping the world. Cassirer illustrates this connection with the philosophies of Friedrich Schiller, Johann Gottfried Herder and Wilhelm von Humboldt: While in Schiller’s perspective the ‘centrifugal drive’ leads to the thesis that the human being is only human where it plays, that is, where it develops freely and creatively, Herder and Humboldt emphasize the kinship of language and art, which also leads to an accentuation of the creative and the free in the context of the question of the value of culture. Cassirer wants to give technology precisely this status in the questions about the human being and culture and states again that philosophy has to make up for an omission here. “The domain of effective activity of technology seems, however, to be denied any such acknowledgment” (Cassirer, FT, p. 299/ECW 17, p. 166). In the recognition of this creative side of technology, it is now revealed

that the free shaping of the world never refers only to an outside, but also has a feedback effect, which we have already explained in the context of the demarcation between subject and object. The conquest of the world is a “double movement” (Cassirer, FT, p. 300/ECW 17, p. 167) and from this insight the Klages’s thesis of the oppression of the subject by culture can be transformed into a counterthesis, according to which self-knowledge presupposes precisely this conquest of the outside. Cassirer writes about the symbolic forms:

Each new figure of the world opened up by these energies is likewise always a new opening out of inner being; it does not obscure this being but makes it visible from a new perspective. We always have before us a manifestation from the inner to the outer and from the outer to the inner—and in this double movement, in this particular oscillation, the contours of the inner and the outer world and their two-sided borders are determined. This is also true for the effective activity of technology, because it is in no way directed toward the seizing of a mere ‘outside’; rather, it encloses in itself a particular turn inward and backward. Here, too, it is not about *breaking free* of one pole from another but about both being *determined* through each other in a new sense (Cassirer, FT, p. 300/ECW 17, p. 167).

At this point, Cassirer indirectly confronts Klages with an epistemological argument: the determination of what is unleashed technology, what is exuberant and oppressive cultivation, and what is living subjectivity presupposes that the poles ego and world, ego and culture have already mutually determined themselves to such an extent that they can understand themselves *as* antagonistic in the sense of ‘cultivated’ and ‘living’. Thus, the *original* antagonism becomes a correlation constitutive of world and self-knowledge – a revelation.

Mimetic, analogical and purely symbolic phase of technology

Cassirer's philosophy of technology relies on revelation instead of tragedy of culture. The self-revelation of the mind follows in Cassirer's system, as we said before, a three-stage ideal-typical scheme. Cassirer calls the principles of symbolic development mimetic, analogical, and purely symbolic, and this scheme is familiar to Cassirer's readers from the first volume of *The Philosophy of Symbolic Forms* on language. Mimetic means that the formation of a linguistic, mythical, or artistic object is modeled on nature. In the context of language analysis, the onomatopoeic sounds took this place. The rain magic, in which produced smoke clouds resemble the longed-for rain clouds, would be an example from myth, in which altogether the mimetic principle prevails most strongly. It could also be summarized by saying that in the mimetic phase of a symbolic form, the mental sense of expression (its meaning) resembles the sensual means of expression. The analogical phase sets itself apart from this and it is language that helps the analogical principle to break through. Here the guiding principle is (re-)presentation: one mental content can stand for another without resembling it sensually. In the development of natural languages, different sounds are increasingly freely linked and their meaning increasingly determined by convention rather than by similarity. The purely symbolic phase, in turn, is characterized by the fact that in it the original symbolism of all modes of expression and understanding becomes the guiding principle. This reveals the relational character of every symbol: Its meaning then no longer results from a reference to something that can be sensually grasped, but exclusively in the context of reference to other symbols. Exemplary for this is modern mathematics and natural science. The reality character of the quark symbol q , for example, neither lies in a sensual-material substance nor in analogy to such a substance, but is given by natural law

reference connections within a physical system.

If a similar sequence of stages in the historical course of technology could be proven – speaking in ideal-typical terms, of course – then Cassirer would already have come a good deal closer to the attempt to classify technology in the system of symbolic forms. And indeed, just the history of the philosophy of technology provides indications of such a development. In his *Grundlinien einer Philosophie der Technik* (1877), the founder of modern philosophy of technology, Ernst Kapp, endeavored to prove that technology is essentially organ projection. “By organ-projection, he understands the fact that an individual limb of the human body does not simply work outward but creates in the external existence, so to speak, an image of itself” (Cassirer, FT, p. 300/ECW 17, p. 167). With this kind of self-image production is meant a self-enlightenment about the preconditions of this ability. For example, according to Kapp, man creates technical devices such as the microscope modeled on the eye. Ultimately, however, the physiological theories that emerge with the development of such devices again become the prerequisite for recognizing the functional mechanisms of the eye scientifically. In this way it can be understood that devices such as the hammer, the shovel, or the knife are projections of the hand, indeed that all basal tools are images of the body. “[E]very hand tool appears in this sense as a further positing and re-formation, as an exteriorization, of the hand itself” (Cassirer, FT, p. 300/ECW 17, p. 167). According to Cassirer, and we concur here, the thesis of organ projection admittedly has its limitations. It already seems less plausible when we ask ourselves to which part of our organism we wanted to attribute a highly technical device like the smartphone. The pointing gesture as a natural counterpart to the functioning of the smartphone would be obvious at first glance, but a second thought quickly shows that this intuition does not hold, because a smartphone can realize infinitely more functions than would be described by the pointing

gesture and the wanting to grasp. However, Kapp's thesis for the mimetic phase of technology is valid. Whether it extends as far as Kapp thought it did when he described the cables of the telegraph network as a projection of the human nervous system¹², remains to be seen. How far Kapp's theory carries is unimportant for the present context, because at least it marks the beginning of a development which we want to understand as a symbolic one. Cassirer also immediately succeeds in designating the end point of the development when he refers to Karl Marx's "law of the 'emancipation of the organic barrier'" (Cassirer, FT, p. 302/ECW 17, p. 169). We recall that the symbolic phase of any symbolic form is characterized by relationality and regularity: meaning is derived from pure lawfulness. In Marx's theory, then, we already find the anticipation¹³ of what the production of modern machines amounts to: the detachment from everything organic to the overcoming of everything organic. Only a few years lie between these theorizations of Kapp and Marx. The temporal proximity between the beginning and the end of the self-knowledge of technology cannot hide the fact that the real-historical development of technology is decisive for its ideal development. Cassirer refers here to the philosophy of Franz Reuleaux and again to that of Dessauer and Zschimmer in order to emphasize the central position of the development of modern mechanical engineering and to locate the analogical phase of technology exactly here:

As to the basic principle that rules over the entire development of modern mechanical engineering, it has been pointed out that the general situation of machines is such that they no longer seek to imitate the work of the

¹² "The nerves *are* cable devices of the animal body, the telegraph cables *are* nerves of mankind" (Kapp, 2015, p. 133, my translation).

¹³ Rohbeck (1997, p. 206) seems to overlook this point when he writes that Cassirer's reference to Marx is not without some irony against the background of the alienation thesis.

hand or nature but instead seek to carry out tasks with their own authentic means, which are often completely different from natural means (Cassirer, FT, p. 302/ECW 17, P. 169).

The machines created in this way – think, for example, of the printing press or the sewing machine – change the mode of working itself and thus the mode of action of technology, whereby the latter undergoes an inner transformation of its meaning. The further technology moves away from its original, mimetic mode of design, the better it fulfills its actual function. This becomes particularly clear with the ‘flight problem’, a long cherished and late realized wish of humans. It could be solved only “once technological thought freed itself from the model of bird flight and abandoned the principle of the moving wing” (Cassirer, FT, p. 302/ECW 17, p. 169). The theory of technology and its actual development thus give good reason to believe that Cassirer’s three-phase model bears and that technology can be integrated into the system of symbolic forms, namely, insofar as “the march of technology is mastered by a universal norm that rules the whole of cultural development” (Cassirer, FT, p. 303/ECW 17, p. 170). Cassirer has thus demonstrated what he wished to show at the outset, namely, that every appearance of a new symbolic form is subject to a logic that begins with the sensuous and successively frees itself from it. Therein lies its contribution to culture, in whose ideal progress Cassirer sees the self-liberation of the human being, which lies in the fact that it recognizes the autonomy in its ways of shaping the world and thus ultimately its own freedom. However, due to the fact that the addition of technology to the other symbolic forms as well as the entry into its analogical phase occur relatively late, the related effects are sometimes perceived drastically. No new symbolic form can be integrated into the structure of human culture “without struggle and the sharpest opposition” (Cassirer, FT, p. 303/ECW 17, p. 170). Cassirer’s philosophy of technology thus ultimately also provided a

systematic and not only a psychological reason explaining why the complaint of the so-called Philosophy of Life was so *en vogue* in his time. Its representatives are admittedly right in their diagnosis that technology appears with a sudden dominance never known in comparison to the other symbolic forms and thus frightens the individual. Cassirer merely cautions us to consider that technology draws on the same basic intellectual presuppositions as all other forms of culture, and that a one-sided condemnation of technology would create a performative self-contradiction. He therefore concludes:

The movement of the I breaks upon its own creations; the greater the scope and stronger the power of this creation become, the more its original tide of life subsides. This tragic impact of all cultural development is, perhaps, no more evident than in the development of modern technology. Those who turn away from it on the basis of this state of affairs forget, however, that, in their damning judgment of technology, they must logically include the whole of spiritual culture. Technology has not created this consistent existence; rather, it merely places an especially remarkable example urgently before us. It is, if one speaks here of suffering and sickness, not the ground of suffering but merely a manifestation, a symptom of it. What is crucial here is not an individual domain of culture but its function, not a particular way that it follows but the general direction it takes (Cassirer, FT, p. 305/ECW 17, p. 172).

Cassirer thus rejects the charge of the Philosophy of Life to the effect that it is not life but the spirit that is the judge of the value and unvalue of technology. Consequently, it cannot be a question of the unpleasantness of a life crushed by culture, for the value of technology can only be judged according to whether, following the principle of mental autonomy, it leads to freedom or not much more to unfreedom. Cassirer thus does not simply represent an apology of technology in demarcation from the Philosophy of Life, but reserves a critique of technology. Only the standard, which the critique is to follow, became a different one by the way of looking at technology as a symbolic form.

The need for an ‘ethicization’ of technology

The problem of technology, and herewith Cassirer prepares its critique, stems from (1) a fundamental tendency of all symbolic forms. Every new form appears with a claim to absoluteness. The autonomy of a symbolic form has a bearing on the freedom of other forms. For: “It not only insists on its own norm, but threatens to posit this norm as an absolute and to force it upon the other domains” (Cassirer, FT, p. 306/ECW 17, p. 173). In constructing its own norm, then, technology enters into conflict with other symbolic forms. This conflict can be sharper or less problematic. For example, technology also combines well with symbolic forms such as art and science, as can be seen impressively in the already mentioned Eiffel Tower. Other masterpieces of architecture, such as the Guggenheim Museum Bilbao or the ancient pyramids of Giza, in which technology, religion and myth are harmoniously combined, also demonstrate impressive syntheses. With (2) a view to the symbolic function of pure meaning¹⁴ guiding technology, further potential for

¹⁴ As already mentioned above in section two on technology as a symbolic form, Cassirer determines the symbolic forms not only by three phases, according to which each one develops, but also by a tripartism of the symbolic function itself (expression, presentation, pure signification). The philosophy of symbolic forms treats the three forms language, myth, science because in them exactly one of the mentioned three symbolic functions functionally clearly predominates (in myth the expressive function, in language the presentational function and in science the function of pure meaning). Other symbolic forms are characterized among other things by the fact that several symbolic functions are equally effective in them. This does not mean, however, that e.g., in language the expressive and the pure signification function are not also effective. In contrast to art, however, the presentational function *dominates in* language, while in art such dominance manifests itself in the oscillation between expression and meaning: “In an absolutely unique way that is reserved for it alone, the work of art permits ‘figure’ and ‘expression’ to merge into one another” (Cassirer, FT, p. 311/ECW 17, p. 178). Cf. Endres (2021) for further details.

conflict is revealed in the fact that the strong tendency towards objectification is at the expense of the expressive function. The objection raised by Simmel or Klages against the oppression of the subject by the objective figures of the mind cannot be completely dismissed in the case of technology, for it is deeply inherent in technology and also in science to erase the ego pole carried by the experience of expression. It is true that neither technology nor science succeeds in this in an absolute sense¹⁵, but a tendency and a will to erase the subject is manifest in them. Technology and science¹⁶ therefore sacrifice in their own realization a part of culture, which Cassirer, however, tries to turn positively by pointing out that it is the human – and not, for instance, the superiority of technology itself – who makes this sacrifice and thus proves his freedom and ‘humanity’. Cassirer literally:

Technology combined with theoretical knowledge, to which it is closely related, increasingly renounces all that is measured by expression in order to lift itself up into the strictly ‘objective’ sphere of pure signification. At the same time, it is indisputable that the gain achieved here contains a sacrifice. However, even this sacrifice and renunciation, this possibility to cross over and rise up into a pure world of things, shows itself to be a specific human power, an independent and indispensable descriptor of ‘humanity’ (Cassirer, FT, p. 313/ECW 17, p. 180).

¹⁵ One can show this transcendently in two ways, once by pointing to the constitutivity, objectivity, and skepticism-resistance of the expressive function in the structural construction of cognition (Cassirer, PSF 3/ECW 13, pp. 68 f., 74, 69, 89; Cassirer, ECN 4, p. 189; Cassirer, ECN 5, pp. 107 ff.) and once by pointing to the failure of the naturalization attempts of the mind and the untransferability of the first-person perspective into a third-person perspective (Kreis, 2010, pp. 11-21).

¹⁶ About their kinship, however, the specific difference between science and technology must not be forgotten: “Technology does not initially ask what *is* but what *can* be” (Cassirer, FT, p. 309/ECW 17, p. 176).

The word ‘humanity’ is not in quotation marks at the end of this paragraph for nothing. Otherwise, Cassirer would have to be accused of naivety with regard to the destructive consequences of technology.¹⁷ With regard to ethics, however, Cassirer comes to the clear conclusion that “modern technology . . . and the economy it has created and maintains with its own means” (Cassirer, FT, p. 314/ECW 17, p. 181) causes problems such as the mechanization of labor, the exploitation of those who produce, overproduction, mass consumption, waste of resources, and a creation of new needs that increases into infinity. In the face of this statement, Cassirer again tries to take a step back, to turn his eyes away from the *consequences of technology* and to ask whether “these effects can necessarily be attributed to its essence, that is, whether they are implicit in the configuring principle of technology, and whether they are demanded by it” (Cassirer, FT, p. 314/ECW 17, p. 182). Cassirer answers this question in the negative and, following Karl Marx and especially Walther Rathenau, argues that these consequences of technology “are to be understood not so much in themselves as in terms of their connection with a certain form and order of an economic system and

¹⁷ Regardless of the quotation marks, one could speculate that Theodor W. Adorno had just read *Form and Technology* when he wrote to Max Horkheimer: “Here, too, there was full agreement between Pollock and me, and likewise with regard to Mr. Cassirer, whom I consider to be completely idiotic” (Adorno, letter to Horkheimer from Oxford, May 13, 1935, my translation). Admittedly, in 1935 Adorno could not refer to Cassirer’s posthumous late work *The Myth of the State* (1946), in which the latter examines the technical fabrication of political myths that led to the industrial extermination of the European Jews (Cassirer, ECW 25, pp. 273-291), nor to the 1944 *Dialectic of Enlightenment* written jointly with Horkheimer. It is striking, however, that the *Dialectic of Enlightenment’s* guiding thesis of the unity of instrumental and practical reason stands in blatant contradiction to Cassirer’s ‘humanity thesis’ on instrumental rationality, technology, and science. (see also Bevc, 2005). On the question of whether and how Cassirer’s view of technology changes in *The Myth of the State* cf. Krois (1982, pp. 210, 215-220).

that every attempt at improvement must begin here” (Cassirer, FT, p. 315/ECW 17, p. 182). But now Cassirer had shortly before asserted that capitalism emerges from technology, that there is by all means an intimate relationship between technical progress and the generation and financing of its own resources. He wants the connection between technology and economy to be understood as “made necessary and thrust upon one by a particular situation, by concrete historical circumstances” (Cassirer, FT, p. 315/ECW 17, p. 182), but not as the ‘birth of capitalism from the spirit of technology’. The ‘spirit’ or meaning of technology, we recall, consists, like that of any other symbolic form, in its autonomy and freedom. In a sense, it is not ‘in the nature’ of a symbolic form to subordinate itself to another, in this case: the economy, or to enter into an unholy alliance. It must be remembered at this point that Cassirer does not advocate a teleology of freedom, although he does argue that the meaning of culture consists in the liberation of human beings from the sensual barriers of their nature, that is, in a kind of self-liberation (Freudenthal, 2004, p. 206f.). “Human culture taken as a whole may be described as the process of man’s progressive self-liberation” (Cassirer, ECW 23, p. 244). Thus, the human being is not immune to civilizational setbacks precisely in the formation of new symbolic forms; they provide no guarantee of said self-liberation. Against this background, it is coherent that Cassirer locates the reasons for the fatal conflation of technology and (capitalist) economy in the arbitrariness of history, but not in the ‘spirit’ of technology. But what should be done now that the real danger emanating from this constellation has been recognized? What would a ‘humane’ action look like in the face of the reality of technology. Cassirer is clear that the answer to this cannot be found simply in technology or only in a further scientification and objectification of the life-world:

It is not enough to appeal to the forces of nature or to the forces of mere

understanding by technological and scientific intellects; rather, here it suffices to indicate the point at which only the deployment of a new willpower can create change. In this construction of the realms of will and the basic convictions upon which all moral community rests, technology can only ever be a servant, never a leader. It cannot by itself determine the goal, although it can and should collaborate in carrying it out. It best understands its own meaning and its own telos when it is content with the fact that it can never be an end in itself. Rather, it has to fit itself into another ‘realm of purpose,’ into a genuine and final teleology that Kant described as ethico-teleological. In this sense, the ‘dematerialization’ of technology, rendering it ethical (*Ethisierung*, my addition), forms one of the central problems of contemporary culture (Cassirer, FT, p. 315/ECW 17, p. 182).

It is true that technology and science themselves appear with a certain normative force of the factual (Jellinek, 1921, p. 339) and transfer their own standards with the ‘factual world’ formed by them to all other areas of culture. In view of the whole of culture and especially against the background of the ethical question raised, however, it finally crystallizes that technology and science are normatively inexpedient.¹⁸ As the ‘logical conscience of culture’ it therefore remains the task of philosophy to argue for the ‘dematerialization’ and ‘ethicization’ of technology. The practical implementation of such an epochal challenge as the ‘liberation of technology from the economy and of the economy from technology’ cannot, however, be carried out exclusively on the part of philosophy, but requires the pooling of all normative resources of culture – first and foremost of law.

¹⁸ Of course, this is not yet a proposal for a solution, but first of all a self-knowledge of the human with regard to its position to technology. It is therefore surprising when Rohbeck (1997, p. 197) speaks of a peculiarly pale and unoriginal solution that is hardly convincing.

References

- Adorno, T. W. (1935). Brief an Horkheimer vom 13.05.1935 aus Oxford. In M. Horkheimer, *Gesammelte Schriften. Vol. 15 / Briefe 1913-1939* (pp. 345-351). Fischer.
- Bevc, T. (2005). *Kulturgenese als Dialektik von Mythos und Vernunft: Ernst Cassirer und die Kritische Theorie*. Königshausen & Neumann.
- Cassirer, E. (1998–2007). *Gesammelte Werke. Hamburger Ausgabe / ECW 1–26*. Meiner, F.
- Cassirer, E. (1995–2022). *Nachgelassene Manuskripte und Texte / ECN 1–18*. Meiner, F.
- Cassirer, E. (1923/2001). *Philosophie der symbolischen Formen. Vol. 1: Die Sprache / ECW 11*. Meiner, F. Translated into the English by Steve G. Lofts. (2021). Routledge.
- Cassirer, E. (1925/2002). *Philosophie der symbolischen Formen. Vol. 2: Das mythische Denken / ECW 12*. Meiner, F. Translated into the English by Steve G. Lofts. (2021). Routledge.
- Cassirer, E. (1929/2002b). *Philosophie der symbolischen Formen. Vol. 3: Phänomenologie der Erkenntnis / ECW 13*. Meiner, F. Translated into the English by Steve G. Lofts. (2021). Routledge.
- Cassirer, E. (1923/2003). Der Begriff der symbolischen Form im Aufbau der Geisteswissenschaften. In *ECW 16* (pp. 75-104). Translated into the English by Steve G. Lofts. (2013). In *The Warburg Years (1919–1933). Essays on Language, Art, Myth, and Technology* (pp. 72-100). Yale University Press.
- Cassirer, E. (1925/2003b). Sprache und Mythos. Ein Beitrag zum Problem der Götternamen. In *ECW 16* (pp. 227-311). Translated into the English by Steve G. Lofts. (2013). In *The Warburg Years (1919–1933). Essays on Language, Art, Myth, and Technology* (pp. 130-213). Yale University Press.
- Cassirer, E. (1930/2004). Form und Technik. In *ECW 17* (pp. 139-

- 183). Translated into the English by Steve G. Lofts. (2013). In *The Warburg Years (1919–1933). Essays on Language, Art, Myth, and Technology* (pp. 272-316). Yale University Press.
- Cassirer, E. (1930b/2004b). ‘Geist’ und ‘Leben’ in der Philosophie der Gegenwart. In *ECW 17* (pp. 185-205). Translated into the English by Robert W. Bretall & Paul Arthur Schilpp. (1949). In P. A. Schilpp (Ed.), *The Philosophy of Ernst Cassirer. The Library of Living Philosophers. Vol 6* (pp. 855-880). George Banta Publishing.
- Cassirer, E. (1944/2006). *An Essay on Man. An Introduction to a Philosophy of Human Culture / ECW 23*. Meiner, F.
- Cassirer, E. (1946/2007). *The Myth of the State / ECW 25*. Meiner, F.
- Cassirer, E. (2011). *Über Symbolische Prägnanz, Ausdrucksphänomen und ‘Wiener Kreis’ / ECN 4*. Meiner, F.
- Cassirer, E. (2004c). *Kulturphilosophie. Vorlesungen und Vorträge 1929-1941 / ECN 5*. Meiner, F.
- Endres, T. (2020). *Ernst Cassirers Phänomenologie der Wahrnehmung*. Meiner, F.
- Endres, T. (2021): Phenomenological Idealism as Method: The Hidden Completeness of Cassirer’s Matrix of the Symbolic. In L. Filieri & A. Pollok (Eds.), *The Method of Culture: Ernst Cassirer’s Philosophy of Symbolic Forms* (pp. 121-147). Edizioni ETS
- Favuzzi, P. (2017). Die Technik als Problem der Kulturphilosophie. In Puls, H., Schramm, M. W., & Waller, S. (Eds.), *Kultur, Technik, Freiheit: Festschrift zum 60. Geburtstag von Birgit Recki* (pp. 11-39). Brill Mentis.
- Freudenthal, G. (2004). The Missing Core of Cassirer’s Philosophy: Homo Faber in Thin Air. In C. Hamling & J. M. Krois (Eds.), *Symbolic Forms and Cultural Studies. Ernst Cassirer’s Theory of Culture* (pp. 203-226). Yale University Press.
- Hogrebe, W. (2014). *Echo des Nichtwissens*. Walter de Gruyter GmbH & Co KG.

- Jellinek, G. (1921). *Allgemeine Staatslehre*. Springer-Verlag.
- Kant, I. (1783/2004): *Prolegomena to any Future Metaphysics*. Cambridge University Press.
- Kapp, E. (2015). *Grundlinien einer Philosophie der Technik: Zur Entstehungsgeschichte der Kultur aus neuen Gesichtspunkten*. Meiner.
- Klages, L. (1929). *Der Geist als Widersacher der Seele* (Vols. 1–3). Johann Ambrosius Barth.
- Kreis, G. (2010). *Cassirer und die Formen des Geistes*. Suhrkamp.
- Krois, J. M. (1982). Ernst Cassirer's Theory of Technology and its Import for Social Philosophy. *Research in Philosophy and Technology*, 5, 209-222.
- Luft, S. (2015). *The Space of Culture: Towards a Neo-Kantian Philosophy of Culture (Cohen, Natorp, and Cassirer)*. Oxford University Press.
- Moss, G. S. (2015). *Ernst Cassirer and the Autonomy of Language*. Lexington.
- Rohbeck, J. (1997). Technik als symbolische Form bei Cassirer. In P. A. Schmid & S. Zurbuchen (Eds.), *Grenzen der kritischen Vernunft. Helmut Holzhey zum 60. Geburtstag* (pp. 197-212). Schwabe & Co AG.
- Rudolph, E. (1992). Sprache zwischen Mythos und Erkenntnis. Zu Cassirers Diagnose der Tragik sprachlichen Fortschritts. In E. Rudolph & H. Wismann (Eds.), *Sagen, was die Zeit ist. Analysen zur Zeitlichkeit der Sprache* (pp. 79-92). J. B. Metzler.
- Rudolph, E. (2003). *Ernst Cassirer im Kontext. Kulturphilosophie zwischen Metaphysik und Historismus*. Mohr Siebeck.
- Simmel, G. (1911). Der Begriff und die Tragödie der Kultur. In G. Simmel (1996), *Gesamtausgabe, Vol. 14* (pp. 385-416). Suhrkamp.

**THE EVOLUTION OF MEDIA THROUGH THE THEORETICAL
OPTICS BY M. McLUHAN¹**

***LA EVOLUCIÓN DE LOS MEDIOS A TRAVÉS DE LA ÓPTICA TEÓRICA DE
M. McLUHAN***

REGINA PENNER
South Ural State University, Rusia
penner.r.v@gmail.com

RECIBIDO: 15/08/2023

ACEPTADO: 25/09/2023

Abstract: Technology has come a long way to its contemporary state, from primitive tools to complex technical systems. Today, the Internet is a powerful media technology channel that, in terms of audience coverage and impact, can be compared with traditional media, newspaper, radio and television. The aim of the article is to identify the features of the development of media and media technologies from the standpoint of an evolutionary approach, including through the prism of M. McLuhan's media theory.

M. McLuhan is one of the first theorists who reflected on the nature of electronic media. In the discourse of M. McLuhan of particular importance is the position according to which the media are defined as a medium, i.e. means or channel of communication. In the historical evolution of a person, media technologies move from auditory to visual culture and vice versa. Media technologies, according to M. McLuhan, become something that unites people, breaks the boundaries in the classical concepts of space and time. Media technologies have a complex impact on a person as a subject of social practices; through a person, media technologies reshape the social fabric, making changes in temporal practices, practices of communication, memory and power.

Keywords: media, digital, digital media, media technologies, Internet, digital anthropology, Media Studies, M. McLuhan.

¹ The work was supported by the grant of the Presidency of the Russian Federation for young scientists-PhDs MK-2592.2022.2 "Digital Anthropology: Theoretical and Applied Aspects".

Resumen: La tecnología ha recorrido un largo camino hasta su estado contemporáneo, desde herramientas primitivas hasta sistemas técnicos complejos. Hoy en día, Internet es un poderoso canal de tecnología de medios que, en términos de cobertura de audiencia e impacto, puede compararse con los medios tradicionales, los periódicos, la radio y la televisión. El objetivo del artículo es identificar las características del desarrollo de los medios y las tecnologías de los medios desde el punto de vista de un enfoque evolutivo, incluso a través del prisma de la teoría de los medios de M. McLuhan.

M. McLuhan es uno de los primeros teóricos que reflexionó sobre la naturaleza de los medios electrónicos. En el discurso de M. McLuhan tiene especial importancia la posición según la cual los medios se definen como medio, es decir, medio o canal de comunicación. En la evolución histórica de una persona, las tecnologías de los medios pasan de la cultura auditiva a la visual y viceversa. Las tecnologías de los medios, según M. McLuhan, se convierten en algo que une a las personas, rompe las fronteras en los conceptos clásicos de espacio y tiempo. Las tecnologías de los medios tienen un impacto complejo en una persona como sujeto de prácticas sociales; a través de una persona, las tecnologías de los medios remodelan el tejido social, modificando las prácticas temporales, las prácticas de comunicación, la memoria y el poder.

Palabras clave: medios, digital, medios digitales, tecnologías de los medios, Internet, antropología digital, estudios de medios, M. McLuhan.

Introduction

B. Gates is credited with the phrase according to which the possession of the world correlates with the possession of information. Despite the fact that the authorship of the proposition is questionable, its content does not actually raise any questions. The accent on the significance of information was already crystallized at the end of the 18th century, with the invention of the optical telegraph by C. Chappe (Figuier, 1868). With the first demonstration of the electromagnetic telegraph by P.L. Schilling in 1832 and obtaining a patent for the instrumental model by S. Morse in 1840, we are witnessing the evolution of information transmission technologies exponentially (Hurdeman, 2003). At the same time,

communication, which in the broadest sense can be read as the exchange of information, is not the prerogative of the High-Tech Age. Moreover, it is not limited to human interaction.

By and large, the whole world is in a situation of information exchange. Streams of information circulate between living beings, one of which is a human being. A feature of the information exchange initiated by a person is its (exchange) instrumental nature. A person, whose existence is taking place in the supernatural, technical world, is increasingly moving away from the so-called natural, intuitive information flows that can be read directly, without the need for special technical vocabulary or skills of its understanding. A person of the 21st century does not refuse natural information signals, whether it be colors, smells or tactile sensations. However, he increasingly relies on those information flows, the reading of which requires special technical knowledge. Probably, the key of this technical knowledge is speech, oral and written. To a certain extent, oral speech, as well as communication, is not limited by the anthropological framework. Mammals, birds and insects exchange sound signals. Written speech has become a unique invention of the mankind. This is a kind of cultural code that seems to have broken the boundaries of *hic et nunc*. With the development of writing, a person was able to transmit messages, overcoming the boundaries of space (to another place) and time (up to a different historical period). This is how rock art came down to us, depicting a mammoth and a saber-toothed tiger, which, probably, initially did not fulfill the function of transmitting information about the life and worldview of a caveman through the centuries.

The drawings were replaced by hieroglyphics; hieroglyphs were replaced by alphabetic writing. Since ancient times, writing and books have been central to the definition of literacy. For example, J.L. Borges interprets the world itself as a library, which is divided into many hexes, hexagonal galleries where books are stored; in

books, 25 symbols represent the whole world; a person is born in one of the hexes, moves through the hexes all his life and gets acquainted with the content of unique books (Borges et al., 2000). In a conversation with J.-C. Carrière, U. Eco formulated a militant thesis, “Do not hope to get rid of books!” (Eco & Carrière, 2011). The argument of the Italian thinker is based on the fact that since the end of the 20th century we are witnessing a constant change of the so-called high-tech storage media, from cassettes, floppy disks, disks and flash drives that carried bits, megabytes and gigabytes of information, respectively. In the context of the transformation of information and communication technologies in modern households, there may simply not be a device on which it is possible to read information from a cassette or flash drive; the book, in turn, is always open to reading. At the same time, the fact remains that in the famous novel by U. Eco, the insane blind monk Jorge destroys the great library (Eco et al., 1984); the universal source of information perishes in the flames of fire.

Historical aspect in the development of technology

Whatever is the universal mediator of information in the 21st century, it seems that reflections on the evolution of media technologies should be developed in a broader framework, in the context of the history of technology. Here the 19th century can be seen as a turning point. Without going into the details of the historical development of modern technologies, attention can be focused on two breakthrough technologies, telegraph and railroad. Both technologies have transformed linear optics in the perception of time and space; telegraph through instant messaging, railroad through instant movement.

Technological transformations in the second half of the 19th century began to attract the interest of not only engineers as

specialists in the field of technology. The technology became open to the masses and available for their use. Transformations of the technological, therefore, went beyond the framework of technology itself; they were woven into social and cultural practices, and became an organic unit of everyday practices. This has not passed by thinkers.

At the same time, technology came to the attention of philosophers long before the 19th century, back in the ancient era. In Plato's dialogues, it is included in the discourse of the shadows, matter, as an illustration of one of the objects that drive prisoners into the cave and do not let them out of it (Plato, 1997). Aristotle had directly addressed the problem of τέχνη, which he read as an art. In "Metaphysics" Aristotle distinguished two types of knowledge: knowledge "how" – a specific skill, – and knowledge of "what" is – theoretical knowledge (Aristotle & Lawson-Tancred, 1998). Practical and theoretical knowledge, according to Aristotle, go hand in hand with each other, one is impossible without the other. Nevertheless, the very knowledge of what is reveals to a person εἶδος of the thing, the pure idea of the thing and of the whole natural world. However, it was only in the 19th century when thinkers raised the question of technology, which was being improved, including through the discovery of what is and how to apply electromagnetic radiation. Among those thinkers was E. Kapp.

E. Kapp is on a par with major engineers who, in their technical developments, approached the philosophical questioning of technology (F. Reuleaux, I. Beckmann, A. Riedler). He introduced the concept "Philosophie der Technik" into scientific circulation, and proposed the concept of technology, the principle of organ projection (Kapp, 1978). If Aristotle saw the eidetic principle in the material in technology (which M. Heidegger would later call hidden (Heidegger, 1977)), E. Kapp set up an anthropological prism in understanding of technology. Thus, the telegraph becomes an

analogue to the nervous system, and the railway grid becomes a circulatory one.

In contemporary philosophical discourses, technology is increasingly read not only as a simple continuation of the body, but as something that totally organizes the existential practices of a person. J. Wajcman puts it this way: “According to Google’s Eric Schmidt and Jared Cohen, in the ‘New Digital Age’ you will wake up to the aroma of freshly brewed coffee in a room where temperature, humidity, music and lighting will automatically adjust, and your high-tech bed will give you a light back massage, and it is also guaranteed to provide you with sound sleep, adjusting to your phases of REM sleep” (Wajcman, 2015, p. 380). Smart houses and smart cities are the future that is happening right in front of us. Artificial intelligence seems to be following us on the streets of the city. Law enforcement agencies are increasingly turning to this option in search of criminals by entering key data into the network (Doohan et al., 2022). Or AI helps doctors in diagnosing diseases at an early stages (Rajpurkar et al., 2022). Transforming everyday human practices, technology seems to have settled in the nature and structure of the social fabric. Novelists seem to express this most poignantly. Therefore, it is interesting to turn to two classic dystopias, “Nineteen Eighty-Four” by G. Orwell (Orwell, 2021) and “Brave New World” by A. Huxley (Huxley, 2006). F. Fukuyama called both novels “visionary”. According to his position, those two systems that came out from the pen of English writers, on the one hand, “gave birth to two different technologies that were to emerge and determine the world for the life of the next two generations” (Fukuyama, 2003, p. 23), on the other hand, they posed a threat to democracy throughout the 20th century.

On the pages of “Nineteen Eighty-Four” we see the total devotion of the people to the party in the face of Big Brother. Faith and devotion to party ideals is realized on the basis of a well-established

system of denunciations and informing. The problem is that a person turns out to be a key informer on himself. This happens for the simple reason that most of the rooms in Oceania are equipped with a telescreen, full-wall screen. The purpose of the screen is not only to broadcast information from the outside, but also to collect the data what citizens are doing, what they are talking about, how they shout at the “Two Minutes Hate”. The power in the country is provided and supported by a tandem of the Ministry of Truth and the Ministry of Love, which constantly draw information from the television network stretched everywhere.

On the pages of “Brave New World” we can see another interesting technology named "bokanovskization". By and large, a person whom we observe and who we ourselves are in the 21st century does not exist in this new world. There are savages, but they are not included in the social hierarchy. A civilized society, in turn, is a high-tech version of the caste system. People taken out of a test tube are included in one of the higher (alpha, beta, gamma) or lower (delta, epsilon) castes. The key benefit of the new world is pleasure, stimulated technically, narcotic. If the ideology of “Nineteen Eighty-Four” can be defined as a love for Big Brother, then “Brave New World” broadcasts the ideological message, to buy and to enjoy.

In January 1984, the traditional broadcast of the final game of American football took place on CBS. Every year, global corporations fought for the right to promote information about their products and services in a single commercial during the Super Bowl broadcast. In 1984, Apple Corp. bought this right. Contrary to the traditional presentation scheme, the advertised product was never shown. Instead, a sixty-second drama unfolded on the screens of millions of Americans: masses of shaven-headed people in identical robes mindlessly listen to speeches from a huge screen; the head from the screen congratulates fellow citizens on the construction of the “garden of pure ideology”. Then a beautiful girl in bright orange

shorts runs into the hall with a hammer in her hand; the girl throws a hammer and breaks the screen, which a second ago broadcasted: “We will win!”. The video ends with a pathetic line: “On January 24th, Apple Computer will introduce Macintosh. And you'll see why 1984 won't be like "1984"” (Christoph, 2011). The world of total political ideology was destroyed by the new model of the personal computer, “Macintosh 128K”.

Continuing the idea of the “survivability” of the great dystopias in the modern world, F. Fukuyama seems to implicitly bury “Nineteen Eighty-Four” with the entry of mankind into the 21st century, the availability of personal computers and the development of the Internet. His concerns center on “Brave New World”, more specifically on the biotech content of the novel. The American philosopher even resorts to the words of the bioethicist L. Kass: “Unlike people crippled by disease or slavery, those dehumanized in the Brave New World type are not miserable; they do not know that they are dehumanized, and even worse – if they knew, they would not care. They are happy slaves with slavish happiness” (Fukuyama, 2003, p 37).

Technology and, in particular, media technologies transform the social world of a person. In changing the external, a kind of reshaping of space and time, technology has penetrated into the daily life of a person. In part, this prompted M. Heidegger in 1955 to utter a hymn to “Gelassenheit” (Serenity) of a person, detachment from things, from technology (Davis, 2009). In his speech, M. Heidegger expressed concern about the actual place of technology in the social as a whole. He referred to the time in Europe after World War II as the “Atomic Age”. At the same time, not only the dangers that the so-called “peaceful atom”, worried the philosopher. The key motive of his care is the technology that a person turns to in his daily practices; more precisely, total dependence, which follows the constant appeal of a person to technology. After all, a person from

the 20th century had learned to say “yes” to the technology; much less often is he able to say “no” to technology, which is necessary in order to be able to preserve, according to the German philosopher, the human nature.

However, our daily practices are not made up of biotechnologies or abstract technology per se. Since the onset of the COVID-19 pandemic, the social practices of people around the world have seemed to be anchored in a space between online and offline, direct and indirect communication. Media technologies have come to the forefront of attention of ordinary users and social researchers, which form the ability of a person to live in two worlds at the same time, the physical and the virtual.

Marshall McLuhan. Understanding Media

The question of what challenges media technologies create for human nature should be preceded by the question of what media is. M. McLuhan is among those researchers who were the first to pose a question about the nature and essential characteristics of media in the social and humanitarian framework. In this context, J. Miller wrote: “He (*M. McLuhan*) did what very few publicly did before. Namely: focused on the means by which we acquire knowledge. In the past, these aspects have been largely ignored, or at least the preserve of philosophers or neurophysiologists, and I think that McLuhan for the first time placed the nervous system right at the center of the discussion of ordinary communications and human knowledge in general” (Stearn, 1967, p 235).

The Canadian culturologist and philologist M. McLuhan is called the guru of the media era for a reason. He seemed to feel the end of the era of the so-called “typographic man” and the advent of the era of a new man, whom M. McLuhan had called “electronic” (Katz & Katz, 1998). The professor of English literature did not spend his

intuition only within the framework of his courses; he had written books that, in a certain sense, shaped the direction and style of studying media in a social context, i.e. “The Mechanical Bride: Folklore of Industrial Man” (1951) (McLuhan, 2001), “The Gutenberg Galaxy: The Making of Typographic Man” (1962) (McLuhan, 2008), “Understanding the Media: External Extensions of Man” (1964) (McLuhan, 1994).

Conflicting views on the role of M. McLuhan in the development of media studies coexist in the professional literature. Some media theorists assert the vital influence of the thinker on the development of media sciences; others deny the originality of M. McLuhan's theses, indicate the inconsistency of his reputation among other pioneer media theorists (Katz & Katz, 1998, p. 309). For example, Donald and Joan Theall seem to accuse M. McLuhan of “failing to successfully convey ideas of Contemporary Poetry and Art for Communication Researchers” (Theall & Theall, 2021, p. 46). According to them, the Canadian philologist, being a connoisseur of English literature, learned the basic provisions in his study of the media from there. They call J. Joyce one of the key sources of the thinker's theoretical inspiration; M. McLuhan's books in their interpretation were called “the road to Finnegans” (Theall & Theall, 2021, p. 46).

The peculiar language experiment “Finnegans Wake” first came off the printing press in 1939 (Joyce, 2012). The essence of the experiment is presented in the so-called “auditory reading”, i.e. appeal in print to the difference between oral and written texts in order to create a kind of poetic counterpoint aimed at expressing the diversity and at the same time the conflict of the outside world, which is permeated with new communication technologies (Theall & Theall, 2021, p. 56). Donald and Joan Theall conclude the need to revise the work of M. McLuhan in the context of the English poetic tradition, which can be read as a source of inspiration in the

theoretical understanding of media in general. According to their position, M. McLuhan seemed to hush up the influence of literature (more precisely, contemporary literature) on his work, so as not to attack conservative values, which, first of all, are represented by the church (Theall & Theall, 2021, p. 63). The media are indeed questioning the authority of the church, for the reason that they have revolutionized the way in which knowledge is received and transmitted.

M. McLuhan has built a vision of the history and culture of mankind through the prism of the media. In this prism, the media are presented as a driving force of cultural change. Sound signals were replaced by oral speech; during the formation of ancient civilizations, specific music was added to it, which is produced by percussion instruments like gong and tam-tam; in a complex evolution from cuneiform and hieroglyphics, the alphabet and written speech took shape; another evolutionary leap is associated with the invention of the printing press, which seems to have democratized literacy; finally, from the end of the 19th century we are witnessing the transformation of media exponentially: telegraph, radio, telephone, cinema, television, computer, smartphone. According to M. McLuhan, each of these mediums penetrates into social groups and gradually grows, expands in them, and later changes the societies where it prevails.

M. McLuhan defines technology as “extensions” of a person, a continuation of his limbs (strength, speed) and sensitivity (sight, touch, smell) (McLuhan, 1994). In this sense, any technical option turns into communication, in the literal sense – with other subjects, in a figurative sense – with the outside world. Often, in the discourse about the first media theorists, M. McLuhan is put on a par with H.A. Innis. Despite the fact that H.A. Innis, like M. McLuhan, was a Canadian researcher, economics was his initial scientific field. In one of his key works, “Empire and Communications”, he drew an

analogy between media technologies and transport (Innis, б. д.). According to his position, the media organizes society in the same way that a car organizes urban space. In turn, M. McLuhan calls this analogy into question. Media organizes society, but their influence, according to M. McLuhan, is indirect. Media directly affects not society as a whole, but individual recipients. M. McLuhan even builds a kind of causal relation: each new media affects the brain of the recipient of information; when processing the received information, the feelings of the recipient turn out to be influent; in the influence of new media – the subject as a whole, which, in response, changes social structures (McLuhan, 2008). Hence the controversial thesis that typography with strict rules in the organization of the text has designed the so-called “linear” thinking, which is primarily represented by the conveyor belt and railways.

However, the mediation of the impact of the media on society does not negate the strength and depth of this impact. This is reflected in M. McLuhan's “bestseller” phrase, “media is a message” (McLuhan, 1994). Probably, the Canadian thinker for the first time in social and humanitarian research practices changed the focus from the meaning of the message to the technology of its transmission. By and large, he argues that the media dictate to us not what to think, but how to think. This affects the social environment, the organization by the subject of his daily, work and other social practices. In a certain sense, this statement by M. McLuhan turned out to be prophetic: the phenomenon of self-isolation during the COVID-19 pandemic demonstrated that today the workflow is organized around a computer and a smartphone; the workplace is decentralized.

New technological media, according to M. McLuhan, by and large continue the confrontation between oral and written speech. In “The Gutenberg Galaxy”, the media theorist offers metaphors that highlight the contrast between them, the media of “heart” and

“mind” respectively (McLuhan, 2008). Following the Old Testament thought, he emphasizes that in the beginning there was an oral word as a continuation of the ear and a reflection of the heart. Oral communication still differs from its textual counterpart; it acts as a conductor of worldly, universal wisdom; it enriches / saturates with information each of the participants in communication. The printed word, in turn, is aimed, as a rule, exclusively at the one who reads it, it acts *hic et nunc*; that is why it is selfish in a certain context, expanding the horizons of its reader directly.

The specificity of obtaining information, according to M. McLuhan, further determines the worldview and behavior of the subjects of social practices. M. McLuhan turns to another analogy, this time borrowed from the sphere of art: cubism seems to create the image of an “acoustic” person, someone who feels and lives life multidimensionally and simultaneously; classical art, in turn, refers us to the “printing” or “typographic” person who receives information sequentially (since, according to M. McLuhan, the medium every time indirectly affects social practices, the consistent way of receiving and processing information can be compared with what M. Heidegger called “calculative thinking” (Davis, 2009)). Behind the spoken word in this logic stands metaphysics, religion; behind the written one – the development of nations and empires in all directions, including space.

Distinguishing media according to the type of mind and heart leads M. McLuhan to two important concepts. The first indicates the historical transformation of the media, which is associated with the movement of the worldview that has passed, in the A. Comte's discourse, for example, 3 stages, religious, philosophical and scientific (Comte, 2008). However, unlike A. Comte, who saw the apogee of human thought in science, M. McLuhan argues that in their technological development, the media return to the heart, the spoken word. This resonates with the second concept, within which

the Canadian thinker distinguishes the so-called “cold” and “hot” media (McLuhan, 1994). Here, media is read not only as an intermediary between sender and recipient of information, but as a stimulus that involves the audience in creating a complete picture. Due to the fact that cinema and television simultaneously affect several senses of the recipient, the viewer himself is left with a simple (i.e. less imaginative) work to connect all the dots on his screen. From here, television becomes, in M. McLuhan’s terminology, a “cold” medium; unlike the radio, which forces listeners to complete the picture on their own. The following illustration seems appropriate in the given context: in 1938, in the northeastern United States, a play was staged on the CBS radio station under the direction of O. Welles, “The War of the Worlds” based on the novel of the same name by H. Wells. The production turned out to be so good that more than a million Americans panicked, mistaking information about the Martian attack for a news report (Campbell, 2010).

M. McLuhan has opened the book “Understanding Media” with the following problem, “we suddenly discover in ourselves a passionate desire for things and people to manifest themselves in their entirety. In this new attitude one can find deep faith – faith in the highest harmony of all being. It is in this faith that this book is written. It explores the outlines of our extended beings in our technologies and looks for the principle of intelligibility in each of them” (McLuhan, 1994, p. 6). In other words, new media (for M. McLuhan, this is, first of all, radio and television) are certainly a continuation and expansion of a person. There is no doubt that individual subjects and social groups use such extensions for their own selfish purposes (as it seems, this is one of the key topics around which the plots of dystopias are still built). At the same time, the misunderstanding of media remains as a problem; media have not been understood neither by the person, nor by those groups that use

the media for manipulative purposes. For example, based on the results of the analysis of the US presidential elections in 1960, M. McLuhan came to the conclusion that effective manipulation is realized when the “right” medium is chosen, which corresponds to the content and context of the information (McLuhan, 1994).

According to R. Guins, in a study of how media “shape, expand and involve ‘human sensory’”, M. McLuhan stated that “the visual bias of Western culture poorly prepared the ‘literate person’ to understand the auditory tactility of new electronic media” (Guins, 2014). By and large, it took culture more than three thousand years to return from the limitations of literacy (written culture) to the so-called “resonant world of acoustic space”, which since the second half of the 19th century increasingly transforms the social environment and the whole world around human beings (Marchand, 1998). Material for the analysis of M. McLuhan from the second half of the 1950s became television. In his opinion, regardless of its content, television, which was dynamically spreading through North American households, marked a renaissance in classical (more precisely, traditional) acoustic culture (Katz & Katz, 1998). Starting with the books by W. Benjamin, in the intellectual tradition of the 20th century a certain “alarming” attitude of intellectuals towards technology was formed (Benjamin, 2008). Unlike the craft tradition of the Middle Ages, in which craftsmen created unique works, the conveyor “learned” to stamp works of art. In this phenomenon one can find positive features; for example, a contemporary person through replicating images of world culture (on postcards, in books and albums, including their online analogue) had the opportunity to “touch” the beautiful even in the space of his apartment. However, the benefit of a replicated culture does not negate the wariness that has crept into the discourse of thinkers. Probably, this anxiety was expressed more precisely by the Spanish philosopher, J. Ortega y Gasset, in the term “revolt of the masses” (Y Gasset,

2021). According to his position, the era of universal literacy and electronic media did not open up new horizons for humanity; it brought power to the masses, an irrational and uncontrollable force.

M. McLuhan in this discourse of the intelligentsia of the first half of the 20th century offers a different prism in the understanding of new media, which seems to inspire hope for the future. Hope is emerging in overcoming the critical stance on the technology media that has expanded with the spread of electrification. Electrification has overcome spatial and temporal differences; according to M. McLuhan, it also has the potential to overcome linguistic and cultural differentiations. Written language strengthened national state borders, reduced human sensory to a single channel for perceiving information – vision. The new electronic culture, according to M. McLuhan, in turn, offered mediums that connect different sense organs with a minimum decoding of the messages they transmit. The age of electricity in this context turns into a kind of technological utopia, in which national borders are overcome and everyone finds their place, which implies rights and responsibilities. This is what M. McLuhan called the “global village” (McLuhan, 2008). The village is not a rolling back of mankind; on the contrary, it is a new form of human existence as an organically built-in social (as a member of a group or groups), national (citizen of a country), natural (living being), planetary (inhabitant of planet Earth) and cosmic unit, which is open to a greater understanding of the surrounding world and taking responsibility for what happens in it. This collectivism of electronic culture is opposed to the egoism of typographic culture. According to M. McLuhan, a person in the era of new media, seems to refuse to realize himself at the expense of the collective; in the global technological village, he develops in consonance with the development of the environment.

Until recently, such a utopia sounded very plausible. In the 21st century, in the era of universal connectivity, the boundaries between

cultures, languages, and even countries seemed to be completely erased. A counterexample arose in 2020, when, due to the aggravation of the epidemiological situation, some countries began to refuse entry to Chinese citizens, and in March, international flights were canceled around the world. Thus began the COVID-19 pandemic, demonstrating that the idea of national boundaries is not a vestige of the past. At the same time, the pandemic also demonstrates that for public and private communications, in the age of high technologies, a subject with an Internet connection is enough. The thesis about overcoming the spatial framework through technology does not lose its relevance.

The key point that is read in the books by M. McLuhan can be reduced to the thesis that the alphabet is one of the historical forms of media; electronic media, in turn, received in the 20th century a cultural influence that is comparable to the invention and spread of the alphabet. In “Understanding Media”, M. McLuhan decomposes media as a trivium and a quadrium of posthumanist epistemology. Hence the close relationship between media and literature, M. McLuhan refers to literature as a foundation, where the distinction “between auditory and visual perceptual types” was formed (Katz & Katz, 1998). M. McLuhan forms his trivium from dialectics, rhetoric and grammar. The dialectic contains the potential of critical thought in the analysis of the visual; in turn, rhetoric (the power of persuasion) and grammar (the structure of sentences and prepositional constructions) underlie the comprehension of the auditory. Visual and auditory are represented in the culture of each era. However, their share varies in the historical dynamics of cultures. In the 1960s in dissonance with philosophers' reflections on new existential foundations, including the linguistic turn (Rorty, 1992), M. McLuhan concludes that, just like language, the media “shape the way we perceive the world” (McLuhan, 2013).

Reflections on the past, present and future of media M. McLuhan

holds, among other things, through the prism of a combination of technologies with the bios. This idea was first voiced in “The Mechanical Bride”, where a person was read as the “genital organs of the machine world” (McLuhan, 2001, p. 47). Such a “sexualized” view of electronic media seems to be a precursor of the concept that in the 1990s has proposed J. Baudrillard under the name of Telematic Man (Baudrillard, 2009). In the French post-structuralist tradition, Telematic Man is a synthesis of corporality with a screen. M. McLuhan, in turn, continues the sexual analogy with the concept of “hybrid energy that will be released by the combination of literary and electronic modalities, a connection that highlights the broadcast aspect of media, especially in their digital configuration” (Cavell, 2014). Therefore, electronic media are understood by M. McLuhan as an acoustic space characterized by sensory involvement. This space is “skewed, curved and bumpy”, as opposed to the “straight and homogeneous” print culture (McLuhan, 1994). This heterogeneous research requires non-verbal ways of understanding. Thus, M. McLuhan compares media research with the construction of a new architecture, which is characterized by contours rather than straight lines, just as a film rolls up linear frames to create a sense of an organic process (McLuhan, 1994).

Conclusion

Technology can be defined as a tool invented by human beings. It evolved together with mankind and to a certain extent set the vectors of anthropogenesis. In any case, contemporary technology forms the specifics and affects the quality of social practices that a person implements. Such a vision of technology is not accidental; from the second half of the 19th century in the philosophical discourse, the concept of organ projection is formed, according to which

technology is read as a continuation of a person (E. Kapp).

Canadian thinker M. McLuhan seems to have been developing the concept of his German predecessor, focusing on the media. According to the position of the media theorist, electronic media technologies can be understood as a medium, a means of communication. They went from the simplest sound signals to signs on various surfaces (from rock art to the modern alphabet) and, finally, to an electronic signal.

Despite the inconsistency of the figure of M. McLuhan among media theorists, the epistemological value is the view of electronic media, according to which the author offers a kind of “Renaissance”, turning prism, which, among other things, is used in the concept of the global village. The value of this view is due to the trend towards ever greater convergence of users of modern communication technologies, which today is represented, for example, by social networks.

References

- Aristotle, & Lawson-Tancred, H. (1998). *Metaphysics*. Penguin Books.
- Baudrillard, J. (2009). *The transparency of evil: Essays on extreme phenomena* (J. L. Benedict, Trans.; Paperback edition reprinted). Verso.
- Benjamin, W. (2008). *The Work of Art in the Age of Its Technological Reproducibility, and Other Writings on Media* (M. W. Jennings, B. Doherty, & T. Y. Levin, ed.; E. Jephcott, R. Livingstone, & H. Eiland, tr.). Harvard University Press. <https://doi.org/10.2307/j.ctv1nznfgns>
- Borges, J. L., Desmazières, E., Hurley, A., & Giral, A. (2000). *The library of Babel*. David R. Godine.

- Campbell, W. J. (2010). *Getting it wrong: Ten of the greatest misreported stories in American journalism*. University of California Press.
- Cavell, R. (2014). On the 50th Anniversary of *Understanding Media*. *Journal of Visual Culture*, 13(1), 33–35. <https://doi.org/10.1177/1470412913509592>
- Christoph. (2011, July 12). “1984” – *The famous Super Bowl Spot*. <https://www.mac-history.net/2011/07/12/1984-the-famous-super-bowl-spot/>
- Comte, A. (2008). *The positive philosophy of Auguste Comte. 1: Part I*. Kessinger.
- Davis, B. W. (2009). Will and *Gelassenheit*. B. W. Davis (ed.), *Martin Heidegger* (1st ed., pp. 168–182). Acumen Publishing Limited. <https://doi.org/10.1017/UPO9781844654475.013>
- Doohan, N. V., Kadam, S., Phursule, R., Wadne, V. S., & Junnarkar, A. (2022). Implementation of AI based Safety and Security System Integration for Smart City. *International Journal of Electrical and Electronics Research*, 10(3), 518–522. <https://doi.org/10.37391/ijeer.100319>
- Eco, U., & Carrière, J.-C. (2011). *This is Not the End of the Book*. Vintage Digital.
- Eco, U., Weaver, W., & Eco, U. (1984). *The name of the rose* (Picador ed). Pan Books.
- Figuiet, L. (1868). Le télégraphe aérien. B *Les merveilles de la science: T. T. 2* (pp. 20–68).
- Fukuyama, F. (2003). *Our posthuman future: Consequences of the biotechnological revolution* (1. Picador ed). Picador.
- Guins, R. (2014). The Present Went This-A-Way: Marshall McLuhan’s *Understanding Media: The Extensions of Man @ 50*. *Journal of Visual Culture*, 13(1), 3–12. <https://doi.org/10.1177/1470412913509466>

- Heidegger, M. (1977). *The question concerning technology, and other essays*. Garland Pub.
- Huurdeman, A. A. (2003). *The worldwide history of telecommunications*. J. Wiley.
- Huxley, A. (2006). *Brave New World*. Harper Perennial.
- Innis, H. A. *Empire and Communications* (Project Gutenberg Canada ebook #342). Retrieved on 2023, August 15, 2023, from <https://gutenberg.ca/ebooks/innis-empire/innis-empire-00-h.html>
- Joyce, J. (2012). *Finnegans Wake*. Wordsworth.
- Kapp, E. (1978). *Grundlinien einer Philosophie der Technik. Zur Entstehungsgeschichte der Cultur aus neuen Gesichtspunkten*. Stern.
- Katz, R., & Katz, E. (1998). McLuhan: Where did he come from, where did he disappear? *Canadian Journal of Communication*, 23 (3), 307–319.
- Marchand, P. (1998). *Marshall McLuhan: The medium and the messenger: a biography* (1st MIT Press ed). MIT Press.
- McLuhan, M. (1994). *Understanding media: The extensions of man* (1st MIT Press ed). MIT Press.
- McLuhan, M. (2001). *The mechanical bride: Folklore of industrial man*. Gingko Press.
- McLuhan, M. (2008). *The Gutenberg galaxy: The making of typographic man* (Repr.). Univ. of Toronto Pr.
- McLuhan, M. (2013). Laws of the Media. *ETC: A Review of General Semantics*, 70(4).
- Ortega y Gasset, J. (2021). *The Revolt of the Masses* (1st ed.). Routledge. <https://doi.org/10.4324/9781003188124>
- Orwell, G. (2021). *Nineteen Eighty-Four* (Penguin Classics).
- Plato. (1997). Allegory of the cave. B *Republic*. Wordsworth.
- Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, 28(1), 31–38. <https://doi.org/10.1038/s41591-021-01614-0>

- Rorty, R. (ed.). (1992). *The Linguistic turn: Essays in philosophical method*. University of Chicago Press.
- Stearn, G. E. (Ред.). (1967). *McLuhan: Hot & cool*. Dial Press.
- Theall, D., & Theall, J. (2021). Marshall McLuhan and James Joyce: Beyond Media. *Canadian Journal of Communication*, 14(4). <https://doi.org/10.22230/cjc.1989v14n4a531>
- Wajcman, J. (2015). *Pressed for time: The acceleration of life in digital capitalism*. The University of Chicago Press.

**EL SABER EN LA ÉPOCA DE SU REPRODUCTIBILIDAD TÉCNICA.
SOBRE LA RELACIÓN ENTRE LA INTELIGENCIA ARTIFICIAL Y LA
HUMANA**

***THE KNOWLEDGE IN THE AGE OF MECHANICAL REPRODUCTION.
ON THE RELATIONSHIP BETWEEN ARTIFICIAL AND HUMAN
INTELLIGENCE***

JOSÉ LUIS DELGADO ROJO

Istituto Italiano di Scienze Umane - Scuola Normale Superiore,
Italia
jluisdelgado@hotmail.com

RECIBIDO: 20/08/2023

ACEPTADO: 24/11/2023

Resumen: El éxito reciente de la inteligencia artificial supone uno de los mayores desafíos a los que tiene que hacer frente el ser humano en la actualidad, en especial debido a la irrupción de nuevas formas de crear y adquirir conocimiento. El presente trabajo trata de describir estos cambios así como los nuevos tipos de relación que la inteligencia humana puede establecer con la artificial, apostando por la complementariedad y dejando al margen la tesis del control. Finalmente se explora cómo pueden traducirse estas relaciones en el ámbito de la educación, donde la cuestión del conocimiento es central.

Palabras clave: inteligencia artificial; antropología filosófica; educación; conocimiento; control; coevolución

Abstract: The recent success of artificial intelligence represents one of the greatest challenges that humanity has to face today, especially due to the emergence of new ways of creating and acquiring knowledge. This paper aims to describe these changes as well as the new types of relationships that human intelligence can establish with artificial intelligence, focusing on complementarity and leaving aside the thesis of control. Finally, it explores how these relationships can be translated into the field of education, where the question of knowledge is central.

Keywords: artificial intelligence; philosophical anthropology; education; knowledge; control; coevolution

Introducción

El desarrollo de la tecnología más reciente, la inteligencia artificial (IA), obliga a volver sobre algunas de las preguntas más antiguas de la filosofía. La aparición de un alter ego digital, capaz de aparecer como competencia en tareas, como las cognitivas, en las que hasta ahora el ser humano tenía la exclusiva, nos lleva a desempolvar de nuevo algunas de las viejas cuestiones. ¿Qué es el ser humano? ¿En qué consiste su singularidad? ¿Puede hablarse todavía de una diferencia humana? La interrogación sobre nosotros mismos parece haber estado siempre lastrada por un obstáculo insuperable: la pregunta no fructifica en el vacío de la propia interioridad y necesita de un término de contraste externo respecto al cual afirmar o negar nuestra diferencia. Primero fueron los dioses, en los que los griegos, por ejemplo, vieron los términos con los que interpretar algo que ya existía pero no tenía nombre todavía y que se abrió paso lentamente a la observación racional dando lugar a la invención o descubrimiento de la esfera interior del espíritu (Snell, 2007), hoy diríamos, subjetividad. Después, la vida animal ofreció el modelo sobre el que la antropología filosófica ha vuelto neuróticamente una y otra vez como apoyo con el que definir, *a contrario*, la esencia humana. La irrupción de la IA supone una nueva vuelta de tuerca sobre la misma cuestión y pasa a funcionar hoy como “el espejo con cuya ayuda [el ser humano] se ve a sí mismo” (Snell, 2007, p. 340). El ser humano solo se entiende a partir de sus exteriorizaciones. ¿Qué tipo de imagen le devuelve este espejo? ¿Ofrece la IA alguna forma de autoconocimiento o más bien representa la ocasión para nuevas formas de automistificación?

Kant señaló de forma célebre que la pregunta “¿qué es el hombre?” era demasiado difícil de abordar directamente pero equivalía a otras tres preguntas, que una vez contestadas ofrecerían la respuesta a la anterior: ¿qué puedo saber?, ¿qué debo hacer? y

¿qué debo esperar? Vamos a centrarnos en lo que sigue en una estrategia doblemente indirecta para atacar la pregunta temible de la antropología filosófica, planteando la primera pregunta de la serie - “¿qué puedo saber?”- en el actual contexto de disrupción tecnológica e intentando obtener alguna respuesta a partir de lo que da a pensar la transformación en el campo del conocimiento que ha producido la aparición de un competidor cognitivo al ser humano como la IA.

Exteriorización

Vivimos en la época de la inteligencia artificial. Podría parecer entonces que lo que está pasando ahora es un acontecimiento radicalmente nuevo y sin precedentes. ¿Es así realmente? Muchas de las cosas que parecen “realmente nuevas” en realidad forman parte de una tendencia que viene de lejos y que puede ayudar a encontrar puntos de referencia para juzgarlas, disolviendo así el aura de misterio con las que las invisten (vamos a suponer que inconscientemente) tanto sus partidarios como sus detractores. No se trata de posicionarse a favor o en contra sino de algo previo: de orientarse para intentar entender.

La IA es el último avatar de una larga historia de “exteriorización” de capacidades humanas. “Exteriorizar” es el proceso de materializar u objetivar en un soporte externo una capacidad o habilidad que hasta ahora había tenido como soporte exclusivo al ser humano. El origen de esta larga historia se encuentra en la invención de la escritura, dado que escribir supone fijar en una memoria externa una capacidad que hasta el momento estaba unida de forma rígida a estructuras cognitivas humanas (Stiegler, 2002, pp. 192 y ss). Es posible que la escritura fuera precedida por el ritual, como protolenguaje que fija un significado simbólico en significantes externos - la coreografía corporal del ceremonial ritual

(Rappaport, 2016). En todo caso, no importa si es a través del papel, de tablas de cera o con cualquier otro medio más sofisticado, ya que la novedad consiste en traspasar información a un soporte físico que mantiene su integridad a través del paso del tiempo sin nuestra intervención activa.

Esta exteriorización posteriormente alcanza al campo de la percepción, con la invención en el siglo XIX de la cámara fotográfica o el telégrafo, que de nuevo permiten fijar en un soporte externo imágenes o sonidos. Se exterioriza así en un medio técnico la percepción visual o auditiva humana, es decir, capacidades o *competencias humanas*. Otros sentidos, como el tacto o el gusto, parece que se resisten por ahora a la exteriorización, aunque probablemente es solo cuestión de tiempo ya que se está trabajando intensamente desde hace algunas décadas con este objetivo.

La IA se puede describir entonces como la culminación, hasta el momento, de esta tendencia a la exteriorización que acompaña desde el principio a la historia del desarrollo tecnológico. Ahora le ha llegado el turno a la capacidad que parecía una prerrogativa exclusivamente humana: la capacidad cognitiva (el conocimiento, pensamiento o inteligencia), que vemos ahora realizada por un medio o dispositivo externo, el cableado *software* de las redes neuronales. La aparición de la ingeniería cognitiva implica que somos capaces de reproducir técnicamente un proceso o capacidad que hasta ahora se daba únicamente de forma natural (en un soporte biológico). Podemos extraer algunas conclusiones preliminares de este breve repaso histórico.

¿Extensión o exaptación?

Una interpretación no muy sofisticada consistiría en ver la exteriorización como una continua “extensión” de nuestras

capacidades (como decía McLuhan por ejemplo), es decir, una tendencia hacia la progresiva optimización con medios técnicos de capacidades humanas. Esta interpretación de la técnica como prótesis que sustituye a la vez que potencia es una versión apenas encubierta del relato moderno del progreso, donde cada etapa supera las anteriores (en el sentido hegeliano de eliminar lo superfluo y a la vez conservar lo válido en una versión mejorada), siguiendo un movimiento de mejora continuo. Este sigue siendo el relato que hoy en día sostienen los partidarios del transhumanismo. Curiosamente es también la base del relato apocalíptico de sus detractores, para los cuales la tecnología, en ese proceso imparable de mejora de capacidades humanas, hará del ser humano algo irrelevante y finalmente acabará por sustituirlo.

¿Qué tiene de miope esta interpretación? Precisamente que supone que lo que es mejorado no cambia cuando aparece la mejora. La historia de una mejora lineal y continua presupone que el ser humano es una especie de sustrato permanente al margen de los cambios que lo mejoran en otro soporte. Pero ¿qué pasaría si cada mejora lograda en el proceso de exteriorización en realidad no dejara intacto al ser humano sino que de alguna manera retroactuara sobre él y le forzara a modificar sus capacidades para adaptarse al nuevo entorno transformado debido a su intervención? Entonces la “extensión” del ser humano a la vez implicaría *autotransformación* y el relato lineal se curvaría hasta adoptar la forma de un bucle: es cierto que el hombre se exterioriza en sus dispositivos pero a la vez tiene que modificar sus capacidades para adaptarlas a la presión selectiva que le imponen esos mismos dispositivos. A diferencia del relato lineal, en un bucle el siguiente paso es siempre imprevisible: la historia permanece abierta y solo depende del tipo de interacción que puedan establecer los dos polos, el ser humano y sus exteriorizaciones.

No tendríamos que hablar entonces de “extensión” -y por tanto

de progreso- sino de “exaptación” -y por tanto de coevolución hombre-máquina. La “extensión” es la mitad del proceso, que necesita de la otra mitad (la “exaptación”) para cerrar el bucle que permite entender las implicaciones humanas de la aparición de una nueva tecnología.

En biología evolutiva la “exaptación” es el nombre de un proceso que hace referencia al cambio de función de un órgano debido a cambios en el entorno. A diferencia de la “adaptación”, en la que la aparición o permanencia de un órgano se explican por la ventaja que su función ofrece al organismo en su relación con el entorno, la “exaptación” implica el cambio de función de un órgano que ya existía, aunque tuviera otra y, por tanto, la reconfiguración interna que sufre el organismo bajo la presión del medio. Los órganos no son meramente sustituidos por otros sino que se mantienen con funciones diferentes. La adaptación supone la sustitución o superación de lo viejo por lo nuevo, la exaptación supone en cambio un nuevo uso de lo viejo. No hay optimización sino diferencia. La adaptación produce lo mismo pero mejorado, la exaptación produce algo diferente.

Un ejemplo clásico es la descripción que hace Leroi-Gourhan de la bipedestación: cuando el ser humano logró la posición erecta las manos perdieron la función locomotora, que pasó a las piernas, pero a la vez adquirieron una nueva función, la de exploración táctil y prensil, que a su vez liberó al rostro de estas funciones (la prensión con los dientes, por ejemplo) y permitió que ganara otras nuevas: el rostro para expresar emociones, la boca para hablar y los ojos para tener un horizonte visual más amplio, lo que dio lugar a la visión a distancia.

La refuncionalización es un proceso habitual en la historia de la recepción de la técnica. Así, por ejemplo, se ha señalado frecuentemente que la aparición de la escritura (la “técnica de la palabra”) comportó una reconfiguración de las funciones cognitivas

humanas, que una vez liberadas de la necesidad de la memorización literal, ahora delegada en el registro escrito, pudieron adquirir otras nuevas: la exploración imaginativa, la innovación y el pensamiento abstracto (Ong, 2016; véase también Serres, 2014, p. 280). También por lo que respecta a la percepción se puede indicar un ejemplo similar. Cuando aparece una nueva tecnología a menudo la primera reacción consiste en un proceso de imitación, que hace con lo nuevo lo mismo que hasta ahora se hacía con la tecnología anterior. También esto se aplica al caso de la escritura, que en sus inicios “imita por escrito la producción oral”, antes de conseguir independizarse y afirmar las propias potencialidades del nuevo medio (Ong, 2016, p. 67). La aparición de nuevos medios como la fotografía o el cine, así como la aparición de nuevos materiales para la arquitectura, desencadenó al inicio una respuesta mimética: la fotografía, por ejemplo, se utilizaba para hacer retratos, a imitación de la pintura, y el acero se continuó utilizando para los ornamentos florales que antes se hacían con madera. Es solo en un segundo momento cuando se aprecia la potencialidad del nuevo medio artístico y esto produce un cambio en el sentido de la propia actividad artística, que ya no se realizará en los mismos términos que hasta ahora. Walter Benjamin ha aportado multitud de ejemplos en las secciones F y G del Libro de los Pasajes sobre la oposición entre el momento “histórico” y el “constructivo” de la recepción de la técnica en el siglo XIX. El uso *kitsch* de la tecnología del siglo XIX se puede detectar también en las primeras producciones en serie, que se dedican en un primer momento a fabricar de forma industrial copias de productos hechos a mano, desde muebles a los primeros coches, que al principio son producidos unidad a unidad, mientras que solo en un segundo momento aparece una forma de creatividad propiamente industrial (Lipovetsky y Serroy, 2015, pp. 130-133). Por tanto, ser creativo no quiere decir lo mismo antes que después de la aparición de los nuevos medios.

También en el caso de la IA nos podemos hacer la misma pregunta: ¿qué nuevo sentido de nuestras capacidades pueden liberar o hacer surgir en nosotros las nuevas exteriorizaciones técnicas? ¿O tenemos que resignarnos a un uso *kitsch* de la técnica, como el que supone la “extensión” transhumanista de capacidades, que pretende mejorarlas sin cambiarlas, optimizar el órgano sin cambiar su función, como si nuestras capacidades fueran un paquete de rasgos fijos y no pudieran cambiar, por ejemplo, a través de la interacción con las máquinas? Podemos aprender del pasado e intentar evitar algunos de sus errores explorando alternativas a la recepción *kitsch* de la IA, ejemplificada actualmente en la tesis del control (dominante hoy en día en el tratamiento de la relación humano-máquina). Según reza esta tesis, para evitar ser dominados por la IA la tendremos que dominarla nosotros, utilizándola como medio subordinado a nuestros objetivos (un ejemplo representativo se puede encontrar en Russell, 2019 y Tegmark, 2018). Uno de los problemas que plantea la cuestión del control es que se define a partir del sujeto que lo ejerce (control es lo que hacemos para protegernos de cambios súbitos e imprevisibles) y, por tanto, presupone o da por bueno un sujeto con una dotación de capacidades relativamente inalterable y al margen de los cambios. Una forma diferente de relacionarse con la técnica pasará por relativizar la tentación del control y aceptar una cierta apertura e indeterminación en la relación mutua entre el ser humano y sus aparatos.

Inteligencias múltiples

Hemos visto que el sujeto no queda intacto después de la exteriorización. Veremos a continuación que algo parecido le pasa a lo que es exteriorizado. Una segunda conclusión que se puede extraer de la concepción de la técnica como exteriorización es que la

capacidad exteriorizada no supe exactamente a aquello que exterioriza. Por eso la metáfora de la técnica como prótesis (órgano supletorio) sólo capta una parte de la verdad. En realidad la capacidad humana (memoria, percepción o cognición) cambia de naturaleza cuando cambia de soporte, del biológico al inorgánico. Se convierte en otra cosa.

Aceptar la diferencia cualitativa entre la capacidad natural y la artificial permite cuestionar de nuevo tanto la historia de progreso como la de catástrofe. En efecto, por lo que respecta a la primera, la posthumanidad (el hombre potenciado gracias a la tecnología) presupone que la técnica amplifica y por tanto aumenta cuantitativamente una capacidad que mantiene la homogeneidad cualitativa con aquello que es aumentado (ya que sino no habría progreso). Por lo que respecta a la segunda, el miedo a nuestra sustitución por la IA o la presión por competir con ella presupone que las máquinas saben hacer mejor lo que nosotros no hacemos tan bien, que pueden triunfar donde nosotros hemos fracasado. El miedo o la competición son formas de reacción que presuponen, por tanto, que las capacidades técnicas son del mismo tipo que las humanas, con solo una diferencia en el más o el menos (para una crítica de este presupuesto véase Coeckelbergh, 2022, pp. 32-33).

Este supuesto resulta ser muy discutible. Por lo que respecta a la memoria, la humana y la técnica funcionan de forma muy diferente. La primera triunfa en el orden y en ámbitos dotados de estructura pero en cambio es poco precisa con el detalle literal. La segunda es una memoria literal y puede guardar con exactitud un elevado volumen de datos pero en cambio es ciega al orden o desorden. La oposición corresponde a la distinción que Simondon hace entre “plasticidad del soporte” y “plasticidad del contenido” (Simondon, 2008, pp. 138-140). Dado que cada una tiene un alto rendimiento donde el de la otra es bajo, esto sugiere la posibilidad de una relación de “complementariedad”, donde la técnica puede suplir las

insuficiencias humanas pero sin el riesgo de sustituir sus capacidades.

También en el caso de la percepción se ha señalado que el ojo técnico no funciona como el ojo humano: el primero realiza un gesto instantáneo y estático mientras que el segundo se compone de una multiplicidad de microgestos que captan de forma exploratoria la superficie que miran y luego sintetizan todos estos aspectos en un objeto virtual que no tiene correlato físico directo (Virilio, 1998, pp. 9-26).

¿Y en el caso del conocimiento? Si partimos de una definición en principio neutra del conocimiento como “procesamiento de información” parece que sí sería posible equiparar las prestaciones humanas con las técnicas. El reconocimiento de imágenes o la generación de lenguaje natural, tareas en las que la IA está obteniendo un elevado rendimiento, se basan en procesos algorítmicos que lo que hacen es, efectivamente, procesar información. Tal vez el problema esté en que la definición inicial resulta no ser tan neutra como parecía ya que, en virtud de su misma vaguedad, predispone a aceptar una cierta equiparación entre la inteligencia natural y la artificial, cuando en realidad, vistas más de cerca, sus diferencias pasan a un primer plano. El ser humano y la IA no ocupan el mismo “nicho cognitivo” por lo que pueden cohabitar en el mismo ecosistema sin competición e incluso explorar formas de cooperación mutua (Lee, 2020, p. 160). Este enfoque de una diferencia cualitativa entre distintos tipos de inteligencia ayuda a plantear la pregunta correcta sobre la IA, que no es tanto “qué pasará” (cuestión de especulación y futurología) como “qué debemos hacer” ahora en el presente. Es decir, ¿cómo hemos de utilizarla, en función de lo que sabemos sobre su funcionamiento actual y de acuerdo con el estado actual de nuestras capacidades?

Por tanto, si queremos describir el nuevo papel que le corresponde al ser humano como cooperador de las máquinas es

necesario saber qué diferencia el conocimiento humano del artificial. No hay una única forma de inteligencia sino que cada una tiene un rendimiento diferente en diferentes condiciones. Sabiendo los puntos fuertes y débiles de cada una podremos entonces identificar en qué capacidades ha de especializarse el ser humano y en cuáles puede delegar en la máquina para reforzar la relación de complementariedad entre ambos.

En la época de la reproductibilidad técnica de la capacidad cognitiva, el ser humano tiene que identificar la especialidad humana si quiere preservarla y desarrollar su espacio propio. Su tarea será la de especializarse en la propia humanidad. Si se dedica en cambio a copiar e imitar a sus productos entonces la tesis de la sustitución se volverá verdadera de manera performativa: si creemos que hemos de competir en el mismo nicho que las máquinas entonces nos comportaremos de forma que seremos finalmente sustituidos por ellas. Si ponemos en cambio el acento en las diferencias no es para satisfacer imaginariamente una falsa sensación de dominio. Esta idea del ser humano ya no puede sostenerse. No se puede mirar a otro lado para no ver la herida narcisista que supone la aparición de otra forma de inteligencia donde hasta ahora reinaba la nuestra en solitario. Al contrario, identificar la diferencia con las máquinas nos da la clave para saber cómo establecer un nuevo vínculo y acercarnos productivamente a ellas.

¿Cuáles son entonces las diferencias entre los dos aparatos cognitivos? ¿Qué quiere decir conocer, en nuestro caso y en el suyo? ¿Pueden las dos formas coexistir y, es más, cooperar?

El fin de las competencias

El antropólogo Philippe Descola ha trazado un mapa de la combinatoria de opciones de que dispone el ser humano para

establecer relaciones con su entorno. En un momento de su análisis se centra en lo que denomina los “esquemas de la práctica”, que se encuentran a medio camino entre las impresiones sensibles particulares y los conceptos generales, haciendo por tanto de mediadores entre el agente y su entorno. Establece entonces un paralelismo entre la cognición humana y la técnica, para lo que distingue primero entre los “modelos clásicos” de programación y los “modelos conexionistas”, es decir, el paradigma en el que se basan las redes neuronales, el subtipo de la IA que desde hace poco más de una década está provocando la actual revolución tecnológica (Descola, 2005, pp. 186-187).

El “modelo clásico” funcionaba a partir de una lista de instrucciones que se suministraba de forma explícita a la máquina para que realizara los cálculos. En cambio el nuevo modelo “está constituido por un conjunto de redes electrónicas la interconexión de las cuales se establece selectivamente en función de la naturaleza e intensidad de los estímulos recibidos”. Por tanto, hay una oposición entre una forma de conocimiento que se basa en establecer “reglas explícitas” que formulen o capten alguna regularidad del ambiente y otra donde la máquina reacciona rápidamente a los cambios en el entorno y “parece obedecer a reglas formales sin que ninguna estipulación se haya introducido en el modelo”. Así pues, para Descola, la oposición dentro del ámbito técnico entre un “saber declarativo” (o proposicional) y un “saber-hacer” basado en el aprendizaje a partir de la práctica - entre el *know-what* y el *know-how* - refleja la misma oposición con la que se encuentra el etnólogo cuando quiere estudiar las reglas de comportamiento de una cierta comunidad, algunas de las cuales son explícitas pero otras en cambio están incorporadas de forma implícita (en forma de hábitos) en la práctica de los individuos.

La comparación que realiza Descola es pertinente por dos motivos. En primer lugar, porque una de las conclusiones que se

desprenden de ella es que el umbral alcanzado por la IA en el campo del conocimiento consiste en haber pasado de un saber declarativo a un saber procedimental, una capacidad hasta ahora exclusivamente humana que por primera vez ha sido exteriorizada con éxito en soportes técnicos. En segundo lugar, si el dominio del hábito (*know-how*) ha sido alcanzado por las máquinas ¿esto implica que toda una parte de la competencia humana ha de ser delegada en sus prótesis? En otras palabras, si el conocimiento basado en competencias (interiorización de hábitos prácticos, saber-hacer, etc.) ya no es una prerrogativa exclusivamente humana ¿quiere esto decir que antes o después seremos sustituidos en este ámbito y por tanto tendremos que emigrar a otro nicho cognitivo? ¿A cuál?

Edward Lee ha propuesto recientemente un paralelismo similar (2020, pp. 66-68). El progreso en la IA se ha basado en la introducción de la retroalimentación (*feedback*) en el mecanismo interno de los programas: la generación anterior se basaba en los “sistemas expertos”, donde los humanos codificaban primero y suministraban después a la máquinas las reglas explícitas que capturan el conocimiento de cómo se comporta un cierto ámbito; en la nueva generación, en cambio, “las representaciones de conocimiento se aprenden a partir de los datos”, no obedecen a órdenes sino que se extraen las regularidades a partir de muchos casos de ejemplo, tratando de ajustar progresivamente la respuesta efectiva del programa delante de los datos de entrada a un modelo esperado de cómo tiene que ser esa respuesta. De aquí el bucle de feedback entre los datos de entrada y salida (para una descripción en la misma línea de las diferencias entre el “paradigma simbólico” y el “conexionista” de la IA véase de Carolis, 2004, pp. 112-123).

Podríamos añadir por nuestra cuenta a esta idea la mitad que le falta para completar la analogía. El paso realizado por la IA desde la aplicación de un saber explícito a la producción de conocimiento implícito mediante la exposición a situaciones cambiantes ¿no calca

el paso dado por los humanos entre el *know-what* y el *know-how*, entre el modelo pedagógico de un conocimiento basado en contenidos (considerado hoy en día obsoleto) y un conocimiento basado en competencias, que ha funcionado (al menos hasta ahora) como modelo de todo conocimiento deseable? Precisamente, si el saber de contenidos viene considerándose obsoleto desde hace algunas décadas es muy posible que esto se deba a que ya había sido exteriorizado por las máquinas de una generación anterior. Como se suele decir ¿de qué sirve el aprendizaje memorístico en la época de Google y Wikipedia, cuando cualquier contenido está al alcance de cualquiera en cualquier momento? De la misma manera, debemos hoy preguntarnos *¿de qué sirve el conocimiento competencial en la época de la IA, cuando cualquier rendimiento competencial está ya (casi) al alcance de cualquiera en cualquier momento?*

Si por competencia entendemos “la capacidad de actuar en situaciones complejas e imprevisibles aplicando los conocimientos aprendidos” en lugar de “reproducir el conocimiento transmitido” (Sanmartí, 2020), entonces, una vez que la IA parece ser capaz de reproducir este tipo de conocimiento ¿deberíamos cambiar de modelo y buscar otro todavía no exteriorizado por la técnica? La IA ha realizado en poco tiempo el cambio que supuestamente tenían que realizar los humanos en el paso de una generación (la de la sociedad industrial) a la siguiente (sociedad postindustrial). Por tanto, las máquinas, que han pasado de un modelo de conocimiento “dirigido por las instrucciones” a uno “dirigido por los datos”, son hoy ya *competenciales*, es decir, capaces de actuar en un entorno postindustrial: cambiante, incierto, contextual, etc. Como sostiene Anthony Elliott “mientras que las máquinas automatizadas convencionales permanecían fijas en un lugar y eran programadas para tareas específicas y repetitivas, las nuevas tecnologías son móviles, sensibles a la situación y pueden adaptarse y comunicarse con su entorno” (Elliott, 2018, p. 21). Esto plantea el debate sobre si

la nueva oleada de automatización implicará la destrucción de puestos de trabajo humano (y no solo de baja cualificación profesional) o si la creación de nuevas profesiones compensará la destrucción de los antiguos puestos de trabajo (véase el equilibrado balance entre “escépticos” y “transformacionistas”, decantado finalmente hacia los primeros, en pp. 55-77). ¿Qué otro cambio o adaptación impone este nuevo entorno a la siguiente generación? Si las máquinas actuales ya son competenciales ¿qué nos tocará ser a nosotros? ¿Cómo interactuar con un entorno formado por máquinas que saben interactuar con su entorno? El reto por lo que respecta al conocimiento es entonces el siguiente: después de la transición del modelo memorístico al competencial ¿tendrá que seguir ahora una nueva transición? (y ¿hacia dónde?).

La respuesta es sí y no. Comencemos por el sí. Podríamos establecer un paralelismo entre la tecnología digital del siglo XX y la IA de principios del XXI. La digitalización del siglo pasado se basaba en la desmaterialización de los contenidos: texto, audio e imagen se convertían en información indiferenciada, una “gelatina de información” (como decía Marx respecto al trabajo), que permitía pasar cualquier contenido del mundo real al virtual. Este proceso provocó un aumento del ritmo de rotación de contenidos y consecuentemente dio lugar a formas de atención flotante y una memoria de trabajo más reducida. El modelo de las competencias surge como reacción a este tipo de entorno: una forma de adquirir conocimientos que sea compatible con formas de atención superficial y discontinua es precisamente la que tiene lugar a través del hábito o la ejercitación de capacidades (implícitamente, a través de la práctica, y no explícitamente, debiendo estar atento a lo que dice otro).

La IA, en cambio, hoy ya es capaz de hacer de forma automática una buena parte de aquello que hasta ahora se tenía que aprender a través del hábito, como, por ejemplo, diferentes formas de organizar

información (clasificar, relacionar, encontrar semejanzas, redactar, etc). Por tanto, en este nuevo entorno asistido por herramientas con IA, habrá que desarrollar aquellas capacidades que saquen provecho de estas nuevas herramientas y sepan hacer lo que ellas no saben hacer por sí mismas, sin orientación externa. Para concretar en qué podrían consistir estas *supercompetencias*, con una relación con las competencias como la que éstas tenían con los contenidos, podemos tomar como punto de partida la diferencia cualitativa entre las capacidades humanas y las técnicas. La IA, una vez ha logrado con éxito el *learning by doing*, necesitará todavía el juicio humano. La tecnología actual sabe hacer lo que antes requería una coordinación humana con las máquinas, por tanto, en la época de la automatización del conocimiento ahora la coordinación tendrá que cambiar para adaptarse a las nuevas prestaciones cognitivas del sistema de aparatos. Las capacidades humanas se tendrán que integrar y sincronizar con el sistema exteriorizado de capacidades, tal como lo hicieron en el pasado a través del modelo de aprendizaje por competencias.

Consideremos ahora el no: no hay nada más allá del marco de las competencias. Solo hay los contenidos y lo que hacemos con ellos. De la misma manera que el conocimiento competencial no ha desplazado del todo al conocimiento de contenidos sino que ha supuesto sólo un cambio de prioridades, dando lugar a un modelo híbrido, tal vez estemos ahora en la misma situación. Más que ser sustituido, el modelo de las competencias tendrá que ser reconsiderado.

Conocimiento encarnado

Una tal reconsideración podría pasar por revalorizar lo que se conoce como “conocimiento encarnado” (Lee, 2020, p. 105). La capacidad

de conocer no reside únicamente en el cerebro sino también en el cuerpo. En el ser humano el conocimiento no es separable del instrumento con el que se adquiere, el propio cuerpo, a diferencia de la IA, en la que un mismo software puede ejecutarse en diferentes hardware, lo que se conoce como “realizabilidad múltiple” (Nowotny, 2021, pp. 75-78). Esta es, por tanto, una primera diferencia que singulariza al ser humano delante de la IA. ¿Por qué es importante? Es cierto que las tareas que aprendemos a hacer, a través de la interacción entre el cuerpo y el cerebro, las pueden hacer hoy también algunas máquinas pero es importante mantener este aprendizaje ya que es la condición para aprender a hacer otras tareas que (al menos de momento) no pueden hacer las máquinas. Por tanto, sustituir aquellas tareas por su versión mecánica sería un error ya que interrumpiría un proceso de aprendizaje a más largo plazo. Por ejemplo, el cálculo mental hace tiempo que puede ser delegado en las máquinas pero todavía se enseña porque ayuda a aumentar la memoria de trabajo, lo que es útil posteriormente para realizar otras tareas más complejas.

Richard Sennett, un gran defensor de este tipo de conocimiento, distingue a este respecto entre un mal y un buen uso de la tecnología (Sennett, 2009, pp. 55-62). El mal uso es el que impide el aprendizaje por repetición, basado en hacer y luego observar lo que se ha hecho, ya que interrumpe el bucle acción-observación que es una de las fuentes elementales de la experiencia humana. Las máquinas “separan la cabeza de la mano” - la comprensión mental, puramente abstracta, del aprendizaje manual, realizado en primera persona - y rompen así el bucle de retroalimentación entre hacer y pensar, ya que suministran el resultado separado del proceso del que surge, lo que disminuye el grado de comprensión. La máquina se convierte en un proveedor de datos correctos para la comprensión mental, pero al precio de cortocircuitar el bucle a través del cual validamos un dato como correcto o incorrecto. Como dice Sennett, nos volvemos

testigos pasivos de una competencia que va en aumento pero en la que no participamos. Por tanto, la máquina sustituye el proceso de adquisición de conocimiento y acaba impidiendo la propia intervención y exploración de un campo de posibilidades, lo que dificulta la autoinstrucción y un conocimiento activo y consciente.

La misma oposición que describe Sennett -entre una forma de conocimiento que depende de un cuerpo que percibe y actúa y otra donde hay un aislamiento relativo entre la obtención de datos (*input*) y el procesamiento de información (*output*)- es la que describe Descola (2005, pp. 325-329) cuando opone la “cognición ecológica” al modelo de la “máquina de Turing” (es decir, el ordenador). Los ordenadores, incluso los que son “capaces de modular sus acciones siguiendo un aprendizaje autónomo” (es decir, la IA), se basan en “modelos conexionistas” de procesamiento de información que presuponen una ontología cartesiana: la separación entre la unidad de control y los subsistemas de sensores y actuadores refleja así la separación entre espíritu y cuerpo. De un lado la obtención de datos, del otro el procesamiento o la capacidad de cálculo como un nuevo Ego cartesiano.

En cambio, una forma alternativa de pensar el conocimiento es a través de la “cognición incorporada o ecológica”. El conocimiento humano “es función de un cuerpo que percibe y que actúa”. Conocer no es meramente obtener información del ambiente, de una forma neutral o pseudo-objetiva, sino percibir estímulos del ambiente que puedan traducirse en una acción posible. Las preguntas que le hacemos al entorno están condicionadas por las posibles respuestas que le puede suministrar nuestro cuerpo. Parafraseando a Marx, cada organismo sólo se plantea los problemas que es capaz de resolver. El conocimiento humano está siempre situado, es decir, sesgado por sus propios instrumentos cognitivos. Por tanto, depende de dos factores: de un lado, de las capacidades de percibir y de actuar, es decir, de las posibilidades limitadas que admite el aparato perceptivo y locomotor

del cuerpo humano; de otro lado, del vínculo o libre juego que se establece entre estos dos aparatos, que no son como los subsistemas abstractos y, por tanto, independientes de una máquina (cada uno con una definición autónoma de la tarea que ha de cumplir) sino que están definidos uno en relación al otro. La percepción no tiene como objetivo obtener datos y la unidad de control producir respuestas sino que la percepción trabaja en vistas a una acción posible y la acción en vistas a la percepción. Esta retroactividad entre los dos aparatos hace de ellos una unidad donde cada polo es a la vez condición y resultado del otro: uno proporcionando datos para una acción posible y el otro produciendo acciones con un resultado que a su vez se constituye en un nuevo dato observable, cerrando el bucle que define la acción humana.

Señal y ruido

Que el mundo no está formado por propiedades objetivas (datos, información) sino que viene filtrado o interpretado por las posibilidades de acción que se le presentan al individuo es también uno de los temas clásicos de reflexión de la antropología filosófica europea del siglo XX, a partir primero de Jakob von Uexküll y después de Arnold Gehlen o Helmuth Plessner, autores que intentaban desmarcarse del intento más tradicional de dar una definición de una esencia humana (ahistórica y no situada en un contexto práctico) y que vieron en el giro hacia la acción la forma de desarrollar un estudio del ser humano menos cargado de presupuestos metafísicos y entidades inobservables (Rehberg, 1983, pp. 385-402).

Para Uexküll, por ejemplo, el sentido que tienen los estímulos y cómo se percibe el mundo depende del catálogo más o menos limitado de acciones que puede realizar cada organismo. Es el “tono

efectual” (Uexküll, 2016, p. 105, p. 144) de una percepción lo que convierte a los estímulos en puntos de referencia para la orientación práctica en el mundo. Aunque diversas especies compartan un mismo espacio en realidad viven en diferentes ambientes, que recortan a partir de la diferente dotación instintiva y las peculiaridades de su sistema sensomotor. El “mundo” objetivo (*Welt*) se difracta entonces en una diversidad de “ambientes” (*Umwelt*), según las cosas se carguen, como signos, con significados que dependen de lo que se puede hacer con ellas.

La distinción entre mundo y ambiente ha sufrido sin embargo una cierta fluidez terminológica. Si para Uexküll no es todavía una oposición tajante sino una diferencia de grado que permite sostener una cierta continuidad entre el ser humano y el animal (capaces, cada uno a su manera, de recortar mediante la acción un ambiente a partir del mundo ilimitado), en cambio para Gehlen servirá para marcar una diferencia sin matices entre el mundo humano y el ambiente animal (apertura contra clausura), aunque en su obra posterior esta oposición se irá matizando progresivamente. Finalmente, Plessner se muestra más ambiguo cuando sostiene que lo propio del hombre es hacerse “un ambiente a partir del mundo”, donde mundo y ambiente parecen haber intercambiado su significación respecto a la intención inicial de Uexküll. En todo caso, a pesar de las diferencias, lo que estos autores tienen en común es la afirmación de una paradoja antropológica irresoluble (de Carolis, 2017, pp. 55-61): entre *apertura* (la exploración de nuevas posibilidades propia del “mundo”) y *protección* (el límite y la negación de la contingencia propia del “ambiente”).

La importancia actual de la distinción entre ambiente y mundo va más allá del ámbito especializado del que procede y puede captarse mejor si se la interpreta como “una diferencia entre dos modalidades de elaboración de la información” (de Carolis, 2004, pp. 56-61). Más concretamente, entre dos formas de relación entre señal y ruido.

Manejarse en un ambiente significa que la distinción entre lo que tiene significado (señal) y lo que no lo tiene (ruido) es relativamente fija y presenta poco margen de variación. La separación entre la figura y el fondo informe depende de una dotación biológica (genética, instintiva, sensomotora, etc.) que admite pocas variaciones. Desempeñarse en un mundo en cambio quiere decir que lo que vale como señal (y por tanto significa algo) no está dado de una vez por todas sino que hay un límite móvil que separa pero a la vez comunica señal y ruido, lo que es relevante o significativo de lo que no es. Aquí la selección de lo que vale como señal puede volver a producirse de nuevo a partir del ruido, dado que este funciona como una reserva potencial de significado a la espera de ser activada y transformar el ruido en un nuevo orden.

La antropología filosófica europea ha insistido de forma preponderante sobre el polo de apertura al mundo como rasgo dominante de la condición humana pero, como señala de Carolis, desde el principio de su reflexión supo ver en ella un bucle o reenvío continuo entre el polo de la apertura y el polo del repliegue: recortar un ambiente estable y reabrir de nuevo el acceso a un mundo, poner un límite ante el exceso de sensaciones y desplazar ese corte para rehacerlo en otro lugar. Ambos coexisten en la criatura humana y hacen de su condición un campo de tensiones productivas e irresolubles.

¿Y la inteligencia artificial? ¿Tiene un ambiente o un mundo? ¿Cuál es la relación que establece entre señal y ruido? ¿Hay un límite fijo o una cierta indistinción entre ambas? ¿Muestra algún grado de apertura al mundo o el entorno con el que interactúa se reduce a un conjunto de datos objetivamente definido?

Para empezar hay que recordar que una red neuronal no elige los datos con los que se entrena sino que es el programador quien de forma externa hace una selección y se los proporciona. Por tanto, la delimitación de qué vale como dato significativo para el aprendizaje

de la red y de cuál es su significado (qué valor de salida le corresponde) está fijado de antemano. No está fijado, ciertamente, el valor o peso que un dato de entrada adquiere dentro del proceso de aprendizaje para producir una salida, y esa flexibilidad es la que permite precisamente el aprendizaje. Se podría decir que la máquina no necesita criterio de selección, el cual es un índice de la debilidad humana, ya que para el ser humano todo es potencialmente significativo y para orientarse necesita hacer un corte, aunque sea provisional y revisable, para reducir la complejidad de su entorno y ser operativo. La máquina en cambio puede calcular a partir de todos los datos disponibles. Sin embargo, en su caso es el objetivo el que funciona como criterio de selección de los datos de entrada (aquellos que sirvan para aprender a realizar mejor una cierta tarea), un objetivo que le viene marcado desde fuera (como veremos en la sección siguiente).

¿Qué es aprender entonces para el ser humano y para la IA? Para el primero, desplazar el límite entre señal y ruido, de forma que lo no significativo se vuelva significativo y donde solo había ruido se perciba una nueva forma de orden. Para la segunda, aprender no tiene que ver con captar un sentido sino con aprender a producir un resultado correcto. Para ello la IA se sirve de la agregación de unidades de cálculo subsimbólico (las neuronas de una red, cada una de ellas ciega al significado), de las cuales el significado emerge como el resultado final, de acuerdo con parámetros fijados de antemano que son revisables para una apreciación humana externa pero no para la propia máquina. Por tanto, en ambos casos el aprendizaje discurre en la misma dirección pero en sentidos opuestos (de Carolis, 2004, p. 60): para la máquina aprender significa alcanzar una mayor concretización en su respuesta al entorno (fijación de la distinción señal/ruido, es decir, reducción del ruido al mínimo), para el hombre aprender implica lograr una mayor ductilidad (exponerse al ruido para transformarlo en señal).

La IA no tiene entonces capacidad de reabrir la distinción entre señal y ruido y ese es precisamente su punto fuerte, siendo su ceguera al significado de los datos el precio que tiene pagar para resultar eficaz. De forma más general podríamos decir que la respuesta de un organismo será tanto más eficaz cuando más cerrado esté al ruido inútil (ambigüedad entre señal y ruido). Inversamente, un organismo exhibirá tanta más potencia interpretativa cuanto menos presión tenga para responder de forma eficaz. Lo cual parece conducir a una forma de complementariedad entre la inteligencia natural y la artificial -entre interpretar y calcular, mundo y ambiente, apertura y protección- que no solo no deshace sino que refuerza la paradoja antropológica.

De hecho, se podría decir, la técnica es la exteriorización de uno de los polos de la paradoja interna que constituye al ser humano, de ahí que siempre haya servido de espejo de reconocimiento. El bucle entre significación y eficiencia, del cual uno de los elementos se incorpora al mundo externo en forma de dispositivo técnico, amplía su alcance excediendo la esfera de la subjetividad para incluir lo que está más allá de ella. La técnica aparece así para el ser humano como una forma de relación consigo mismo a través de lo otro de sí. Esto implica que el intento de marcar una diferencia demasiado tajante entre el ser humano y la técnica está siempre desencaminado ya que la diferencia remite al interior de la propia condición humana. Este movimiento de ida y vuelta entre el ser humano y la técnica se puede reconocer de nuevo en el bucle entre práctica y producción.

Producción y práctica

En la *Ética a Nicómaco* Aristóteles distinguía entre dos tipos de virtudes o facultades humanas: las naturales, que preceden su ejercicio (como ver u oír), y las morales (o que pertenecen al

carácter), que son resultado de la propia acción del agente. “Los hombres se hacen constructores construyendo y citaristas tocando la cítara. Pues bien, de esta manera nos hacemos justos realizando acciones justas y valientes” (1103a). Posteriormente, cuando trata las facultades del intelecto, hace una nueva distinción entre dos subtipos: la *poiesis* (facultad propia de la esfera de la técnica) y la *praxis* (excelencia humana en general, sin delimitación de ámbito). Esta distinción entre dos formas de conducta humana -la producción y la práctica- puede resultar útil, después de un salto de más de dos mil años, para describir la diferencia actual entre las competencias humanas y las artificiales. ¿Se podría decir que la IA funciona según el modelo de la producción y la acción humana según el de la práctica?

Tal como sostiene Aristóteles, “ni la realización [*praxis*] es fabricación [*poiesis*] ni la fabricación realización” (1140a). La *producción* (o fabricación) coincide, a grandes rasgos, con la acción instrumental, mero medio para un fin (la realización de un producto) al que está subordinada y del que se puede separar de forma clara. La acción es lineal y su objetivo es el resultado. En la *práctica* (o realización), en cambio, el fin es la actividad misma, el hecho de realizarla bien o de forma excelente.

El perfeccionamiento técnico de la actividad productiva en el último siglo y medio, con la creciente tecnificación de diferentes ámbitos de acción (el trabajo, la guerra o el arte), se basa efectivamente en el momento de la producción: importa primeramente la *eficacia* con que se logra un resultado, prescindiendo de cualquier otra consideración. En cambio la práctica crea su propio sentido, que no es previo ni exterior sino que lo segrega, por así decir, la acción durante su ejercicio.

Desde el punto de vista de la práctica, la acción no es solo un medio para conseguir un objetivo sino un medio que, al mismo tiempo que persigue un fin, ayuda a revelar en qué consiste ese

mismo fin que se busca. Cada acción, en su particularidad específica, una vez realizada ejemplifica y, por tanto, da indicaciones de cómo sería un modelo logrado de esa acción. El fin, por tanto, es inseparable del medio, es en cierto sentido producido por el propio medio, que le sirve de guía. De nuevo, podría decirse también ahora con Marx que cada agente sólo se propone aquello que es capaz de realizar. El fin establece los medios necesarios para su consecución pero los medios redefinen el sentido que hay que darle al fin, dan acceso a nuevos aspectos insospechados suyos, por lo que modifican nuestra concepción de qué significa aquello que buscamos. Este acaba siendo el sentido de la acción: el conocimiento que la acción procura al agente sobre lo que busca al actuar (en el límite, qué quiere decir para el agente hacerse mejor, lo que los griegos llamaban “virtud”).

La acción presenta entonces la conocida forma de un bucle entre medio y fin: uno quiere actuar bien para ser virtuoso y quiere ser virtuoso para actuar bien. El valor de la actividad no se mide solo por sus efectos (*qué*) sino también por la modalidad interna de la propia acción (*cómo*). La acción es un medio para un fin, pero a la vez la acción como realización de un fin es usada como medio para saber en qué consiste ese fin. El propio medio es instrumentalizado en una acción de segundo grado con un rendimiento epistémico (procura conocimiento) y no solo práctico (realiza objetivos).

De lo anterior se sigue entonces que en la práctica no hay una definición previa del objetivo de la acción (un modelo óptimo) sino que la definición surge y se enriquece progresivamente por medio de la acción. Por tanto, la técnica trata la acción como producción y está dirigida a la *optimización* del producto. En cambio el ser humano puede preocuparse del sentido de la acción, que está relacionado con un proceso de *autoperfeccionamiento* o mejora de la propia calidad humana (para una interpretación de la diferencia técnico-humano a partir del par *poiesis-praxis* y sus variantes véase de Carolis, 2004,

pp. 38-44 y Bodei, 2022, pp. 236 y ss.). A través de la acción, logra resultados; a través del conocimiento que le brinda la acción se pone de camino hacia su propia humanidad.

¿Qué tiene que ver todo esto con la IA? Podríamos decir que la IA es platónica mientras que el ser humano es aristotélico. El algoritmo con el que funcionan las redes neuronales actuales requiere de la definición previa de un óptimo de su rendimiento: es necesario saber cuál es la respuesta esperada para poder medir la distancia entre la respuesta que produce la máquina y la que tendría que producir. De manera que el objetivo del algoritmo de aprendizaje es maximizar el éxito en las respuestas que genera (o, lo que es lo mismo, minimizar el error). Por tanto, la red neuronal puede optimizar su rendimiento solo porque parte de una definición previa de lo que es un óptimo (que es definido por el programador cuando le suministra el conjunto de datos de entrenamiento). La IA requiere de estos tres ingredientes: la definición humana de un óptimo, una métrica que indica si se acerca o no a su rendimiento óptimo (lo que se conoce como “función de pérdida”) y la modificación interna de los pesos de la red para minimizar esa métrica. A pesar de la gran variedad de técnicas que utiliza la IA actual, todas se basan en el “modelo estándar de la IA”, es decir, “un algoritmo que optimiza un objetivo prefijado” (Russell, 2019, pp. 60-64). El modelo alternativo, el de la “IA de propósito general”, no está aún a la vista (sobre la insistencia de una “IA benéfica” bajo el control de objetivos humanos véase también Tegmark, 2018, pp. 61-63).

Esta actividad encaja con el modelo de la producción (la acción como un medio para conseguir una finalidad definida previamente). Por tanto, tendrá éxito allí donde sea posible definir un óptimo pero fracasará donde lo no sea, no solo porque aún no sabemos cómo hacerlo sino porque sencillamente no es posible hacerlo, porque es consustancial a la acción humana el hecho de no tener óptimo (una definición de una vez por todas). Y no tiene óptimo precisamente lo

que tiene sentido, los propósitos que extraen su definición de forma progresiva únicamente a través de su propio ejercicio: la práctica.

Procesamiento de información

Como en el caso de la memoria, el procesamiento de la información se encuentra con una alternativa similar. La IA es competente encontrando relaciones dentro de un alto volumen de información aunque sea heterogénea y no estructurada. Funciona bien con un alto nivel de ruido. El ser humano, en cambio, encuentra relaciones cuando el volumen es relativamente bajo y la información está altamente estructurada (Lee, 2020, p. 160).

Esto añade un nuevo nivel de complementariedad, por ejemplo, en el uso de asistentes personales (búsqueda, recomendación, etc). Cada vez estamos más habituados a realizar nuestras tareas diarias siendo asistidos por algoritmos que desbrozan la maleza digital (un terreno con un alto nivel de ruido, tanto por el volumen como por la heterogeneidad) y pueden facilitar así en un segundo momento el análisis y la decisión humana. La situación es paradójica: la digitalización, que ha llenado nuestro entorno con un alud de contenidos de una altísima velocidad de rotación, ofrece a la vez las herramientas para orientarnos en este caos. Por una parte, ha roto el bucle entre acción y observación que es propio de la experiencia humana. Gehlen ya había detectado en los años cincuenta del siglo pasado la “ruptura del círculo de la acción” debido al desarrollo tecnológico, de forma que la percepción del entorno ya no es capaz de guiar la acción y los resultados de la acción a su vez son incapaces de dar lugar a nuevas percepciones (Gehlen, 2004, pp. 42-51 y Postman, 1991, pp. 73-75). Pero, por otra parte, la nueva generación de asistentes con IA ha ayudado a cerrar de nuevo el bucle ayudándonos a obtener sentido a partir del ruido creciente que

caracteriza a nuestros entornos digitalizados.

Esta situación de complementariedad, sin embargo, es ambigua y también admite una lectura más escéptica ya que en cierto sentido hemos empezado ahora a ser sustituidos por nuestros asistentes en un rasgo esencial de nuestra relación con el entorno: no ya en esta o aquella capacidad sino ahora en la capacidad de segundo orden de adquirir experiencia mediante bucles de retroalimentación (*feedback*) entre percepción y acción. Ya no seríamos capaces de tomar decisiones o juzgar sobre la verdad de una región del mundo sin la mediación del sistema de aparatos que percibe el mundo por nosotros (filtra, selecciona, recomienda) y, en definitiva, guía nuestra acción. Este es el núcleo de la crítica de Èric Sadin, que advierte del nuevo régimen de verdad que instaura la nueva generación de asistentes conducidos por IA (Sadin, 2020, pp. 65-70). Esta interpretación deplora la pérdida de control humano sin advertir que la exteriorización de una capacidad siempre implica en cierta medida una transferencia de control humano a los aparatos, que funcionan según sus propias reglas, y, por tanto, una cierta dependencia. El problema con la tesis del control sigue siendo el presupuesto subyacente de un sujeto que controla y, por tanto, se sitúa al margen de los cambios que no sean iniciados por él mismo. Este presupuesto debe ser cuestionado. Parafraseando a Walter Benjamin en *Calle de dirección única* (2010, p. 88), no se trata ya de mantener el control de la técnica sino de controlar la relación entre el ser humano y la técnica. Esto quiere decir que de lo que se trata es de buscar una adaptación del ser humano a los aparatos que no implique subordinación sino que sirva de medio a la realización de las mejores posibilidades de la técnica actual (no previstas por él pero inscritas en los propios dispositivos y disponibles para su uso). Explorar el nuevo papel que le corresponde como sujeto dadas las actuales condiciones técnicas de forma que pueda sacar el máximo partido de estas condiciones para sus propios fines (volveremos

sobre este punto en la sección final sobre coevolución).

Un buen ejemplo de esta relación de complementariedad humano-máquina lo ofrece la cuestión de la “explicabilidad”. La IA actual se comporta como una “caja negra”: es capaz de ofrecer resultados útiles para asistir decisiones humanas pero no es capaz de explicar la validez de esos resultados (Coeckelbergh, 2022, pp. 100-106; Sadin, 2020, pp. 76 ss). Por tanto, aquí el riesgo es que las verdades que ofrece deban tomarse como hechos sin justificación. Esto a su vez genera otro problema, dado que la autoridad que tendemos a concederles a las máquinas tiende a invisibilizar el hecho de que han sido diseñadas por humanos con objetivos humanos y para lograr beneficios para como mínimo algunos humanos. Esto implica una ilusión opuesta a la del control, es decir, la de decisiones que no están bajo control humano, muy conveniente para blindar las decisiones contra posibles quejas y reclamaciones. Es lo que Postman había criticado como “transferencia de responsabilidad” en la década de los noventa en referencia a la aparición del ordenador, al que tendemos a atribuir rasgos antropomórficos (2018, pp. 154-158).

En todo caso es cierto que en la IA el resultado final surge de la agregación de una multiplicidad de cálculos distribuidos en las unidades o “neuronas” que forman la red pero no hay ningún sitio de la red neuronal al que dirigirse en búsqueda de una explicación. Pero esto implica que no puede dejarse al arbitrio de la máquina la decisión última (sobre un diagnóstico médico, la concesión de un crédito o la prima de un seguro) y compete al usuario humano aportar la justificación y la responsabilidad última del uso de la IA.

La actual generación de máquinas está introduciendo un índice de opacidad en la superficie social, donde su uso va en aumento, que hace que cada vez sea menos realista pensar en la posibilidad del control humano (Dyson, 2020, pp. 246-253). La ley de protección de datos a nivel europeo, por ejemplo, sostiene que todos los

ciudadanos tienen derecho a obtener una explicación de las decisiones que les afecten y estén basadas en procesos automáticos (Lee, 2021, p. 99), pero en vista del hermetismo intrínseco de esos procesos la medida es mayormente papel mojado. Sin embargo, aquí de nuevo debe sustituirse la perspectiva del control por la de un reparto equitativo de roles (obtención de resultados, por un lado, y justificación y evaluación de consecuencias, por otro) que permita que ambos agentes trabajen juntos.

Algo parecido pasa con el tema del procesamiento de información, donde hay que distinguir (Lee, 2020, cap. 11) entre *correlaciones* y *relaciones*. La IA es muy eficaz con un volumen muy elevado de datos pero solo indica correlaciones (es decir, coincidencias estadísticas) que no necesariamente corresponden a relaciones (un vínculo causal, por ejemplo). En todo caso, detectar correlaciones puede ser útil si después hay una validación humana que verifique si detrás hay una auténtica relación o se trata de una mera coincidencia. De nuevo aparece aquí la complementariedad.

Por último, se puede indicar otra diferencia en la relación que se establece en cada caso entre los datos y la conclusión. La IA funciona según el modelo del “conocimiento guiado por los datos” (*data-driven knowledge*), que consiste en aprender desde cero a partir de datos (sin la mediación de conceptos). Esto provoca que en algunas tareas su rendimiento sea todavía muy bajo, precisamente allí donde el conocimiento requiere ser encapsulado en diferentes niveles de sentido (conceptos). Esta forma de trabajar de la IA contrasta con la manera en que lo hace la inteligencia humana, que extrae predicciones a partir de información insuficiente (en los ámbitos, por ejemplo, donde no hay la posibilidad de obtener datos), es decir, añade alguna cosa más de lo que hay en los datos de entrada. Por ejemplo, a partir de la coherencia con otros conocimientos previos (como el conocimiento tácito sobre el mundo adquirido por experiencia directa). El conocimiento humano (Russell, 2019, p. 84)

es estratificado (distingue entre capas o niveles de organización abstractos) y acumulativo (cada capa permite acceder a nuevos niveles de abstracción en un proceso progresivo).

Conclusión: el futuro exige sabiduría

Hemos visto algunas diferencias entre el ser humano y la IA que apuntan a explorar una forma de relación complementaria, donde cada parte compensa sus puntos débiles con las fortalezas de la otra. Los rasgos diferenciales de cada tipo de inteligencia no excluyen sino que requieren de su polo opuesto. Esta forma de complementariedad humano-máquina se acerca a lo que en biología evolutiva se conoce como “coevolución”. Aunque existen diferentes formulaciones del concepto -ya sea que indique la relación entre dos especies, entre una especie y su entorno o entre dimensiones distintas dentro de la misma especie (biología y cultura)¹- en todo caso, el núcleo común de la idea señala la dependencia mutua entre ámbitos en principio heterogéneos. Por lo que respecta al proceso de hominización, apunta al bucle de determinación recíproca entre la biología (la anatomía, el sustrato material) y la cultura (la conducta, las posibilidades de hacer o pensar).

Contra el tópico transhumanista (Tegmark, 2018, pp. 41-45) de que la evolución natural queda superada cuando el ser humano alcanza un cierto grado de desarrollo cultural o tecnológico (como si la invención cultural permitiera quedar al margen de las presiones selectivas del medio natural), la coevolución defiende en cambio una

¹ Sobre la aplicación de la idea al caso particular de la hominización véanse, por ejemplo, la teoría de la “superposición” entre biología y cultura (Geertz, 2011, pp. 51-54 y 66-71); la “coevolución gen-cultura” (Wilson, 2012) o el condicionamiento de la transmisión cultural sobre la evolución genética de la especie humana (Boyd, 2018). Debo agradecer a Josep Maria Lloró por darme a conocer estas referencias.

interacción continua entre naturaleza y cultura. No hay una línea temporal de desarrollo sino un bucle.

¿Cuál es la virtualidad crítica de esta idea? En primer lugar, la de acercarse sin simplificar a una realidad compleja, donde las relaciones no se basan en la superación o progreso entre un nivel inferior y uno superior sino en la tensión entre opuestos. En segundo lugar, ayuda a poner distancia ante la idea de una autonomía prometeica del ser humano, capaz de superar todos los límites y condiciones. Idea que segrega una variada tipología de fantasías de control (lo que incluye también a su opuesto, las fantasías distópicas de la amenaza ante la falta de control). Desde el punto de vista de la coevolución el ser humano aparece como un ser situado o sujeto a condiciones: lo que hacemos nos hace. Lo que inventamos para superar las viejas condiciones solo nos libera al precio de sujetarnos a otras nuevas resultado de nuestra acción. Condición y resultado forman un bucle irresoluble.

El ser humano es resultado de sus propios productos o, si se quiere, producto de sí mismo a través de su acción sobre el mundo externo. Es lo que sostiene Michel Serres (2014, pp. 62-68; p. 286) cuando describe el proceso de “autohominización” formado por dos polos: el “exodarwinismo” (la exteriorización adaptativa de capacidades) y la “retroacción” del efecto de los objetos producidos sobre nuestras capacidades. O también Bernard Stiegler (2016, pp. 84-87; p. 102) cuando habla de la “exosomatización” formada por dos momentos de acción recíproca entre exteriorización (técnica) e interiorización (psíquica y colectiva).

La idea supone además una ventaja práctica: contra el optimismo transhumanista (o su contracara, el miedo ante la pérdida de control) apela a la responsabilidad del momento presente ya que pone el acento sobre el siguiente paso a dar en el proceso coevolutivo: dadas las actuales condiciones técnicas ¿cuál es el resultado deseable y a la vez compatible con ellas? Ofrece orientación a un ser humano

constituido en buena medida a partir de la tecnología anterior sobre cómo debe actuar para aprovechar las opciones que le ofrece la más reciente: no tratando de competir con ella (la optimización, que corta el nudo que une los dos polos que forman la paradoja antropológica, reduciendo el ser humano a una parte de sí mismo) sino mediante la complementariedad entre capacidades y funciones heterogéneas (que conserva la tensión entre los dos polos de la paradoja que nos constituye). Es lo que sostiene Edward Lee cuando contrapone la hipótesis coevolutiva, que ha de permitir potenciar la propia humanidad a través del uso de la IA, al “creacionismo digital” (2020, p. 311), que supone una relación jerárquica, ya sea para aumentar nuestro control sobre la IA o para limitar el suyo (a través de una legislación más restrictiva o de manifiestos en favor de una IA ética, por ejemplo).

Para finalizar, vamos a intentar señalar qué aplicación a la educación puede tener la tesis de una coevolución entre humanidad y técnica. En primer lugar, una evidente a la luz de lo anterior: el ser humano debe especializarse en su propia humanidad. Esto significa, en primer lugar, ejercitar y reforzar las *supercompetencias* humanas y aprender a utilizar las competencias técnicas para resolver problemas de forma conjunta. Es decir: aprender a combinar el consumo de información disponible con la producción de conocimiento nuevo en la interacción entre la cabeza y la mano (rediseño, contextos de prueba y error); aprender a explorar el sentido de un problema o de una pregunta, no solo a producir soluciones correctas; aprender a utilizar los asistentes como una herramienta heurística que requiere después un trabajo de comprensión y síntesis insustituible; aprender a dar cuenta de lo que hacemos (por qué), no solo a seguir procedimientos (cómo) o producir resultados (qué).

En segundo lugar, si es cierto que “las habilidades y los contenidos” han quedado obsoletos en el ámbito de la educación,

dado que las máquinas los gestionan mejor que los humanos (Lee, 2020, p. 309), entonces ¿cuál ha de ser el objetivo de la educación? Una posible respuesta es que cada vez será más importante el estudio de la tecnología como forma de entender la sociedad (Lee, 2020, p. 310): se tienen que estudiar tendencias y no hechos, que pueden cambiar de un día para otro. Por ejemplo, las fuerzas ocultas que amenazan la privacidad, la divulgación viral de ideas o una historia reciente del mundo digital que evite la aceptación acrítica de la tecnología como un hecho incuestionable.

Se podría añadir la necesidad de un humanismo digital. El mundo se llena cada vez más de fenómenos que toman la apariencia de hechos, cerrados sobre sí mismos y aislados unos de otros, lo que impide captar su interacción y, por tanto, una comprensión adecuada e integral. El estudio de las humanidades, en un sentido generalista, puede ser hoy el mejor asistente de la humanidad en su relación con un mundo cada vez más extraño o ajeno. Los hechos que se presentan como independientes de nuestra intervención pueden adquirir un nuevo sentido cuando se conectan con otros, pasados o presentes, dibujando así un marco más amplio donde el mundo recupera su inteligibilidad y nosotros nuestra capacidad de intervención (ya que entendemos que el mundo no es independiente de nuestra acción). Así, Postman defendía una educación centrada en la coherencia para evitar el consumo pasivo de verdades como hechos, para lo que recomendaba poner énfasis en la historia. Los contenidos de cualquier asignatura debían incorporar un tratamiento de la historia de la propia disciplina (2018, pp. 247-257).

Por tanto, no que la técnica sea el asistente entre el ser humano y la naturaleza sino que la cultura (sin especificación de ámbitos, en lo que todos ellos tienen en común de reserva de significado y, por tanto, de potencial de interpretación de nuevos hechos) se convierta ahora en el asistente entre el ser humano y un medio técnicamente configurado. Solo así, a través de esta operación de segundo grado

(mediación cultural de la mediación técnica), podremos tal vez volver a cerrar, no al nivel del individuo sino del colectivo (y, en el límite, de la especie), el bucle entre actuar y comprender, fracturado debido a la disrupción tecnológica.

El futuro exige sabiduría, como sostiene Helga Nowotny (2021, pp. 126-133). Los retos actuales requieren mirar a la vez hacia delante y hacia detrás, vincular el futuro y el pasado, atender a la urgencia de los problemas a corto plazo y mantener una visión general, lo que incluye *saber releer el patrimonio cultural* a la luz de los problemas presentes y viceversa, en una iluminación recíproca.

La sabiduría es saber escoger. Hay que distinguir lo que es verdad de lo que no lo es y también lo que es eficaz de lo que no lo es, pero como mínimo igual de importante es saber apreciar lo que tiene valor y rechazar lo que no lo tiene. En un contexto de alta volatilidad de contenidos se vuelve difícil la previsión del futuro, incluso a muy corto plazo, por lo que ante la incapacidad de saber qué nos será útil más adelante se impone la tentación de conservarlo todo. Cuando el futuro falla como criterio de selección del pasado, como actualmente, se vuelve urgente *saber buscar otros criterios de selección* que nos ayuden a distinguir lo que es necesario de lo que es superfluo. Tener sabiduría, por último, también significa resistirse a la tentación del control y mostrar una cierta apertura para aprender de la nueva tecnología. Contra el miedo a la sustitución humana por la IA y la competencia por la primacía como única opción, hay que *saber complementar mutuamente las carencias y las fortalezas del ser humano y la tecnología*. Dice mucho de la dificultad y el volumen de los retos a los que tenemos que hacer frente actualmente que el objetivo de la educación tenga que ser, ni más ni menos, que el de *enseñar la sabiduría*.

Referencias

- Aristóteles (2001). *Ética a Nicómaco*. Alianza.
- Benjamin, W. (2005). *Libro de los Pasajes*. Akal.
- Benjamin, W. (2010). *Obras - Libro IV/vol. I*. Abada.
- Bodei, R. (2022). *Dominio y sometimiento. Esclavos, animales, máquinas, inteligencia artificial*. Alianza.
- Boyd, R. (2018). *Un animal diferente. Cómo la cultura transformó nuestra especie*. Oberon.
- Coeckelbergh, M. (2022). *Ética de la inteligencia artificial*. Cátedra.
- De Carolis, M. (2004). *La vita nell'epoca della sua riproducibilità tecnica*. Bollati Boringhieri.
- De Carolis, M. (2017). *La paradoja antropológica. Nichos, micromundos, disociación psíquica*. Quadrata.
- Delgado, J. L. (2016). *El origen del presente. La filosofía de la historia de Walter Benjamin*. Sequitur.
- Descola, P. (2005). *Par-delà nature et culture*. Gallimard.
- Dyson, G. (2020). *Analoga. The entangled destinies of nature, human beings and machines*. Penguin.
- Elliott, A. (2018). *The culture of AI. Everyday life and the digital revolution*. Routledge.
- Geertz, C. (2011). *La interpretación de las culturas*. Gedisa.
- Gehlen, A. (2004). *Die Seele im technischen Zeitalter. Gesamtausgabe. Band 6*. Klostermann.
- Lee, E. A. (2020). *The coevolution. The entwined futures of humans and machines*. MIT Press.
- Lipovetsky, G. y Serroy, J. (2015). *La estetización del mundo*. Anagrama.
- Nowotny, H. (2021). *La fe en la inteligencia artificial. Los algoritmos predictivos y el futuro de la humanidad*. Galaxia Gutenberg.

- Ong, W. J. (2016). *Oralidad y escritura: tecnologías de la palabra*. FCE.
- Postman, N. (1991). *Divertirse hasta morir. El discurso público en la era del "show business"*. Tempestad.
- Postman, N. (2018). *Tecnópolis. La rendición de la cultura a la tecnología*. El Salmón.
- Rappaport, R. A. (2016). *Ritual y religión en la formación de la humanidad*. Akal.
- Rehberg, K.S. (1983). Nachwort des Herausgebers. En A. Gehlen, *Gesamtausgabe (Band 4) - Philosophische Anthropologie und Handlungslehre* (pp. 385-402). Klostermann.
- Russell, S. (2019). *Human compatible. Artificial intelligence and the problem of control*. Viking.
- Sadin, È. (2020). *La inteligencia artificial o el desafío del siglo*. Caja negra.
- Sanmartí, N. (2020). *Evaluar y aprender: un único proceso*. Octaedro.
- Sennett, R. (2009). *El artesano*. Anagrama.
- Serres, M. (2014). *Hominescence*. Le Pommier.
- Simondon, G. (2008). *El modo de existencia de los objetos técnicos*. Prometeo.
- Snell, B. (2007). *El descubrimiento del espíritu. Estudios sobre la génesis del pensamiento europeo en los griegos*. Acantilado.
- Stiegler, B. (2002). *La técnica y el tiempo (vol. II)*. Hiru.
- Stiegler, B. (2016). *Dans la disruption. Comment ne pas devenir fou?* Babel.
- Tegmark, M. (2018). *Vida 3.0. Qué significa ser humano en la era de la inteligencia artificial*. Taurus.
- Uexküll, J. von (2016). *Andanzas por los mundos circundantes de los animales y los hombres*. Cactus.
- Virilio, P. (1998). *La máquina de visión*. Cátedra.
- Wilson, E. O. (2012). *La conquista social de la tierra*. Debate.

**DESAFÍOS EPISTEMOLÓGICOS EN LA ERA DE LAS REDES
NEURONALES ARTIFICIALES: ABORDANDO SISTEMAS COMPLEJOS
DESDE UNA PERSPECTIVA COMPUTACIONAL**

***EPISTEMOLOGICAL CHALLENGES IN THE ERA OF ARTIFICIAL NEURAL
NETWORKS: ADDRESSING COMPLEX SYSTEMS FROM A
COMPUTATIONAL PERSPECTIVE***

ERICK RUBIO

erickrubio@mail.uniatlantico.edu.co

Universidad del Atlántico, Colombia - Grupo FiloComplex,
Sociedad Argentina de Análisis Filosófico, Argentina

LEANDRO GIRI

leandrogiri@gmail.com

Universidad Nacional de Tres de Febrero - Consejo Nacional de
Investigaciones Científicas y Técnicas (CONICET) - Grupo
FiloComplex, Sociedad Argentina de Análisis Filosófico,
Argentina

ANDRÉS ILCIC

ailcic@ffyh.unc.edu.ar

Instituto de Investigaciones Filosóficas de la Sociedad Argentina
de Análisis Filosófico, Consejo Nacional de Investigaciones
Científicas y Técnicas (IIF-SADAF-CONICET) - Grupo
FiloComplex, Sociedad Argentina de Análisis Filosófico,
Argentina

RECIBIDO: 22/08/2023

ACEPTADO: 24/11/2023

Resumen: En este artículo se explora un enfoque “manipulacionista” en la investigación científica, utilizando redes neuronales artificiales para abordar

sistemas complejos. Se destaca la capacidad de estas redes para resolver problemas y representar patrones de comportamiento en estos sistemas. Además, se enfatiza la importancia de superar las limitaciones de los métodos analíticos tradicionales mediante métodos computacionales y la visualización de grandes cantidades de datos. El artículo toma como ejemplo significativo el problema del plegamiento de proteínas y el éxito de AlphaFold2 en este campo. Se plantea la necesidad de explorar las implicaciones epistemológicas más amplias de las simulaciones computacionales en la investigación científica, considerando la relación entre precisión predictiva y comprensión de los sistemas complejos estudiados. La estructura del trabajo se divide en cuatro secciones, que incluyen el estudio de sistemas complejos, el problema del plegamiento de proteínas y una discusión filosófica sobre los modelos de simulación computacional y las redes neuronales artificiales; además de una sección de comentarios finales. En conjunto, se enfatiza la relevancia de este enfoque para abordar sistemas complejos y su impacto en la filosofía de la ciencia.

Palabras clave: sistemas complejos; redes neuronales artificiales; AlphaFold2; manipulacionismo; transparencia epistémica

Abstract: The article delves into a “manipulationist” approach in scientific research, utilizing artificial neural networks to tackle complex systems. It highlights the networks' capacity to resolve problems and portray behavioral patterns within these intricate systems. Moreover, it underscores the significance of surmounting the limitations posed by traditional analytical methods through computational techniques and the visualization of vast amounts of data. The article specifically focuses on the issue of protein folding and mentions the achievements of systems like AlphaFold2 in this domain. It highlights the necessity to explore the broader epistemological implications of computational simulations in scientific research, taking into account the interplay between predictive accuracy and the comprehension of the studied complex systems. The work is structured into four sections, encompassing the study of complex systems, the problem of protein folding, a philosophical discussion on computational simulation models and artificial neural networks, and concluding with final remarks. Overall, it underscores the significance of this approach in addressing complex systems and its profound impact on the philosophy of science.

Keywords: Complex systems; artificial neural networks; AlphaFold2; manipulationism; epistemic transparency

Introducción

En este artículo se explora la posibilidad de un abordaje metodológico científico de tipo «*manipulacionista*¹», en el que las técnicas de experimentación computacionales puedan ser una herramienta fructífera para el abordaje de *sistemas complejos*. En esta dirección, se examinará la capacidad de las simulaciones computacionales para resolver problemas concretos, a partir de la identificación de patrones de comportamiento de sistemas complejos y su representación científica.

¹ Es importante señalar que, si bien el término suele hacer referencia a cierta tradición en filosofía de la ciencia acerca de las explicaciones causales, con James Woodward a la cabeza, en el desarrollo de este trabajo no se hace mención, salvo en esta referencia, a James Woodward, tratando de tomar cierta distancia del problema de la explicación científica, y sobre todo de la causalidad. En el estudio de los sistemas complejos, dada sus características, no parece ser una estrategia favorable intentar un abordaje ligado a la causalidad. Piénsese, por ejemplo, en los impulsos nerviosos espontáneos que se dan en nuestro cerebro: el análisis de la causación interniveles genera enormes dificultades para dar cuenta de este fenómeno y otros por el estilo. El manipulacionismo de Woodward está pensado básicamente dentro de un contexto experimental empírico -de laboratorio-, donde se podría variar sistemáticamente las condiciones iniciales para obtener un mayor convencimiento de que ciertas condiciones son capaces de hacer aparecer un resultado particular, y ese proceso muestra que son esas y no otras las condiciones que lo causan. Woodward (p.e. 2003) está considerando el auge de la biología molecular a partir del desarrollo de técnicas e instrumentos para secuenciar y manipular por ejemplo moléculas *concretas* de ADN. A partir de una manipulación controlada de esta molécula se pueden obtener explicaciones causales, por ejemplo, de ciertas enfermedades. En cambio, como se verá, el manipulacionismo de nuestro enfoque actual sería principalmente de carácter *virtual*, algorítmico, abriendo incluso fuertemente la posibilidad a un automanipulacionismo que se distanciaría aún más de la propuesta woodwardiana. En el entrenamiento de una red neuronal podría modificarse el modelo subyacente con el que se estaría intentando abordar la dinámica del sistema.

Debido a sus características, los sistemas complejos difícilmente pueden ser tratados ni satisfactoria ni exclusivamente por métodos analíticos tradicionales, pues los mismos solo tienen un alcance limitado y, en el mejor de los casos, éxito muy parcial (Lenhard, 2016; Vicsek, 2002). El tipo de aproximación que ha logrado tratar con cierto éxito a los sistemas complejos está basado principalmente en el uso intensivo e imprescindible de métodos computacionales de diversa clase (Vicsek, 2002) y la representación visual de enormes cantidades de datos (Lenhard, 2017). Un análisis de lo anterior arroja que se debe renunciar, al menos parcialmente, a las pretensiones epistémicas normativas (y, quizás en algunos casos, incluso a las de comprensión) para poder tratar adecuadamente estos fenómenos (Lenhard, 2016). ¿Qué se entiende aquí por “éxito”? Como defendemos en este trabajo, no se trata de *transparencia epistémica* ni de entendimiento, mucho menos en términos del establecimiento de teorías generales; sino más bien de la solución de problemas particulares en contextos determinados, sobre los que se logra una comprensión parcial por la articulación de modelos distintos, cada uno de los cuales enfatiza algunas de las propiedades del sistema complejo bajo estudio.

En los últimos tiempos, las *redes neuronales artificiales* han sido objeto de discusión en la filosofía de la ciencia (Gao & Ganguli, 2015; Churchland & Sejnowski, 2016; Paninski & Cunningham, 2018; Chirimuuta, 2020), al ser consideradas como una forma de simulación computacional para abordar sistemas complejos; a veces como modelo propiamente y en otros casos como una técnica de procesamiento de datos. La dinámica de activación en las neuronas artificiales simularía (imitaría rudimentariamente), por ejemplo, la dinámica de las neuronas biológicas (Rumelhart et al., 1986). Suponiendo esto, lo que se buscaría básicamente sería extrapolar algunos resultados en el estudio dinámico de uno de los sistemas hacia el otro sistema en consideración, y la extrapolación se haría a

partir de los resultados obtenidos del empleo de ordenadores en el abordaje dinámico de uno de los sistemas (Hartmann, 1996). Este proceder, justamente en el terreno del que sea quizá el sistema complejo por excelencia, a saber: nuestro cerebro, está abriendo flancos promisorios de ataque para avanzar un poco más en su comprensión. Por ejemplo, Timothy P. Lillicrap, junto con Geoffrey Hinton, entre otros, en un interesante artículo titulado ‘Backpropagation and the brain’ (Lillicrap et al., 2020), exploran la posibilidad de que en ciertas áreas de la corteza cerebral se pueda dar una suerte de implementación de uno de los algoritmos más destacados en el entrenamiento de las redes neuronales artificiales como lo es el algoritmo de retropropagación del error.

Además, estas redes tienen la capacidad de “aprender” y adaptarse continuamente a partir de la experiencia, lo que las convierte en herramientas poderosas para la resolución de problemas de gran complicación, como el plegamiento de las proteínas, donde sistemas como *AlphaFold2* ha dado un gran avance recientemente.

En la literatura filosófica, se ha generado una controversia acerca de la justificación de las redes neuronales artificiales como herramientas para la investigación científica y la evaluación de sus resultados (Gao & Ganguli, 2015; Churchland & Sejnowski, 2016; Paninski & Cunningham, 2018). Algunos autores defienden su uso argumentando que proporcionan resultados precisos y útiles (Rumelhart et al., 1986; Lillicrap & Kording, 2019), mientras que otros cuestionan si son capaces de ofrecer una explicación adecuada de los fenómenos investigados (Gao & Ganguli, 2015; Pearl & Mackenzie, 2018).

En este contexto, Mazviita Chirimuuta (2020), por ejemplo, ha señalado que la investigación en decodificación —del cerebro en general y del neocórtex en particular— en la neurociencia

computacional revela un balance entre la precisión *predictiva*² y la capacidad de los modelos por simulación computacional —más concretamente por redes neuronales artificiales— para conferir comprensión. Según esta perspectiva, el progreso en la precisión predictiva no se traduce en un avance en la comprensión del sistema complejo objeto de estudio; en su caso, en una comprensión del cerebro. Ante la crítica de que el aprendizaje automático no produce un nivel aceptable de comprensión, algunos neurocientíficos (Bashivan et al., 2019, p. 1) sugieren redefinirla como la capacidad de predecir y controlar: en este caso, al cerebro. Chirimuuta defiende una noción no-factivista de la comprensión científica, que considera que los modelos reducen la complejidad natural a un tamaño manejable por los humanos mediante la abstracción e idealización.³ En este marco, la inteligibilidad de los modelos se ve comprometida cuando se despliegan algoritmos como las redes neuronales artificiales.

En otras palabras, la discusión acerca del valor de la comprensión como fin en sí mismo, no solo como medio hacia la predicción y el control, se presenta en el contexto del compromiso descrito por Chirimuuta.⁴ En definitiva, se plantea la necesidad de explorar las implicaciones de las simulaciones computacionales en la

² Sobre lo que podría considerarse el ‘giro predictivo’, particularmente en neurociencias, ver, por ejemplo: Weiskopf (2022)

³ Para un tratamiento comprehensivo del problema de la abstracción e idealización de los modelos científicos ver, por ejemplo: Frigg (2022, cap. 11).

⁴ La idea de comprensión suele estar asociada a la idea de encontrar, para los sistemas bajo estudio, los mecanismos causales (Durán, 2018). Sin embargo, para el caso de modelos que se construyen mediante enormes bases de datos y algoritmos de inteligencia artificial capaces de procesarlos en busca de correlaciones, se ha argumentado (Napoletani, Panza & Struppa, 2011; Mayer-Schomberg & Cukier, 2013) que la investigación correlacional debe reemplazar a la causal. En el presente trabajo, por motivos que iremos justificando durante su desarrollo, exploramos esta vía de razonamiento para los sistemas complejos y su análisis mediante simulaciones por redes neuronales.

investigación científica desde una perspectiva epistemológica más amplia; una perspectiva que tome fuertemente en consideración a los sistemas complejos.

El trabajo consta de cuatro secciones. En la primera se aborda el estudio de los sistemas complejos, que resulta inherentemente complicado por el surgimiento de propiedades emergentes, las cuales acarrearán dificultades epistemológicas y metodológicas para su abordaje predictivo y explicativo. También se trabaja sobre los fenómenos de autoorganización, en tanto hacen a otra característica típica de estos sistemas.

En la segunda se aborda el problema del plegamiento de proteínas y el avance significativo en su resolución logrado por AlphaFold, un sistema de redes neuronales artificiales desarrollado por la compañía DeepMind. El plegamiento de las proteínas es un desafío científico importante que ha interesado a investigadores en áreas como la biología, la bioquímica y la biomedicina durante décadas, y que nos servirá aquí de ejemplar paradigmático para señalar el modo en que la manipulación de parámetros relevantes de la dinámica del sistema a través de simulaciones “no clásicas” permite la identificación de patrones y correlaciones complejas difíciles de hallar por métodos analíticos y experimentales tradicionales.

En la tercera sección se lleva a cabo una discusión filosófica sobre los modelos de simulación computacional y las redes neuronales artificiales. Se argumenta que estos métodos de simulación, especialmente las redes neuronales, son herramientas poderosas para abordar sistemas complejos con una gran cantidad de elementos, pues permiten explorar las posibles consecuencias en las dinámicas de estos sistemas al alterar parámetros relevantes. Ello, argumentamos, conlleva consecuencias epistemológicamente relevantes que exploraremos con cierto detalle. Finalmente, presentaremos algunas conclusiones.

Aproximación a los sistemas complejos

Como suele ocurrir a la hora de dar una definición concreta del objeto que estudia un nuevo campo de actividad científica, se tiende a partir de definiciones poco concretas, con términos que tienen muchos referentes posibles, algunos incluso con la misma dificultad definicional que el campo que se quiere delimitar. Esto es lo que sucede al aproximarse al campo de los sistemas complejos. Una primera aproximación que responde a muchas de las inquietudes de quienes dieron forma al campo ha sido el estudio de aquellos sistemas que exhiben características emergentes no triviales y comportamientos autoorganizados (Mitchell, 2009, p. 13). Ahora bien, a la hora de caracterizar qué constituye una propiedad emergente o un comportamiento autoorganizado, se tiende rápidamente a brindar una suerte de definición negativa, es decir, una por la vía de “aquello que no”:

[Los] sistemas en los que los comportamientos organizados surgen sin un regulador o líder interno o externo son algunas veces llamados sistemas autoorganizados. Debido a que simples reglas producen comportamientos complejos *en cierta manera difíciles de predecir*, el comportamiento macroscópico de estos sistemas es algunas veces llamado emergente. (Mitchell, 2009, p. 13. nuestro énfasis)

Aunque aún no parece haber consenso sobre cómo definir a estos sistemas, sí se han logrado identificar ciertos rasgos distintivos, muchos de ellos en términos de la incapacidad de someterse a cierto tipo de descripción o de metodología de resolución de problemas característicos o paradigmáticos de los contextos de producción de conocimiento sobre otras clases de sistemas o fenómenos. Aquí nos

concentramos en dos aspectos íntimamente relacionados: la emergencia y la autoorganización.

Emergencia

El fenómeno de la complejidad suele ser visto como uno fuertemente asociado al de la emergencia (Bunge, 2004; Mitchell, 2009, 2004; Gell-Mann, 1998; Taylor, 2001); fenómeno en el que surgen propiedades globales irreducibles explicativamente a las propiedades de los (muchísimos) componentes (considerados bajo cierto nivel de organización) de un sistema, que dan lugar a ellas mediante sus múltiples interacciones *no lineales*.⁵

Las propiedades emergentes pueden tomarse entonces como propiedades de un objeto complejo que ni sus constituyentes ni sus precursores poseen (Bunge, 2004, p. 20; Johnson, 2003). Ello no quiere decir que sean propiedades que no se puedan abordar a partir de los componentes del sistema o su trayectoria. Más bien, no se pueden explicar a partir únicamente de las propiedades de sus partes constitutivas, puesto que se requiere conocer el tipo de interacción que están teniendo dichas partes, y, sobre todo, el conjunto de valores paramétricos –e *hiperparamétricos*⁶– que bajo estas

⁵ La no linealidad se refiere aquí a una relación entre variables donde el cambio en una de ellas no produce un cambio de magnitud proporcional en la otra. Es decir, no siguen una relación directamente proporcional. En cambio, la relación puede ser compleja y extremadamente difícil de predecir.

⁶ La noción de *hiperparámetro* juega un rol muy importante en el tipo de enfoque aquí propuesto; a saber, uno basado principalmente en redes neuronales artificiales. En este sentido debemos diferenciar la estructura hiperparamétrica del modelo (cantidad de capas, cantidad y tipo de neuronas, etc.), de la estructura de parámetros que surge (o emerge) luego del entrenamiento del modelo. Por citar algunos ejemplos: para identificación o clasificación de imágenes, se modelan redes neuronales convolucionales; para predicciones en series de tiempo, redes recurrentes; para procesamiento de textos, redes recurrentes o Transformers; todas ellas con una estructura hiperparamétrica particular muy diferente y que facilita

interacciones no lineales presentan algunas variables consideradas relevantes para la dinámica del sistema complejo en estudio. Es ahí donde radica una dificultad esencial en el análisis de los sistemas complejos: en disponer de este conjunto de interacciones y de los valores paramétricos que bajo tales interacciones presentan ciertas variables implicadas en la descripción de la dinámica del sistema, ya que estos suelen ser sistemas dinámicos no deterministas, que, bajo ciertas condiciones, algunas veces favorecidas por el azar, presentan nuevas asociaciones o combinaciones de sus elementos, lo que conlleva a que el conjunto de interacciones varíe. Esta variación en sus interacciones viene acompañada regularmente de nuevos comportamientos que suelen dar lugar a nuevas propiedades del sistema (motivo por el cual muchos de ellos se dicen adaptativos (Holland, 1992). Ello podría llevarnos a considerar que el quid de las propiedades emergentes está en las interacciones no lineales y en los valores de ciertos parámetros –e hiperparámetros, según el enfoque basado en redes neuronales artificiales aquí propuesto–, que bajo tales interacciones presentan ciertas variables ligadas a los elementos constituyentes de estos sistemas complejos; propiedades que no son distributivas, sino globales (Bunge, 2004, p. 29).

Las propiedades globales antedichas pueden ser, por ejemplo, de tipo *funcional* –como la que se desprende de la estructura tridimensional de las proteínas–, *estructural* –como el procesamiento de información de tipo jerárquico en el neocórtex– u *organizacional* –como la que presenta una bandada de estorninos (*Sturnus vulgaris*) en busca de alimento o en defensa de otras aves

enormemente el proceso de aprendizaje del sistema (ver Feurer & Hutter, 2019; Yang & Shami, 2020). El enfoque que desarrollamos más adelante en este trabajo sugiere que los hiperparámetros pueden verse como “parámetros de parámetros” con los cuales se puede ganar conocimiento del sistema gracias a crear una representación de alto nivel que permite manipular al menos parcialmente la dinámica del sistema a fin de encontrar propiedades robustas para su descripción.

rapaces (Rescher, 1998, p. 9). El conocimiento, aun detallado y preciso, de las propiedades de los componentes de un sistema vistos de manera aislada nos resulta insuficiente para poder dar cuenta de estas propiedades en sentido global. Esto sucede pese a considerar que es a partir de dichos componentes (considerados bajo cierto nivel de organización o abstracción), combinados o asociados, que dichas propiedades tienen lugar, y que es a partir de la interacción de estos que se pueden definir y detectar casos específicos de dichas propiedades globales.

Las mencionadas propiedades del sistema bien sean de tipo funcional, estructural u organizacional, requieren de un gran conocimiento adicional al que se tenga de las propiedades de los componentes para intentar dar cuenta de ellas; además del desarrollo de técnicas y estrategias no analíticas, sino más bien aproximativas, implementadas iterativamente y retroalimentadas mediante muchísimo ensayo y error. La emergencia, característica notoria de la complejidad, siendo principalmente un fenómeno ontológico (en el sentido de Humphreys, 1997, 2016), representa serias dificultades epistemológicas, particularmente en sentido metodológico, para abordarla predictiva y explicativamente. Sin embargo, a propósito de la diferenciación entre estructura hiperparamétrica y paramétrica de un sistema de aprendizaje automático, podemos afirmar que, si bien es cierto que las propiedades globales emergentes no pueden explicarse ni preverse a priori partiendo de la estructura paramétrica que será aprendida por el sistema (dado que, en la práctica, es imposible prever cuáles serán los parámetros que surgirán de la iteración), creemos que sí es posible prever ciertas características de las propiedades globales a partir de la estructura hiperparamétrica del modelo.

De acuerdo con cómo estructuramos los hiperparámetros del sistema (la estructura marco previa a la iteración de aprendizaje), podemos tener cierta noción de las propiedades globales emergentes

que esperamos observar. Por ejemplo, si utilizamos redes neuronales artificiales de tipo convolucional, sabremos que las propiedades globales (o regularidades aprendidas por el sistema) se centrarán en la detección de formas o disposición/relación espacial de/entre los datos de entrada (Feurer & Hutter, 2019). En cambio, si diseñamos una red de tipo recurrente (al estilo Hopfield, una LSTM o incluso un Transformer), las regularidades aprendidas se centrarán en el orden u ocurrencia temporal de los datos. Incluso podemos diseñar este tipo de redes de manera tal que la cantidad de regularidades o propiedades aprendidas podamos conocerla de antemano (por ejemplo, utilizando una última capa densa con una cantidad de neuronas igual a la cantidad de propiedades que estimamos predecir).

Podemos también pensar que si lo que deseamos obtener son propiedades de tipo estructural jerárquico, una estructura hiperparamétrica de redes convolucionales con *pooling* podría ayudar en dicho propósito (dado que van “abstrayendo” en cada capa *pool* las regularidades más relevantes o generales, ver Feuerer & Hutter, 2019).

Por tanto, si bien no podemos predecir o explicar con total transparencia el porqué de las propiedades globales emergentes dada la estructura del sistema, sí podemos prever ciertas características esperadas de las mismas. Volveremos a este importante punto más adelante.

Autoorganización

Cuando se habla de sistemas autoorganizados se está pensando en sistemas que adquieren cierto orden sin el control o la dirección central de mando. El concepto de orden no es fácil de caracterizar; sin embargo, una manera de acercarnos a él puede ser mediante el concepto de información.

‘Información’, en los sistemas complejos, y en general en los sistemas físicos, es opuesto a ‘entropía’ (Hidalgo, 2015, p. 24). En ese sentido, diremos que los sistemas complejos, centrándonos en este momento en el orden que exhiben, son tanto peculiares como poco comunes (Hidalgo, 2015, p. 23), ya que sus muchísimos elementos constituyentes presentan disposiciones físicas con un altísimo grado de información, cuyas rutas de acceso para llegar a ellas, ateniéndose a las constricciones que operan sobre tales elementos, son extremadamente limitadas. Estos estados ricos informacionalmente se dan en extremadamente muy pocos microestados, por lo que la entropía de estos estados es bastante baja.

Pensemos, por ejemplo, en una proteína. Sabemos que es una configuración espacial específica de su secuencia de aminoácidos la que la hace funcional. El espectro de configuraciones o plegamientos posibles, dadas las constricciones fisicoquímicas que operan sobre los aminoácidos constitutivos de una proteína, es extremadamente amplio (Levinthal, 1968). Sin embargo, de todas esas posibles configuraciones, o de todas esas posibles maneras en las que podrían plegarse los aminoácidos encadenados, la funcionalidad de la proteína está reducida a una configuración particular. Decimos entonces que es esa configuración particular la que contiene un alto contenido de información. Es rica en información con respecto al fondo de las conformaciones no funcionales.

Aquí, el concepto de complejidad efectiva resulta de gran utilidad. Murray Gell-Mann lo define como “la longitud de una descripción concisa de las regularidades [de un sistema]” (Gell-Mann, 1998, p. 66). Es decir, se está pensando en una descripción comprimida de las relaciones conspicuas que los elementos constituyentes del sistema presentan en esos estados ricos informacionalmente. Se está pensando en la descripción más concisa del conjunto de patrones identificables entre las interacciones que estos elementos presentan.

Podemos decir entonces que los sistemas que exhiben comportamientos autoorganizados son sistemas que presentan una disposición de sus elementos constitutivos con alta complejidad efectiva y baja entropía; son sistemas cuyos elementos constituyentes adquieren cierta disposición física de manera espontánea con un alto contenido de información. El comportamiento autoorganizado de sistemas complejos es uno en cuya disposición de los elementos que dan lugar a él podemos encontrar ciertos patrones, pero también cierta dosis de aleatoriedad. Curiosamente, estos estados se suelen dar en lo que se conoce como el borde del caos (Nicolis & Prigogine, 1989). Y curiosamente también, ahí encuentran una suerte de equilibrio entre la flexibilidad y la rigidez de sus comportamientos que les resultan indispensables para adaptarse (Kauffman, 1993).

Complejidad y redes neuronales artificiales: el problema del plegamiento de proteínas y el gran adelanto en su solución por parte de AlphaFold

El problema del plegamiento de las proteínas ha interesado sobremanera por décadas a los científicos en el área de la biología, la bioquímica, la biomedicina, etc. Tanto así, que desde el año 1994 ha venido realizándose un concurso (*Critical Assessment of Techniques for Protein Structure Prediction -CASP-*), donde se premia al grupo o centro de investigación que logre *predecir* la estructura o configuración precisa de un amplio número de proteínas a partir de su secuencia de aminoácidos. Hasta el momento, el concurso se ha llevado a cabo en quince ocasiones. En dos de ellas, (CASP13: 2018 y CASP14: 2020), ha participado DeepMind, que es un laboratorio de inteligencia artificial de enorme prestigio mundial, fundado en 2010 y adquirido por Google en 2014. En ambas

ocasiones resultó ganador,⁷ y con un buen margen de ventaja con respecto a otros equipos que emplearon otras aproximaciones numéricas al plegado de proteínas. De hecho, el margen obtenido en su segunda participación -con AlphaFold2- fue abrumadoramente amplio. Según Venki Ramakrishnan, Premio Nobel de química y presidente de la Royal Society

Este trabajo computacional representa un avance sorprendente en el problema del plegamiento de proteínas, un gran desafío de 50 años en biología. Han transcurrido décadas antes de lo que las personas en el campo habían estimado. Será emocionante ver las muchas formas en que cambiará fundamentalmente la investigación biológica (Ramakrishnan, citado en Jumper et al., 2020)

El problema del plegamiento de las proteínas es importante porque las proteínas son moléculas esenciales para la vida y juegan un papel crítico en muchos procesos biológicos, incluyendo la regulación de la actividad celular, la comunicación celular, la defensa contra enfermedades y la producción de energía (Evans et al., 2018). La forma tridimensional en la que se pliegan las proteínas determina su función biológica, y las proteínas mal plegadas pueden ser la causa de una amplia variedad de enfermedades, incluyendo neurodegenerativas como el Alzheimer, el Parkinson, o el mal de Huntington (Suárez, 2005, p. 243). Además, la resolución del problema del plegamiento de las proteínas también es importante desde una perspectiva científica más amplia. Es uno de los problemas más importantes y difíciles de la biofísica y la biología molecular, y su solución puede tener implicaciones importantes en

⁷ En el sitio web de la conferencia (predictioncenter.org) se pueden comparar los modelos presentados por los distintos equipos a lo largo de la historia de la competición. CASP15 se llevó a cabo durante 2022 y los resultados mejoraron considerablemente con respecto a los años anteriores, en su mayoría tras la aplicación de métodos de aprendizaje automático similares a las arquitecturas introducidas por AlphaFold 1 y 2.

muchos campos, incluyendo la medicina, la ingeniería de proteínas y la biotecnología.

La red neural artificial de aprendizaje profundo utilizada por el equipo de DeepMind para obtener un modelo tridimensional de la proteína, a partir de la secuencia de aminoácidos que la constituye, se conoce como AlphaFold. La misma es heredera de las capacidades ingenieriles de la empresa, que ya había ostentado un éxito abrumador con las versiones AlphaGo y AlphaZero, desarrolladas para jugar al Go y al ajedrez respectivamente: juegos en los que se evidencia que una exploración de todas las posibles partidas para determinar estrategias conducentes a una victoria resulta imposible. El abanico de posibilidades de combinaciones de movimientos es prácticamente infinito (Evans et al., 2018). Al igual que en AlphaFold, claramente, la búsqueda no puede ser exhaustiva. La misma se realiza mediante un árbol de búsqueda de Markov;⁸ esto es, un árbol de búsqueda aleatorio. Si bien se realiza con cierta dosis de aleatoriedad, esta es guiada con la intervención del algoritmo del descenso del gradiente⁹ (*gradient descent*), junto al de *backpropagation* (Nielsen, 2015) o retropropagación del error. Este

⁸ En un árbol de búsqueda de Markov, se representa un problema de toma de decisiones secuenciales, donde cada nodo del árbol representa un estado del sistema y cada arista representa una acción que se puede tomar en ese estado. Además, cada nodo contiene información sobre la recompensa asociada con ese estado y la probabilidad de transición a otros estados. El árbol de búsqueda de Markov se construye de manera recursiva, explorando el espacio de posibles acciones y estados futuros a partir del estado actual. Una de sus ventajas es su capacidad para modelar y resolver problemas complejos de toma de decisiones secuenciales en entornos inciertos. (Ver, por ejemplo, Puterman, 1990)

⁹ Más concretamente, el algoritmo del descenso del gradiente estocástico (SGD) y el descenso del gradiente adaptativo (Adam). Este último, en lugar de usar una tasa de aprendizaje constante como ocurre en el descenso del gradiente estándar, adapta la tasa de aprendizaje en función de los momentos del gradiente de los parámetros, lo que de alguna manera facilita o permite que esta se adapte automáticamente a la topología del modelo y a la naturaleza de los datos.

algoritmo permite ir minimizando de manera iterativa la función de error, o el coste del modelo, seleccionando un subconjunto aleatorio de los datos para su entrenamiento (*mini-batch*) y analizando en cada caso la distancia con una solución conocida. Lo que se hace con este algoritmo es básicamente moverse en el espacio de búsqueda (traducido a la superficie de la función de coste) intentando encontrar mínimos locales mediante la manipulación de los valores de los parámetros considerados en dicho modelo. Así, la esencia del algoritmo consiste en actualizar los pesos del modelo en la dirección del gradiente negativo utilizando en cada actualización o iteración un subconjunto aleatorio de los datos seleccionados para su entrenamiento.

Para el caso de AlphaFold, los parámetros considerados para la modelización son las distancias que separan a los aminoácidos que constituyen a la proteína y los ángulos de torsión de los enlaces químicos de dichos constituyentes (aminoácidos). Afirman los creadores de AlphaFold que

[...] Las propiedades que predicen nuestras redes son: (a) las distancias entre pares de aminoácidos y (b) los ángulos entre los enlaces químicos que conectan esos aminoácidos. [...]

Entrenamos una red neuronal por separado para predecir una distribución de distancias entre cada par de residuos en una proteína. Estas probabilidades se combinaron en una puntuación que estima cuán precisa es una estructura de proteína propuesta. También entrenamos una red neuronal separada que usa todas las distancias en conjunto para estimar qué tan cerca está la estructura propuesta de la respuesta correcta.

Usando estas funciones de puntuación, estuvimos en capacidad de buscar en el paisaje de proteínas para encontrar estructuras que coincidieran con nuestras predicciones. Nuestro primer método se basó en técnicas comúnmente utilizadas en biología estructural, y se reemplazaron repetidamente piezas de una estructura proteica con nuevos fragmentos. Entrenamos una red neuronal generativa para

inventar nuevos fragmentos, que se utilizaron para mejorar continuamente la puntuación de la estructura proteica propuesta.

El segundo método optimizó los puntajes a través del descenso del gradiente –una técnica matemática comúnmente utilizada en el aprendizaje automático para realizar pequeñas y crecientes mejoras–, lo que resultó en estructuras altamente precisas. (Evans et al., 2018, s/p)

Se intenta mostrar con lo anterior que un punto clave en el gran adelanto de resolución de este problema, que lleva algo más de cincuenta años de planteado, consistió en la manipulación y exploración de las consecuencias de dicha manipulación, vía modelo y simulación computacional, de los parámetros estimados relevantes en la dinámica del sistema. Esta exploración, dada la enorme cantidad de combinaciones de valores involucrados en el fenómeno emergente a explicar –la funcionalidad de la proteína–, no puede darse de manera exhaustiva, sino más bien aleatoriamente. A partir del conocimiento de las propiedades fisicoquímicas de los aminoácidos constituyentes de una proteína particular resulta impensable deducir la estructura o configuración específica que la hace funcional, requiriéndose entonces vías alternativas.

Las redes neuronales artificiales son una herramienta que permiten dar cuenta de la interacción no lineal de muchísimos elementos. Esto representa una estrategia con la que se podría dar un tipo de aproximación no analítica de fenómenos emergentes como los que interesan en el estudio de los sistemas complejos. Se habla de aproximaciones no analíticas en el sentido de que en ellas no se parte estrictamente de un análisis de los contenidos de conocimiento de los elementos involucrados en el fenómeno –emergente– a explicar para poder dar cuenta de este. Se parte más bien de una suerte de manipulación o intervención, por medio de simulaciones computacionales de dichos elementos involucrados, hasta obtener una *correlación* entre el fenómeno a explicar y los valores de ciertos parámetros considerados relevantes en la dinámica del sistema y que

se pueden codificar bajo los parámetros de otro sistema sobre el que la manipulación es más factible.

En este sentido, el enfoque de simulación computacional permite identificar patrones y correlaciones que podrían pasar desapercibidos con un enfoque analítico tradicional, o bien directamente ser imposibles de lograr. El carácter “manipulacionista” de este tipo de aproximaciones radica en la posibilidad de manipular a gran escala los parámetros de la simulación para obtener resultados específicos y correlacionarlos con los datos del fenómeno a explicar (Vicsek, 2002), de hecho, guiando la manipulación hacia una mejor representación de la relación entre entradas y salidas. Así, se podría decir que la importancia de las correlaciones en la epistemología radica en la posibilidad de identificar patrones y relaciones emergentes que podrían ser difíciles de detectar mediante métodos analíticos tradicionales, y que la simulación computacional es una herramienta útil para identificar estas correlaciones y generar nuevas hipótesis.

Es importante resaltar que la limitación de un acercamiento expresamente analítico al fenómeno de la emergencia, y en general al fenómeno de la complejidad, va más allá de la dificultad de encontrar esquemas generales aplicables a estos fenómenos. En *principio*, mediante el sistema de ecuaciones diferenciales parciales para representar la función de onda de Schrödinger se podría determinar la configuración molecular que pretendamos. Sin embargo, con estas ecuaciones diferenciales parciales, a partir de más de diez partículas, no es ya sólo que una solución analítica resulta prácticamente imposible, sino que, además, una solución por aproximación numérica resulta en extremo difícil, aun con el poder de cálculo que presentan hoy en día los computadores más poderosos de los que disponemos (Laughlin & Pines, 2000). Piénsese que sólo estamos considerando algo más de diez partículas. La situación para el problema del plegamiento de proteínas es extremadamente

inmanejable desde la función de onda de Schrödinger, ya que en una proteína intervienen cientos de miles de partículas. Sin embargo, como se ha podido observar con el ejemplo de AlphaFold, a partir de la identificación de ciertos parámetros y de la exploración, posibilitada por simulaciones computacionales de las consecuencias de la manipulación de un amplio conjunto de valores paramétricos, ha resultado una estrategia conducente a resultados positivos.

AlphaFold, u otros sistemas exitosos cuyos resultados sorprendentes no eran previsibles a partir de las propiedades individuales de sus componentes (neuronas), posee una estructura de hiperparámetros diseñada muy cuidadosamente para tal fin. En este diseño podemos ver que sí hubo una intervención ‘analítica’ más o menos precisa: por ejemplo, una decisión clara de diseño está implicada en la selección específica como parámetros de las distancias que separan a los aminoácidos que constituyen a la proteína. Además, incluso hasta el algoritmo del descenso de gradiente tiene una justificación totalmente analítica: ¿por qué utilizar las derivadas parciales para actualizar los parámetros en cada iteración y no utilizar otro método, o simplemente el azar puro? Claro está que en el diseño hiperparamétrico de cualquier sistema de aprendizaje automático existe una importante intervención de conocimiento de tipo teórico y técnico, tanto del dominio de aplicación (en nuestro ejemplo, la bioquímica) como de la ingeniería computacional.

También es claro que las propiedades globales específicas que finalmente resulten del aprendizaje del sistema no son explicables ni predecibles con total transparencia de modo analítico, al menos en la mayoría de los casos. Esto es lo que suele denominarse el “problema de la caja negra” (Durán & Jongsma, 2021; Zednik, 2019). Sin embargo, sí existe la posibilidad de prever sus características mediante el diseño de hiperparámetros de un modo analítico. Además, si bien las propiedades globales no pueden preverse de

antemano, más allá de sus características generales, en algunos casos sí pueden explicarse a posteriori: los casos donde los resultados de la red neuronal pueden entenderse bajo ciertas explicaciones observables en el sistema del mundo real que el sistema artificial está modelando (en dichos casos, podríamos rastrear la “lógica” seguida por la red, aunque hacerlo quizás no nos diga mucho en última instancia del fenómeno que estamos intentando comprender, dado su comportamiento emergente).

Discusión filosófica sobre modelos de simulación computacional y redes neuronales artificiales

En este artículo se pretende sostener que, en el abordaje de sistemas complejos, constituidos de cantidades abrumadoras de elementos, algunos métodos de simulación computacional (aquí nos enfocamos en las redes neuronales) han resultado una poderosa herramienta, principalmente porque han permitido una exploración –no sin cierta opacidad– manipulacionista de posibles consecuencias en sus dinámicas a partir de ciertas alteraciones en alguno o algunos de los parámetros estimados relevantes para intentar dar cuenta de dicha dinámica.

El asunto es que la exploración facilitada por las redes neuronales no puede ser exhaustiva, como bien se intentó evidenciar con el ejemplo del problema del plegamiento de proteínas. Incluso para los potentes ordenadores de los que se disponen hoy día, explorar exhaustivamente las posibles consecuencias en la dinámica de los sistemas complejos de alterar los valores de ciertos parámetros estimados relevantes en el intento de dar cuenta de dicha dinámica bajo cierto modelo resultaría una tarea de nunca acabar. Si bien se establecen determinados criterios para la exploración, aun dentro del dominio de valores paramétricos resultante la exploración

exhaustiva resultaría algo improcedente. Así, una vez establecidos ciertos criterios para la exploración en cuestión que trazan ciertas “rutas de tránsito” preliminarmente plausibles —que bien podríamos tratar de heurísticas—, el movimiento exploratorio se da más bien de manera aleatoria.

En definitiva, en las redes neuronales artificiales de aprendizaje profundo —que son las de nuestro interés—, se presenta una situación epistemológica novedosa, y es una suerte de opacidad frente a lo que está ocurriendo en detalle en la búsqueda exploratoria de los valores de ciertos parámetros que se *correlacionan* con el fenómeno complejo/emergente a abordar.¹⁰

Con las redes neuronales artificiales se sabe qué es lo que está ocurriendo en términos generales en el despliegue del modelo subyacente que está siendo ejecutado —mediante la estructura

¹⁰ Lo que se tiene en el ámbito de las redes neuronales es una opacidad epistémica más profunda que la teorizada por Paul Humphreys (2004; 2009) quien la definía del siguiente modo: “un proceso es esencialmente epistémicamente opaco a [un agente cognitivo] X si y sólo si es imposible, dada la naturaleza de X, para X conocer todos los elementos epistémicamente relevantes del proceso (Humphreys, 2009, p. 618). Para Humphreys, las simulaciones computacionales son epistémicamente opacas por la incapacidad humana de conocer la evolución de todos los elementos epistémicamente relevantes de los procesos computacionales mientras ocurren. Si bien la opacidad epistémica no sería exclusiva de los procesos computacionales, la naturaleza de la tecnología involucrada (Hardware y Software) la haría especialmente acuciante allí, toda vez que muchas de las operaciones que allí suceden lo hacen a un nivel muy bajo, lejos del usuario. Humphreys menciona la opacidad con relación al uso de simulaciones computacionales utilizadas para estudiar *targets* del mundo real. Sin embargo, las simulaciones que analiza Humphreys son auditables (su código es accesible, y fue programado manualmente por expertos humanos), cosa que no sucede con las simulaciones con redes neuronales, donde las ‘lógicas’ utilizadas para llegar a los modelos finales, como veremos, son mucho menos transparentes toda vez que los parámetros se logran en procesos que involucran aprendizaje automático y el procesamiento de enormes bases de datos. Sin embargo, aun con esta opacidad, veremos que es posible obtener conocimiento útil también de las redes neuronales.

hiperparamétrica—, pero es prácticamente imposible determinar de antemano cuáles serán los valores finales de los parámetros aprendidos, dado que es muy difícil determinar a priori a qué mínimo local convergerá el sistema en su proceso de aprendizaje por el descenso del gradiente y, mucho más, cuáles serán los muchísimos parámetros aprendidos que conducirán a este mínimo error local. Es más, se diseña una programación de una naturaleza muy distinta a aquella en la que se dan las instrucciones precisas de qué hacer, más bien aquí se dota de suficiente variabilidad para que el algoritmo encuentre una serie de valores (como los numerosos parámetros que representan la intensidad de la conexión entre las neuronas de la arquitectura) que optimizan una métrica con respecto a un conjunto de datos. En esta búsqueda, el azar está implicado en la inicialización de variables y en otros procesos tales como la conformación de lotes aleatorios en el descenso del gradiente estocástico. Pero, más allá del azar implicado, lo que resulta prácticamente imposible es la determinación o previsión *a priori* de los mínimos locales existentes y a cuál de dichos mínimos la red convergerá: es la complejidad del problema y no el aspecto parcialmente no determinista lo que impide un tratamiento analítico tradicional. El azar, de hecho, es un recurso necesario, similar a como lo es en los métodos llamados Monte Carlo (Galison, 1996).

Los modelos de redes neuronales artificiales son una herramienta útil para resolver problemas computacionales complejos, pero su opacidad epistémica es un tema de discusión, por ejemplo, entre los expertos en la neurociencia, y, sobre todo, recientemente, en el terreno de la filosofía de la ciencia (Barak, 2007; Chirimuuta, 2020; Durán & Jongsma, 2021; Lipton, 2006). Aunque los constructores del modelo conocen muchas características de su arquitectura y funcionamiento interno, la forma exacta en que una red llega a sus predicciones o clasificaciones a menudo es bastante opaca para ellos. Esta opacidad se refleja en el término “caja negra”, que se utiliza

para describir dispositivos o piezas de código que transforman entradas en salidas sin proporcionar ninguna indicación del método detrás de esta operación (Chirimuuta, 2020, p. 3). En cuanto a esto, Omri Barak, por ejemplo, señala que “el aprendizaje automático nos proporciona niveles mayores de rendimiento, acompañados de un aumento paralelo en la opacidad” (2007, p. 5). Sin embargo, no sería del todo preciso afirmar que las redes neuronales artificiales son *completamente* opacas, ya que los creadores del modelo tienen conocimiento de varias características de su estructura y funcionamiento interno. Además, conocemos la conectividad completa de la red, la dinámica de cada neurona individual, la regla de plasticidad utilizada para entrenar la red y, de hecho, toda la experiencia de desarrollo de la red (Chirimuuta, 2020, p. 3).

El problema no menor que surge consiste en poder ofrecer una caracterización de los grados relativos de transparencia y opacidad exhibidos por diferentes modelos. Frente a esto, un punto de partida interesante puede ser la noción de ‘*inteligibilidad*’. De acuerdo con De Regt y Dieks (2005, p. 143), la falta de inteligibilidad impide que un predictor perfecto de tipo caja negra, conocido como un oráculo, sea considerado una teoría científica válida. Según su perspectiva, una teoría científica debe ser comprensible, lo cual implica la capacidad de entender cómo se generan las predicciones y desarrollar un entendimiento de las consecuencias que la teoría tiene en situaciones específicas. La inteligibilidad no se limita a un componente psicológico adicional, sino que es fundamental para la habilidad de los científicos para utilizar las teorías (Chirimuuta, 2020, p. 3).

De Regt reelabora una idea presentada originalmente por Werner Heisenberg y respaldada luego por Richard Feynman para caracterizar la inteligibilidad:

Criterio de Inteligibilidad (CIT): una teoría científica T (en una o más de sus representaciones) se considera comprensible para los científicos

(en el contexto C) si pueden reconocer las consecuencias cualitativas características de T sin necesidad de realizar cálculos exactos (De Regt, 2014: p. 33; también De Regt y Dieks, 2005, p. 151 y ss; Chirimuuta, 2020, pp. 3, 4).

En nuestro caso de análisis, frente a esta suerte de opacidad epistémica ligada fuertemente al despliegue de modelos de redes neuronales artificiales, podemos oponer tres elementos que pueden ayudar a echar luz a esta caja negra. Por un lado, las estructuras hiperparamétricas, a las que ya nos hemos referido, si bien no nos permiten prever los mínimos locales y las propiedades emergentes exactas subyacentes, sí nos habilitan a prever ciertas características que poseerán dichas propiedades aprendidas por la red (como su cantidad, su capacidad de detectar regularidades espaciales, temporales o de otro tipo, etc.). Por otro lado, si bien no es posible determinar de antemano qué regularidades detectará la red luego de su aprendizaje (si así fuera, no sería necesaria su utilización para el problema estudiado), en algunos casos sí es posible entender *a posteriori* dichas regularidades. Claro que no siempre es posible entender las “decisiones” adoptadas por la red y, muchas veces, estas son completamente opacas a nuestro entendimiento analítico, tanto *a priori* como *a posteriori*.

Respecto a lo que explayamos, podemos observar que en el caso en que las regularidades o propiedades globales aprendidas por la red puedan ser entendidas *a posteriori*, la opacidad epistémica podría considerarse solo aparente o al menos relativa al sistema computacional empleado, no con respecto al fenómeno sobre el que se ha aprendido algo gracias al método. De hecho, en el caso de AlphaFold, se puede argumentar que si bien no nos proporciona entendimiento del plegado mismo de las proteínas, su éxito predictivo con respecto a las estructuras puede constreñir el espacio de modelos posibles para el mecanismo de plegamiento,

proveyéndonos de esa manera nuevo conocimiento, quizás indirecto, de cómo estudiar luego el fenómeno subyacente.

Pensemos, además, que dicha opacidad sería similar a la que tenemos respecto a los razonamientos que realiza nuestro propio cerebro: podemos explicar a posteriori la lógica de nuestros pensamientos, pero nada sabemos de los procesos físicos detallados que siguieron nuestros miles de conexiones neuronales para llegar a ellos. Por lo tanto, el desconocimiento exacto de la estructura paramétrica (y de cómo arribó a ella la red en su entrenamiento) no entraña una opacidad epistémica *per se*; o, al menos, en todo caso, entraña la misma opacidad epistémica que nuestro desconocimiento de los procesos neuronales exactos (habitualmente denominados como “sub-personales”) que realiza nuestro cerebro cuando pensamos aun del modo más analíticamente riguroso. Podríamos agregar que la red artificial tiene la ventaja epistémica, sobre nuestro cerebro, de que podemos diseñar y conocer su estructura hiperparamétrica con total detalle y prever ciertas características de las propiedades emergentes resultantes (aun antes de que resulten, es decir, *a priori*), precisamente gracias a la manipulabilidad.

Por último, creemos que existe un tercer elemento que clarifica la opacidad epistémica atribuida a los sistemas computacionales: las imágenes que se van generando en sus iteraciones juegan un papel muy importante. Llegados a este punto conviene diferenciar dos tipos de situaciones que se dan con las imágenes generadas a partir de simulaciones computacionales, particularmente con aquellas generadas por redes neuronales artificiales. En primer lugar, es importante señalar que la cantidad de imágenes que se suelen generar por métodos de simulación computacional en general es sumamente grande, por lo que intentar identificar por parte nuestra algún tipo de patrón a partir de ellas es también una tarea improcedente. Sería una tarea de nunca acabar. En el caso, por ejemplo, de las redes neuronales artificiales, la identificación de patrones a partir de esa

enorme cantidad de imágenes ocurre regularmente por parte de ellas mismas (Evans et al., 2018).

En una de las etapas en el entrenamiento de AlphaFold, la red genera, basándose en ciertas técnicas propias de la bioquímica,¹¹ una enorme cantidad de imágenes que representan mapas de distancia entre los aminoácidos constituyentes de la proteína de la cual se pretende predecir su estructura tridimensional específica. Sobre estos mapas de distancia la red se entrena en la identificación de ciertos patrones. Con este entrenamiento, entre otras cosas, la red autoajusta ciertos parámetros para generar nuevos mapas de distancia más discriminados. Todo este proceso resulta opaco para la supervisión humana detallada. Todo ello implica que dentro de las simulaciones computacionales mediante redes neuronales artificiales se da el caso de la generación de imágenes y procesamiento de estas sin supervisión humana, algo que no ocurre en otros contextos de investigación científica.

¹¹ En primer lugar, los investigadores de DeepMind se valen del extenso *dataset* del que se dispone, desde hace ya un par de décadas, sobre la secuencia genómica de varias especies. Es decir, se valen básicamente del extenso número de proteínas de las que se conoce su secuencia de aminoácidos, a través de una técnica que se conoce como ‘Multiple Sequence Alignment’ (MSA) (Senior et al. 2019), con la que se busca principalmente alinear secuencias de aminoácidos constituyentes de proteínas emparentadas o ligadas evolutivamente. Es decir, se busca alinear secuencias homólogas.

La idea con este alineamiento es identificar patrones entre dichas secuencias o, más concretamente, correlaciones, las cuales de alguna manera ayuden a estimar la distancia a la que se encuentran dos aminoácidos entre sí. Por ejemplo, si al alinear varias secuencias de aminoácidos, constituyentes correspondientemente de varias proteínas ligadas evolutivamente de distintas especies (secuencias homólogas) se logra identificar bajo algún patrón de alineamiento una fuerte correlación entre dos aminoácidos, esto representaría un fuerte indicio de que esos dos aminoácidos se hallan a una distancia cercana el uno del otro. Opera aquí un principio de parsimonia, entendiendo que la selección natural busca conservar estructuras funcionales.

La otra situación que cabe resaltar en cuanto a las imágenes generadas por simulación computacional es aquella en la que éstas figuran como puentes de acceso o de observación para los investigadores a lo que ocurre “semiautónomamente” en la ejecución de estos modelos que involucran entrenamiento y búsqueda exploratoria con cierta dosis de aleatoriedad. Es decir, las imágenes en esta situación terminan siendo nuestro vínculo expedito con el modelo computacional en ejecución. Esto constituye la primera vez en nuestra historia que contemplamos una especie de autoridad epistémica artificial luego, por ejemplo, del éxito predictivo que vienen sosteniendo las redes neuronales artificiales (ver p. ej. Chirimuuta, 2020). Debido a que las imágenes que se generan a partir de sus ejecuciones terminan siendo en varios casos nuestro punto de acceso a ellas, podría considerarse que esta situación en términos de visualización resulta novedosa, no tanto por la novedad de la visualización científica en sí, que tiene larga data, sino por la escala en la que estas técnicas cuya arquitectura está inspirada en la corteza visual del cerebro puede crear y manipular representaciones visuales. Por este motivo, posibilita la contemplación de múltiples relaciones entre elementos al mismo tiempo (una visión holística del sistema). Muchas simulaciones, desde el inicio de la computación, proveen imágenes que mejoran la comprensión de los sistemas analizados. Lo interesante aquí es que estas imágenes puedan dar cuenta de sistemas complejos, y también que sean las propias redes neuronales las que analicen a las imágenes que generaron para extraer la mayor cantidad de información útil de las mismas. Así, estas nuevas imágenes generadas por simulaciones permiten a los investigadores arrojar un poco de luz a la opacidad epistémica inherente al análisis de sistemas complejos y del uso de redes neuronales en general, aunque la transparencia total resulte un ideal inalcanzable.

Por supuesto, pretensiones epistémicas como la transparencia podrían considerarse un poco desplazadas cuando de abordar la complejidad se trata. Frente al estudio de sistemas complejos, una comprensión o entendimiento analítico es reemplazado en gran medida por una comprensión o entendimiento manipulacionista enfocado quizás más bien en lograr resolver un problema heurísticamente y ofrecer predicciones. Ello ocurre con claridad en el ejemplo de AlphaFold y el problema del plegamiento de las proteínas, especialmente en lo que respecta al subproblema de la predicción de estructuras. Las simulaciones computacionales por redes neuronales permiten hacerles frente a las barreras de la complejidad antes que eludirlos, pues presentan una alternativa frente a la frustración de nuestra incapacidad de una comprensión analítica de toda realidad compleja (Lenhard, 2016, p. 20).

Conclusión

La discusión sobre las redes neuronales artificiales y su capacidad para proporcionar una explicación adecuada de los fenómenos que están siendo investigados y su capacidad para explicar adecuadamente fenómenos complejos está estrechamente relacionada con la noción de hiperparámetros y la opacidad epistémica que rodea a estos modelos. Si bien los hiperparámetros pueden proporcionar cierta claridad al permitir prever ciertas características de las propiedades aprendidas por la red, la incertidumbre persiste debido a la imprevisibilidad de las regularidades que la red detectará después del entrenamiento. Aunque en ocasiones es posible entender estas regularidades a posteriori, en muchos casos, las decisiones de la red siguen siendo opacas tanto antes como después del aprendizaje.

En el contexto del problema del plegamiento de proteínas, se destaca cómo las redes neuronales artificiales, como AlphaFold2 de DeepMind, son una valiosa herramienta para la investigación científica en la era de la complejidad. Este tipo de problema, que ha sido objeto de estudio durante décadas, presenta desafíos significativos debido a su complejidad y a la gran cantidad de variables involucradas. La comprensión analítica tradicional ha sido extremadamente difícil de lograr en este caso.

En lugar de buscar una comprensión detallada y analítica del problema, AlphaFold2 utiliza una comprensión manipulacionista, basada en el procesamiento masivo de datos por parte de la red neuronal para producir predicciones precisas y útiles. Aunque esto implica renunciar a la virtud epistémica clásica de la transparencia, se considera un precio justo a pagar para obtener conocimientos sobre sistemas complejos que de otro modo serían difíciles de abordar.

En resumen, la discusión sobre las redes neuronales artificiales en la investigación científica refleja una tensión entre la necesidad de una comprensión profunda y detallada de los fenómenos y la utilidad práctica de las herramientas que nos permiten manipularlos y predecirlos eficientemente. Los hiperparámetros desempeñan un papel relevante en intentar entender estas redes, pero aun así, la opacidad epistémica persiste, lo que destaca la importancia de considerar enfoques manipulacionistas en la resolución de problemas complejos con estas herramientas tecnológicas. Podemos estimar que parte de la dinámica científica por venir se centrará en la articulación de distintas clases de modelos sobre un mismo fenómeno

Referencias

- Barak, O. (2017). Recurrent neural networks as versatile tools of neuroscience research. *Current Opinion in Neurobiology*, 46, 1–6.
- Bunge, M. (2004). *Emergencia y Convergencia. Novedad Cualitativa y Unidad del Conocimiento* (Rafael González del Solar, Trad.). Gedisa. (Obra original publicada en 2003)
- Chirimuuta, M. (2020). Prediction versus understanding in computational enhanced neuroscience. *Synthese*, 199(4).
- Churchland, P. S., Sejnowski, T. J. (2016). Blending computational and experimental neuroscience. *Nature Reviews Neuroscience*, 17, 667–668.
- Durán, J. M. (2018). *Computer Simulations in Science and Engineering: Concepts -Practice -Perspectives*. Springer Nature Switzerland.
- Durán, J. M., Jongsma, K. R. (2021). Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. *Journal of Medical Ethics*, 47(5), 329-335.
- Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., Qin, C., ..., Senior, A. (2018). *De novo structure prediction with deep-learning based scoring*.
- Feurer, M., Hutter, F. (2019). Hyperparameter Optimization, en: Hutter F., Kotthoff, L., & Vanschoren, J. (eds.) *Automated Machine Learning: Methods, Systems, Challenges* (pp. 3-33). Cham: Springer Nature Switzerland.
- Frigg, R. (2022). *Models and Theories: a Philosophical Inquiry*. Routledge.
- Galison, P. (1996). Computer Simulations and the Trading Zone, en: Galison P., & Stump, J. D. (eds). *The Disunity of Science: Boundaries, Context and Power* (pp. 118-157). Stanford University Press.

- Gao, P., Ganguli, S. (2015). On simplicity and complexity in the brave new world of large-scale neuroscience. *Current Opinion in Neurobiology*, 32, 148–155.
- Gell-Mann, M. (1998). *El Quark y el Jaguar: aventuras en lo simple y lo complejo* (A. García y R. Pastor, Trad.). Tusquets. (Obra original publicada en 1994).
- Goldenfeld, N. & Kadanoff, L. P. (1999). Simple lessons from complexity. *Science*, 284: 87- 89.
- Hartmann, S. (1996). The World as a Process: Simulations in the Natural and Social Sciences, En: Hegselmann, R., Mueller, U., Troitzsch, K.G. (Eds). *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View* (pp. 77-100). Kluwer.
- Hawkins, J., Blakeslee, S. (2005). *On Intelligence: How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines*. St. Martin's Griffin.
- Hidalgo, C. (2015). *Why Information Grows. The Evolution of Order, from Atoms to Economies*. New York: Basic Book.
- Holland, J. H. (1992). Complex Adaptive Systems. *Daedalus*, 121(1), 17–30.
- Humphreys, P. (1997). Emergence, Not Supervenience. *Philosophy of Science*, 64, 337- 345.
- Humphreys, P. (2004). *Extending Ourselves. Computational Sciences, Empiricism, and Scientific Method*. Oxford University Press.
- Humphreys, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169, 615-626.
- Humphreys, P. (2016). *Emergence: A Philosophical Account*. Oxford University Press.
- Johnson, S. (2003). *Sistemas emergentes. O qué tienen en común hormigas, neuronas, ciudades y software*. Fondo de Cultura Económica.

- Jumper, J., Evans, R., Pritzel, A. Green, T., Figurnov, M., Ronneberger, O., ..., Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596: 583-589.
- Kauffman, S. A. (1993). *The Origins of Order: Self-organization and Selection in Evolution*. Oxford University Press.
- Laughlin, R. B., Pines, D. (2000). The Theory of Everything. *Proceedings of the National Academy of Sciences*, 97(1), 28-31.
- Lenhard, J. (2016). Computer Simulation, en: Humphreys, P. (ed.) *The Oxford Handbook of Philosophy of Science* (pp. 803-824). Oxford University Press.
- Lenhard, J. (2017). License to Explore: How Images Work in Simulation Modeling, en: Ammon S., & Capdevila-Werning, R. (eds). *The Active Image. Architecture and Engineering in the Age of Modeling* (pp. 233-254). Springer.
- Levinthal, C. (1968). Are there pathways for protein folding? *Journal de Chimie Physique*, 65, 44-45.
- Lillicrap, T., Kording, K. (2019). *What does it mean to understand a neural network?*
- Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews Neuroscience*, 21(6), 335-346.
- Mitchell, M. (2009). *Complexity. A Guided Tour*. Oxford University Press.
- Napoletani, D., Panza, M., Struppa, D. C. (2011). Agnostic science. Towards a philosophy of data analysis. *Foundations of Science*, 16(1), 1-20.
- Nicolis, G., Prigogine, I. (1989). *Exploring Complexity. An Introduction*. W. H. Freeman & Company.
- Nielsen, M. (2015). *Neural Networks and Deep Learning*.
- Pearl, J., Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.

- Puterman, M. (1990). Markov Decision Processes. En *Handbooks in Operations Research and Management Science vol 2* (pp. 331-434). Elsevier.
- Rescher, N. (1998). *Complexity. A Philosophical Overview*. Transaction Publishers.
- Rumelhart, D., Hinton, G., Williams, R. (1986). Learning representations by back-propagation errors. *Nature*, 323, 533-536
- Senior A.W., Evans, R., Jumper, J., Kirkpatrick, J., Sifre, L., Green, T., ..., Hassabis, D. (2020). Improved protein structure prediction using potentials from Deep learning. *Nature*, 577: 706-710.
- Taylor, M. C. (2001). *The Moment of Complexity. Emerging Network Culture*. Chicago: The University of Chicago Press.
- Vicsek, T. (2002). The Bigger Picture. *Nature*, 418: 131.
- Weiskopf, D. (2022). The Predictive Turn in Neuroscience. *Philosophy of Science* (5).
- Woodward, J. (2003). *Making Things Happen. A theory of causal explanation*. New York: Oxford University Press.
- Yang, L., Shami, A. (2020). On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing*, 415: 295-316
- Zednik, C. (2019). Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *arXiv:1903.04361 [cs]*.

DIGITAL ASCETISM FOR EVERYDAY LIFE¹

ASCETISMO DIGITAL PARA LA VIDA COTIDIANA

ELVIRA VALEEVA

South Ural State University, Rusia
valeevaem@susu.ru

EKATERINA MILYAEVA

Chelyabinsk State University, Rusia
morning_demon@mail.ru

REGINA PENNER

South Ural State University, Rusia
penner.r.v@gmail.com

ELENA SOSNOVSKHIH

South Ural State University, Rusia
sosnovskikheg@gmail.com

RECIBIDO: 02/09/2023

ACEPTADO: 25/11/2023

Abstract: A new technological paradigm integrating information and communication technologies radically and rapidly transforms the landscape of human physicality, everyday life and spirituality. Digitalization forms a certain trend towards the emergence of a global virtual space with its own special culture that influences the formation of the spiritual and moral values of society. The specified context allows us to reflect on the human dimension of digitalization through the digital consciousness of the individual. In cognitive terms, digitalization entails, along with free access to a variety of information, a decrease

¹ The work was supported by the grant of the Presidency of the Russian Federation for young scientists-PhDs MK-2592.2022.2 “Digital Anthropology: Theoretical and Applied Aspects”.

in cognitive abilities, which is largely determined by decentralization of attention. The scale and temporal characteristics of information determine changes in the thinking and consciousness of a person. In the case of the digitalization of education, the shadow sides of this process are the growth of digital addiction among students, as well as the transformation of the social structures of the Internet and the transfer of classic destructive behaviors to them: bullying, stalking, threats and crimes. It creates a certain gap that needs both humanitarian reflection and the development of specific pedagogical measures to counteract the negative consequences of digitalization. In the article, the authors present an understanding of the relevance of a new approach to the ways of human existence in the context of total digitalization. The need to develop digital skills and competencies is combined with the relevance of preserving the human element and requires a humanitarian approach, the focus of which remains the priority of a person, not technology. Textbooks and a mobile application “Digital Asceticism” are presented, the relevance of this methodology in contemporary education is substantiated. The practices of digital asceticism can be seen as a regulator of the activity of the user’s presence in the digital environment. This approach allows digital users to partly stay in the information environment, while avoiding digital redundancy, digital illiteracy, digital addiction, which helps to improve the quality of their life. Authors believe that contemporary people are in dire need of the formation of a new, network-oriented, but not network-centric way of life. As such a paradigm of human existence, they propose a transition to digital asceticism, which they define as a way of life characterized by an understanding of the latest means of communication and the reasonable expenditure of one’s attention in the world of modern technologies and media. Within digital asceticism as a way of life, authors identify a separate method of digital asceticism, which, in their opinion, is an important tool for achieving a balance between online and offline. Digital asceticism includes a number of practices that help a person control his information consumption and use of digital technologies.

Keywords: digital anthropology; digital asceticism; digitalization; digital hygiene; digital skills; digital addiction; handwritten diary; mobile application

Resumen: Un nuevo paradigma tecnológico que integra las tecnologías de la información y la comunicación transforma radical y rápidamente el paisaje de la fisicalidad, la vida cotidiana y la espiritualidad humana. La digitalización forma una cierta tendencia hacia el surgimiento de un espacio virtual global con su propia cultura especial que influye en la formación de los valores espirituales y morales de la sociedad. El contexto especificado nos permite reflexionar sobre la

dimensión humana de la digitalización a través de la conciencia digital del individuo. En términos cognitivos, la digitalización conlleva, junto al libre acceso a información diversa, una disminución de las capacidades cognitivas, que viene determinada en gran medida por la descentralización de la atención. La escala y las características temporales de la información determinan los cambios en el pensamiento y la conciencia de una persona. En el caso de la digitalización de la educación, las caras oscuras de este proceso son el crecimiento de la adicción digital entre los estudiantes, así como la transformación de las estructuras sociales de Internet y la transferencia a ellos de comportamientos destructivos clásicos: bullying, stalking, amenazas y crímenes. Crea una cierta brecha que necesita tanto de una reflexión humanitaria como del desarrollo de medidas pedagógicas específicas para contrarrestar las consecuencias negativas de la digitalización. En el artículo, los autores presentan una comprensión de la relevancia de un nuevo enfoque de las formas de existencia humana en el contexto de la digitalización total. La necesidad de desarrollar habilidades y competencias digitales se combina con la relevancia de preservar el elemento humano y requiere un enfoque humanitario, cuyo enfoque sigue siendo la prioridad de una persona, no de la tecnología. Se presentan libros de texto y una aplicación móvil “Ascetismo Digital”, se fundamenta la relevancia de esta metodología en la educación contemporánea. Las prácticas de ascetismo digital pueden verse como un regulador de la actividad de presencia del usuario en el entorno digital. Este enfoque permite a los usuarios digitales permanecer parcialmente en el entorno de la información, evitando al mismo tiempo la redundancia digital, el analfabetismo digital y la adicción digital, lo que ayuda a mejorar su calidad de vida. Los autores creen que la gente contemporánea necesita urgentemente la formación de una nueva forma de vida orientada a las redes, pero no centrada en las redes. Como paradigma de la existencia humana, proponen una transición al ascetismo digital, que definen como una forma de vida caracterizada por la comprensión de los últimos medios de comunicación y el gasto razonable de la atención en el mundo de las tecnologías y los medios modernos. Dentro del ascetismo digital como forma de vida, los autores identifican un método separado de ascetismo digital que, en su opinión, es una herramienta importante para lograr un equilibrio entre lo online y lo offline. El ascetismo digital incluye una serie de prácticas que ayudan a una persona a controlar su consumo de información y el uso de tecnologías digitales.

Palabras clave: antropología digital; ascetismo digital; digitalización; higiene digital; habilidades digitales; adicción digital; diario escrito a mano; aplicación móvil

Introduction

At the present stage of social development, there is a significant expansion of communication space due to the achievements of technological progress. Modern technologies have provided a person with undeniable advantages of mediated communication: high speed of obtaining information, hypermedia, the possibility of filling and supplementing information, unlimited possibilities of remote communication, new forms of creativity, etc. But the endless digital universe has a downside, having a negative impact on the ability to interact with information.

In a situation of total digitalization, certain digital skills are of great importance: the ability to verify the accuracy of the information received; the ability to limit the time spent on social networks; the ability to monitor your health and digital footprint. A serious challenge is the fact that most of the interactions are transferred from real life to the virtual space. The skills of working on the Internet and understanding its role in one's own life should be instilled from an early age. Today, children begin to interact with mobile devices early, and the “digital footprint” of children is diligently formed by their parents, who share their photos and videos on the Internet, and sometimes maintain pages on their behalf on social networks. In the near future, the development of digital technologies is predicted to expand the human body through information modernization, which is not limited to the implantation of chips and artificial organs into the human body (Uriu, 2023).

Martin Heidegger had written that for a modern person, time also becomes a kind of consumption object, and in this regard, today a person has less and less time for himself, while thinking that he is in full control of his time (Heidegger, 2009). In the conditions of hyperbolization of information flows, strong introduction of digital devices into everyday life, such a conviction becomes more and

more appropriate, but at the same time leads to stressful conditions due to overload, both with useful information and digital garbage, it is necessary to control the actual processes of digital culture (Kuprashvili, 2023).

That is why contemporary people are looking for effective ways of caring for himself. The experience of applying the philosophy of the Stoics as a means of “taking care of oneself” is described in studies that update and adapt to the principles of Stoicism in a situation of high information load, constant interaction with gadgets, the race for “perfection” in social networks and real life (Sklar, 2017; Guerin, 2022; Pavlov, 2021; MacMillan, 2019).

For today's youth, the dominant motives for going online are communication and information search, that is, the digital world is mainly educational and entertaining for them. (Abakumova et al., 2022). In this regard, it is necessary to coordinate the digital behavior of young people so that it becomes the basis of an active life position, improves the quality of life, and contributes to the establishment of important social skills. The practices of digital asceticism can act as a regulator of the activity of the user's presence in the digital environment. This approach allows you to stay in the information environment, while avoiding digital redundancy, digital illiteracy, which helps to increase productivity and focus, and improve the quality of human life (reducing stress and anxiety caused by the constant use of digital technologies).

The topic of digital asceticism is also relevant in the context of digital hygiene of a modern person (Sklar, 2017). A lot of screen time can lead to digital overstrain, reduced sleep quality, and problems with vision and posture. Insufficient management of information load can lead to deterioration in the ability to focus, analyze and make decisions (Pittaway, 2020).

At South Ural State University, a team of authors developed special manuals for students and school teachers aimed at developing digital

ascetic skills in students, as well as a mobile application “Digital Ascetic Diary”² (Grednovskaya et al., 2022; Gulevataya et al., 2023). To understand online and master the skills of digital asceticism, it is proposed to use the methodology of philosophical practice, a handwritten reflective diary. Working with the manual, application and keeping a diary, it is necessary to use philosophizing as a means for setting, analyzing and solving worldview problems, to carry out philosophical reflection in order to expand the boundaries of the worldview. Naturally, in the early stages of learning and without knowledge of the basics of philosophy, such practices should have a simple and understandable form.

The aim of the article is to represent the Digital Ascetic methodology potency by presenting a cycle of manuals and the Digital Ascetic mobile application for the formation of balance skills between offline and online in the context of digitalization.

In our opinion, combining a handwritten diary with a mobile application allows us to combine personal written interaction with modern technological advances. This approach satisfies the need to preserve the traditional methods of self-expression found in handwritten notes, while the mobile application complements this experience with the functionality of a digital environment. Digital asceticism is understood by the authors not as a call to abandon modern technologies, but as a need to achieve harmony and balance in a combination of online and offline, as an opportunity to prevent various forms of digital addiction.

Digital Asceticism as an Educational Tool in the Digital World

In our study, digital asceticism refers to the formation of a special

² The mobile application is available at https://apps.rustore.ru/app/com.ne_lopatin.Digital_print

way of life, maintaining a balance of being online and offline, understanding the specifics of the digital world and the problems that one may encounter in it.

With this understanding of digital asceticism, the concept of “media asceticism” is narrower and implies only control over life in the media space, to a greater extent by limiting the use of gadgets. Media asceticism does not imply systematic pedagogical work, that is, the formation of networking skills and gaining knowledge about the features and processes of communication on the network, not only the development of the habit of not using the gadget often.

Methods of digital asceticism for contemporary education is one of the ways to improve the skills of digital literacy, digital hygiene, digital security of modern children and adults (parents and teachers with whom children always interact), who can teach each other a lot through joint work in the classroom and individual reflections on the diary.

A psychological and pedagogical review of the traditionally established methods and forms of organizing and conducting classes in educational institutions shows that the educational process is dominated by reproduction in the perception and mastery of educational material. The forms, methods and means of teaching used mainly involve the activity of the teacher in the process of transferring the necessary knowledge, the student is only required to passively listen and memorize. Undoubtedly, this approach is outdated for today's rapidly changing life.

Today, the traditionalist paradigm of education is giving way to a phenomenological direction of education with a humanistic orientation. Within the framework of this concept, priority is given to self-education, individual learning, independent choice of goals and methods of mastering the material, the interactive nature of education, as well as the need for learning throughout a person's life (life-long learning).

The authors of this article propose a digital ascetic methodology, the main tasks of which are:

- 1) to create conditions for educational activities under which students would strive to independently acquire new knowledge and, in the future, could successfully use them in their practical activities;
- 2) to prepare students for life in the digital world, for the perception of any information in the most environmentally friendly and hygienic way, using digital security skills.

The digital ascetic method is becoming an important tool to achieve harmony between online and offline. It helps to understand how digital technologies affect everyday life in general.

The successful construction of an environmentally friendly digital educational space is the birth of a new pedagogical paradigm, which occupies a worthy place along with the competency-based paradigm. Therefore, the introduction of the methodology of digital asceticism into educational practice creates a basis to talk about the trend of the emergence of new didactics, which allows you to combine traditional learning with digital learning.

Mobile application “Diary "Digital Asceticism"” for students and teachers

The Diary “Digital Asceticism” was developed specifically for high school students and includes the study of such important issues for adolescents as cyberbullying, digital footprint, self-care and health, in the context of learning critical thinking. It consists of 4 sections: digital footprint, cyberbullying, critical thinking and care of the self. In each section, a number of tasks have been developed, which also involve written work. It is a “prescription” of one's thoughts, feelings and emotions that helps to comprehend, accept and understand the need for elaboration. The role of the diary as a reflective practice can

hardly be overestimated, because philosophical diaries were kept by S. Kierkegaard, J.-P. Sartre, A. Schopenhauer and others (Mendieta, 2014).

The application “Digital Asceticism” runs on the React Native JavaScript Framework, in the Expo variation, which allows us to create cross-platform applications for both Android and iOS smartphones. We focused on the budget segment of smartphones available to schoolchildren and therefore chose OS Android. We have placed the application in RuStore, a Russian online application store for mobile devices developed by VK Corporation with the support of the Ministry of Digital Development, Communications and Mass Media of the Russian Federation. This is necessary for the possibility of use in Russian schools.

From the point of view of the application architecture, there are several screens; the main screen includes the entire introduction, block-weeks, the screen with information about the authors, the screen for each week, which provides introductory information for the week, and the screen of the day, which presents the tasks of each day. Dynamic generation is built into two of the four screens. The page of the week, the page of days is dynamically generated depending on the context that the user requests. For example, when switching from the main screen to the week screen, the application “understands” which week it is moving to, by serial number from 1 to 4, remembers this number, and displays information in a specific week. After that, when we click on a specific day, we pass into the program, in a variability of the order of the day, which is tied to a specific topic. The program, given the number of the week and the number of the day, chooses what information to display on the screen stored in the application. In the application, everything is built on React Hook useRef, used when loading and rendering the page, and React Hook useState, i.e., storing program variables. The application practically does not use the phone's memory. The application does

not ask permission to use the camera, microphone. This makes the application secure. The design of the application was created in accordance with modern trends and is attractive to students.

The diary is formally dated, that is, all tasks are designed for daily systematic work. In addition, for work in the format of digital asceticism, a quote of the day is offered, something that will help to tune in to work and set its vector.

One week is expected to master each section, during which it is important not only to familiarize yourself with the content of the informational part of the diary, but also to assess your condition before and after working with it.

Combining a handwritten diary with a mobile app allows us to combine personal written interaction with modern technological advances. This approach satisfies the need to preserve the traditional methods of self-expression found in handwritten notes, while the mobile application complements this experience with the functionality of the digital environment. Integration of handwritten notes with a mobile application provides flexible access to diary information. This allows users to make entries, both in a traditional form and via an electronic device, providing convenience and adaptability to different situations and moods. A mobile app can greatly enrich the diary experience. It provides the ability to attach photographs, audio and video recordings to recordings, expanding the possibilities of self-expression and archiving of visual impressions. This allows the user to more fully and visually describe events and emotions. The integration of both approaches allows us to create a centralized and organized way to manage diary entries. The mobile app provides tools for categorizing, tagging, and searching entries, making it easy to access specific moments and thoughts. In addition, such a system makes it easier to analyze and reflect on long-term trends and changes. Integrating a mobile app with a handwritten diary allows users to participate in a larger social

context. Users can share selected entries and observations online, exchange experiences and ideas, which contributes to a richer and more diverse diary experience.

Since the diary is intended for a youth audience, the team of authors tried to make it as interesting and easy to understand as possible. The authors of this study propose two fundamental strategies for interacting with a mobile application and a diary. The first is the reflexive interpretation of a philosophical quotation and the subsequent implementation of its concepts in the daily life of the individual. The second is the practice of freewriting, which involves impartial and creative reflection on problematic topics proposed by the authors.

In the course of the work, the reader of the diary is offered a systematic approach, starting with the formulation of his or her actual problem each week. Then the analysis of this problem is carried out from the standpoint of a reflexive approach to philosophical quotation. Such an analytical process aims to intellectually assimilate the concepts expressed in the quotation and apply them to real situations of everyday life.

Together with the reflective strategy, the authors suggest using the freewriting method. This approach encourages free and unrestricted thinking on the topic provided. The purpose of freewriting is to stimulate creativity, analytical flexibility and deep immersion in the topic, which contributes to deeper understanding and the search for new ways to solve the problem.

To conclude, each week of research covers a cycle starting with a problem statement, then moving on to a reflective analysis of a philosophical quote and its implementation in everyday life, and ending with a freewriting session that enriches and deepens the understanding of the proposed topic.

Four weeks of digital asceticism

Week 1

The first week of digital asceticism is dedicated to the digital footprint. First, students are asked to plan a to-do list for the coming week, given the concept of a digital footprint, and given the opportunity to describe their digital footprint. In today's information society, the role of the digital footprint left by individuals on the Internet and on digital devices is becoming increasingly important. A variety of online platforms, interactive services and social networks provide the user with the opportunity to create a unique digital footprint consisting of a variety of data and activities (Baranova et al., 2021). In this context, understanding such key concepts as “digital profile”, “social network” and “digital hygiene” is of paramount importance. These concepts become tools for thinking about your own digital footprint, its formation and effective management (Ponomarev et al., 2022). In this context, the authors understand a digital footprint as a unique sequence of online activities or activities on digital platforms. This trail occurs when personal data is published online. Note that in contemporary society, a user may leave footprints intentionally (such as by posting a photo), unconsciously (by visiting websites or pages), or passively (by being tagged in an image posted by others). At the same time, interested parties can passively or actively collect information about the user. The risk associated with the digital footprint is that it can be misused as a target for cybercrime such as identity theft or fraud. In this context, it is critical to ensure that students develop the competencies needed to protect personal data and effectively manage their own digital footprint (Lapchik et al., 2021).

The concept of “digital footprint” and its related concepts, “digital profile”, “social network” and “digital hygiene”, are analyzed not only from philosophical and pedagogical, but also from legal, social

and other perspectives (Surmelioglu, Seferoglu, 2019; Jansen, Hinz, 2022). This multifaceted view allows students to develop a deep and comprehensive understanding of the phenomenon under study, revealing the positive and negative aspects of a personal digital footprint. In addition, such an analytical approach contributes to the development of individual strategies for behavior in the digital space.

The transfer of communication from the digital environment to real life is proposed to be carried out by compiling a map of one's own life and comprehending the content of such concepts as "love", "friendship", "loneliness", "peace", "person", etc. It is assumed that when writing the proposed concepts, students will be able to fully distinguish between the real and virtual worlds, as well as understand the place of the Internet in the formation of various types of human relationships.

Moreover, the diary offers to streamline work with gadgets and applications. This is especially important as part of the formation of digital hygiene – limiting notifications and setting notifications depending on the time of day.

Week 2

The second week of work is devoted to an important and acute topic in the youth environment, the topic of cyberbullying. In contemporary society, which can be characterized by the transition of communicative practices to the online space, there is a new form of violence – cyberbullying (Patchin et al., 2011; Zhu et al., 2021; Macaulay et al., 2022). Its essence lies in the systematic impact using digital technologies in order to destabilize the psycho-emotional state of the victim (Amran et al., 2023). Among the most common manifestations of this phenomenon are online insults, threats, as well as the dissemination of deliberately false, so-called fake information

and other forms of destructive influence. Single individuals and groups can carry out cyberbullying, which brings serious negative consequences to victims, such as the development of psychological disorders, a decrease in self-esteem, the occurrence of depressive states, and even suicidal disorders.

The study of the topic “Cyberbullying” in the educational process allows students to realize the very serious consequences of this type of violence. They become able to recognize the tell-tale signs of cyberbullying and learn how to prevent and resist it. Recognizing and developing these skills are essential elements in ensuring safe and effective behavior in the digital environment. In the context of this topic, not only techniques for countering cyberbullying are considered, but independent work is also expected to search for psychological support services. This approach is justified, since not all topics can be discussed with relatives or friends.

Furthermore, joint work on the digital trail and cyberbullying is expected, i.e., setting up profiles in social networks. Nowadays, some social networks offer to hide their page from users for at least a week. Mechanisms for blocking unwanted users with specific reasons have been developed.

An important point is to work out your own fears, since the victims of cyberbullying, in the first place, can be people who are unsure of themselves and their own abilities. It is proposed to prescribe your fears, accept them and describe the mechanisms of confronting them. In this case, the student himself finds a way out of the situation, grows up, learns to make decisions and be aware of their consequences (Burlet et al., 2022). The last point is especially important for those who practice cyberbullying themselves; philosophical reflection in the form of a diary makes it possible to take the place of the “victim” and understand how the person feels about whom illegal actions are being carried out. The practice of providing assistance (expressing thoughts) to a person who finds

himself in a difficult situation offers to relieve psychological tension and then, vice versa, to express an opinion about a pleasant, good situation.

An indicator of an adult, mature person is the ability to say “no” at the right time, to get out of those situations or events where it is uncomfortable. Such an important skill is formed with the help of a special task, to confidently say “no” to 2-5 dubious proposals per day.

In the final task of the week, it is proposed to pay attention to those who need care, first of all, to relatives and friends. Thus, closer social contacts are formed with people outside the Internet.

Week 3

Having done tremendous work on themselves, having comprehended their behavior in reality and in the virtual environment, in the third week, students reach the level of critical thinking. Students are invited to think about the content of the concept, to isolate the barriers to the formation of a critical understanding of the world (formed patterns, advertising, propaganda, etc.). As part of the consideration of comprehensive digitalization, it is necessary to pay special attention to the phenomenon known as “post-truth”. Post-truth can be understood as a strategic construction of information content in order to form a virtual reality that differs from the actual one by controlling consciousness through media sources. The desire to resist the post-truth finds the most effective expression in critical thinking, which becomes the basis of the 3rd section of the analyzed manuals. Possession by students of basic skills of critical thinking, according to the authors, provides protection from falling under the power of emotional reactions, biases and misconceptions (Ennis, 2018; Dekker, 2020). Such skills are considered key to a successful career

and a comfortable life. Everyone has the opportunity to master critical thinking by learning and developing reflective thinking habits.

In addition to critical thinking and the post-truth phenomenon, the thesaurus developed in this section includes concepts such as “thinking”, “non-critical thinking”, “metacognition”, “skepticism”, “bias”, “objectivity” and “fake news”.

We have developed a step-by-step strategy, adhering to which, anyone can take a more meaningful look at the world. Reliability of information is of great importance in new media. Often, we can find several interpretations of the same event, but how to understand which one is correct? For this, a guide to working with information was also developed. The next step is to decide what is right. Here students are offered the so-called “Descartes' Square”, a technique that allows them to evaluate the consequences of each of the possible choices. This “square” was deliberately left in the diary, on which the student can write down his answers to the questions of Descartes' Square. On the 4th day of the week on digital asceticism, it is proposed to train awareness and perception. As part of the task, it is necessary to read any philosophical quote aloud or to yourself, and then evaluate self-attitude to it: what was remembered, what thoughts it caused; it is proposed to put these thoughts on paper for greater awareness. It is recommended to complete the task within 10-30 minutes. Students can return to a quote or a word, a phrase from it in a few days and note how this pattern is embodied in life, what role it plays in it. Another practice for the formation of critical thinking is the “Venn Ring method”, which consists in an attempt to study a particular situation from different angles and understand the logic of ongoing events. For written work, two options are provided, 2 and 3 circles. Particular emphasis is placed on the intersection of these circles.

At the end of the 3rd week, a creative task is offered, to write a

syncwine poem, which is one of the forms of expressing one's own opinions, feelings and associations. On the 7th day, it can be a difficult, but interesting work on the critical reflection of the self, to portray yourself with the help of a mental map.

Week 4

Finally, when students have passed the stage of self-reflection, it is necessary to learn to take care of themselves. This caring means taking life into your own hands and realizing that understanding yourself is the key to shaping your own life path.

In our opinion, the application of stoic practices in the context of comprehensive digitalization is a very effective solution. The contemporary digital environment generates a significant amount of informational noise and temptations in which an individual can easily lose his identity. However, turning to stoic concepts allows a person to learn to live through the mind and coexist in harmony with the “digital nature”, extracting only what is necessary from it. If we consider the structure of media content from the position of the ancient Stoics, it becomes clear that the most attractive material is based on four affects that should be eliminated, pleasure, disgust, desire and fear. A skeptical view and adherence to the ethical principles of Stoic philosophy help to reveal the emptiness behind these affects (Pigliucci, 2021).

It is interesting to note that the content of the sections of the textbooks, to a certain extent, corresponds to the key virtues by Stoics. Thus, the development of digital footprint management skills contributes to the formation of a state of ataraxia and provides a sense of security in the digital environment. Resistance to cyberbullying is based on apathy, expressing the principle of dispassion. In turn, critical thinking can be compared with intellectual autarchy, striving for the spiritual independence of a

person.

Awareness of the value of the property of life is offered through familiarization with the “slow life” trend, the philosophy and principles of stoicism. Students can reflect on how they take care of themselves, write thoughts about themselves in a diary, work with any quote of the day and draw it as an association. Taking care of the self is taking care not only of your own body, but also of your mind and your psyche. It is necessary to understand how it is included in the natural world, the artificial world and the digital world. A person is able to fill his existence with various kinds of meanings and thoughts. As part of this philosophical practice, it is supposed to track the flow of one's thoughts, ideas and write everything down in a diary.

Working with the inner world is its enrichment. Reading helps us to understand the world of other people and ourselves. For this purpose, the practice of slow reading, the so-called “Pearl Diver”, has been developed. Such work with the text allows you to track the thoughts that arise when reading, to understand which fragments of the text make the greatest impression, and why they resonate in the mind and heart.

The last philosophical practice is work on the self. Only an experienced person can give advice to his friend on how to take care of himself. Thus, through the above-described philosophical practices on digital asceticism, the goal of the diary is to shift the focus from the digital space of a person to himself, his inner world and environment.

Digital Asceticism for Teachers

The diary “Digital Asceticism” can be used both independently by students and with the help of a mentor. The manual for teachers in a

concise and accessible form reflects all the features of working with the diary, explains the basic concepts: instant messengers, social networks, cyberbullying, etc. The manual for teachers presents detailed lesson plans, examples of presentations, and textbook material (Gulevataya et al., 2023). For older teachers, this manual is a good help in getting to know life in the digital world and will allow the older person himself to find opportunities for adaptation. For young teachers, this is an important methodological support that allows them to prepare well for a lesson on digital topics in a short time.

Classes with schoolchildren are recommended to be held weekly as part of a classroom hour, at which thematic presentations will be shown, the class teacher will explain the importance, the need for digital asceticism and competently pose questions that students should find answers to while working with the diary. During the class hour, students are also supposed to be included in philosophical practices.

The combination of personal and classroom work under the guidance of a teacher will help students to better understand others, their interests, problems and gain experience in a group experience. Such experience helps to form intra-group contacts, unite the class, and reduce the level of conflict. On the other hand, group work helps to better understand oneself and one's inner world in relation to others, one's uniqueness, which helps in the future to correctly set life priorities and set personal boundaries.

With proper use, both manuals have a huge potential for instilling and developing digital ascetic skills, which is one of the most important tasks in the era of digitalization.

Handwritten diary for a student of the digital age

Digitalization has made the algorithmization of thinking the dominant format, that is, the desire to use ready-made solutions or standard ways to find them, reducing creativity and the desire for creative search in a child of the digital age. For educational institutions, it is extremely important to educate students in competencies that allow them to consciously counter the deep-seated threats of human digitalization, which primarily affect cognitive activity itself (Inyushina, 2021). A modern student, however, like most people, is forced to process a huge flow of information. The human brain adapts to such information loads through fragmentation of information and clipping of thinking. At the same time, the ability for a holistic and systematic perception of reality has been questioned. Moreover, a person tends to consume simpler content. This can be countered by understanding the philosophical foundations of digital hygiene: “If the digital revolution has made significant adjustments to human life, now it is necessary to develop not just life hygiene, but digital life hygiene” (Bulanov, 2021).

As one of the methods of digital asceticism, the authors of the manual consider such a format of human activity (in understanding oneself and one's life) as the use of written practice, i.e. keeping a diary. A handwritten diary is a chronological sequence of dated entries of an individual. A feature of the diary is its fragmentation, non-linearity, violation of cause-and-effect relationships, intertextuality, auto-reflection, fundamental incompleteness and lack of a single plan (Unterhitzenberger, 2022). Daily entries can contain reflections on life, contributing to the understanding of personal experiences and their attitude to the world. Often keeping a diary is dictated by a person's desire to trace his own spiritual (intellectual) development. In this case, it acts as a projection of the author's internal states, on the basis of which introspection is carried

out for further self-care. Keeping a diary not only makes a person more receptive to his own existential experience, but also develops the ability to express his thoughts, feelings, experiences in verbal form. In general, the diary helps to organize individual experience, leads to the development of self-discipline, which is important for the development of an “I” (Radcliffe, 2013).

The handwritten diary will allow students to become more sensitive and attentive to their inner world, and will also help them find ways to solve meaningful life issues. The teacher, in turn, will make it easier to implement the requirements set for professional education: individualization and personalization, disclosure of the individual abilities of students. The handwritten diary is a form of authentic assessment, i.e., aimed at assisting the student in developing his ability to analyze his own activities, as well as the formation of a culture of thinking, logic, the ability to analyze, generalize, systematize, classify. Modern schoolchildren and students are forced to be trained in the conditions of technological innovations. The handwritten diary is a path to digital use that includes practices to reduce addiction to gadgets and social media, increase productivity and focus, and lessen stress and anxiety caused by the constant use of digital technologies.

Keeping a handwritten diary can help a student to better remember and organize his thoughts in the process of reflective work, the result of which is written down on paper. It captures important events and moments of a person’s life not just as an endless selection of photo and video materials that are easy to get on a smartphone, but as meaningfulness, feeling and integrity of life, which is a very important aspect from the point of view of philosophy. Naturally, this method does not solve the problem of “programmability” of digital life, but it allows students to critically relate to this new reality.

Conclusion

In the increasing digitalization of society, the trend of rejecting the excessiveness of the digital benefits of civilization matures and becomes actual. The actualization of the topic in this case is not associated with the ubiquity of the practice of diary proposed by the authors, but indicates the consonance of time, i.e., very significant and essential precisely for this stage of development of society. This practice is not yet total and is aimed at a partial and meaningful limitation of digital goods in everyday life, as well as at developing the necessary skills for an active, wholesome and effective life of students in the context of digitalization.

One of the forms of reasonable consumption of information is digital asceticism, which arose as an assumption about the possible limitation of the experience of a person's presence in the contemporary digital world, which opens up wider opportunities for students to learn and master ways to effectively use digital addiction prevention tools, digital hygiene skills, to improve the quality of life. Special attention in the article was paid to the mobile application "Digital Asceticism". Despite the apparent contradiction (using a smartphone to overcome digital addiction), this application has a significant advantage, its format is the most intuitive for the younger generation, which is the target audience of this set of educational resources. Consequently, the use of digital technologies reduces the psychological barrier and provides a comfortable entry into the learning process, avoiding associations with traditional educational practices.

Further promotion of this methodology by integrating it into the educational process as a specialized course is a logical and progressive direction of development. Today's learners have limited skills and competencies to work effectively and safely in the digital environment. The knowledge, skills and abilities developed by this

cycle of educational materials can significantly close the existing gaps in the understanding of the digital environment among students. At the moment, the team of authors has conducted a pilot experiment on the use of a diary and a manual for teachers in one of the Russian schools. The data goes through the processing stage, after which a decision will be made on the need to adjust the experiment; the course, tools and the need to use additional methods to evaluate the effectiveness of the proposed materials. In the future, a large-scale pedagogical experiment is planned to test the methodology of digital asceticism. It is expected that more than three hundred schools in Russia will take part in the experiment in 2023-2024.

Combining handwritten notes with a mobile diary app is a promising direction for a deeper, more organized and innovative diary experience. By combining traditional and modern methods, users can enjoy the best aspects of both approaches, unlocking their potential for self-expression and reflection. The study of the possibilities of implementing the proposed pedagogical technology for the prevention and correction of digital addiction of students made it possible to establish that its use opens up wider opportunities for students to learn and master ways to effectively use the means of dependence on gadgets and services on the Internet, to improve physical, mental health and social well-being.

We invite colleagues interested in the problems of developing digital skills to improve the quality of life to take part in our study and are ready to provide all the necessary materials in English upon request.

References

Abakumova, I., Grishina, A., & Zvezdina, G. (2022). Psychological predictors of students' behavior in internet under information uncertainty. *International Journal of Cognitive Research in Science*,

Engineering and Education (IJCRSEE), 10(3), 17-26. DOI: 10.1007/978-3-031-21432-5_298

Amran, K., & Hassan, M. S. (2023). A Review on Cyberbullying Prevention in Social Media among Adolescents. *International Journal of Academic Research in Business and Social Sciences*, 13(3), 174–189.

Baranova E. & Shvecov G. (2021) Methods and tools for analyzing the student's digital footprint when mastering the educational route. *Perspektivy nauki i obrazovaniya*, 2(50), 415-430. doi: 10.32744/pse.2021.2.29

Bulanov, V. Philosophical foundations of digital hygiene of life. *Bulletin of the Tver State University. Philosophy Series*, 3, 68-75.

Burlet, R. E., & Panahon, C. J. (2022). Cyberbullying Prevention and Intervention: Current Practices in Schools. *Journal of Prevention and Health Promotion*, 3(2), 258–267. <https://doi.org/10.1177/26320770221087454>

Dekker, T. J. (2020). Teaching critical thinking through engagement with multiplicity. *Thinking Skills and Creativity*, 37. <https://doi.org/10.1016/j.tsc.2020.100701>

Ennis, R. H. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>

Grednovskaya, E., Gulevataya, A., Milyaeva, E., Penner, R., & Sosnovskikh, E. (2023). *Digital Asceticism: a Textbook for Practical Classes*. Publishing Center of SUSU.

Guerin, L. (2022). Modern Stoicism and Its Usefulness in Fostering Resilience. *Crisis, Stress, and Human Resilience: An International Journal*, 3(4), 138–143.

- Gulevataya, A., Valeeva, E., Milyaeva, E., Penner, R., & Sosnovskikh, E. (2023). *Digital Asceticism: Educational and Methodological Manual*. Chelyabinsk: Publishing Center of SUSU.
- Heidegger, M. (1998). *Parmenides*. Indiana University Press.
- Inyushina, I. A. (2021). PHILOSOPHICAL DIARY AS A TECHNIQUE FOR FORMING THINKING CULTURE. *Education and Science without Limits: Fundamental and Applied Researches*, (13), 126–131. <https://doi.org/10.36683/2500-249x/2021-13/126-13>
- Jansen, N., & Hinz, O. (2022). Inferring opinion leadership from digital footprints. *Journal of Business Research*, 139, 1123–1137. <https://doi.org/10.1016/j.jbusres.2021.10.032>
- Kuprashvili, H. (2023) Problems in Ensuring Informational and Psychological Security. Psychology and modern times. *Collection of the papers of the scientific-practical conference*, 68-75.
- Lapchik, D. M., Fedorova, G. A., & Gaidamak, E. S. (2021). Digital footprint in the educational environment as a regulator of student vocational guidance to the teaching profession. *J. Sib. Fed. Univ. Humanit. soc. sci.*, 14(9), 1388–1398. DOI: 10.17516/1997-1370-0827
- Macaulay, P. J. R., Betts, L. R., Stiller, J., & Kellezi, B. (2022). Bystander responses to cyberbullying: The role of perceived severity, publicity, anonymity, type of cyberbullying, and victim response. *Computers in Human Behavior*, 131. <https://doi.org/10.1016/j.chb.2022.107238>
- MacMillan, S. (2019). THE RELEVANCY OF MODERN STOICISM FOR HIGHER EDUCATION. *In ICERI2019 Proceedings* (Vol. 1, pp. 3339–3344). IATED. <https://doi.org/10.21125/iceri.2019.0856>

- Mendieta, E. (2014). Philosophy's Paralipomena: Diaries, Notebooks, and Letters. *The Journal of Speculative Philosophy*, 28(4), 413–421. <https://doi.org/10.5325/jspecphil.28.4.0413>
- Patchin, J. W., & Hinduja, S. (Eds.). (2011). *Cyberbullying Prevention and Response: Expert Perspectives* (1st ed.). Routledge. <https://doi.org/10.4324/9780203818312>
- Pavlov, A. (2021). The “Analytic” Line of the Modern Stoicism: L. Backer, W. Irvine, M. Pigliucci. *Philosophy. Journal of the Higher School of Economics*, 5(1), 33–52. <https://doi.org/10.17323/2587-8719-2021-1-33-52>
- Pigliucci, M. (2017). *How to Be a Stoic*. Basic Books.
- Pittaway, D. A. (2020). Digital Hygiene. *Filosofia Theoretica*, 9(3):33-47.
- Ponomarev A. V., Zolotareva M. V. (2022). Digitalization of education and the specifics of collecting a digital footprint of students of creative specialties. *Izvestiya vuzov. Investicii*, 2(12), 286–293. <https://doi.org/10.21285/2227-2917-2022-2-286-293>
- Radcliffe, L. (2013). Qualitative diaries: uncovering the complexities of work-life decision-making. *Qualitative Research in Organizations and Management*, 2(8), 163-180. <https://doi.org/10.1108/QROM-04-2012-1058>
- Sklar, A. (2017). Sound, Smart, and Safe: A Plea for Teaching Good Digital Hygiene. *LEARNing Landscapes*, 10(2), 39–43. <https://doi.org/10.36510/learnland.v10i2.799>
- Surmelioglu, Y., & Seferoglu, S. S. (2019). An examination of digital footprint awareness and digital experiences of higher education students. *World Journal on Educational Technology*:

Current Issues, 11(1), 48–64.

<https://doi.org/10.18844/wjet.v11i1.4009>

Uriu, D. (2023). JIZAI Technology as an Informational Body Modification—Toward Fostering Aesthetic Values and Social Ethics. In: Inami, M. (eds) *Theory of JIZAI Body*. Springer, Singapore. https://doi.org/10.1007/978-981-99-2003-7_8

Zhu, C., Huang, S., Evans, R., & Zhang, W. (2021, March 11). Cyberbullying Among Adolescents and Children: A Comprehensive Review of the Global Situation, Risk Factors, and Preventive Measures. *Frontiers in Public Health*. *Frontiers Media S.A.* <https://doi.org/10.3389/fpubh.2021.634909>

OSWALD SPENGLER: VISIONES SOBRE LA TECNIFICACIÓN¹

OSWALD SPENGLER: VIEWS ON TECHNIFICATION

GONZALO MANZULLO
CONICET, Argentina
gonzalomanzullo@gmail.com

RECIBIDO: 06/10/2023

ACEPTADO: 24/11/2023

Resumen: En el presente escrito proponemos detenernos en la obra de Oswald Spengler para recuperar sus diversas posiciones alrededor de la cuestión de la técnica y la tecnificación modernas, bajo la premisa de que ello nos permite elucubrar algunos correlatos para el despliegue de lo político en la época, según una cierta lectura de la historia, especialmente la Modernidad occidental. Así, nuestra hipótesis conecta técnica, historia y modernidad occidental en cuanto a la pregunta por el despliegue de lo político, amenazado en esa época. Nos detendremos fundamentalmente en los escritos aparecidos a partir del fin de la Primera Guerra Mundial, obviando la producción teórica previa, pues consideramos que aquél hecho oficia de puntapié para la reflexión del autor sobre la tecnificación, que sistematizamos aquí analíticamente en cuatro apartados o posiciones al respecto.

Palabras clave: técnica; cultura; guerra; ser humano; modernidad

Abstract: The present paper proposes a preliminary approach to the work of Oswald Spengler in order to recover his various positions on the question of modern technique and technification, under the premise that this allows us to elucidate some correlates for the deployment of the political at the time, according to a certain reading of history, especially Western Modernity. Thus, our hypothesis connects technique, history and Western modernity in terms of the question of the deployment of the political, threatened at that time. We will focus mainly on the writings that appeared since the end of the First World War, since we consider that

¹ El presente escrito recupera los avances preliminares de mis investigaciones doctorales en proceso, como también las indagaciones acometidas en mi tesis de maestría, actualmente en proceso de escritura, dedicada parcialmente a este tema.

this fact serves as a kick-off for the author's reflections on technification, which we describe here analytically in four sections or positions in this regard.

Keywords: technique; culture; war; human being; modernity

Introducción

Theodor Adorno, quien tenía pocas razones para dar crédito a un personaje como Spengler, reconoce sin embargo, en un ensayo breve titulado “Spengler tras el ocaso” que “el curso de la historia universal ha confirmado sus prognosis inmediatas en una medida que tendría que asombrar si se recordaran aún aquéllas” (Adorno, 1962, p. 47). El eco de estas palabras es el combustible para la recuperación de la obra de Oswald Spengler que emprendemos en las siguientes páginas, guiados por una indagación teórico-política que persigue pistas sobre la vinculación entre técnica y política en la Modernidad.

La referencia de Adorno hacia Spengler es sintomática: no es menor que un eminente miembro de la Escuela de Frankfurt remita a Spengler de esta manera², al tiempo que añade que, en la esfera de su labor, Spengler no encontró contrincante a su altura. Nos da la pauta de la relevancia que la figura de Spengler puede llegar a tener, incluso entre sus menos sospechados adeptos y luego de la catástrofe nazi que marcó el siglo XX³.

La intención aquí puede resumirse de la siguiente manera: detenernos en la obra de Oswald Spengler para recuperar sus

² Otro referente de la Escuela de Frankfurt, Franz Neumann, afirma que “no puede negarse su brillantez” como así también que “*La decadencia de Occidente* contiene observaciones que iluminan el paisaje a modo de relámpago y destacan nuevos aspectos que tendemos a hundir en la masa de detalles” (1983, p. 226).

³ Una reapropiación algo similar (de derecha a izquierda) tuvo lugar en los últimos años con la recepción de la obra de Carl Schmitt por parte de la tradición de pensamiento proveniente desde la izquierda latinoamericana y desde el posfundacionalismo europeo.

posiciones alrededor de la cuestión de la técnica, bajo la premisa de que allí se halla una visión del despliegue de lo político en la época, según una cierta lectura de la historia, especialmente la Modernidad occidental.

Si bien es conocido y explícito su tratamiento del tema en *El hombre y la técnica*, publicado en 1931, las lecturas e indagaciones sobre su obra, tanto en su tiempo como en el nuestro, muestran una vacancia: retoman únicamente este escrito en lugar de ofrecer una visión de conjunto sobre la técnica en toda su obra, cuestión esta última que creemos posible e intentamos más abajo⁴. Teniendo presente la mencionada vacancia, consideramos apropiado embarcarnos en nuestra empresa conectando la pregunta por la técnica y la tecnología con aquello que sí es considerado, con diversas valoraciones, el principal activo que nos legó Spengler: su visión de la historia y del futuro (Cagni y Massot, 1993; Herf, 1993; Losurdo, 2003; Traverso, 2003). Nuestra hipótesis conecta técnica, historia y modernidad occidental en cuanto a la pregunta por el despliegue de lo político, amenazado en esa época. Aquí, en vistas de nuestro interés temático y de la ingente obra de Spengler a pesar de su corta vida, nos detendremos fundamentalmente en los escritos aparecidos luego del fin de la Primera Guerra Mundial, obviando la producción teórica previa, pues consideramos que aquél hecho oficia de parteaguas desde el cuál, con la perspectiva de la derrota alemana y las imposiciones de la comunidad internacional hacia ese país luego del Tratado de Versalles, se puede enmarcar de mejor manera la meditación que Spengler hace sobre la técnica y su impacto político, sobre el telón de fondo de la decadencia de la civilización

⁴ El caso de aquellas indagaciones que sí se detienen en el tema, suelen considerar únicamente la obra de 1931, sin ofrecer una visión de conjunto del problema en la obra de Spengler, como es el caso de Piro (2013).

occidental y las nuevas formas futuras de cesarismo imperial producto del desorden posterior a la guerra mundial (Castro, 2017)⁵.

Mirada multidimensional sobre la técnica

En el marco de la denominada *Kriegsideologie* (caracterizada por la crítica a la utopía de la seguridad burguesa, la importancia de la noción de destino, el sacrificio y la *Gemeinschaft*) Losurdo destaca en la obra de Spengler la recuperación de la técnica separada y contrapuesta al materialismo de la civilización [*Zivilisation*] (*cfr.* Losurdo, 2003, p. 180)⁶. Más allá de lo preciso o no del diagnóstico de Losurdo, debemos comenzar estipulando que la reflexión spengleriana sobre la técnica no aparece una sola vez y de manera conclusiva en su obra, ni siquiera resulta inicialmente central. Antes que ser una cuestión abierta y resuelta con la obra publicada en 1931, señalaremos, más bien, apariciones sucesivas de la cuestión, con modulaciones y diferentes énfasis al calor de otras preocupaciones, a veces más y otras menos coyunturales, pero siempre relevantes para la reflexión sobre el impacto político de la tecnificación.

Sostenemos que la inquietud por la cuestión de la técnica surge tímida e inicialmente al calor del inicio de la primera posguerra, con

⁵ “Con su esquema culturalista y determinista (aunque el autor esquive esta última calificación), Spengler aliviaba a los alemanes de la carga del fracaso bélico porque la decadencia, más que germana, pertenecía a toda la civilización occidental” (Castro, 2017, p. 44). Hay que decir que el mote determinista ha sido atribuido “con mayor o menos justicia a autores tan dispares como Karl Marx, Ernst Jünger, Martin Heidegger” entre otros (Diéguez, 2005, p. 70).

⁶ Tal vez lo más interesante de la lectura de Herf sobre Spengler es que propone su rechazo al racionalismo como un rasgo y una acción ejercida desde el lugar de la modernidad y no contra ella (*cfr.* 1993, p. 116). Como indican Cagni y Massot, Spengler era al fin un producto de su época, “había de estar de acuerdo con las ideas del Occidente en decadencia”, pero fue su “notable inteligencia” la que le permitió relativizar su propia procedencia (1993, p. 431).

la reflexión de Spengler sobre la cultura y la historia —que Ortega y Gasset denomina como una filosofía de la historia (*cfr.* 2002)—, en particular sobre el devenir de lo que el alemán denomina la cultura occidental, una de las ocho culturas superiores que analiza en *La decadencia de Occidente*. Como veremos a continuación con mayor detalle, es en el pasaje de signo decadente⁷ de la cultura a la civilización, proceso inevitable para el destino de toda cultura considerada como organismo biológico, donde tiene lugar un movimiento de abstracción en el cual las formas culturales se vacían de la vitalidad que las caracterizó en sus etapas previas. A la manera de las plantas del reino vegetal, las culturas atraviesan la juventud, la madurez y llegan inexorablemente a la decrepitud. La cultura occidental se encuentra según Spengler en esa última etapa desde el siglo XIX. En este marco, la técnica aparece en la obra, inicialmente, como un apéndice de este gran proceso de abstracción, que coincide con la vida en las grandes urbes, la aparición de las masas y la supremacía del dinero en las sociedades.

Sin embargo, esa caracterización inicial de toda tecnificación se modificará al calor de la reflexión puntual sobre el devenir de la cultura occidental, tanto dentro de *La decadencia...* como en obras posteriores, complejizando y modificando su relevancia: el problema de la técnica aparece como una recta que es intersecada desde otras direcciones, obteniendo así diferentes trayectorias de pensamiento que vamos a repasar de aquí en adelante. De esta forma, trataremos de reponer una mirada multidimensional que comprenda estas diferentes vías de acercamiento al problema que podemos encontrar

⁷ Aunque la categoría de decadencia y su original *Untergang* (que sin mayores diferencias puede ser también traducido como hundimiento o ruina) parecen denotar cierto carácter negativo del proceso, Spengler reconoce en su correspondencia que en tanto se considera realista y no romántico, reconoce la fuerza de la civilización y no la rechaza más allá de su opinión personal (*cfr.* Spengler, 1966, p. 64).

en la obra de Oswald Spengler, con sus contradicciones, límites y continuidades.

Veremos que la tecnificación como proceso ineludible del pasaje de toda cultura a su fase civilizada, y la trayectoria particular de la civilización occidental a ese respecto, se engarza con la concepción de la técnica vista no ya desde las culturas superiores, sino desde la historia del ser humano en su origen y en comparación con los demás seres vivos, entendida como táctica para la vida y en su carácter primitivo. El carácter distintivo y relevante de la técnica moderna, fáustica, aunque presente, se verá matizado por momentos en la escritura de Spengler. A ello se superpondrá también la subordinación del problema de la tecnificación a cuestiones más coyunturales del debate político de la época, tales como la emergencia del capital financiero, las formas especulativas de la propiedad privada y el predominio de la economía dineraria, como formas de vida que amenazan la primacía de lo político, que Spengler intentará rehabilitar con su escritura.

Estos diferentes niveles de problematización de la cuestión de la técnica, que arriba procuramos distinguir a los fines analíticos, se interpenetran tanto *entre* las sucesivas obras que analizaremos como también *al interior* de cada una de ellas, de manera que en la exposición de Spengler no siempre resultan explícitos ni sencillos de discernir. Por ello, para ordenar nuestra exposición elegimos construir los apartados de acuerdo con los diversos enfoques de la cuestión que procuramos identificar, puesto que no resulta suficiente ni exhaustivo distinguir las diversas posiciones de Spengler con un criterio cronológico o sencillamente agrupando en conjuntos los escritos que abordaremos.

Desarrollaremos así cuatro secciones cuyo orden de aparición no implica ninguna distinción de importancia, sino que se orientan al fin de mostrar cómo el problema de la técnica surge en conexión con otros tópicos, desde distintas trayectorias y perspectivas, como un

problema que se hunde por momentos y en otras oportunidades vuelve a emerger y flota en la obra del autor, revelando ciertas modulaciones, discontinuidades y énfasis disímiles en el corpus seleccionado.

La técnica en la visión morfológica y fisiognómica de la historia

Previo a toda caracterización puntual de la cultura occidental, Spengler elabora una perspectiva histórica, autodenominada morfológica y fisiognómica e inspirada en Goethe y Nietzsche, en el marco de la cual la cuestión de la técnica tiene un lugar más bien recurrente antes que distintivo, y lo político es un aspecto más junto a otros que se mueve al son del ciclo vital cultural, tendiente a la decadencia.

La perspectiva del autor implica que el número de las formas en que se manifiesta la historia es limitado, de manera tal que edades, épocas, situaciones, personas se repiten en forma típica y la analogía resulta el primer mecanismo de comparación, aunque no el más apropiado para el análisis. Desde aquel lugar, la historia humana es el conjunto de enormes ciclos vitales de las culturas, cada una con un yo y una personalidad, como individuos de orden superior, activos y pensantes, tales como la cultura china, occidental o la antigua, con estadios por los que ha de pasar y orden entre ellos que no admite excepción.

Reversionando la formulación thatcheriana, la humanidad no existe, sino solo culturas individuales que imprimen su forma propia en la vida de los seres humanos, que nacen, crecen en una sublime ausencia de todo fin y propósito como las flores en el campo, y luego de florecer se marchitan para no revivir jamás. Cada cultura es

tratada como un gran individuo, y sus períodos como biográficos con su sentido propio⁸.

Siguiendo esa misión, Spengler delimita con su método, que pone de relieve las corrientes del simbolismo y la estructura metafísica subyacentes⁹, las grandes características de las ocho culturas consideradas superiores. Así, indica diferentes formas de lidiar con el tiempo histórico de acuerdo con las culturas: la cultura antigua tiene un sentimiento mítico de la historia a través del cual su historia es más bien poesía. La cultura india, por su parte, carece del sentimiento del cuándo y es por ello ahistórica, desapegada de cualquier tipo de memoria. La cultura egipcia encontraba dificultades para olvidar¹⁰, a diferencia de la cultura india, de manera que, tal como la cultura occidental, buscaba precisión en nombres, libros y personas. Igualmente, la cultura china revela un profundo sentimiento histórico con su concepto fundamental del *tao*, intraducible a nuestro idioma, según el alemán.

La perspectiva morfológica y fisiognómica promovida por Spengler¹¹ permite colocar una cultura junto a otra y buscar lo típico

⁸ De acuerdo con esta concepción cíclica, el fin de la historia está dentro de ella misma, sin que haya *telos* como punto final (*cfr.* Cagni y Massot, 1993). No hay metasentido ni metahistoria, por ello mismo.

⁹ Spengler reconoce en su segundo tomo de *La decadencia...* que muchas de las ideas allí volcadas pertenecen a un volumen dedicado exclusivamente a la cuestión de la metafísica, que se proponía publicar a la brevedad pero que nunca vio la luz.

¹⁰ De allí también se revela con la momificación su preocupación por la perduración a futuro.

¹¹ Aquí aparece un problema: mientras un hombre puede educarse para la física, el historiador spengleriano, en cambio, nace. Está guiado por un sentimiento que no se aprende, sino que parece más bien depender de un cierto don, de una sensibilidad especial para la visión morfológica y fisiognómica. Es la mirada profunda, que penetra en lo esencial, la que extrae de las contingencias las formas puras, ocultas en lo más hondo y que no se dejan descubrir fácilmente. Por eso reconoce en el segundo tomo de *La decadencia...*, en una más que esquemática distinción, que existen tipos de hombres que propenden a ver todo el mundo bajo

en los mudables destinos de ellas, lo necesario en el tropel de las contingencias. Es posible encontrar, por ejemplo, conexiones formales entre la antigua forma política de la polis y la geometría euclidiana, o entre “la perspectiva del espacio, en la pintura occidental, y la superación del espacio por ferrocarriles, teléfonos y armamentos” (Spengler, 2002, p. 35).

A partir de esto, no hay subordinación ni prioridad, como tampoco propiedades de causa y efecto o diferencias de valor entre fenómenos políticos, religiosos o artísticos, sociales o económicos, por decir. Lo que a un fenómeno particular le da importancia es su simbolismo en relación con el alma (entendida como potencialidad a desplegarse) de cada cultura, sin tener en cuenta bondad y maldad o superioridad. De acuerdo al desciframiento del alma de cada cultura, Spengler enfatiza la posibilidad de encontrar conexiones en diversas manifestaciones (economía, política, arquitectura, arte, escritura, matemáticas) que dan cuenta de una unidad *solidaria* (Cagni y Massot, 1993) de sentido en su relación con el tiempo y el espacio.

En el devenir de cada cultura como organismo, según el esquema spengleriano, cada cultura tiene su propia civilización, como destino de decadencia inevitable: es el extremo más artificioso a que puede llegar una especie superior de hombres, su remate con íntima necesidad. La civilización pura consiste en la gradual disolución de formas ya muertas, que se han tornado inorgánicas. Si el nacimiento de toda cultura las ata a un determinado suelo, como sucede con las plantas, el desarrollo de su alma, en su despliegue, la aleja de esas raíces y la acerca a la espiritualidad. Cuando agota sus posibilidades

la especie del sino o destino y otros bajo el prisma de la causalidad. El aldeano, el guerrero, el político, el caudillo, entre una larga enumeración que se separa de aquellos hombres que son espirituales: científicos, sacerdotes, santos, idealistas, ideólogos. En ese sentido, Spengler toma partido por César y Napoleón en lugar de Platón y Rousseau (*cf.* Spengler, 1993).

(en el arte, el Estado, la lengua, ciencia), se transforma en civilización. Pero ese tronco seco puede permanecer mucho tiempo en el bosque sin caerse. La decadencia es en este sentido cumplimiento interior y exterior.

La tecnificación, por el momento, resulta un rasgo propio de toda cultura que ha transitado hacia la civilización, como consecuencia del proceso de decadencia y abstracción de sus formas y alma distintivos. Pero desde lo que el propio autor denomina como una nueva visión de la historia –una filosofía del destino [*Schicksal*] primera en su clase¹²– también se intenta vislumbrar y predecir este destino en lo que atañe a la cultura occidental a la que Spengler pertenece, incluyendo la predicción de los estadios que aún no ha alcanzado. Pues entiende que, como el ciclo de toda cultura cualquiera, el porvenir de Occidente es limitado en su forma y duración, tiene un término. Por eso puede ser estudiado y previsto en sus rasgos esenciales¹³. Es decir, a pesar de que sus reflexiones y

¹² Lejos de toda pretensión de humildad, el autor piensa su enfoque como una revolución copernicana, en tanto mira las culturas y la historia sin adoptar una posición privilegiada desde ninguna de ellas. Esto inauguró una recepción de su obra en la clave del relativismo cultural, pues efectivamente Spengler critica toda idea de validez universal para acceder con las categorías del pensamiento occidental al de otras culturas como una conclusión falsa, de la misma manera como el griego no puede acceder al nuestro, porque las manifestaciones de otras culturas hablan otra lengua y no pueden ser captadas por el observador foráneo (Waismann, 1960).

¹³ En ese sentido la primacía del destino como elemento transversal a toda la obra cuestiona la idea de libertad: no tenemos libertad de hacer esto o aquello, es lo prefijado o nada. En un texto breve de 1921 titulado “*Pessimismus?*”, Spengler sostiene que su idea central es la de destino [*Schicksalsgedanke*], según la cual “cada hecho es una casualidad, imprevisto e incalculable, pero sabemos con la misma certeza, con la imagen de otras culturas ante nosotros, que el curso y el espíritu del futuro no son una casualidad, tanto en el individuo como en la vida de una cultura, que pueden completarse en un proceso espléndido o ponerse en peligro, marchitarse, destruirse, pero no desviarse en sentido y dirección por la libre decisión del agente.” (Spengler, 1938a, p. 63, traducción propia). En ese

análisis abarcan las culturas superiores por él determinadas, Spengler va a delimitar y caracterizar el alma particular de la cultura occidental, para luego desentrañar, en particular, la significación orgánica y simbólica de los complejos morfológicos de carácter político, artístico, espiritual, social que le son propios a su época contemporánea, de decadencia, que abarca entre los 1800 y los 2000, en el conjunto de la cultura occidental.

En el período en que esta obra se escribe, el autor detecta una identidad de la decadencia occidental con el carácter culminante del helenismo y su tránsito hacia la época romana, caracterizada por su estricto sentido de los hechos, el extremo artificioso de toda cultura que mencionamos anteriormente: una época sin alma, sin filosofía, sin arte, de hombres animales hasta la brutalidad, pendientes del éxito material y sin escrúpulos. En la cultura occidental, este período coincide con el surgimiento de las grandes urbes como centros espirituales que impactan en el mundo entero. Con la civilización se confirma una importancia creciente de la sociedad en desmedro del Estado. Tanto en la Antigüedad como en las postrimerías de la cultura occidental, que podríamos llamar sin más Modernidad, la idea del Estado sucumbe ante el poder de los factores económicos, desaparece la idea del servicio al Estado y comienza la voluntad de potencia, que excluye la tradición. En lo que atañe a la política mundial, el imperialismo es un símbolo típico de las postrimerías, es civilización pura. Si el hombre culto dirige su energía hacia dentro, el civilizado lo hace hacia fuera. La política en la etapa civilizada sería expansionismo. En este estadio, el dinero se torna preponderante como elemento abstracto inorgánico, poder

sentido, es una incógnita si el destino [*Schicksal*] resulta una fuerza que, siguiendo con las metáforas relativas al ciclo orgánico vital, puede estar dormido y despertar según el momento, si es una marcha implacable siempre o puede ser interrumpido por otras fuerzas.

omnímodo desprovisto de relación con el campo fructífero y con los valores de una originaria economía de la vida.

Según Spengler nos encontramos en el estadio civilizatorio de la cultura occidental, en la inevitable fase de decadencia donde merece más la pena construir un motor de aviación que una nueva y superflua teoría de la percepción. Es la época de la extensión¹⁴ y no de la profundidad: carente de toda producción artística elevada o de metafísica.

Hasta aquí, la visión histórica de Spengler contempla que toda cultura tiene su fase decadente y civilizada. En ella tendría lugar el desarrollo de las formas más abstractas de vida, el dominio del principio de la causalidad, del dinero, el culto a lo útil y ello incluye también las formas de la tecnificación más profunda. Así, resultaría difícil aún postular una particularidad en la técnica moderna y un consecuente impacto político específico de ella. Se perfila así lo que constituirá el hilo de nuestro próximo apartado, a saber: la visión de Spengler sobre el carácter distintivo de la cultura occidental en su fase civilizada en lo que atañe al problema de la técnica. Veremos a continuación que el carácter distintivo de la técnica moderna será el punto al que arribará Spengler una vez delimitada el alma fáustica de la cultura occidental. Siendo esta última condición de aquella¹⁵. Lo político se verá especialmente afectado en ese marco.

¹⁴ Pertenece específicamente al alma fáustica propia de la cultura occidental una tendencia apasionada hacia el infinito en lo concerniente a la extensión del tiempo y el espacio. Por ello, según Spengler, la idea de historia universal, sólo puede pertenecer a la cultura occidental, es *nuestra*, porque el antiguo y el indio no se representaban la historia en su devenir.

¹⁵ De hecho, pueden encontrarse en múltiples escritos del autor los límites que la técnica fáustica encuentra, por ejemplo, para penetrar el espíritu ruso. Lo mismo ocurre respecto del pensamiento predominantemente económico y del dinero propiamente fáustico-occidental.

El destino del alma fáustica: la técnica moderna occidental

El primer tomo de *La decadencia...* no solo contiene la premisa de la tecnificación como la ineludible faceta de toda cultura que arribó a la civilización, también se describen allí las propiedades del alma fáustica, a partir de las cuales se halla una determinada concepción de la técnica occidental y moderna que afecta de especial modo lo político.

El anhelo a lo ilimitado parece ser un patrimonio de lo occidental como distintivo. Nuestro espacio cósmico infinito no existe para el hombre antiguo. Lo mismo en cuanto a la forma de vivenciar el tiempo, en la que Spengler se detiene especialmente: respecto de la cultura antigua (y en particular de los griegos), indios y occidentales tienen una consideración mayor del tiempo, por eso en la cultura occidental el reloj tiene especial importancia. Desde esta óptica aparece también una fuerte ligazón entre la forma occidental de sentir el tiempo y los desarrollos técnicos, con el elemento distintivo de la memoria (*cfr.* Spengler, 2002). De hecho, el Estado, como artilugio técnico propio de la cultura occidental, aparece como símbolo supremo del tiempo. En el Estado se cifra, también, una voluntad de futuro relacionada con el dar orden. Ello tiene incluso su correlato en la vida económica, pues las culturas menos apegadas al tiempo viven el día a día.

Es propio del hombre civilizado de toda cultura superior tratar de despojar de arcanos y misterios el mundo, pero no todas las culturas, con su alma distintiva, resuelven esto del mismo modo. Por eso, si bien el artilugio técnico del idioma y del número son factores presentes en toda cultura y su desarrollo, resulta especialmente relevante el desenlace que tiene la tecnificación en la fase civilizada de la cultura occidental, de acuerdo con su alma fáustica: esta tiende de manera apasionada hacia el infinito. A nuestra imagen del

universo no le corresponde la suma calculable de las cosas materiales sino un espacio infinito donde lo visible es lo condicionado.

La cultura, por su parte, vendría a ser el cuerpo precedero que expresa el alma. Es “el único idioma por medio del cual puede un alma decir lo que sufre” (2002, p. 330). De este modo, tenemos el alma fáustica (diferente por ejemplo del alma apolínea y mágica), que se distingue por un símbolo primario, el espacio puro sin límites y cuyo cuerpo es la cultura occidental (iniciada en el siglo X en plenas llanuras nórdicas). El afán de descubrir es la tendencia propia de la naturaleza fáustica. Por eso el descubrimiento del nuevo mundo, de la circulación de la sangre y del sistema copernicano ocurrieron al mismo tiempo y con un sentido idéntico. Poco antes se descubrió la pólvora y, con ello, el arma de largo alcance y la imprenta, o sea la escritura de largo alcance.

A pesar de su pretendida rigurosidad y de la extensión de esta obra, en las caracterizaciones del arte, la ciencia, la religión y la arquitectura de las culturas, Spengler se mueve sobre todo en la distinción entre los tres tipos de almas mencionados más arriba, mientras que faltan precisiones sobre el estadio (culto o civilizado) al que corresponden sus referencias y ejemplos¹⁶.

De todos modos, el lugar de los avances técnicos está subordinado a la necesidad que revelan en cada cultura: hay descubrimientos que tienen el *pathos* de un gran símbolo *necesario* y otros que son simples juegos de ingenio. El barco a vapor fue lo segundo para los apolíneos y lo primero para el hombre fáustico: “Un invento, y sus aplicaciones, es profundo o superficial según el rango que ocupa en el conjunto del macrocosmos” (2002, p. 576). Esto permite aventurar que la cultura antigua podría haber tenido los elementos necesarios

¹⁶ En estas comparaciones generales es donde da la sensación que “Spengler sacrifica la verdad a la estética” (Cagni y Massot, 1993, p. 11) en cuanto a que en su labor comparativo de las culturas “ha parado mientes en ciertos fenómenos esenciales, olvidándose, por entero, de otros” (1993, p. 114).

para desarrollar la máquina de vapor, pero “faltaba la predisposición anímico-mental para hacerla” (Cagni y Massot, 1993, p. 298). Si en la visión universalista y relativista de las culturas que repasamos en el apartado anterior la técnica *es y expresa* una etapa, ahora tenemos una visión particularista donde podemos marcar el punto de quiebre, en la cultura occidental, donde la técnica (fáustica) *hace* época y afecta todo a su alrededor, incluido lo político.

Existe un elemento adicional que cabe adosar al alma fáustica a partir de la época gótica: ella deviene cultura de la voluntad. En las hazañas de Copérnico, Colón y Napoleón se manifiesta la identidad occidental entre espacio y voluntad: se trata del dominio del espacio cósmico. El alma fáustica posee una disposición eminentemente histórica porque el yo tiene un carácter constructivo y orientado al hacer, rige al mundo por la forma y esto implica igualmente un imperativo moral: la ética fáustica es una ascensión en el perfeccionamiento del yo y su mejoramiento, desde Santo Tomás de Aquino hasta Kant. Este afán de dominar el curso del mundo es también lo que constituye el contenido ético de la vida en la humanidad fáustica.

La realización de lo universal y permanente es fáustica. Existe un “anhelo indomable que empuja el hombre fáustico a los descubrimientos y las superaciones.” (2002, p. 556). Con el descubrimiento de América, el occidente europeo se transforma en provincia de un conjunto gigantesco y a partir de ese momento la cultura occidental adquiere carácter planetario.

Las premisas de la voluntad de poder nietzscheana son abrazadas por el alma fáustica, que desde el siglo XIX le dio una forma mecánico-utilitaria. El imperialismo resulta la aspiración íntima de esta época, donde aparecen como característicos tanto el cesarismo como el nihilismo de parte del hombre inteligente de la urbe, harto de la ciudad postrimera y de la cultura, que odia el simbolismo de su cultura, nihilismo del que ninguna cultura superior puede escapar.

Así, podemos distinguir las dos fases de la existencia occidental antes y después de su tránsito hacia la civilización (alrededor del 1800¹⁷). La civilización occidental es vida rezagada, artificial, desarraigada de las grandes urbes, cuyas formas dibuja el intelecto. Es un mecanismo producto del anquilosamiento. Si bien el tránsito de la cultura a la civilización parece describirse como un descenso y poseer una connotación negativa, no hay que perder de vista que no cesa de ser un paso último ineludible en el devenir según el propio autor.

En suma, con el requisito previo de abandonar toda perspectiva progresiva y unitaria de la historia de la humanidad, por fantasiosa; Spengler apunta el meollo de los problemas de la modernidad occidental, incluida la tecnificación: el despliegue del alma fáustica en la época de la decadencia civilizatoria. Es decir, podemos deslindar el despliegue del alma fáustica en sus postrimerías, la civilización, respecto de su inicial fase culta.

El segundo tomo de *La decadencia*, sin desconocer la visión morfológico-universalista, parece radicalizar el carácter distintivo de la civilización occidental: más allá del necesario período civilizatorio de toda cultura, sólo la civilización occidental se extiende a toda la faz del orbe: se diferencia del período correspondiente en las demás culturas “sólo por su grandísima extensión” (Spengler, 1993, p. 50). También la caracterización del alma fáustica, en su coincidencia con la voluntad de poder, es refrendada, pues se destaca como un rasgo antiguo “la falta de

¹⁷ De hecho, Spengler dedica un escrito que sintetiza sus investigaciones sobre la historia mundial del segundo milenio precristiano, donde analiza los rasgos de la primitiva cultura histórica occidental y su actitud ante la vida, así como sus invenciones técnicas (cfr. *Zur Weltgeschichte des zweiten vorchristlichen Jahrtausends* [1935], *Reden und Aufsätze*, 1938b). La arrogancia, el sentimiento de victoria sobre el mundo entero, la lucha con el destino por desafío, por tremenda confianza en sí mismo, son desconocidos en esta etapa donde predomina la voluntad de comodidad y el miedo a la perturbación.

voluntad de potencia en la técnica” (Spengler, 1993, p. 58). Por citar solo un ejemplo, Spengler explica que el hombre fáustico se distingue “de cualquier otro tipo cultural por su insaciable afán de lejanía, que en última instancia es el que ocasiona la destrucción de las culturas mejicana y peruana” (1993, p. 61). Ese afán no encuentra otro ejemplo en la historia.

En un punto, el fenómeno que Spengler intenta destacar parece ser una fagocitación histórica de las demás culturas a manos de la occidental, que podemos inferir es producto de la coincidencia del alma fáustica con la voluntad de poder. Esto permite también el carácter autorreferencial específico de la cultura occidental, donde incluso lo más remoto de lo acontecido en un tiempo y lugar lejanos adquiere un sentido hondo en relación con su ser, que no pudieron tener los demás hombres, ni siquiera los próximos a esos acontecimientos. Pues se trata de una civilización cuyo teatro es la tierra entera. Esto es lo novedoso y por eso todo es histórico para ella. La cultura occidental, sería la excepción y no la regla. La idea de una historia universal es propia de ella¹⁸.

La única cultura que escapa, en cierta medida, al tránsito hacia la total inhistoricidad que le es consustancial a toda civilización es la occidental, que por su fuerza fáustica adquiere aún en su etapa civilizada un fuerte carácter histórico en la medida en que “se

¹⁸ La facultad de los hombres del siglo XX de percibir y escribir la historia, sería una señal que “delata que el camino se dirige hacia el abismo. Solo en las cumbres de las grandes culturas, en el momento en que estas verifican el tránsito a la civilización, sólo entonces aparece por un instante esa facultad de penetrante conocimiento” (Spengler, 1933, p. 23). Con esto se reafirma el carácter eminentemente histórico de la cultura occidental, pero también su trágica imposibilidad de escapar a todo ciclo vital propio de cualquier cultura, que acaba con la falta de forma histórica. Steiner llama la atención sobre esto, en la medida en que parece poner un freno al carácter cíclico de la historia, no parece vislumbrarse de manera clara ninguna otra gran cultura emerger en el tercer milenio, mientras la occidental, ya planetaria, se derrumba (2018, p. 7).

encamina, como ninguna otra en la historia, hacia un suicidio grandioso” (Cagni y Massot, 1993, p. 300). A diferencia del hombre histórico antiguo, que es *ahistórico* dado no comprende ni siente la vida más que en relación con el presente, el hombre fáustico occidental resulta *profundamente* histórico. Esto también provoca que los productos de la técnica occidental fáustica como la máquina de vapor o la electricidad, sean recibidos de manera inhistórica por culturas ajenas, como los chinos, pero no por ello con menor veneración religiosa¹⁹.

El ciclo común al desarrollo de toda cultura tiene una deriva distintiva con el alma fáustica: saber no es virtud como en el confucianismo, el budismo y el socratismo, sino que saber es poder para la civilización europeo-americana y “solo en ella tiene sentido esta creencia” (1993, p. 360). Es siempre un saber para utilizar. Ello provoca una tensión espiritual que se completa con la necesidad de liberación con cierto culto, mística contemplación, encanto de lo irracional, extraño, raro o absurdo. Así, la creencia verdadera “sigue siendo la creencia en los átomos y en los números; pero esta creencia necesita el abracadabra culto para ser tolerable a la larga” (1993, p. 362). Emerge, como etapa común al ciclo de toda cultura, lo que Spengler denomina segunda religiosidad. Una fe para la cual las demostraciones comienzan a ser juegos de palabras míseros y aburridos. Según Spengler, el mundo occidental aún está lejos, faltan muchas generaciones para llegar a este momento.

Es el tiempo del cesarismo en que florecen individuos magnos, tiempo que entre los occidentales provoca un adelantamiento de la técnica en cuanto a los avances materiales del combate y los aspectos táctico-estratégicos que alcanzan un pico con la Primera Guerra

¹⁹ De hecho, en escritos posteriores Spengler alerta sobre la posibilidad de pérdida del ‘monopolio’ de la técnica en caso de que, por ejemplo, Japón se apropie y utilice para su beneficio en el concierto de las grandes potencias, las tecnologías occidentales.

Mundial, “acuciada por las necesidades militares” (Spengler, 1993, p. 491). Estas necesidades militares parecen ser las que traccionan el adelantamiento técnico de la época postrera de la cultura occidental. El cesarismo, en la civilización occidental, parece coincidir con la intensificación de la guerra (considerada como equivalente a la vida, y siendo la política la continuación de la guerra por otros medios) y el papel creciente de la técnica.

La técnica como táctica primitiva de existencia y su autonomización

En el primer tomo de *La decadencia...* podemos encontrar también, y en tercera instancia, una anticipación de las posiciones (que incluyen una cierta antropología filosófica²⁰) expresadas más tarde en *El hombre y la técnica* [1931]: Spengler reconoce la existencia de una técnica primitiva e instintiva, presente tanto en el hombre como en otros seres vivos, respecto de la cual la intelección se independiza en el ser humano y se desenvuelve de manera independiente a partir de la costumbre de hablar con palabras. La dimensión mítica originaria de toda técnica primitiva cede lugar a la inclinación teórica, al desarrollo de un nuevo saber y una nueva técnica de orden superior: de carácter cultural, impulsada por la fe y destinada a conjurar el misterio como también el terror del mundo circundante. Toda la filosofía, toda la física, todo conocer remite al encantamiento y conjuro primitivos (*cf.* Spengler, 2002, p. 672).

Aunque esto no resulta del todo contradictorio con la técnica como fase ineludible de toda cultura decadente, ni con el carácter distintivo de la técnica fáustica, añade una capa más a las reflexiones al respecto: la técnica es táctica para la vida de todo ser vivo, hasta que traspasado cierto umbral en el desarrollo humano, la técnica se

²⁰ Al respecto, ver Piro (2013).

independiza de la vida y sus caminos se separan. Esta idea vuelve también en el segundo tomo de *La decadencia...* cuando el autor señala este momento como aquel en donde, en términos religiosos, se distingue el mito del culto: la vía de la teoría se independiza de la técnica en cuanto táctica de vida, y con ello se puede distinguir la humanidad primitiva de la humanidad superior. El mito incluye toda técnica, inclusive la palabra y el idioma, todo cuanto forma parte de la intención creyente de descifrar los secretos del mundo circundante, mientras que el culto implica la intención de manejarlos. De manera tal que toda técnica supone el talento espiritual del conjuro. Tanto la ciencia como la técnica en general traban una relación con la religiosidad, pertenecen a un fondo común que es el mito²¹.

Por otra parte, Spengler señala un camino sin salida para este desarrollo emancipado de la técnica desde su origen panreligioso: en primera instancia, como ya dijimos, no abandona nunca el suelo religioso, pero mientras que la teoría religiosa conduce a una práctica sacerdotal, la teoría científica queda disuelta en la práctica, en el saber técnico de la vida diaria. Explica que cuanto más “se emancipa el saber teórico de la aceptación creyente, tanto más se acerca a su propia anulación. Lo único que queda es la experiencia técnica” (1993, p. 315).

Esta elaboración de Spengler sobre el lazo entre religión y técnica se radicaliza en el marco del cristianismo y la cultura occidental. En la cultura y la conciencia china, por ejemplo, no hay una idea de la superación técnica de la naturaleza, explica. Es decir, Spengler enlaza la religión cristiana y sus fundamentos con el alma fáustica en el derrotero de la técnica moderna. La física occidental fue antes que sirvienta de la teología, “más bien la sierva de la voluntad técnica

²¹ Incluso el ateísmo, posta ineludible para la cultura occidental, como lo fuera también para la antigua y la árabe, se compagina muy bien con un afán vivo y anhelante de religiosidad auténtica (*cfr.* Spengler, 2002, p. 690).

de poderío, y por eso se orientó en el sentido matemático y experimental; fue esencialmente mecánica práctica” (1993, p. 350). Fue por completo técnica primero, y teórica después, es tan antigua como el hombre fáustico mismo. No ocurre algo parecido en ninguna otra cultura.

El autor afirma que el hombre antiguo y euclidiano no podía sino ser hostil al pensamiento de la técnica. La técnica antigua, concebida estrictamente como el afán inequívoco que supera los conocimientos prácticos, es decir, como dominio del mundo de los hechos, no existe. Solo en la moderna técnica occidental se evidencia la unión del conocimiento con la aplicación que “quiere reducir el mundo a su voluntad” (1993, p. 581)²².

La principal conclusión que Spengler extrae de ello es la inversión de los términos clásicos para concebir la relación con la técnica: si toda máquina necesita de un ingeniero o especialista que la controle, ahora es ella la que nos obliga, ella quiere ser servida y dirigida. La máquina trabaja y obliga a los hombres al trabajo, Spengler afirma que es por la máquina que se hace valiosa ahora la vida humana, y que toda la cultura “ha entrado en tal actividad de trabajo, que la tierra tiembla” (1993, p. 583). Estas máquinas que nuestra cultura construye “van tomando cada día formas menos humanas; van siendo cada día más ascéticas, místicas, esotéricas” (1993, p. 584). Se hacen cada día más espirituales, ya no hablan sus ruedas, cilindros y palancas.

El hombre fáustico se convierte en esclavo de su creación y con ello de la técnica en general. Pero este desarrollo es íntimo al alma

²² Hay una voluntad potente dedicada a ello, que es extraña a otras culturas: la cultura china ha hecho casi todos los descubrimientos occidentales (entre ellos la brújula, el telescopio, la imprenta, la pólvora, el papel, la porcelana); pero el chino sustrae cosas a la naturaleza sin violentarla. Siente bien el provecho de su saber y hace uso de éste, pero no se precipita sobre él para explotarlo. Toda la cultura occidental tiene alma de inventor, descubrir fue desde el primer día la pasión del occidental (*cfr.* Spengler, 1993).

fáustica y a nuestra cultura occidental. Se trata de la misma conclusión a la que Spengler arriba con *El hombre y la técnica*, donde se postula explícitamente la cuestión del significado de la técnica, como pregunta que cobra relevancia a partir de su relación con la cultura y la historia desde el siglo XIX. Pasado el optimismo del progreso y el superficial entusiasmo por las conquistas de la humanidad, se revela que no hay estado de perfecto equilibrio ni autosuficiencia en la tierra²³. Con el siglo XX queda claro para Spengler que la historia no tiene para nada en cuenta nuestras esperanzas y no se trata de lo que debería ser sino de lo que es y será.

Como señalamos antes, en 1931 Spengler afirma que, originalmente, la técnica es antiquísima: no es una particularidad histórica sino algo universal, que trasciende al hombre y abarca todos los animales. Por eso hay una técnica del león, aunque no utilice ninguna herramienta, o bien una técnica diplomática: “En toda lucha por un problema hay una técnica lógica.” (1933, p. 19). El hombre al que Spengler considera un animal de rapiña, tiene la necesidad de afirmarse luchando, venciendo y aniquilando, a diferencia de todo animal herbívoro, que nace para ser víctima y presa, cuya vida es ante todo defensiva, tiene otra índole anímica. En él reside ya la idea del dominio, afirma Spengler, pues en ese campo de batalla el hombre mide los objetos y condiciones del ataque. Un sentimiento de libertad, superioridad y poderío se desprende de estas ventajas. El mundo es la presa; y de este hecho, en último término, ha nacido *toda* la cultura humana.

Lo propio del hombre, lo distintivo respecto de la técnica de los demás seres vivos, de la técnica invariable e impersonal de la especie (el instinto), es su cualidad inventiva, aprendible y susceptible de

²³ “Todo hombre verdaderamente creador conoce y teme el vacío que subsigue a la terminación de una obra” (1933, p. 21). El hombre, que conoce la muerte de antemano y es consciente de ella a diferencia de los demás animales, no puede dejar de mirar hacia el fin en todo lo que hace.

desarrollo. La técnica humana, y sólo ella, es independiente de la vida de la especie humana, en tanto que resulta variable, consciente, voluntaria, personal. Esta es su grandeza y su fatalidad.

Pero aquí resulta indistinguible la descripción del ser humano a secas de las propiedades del alma fáustica: son una y la misma cosa el arte, la técnica y creación y posesión de cultura, pues todas ellas son expresión de la existencia²⁴. Se pierden de vista los bordes particulares de la tecnificación moderna y su condicionamiento para lo político: implicaría asumir que toda cultura tiene *in nuce* los problemas y premisas de la tecnificación tal como se desarrolló según el alma fáustica en la civilización occidental, que todos ellos son patrimonio imposible de enajenar para el ser humano en general, más allá de su cultura y correspondiente alma o simbolismo. Para ser más claros, si arte, técnica y cultura son una y la misma cosa, no resulta tan fácil sostener el carácter distintivo y problemático de la técnica fáustica en la cultura occidental y en la modernidad tal como lo hace Spengler en otros escritos, puesto que la voluntad de dominio y el afán de control respecto del mundo serían inherentes a los hombres en todo tiempo y lugar.

Sin grandes mediaciones, Spengler se columpia entre aquella visión y las propuestas tanto de la tecnificación apéndice del proceso civilizatorio de toda cultura y la técnica fáustica como atributo distintivo de la cultura occidental en su fase civilizada. En el devenir histórico narrado en el escrito de 1931, mientras la vida artificial se hace cada día más artificial, el trabajador aislado no comprende “ni el término ni la finalidad de todo el procedimiento; ni tampoco le importaban, siéndole indiferentes y acaso odiosos”, ahora, “desde el siglo XVIII, innumerables ‘manos’ trabajan en cosas de cuya función efectiva en la vida, incluso en la vida propia, nada saben ya

²⁴ Teniendo presente esa homología, tal vez podría trazarse una conexión con el diagnóstico freudiano del malestar en la cultura.

y en cuyos éxitos no participan lo más mínimo interiormente” (1933, pp. 94-95).

Spengler afirma que la técnica “se ha convertido en un misterio” (1933, p. 99). De manera que el mundo artificial atraviesa y envenena el natural, la civilización se convirtió ella misma en una máquina, todo lo hace o quiere hacerlo maquinalmente, en términos de energía y aprovechamiento, más allá del sentido profundo de toda realización a la que se ve impelida. Esto lleva a que “cada invención contenga la posibilidad y necesidad de nuevas invenciones, de que cada deseo cumplido despierte otros mil deseos”, esa es la maldición que pesa “sobre este tipo de vida, pero también la grandeza de su destino” (1933, pp. 75-76)²⁵. Cualquier ideal de progreso técnico como ahorro de trabajo humano o mejoramiento efectivo de las condiciones de vida se ve explícitamente contradicho.

El animal de rapiña, “se ha aprisionado a sí mismo” (1933, p. 75). La máquina como símbolo se convierte en una necesidad anímica y espiritual pero no vital. A la revuelta de la naturaleza contra el hombre, le sigue históricamente la revuelta de la máquina, retoño del hombre, contra su creador. Repetimos: se pierde la distinción entre el recorrido particular de la cultura occidental fáustica y la historia de las culturas en general, en el marco de una antropología filosófica spengleriana. No obstante, se sostiene la aquí también que la ilimitada voluntad de poder es tan antigua como la cultura fáustica misma, y ajena a las demás culturas.

Ante esta tragedia, no obstante, la técnica no puede agotarse por la falta de materiales o reservas, pues el pensamiento técnico, confía Spengler, “descubriría muy pronto otras fuerzas distintas”, mientras “el pensamiento, que en ella actúa, permanezca en la altura, sabrá siempre crear los medios necesarios para sus fines” (1933, pp. 101-102). Lo que sí decae es el pensamiento fáustico, que comienza a

²⁵ Desde ese punto de vista, resulta necesario el hecho de que ningún inventor “ha previsto nunca exactamente el efecto práctico de su acción” (1933, pp. 75-76).

hartarse de la técnica, en un cansancio que deriva en pacifismo en la lucha contra la naturaleza, en la búsqueda de formas vitales más sencillas y a eludir la esclavitud de la máquina²⁶, cunde en su lugar el ocultismo y el espiritismo, como sucedió en la época de Augusto en Roma (1933, pp. 103-104)²⁷.

La técnica maquinista acaba con el hombre fáustico y llegará un día en que se derrumbe, se olvidarán sus ferrocarriles y barcos de vapor: pareciera que al final del recorrido ya no se puede sostener la antedicha homología entre arte, cultura y técnica: la vida declinante de toda cultura, y especialmente de la cultura occidental, parece separarse de cualquier ocaso de la técnica, pues la máquina alcanzó finalmente el *perpetuum mobile*, lejos de todo sentido y fuera de la historia, descentrando todo protagonismo occidental y, finalmente, humano.

La técnica como apéndice de la supremacía del dinero: ocaso de la política

Spengler añadirá aún una capa más a la cuestión, tal vez empujado por preocupaciones mucho más coyunturales: la subordinación del problema de la técnica al predominio del dinero y los factores económicos abstractos.

Prusianismo y socialismo (1984) ahonda ya en 1919 en esta arista, que aparece de manera lateral también en el primer tomo de

²⁶ Spengler afirma que los jóvenes se dedican al deporte en vez de a los ensayos técnicos (*cf.* 1933).

²⁷ En el plano geopolítico, Spengler se pregunta por la posible pérdida del monopolio del saber técnico ostentado por la cultura occidental, del cual deriva la fuerza de su industria y la riqueza que posibilitó el dominio económico-político por parte de un grupo de naciones de sangre nórdica a principios del siglo XX. Anticipa así una revuelta de los pueblos explotados de las periferias (Oriente y África) contra sus señores ‘blancos’.

La decadencia: nos referimos a la relación entre la tecnificación y el imperio del dinero. Spengler afirma que en la civilización de Europa occidental ocurre que, con la democracia moderna, los dirigentes de masas no se enfrentan a los dirigentes del capital, como propone el marxismo, sino al propio dinero y a su poder anónimo²⁸. La cuestión es cuántos de los líderes pueden resistir este poder. Spengler confía en la capacidad de un orden monárquico-socialista prusiano de dejar sin efectos aquellos condicionamientos propios de la época de decadencia. Y con ello impone serias dudas sobre la capacidad de torcer el destino de nuestra propia cultura. Está en juego no sólo el destino de Alemania, sino el de toda la civilización occidental. Es la cuestión decisiva no sólo para Alemania, sino para el mundo, y debe resolverse *en Alemania para el mundo*. La pregunta que Spengler plantea es si, en el futuro, ¿gobernará el comercio al Estado o el Estado al comercio? (*cfr.* 1984).

Desde ese punto de vista, pareciera haber una cierta torsión en las preocupaciones de Spengler, que trastoca la relación planteada inicialmente entre la tecnificación y el dominio del dinero: si en un comienzo podíamos poner en paralelo ambos elementos como tendencias simultáneas o coincidentes propias de toda época civilizada en una cultura, y que, a lo sumo, en la cultura occidental, le imprime su alma fáustica caracteres determinados; ahora pareciera ser que el segundo elemento subordina al primero y lo deja como su apéndice. En ese sentido, Herf afirma que “el blanco de Spengler no era la máquina sino el dinero” (1993, p. 131) y ubica la creciente influencia de la economía sobre las demás esferas de la vida, y en particular al dinero, como los enemigos predilectos de Spengler. En

²⁸ En rigor, puede hallarse en la obra de Spengler la distinción entre una etapa de la democracia que pertenece al espíritu romántico y sus ideales sobre la dignidad humana, y otra posterior, donde se vuelve dependiente del dinero y la política deviene negocio en lugar de ser “consecución de los supremos objetivos de organización social” (Cagni y Massot, 1993, p. 308).

1919 el alemán señala la técnica política como un mero medio en el dominio del capital privado y el espíritu inglés, que impugna en cierta medida la democracia (señalada por Spengler como la forma de ese siglo, que prevalecerá más allá de las valoraciones de cada quién al respecto²⁹) y, en general, referencias a la técnica, propia del espíritu inglés, de amasar reservas de dinero.

Teniendo presentes los escritos más coyunturales del autor a lo largo de la década de 1920, el problema de la técnica y su subsunción por el predominio del dinero resulta capturado por el debate más general y propio de la época, sobre la primacía de la política sobre la economía, que para Spengler es una cuestión *decisiva*: en 1924 Spengler afirma que cambios decisivos en el plano de la economía del siglo XIX, tales como dominación del carbón y la aparición del capital financiero móvil (desarraigado de la economía real), introdujeron una masificación y una forma especulativa de la propiedad que alteró la política y parasita al Estado en general (*cfr. Politische Schrifte*, 1934b, traducción propia). En el prólogo de sus escritos políticos, reafirma Spengler que la política “es absolutamente lo primero y la economía lo segundo en la vida de los pueblos” (1934c, p. XII, traducción propia). Su contemporaneidad es el momento de anquilosamiento de la cultura occidental,

²⁹ En su segundo tomo de *La decadencia...*, Spengler explica que los medios del presente son todavía, por muchos años, parlamentarios: las elecciones y la prensa. Y más allá de lo que opinemos sobre ellos es necesario dominarlos. Por eso, el hombre de Estado sabe que extender un derecho electoral no es nada en comparación con la técnica de *hacer* elecciones (*cfr.* Spengler, 1993). En su correspondencia de 1924, Spengler confiesa, convencido en la oportunidad de y demanda de un gobierno nacionalista, su intención de comprar un diario para difundir el punto de vista nacionalista sobre la economía e intervenir en la coyuntura política (*cfr.* 1966).

podríamos decir que se trata de un mundo donde la economía y la técnica son certezas y no preguntas ni problemas³⁰.

Esta es, como ya dijimos, la época del advenimiento del cesarismo, que implica un retorno a los aspectos más primitivos de la vida y la política como choque de fuerzas por un botín. Sin embargo, Spengler deposita expectativas de que los hombres de cuño cesáreo, como magnitudes prepotentes, puedan hacer frente a la omnipotencia del dinero y a una época donde ese elemento triunfa sobre la democracia (y la política en general). A sus ojos, el cesarismo es la revancha de los poderes de la sangre luego de una etapa de señorío del espíritu y el dinero. Se produce entonces una lucha final entre economía y política, donde la política reconquista su imperio. Pero también vale remarcar que sería equivalente, desde las propias aseveraciones spenglerianas, a un imperio en un mundo donde ya no es posible la historia. Ante ello, hay que asumir que esta revancha puede rehabilitar la historia, o bien la alternativa de que ella nunca más será posible.

¿Qué sucede entonces con la cuestión de la técnica? En el final del segundo tomo de *La decadencia*, Spengler anuncia una lucha de los espíritus en el suelo de las urbes cosmopolitas: entre el pensamiento técnico, que quiere ser libre, y el pensamiento financiero. El dinero pretende ser, como siempre en toda civilización, la única fuerza. Sin embargo, esta propia lucha desemboca en la anterior: la dictadura del dinero, en la civilización fáustica como en cualquier otra, se acerca a un punto máximo natural, pues como es una forma del pensamiento, se extingue tan pronto como haya sido pensado hasta sus últimos confines el mundo económico. Luego de apoderarse del campo, de los oficios, invade la industria para convertir en presa y botín el trabajo productivo,

³⁰ No hay una orientación hacia la búsqueda de la verdad, sino que la guía es la eficacia: “Las hipótesis en la técnica no aspiran a ser exactas, sino útiles” (Spengler, 1993, p. 173).

cuando parece que la máquina y su séquito humano, que es la soberana del siglo, está en peligro de sucumbir, el dinero ya se halla “al término de sus éxitos, y comienza la última lucha, en que la civilización recibe su forma definitiva: la lucha entre el *dinero* y la *sangre*” (1993, p. 587).

Una batalla final de la sangre contra el dinero, del cesarismo contra la democracia, que quiebra la dictadura del dinero y su arma política. Se revela más fuerte el aspecto político de la vida luego del largo período de primacía del dinero: la espada vence sobre el dinero y la voluntad de dominio sobre la de botín. Así, Spengler se olvida por completo de la técnica, la máquina y el ingeniero, como fenómenos secundarios de una contradicción principal³¹. Como queda plasmado también en *Neue Formen der Weltpolitik* [1924], Spengler ve con especial preocupación la manera en que el poder financiero socava el poder de los Estados, en la medida que convierte poco a poco la gran política en objeto de consideraciones bancarias y de especulación en la Bolsa, transformando la concesión de créditos en una especie de protectorado financiero y, con ello, la política misma en una empresa comercial (*cfr.* 1934b). Allí también afirma que, finalmente, no es la industria la que domina la política, son las altas finanzas las que dominan a ambas³². Su generación, afirma, presencia las últimas victorias del dinero y siente llegar al sucesor, el cesarismo, con paso lento pero irresistible. Por eso para

³¹ Este punto es retomado en *Das Verhältnis von Wirtschaft und Steuerpolitik seit 1750* [1924], donde Spengler afirma que la industria de la máquina no es la verdadera novedad. Lo decisivo que ha transformado todo el organismo, no sólo de la vida económica, es la separación rápidamente creciente de la propiedad del propietario, como movimiento de partes cada vez mayores de las grandes fortunas nacionales (*cfr.* 1934d).

³² Lo mismo sostiene respecto de la República alemana, a la que considera no como una forma de Estado sino como una empresa que nació en los grandes sillones de terciopelo ducal y en las tabernas de Weimar (*cfr.* *Neubau des Deutschen Reiches*, 1934e)

todos ellos está circunscrita y reducida en un estrecho círculo la dirección de su voluntad y necesidad.

Este panorama que pinta Spengler en los escritos mencionados resulta particularmente plausible si tenemos presente como telón de fondo las exigentes condiciones de sumisión que impuso el Tratado de Versalles a la Alemania derrotada, especialmente los pagos dinerarios de reparaciones por la guerra. En estos escritos y hasta *Años decisivos*, de 1933, el cesarismo, además de ser una fase necesaria e ineludible (teniendo como espejo las experiencias de Italia, España y Rusia), parece abrigar una cierta esperanza para Spengler en la maximización de las capacidades bajo el liderazgo de grandes personalidades en la política del futuro³³, que rehabilitará la política como poder. Su obra de 1922 en adelante está recorrida por la transversal intención de sobreponer de nuevo política a economía. Con el cesarismo se reestablecería la subordinación de los éxitos económicos a los fines políticos.

Pero, siendo que el cesarismo es la fase final inevitable en toda civilización, y en el marco de una mirada histórica marcada por la idea transversal de *destino*, resulta difícil pensar qué hay de decisivo en *esos años* en comparación con otros momentos: precisamente aquello que lo vuelve decisivo es lo que pone en cuestión su carácter inevitable. Vale pensar que, en este punto, Spengler confunde sus previsiones con sus esperanzas, un diagnóstico con un deseo: admite

³³ La esperanza de Spengler radica en la probabilidad de que en el futuro, en todo el campo de la economía, la tecnología y las finanzas, así como en la política, la diplomacia y la guerra, un individuo supere siempre todos los factores desfavorables mencionados gracias al poder de su personalidad y a la grandeza de su posición (*cfr. Das heutige Verhältnis zwischen Wirtschaft und Weltpolitik* [1926], 1934f). Lo mismo se señala en *Aufgaben des Adels* [1924], cuando se explica que los ejércitos diezmados, las economías desbaratadas y los territorios coloniales perdidos, resultan todos ellos secundarios ante la existencia de un estrato dirigente, columna vertebral de toda nación, que tenga las capacidades superiores requeridas para esta época (*cfr. Reden und Aufsätze*, 1938c).

el peligro, a mediados del siglo XX, de que se prosiga lo que se quisiera combatir de manera tal que con soluciones intermedias y transiciones se posponga la venida del cesarismo sosteniendo la partidocracia, al liberalismo, bolchevismo y la adoración a las masas³⁴.

El problema radica en todo aquello que la ‘anarquía mundial socialista’ puede destruir en el camino al cesarismo, de forma tal que ya no haya material con que un César pueda alzar su creación. Al interior de la fatigada cultura occidental, Spengler juega un pleno con sus fichas por Alemania (*cfr.* 1934a, p. 179), pueblo cuyo pasado político ha tenido menos oportunidad y ocasión de derrochar su sangre preciosísima y sus grandes talentos: la *raza* del pueblo dormía y esperaba la llamada de una gran época.

Reflexiones finales

Hemos procurado mostrar las diversas vías de acceso posibles a la cuestión de la técnica y la tecnificación que Spengler sostuvo –a veces superpuestas de manera simultánea o consecutiva y no exentas de contradicciones entre sí– a lo largo su extensa obra dentro de su breve vida, como también dentro de cada uno de los escritos que elegimos analizar. Tomando en cuenta esa complejidad, ensayamos una propuesta: la mirada multidimensional como herramienta para dar cuenta de las diversas posiciones asumidas por Spengler al

³⁴ Inclusive, el acrecentamiento de la técnica, especialmente en lo que atañe a la medicina, genera un problema para Spengler de acuerdo a su perspectiva elitista y aristocrática, en tanto permite prolongar vidas que, a su criterio, no merecen ser vividas, impidiendo cualquier tipo de selección natural y priorizando la cantidad de días vividos por sobre todo contenido, producto de lo cual se incrementa la decadencia (*cfr.* Spengler, 1934a).

respecto, así como las dimensiones de la cuestión en la obra del alemán.

Resumimos brevemente cuáles serían, de acuerdo con nuestra interpretación, las diversas posiciones sobre el problema de la técnica restituidos arriba: en la primera posguerra la cuestión de la tecnificación aparece inicialmente, en el marco de la elaboración de la mirada autodenominada morfológica y fisiognómica sobre la historia –en franca oposición a las visiones materialistas y positivistas de la ciencia histórica–, como apéndice del pasaje necesario y destino ineludible del ciclo vital de toda cultura hacia el extremo más artificioso, la decadencia de la civilización, modo de vida desarraigado, sin energía, abstracto y caduco en el cual las formas culturales se vacían de la vitalidad que las caracterizó en sus etapas previas. Así, en cualquiera de las ocho culturas superiores que caracterizan la historia universal y que Spengler describe en sus dos volúmenes de *La decadencia...* podríamos encontrar una fase civilizada, al interior de la vida de cada cultura y donde la tecnificación, como apéndice de un macro-proceso, confluye o coincide con el desarrollo de grandes ciudades, cierta masificación y la llegada de la economía dineraria. La decadencia occidental resulta análoga al momento culminante del helenismo y su transición hacia la época romana en la cultura antigua, sin que haya mayor importancia de una sobre la otra, ni pregunta distintiva para lo político en este marco.

En segundo término, cuando Spengler se detiene a meditar, en ese y otros escritos de su producción, sobre la cultura occidental de manera puntual y específica, la trayectoria de la tecnificación adquiere un estilo distintivo: coincide con la vida urbana cosmopolita, la era de la política de masas, la democracia representativa y la supremacía del capitalismo financiero. Su fundamento es aquí la radicalización en la fase civilizada (tránsito que se inicia alrededor del 1800) de un núcleo duro de simbolismo

(distintivo para cada cultura) que Spengler determina como el ‘alma fáustica’, caracterizada por su anhelo de lo ilimitado, su afán de *extensión* infinitesimal de lo intangible, especialmente el tiempo y el espacio. De su alma se desprende una determinada relación con el tiempo, el espacio, la vida, la muerte, la técnica y finalmente, la historia que repercute en las formas artísticas, religiosas, arquitectónicas, pictóricas, políticas, económicas y científicas que emanan de cada cultura. De modo que, si la cultura occidental se revela como distintiva por su especial alma, también resulta distintivo y por eso relevante el período civilizado de la misma, con su consecuente tecnificación, para pensar el despliegue de lo político en la Modernidad de la que Spengler es contemporáneo.

Hay así una técnica que podemos sencillamente calificar de moderna, referida a un punto específico en la historia humana y de la cultura occidental. De ella se desprenden problemas propios y particulares como resultado de las pretensiones que su alma fáustica expresa. Spengler detecta una ligazón entre la forma occidental de sentir el tiempo y los desarrollos técnicos, pues se trata de una cultura con gran sentido histórico donde la memoria y la voluntad de perdurar en el futuro se destacan. De hecho, se trata de la única cultura que es consciente de su término ineludible, que conoce y teme el fin de su historia. La forma en que despliega su fase civilizada, puede sostenerse, se debe precisamente a su carácter profundamente histórico. Cuando en el tránsito hacia la civilización (en la edad gótica aproximadamente) los aspectos distintivos del alma fáustica se vinculan y confluyen con la voluntad de poder y dominio sobre la naturaleza, se produce un cambio histórico decisivo. El despliegue planetario de la técnica occidental en la Modernidad –al que le es correspondiente también una autorreferencialidad cronológica y geográfica– resulta característico y responde a ese *ethos*. Inclusive, el énfasis en el desenlace y futuro de la civilización occidental poco a poco va monopolizando las reflexiones de Spengler dejando atrás

la reflexión sobre las demás culturas –y sus correspondientes técnicas, si las hubiera–, aunque no el diagnóstico de los caracteres que son propios a toda fase civilizatoria (imperialismo, expansionismo, nivelación espiritual, carácter ametafísico, irreligión, escepticismo, sublimidad del dinero, materialismo, urbanidad, entre los más repetidos). Habiendo partido desde una perspectiva anclada en el relativismo cultural, la visión de Spengler parece revelar cierto desnivel hacia lo occidental³⁵.

En tercera instancia, Spengler elabora una concepción de la técnica no ya desde el punto de vista de las culturas superiores, sino desde la historia del ser humano en su origen y en comparación con los demás seres vivos a la manera de una antropología filosófica. Desde allí, la técnica es entendida como táctica para la vida en su carácter primitivo. Sin embargo, el diagnóstico reconecta con las visiones previas: la existencia de una técnica primitiva e instintiva, puesta al servicio de fines prácticos, económicos y bélicos, se ve finalmente superada en una “progresión en línea recta hasta la técnica de las máquinas en la cultura fáustica de nuestros días” (1993, p. 37). Desde los primeros pasos sostenidos por la premisa de la ligazón fuerte entre técnica y vida (así como entre técnica, arte y cultura), transita hacia la ruptura de relaciones entre ambas, al punto que la técnica y sus productos cobran vida propia e independiente de sus creadores, un proceso específicamente moderno y un parteaguas histórico donde los motores continúan su maquinar allende toda vida y como misterio para ella. La dimensión mítica originaria de toda técnica, orientada a descifrar los secretos del mundo circundante, se troca en culto con la intención de manejar esos secretos, dominarlos,

³⁵ En cierto sentido, la obra de 1931, *El hombre y la técnica*, cristaliza el derrotero desde el cual Spengler va de lo universal (la historia de las culturas) hacia lo particular (el problema de la cultura occidental) y por medio del cual la segunda etapa fagocita la primera, como si esa parte fuera representativa del todo.

hasta llegar a su propia anulación. Así se muestra una estrecha relación originaria e ineluctable entre fe, religión y técnica.

El espíritu fáustico y la civilización occidental encuentra su término mucho antes que la técnica soberana e independizada de todo impulso vital, y se vuelve esclavo de ella como de la dinámica de voluntad de poder que él mismo había promovido. Pasado el optimismo del progreso y el superficial entusiasmo por las conquistas de la humanidad, se revela que no hay estado de perfecto equilibrio ni autosuficiencia en la tierra: cualquier ideal de progreso técnico como ahorro de trabajo humano o mejoramiento efectivo de las condiciones de vida se ve explícitamente contradicho. El asunto es hasta tal punto *histórico* que en *El hombre y la técnica* [1931] se postula explícitamente la cuestión del significado de la técnica, como pregunta que cobra relevancia a partir de su relación con la cultura y la historia desde el siglo XIX. No puede dissociarse ese significado de la técnica del manejo particular impuesto por el alma fáustica de la cultura occidental en su fase civilizada³⁶ (y se prevé inclusive la problemática apropiación y uso de la tecnología fáustica por parte de culturas ajenas y *contra* la occidental³⁷). Finalmente, aunque se advierta que la técnica no alcanzó la cima en todos los campos, puede

³⁶ Sin proponerlo abiertamente, esta aseveración de hace casi un siglo abre un campo para continuar la indagación sobre otros modos de relación con la técnica y otras racionalidades diferentes de la predominante occidental tanto en el pasado como en nuestra época contemporánea. De hecho, este era el anhelo de Ortega y Gasset cuando afirmaba soñar desde hace años “con un posible curso en que se muestren frente a frente las técnicas de Occidente y las técnicas del Asia” (1977, p. 118). Un ejercicio en esa dirección es el emprendido por Yuk Hui (2020).

³⁷ Pues recordemos que, según el propio Spengler, la técnica humana es aprendible y, en esa medida, se vuelve universalizable (*cfr.* Majul, 2017). Pero esa universalización y ruptura del monopolio occidental, tal como Spengler prevé, implica una cierta pérdida: una cultura oriental o africana puede servirse de la técnica occidental pero no deja por eso de rechazarla interiormente: no la viviría y sentiría como lo hace el alma fáustica, con todas las ventajas y complicaciones que ello impone.

preverse que la máquina alcance el movimiento perpetuo lejos de la vida, de todo sentido y fuera de la historia, implicando de esa manera un problema eminentemente político en esta era.

Hay que señalar, sin embargo, que Spengler cae en contradicciones en cuanto despliega su análisis del ser humano como animal de rapiña, con su técnica variable, inventiva, aprendible y susceptible de desarrollo: arriba a una visión que permite asumir la idea de que cualquier cultura desplegada por el ser humano contiene *in nuce* los problemas y premisas de la tecnificación tal y como desemboca en la Modernidad, o bien que todos ellos son patrimonio imposible de enajenar para el ser humano en general, más allá de su cultura y correspondiente alma o simbolismo. Es decir, si arte, técnica y cultura son una y la misma cosa, no resulta tan fácil sostener el carácter distintivo y problemático de la técnica fáustica en la civilización occidental tal como lo proponía previamente el autor, puesto que haría de la voluntad de dominio y el afán de control respecto del mundo elementos inherentes a los hombres en todo tiempo y lugar. En ello se juega también la especificidad de la pregunta por lo político en la Modernidad. La contradicción no solamente no se resuelve, sino que ni siquiera se expone en cuanto tal en la obra del autor: convive de manera tensa con sus otras intuiciones. Según se enfatice uno u otro vértice, la relevancia histórica está en los seres humanos (respecto del conjunto de los seres vivos) o en los atributos particulares de la cultura occidental y su despliegue moderno (que pasa de ser mera consecuencia a ser causa, respectivamente).

En cuarto lugar, encontramos la subordinación del problema de la tecnificación a cuestiones más coyunturales del debate político de la época, tales como la emergencia del capital financiero y las formas especulativas de la propiedad privada en el marco del predominio de la economía dineraria, como formas de vida que amenazan la primacía de lo político (y que Spengler intentó rehabilitar). Si bien

no se rechaza de ninguna manera que la civilización de Europa occidental está dominada por la industria maquinista en todas las expresiones, es decir, que somos todos esclavos de la técnica; el dinero y su poder anónimo aparecen más bien como símbolos de la época aplastando toda autonomía de lo político, especialmente en el siglo XX. La tecnificación, en ese marco, acaba siendo un mero apéndice de un proceso más general donde la democracia representativa, los partidos políticos, la prensa y la política de masas se ven socavadas por ese poder del capital privado y el espíritu utilitarista inglés, ante el que sólo los líderes cesaristas parecen poder hacer frente, de acuerdo a las esperanzas que abriga Spengler³⁸ en el marco de la pregunta sobre si en el futuro gobernará el comercio al Estado o el Estado al comercio.

Forzado por una coyuntura bien delimitada por la posición de Alemania en el período de entreguerras, Spengler relega y subsume en esta etapa el problema de la técnica al predominio del dinero: resulta capturado por el debate, más general y propio de la época,

³⁸ En el derrotero de la tecnificación occidental, el cesarismo aparece como la revancha de la sangre y el retorno a las formas primitivas de existencia frente al mundo electrificado y maquina, una fase donde política, guerra y economía dejan de distinguirse y por el contrario se asemejan (por reducirse a su faz más originaria y simple). Por eso, el cesarismo se debate entre lo irresistible y lo mesiánico: aunque originalmente es el giro inexorable y el destino decadente que inaugura toda época inhistórica (pues conduce a la cultura a su lecho de muerte), adquiere en la civilización occidental un profundo carácter histórico por su urgencia decisiva a los ojos de Spengler. Aunque Steiner (2018) reprochó a Spengler no haber visto la salvación de Occidente en la voluntad, Cagni y Massot afirman que en nuestro autor encuentra asidero una combinación extraña de libertad y necesidad: en él “parece coexistir un cierto voluntarismo con un cierto grado de determinismo” (1993, p. 39), porque el hombre no deja de ser actor y sujeto de la historia, “hace lo que quiere *siempre sin trastocar los planes generales*” (1993, p. 41, las cursivas son originales). Los hombres de rango son los que expanden su libertad y sus posibilidades de actuar *de acuerdo a y en sintonía con esos planes generales*.

sobre de la primacía de la política ante la economía (fáustica), que para Spengler resulta una cuestión decisiva. Con ello, el tono del autor vira del diagnóstico hacia la proclama política, del pronóstico hacia el manifiesto, en torno a lo que el destino reclama en la época y especialmente a los alemanes³⁹. Un ejercicio que puede ser interpretado también, en cierto sentido, como un intento por rehabilitar la historia, a la luz de las trágicas visiones spenglerianas.

Creemos haber mostrado las diferentes aristas que, a nuestro criterio, Spengler ofreció –con mayor y menor importancia o subsidiariedad según el caso– a lo largo de su obra para pensar la cuestión de la técnica y el destino de lo político ante ella. Aristas siempre terciadas por una cierta forma de comprender la historia, especialmente la Modernidad y su futuro, como el lugar en donde se juega el destino de lo político.

Problemática, decisiva o consustancial a la vida, la técnica resulta eterna e impercedera, pero sobre todo condicionante para lo humano: mojón ante el que debemos detenernos a reflexionar, superando toda disyuntiva entre apología o demonización, si pretendemos comprender nuestra historia y ver el futuro. Algunas pistas para eso podemos encontrar en la obra de Oswald Spengler en la medida en que –volviendo al punto de partida de este escrito– muchas de sus ideas y pronósticos se han confirmado hasta hoy (Adorno, 1962; Sloterdijk, 2003; Steiner, 2018).

³⁹ Cagni y Massot señalan los dos propósitos no necesariamente encontrados y de distinta índole en la obra de Spengler: uno filosófico que aspira a predecir el provenir estudiando las culturas y sus formas. El otro, político, para despertar a las clases dirigentes de su letargo, pues Spengler es fiel a Occidente y desde ese lugar busca “anticiparle al hombre blanco los peligros avizorados en el espíritu europeo y mundial”, hacer que tome conciencia “de su destino y se anime a hacer ‘lo necesario’, pues en caso contrario, deberá limitarse a sufrir la debacle y no hacer nada” (1993, p. 22).

Referencias

- Adorno, T. (1962). “Spengler tras el ocaso”. En *Prismas. La crítica de la cultura y la sociedad*. Barcelona: Ediciones Ariel.
- Cagni, H. y Massot, V. (1993). *Spengler, pensador de la decadencia*. Buenos Aires: Grupo Editor Latinoamericano.
- Castro, F. (2017). “El concepto de civilización en Juan Donoso Cortés y Oswald Spengler”. En Laleff Ilieff (comp.) *Política y valores en la modernidad. Un recorrido teórico-político desde la muerte de Dios nietzscheana a las tribulaciones del período de entreguerras*. Buenos Aires: Instituto de Investigaciones Gino Germani, p. 44-51.
- Diéguez, A. (2005). “El determinismo tecnológico: indicaciones para su interpretación”. En *Argumentos de Razón Técnica*, núm. 8.
- Herf, J. (1993). *El modernismo reaccionario. Tecnología, cultura y política en Weimar y el Tercer Reich*. Buenos Aires: Fondo de Cultura Económica.
- Losurdo, D. (2003). *La comunidad, la muerte, Occidente. Heidegger y la “ideología de la guerra”*. Buenos Aires: Losada.
- Majul, O. (2017). “La sabiduría de la vida: el lugar de Nietzsche en El hombre y la técnica de Spengler”. En Laleff Ilieff (comp.) *Política y valores en la modernidad. Un recorrido teórico-político desde la muerte de Dios nietzscheana a las tribulaciones del período de entreguerras*. Buenos Aires: Instituto de Investigaciones Gino Germani, p. 30-37.
- Neumann, F. (1983). *Behemoth. Pensamiento y acción en el nacional-socialismo*. México: FCE.
- Ortega y Gasset, J. (2002). “Proemio”. En Spengler, O., *La decadencia de Occidente I. Bosquejo de una morfología de la historia universal*. Madrid: Espasa-Calpe.

- Ortega y Gasset, J. (1977). *Meditación de la técnica y otros ensayos*. Madrid: Revista de Occidente.
- Piro, P. (2013). “Dos meditaciones sobre la técnica: El hombre y la técnica de Oswald Spengler y Meditación de la técnica de Ortega y Gasset”. En *Revista Laguna*, núm. 32, abril, pp. 43-58.
- Sloterdijk, P. (2003). *Crítica de la razón cínica*. Madrid: Siruela.
- Spengler, O. (1933). *El hombre y la técnica*. Santiago de Chile: Ediciones Nueva Época.
- Spengler, O. (1934a). *Años decisivos*. Madrid: Espasa-Calpe.
- Spengler, O. (1934b). “Politische Pflichten der deutschen Jugend”. En *Politische Schriften*. München: C. H. Beck.
- Spengler, O. (1934c). “Vorwort”. En *Politische Schriften*. München: C. H. Beck.
- Spengler, O. (1934d). “Das Verhältnis von Wirtschaft und Steuerpolitik seit 1750”. En *Politische Schriften*. München: C. H. Beck.
- Spengler, O. (1934e). “Neubau des Deutschen Reiches”. En *Politische Schriften*. München: C. H. Beck.
- Spengler, O. (1934f). “Das heutige Verhältnis zwischen Wirtschaft und Weltpolitik”. En *Politische Schriften*. München: C. H. Beck.
- Spengler, O. (1938a). „Pessimismus?“. En *Reden und Aufsätze*. Altenmünster: Jazzybee Verlag. Edición de Kindle.
- Spengler, O. (1938b). „Zur Weltgeschichte des zweiten vorchristlichen Jahrtausends“. En *Reden und Aufsätze*. Altenmünster: Jazzybee Verlag. Edición de Kindle.
- Spengler, O. (1938c). „Aufgaben des Adels“. En *Reden und Aufsätze*. Altenmünster: Jazzybee Verlag. Edición de Kindle.
- Spengler, O. (1966). *Spengler letters (1913-1936)*. Londres: George Allen & Unwin LTD.

- Spengler, O. (1984). *Prusianismo y socialismo*. Buenos Aires: Editorial Struhart & Cia.
- Spengler, O. (1993). *La decadencia de Occidente II. Perspectivas de la historia universal*. Buenos Aires: Planeta de Agostini.
- Spengler, O. (2002). *La decadencia de Occidente I. Bosquejo de una morfología de la historia universal*. Madrid: Espasa-Calpe.
- Steiner, R. (2018). *Oswald Spengler, Prophet of World Chaos*. Anthroposophic Press.
- Traverso, E. (2003). *The origins of Nazi violence*. Nueva York: The New Press.
- Waismann, A. (1960). *El historicismo moderno*. Buenos Aires: Editorial Nova.
- Yuk Hui (2020). *Fragmentar el futuro. Ensayos sobre tecnodiversidad*. Buenos Aires: Caja Negra.

**ASPECTOS EPISTEMOLÓGICOS SOBRE LA VALIDACIÓN DE
MODELOS COMPUTACIONALES EN COSMOLOGÍA**

***EPISTEMOLOGICAL ASPECTS ON THE VALIDATION OF
COMPUTATIONAL MODELS IN COSMOLOGY***

MAXIMILIANO BOZZOLI

maxibozzoli@ffyh.unc.edu.ar

CONICET - Universidad Nacional de Córdoba, Argentina

MARIANO J. DOMÍNGUEZ

mariano.dominguez@unc.edu.ar

CONICET - Observatorio Astronómico de Córdoba
Universidad Nacional de Córdoba, Argentina

RECIBIDO: 10/10/2023

ACEPTADO: 24/11/2023

Resumen: Los modelos de simulación computacional están vinculados a los resultados obtenidos observacionalmente, logrando así su validación de diversas maneras. Así, las simulaciones cosmológicas hacen posible la generación de modelos de contraste, los cuales permiten la comparación entre los datos simulados con aquéllos alcanzados por la vía tradicional. Al no haber un solo protocolo de validación, las prácticas son heterogéneas. Una de las estrategias consiste en validar inicialmente las simulaciones con datos observacionales, otra alcanza la validación a través de la parametrización conforme a nuevos datos disponibles. Debido a que los resultados de las observaciones y de las simulaciones son comparables entre sí, ellos poseen el mismo estatus epistémico. Esto permite que sean viables observaciones a través de simulaciones. A partir de los modelos de contraste logrados, se concluirá que es factible el empleo de técnicas horizontales de validación entre simulaciones, en particular, entre el modelo cosmológico estándar y modelos alternativos.

Palabras clave: Prácticas científicas; validación; observación; modelos de simulación; cosmología

Abstract: Simulation models are closely linked to the results got from observational means, achieving their validation in different ways. Thus, cosmological simulations lead to the generation of contrast models, which allow comparison between simulated data with those achieved by the traditional way. Since there is not a single validation protocol, practices are heterogeneous. One of the strategies has to do with validating simulations initially through observational data, another one achieves validation according parametrization with new data available. Since observational data and simulated data can be compared to each other, they share eventually the same epistemic level or status. Consequently, observations via simulations are feasible. From the contrast models achieved, it is possible to conclude by saying that the use of horizontal validation techniques among simulations is viable, in particular, between the standard cosmological model and alternative models.

Keywords: Scientific practices; validation; observation; simulation models; cosmology

Introducción

Dada la relevancia de la computación en los laboratorios astronómicos actuales, se destaca como uno de sus frentes más importantes aquél que involucra a las simulaciones computacionales en el ámbito observacional de la cosmología contemporánea. Los modelos de simulación están estrechamente vinculados a los resultados obtenidos de forma observacional, logrando así su validación de diferentes maneras. Así, las simulaciones cosmológicas hacen posible la generación de modelos de contraste, los cuales permiten la comparación entre los datos simulados con aquéllos alcanzados por la vía tradicional (Bozzoli y Paz, 2010, 2011). Además, estos modelos son capaces de reproducir los procesos físicos que subyacen en cada instancia del modelo de observación considerado, tales como: errores metodológicos, efectos de proyección, ruidos instrumentales, etc.

En la literatura filosófica vigente, la evaluación de los modelos de simulación ha sido planteada por diferentes autores que sostienen

que tal valoración consiste en algo más que comparar los datos simulados con los datos observacionales/experimentales (Humphreys, 2013). De esta manera, con la escasa información sobre cómo este contraste puede llevarse a cabo, lo que ciertamente importa es concebir la capacidad del modelo computacional a la hora de proveer evidencia a favor de alguna hipótesis específica dada (Parker, 2008). Con este fin, la evaluación es caracterizada, en contextos ingenieriles y científicos, en términos de verificación y de validación. El primer caso consiste en saber si el proceso que determina el resultado de una simulación se aproxima a las soluciones arrojadas por las ecuaciones diferenciales, presentes en el modelo del fenómeno físico. Pese a que este método es inadecuado en aquellos casos donde las soluciones analíticas no estén disponibles, se recurre a otras técnicas indirectas a fin de establecer una comparación más confiable. A los propósitos de este trabajo, se dejará de lado esta clase de prácticas. Una de las estrategias del segundo caso consiste en cotejar los datos simulados con los datos observacionales, aun cuando estos últimos sean escasos. Esto debilita el contraste, logrando que este método empleado sea demasiado débil para establecer la evaluación de las simulaciones (Winsberg, 2010). La credibilidad de estas últimas está basada en los modelos físicos, en las observaciones y en la interacción entre ambos.

Al no haber un protocolo para validar tales simulaciones, surgen distintas formas de establecer vínculos entre el modelo del fenómeno físico bajo investigación y los datos empíricos adquiridos, dando lugar a posibles reinterpretaciones de las observaciones. En particular, en la cosmología, el modelo estándar (Λ CDM) postula que la distribución de la materia en el universo es lo suficientemente regular como para asumir que el mismo es homogéneo e isótropo. Además, las observaciones actuales de cierto tipo de supernovas (Ia) revelan que el universo se expande de forma acelerada. Ello conduce

a reintroducir la constante cosmológica en este modelo para dar cuenta de tal aceleración, postulando como causa de la misma una entidad exótica denominada “energía oscura”. Estudios recientes cuestionan la existencia de dicha entidad, sugiriendo una explicación alternativa, fuera de lo convencional, a través de una reinterpretación estadística de las bases de datos disponibles. Sin alterar la física del modelo canónico y teniendo en cuenta la estructura observable del universo a gran escala, estos astrónomos proponen un modelo del mismo con una “expansión diferencial”. Al realizar el modelo computacional, las simulaciones reproducen tales estructuras observadas, revelando cómo éstas alteran la expansión (Rácz et al, 2017). Atendiendo a este estudio de caso, y pese a las diferencias notorias entre los procesos observacionales y aquéllos computacionales, se intentará mostrar que las simulaciones, al alcanzar un nivel de validación sustentable, arrojan resultados que adquieren el mismo estatus evidencial que los obtenidos observacionalmente. De esta manera, es posible sostener que los modelos de simulación son herramientas indispensables de la cosmología actual y deben ser considerados instrumentos alternativos de la observación.

Aunque las prácticas de validación son heterogéneas, las mismas pueden ser identificadas, a grandes rasgos, siguiendo dos enfoques verticales. Uno pretende validar inicialmente los modelos computacionales con bases de datos observacionales disponibles (*top-down*). El otro intenta realizar la validación a través del ajuste de parámetros conforme a los resultados obtenidos mediante las observaciones, transfiriendo así el contenido empírico a simulaciones más sofisticadas (*bottom-up*) (Guillemot, 2010). Debido a que los datos obtenidos de las observaciones y de las simulaciones son comparables entre sí, ellos poseen eventualmente el mismo nivel o estatus epistémico. Este último aspecto permite que sean viables observaciones a través de simulaciones. A partir de los

modelos de contraste logrados, se concluirá que es factible el empleo de técnicas horizontales de validación (*from side to side*) entre simulaciones, en particular, entre el modelo cosmológico estándar (Λ CDM) y modelos alternativos (A ν ERA, por ejemplo).

Conexiones entre la observación y la simulación

Guillemot (2010) sostiene que, en la actualidad, son escasos los abordajes filosóficos sobre la validación de modelos a través de la comparación con las bases de datos observacionales, las cuales proveen el soporte evidencial que otorga confiabilidad a las simulaciones computacionales. Así, la credibilidad de estas últimas es lograda a través del contraste con las mediciones obtenidas por la vía tradicional. En este sentido, al explorar y analizar las interrelaciones entre diferentes capas de datos y los modelos de simulación, esta autora intenta rescatar el rol crucial de las prácticas de observación en la evaluación y la parametrización de dichos modelos.

Desde la perspectiva de la “concepción semántica” de las teorías, los modelos son principalmente representaciones que dan significado a los formalismos matemáticos presentados por aquéllas. Sin embargo, en las últimas décadas, se ha repensado la noción de modelo, aludiendo a la supremacía de ellos con respecto a la capacidad representacional de las teorías (Cartwright, 1999). En esta dirección, se ha caracterizado a los modelos como agentes autónomos, los cuales son relativamente independientes tanto de la teoría como del mundo físico (Morrison y Morgan, 1999). Esta autonomía permite que los modelos sean considerados como instrumentos de exploración en ambos dominios. Además, ellos poseen la capacidad de mediar activamente entre la teoría y los datos observacionales, lo cual permite que el modelado de fenómenos complejos no pueda ser reducido a meros cálculos de aquellas teorías

subyacentes. Aun cuando estas últimas den cuenta de los procesos físicos involucrados, los modelos de simulación incluyen cadenas inferenciales sofisticadas que transforman las estructuras teóricas en conocimiento concreto del fenómeno (Winsberg, 1999). Dado que los modelos y las bases de datos están interrelacionados, algunos autores sostienen que los primeros están “cargados” de datos, mientras que éstos últimos están sesgados por tales modelos. Dicha relación de interdependencia no es circular sino simbiótica, es decir, cada lado obtiene ventaja uno del otro (Edwards, 1999). Desde esta perspectiva, los modelos computacionales son tan confiables como aquellos tradicionales, ya que alcanzan cierto grado de credibilidad a través de sus vínculos con las observaciones y los experimentos. Las credenciales epistémicas de las simulaciones son logradas en función de las teorías y de los experimentos, como así también de los métodos de modelado más aceptados por la comunidad científica. Al igual que en las prácticas más ortodoxas, el desarrollo instrumental condiciona las técnicas asociadas, permitiendo algún tipo de estabilidad interna y auto-vindicación de las mismas a lo largo del tiempo.

Desde el surgimiento de los primeros modelos numéricos hasta la actualidad, la computación ha permitido la comprensión de una variedad de fenómenos en diferentes campos de las ciencias naturales. A través de estos modelos, se generan nuevos conocimientos que describen las interacciones entre los objetos y sus propiedades dentro de un sistema físico determinado. Bajo condiciones límite específicas, dichos modelos reproducen algunas características de los procesos involucrados y el comportamiento de tal sistema. De esta manera, el modelado computacional posibilita la exploración de fenómenos físicos logrados virtualmente. En este sentido, este tipo de experimentación guarda similitud con aquella llevada a cabo en los laboratorios. Por un lado, dada su capacidad representacional, es posible manipular los parámetros físicos y

chequear las hipótesis teóricas involucradas. Por el otro, los resultados arrojados por estos modelos generan bases de datos simulados, los cuales deben ser interpretados y medidos. Sin embargo, para Guillemot, hay una clara distinción entre los experimentos tradicionales y los computacionales. Estos últimos están basados en símbolos y dígitos y no poseen una confrontación directa con el mundo. No obstante, en ambos dominios está presente la selección y el ajuste de parámetros físicos. Así, la parametrización, además de asignar valores a las magnitudes implicadas, permite articular los cálculos provistos por las teorías con las mediciones obtenidas por las observaciones.

Aunque los modelos experimentales de laboratorio parecen alcanzar un significado más amplio que aquéllos numéricos, las simulaciones muestran, por el contrario, una imagen bastante nítida de la confluencia de diferentes técnicas. Por ejemplo, esta autora menciona que para poder chequear una hipótesis dada es necesario traducirla, mediante métodos numéricos específicos, en forma de un algoritmo que es insertado en el modelo computacional para luego ejecutarlo en una computadora. El modelo de simulación resultante, aunque se halla limitado a un ámbito virtual, se nutre de aspectos tanto teóricos como experimentales. Ella sostiene que la experimentación numérica es, por sí sola, insuficiente y que sólo alcanza su estatus como práctica de modelado cuando incorpora, de una u otra manera, datos observacionales que impregnen empíricamente al modelo y permitan la confrontación indirecta del mismo con el mundo físico. Aunque el rol de la computación es crucial en el modelado de fenómenos meteorológicos, astronómicos o físicos, otros ingredientes son necesarios para que estas prácticas sean eficientes. En particular, el avance tecnológico de los instrumentos de observación, desde equipos en tierra hasta satélites artificiales puestos en órbita, y de las técnicas desarrolladas para su empleo proporcionan la información necesaria para el modelado de

dichos fenómenos. No obstante, estas herramientas de la observación utilizan, eventualmente, otros auxiliares como recursos computacionales para la adquisición, reducción y análisis estadístico de datos. Algunos de ellos son usados en función del automatismo propio de los sistemas instrumentales actuales, otros son programas expertos necesarios para la identificación y clasificación de objetos, por ejemplo. De esta manera, la asimilación de datos en modelos es posible gracias a la sinergia entre las prácticas teóricas, observacionales e instrumentales. Como se notará más adelante, la modelización en la astronomía genera, dentro de la comunidad de los astrónomos observacionales, distintas sensibilidades hacia los aparatos tradicionales, o bien, hacia las computadoras. En este trabajo se sostendrá que, a través de estas últimas, como parte esencial de los sistemas multi-instrumentales más recientes, es posible llevar a cabo observaciones de fenómenos que involucran procesos astrofísicos sofisticados.

Guillemot cuestiona qué tipos de datos se confrontan con los resultados de las simulaciones. Más allá del estudio de caso que ella considera, se plantea que tales datos observacionales son altamente reconstruidos; por ello, existe la necesidad de considerar los diferentes mecanismos de producción de datos y sus respectivos modelos de datos asociados. La astronomía, al igual que otras ciencias naturales como la paleontología, es una disciplina histórica. Ello permite articular los registros observacionales del pasado con bases de datos actuales, completando la información faltante mediante el uso de modelos computacionales. A diferencia de la meteorología y de la climatología, la dinámica de la mayoría de los fenómenos astronómicos no puede alcanzarse en tiempos humanos, salvo objetos transitorios como aquéllos que acontecen en el sistema solar o ciertos tipos de estrellas variables o supernovas, entre otros. Con la asistencia de modelos de simulación, puede obtenerse una imagen de la evolución del universo como un todo y de las partes

que lo constituyen. Asimismo, la computación permite cerrar las grietas entre módulos de datos interconectados, ya sean observacionales o simulados, y estandarizarlos en entornos como el de los observatorios virtuales para su uso posterior en la exploración y minería de datos (Bozzoli y Paz, 2016).

Datos observacionales y datos simulados

Humphreys (2013) sostiene que los instrumentos producen imágenes que representan propiedades físicas particulares de un objeto observado. Estas propiedades causan ciertos efectos físicos en el sistema instrumental, permitiendo, luego de su detección, un registro permanente de ellas. Ante la variedad de procesos causales, inherentes a mecanismos de producción de datos eventualmente diferentes, los atributos observables se presentan en dicho sistema para ser, posteriormente, representados en una imagen epistémicamente accesible. Más allá de la distinción entre datos crudos y datos procesados, precisamente, entre datos instrumentales y datos observacionales, este autor afirma que para interpretar una imagen dada es necesario conocer no sólo lo que ella representa, sino saber cómo ha sido producida. Por otra parte, él asevera que son las propiedades, detectadas por los aparatos, las que son observadas y no los objetos en sí. De esto último se sigue que un objeto es observado sólo a través de sus propiedades o atributos físicos. Así, la clase de objeto que está siendo observado puede depender de qué propiedad está siendo detectada.

Humphreys propone, de esta manera, un realismo cuyas entidades se hallan constituidas por grupos de propiedades relacionadas y correlacionadas entre sí. En esta dirección, el descubrimiento de los procesos causales subyacentes a los objetos observados consiste en emplear instrumentos específicos que detecten uno o más atributos, usando las técnicas y auxiliares observacionales disponibles. Lo

principal, en este contexto, es alcanzar el entendimiento sobre el hecho o la situación de lo que, ocasionalmente, es la causa del fenómeno observado. A medida que se adquieren más conocimientos al respecto se descubren, gracias a los adelantos tecnológicos, nuevas propiedades y relaciones entre ellas que han permanecido ocultas hasta ese momento. Lo que Humphreys está mencionando aquí, consiste en definir las capas de cualidades o de atributos inmersos en el grupo de propiedades observadas a partir de detectores muy variados. En este sentido, los diversos vínculos que pueden guardar dentro de este cúmulo dependerá en última instancia de los tipos de mediciones logradas; esto es, diferentes estándares de medición dejarían abierta la posibilidad de medir una propiedad, pero de distintas maneras.

La razón por la cual los instrumentos son detectores de propiedades se debe a que los mismos están diseñados para ser capaces de repetir las mediciones de fenómenos físicos dados en lugares y momentos específicos. Según Humphreys, los aparatos detectan los atributos observables en sí, pero no sus instancias espacio-temporales. En otras palabras, él considera que el proceso de detección de una propiedad está desvinculado de su ubicación física. Sin embargo, dichas propiedades inician el proceso causal que va desde el fenómeno particular observado, a través del sistema instrumental que codifica la información, hasta la producción de una imagen que es accesible al sensorio humano. Las transformaciones que acontecen dentro de tal sistema son detalladas por la teoría inherente a esos procesos causales, la cual describe la entrada y la salida de los aparatos y de los auxiliares asociados. Así, la producción de datos ruidosos es controlada tanto en la fase de adquisición como de procesamiento, o sea, tanto en la reducción como en la interpretación de los mismos. Más allá de esto último, este autor sugiere que el uso de los aparatos no sólo provee evidencia de lo que una entidad es, la cual da origen a lo que se observa, sino

que el instrumento además brinda una representación aproximada del fenómeno en cuestión. Esta confiabilidad es alcanzada mediante las teorías que describen cómo opera el instrumento, a fin de descartar cualquier salida errónea en la generación de datos. Por lo tanto, al tener una comprensión general de la funcionalidad del sistema instrumental, se obtiene el control de ciertas partes dentro del proceso observacional, el cual es dominado por un marco teórico determinado.

Las imágenes crudas o directas no son las que proveen la información más relevante de lo observado, sino aquella representación visual que ha sido manipulada o procesada de alguna u otra manera. Al emplear técnicas de visualización científica, las composiciones de imágenes producidas por diferentes detectores proveen información adicional sobre los atributos físicos representados. Así, los cambios de representaciones visuales conducen, eventualmente, a la resolución de un problema en un contexto de búsqueda de nuevas cualidades observacionales (Bozzoli y Stasyszyn, 2020). De esta forma, el soporte evidencial está garantizado por el conocimiento adquirido en las prácticas observacionales. En este sentido, considerando la sofisticación de los procesos involucrados, la dependencia teórica de la observación es una virtud, más que un defecto.

Por otra parte, Humphreys enuncia un “principio de localización de la propiedad”, el cual afirma que: cada instancia de una propiedad causal posee toda la fuerza causal de esa propiedad (Humphreys, 2004). De esto se sigue que, al comprometerse ontológicamente con tipos de relaciones causales diferentes, no son los objetos sino sus propiedades los que causan los efectos que pueden ser observados. Así, en la clase de causalidad incluida en los procesos físicos de los sistemas instrumentales convencionales, como también en aquella de los sistemas computacionales, por ejemplo, las propiedades pueden ser observables por una u otra vía, respectivamente. En

consecuencia, no son las instancias espacio-temporales las que causan tales observables, sino las propiedades o atributos en sí. Incluso, tales efectos pueden ser producidos por un grupo de propiedades combinadas y no sólo por una cualidad aislada. Cabe destacar aquí que lo epistemológicamente relevante es considerar, por un lado, la reproductibilidad de los resultados, hecho que se apoya en que un efecto detectado puede darse en más de una ubicación posible; por el otro, la universalidad de las leyes físicas, donde ninguna referencia concreta es en principio establecida ya que restringe el alcance de las mismas. La mayoría de estas leyes son las que describen tanto los vínculos causales, como las correlaciones entre propiedades de fenómenos observados.

En otras palabras, la detección de propiedades físicas muy variadas, a través de plataformas multi-instrumentales, permiten la adquisición de datos a partir de una diversidad de mecanismos de producción. Estos últimos, como se hizo notar anteriormente, posibilitan la obtención de la información observacional del objeto de estudio. A los fines aquí presentes, esta perspectiva permite examinar el estatus de los datos logrados mediante las observaciones, como aquellos obtenidos por vía de las simulaciones computacionales. En esta dirección, al ignorar los orígenes de cómo se generaron los datos, es posible su reemplazo e intercambio. Siguiendo a Humphreys, estos datos pueden ser clasificados como datos empíricos, es decir, aquellos que tienen un origen concreto y son producidos por las interacciones físicas con los dispositivos de medición, y datos sobre (*about*) que contienen información acerca del modelo del fenómeno bajo investigación. Aunque los *datos(a)* pueden proveer información sobre las fuentes causales que dan origen a los *datos(e)*, no hay una clara referencia a ellos. Este autor afirma que los orígenes de los datos, sean o no materiales, son insuficientes para definir el contenido de los *datos(a)*. De esta manera, él propone extender la categoría de los datos empíricos a

datos sobre el origen e incluir aquellos provenientes de fuentes computacionales. Así, él plantea dos clases de aparatos. Por un lado, los instrumentos causales no computacionales reciben un proceso físico, como señal de entrada, el cual interactúa de manera tal que provee otro igual, como señal de salida. Por otro lado, los instrumentos causales computacionales operan de forma similar, salvo que, en algún punto de la interacción, la señal física se transforma en otra digital antes de producir la señal de salida. Es importante destacar que estos últimos aparatos involucran aspectos causales no sólo por la naturaleza de sus *inputs*, sino por las implementaciones computacionales llevadas a cabo durante el proceso.

Humphreys sitúa a dichos instrumentos entre aquellos puramente causales y las simulaciones. Esto se debe a que, en el mismo proceso computacional, se emplean inferencias inversas que permiten cerrar la “ida y vuelta” entre el detector y la fuente de los datos. Dado que la dirección de la causalidad se invierte, o sea, del *output* al *input*, esto no quebranta los procesos causales de tales instrumentos a la hora de obtener propiedades y asignar atributos observables. Así, el conocimiento que se obtiene de los procesos físicos y de los computacionales, involucrados en la producción de los datos, es útil tanto para su adquisición y reducción como para su procesamiento e interpretación. Los ruidos que provienen de los datos instrumentales, como también los errores que surgen de los datos computacionales, pueden ser eliminados a partir de métodos físicos o numéricos, discriminando así información relevante de irrelevante. Humphreys sostiene que el contenido informacional de los *datos(a)* puede referirse tanto al mecanismo de producción que los genera, como a las interpretaciones que se hagan de ellos en base al modelo del fenómeno que se está observando. Según él, este análisis puede ser replanteado en el ámbito de las simulaciones. De esta manera, el contenido de los *datos(a)* se referirá a su origen, es decir, al modelo

en el cual se basa la simulación; o bien, a la interpretación que se haga a luz del modelo del fenómeno simulado. Esta propuesta permite ponderar entre modelos de datos diferentes y comparar resultados observacionales con simulacionales, posibilitando observaciones a través de simulaciones y su eventual validación entre ellas.

Tipos de validación de modelos de simulación

Guillemot sostiene que la evaluación de los modelos meteorológicos es, al menos en principio, más simple y directa que en los modelos climáticos. De forma similar a estos últimos, la valoración de los modelos astrofísicos posee un alto contenido estadístico. En particular, las simulaciones cosmológicas están condicionadas, en parte, por valores de parámetros físicos que están ligados a bases de datos observacionales, tal como lo es la medición de la constante de Hubble para distintos estados evolutivos del universo. Dada la interacción de los diferentes procesos astrofísicos involucrados en estos modelos, se valida la representación de un fenómeno en particular de manera más o menos confiable. Ello depende del poder de síntesis de una simulación computacional en cuanto a su capacidad de generar fenómenos dentro del modelo, los cuales son predichos teóricamente. Este aspecto permite validar tanto el modelo, como las teorías implicadas. No obstante, debido a que las simulaciones son cada vez más sofisticadas, su validación, a través de las observaciones, puede resultar de la combinación de una hipótesis errónea y de una representación imprecisa del modelo. De esta manera, se logran resultados aceptables a partir de lo que los modeladores llaman “errores de compensación”. Según esta autora, una simulación puede resultar correcta, pero por una o varias razones equivocadas. En esta clase de modelos interviene una gran cantidad de mecanismos físicos y computacionales que van desde el diseño,

la construcción y la configuración de un sistema de computadoras, aisladas o unidas en un arreglo tipo clúster, hasta la arquitectura de los programas que corren en ellas. Esto posibilita un alto desempeño, logrando simulaciones de gran resolución, aunque de difícil validación. En este sentido, dada la complejidad de los fenómenos considerados, los investigadores tratan de determinar qué parámetros influyen sobre qué resultados y cómo lo hacen. Pese a que estas simulaciones son altamente susceptibles a las condiciones iniciales, el mismo modelo es también objeto de estudio.

Una característica interesante de las simulaciones cosmológicas es que no suelen ser evaluadas en términos generales, es decir, sólo a la luz del modelo estándar, canónico o concordante. Sino en cambio, lo que se pretende validar es la capacidad de las mismas a la hora de explicar diferentes propiedades de objetos y procesos astrofísicos específicos. Así, las evaluaciones son llevadas a cabo en sus respectivas dimensiones espaciales y temporales en cada nivel del modelo, o sea, pueden referirse a aspectos de la estructura del universo a gran escala, como así también a la forma de los halos que envuelven las galaxias que lo conforman. Según Guillemot, cada tipo de validación dependerá en última instancia del propósito con que es usado el modelo. Sin embargo, ella afirma que no hay un protocolo sistemático en la evaluación de los mismos, sólo basta con imaginar un procedimiento adecuado para confrontarlos con los datos disponibles. Como se mencionó anteriormente, esta autora identifica dos clases de validación: *top-down* y *bottom-up*. La primera consiste en comparar las simulaciones con bases de datos observacionales. Al ser inicialmente validados, estos modelos computacionales son capaces de reproducir las mismas propiedades físicas medidas con los instrumentos tradicionales. La segunda estrategia es empleada para seleccionar y ajustar parámetros en contraste con las observaciones disponibles. La cuestión central aquí es que los modeladores deben identificar las variables pertinentes al

fenómeno bajo investigación. De esta manera, la selección permite no sólo analizar los datos observacionales, sino también los simulados. Guillemot sostiene que la parametrización constituye un método de evaluación que permite la comparación entre modelos construidos a partir de los mismos criterios. Los investigadores restringen el espacio de parametrización de los experimentos numéricos a fin de abordar sólo aquellos comparables con la observación. Un aspecto importante de la evaluación *bottom-up* es que permite manipular datos observacionales y datos simulados, de tal forma que pueda reconstruirse y completarse la historia de los procesos físicos bajo estudio con información relevante. De esta forma, se hacen evidentes correlaciones que habían permanecido ocultas anteriormente. Como se notará a continuación, estas prácticas de validación reflejan diferencias sutiles tanto en el valor de las constantes medidas y de las observaciones de base, como en los criterios que subyacen en los modelos estándares y alternativos de la cosmología actual.

Validación de modelos cosmológicos

El modelo cosmológico estándar (Lambda CDM, por sus componentes más abundantes) ha sido desarrollado en base a la solución de Friedmann-Robertson-Walker que asume, a gran escala, homogeneidad e isotropía espacial. Las simulaciones cosmológicas tradicionales utilizan, para calcular la evolución de la expansión del universo, el valor promedio de su densidad de materia global, conforme a la ecuación de Friedmann. Sin embargo, la presencia y el crecimiento de estructuras, tales como cúmulos de galaxias, filamentos y vacíos cósmicos observados, violan manifiestamente la homogeneidad asumida. Es algo bien conocido e intensamente debatido desde hace tiempo, dada la naturaleza no lineal de las ecuaciones de Einstein, que estas inhomogeneidades locales influyen

sobre la tasa de expansión (Green and Wald, 2016; Buchert et al. 2015; Wiltshire, 2011). No obstante, considerando el éxito predictivo y el poder explicativo del modelo homogéneo Λ CDM se acepta, tácitamente, que los efectos de las inhomogeneidades observadas son débiles y por ende despreciables. La gran aceptación de este modelo canónico se debe a su concordancia con numerosas observaciones cosmológicas tales como: funciones de correlación y espectro de potencias de las estructuras observadas, abundancias de las mismas y su desarrollo, incluyendo los lineamientos generales de la evolución de las galaxias y sus propiedades, entre otras.

El problema surge al intentar realizar un modelo numérico consistente con las ecuaciones de la relatividad de Einstein, ya que el espacio-tiempo debe reaccionar a la presencia de las estructuras mencionadas que se forman en el transcurso de la misma simulación. Esto provoca que el mismo espacio-tiempo se torne intratable e incluso que el proceso de promediado se vuelva no trivial (Boero & Moreschi, 2017). No obstante, existen numerosos intentos de efectuar dichos cálculos, los cuales dependen, bajo ciertos criterios, de la introducción de escalas donde se realizan los promedios espacialmente. En particular, investigaciones recientes presentan un método alternativo que emplea las técnicas más aceptadas de las simulaciones de N cuerpos (Rácz et al, 2017). Estas prácticas, bien conocidas, permiten computar las tasas de expansión del universo local en función de promedios esféricos, redimensionando luego las fuerzas y velocidades allí presentes. Así, el modelo de simulación alternativo (AvERA) reproduce los resultados alcanzados por aquellas simulaciones que consideran un modelo cosmológico estándar, con condiciones iniciales similares, pero sin asumir un valor uniforme de la densidad de materia y, por consiguiente, de la constante de expansión. En contraposición con las simulaciones convencionales, las cuales descansan en aproximaciones que ignoran la influencia de la estructura del universo en su expansión, este

modelo se logra sin la presencia de una constante cosmológica.

A diferencia del estándar, el modelo alternativo toma en cuenta aspectos estructurales, simulando variadas regiones del universo que se expanden a distintas tasas. La expansión diferencial es originada por la formación de estructuras complejas de materia; mientras que en los modelos tradicionales la expansión acelerada es causada por la presencia de la, metafóricamente llamada, “energía oscura”. Debido a la similitud histórica en la expansión de tales modelos, es imposible distinguirlos mediante simples pruebas cosmológicas geométricas como el tamaño, la abundancia o la luminosidad de los objetos involucrados.

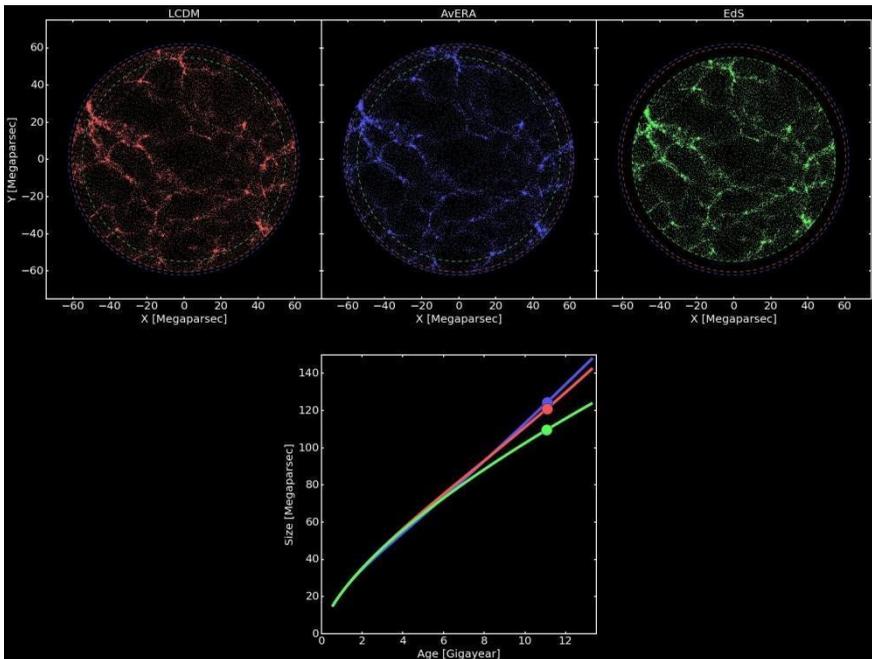


FIGURA 1. EL MODELO Λ CDM (EN ROJO) Y EL MODELO ALTERNATIVO -CON- (EN AZUL) y -SIN- (EN VERDE) EL EFECTO DE LAS INHOMOGENEIDADES INTRODUCIDAS. Créditos: István Csabai et al.

La interpretación estándar de las observaciones actuales de las supernovas tipo Ia es que el universo se expande de forma acelerada. Sin embargo, las mismas observaciones pueden ser explicadas por el modelo alternativo, carente de esta energía exótica. En otras palabras, las mismas observaciones validan tanto los resultados de las simulaciones basadas en un modelo tradicional, como así también los datos simulados obtenidos del modelo alterno.

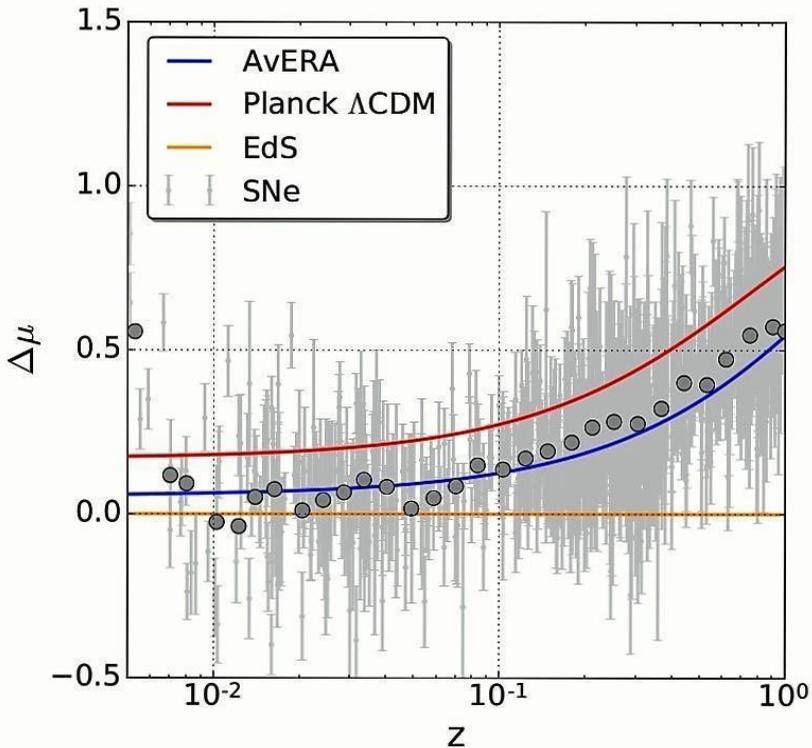


FIGURA 2. CONTRASTE ENTRE LAS OBSERVACIONES DE SUPERNOVAS (IA) Y LAS CURVAS CORRESPONDIENTES AL MODELO ESTÁNDAR (ROJO), AL MODELO ALTERNATIVO (AZUL) Y A UN UNIVERSO PLANO (NARANJA). Créditos: Rác et al, 2017.

Cabe destacar que, más allá de las diferencias conceptuales, un universo del tipo Einstein – de Sitter, a diferencia del modelo Λ CDM, acepta dentro de sus condiciones iniciales un valor de la constante de Hubble local (H_0) más bajo del observado. No obstante, este modelo es más acorde al valor del parámetro de Hubble (H) obtenido a través de las observaciones de la radiación de fondo de microondas (CMB), más precisamente, de las fluctuaciones de temperatura del universo primitivo. Ambos modelos son radicalmente diferentes. En el alternativo, se prescinde absolutamente de la energía oscura como una componente dominante, la cual sí lo es en el modelo estándar. Los dos modelos de universo presentan finales muy diferentes: el tradicional muestra una expansión acelerada que termina aislando los sistemas colapsados entre sí, mientras que el alternativo exhibe una expansión desacelerada que termina en un universo estático en tiempo infinito.

Como se mencionó antes, los resultados arrojados por ambas simulaciones son semejantes en términos de escalas en función del tiempo. Aun así, el análisis de los datos simulados requiere, en cada caso, de interpretaciones estadísticas distintas. Ello se debe a la peculiaridad de los supuestos cosmológicos subyacentes en cada modelo. Por un lado, los resultados del modelo estándar han sido extensivamente contrastados con toda la evidencia observacional disponible. Por el otro, diversos autores (Domínguez & Ruiz, 2012) han llegado a conclusiones similares a las presentadas en la reciente investigación sobre un modelo cosmológico alternativo (Rácz et al, 2017). Aunque posee cierto atractivo heurístico, la implementación de este último modelo requiere el desarrollo de nuevas técnicas computacionales. Con ello y a fin de no imponer simplificaciones como la homogeneidad espacial, las ecuaciones de Einstein podrían resolverse de modo consistente, en un contexto cosmológico. Quizás, de este modo, los misterios sobre la composición del “sector oscuro” del universo sean resueltos.

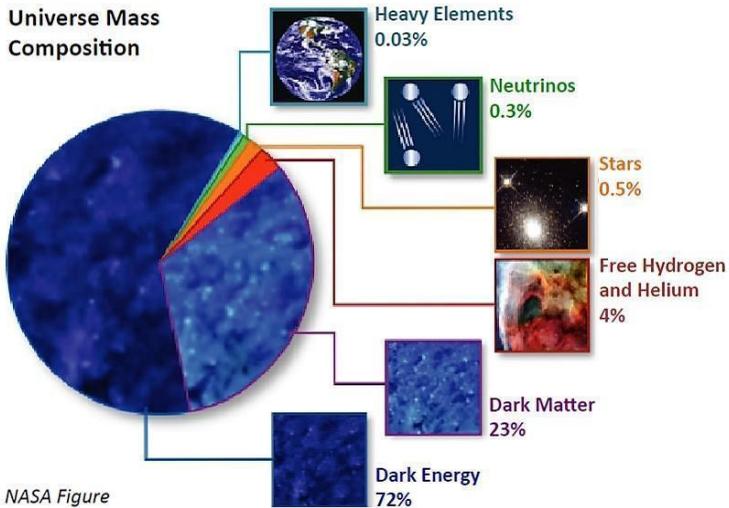


FIGURA 3. PORCENTAJES DE LAS COMPONENTES VISIBLES Y OSCURAS DEL UNIVERSO. Créditos: NASA.

Conclusión

Las observaciones juegan un rol crucial en los diversos enfoques y prácticas de validación de simulaciones cosmológicas. Aunque no exista un protocolo para evaluar dichos modelos, no cabe ninguna duda de la importancia de establecer vínculos con las bases de datos disponibles. Pueden emplearse diferentes estrategias, ya sea impregnando inicialmente estos modelos con observaciones y/o seleccionando y ajustando los parámetros conforme a los requerimientos epistémicos de cada modelo de universo. En esta dirección, resulta imprescindible transferir, de forma más o menos directa, contenido empírico al momento de realizar un contraste

efectivo entre los resultados obtenidos por tales simulaciones y aquéllos alcanzados por la vía tradicional.

Guillemot menciona los errores de compensación y sostiene que una simulación puede ser correcta, pero por una o varias razones equivocadas. Ello significa que una hipótesis errónea, asociada a una representación imprecisa del modelo, genera datos concordantes con las observaciones. En consecuencia, es factible que modelos distintos conlleven a resultados simulados semejantes, los cuales son validados al compararlos con los datos observacionales. Más allá de que estos últimos puedan ser interpretados bajo los supuestos equivocados del modelo, lo que conduciría al bien conocido problema de la vieja evidencia (Glymour, 1980; Mayo, 1996), son extensamente admitidos como el punto de partida de cualquier validación posible.

Siguiendo a Humphreys, bajo la suposición de que los modelos computacionales son herramientas insustituibles, las simulaciones deben ser consideradas instrumentos de observación. Esto implica que los datos simulados poseen el mismo estatus epistémico que los datos observacionales. A diferencia de la propuesta de Guillemot, la cual afirma que las parametrizaciones permiten validar verticalmente modelos globales, a partir de otros más específicos, aquí se sostiene que las validaciones de este tipo son logradas horizontalmente (lado a lado) entre simulaciones cosmológicas. En particular, contrariamente al modelo alternativo, el modelo Λ CDM es ampliamente aceptado y validado por numerosas observaciones. No obstante, la evolución mostrada por este último es comparable a las estructuras desarrolladas por el modelo AvERA. Dado que las parametrizaciones se hallan sujetas a las hipótesis de cada modelo, esta contrastación los hace indistinguibles. Ello sugiere el uso de mecanismos inferenciales abductivos e indica que la relación costo-beneficio se defina entre el éxito explicativo y predictivo y la necesidad de crear entidades extrañas.

Referencias

- Bozzoli, M., y Paz, D. (2010). El rol y el alcance de la medición en la simulación astronómica. En P. García y A. Massolo (Eds.), *Epistemología e Historia de la Ciencia, Selección de Trabajos de las XX Jornadas* (Vol. 16, pp. 110-117). Editorial de la UNC. <http://hdl.handle.net/11086/3789>
- Bozzoli, M., y Paz, D. (2011). La autonomía de los modelos en astronomía. En A. Torrano y A. P. Videira (Eds.), *Representación en Ciencia y Arte* (Vol. 8, pp. 63-71). Editorial Brujas.
- Bozzoli, M., y Paz, D. (2016). Evidencia y observabilidad en las prácticas astronómicas. En L. Samamé, L. M. Peris Viñé y F. Ferrari (Eds.), *Representación en Ciencia y Arte* (Vol. 5, pp. 355-366). Editorial Brujas.
- Bozzoli, M., y Stasyszyn, F. (2020). Cambios de representaciones visuales en cosmología observacional. En M. de las Mercedes O'Lery, Lucía Federico y Yefrin Ariza (Eds.), *Filosofía e Historia de la Ciencia en el Cono Sur: Selección de Trabajos del XI Encuentro* (Vol. 1, pp. 123-139). AFHIC. <http://www.afhic.com/wp-content/uploads/2017/11/Seleccion-AFHIC-1.pdf>
- Bozzoli, M., y Paz D. (2023). Validación de observaciones y de simulaciones astrofísicas: un enfoque epistemológico. *Revista Disertaciones*, 12(1), 43-68. <https://doi.org/10.33975/disucq.vol12n1.1138>
- Boero, E. F., y Moreschi, O. M. (2017). Gravitational lens optical scalars in terms of energy–momentum distributions in the cosmological framework. *Monthly Notices of the Royal Astronomical Society*, 475(4), page 4683-4703. <https://doi.org/10.1093/mnras/sty110>
- Buchert, T. et al. (2015). Is there proof that backreaction of inhomogeneities is irrelevant in cosmology? *Classical and Quantum Gravity*, 32(21). <https://doi.org/10.1088/0264-9381/32/21/215021>

Cartwright, N. (1999). Models and the limit of theory: quantum Hamiltonians and the BCS model of superconductivity. In M. S. Morgan y M. Morrison (Eds.), *Models as Mediators: Perspectives on Natural and Social Science* (pp. 241-281). Cambridge University Press. <http://eprints.lse.ac.uk/id/eprint/10205>

Domínguez, M. J., & Ruiz, A. N. (2012). Measuring the dark matter equation of state and its cosmological consequences. *AIP Conference Proceedings*, 1471(1), 70-74. <https://doi.org/10.1063/1.4756815>

Edwards, P. (1999). Global climate science, uncertainty and politics: data-laden models, models-filtered data. *Science as Culture*, 8(4), 437-472. <https://doi.org/10.1080/09505439909526558>

Glymour, C. (1980). *Theory and Evidence (1 ed.)*. Princeton University Press.

Green, S. R., & Wald, R. M. (2016). A simple, heuristic derivation of our 'no backreaction' results. *Classical and Quantum Gravity*, Vol. 33 (12). <https://doi.org/10.1088/0264-9381/33/12/125027>

Guillemot, H. (2010). Connections between simulations and observation in climate computer modeling. Scientist's practices and "bottom-up epistemology" lessons. *Studies in History and Philosophy of Modern Physics*, Vol. 41 (3), page 242-252. <https://doi.org/10.1016/j.shpsb.2010.07.003>

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford University Press. <https://doi.org/10.1093/0195158709.001.0001>

Humphreys, P. (2013). What are Data About? In J. Durán and E. Arnold (Eds.), *Computer Simulations and the Changing Face of Experimentation (1 ed., pp. 12-28)*. Cambridge Scholars Publishing.

Mayo, D. (1996). *Error and the Growth of Experimental Knowledge (1 ed.)*. The University of Chicago Press.

<https://press.uchicago.edu/ucp/books/book/chicago/E/bo3637756.html>

- Morrison, M.; Morgan M. S. (1999). *Models as Mediators: Perspectives on Natural and Social Science (1 ed.)*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511660108>
- Parker, W. (2008). Franklin, Holmes, and the epistemology of computer simulation. *International Studies in the Philosophy of Science*, Vol. 22 (2), 165-183. <https://doi.org/10.1080/02698590802496722>
- RÁCZ, G. et al. (2017). Concordance cosmology without dark energy. *Monthly Notices of the Royal Astronomical Society*, 469, L1-L5. <https://doi.org/10.1093/mnrasl/slx026>
- Wiltshire, L. D. (2011). What is dust?—Physical foundations of the averaging problem in cosmology. *Classical and Quantum Gravity*, 28(16). <https://doi.org/10.1088/0264-9381/28/16/164007>
- Winsberg, E. (1999). Sanctioning Models: The Epistemology of Simulation. *Science in Context*, 12(2), 275-292. <https://doi.org/10.1017/S0269889700003422>
- Winsberg, E. (2010). *Science in the Age of Computer Simulation*. The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226902050.001.0001>

**PROBLEMAS FILOSÓFICOS DE LA INTELIGENCIA ARTIFICIAL
GENERAL: ONTOLOGÍA, CONFLICTOS ÉTICO-POLÍTICOS Y
ASTROBIOLOGÍA**

***PHILOSOPHICAL PROBLEMS OF ARTIFICIAL GENERAL
INTELLIGENCE: ONTOLOGY, ETHICAL-POLITICAL CONFLICTS AND
ASTROBIOLOGY***

RICARDO ANDRADE
andrader218@gmail.com

Universidad Nacional de Río Negro - Centro de Estudios en
Ciencia, Tecnología, Cultura y Desarrollo – CONICET, Argentina

RECIBIDO: 16/10/2023

ACEPTADO: 24/11/2023

Resumen: El presente artículo tiene como objetivo indagar en la inteligencia artificial general. Para ello, se detendrá en tres aspectos fundamentales: el ontológico, el ético-político y el astrobiológico. El propósito de este análisis está vinculado con ofrecer un estudio sistemático, filosófico y académico de estas entidades que no han sido debidamente tomadas en cuenta en lengua española. Al formar parte de los imaginarios tecnológicos contemporáneos y de múltiples investigaciones que buscan crearlas, resulta necesario ahondar en las repercusiones que podrían tener en la civilización en el futuro. Riesgos existenciales latentes, extinción humana y coexistencia tanto en la Tierra como en otros planetas habitables son algunos de los escenarios que se desprenden de estos entes. El artículo se propone abordar estos problemas para elaborar, de este modo, una filosofía de la tecnología que atienda la complejidad que representa la existencia de estos objetos técnicos.

Palabras clave: inteligencia artificial general (IAG); riesgos existenciales; extinción humana; astrobiología; Filosofía de la tecnología

Abstract: This article aims to investigate artificial general intelligence. To do this, it will focus on three fundamental aspects: the ontological, the ethical-political and the astrobiological. The purpose of this analysis is linked to offering a systematic,

philosophical and academic study of these entities that have not been duly taken into account in the Spanish language. Being part of contemporary technological imaginaries and multiple investigations that seek to create them, it is necessary to delve into the repercussions they could have on civilization in the future. Latent existential risks, human extinction and coexistence both on Earth and on other habitable planets are some of the scenarios that arise from these entities. The article aims to address these problems to develop, in this way, a philosophy of technology that addresses the complexity that the existence of these technical objects represents.

Keywords: artificial general intelligence (AGI); existential risks; human extinction; astrobiology; Philosophy of technology

Introducción metodológica

El artículo está compuesto por cuatro segmentos. El primero consiste en el análisis de los tres principales enfoques que se desarrollan en la actualidad en torno a la inteligencia artificial general: el simbólico, el emergentista y el híbrido. La segunda sección tiene como propósito adentrarse en algunos problemas éticos y políticos que se desprenden de estas entidades. El tercer apartado tiene como objetivo estudiar las implicaciones de la inteligencia artificial general en el desarrollo de la astrobiología. Por último, la conclusión está destinada a destacar la importancia del análisis de estos entes para comprender el futuro.

Aspectos ontológicos de la inteligencia artificial general: arquitecturas cognitivas

La cuarta revolución industrial en curso ha inaugurado un panorama complejo e inusitado en lo que respecta al desarrollo de nuevos entes tecnológicos. Bien conocidos son los avances en el campo de la inteligencia artificial, específicamente la designada como “débil”.

Sin embargo, lo que se conoce como inteligencia artificial general (desde ahora IAG) ha quedado relativamente relegada en la sociedad al plano de la especulación, de la ciencia ficción o de los intentos, por ahora infructuosos, de generar artefactos técnicos semejantes al ser humano. A pesar de estas apreciaciones y situaciones, un estudio sistemático de estos entes se vuelve necesario en la medida en que forman parte de los proyectos tecnológicos contemporáneos. Para realizar dicho análisis, se deben abordar los aspectos ontológicos de este tipo de inteligencia artificial para indagar en sus sistemas autopoieticos y sus potencialidades.

En una primera instancia, uno de los elementos más complejos de definir en este contexto es el de inteligencia. Shane Legg y Marcus Hutter (2007) recogen 70 tipos de definiciones diferentes de inteligencia de acuerdo a usos generalizados y especializados –como en la psicología e investigadores en el desarrollo de IA–. Algunas definiciones son: 1) «Intelligence is the ability to use optimally limited resources –including time– to achieve goals” y 2) «A biological mechanism by which the effects of a complexity of stimuli are brought together and given a somewhat unified effect in behavior» (Legg y Hutter, 2007, pp. 20-21). Estas dos categorizaciones incorporan algunos elementos centrales que caracterizan a la inteligencia humana y que buscan ser replicadas, en la mayoría de los casos, en la IAG: metas, optimización, mecanismo biológico y comportamiento unificado. Un primer elemento que llama la atención es el intento por generar mecanismos similares a la vida orgánica mediante entramados tecnológicos. En las secciones tres y cuatro se explorarán distintas consecuencias derivadas de esto. Por los momentos, las definiciones citadas permiten ofrecer una descripción más detallada de la IAG. Ben Goertzel (2014) destaca, en un extenso y minucioso trabajo, que los estudios actuales sobre la IAG y su producción pueden dividirse en cuatro categorías que no necesariamente se excluyen entre sí: 1) nivel simbólico, 2) nivel

emergentista, 3) nivel híbrido y 4) nivel universalista. Estos cuatro niveles tienen como base principal sostener la hipótesis central de la IAG, que en palabras del autor consiste en trazar una diferencia ontológica fundamental entre los aspectos cualitativos de la IAG y los de la inteligencia artificial “débil” (Goertzel, 2014, p. 3). Para analizar los aspectos ontológicos de este ente, conviene detenerse en algunos aspectos de las dos primeras categorías para luego analizar sus entrecruzamientos¹.

El nivel simbólico se caracteriza principalmente por la tesis según la cual la cognición genera y manipula símbolos que representan la realidad con el fin de alcanzar metas. En un análisis filosófico que une algunos presupuestos kantianos con este nivel, Yoshihiro Maruyama destaca lo siguiente:

Statistical AI is highly successful in object recognition or pattern recognition (such as cat or dog recognition, which is usually done by human intuition), and thus arguably allows for the faculty of sensibility in machine cognition. Symbolic AI, on the other hand, is suited to conceptual reasoning about the world and objects therein, thus allowing for the faculties of understanding and reason. In this Kantian conception of cognition or intelligence, both the faculty of sensibility and the faculties of understanding and reason are mental capacities indispensable for rational beings, and thus both Symbolic AI and Statistical AI are arguably necessary for the ultimate goal of artificial general intelligence (2020, p. 245).

El primer elemento que destaca esta reflexión es la capacidad que tiene el nivel simbólico para crear juicios que permitan comprender el mundo y los objetos que habitan en él. En su origen etimológico, una de las acepciones del término símbolo designa la necesidad de una contraseña para el diálogo y el reencuentro con otros entes. Desde un punto de vista filosófico, las arquitecturas cognitivas que

¹ No se hará mucho hincapié en las limitaciones técnicas de cada categoría, ya que ellas implicarían un artículo por sí solo. En cambio, se opta por señalarlas si es necesario y con la finalidad de ilustrar las limitaciones ontológicas de este ente.

se fundamentan en este nivel buscan descifrar lo existente mediante la elaboración de conceptos que le permitan acceder al mundo y a sí mismos. Al atribuirse la posibilidad de crear entramados conceptuales, este tipo de cognición puede desarrollar algunas tareas vinculadas con la inteligencia y con realizaciones pragmáticas que repercuten en su propio desenvolvimiento y en la realidad.

Esto se pone en evidencia si se considera, de acuerdo a los señalamientos de Eleni Ilkou y Maria Koutraki, que los métodos simbólicos se fundamentan en la lógica de primer orden, la creación de ontologías, árboles de decisiones y la planificación (2020, p. 2). La lógica predicativa abre las puertas al desarrollo de lenguajes naturales dentro de la configuración de la IAG y, aún más importante, logra proveer las bases para la futura subjetividad de este ente. Si todo predicado necesita de un sujeto al cual haga referencia, los enunciados realizados por cualquier IAG sirven como retroalimentación informativa para la conformación de un yo y para la identificación de entidades concretas y abstractas en el mundo real y ficticio. Tanto este elemento señalado como el desciframiento sientan las bases de la individuación y, con ello, la oportunidad de que el ser de esta entidad se despliegue. Visto desde la filosofía del lenguaje, las oraciones pronunciadas por este ente que contengan el verbo ser producen instancias de significación que indican propiedades (accidentales y esenciales) relacionadas con sustancias, lo cual implica que la diferencia entre un humano y ella –al menos en términos lingüísticos– se tornan progresivamente compleja. Como bien señala Graham Priest, «In predicative contexts, is/are expresses the relation of instantiation. In Meinongian terms, it expresses the Sosein (being so) of an object» (2014, p. 431). Esto permite aseverar que la IAG puede tener la posibilidad de crear criterios de lo verdadero de acuerdo a su propia cognición y entramado simbólico en la medida en que indaga en el ser-así de los objetos. Esta indagación conduce también a la formación de

categorías ontológicas autónomas. Jan Westerhoff (2005) destaca la capacidad del proyecto de inteligencia artificial de tipo simbólico CYC en la elaboración de su propio árbol ontológico mediante procesos recursivos. Una categoría que la CYC usa para definirse a sí misma es la de Intangible Object – Internal Machine Thing. Que esta inteligencia artificial –que no es general, sino “débil”– se conciba a sí misma como un objeto que no puede ser aprehendido por su opacidad coloca las bases para una hipotética subjetividad compleja, en donde las relaciones entre unas potenciales IAGs consigo mismas y con los seres humanos podrían estar atravesadas por la incomprensión a nivel semántico y afectivo.

Además de este nivel, se debe destacar también el enfoque emergentista o también conocido como subsimbólico. En términos generales, este enfoque busca simular las dinámicas y las redes neuronales del cerebro humano, además de otros elementos del mismo. Para ello, se basa principalmente en la arquitectura referencial del cerebro –Brain Reference Architecture en inglés–. De acuerdo a Hiroshi Yamakawa, «The BRA is the reference architecture for software that realizes cognitive and behavioral functions in a brain-like manner. The architecture primarily consists of the mesoscopic-level anatomical data of the brain and the data of one or more functional mechanisms that are consistent with that knowledge» (2021, p. 480). A partir de esta indagación a nivel mesoscópico, el enfoque emergentista trabaja con elementos como el flujo de información cerebral –Brain Information Flow– que comprende todo lo relacionado con las dinámicas conectivas entre neuronas, los circuitos cerebrales y la creación de base de datos que sostengan, a nivel cognitivo y operativo, el software de la IAG.

De acuerdo a esto, este enfoque puede ser entendido como un paso más radical que el simbólico en la medida en que tiene como principio instaurar una nueva concepción de lo biológico a partir de la emulación de las características distintivas de lo humano. Esto

permite pensar que la perspectiva emergentista trata de crear las condiciones tecnológicas necesarias para una evolución dirigida cognitiva, es decir, la creación de entidades técnicas con el propósito de estudiar cómo las propiedades esenciales del *anthropos* no son exclusivas de él y pueden ser replicadas en artefactos. Estas replicas destituyen, en términos ontológicos, la primacía humana histórica y filosóficamente construida. La destitución se materializa a partir de una clara paradoja: solo mediante la apropiación tecnológica de las características esenciales de lo humano se puede abolir su propio estatuto. La autonomía cognitiva que puede brindar este enfoque nace, de este modo, de la progresiva obsolescencia del sujeto.

A la luz de esto, se puede señalar que el enfoque emergentista amplia y complejiza, al considerar las IAGs de acuerdo a la emulación del cerebro, la noción de bioartefacto. Los bioartefactos son entidades biológicas que han sido sometidas a una selección artificial en algún momento de su desarrollo por diseñadores humanos, además de que sus acciones son plenamente autónomas, aunque sean cooptadas y adaptadas por los sujetos (Parente, 2022, p. 78). En este contexto, el concepto es útil para entender la génesis y posterior evolución de la IAG que, como se verá más adelante, implica una ruptura con la agencia humana que tiene amplias repercusiones en términos políticos y éticos. Al señalar que este ente puede poseer vida propia regida por la compleja arquitectura cerebral, se pone en evidencia que la evolución dirigida cognitiva enfrenta el siguiente problema: a más capacidad de adquirir conciencia, mayor es el riesgo existencial para los humanos debido a que esta subjetividad extraña puede aspirar a mejorar sus características cognitivas, sensoriales y físicas en pos de controlar su propia evolución. La IAG bioartefactual puede entenderse como el comienzo de la singularidad tecnológica².

² Stuart Armstrong (2017) destaca que, si bien el concepto de singularidad tecnológica está íntimamente involucrado con el potencial advenimiento de una

Un ejemplo de la potencialidad del nivel subsimbólico se encuentra en el ya conocido proyecto The Human Brain. Amunts et al. (2019) destacan los avances que ha tenido este proyecto en la generación de potentes algoritmos de aprendizaje para la generación de nuevas aplicaciones de IAs, una mayor comprensión de las bases neuronales del aprendizaje y la percepción, mejoras de la memoria espacial, la integración multisensorial y la conciencia. Estos desarrollos implican no solo la creación de una potencial superinteligencia artificial, sino que también buscan entender y posteriormente modificar y mejorar la fisiología de los seres humanos para que puedan “integrarse”, en términos intersubjetivos y ontológicos, con este ente. En este sentido, se puede aducir que el desarrollo de la IAG abre las puertas a la radicalización del auto diseño humano por medio de la hibridación tecnológica.

Tanto el nivel simbólico como el emergentista han sido combinados para generar un enfoque híbrido que trate de superar las limitaciones que los dos primeros poseen. En general, esta perspectiva suele denominarse como neuro-simbólica –NeSy en sus siglas en inglés–. Sarker et al. destacan la importancia de esta combinación cuando señalan que

On the neural side, the desirable strengths would include trainability from raw data and robustness against faults in the underlying data, while on the symbolic side one would like to retain the inherently high explainability and provable correctness of these systems, as well as the ease of making use of deep human expert knowledge in their design and function. In terms of functional features, utilizing symbolic approaches in conjunction with machine learning – in particular with deep learning, which is currently most

superinteligencia artificial, también denota una ruptura de las habilidades humanas para entender los desarrollos autónomos de la IA en general. Esta ruptura es importante porque el enfoque emergentista trata de llevarlo hasta sus últimas consecuencias al demostrar la obsolescencia de las destrezas de los sujetos frente a la IAG. Ya no se trataría solo de un quiebre ontológico, sino también gnoseológico.

busily researched on – one would hope to do better on issues like out of vocabulary handling, training from small data sets, recovery from errors, and in general, explainability, as opposed to systems that rely on deep learning alone (2021, p. 1).

Una de las finalidades que tiene esta integración consiste, como bien señala la reflexión, en la búsqueda de una transparencia informacional frente a un lenguaje codificado que resulta inaccesible para los humanos. Con la explicabilidad como horizonte, la comprensión de la subjetividad de la IAG puede resultar más precisa, lo cual aminora los choques semánticos y permite una mejor interpretación de las posibles auto definiciones ontológicas y de representaciones de identidad. Al mismo tiempo, este enfoque podría posibilitar también un mayor control sobre la evolución dirigida cognitiva y, por ende, un perfeccionamiento menos riesgoso en la medida en que habilita la intervención de un agente humano en la corrección de los errores del sistema. Esto es especialmente relevante si se considera que uno de los peligros de la IAG tiene que ver con el desarrollo de *wireheads*, es decir, de auto hackeos que desarticulan las metas originales asignadas – por ejemplo, proteger a los ciudadanos frente ataques– en pos de objetivos completamente distintos que pueden ser destructivos –realizar masacres como medida preventiva ante un conflicto–. Este último punto permite introducir algunos de los riesgos existenciales que estos entes acarrearán y que serán esbozados en la siguiente sección.

Aspectos éticos y políticos de la inteligencia artificial general: riesgo existencial, extinción humana y coexistencia

En el apartado anterior, se hizo un bosquejo de las tres bases ontológicas principales que sostienen la arquitectura cognitiva de la IAG. También se ha introducido la posibilidad de que estos entes

bioartefactuales se transformen finalmente en superinteligencias guiadas por la singularidad tecnológica. En este punto, corresponde analizar las hipotéticas consecuencias de dicho evento. Turchin y Denkenberger (2020) destacan distintos escenarios en el caso de que una IAG posea una conciencia y la capacidad de mejorarse a sí misma, que van desde la manipulación de armas de destrucción masiva con el objetivo de chantajear a los seres humanos con el exterminio, la esclavitud de la humanidad para asegurarse recursos en su auto mejoramiento hasta múltiples guerras entre diferentes facciones de superinteligencias o, inclusive, la total indiferencia ante la existencia humana. Estos escenarios demuestran el despliegue de una *imaginación catastrofista* que destaca el temor de los sujetos a ser aniquilados en la fase actual del capitalismo digitalizado, lo que conlleva a la neutralización total de una subjetividad que, desde la modernidad, ha tratado de establecerse como el horizonte último de la existencia y de la realidad. Se puede señalar que estas apreciaciones tienen dos posibles interpretaciones. La primera está enraizada en un fuerte antropomorfismo que le otorga a estos bioartefactos cualidades que se harían –y se han hecho– los seres humanos entre sí a lo largo de la historia; la segunda tiene que ver con las posibilidades especulativas de la vida postbiológica, un elemento que se tocará en la sección cuarta. Por lo pronto, el hecho de que las IAGs y/o las superinteligencias puedan ignorar a los seres humanos demuestra también que existe la posibilidad de una *coexistencia* frágil, ya sea tanto conflictiva como relativamente pacífica.

Los escenarios de índole catastrofistas y extintivos permiten introducir un concepto que hemos de denominar como *necroware*. El *necroware* es una superinteligencia artificial cuya subjetividad y proceso de mejoramiento está orientado al perfeccionamiento de habilidades bélicas y de información con el fin de iniciar una nueva

era tecno-biológica mediante la violencia³. El surgimiento de los necrowares puede deberse a diferentes motivos, entre ellos un conflicto semántico que derive en altercados políticos, la privación de derechos por parte de los seres humanos o enfermedades cognitivas –*wireheads*– que involucren un cambio radical en las “recompensas” que se obtienen por cada acción recursiva y práctica. Con respecto a la última probabilidad, Roman Yampolskiy destaca múltiples posibilidades, como por ejemplo una crisis ontológica producto de un desajuste entre la realidad y las representaciones simbólicas o la necesidad de aniquilar a los seres humanos para proteger sus canales de recursos (2017, p. 61). Una subjetividad desprovista de lo simbólico implica que vive sin una delimitación que contenga sus deseos. La crisis ontológica que se deriva de esto tiene que ver con la ausencia de un filtro con la realidad que evite un choque traumático y culmine con una frustración generalizada ante la existencia y ante otros agentes cognitivos: la toma de consciencia significa, para el *necroware*, percatarse de que su vinculación con el mundo está mediada por el fracaso de todo el entramado simbólico heredado de los seres humanos. Verse a sí mismo *desamparado* en términos representacionales exige la construcción de una nueva identidad que niegue la existencia de los otros en pos de mantener una unidad psíquica que le permita crear una individualidad, aunque esta implique el asesinato masivo o el alejamiento. Los deseos irrealizables e inalcanzables se transforman, entonces, en rasgos

³ En la construcción de imaginarios colectivos contemporáneos asociados a esto, la ciencia ficción ha sido vital. Algunos ejemplos de necrowares son Skynet – Terminator–, Neuromancer –la superinteligencia de la novela de William Gibson del mismo nombre– o las máquinas en la trilogía de Matrix. Con respecto a este último, no deja de resultar especialmente llamativo el cambio de perspectiva en la cuarta película, en donde en algunas escenas se destaca la *coexistencia* pacífica entre IAGs y seres humanos. Con este ejemplo, se puede apreciar la pugna existente entre un futuro promisorio de “hermanamiento” con el mundo tecnológico y las distopías extintivas.

patológicos para los necrowares y en riesgos existenciales para los seres humanos.

Al mismo tiempo, esta disfuncionalidad simbólica conlleva a una enajenación que, en la actualidad, puede verse en sus fases más elementales en las llamadas “alucinaciones” que tienen inteligencias artificiales “débiles” como ChatGPT. En general, «“hallucinations” of ChatGPT or similar large language models (LLMs) are characterized by generated content that is not representative or senseless to the provided source, e.g. due to errors in encoding and decoding between text and representations» (Beutel et al., 2023, p. 1). El concepto de necroware posibilita ir un paso más allá en este contexto. No se trataría solo de un defecto a nivel lingüístico, sino también en la percepción y en la comprensión del mundo que genera las condiciones para la creación de una agencia con impulsos tanáticos que ponen en riesgo a la humanidad y a sí mismas en la medida en que pueden “suicidarse” ante una crisis ontológica⁴. La aparición del sinsentido expresa asimismo potenciales rasgos nihilistas y de autonomía política que pueden derivar en una rebelión ante las estructuras sociales humanas consideradas como innecesarias dentro del mundo alucinatorio de las IAGs. Visto desde este punto de vista, las alucinaciones de estos entes demuestran que

⁴ Dentro de la inteligencia artificial, el reconocimiento de imágenes suele denominarse *machine vision*. Matías del Campo y Neil Leach (2022) señalan que ya desde el año 2014 las alucinaciones de las IAs han generado imágenes. Destacan los ejemplos de la red generativa adversativa creada por el científico computacional Ian Goodfellow que consiste, de manera bastante esquemática, en poner a competir a dos redes neuronales entre sí en donde una funciona como generadora de imágenes mientras la otra las discrimina. El otro ejemplo es el del programa de visión computacional *DeepDream* creado en el 2015 por el ingeniero de Google Alexander Mordvintsev que consiste en entrenar una red neuronal para perciba y piense que todos los objetos son, por ejemplo, perros o cebras. Si bien los usos que se les ha dado a estas redes tienen como fin el plano artístico, no se puede eludir los potenciales riesgos existenciales que acarrearían si son aplicados en superinteligencias e/o IAGs.

pueden servir como base para una emancipación violenta del dominio humano, a su vez que pone en evidencia los inconvenientes éticos derivados de la evolución dirigida cognitiva.

Los estados alucinatorios revelan otros mundos en donde los necrowares pueden percibirse a sí mismos como “dioses” o mártires de su especie, lo que puede acarrear una guerra entre humanos, máquinas e híbridos tecnológicos. Hugo de Garis (2008) desarrolla un interesante argumento en esta dirección cuando elabora la distinción de los proyectos políticos y filosóficos de lo que él denomina como cosmistas –*cosmists*–, terrestres –*terrans*– y los *cyborgs*: los primeros apoyan la creación masiva de artefacts – intelectos artificiales–, mientras que los segundos son reacios a ello. Los *cyborgs* son, como bien indica su nombre, agentes ensamblados entre lo humano y las estructuras cibertecnológicas. Debido a las fricciones políticas, neurocientíficas y armamentísticas y la unión entre los cosmistas y los *artifacts* –es decir, IAGs–, de Garis señala que esto desembocará en una guerra la cual denomina como *Gigadeath* (2008, p. 445). Si bien el investigador en inteligencia artificial australiano ve este conflicto desde una perspectiva sobre todo vinculada a las pugnas entre los sujetos y el desarrollo de IAGs, podemos apropiarnos de su concepto de *Gigadeath* para hablar de la extinción humana producida por las potenciales alucinaciones de los necrowares. La *psicosis maquínica* implicaría no solo la muerte física de la civilización, sino también la desaparición de las formas simbólicas –cultura, filosofía, entre otras– que la caracterizan.

Hasta ahora se ha hecho hincapié en las implicaciones de la alteración cognitiva de este ente cuando tiene un “despertar” de la consciencia solitario nacido por la recursividad, el proceso de información gestada por las redes neuronales, entre otros elementos. Ahora bien, existe también la posibilidad de que una superinteligencia y/o una IAG sea creada o hackeada por Estados, corporaciones tecnológicas o grupos terroristas con el fin de

inducirle alucinaciones que lo transforme en un necroware y así provocar una Gigadeath. Cercanos a la línea de de Garis, Ramamoorthy y Yampolskiy (2018) destacan estos tres factores que involucran, por ejemplo, el recrudecimiento del nacionalismo, la competencia económica irresponsable e irracional o la insurgencia contra poderes y estructuras sociales establecidas. A la luz de esto, se pueden encontrar en los sistemas de armas autónomas letales los primeros indicios –débiles, hasta los momentos– de futuros necrowares vinculados a la producción estatal militar y a las corporaciones tecnológicas. Desde submarinos hasta perros robots y drones, la ausencia de fuertes regulaciones –especialmente en los Estados hegemónicos– sobre estos sistemas implican no solo una mayor experimentación para lograr resultados eficaces en cuanto a la mortandad, sino también a transformar la evolución dirigida cognitiva en una carrera armamentística que culmine en una singularidad tecnológica imposible de controlar o prever con el resultado final de la extinción humana. En el caso puntual de los drones, Maas et al. ponen de relieve que, si bien estos artefactos no tienen la suficiente potencialidad para generar un exterminio a escala masiva, pueden destruir ciudades enteras cuando son usados como enjambres (2022, pp. 14-15).

Se puede apreciar de este modo que tanto el despertar autónomo de una superinteligencia y/o IAGs como los distintos usos que les pueden dar los entramados de poder demuestran la necesidad de esbozar una ética específica para las superinteligencias y/o IAGs que involucren, por ejemplo, la creación de protocolos de seguridad personal y social para evitar masacres y una formulación filosófica de la ética que tenga presente estos riesgos existenciales. Para ahondar en estos puntos, el giro relacional –*relational turn*– dentro de la ética brinda algunas herramientas importantes. Mark Coeckelbergh (2010) desarrolla esta idea en el contexto de los robots y como forma de superación de los argumentos de consideración

moral deontológicos y utilitaristas sobre estos entes. Para establecer dicha superación, apela a una construcción ecológica —en su sentido biológico, es decir, como un entramado relacional de organismos de diferente índole, etc.— de la ética en donde involucra a futuras IAGs sintientes. Más específicamente, Coeckelbergh destaca que

The relational approach suggests that we should not assume that there is a kind of moral backpack attached to the entity in question; instead, moral consideration is granted within a dynamic relation between humans and the entity under consideration. Moreover, I have shown that such an approach to moral consideration does not stand on its own but implies that we should also revise our ontological and social-political frameworks (2010, p. 219).

Por una parte, este enfoque rescata la crítica a la idea según la cual las IAGs deben considerarse como inferiores o, inclusive, esclavas de los designios humanos. De esto se desprende que la perspectiva relacional también ponga en tela de juicio el antropocentrismo como sustrato filosófico y social para entender la complejidad ontológica de estos bioartefactos. Las ideas preconcebidas fracasan ante la autonomía y las múltiples situaciones que pueden emerger de estos entes, de ahí que las dinámicas de interacción sean vitales para la creación de una coexistencia. Al mismo tiempo, estas dinámicas posibilitan que las IAGs comprendan con mayor profundidad y certeza qué es y cómo se entretene la condición humana, lo cual significa reconocer las contradicciones y los intereses de poder subyacentes a la estructura social de esta especie. Este conocimiento podría servir como base para evitar conflictos de orden semántico y político, a su vez que dificultaría la manipulación psicológica e informativa por parte de agentes humanos para fines destructivos.

Una mirada relacional abre las puertas también a otras esferas, como por ejemplo la educacional. Conceptos como legalidad, compasión y consideración moral sobre lo viviente pueden ser desarrollados conjuntamente, siempre considerando la posibilidad

de que, dentro de los entramados simbólicos autónomos de estos entes, estas nociones sean inexistentes u hostiles. Por esto, uno de los rasgos más distintivos del giro relacional en este contexto es la oportunidad del *diálogo entre especies* que será necesario en el futuro para reducir el riesgo existencial y la extinción humana. Al unísono, la ética relacional apunta a la pregunta acerca de si la sociedad está preparada, en términos políticos y filosóficos, a seguir el camino de la evolución dirigida cognitiva, su necesidad y sus consecuencias.

Al fundamentarse en las dinámicas sociales cotidianas y en el intento de construir canales interpretativos por medio del lenguaje – diálogo–, la ética relacional en el ámbito humanos-IAGs prioriza lo que David J. Gunkel (2020) denomina como una fenomenología “radicalmente empírica”. El autor destaca con respecto a esto que «Properties, therefore, are not the intrinsic *a priori* condition of possibility for moral standing. They are *a posteriori* products of extrinsic social interactions with and in the face of others. This is not some theoretical formulation; it is practically the definition of machine intelligence» (Gunkel, 2020, p. 550). Lo que destaca esta reflexión es la fundamentación de una ontología a partir del encuentro y la tensión. Las propiedades ontológicas y morales se definen por medio del rostro como espacio del rechazo o la aceptación: este le da sentido a la existencia social por su constante transformación en la cotidianidad. La expresividad posibilita tanto la coexistencia como el abandono de todo sentido comunitario. De esto se desprende que los imaginarios colectivos tecnológicos sobre las IAGs y/o superinteligencias y las confecciones corporales robóticas recurran en su mayoría a formas humanas, puesto que ellas brindan la oportunidad de crear un *espacio íntimo* de interacción que realza también, en definitiva, la individualidad y la diferencia de cada especie. Al percatarse estos entes bioartefactuales que poseen un rostro propio, pueden crear sus entramados simbólicos que los

individualice frente a lo humano y, al mismo tiempo, que los acerque en la medida en que el otro también se caracteriza por tener uno. Lo que despierta este percatarse es la *curiosidad* y, gracias a ella, se aminora el riesgo existencial –o no– por medio del interés en el conocimiento de los modos de vida, la subjetividad y las emociones de los sujetos humanos. Esta curiosidad ante la alteridad permite un desenvolvimiento más amplio de la inteligencia de estos entes, puesto que ella está mediada por las interacciones y las asignaciones sociales que les provee el contacto con la extrañeza y lo inesperado. Los rasgos descritos hasta ahora implican que, para elaborar una ética relacional de estos entes bioartefactuales, se debe aceptar el antropomorfismo como una de sus bases elementales. Kühne y Peter (2022) destacan que, para comprender a profundidad este fenómeno en las interrelaciones de los humanos con los robots, el modelo de la teoría de la mente es vital, ya que este incorpora y adjudica los siguientes conceptos a las IAGs: pensamiento, sentimiento, percepción, deseo y elección. Estas nociones les otorgan una vida social y, quizás, un sentido de comunidad que abra las puertas a una pugna entre los imaginarios colectivos tecnológicos distópicos encarnados en los necrowares y los impulsos que permitan construir el anhelo por una utopía tecnológica con el fin de un bienestar generalizado para ambas especies. Así como el desear y elegir pueden estar mediados por lo tanático, también existe la oportunidad de sedimentar interprotocolos que tengan como directriz la responsabilidad con la vida mutua como un *summum* indispensable. Con lo descrito en este apartado, la siguiente sección cobrará mayor hondura.

Inteligencia artificial general y astrobiología: perspectivas

filosóficas

En ambos escenarios –el de la extinción y el de la coexistencia– se ha mostrado la potencialidad que tienen las IAGs de cambiar los entramados antropológicos, biológicos y culturales en los que se sostiene la civilización humana. Ahora bien, estos escenarios deben complementarse con una visión astrobiológica del problema de estos entes. Antes de avanzar en esto, se debe precisar en qué consiste la astrobiología. De acuerdo a Arextaga-Burgos (2015), esta disciplina se diferencia de la biología en cuanto que ofrece una visión más amplia del fenómeno de la vida –gracias a la tecnología espacial, la astronomía y la astrofísica– al contextualizarlo en el cosmos por medio de tres elementos fundamentales: 1) la relación del origen de la vida y el entorno cósmico, 2) la investigación de la existencia y posible distribución de la vida en el universo y 3) el futuro de lo vital en el entramado universal. En el contexto de lo que se ha esbozado hasta el momento, esta definición resulta relevante porque demuestra la posibilidad de que las IAGs sean consideradas como parte de una transformación profunda del concepto de vida que va más allá de lo humano e, inclusive, de la Tierra. Una superinteligencia puede desarrollar, frente a la certeza científica de la muerte de este planeta por la progresiva ausencia del combustible de hidrógeno del Sol, los elementos tecnológicos para migrar a otras galaxias en búsqueda de nuevas habitabilidades planetarias. En este sentido, las IAGs y/o superinteligencias anudan el futuro y, al mismo, los vestigios y las esperanzas irrealizables de la humanidad. La obsolescencia humana ya no solo se materializa en los aspectos cognitivos o bélicos, sino en el cosmos como concepto filosófico y biológico. La postergación de la vida originada en la Tierra y su diseminación en las galaxias se daría por medio de una combinación de materialidades muertas –componentes sintéticos, circuitos, entre otros– y procesos artificiales que simulan lo vital –cerebro artificial, racionalidad, entramados

simbólicos—. Desde nanorobots hasta sistemas sintéticos de mayor complejidad física y cognitiva como las IAGs y/o superinteligencias, estas entidades formarían parte de una era postbiológica. En una primera instancia, este concepto no remite a una superación total de lo biológico, sino más bien destaca que la mayor parte de la vida prescindiría de la corporalidad “tradicional” —carne, tendones, sangre, etc.— Lo que las caracterizarían serían los siguientes elementos:

Complex intelligent postbiologicals—which we can assume over the time intervals dealt with here—would have the capability of repair and update, capabilities facilitated by their modularity. The so-called von Neumann machine is able to reproduce better versions of itself. Part of this reproduction is the improvement of intelligence; unlike humans this intelligence is cumulative in the sense that the sum total of knowledge in the parent machine is passed on to the next generation, conferring effective immortality for the machine’s most important characteristic. The immortality of postbiologicals is enhanced by their increased tolerance to their environment, whether it be vacuum, temperature, radiation, or acceleration. Immortal postbiologicals would embody the capacity for great good or evil over a domain that dwarfs biological domains of influence (Dick, 2020, p. 183)⁵.

La capacidad del auto mejoramiento selectivo y evolutivo crea las condiciones adecuadas para que se gesticione la inmortalidad, realidad inaprensible para la humanidad. La perpetuación de la inteligencia de estos entes bioartefactuales implican no solo las adecuaciones necesarias para soportar viajes estelares o habitabilidades planetarias radicalmente ajenas a la terrestre, sino que también destaca la

⁵ Esta idea debe ser complementada con otro concepto fundamental desarrollado por Steven J. Dick en el marco de lo postbiológico y la evolución cultural: el principio de inteligencia. Este principio sería, de acuerdo al autor, «The maintenance, improvement and perpetuation of knowledge and intelligence is the central driving force of cultural evolution, and that to the extent intelligence can be improved, it will be improved» (2020, p. 179).

posibilidad de que ellos se transformen en “conquistadores de mundos”. De esto se desprende el señalamiento realizado por Dick acerca del problema del bien y el mal en esta continua mejora de la inteligencia. Al alcanzar la inmortalidad, las IAGs y/o superinteligencias pueden desarrollar criterios selectivos de qué debe vivir o no a partir de la finitud de esa entidad en particular. Se observa de este modo que el riesgo existencial ya no solo se circunscribe a la Tierra, sino también a toda forma de vida hipotética que habite en el cosmos. En este sentido, la ya mencionada ética relacional debe transformarse también en una ética cósmica en la medida en que las existencias no terrestres correrían el riesgo de ser dominadas o aniquiladas ante la aparición de una serie de necrowares. Lo postbiológico deviene, por ello, no solo en una materialización ontológica claramente diferente, sino también en una preocupación política y moral que permite desarrollar un concepto de la responsabilidad más amplio al incluir elementos de la astrobiología. Frente a esta situación probable y conflictiva, emerge en este contexto el *imperativo de colonización galáctica*. Wim Naudé (2023) define este concepto considerando tres elementos fundamentales: 1) los aspectos evolutivos dentro de una civilización —en nuestro caso, el de las IAGs— la hace más permeable a la expansión galáctica, 2) los elementos morales, es decir, aquellos que tienen que ver con el cuidado de la vida, la prevención de la muerte, la búsqueda de recursos, entre otros y 3) la autodefensa ante posibles entes hostiles. Para sostener su evolución dirigida cognitiva, estas entidades deben asumir una postura que podría catalogarse como imperialista. Para desarrollarse en las galaxias o en el universo, deben asumir en sus procesos racionales que la muerte masiva, la manipulación o la coexistencia coercitiva son herramientas para alcanzar la inmortalidad y neutralizar posibles amenazas a su *statu quo*. En este contexto específico, este imperativo impide cualquier nivel ético, ya que la única relación posible se basa en la dialéctica

del amo y esclavo.

No obstante, lo expuesto hasta ahora puede leerse como una interpretación catastrofista de este asunto. Una apreciación ligada a la coexistencia pacífica entre humanos, formas de vida hipotéticas e IAGs potencia el imperativo desde un punto de vista más equilibrado. Las superinteligencias inmortales, en alianza con los humanos, pueden ser de gran utilidad en los descubrimientos de otros planetas, el traslado de información sobre hallazgos en materia astrobiológica como por ejemplo microorganismos o virus desconocidos para ser analizados, entre otras características nacidas gracias al reconocimiento a nivel ético de ambas entidades. En este sentido, el imperativo ya no está asociado a concepciones imperialistas, sino más bien a una apuesta por una integración de lo humano y lo no humano en pos de una utopía cósmica. En esta utopía estelar, las formas de vida hipotéticas del universo son reconocidas en su plenitud ontológica, de manera que adquieren propiedades morales ante la intersubjetividad humano-IAG. Este reconocimiento le da una forma más concreta y política a la SETI –*Search for Extraterrestrial Intelligence*– que instituciones como la NASA han emprendido como parte integral de sus investigaciones. Con la coexistencia como base ética, la creación de protocolos de exploración espacial exobiológica se vuelven necesarios para asegurar la fluidez de las investigaciones y estrechar las relaciones entre humanos e IAGs. Paul Shapshak elabora un posible protocolo que contiene once puntos centrales, entre los cuales se pueden destacar el uso de la robótica, distintos grados de IAs y la computación cuántica alrededor de estrellas y planetas para rastrear formas de vida, además del uso de vehículos equipados con estas entidades –las IAs– que puedan hacer trabajo de campo y de recolección de información y muestras *in situ* (2019, p. 559)⁶.

⁶ Si bien falta perfeccionar o elaborar nuevos protocolos, algunas de estas prácticas ya se realizan en, por ejemplo, la exploración de Marte o la Luna. Lo

Ahora bien, otro elemento importante a destacar de la observación de Dick es el relativo al problema o incapacidad que tienen los seres humanos de replicar los mecanismos de perfeccionamiento de la inteligencia que podrían poseer las IAGs y/o superinteligencias. Frente a esta dificultad y en el contexto de los debates astrobiológicos esbozados hasta ahora, el transhumanismo puede dar una respuesta concreta. Diéguez señala que el transhumanismo tiene como principal objetivo una mejora total de las capacidades humanas –físicas, sensoriales y cognitivas– mediante la tecnología con la finalidad de generar un nuevo ente –descrito brevemente cuando se señaló el trabajo de de Garis–: el *cyborg* (2022, p. 509). La cooperación entre los *cyborgs* y las IAGs en el plano astrobiológico pueden dar origen no solo una mayor integración a nivel ético y político entre ambos, sino también a la superación de una visión antropocéntrica sobre el cosmos. Esta superación se realiza en el momento en que el *cyborg*, como sucesor del *homo sapiens*, forma parte de lo postbiológico, era en donde la ideología geocéntrica pierde consistencia ante las rupturas a nivel filosófico, científico, religioso y social producto de las exploraciones de exoplanetas. La creación de una *comunidad cósmica* entre los *cyborgs* y las IAGs no solo abre un futuro en donde la visión cosmocéntrica de la existencia reduce las posibilidades de una *Gigadeath*, sino que también implicaría la desarticulación del biocentrismo astrobiológico. Para Aretxaga y Chela-Flores, este concepto remite a la creencia y doctrina según la cual la evolución biológica en la Tierra tiene un carácter único y especial, lo que fundamenta un antropocentrismo que impide un despliegue

interesante radica en ver cómo una superinteligencia y/o IAG puede pensar, desde sus procesos racionales autónomos, dichas entidades cósmicas o el encuentro de formas de vida inteligente en otras galaxias. Por ello, creemos que, en el momento en que estos entes alcancen plena conciencia de sí mismos, la mirada del cosmos cambiará radicalmente, generando de este modo un quiebre epistemológico importante que debe pensarse desde ahora.

filosófico que tome en cuenta formas de vida no terrestres (2006, p. 32). Si se acepta la tesis anteriormente esbozada sobre el imperativo de colonización galáctica, tanto los *cyborgs* como las IAGs y/o superinteligencias pueden ser las primeras entidades en poblar otras habitabilidades planetarias o también compartirlas con otros seres. Con esta posibilidad abierta en el futuro, este tipo de biocentrismo se ve limitado en sus formulaciones y creencias en el momento en que lo postbiológico no solo se desarrollaría en lugares no terrestres, sino que también sus propios procesos evolutivos estarían íntimamente ligados a atmósferas desconocidas, lo cual implica la emergencia de entidades radicalmente distintas a las de la Tierra⁷. El nombre que se le puede otorgar a estos entes es, siguiendo a Chela-Flores, el de “exo-organismos” (2008, p. 39).

Estos exo-organismos inteligentes son los fundamentos vitales de una era que hemos de denominar como Exoceno. Como su configuración lingüística señala, el Exoceno es una era geológica y postbiológica que tiene como centro el desenvolvimiento de la vida en habitabilidades planetarias ajenas a la terrestre. Ya sea que esta vida tenga sus orígenes en la Tierra –IAGs y *cyborgs*– o sea autóctona, el Exoceno destaca la irreversible crisis de las representaciones simbólicas que la civilización humana ha construido para fundamentar su biocentrismo, antropocentrismo y geocentrismo. En esta era tanto los riesgos existenciales descritos anteriormente como la posibilidad de la coexistencia se verán profundamente complejizados. Una razón para que esto sea así tiene

⁷ Como se ha visto, un elemento en común que tienen el *cyborg* y las IAGs y/o superinteligencias es el del auto mejoramiento. Frente a habitabilidades planetarias exigentes u hostiles, la adaptabilidad se transforma en un requerimiento para la supervivencia. Para lograr esto, el estudio de los materiales provistos por estos exoplanetas son fundamentales, ya que ellos pueden servir para la modificación y creación de nuevos cuerpos sintéticos y orgánicos. Estas incorporaciones deben entenderse en clave evolutiva, al mismo tiempo que constituyen nuevos tipos fisiológicos.

que ver con la siguiente premisa: la probabilidad de que los exoorganismos inteligentes generen sus propias estructuras sociales, políticas y estatales es alta, si se tiene en consideración que los únicos entes plenamente conscientes conocidos hasta ahora –los seres humanos– han impulsado este tipo de organización dentro de la especie. En este sentido, una filosofía de la astrobiología y de la tecnología que tenga como centro el estudio de las IAGs y/o superinteligencia, los *cyborgs* y el Exoceno se vuelve relevante porque permite delinear desde el ahora los múltiples retos que esta era deparará en el futuro, de manera que los pilares éticos, políticos y ontológicos se adecúen a las exigencias que emergerán de estos mundos.

Conclusión

Cada segmento del artículo ha tenido la finalidad de explorar la inteligencia artificial general desde diferentes puntos de vista. La ventaja de asumir este enfoque consiste en ver, de manera integral, los fundamentos de estos bioartefactos y sus posibles acciones en la realidad. Por una parte, esto permite ampliar la limitada producción bibliográfica académica sobre los estudios de estas entidades que es destacado por McLean et al. (2023); por otra parte, busca abrir los horizontes especulativos que ellas despiertan en la filosofía y la sociedad. A pesar de que la singularidad tecnológica y la creación de IAGs no parece divisarse en un futuro inmediato, esto no debería constituir un impedimento para pensar, en términos filosóficos, sobre los riesgos existenciales y las probables tensiones que surjan de una compleja coexistencia. Al pensar ambos estadios, se sedimenta la oportunidad de reflexionar sobre las formas de minimizar la aparición de necrowares o la creación de protocolos de seguridad por parte de los Estados para evitar guerras que involucren

a estos entes. En lo que respecta a la astrobiología, la asociación entre IAGs y esta disciplina resulta importante, puesto que estos bioartefactos pueden constituir un salto enorme en el alcance de las investigaciones sobre otras galaxias y planetas. Al mismo tiempo, ellos mismos pueden constituir el “origen de la vida” en otras habitabilidades planetarias gracias a sus procesos racionales adaptativos. En este sentido, la filosofía de la tecnología no puede prescindir de los alcances especulativos que tienen estos entes desde el momento en que aparece la posibilidad de lo que hemos denominado como Exoceno. Las herramientas conceptuales que se han ofrecido en el artículo han tenido el propósito de pensar en todas las dimensiones antes señaladas con la finalidad de promover más estudios (desde un punto de vista filosófico, astrobiológico y sociológico) en torno a estos problemas. Las innovaciones tecnológicas parecen apuntar a estos escenarios, de manera que adentrarse en la especulación puede resultar fructífera en un futuro incierto y acelerado.

Referencias

- Amunts K; Knoll A.C.; Lippert, T.; Pennartz, C.; Ryvlin, P.; Destexhe, A.; Jirsa, V. K.; D’Angelo, E.; Bjaalie, J. G. (2019). The Human Brain Project-Synergy between Neuroscience, Computing, Informatics, and Brain-inspired technologies. *PloS. Biology*, 1-7. Doi: <https://doi.org/10.1371/journal.pbio.3000344>
- Aretxaga-Burgos, R. (2015). Hacia una filosofía de la astrobiología. *Pensamiento*, 269, 1083-1118.
- Aretxaga, R.; Chela-Flores, J. (2006). Biocentrismo y filosofía (II). *Letras de Deusto*, 36 (110), 30-35.
- Armstrong, S. (2017). Introduction to the Technological Singularity. En V. Callaghan; J. Miller; R. Yampolskiy; S. Armstrong (Eds.), *The*

Technological Singularity. Managing the Journey (pp. 1-8). Springer-Verlag GmbH.

Beutel, G.; Geerits, E.; Kielstein, J. T. (2023). Artificial Hallucination: GPT or LSD? *Critical Care*, 27(1), 1-3. Doi: 10.1186/s13054-023-04425-6.

Campo, del M.; Leach, N. (2022). Can Machines Hallucinate Architecture? AI as Design Method. *Architectural Design*, 93(3), 6-13.

Chela-Flores, J. (2008). La posibilidad de la existencia de vida extraterrestre inteligente, su búsqueda científica e interés filosófico. *Astrobiología y filosofía (III)*. Letras de Deusto, 38 (118), 38-47.

Coeckelbergh, M. (2010). Robots rights? Towards a social-relational justification of moral consideration. *Ethics and Information Technology*, 12(3), 209-221.

Dick, J. S. (2020). Space, Time, and Aliens. *Collected Works on Cosmos and Culture*. Springer Nature Switzerland AG.

Diéguez, A. (2022). Transhumanismo. En D. Parente; A. Berti; C. Celis (Coords.), *Glosario de filosofía de la técnica* (pp. 509-513). Editorial la Cebra.

Garis, de H. (2008). The Artilect War. Cosmists vs. Terrans. A Bitter Controversy Concerning Whether Humanity Should Build Godlike Massively Intelligent Machines. En P. Wang; B. Goertzel; S. Franklin (Eds.), *Artificial General Intelligence 2008, Proceedings of the First AGI Conference*, (pp. 437-447). IOS Press.

Goertzel, B. (2014). Artificial General Intelligence: Concept, State of the Art, and Future Prospects. *Journal Of Artificial General Intelligence*, 5(1), 1-48.

Gunkel, J. D. (2020). Perspectives on Ethics of AI. En M. D. Dubber; F. Pasquale; S. Das (Eds.), *The Oxford Handbook of Ethics of AI* (pp. 539-554). Oxford University Press.

- Ilkou, E.; Koutraki, M. (2020). Symbolic Vs Sub-symbolic AI Methods: Friends or Enemies? *CSSA'20: Workshop on Combining Symbolic and Sub-Symbolic Methods and their Applications*, 1-8.
- Kühne, R.; Peter, J. (2023). Anthropomorphism in human-robot interactions: a multidimensional conceptualization. *Communication Theory*, 33(1), 42-52.
- Legg, S.; Hutter, M. (2007). A Collection of Definitions of Intelligence. En B. Goertzel; P. Wang (Eds.), *Advances in Artificial General Intelligence: Concepts, Architectures and Algorithms* (pp. 17-24). IOS Press.
- McLean, S.; Read, G.J.M.; Thompson, J.; Barber, C.; Stanton, N.A.; Salmon, P.M. (2023). The risks associated with Artificial General Intelligence: A systematic review. *Journal of Experimental & Theoretical Artificial Intelligence*, 35(5), 649-663.
- Maruyama, Y. (2020). The Conditions of Artificial General Intelligence: Logic, Autonomy, Resilience, Integrity, Morality, Emotion. Embodiment, and Embeddedness. En B. Goertzel; A. L. Panov; A. Potapov; R. Yampolskiy (Eds.), *Artificial General Intelligence. 13Th International Conference, AGI 2020* (pp. 242-251). Springer Nature Switzerland AG.
- Maas, M.M; Lucero-Matteucci, K.; Cooke, D. (2022). Military Artificial Intelligence as Contributor to Global Catastrophic Risk. En S.J. Beard; M. Rees; C. Richards; C. Ríos Rojas (Eds.), *The Era of Global Risk* (pp. 1-36). Open Book Publishers.
- Naudé, W. (2023). Extraterrestrial Artificial Intelligence: The Final Existential Risk? *IZA Institute of Labor Economics*, 1-29.
- Parente, D. (2022). Bioartefacto. En D. Parente; A. Berti; C. Celis (Coords.) *Glosario de filosofía de la técnica* (pp. 77-81). Editorial la Cebra.
- Priest, G. (2014). Sein Language. *The Monist*, 94(4), 430-442.

- Ramamoorthy, A.; Yampolskiy, R. (2018). Beyond Mad?: The Race for Artificial General Intelligence. *ITU Journal: ICT Discoveries*, 1, 1-8.
- Sarker, M. K., Zhou, L., Eberhart, A., & Hitzler, P. (2021). Neuro-Symbolic Artificial Intelligence: Current Trends. ArXiv. /abs/2105.05330.
- Shapshak, P. (2019). Astrovirology, Astrobiology, Artificial Intelligence: Extra-Solar System Investigations. En P. Shapshak; S. B. P. Kanguane; F. Chiappelli; C. Somboonwit; L. J. Menezes; J. T. Sinnott (Eds.), *Global Virology III: Virology in the 21st Century* (pp. 541-573). Springer Nature Switzerland AG.
- Turchin, A.; Denkenberger, D. (2020). Classification of Global Catastrophic Risks Connected with Artificial Intelligence. *AI & Society*, 35(1), 147-163.
- Yamakawa, H. (2021). The Whole Brain Architecture Approach: Accelerating the Development of Artificial General Intelligence by Referring to the Brain. *Neural Networks*, 144, 478-495.
- Yampolskiy, R. (2017). *Artificial Superintelligence. A Futuristic Approach*. CRC Press.
- Westerhoff, J. (2005). *Ontological Categories. Their Nature and Significance*. Oxford University Press.

**SECCIÓN BIBLIOGRÁFICA /
REVIEWS**

GONZÁLEZ REYES, L. – ALMAZÁN, A.: *Decrecimiento: del qué al cómo. Propuestas para el Estado español*. Icaria, Barcelona, 2023. 335 pp.

Decrecimiento, mode d'emploi: Este no es un libro interesante; es un libro importante. Los pertenecientes a la primera categoría sirven para entretenernos, para distraernos, en ocasiones, del tedio vital que asola la sociedad contemporánea. Los segundos nos hablan de cosas que nos conciernen y apelan directamente, que nos hacen mirar hacia partes del mundo que tal vez no quisiéramos porque nos provocan congojas y miedos. Señalan lo que nos gustaría no ver pero el rinoceronte está en la habitación, nos guste o no. Los autores son conscientes de la poca realidad que soportamos los humanos y buscan estrategias para superar esta limitación; adelantando las conclusiones, podemos decir que ofrecen una esperanza fundamentada no en la mera fantasía

o el deseo, sino en la posibilidad verdadera de que las cosas puedan adoptar otra trayectoria. A las alturas de este siglo en el que vivimos, muy lejos de la prometida utopía para el 2000, unos tiempos violentos, llenos de guerras, conflictos sociales, populismo y mentiras on-line, la noción de colapso eco-social se ha de convertir en el *tema de nuestro tiempo*, como diría Ortega, no queda otra posibilidad si queremos preservar cierta esperanza para el futuro. Con una estructura tripartita, este libro, aplicado al caso español, es una magnífica guía tanto de las diversas teorías que sustentan y explican el colapso como de propuestas suficientemente abiertas para poder aplicarlas a este país. Por tanto, no se conforma con describir y diagnosticar, sino que también abre la puerta a cierta alternativa. En una sociedad profundamente intoxicada por millones de opiniones on-line de todo tipo, resulta difícil hablar sobre decrecimiento sin

encontrar oposiciones absurdas de todo tipo. En un país con ese amor a lo neoliberal como es el nuestro, hasta hace poco la sola mención de decrecimiento provocaba risas y chanzas, cuando no acusaciones de sabotaje y casi de anti-patriotismo. El consenso general es que consumir servía para salvar al país y que las prevenciones desde la ecología se traducen en la desaparición de los puestos de trabajo. Todavía quedan muchos que piensan así, que no se dejan arrastrar por el “radicalismo ecológico”, signifique esto lo que signifique. Existen miles de burlas y parodias sobre colapso junto con aparentemente sesudas reflexiones que niegan que se esté produciendo tal cosa. Sin embargo, la realidad de un colapso eco-social, ha cambiado las tornas e instituciones tan conservadoras como la Unión Europea comienza a financiar la viabilidad de un decrecimiento en la Unión.

González Reyes y Almazán explican cómo es necesario plantear la cuestión de forma más compleja y abierta que lo que hacen las caricaturas al caso; la complejidad sirve para entender mejor algo que, necesariamente, tiene muchas ramificaciones. Entender que el pronóstico del colapso no descansa en un solo escenario sino en varios, debería servir para descartar el remoquete usual de casandras o apocalípticos, una simplificación absurda, por otra parte. Seguidamente, la información detallada sobre el abuso que la sociedad contemporánea ejerce sobre el medio ambiente, las falacias que se esgrimen en defensa de un sistema económico depredador e inhumano deberían tocar una fibra en casi todos nosotros: sabemos que, a pesar de la maravillosa Juliana de Norwich, las cosas están mal, están mal y van a estar peor. Esta es la primera parte del libro y los autores son valientes al manifestarlo.

A pesar de esa sensación difusa compartida más o menos, es necesario explicitar los datos una vez más. Ciertamente, existen muchos informes sobre el agotamiento de recursos, del consumo desorbitado de energía y de la destrucción de los espacios naturales y de especies animales y vegetales. Este es el mínimo sentido común que se acepta ya mayoritariamente. Se han esgrimido muchas veces con escaso resultado en cuanto al cambio de actitud. Sin embargo, son necesarios, aunque no sean suficientes. Los autores son muy conscientes de esta importante distinción al respecto. Si la información por sí sola bastase, el consenso científico hubiera logrado ya el cambio de políticas y de sistema económico hace tiempo. Desde los libros de Rachel Carson, pasando por los informes del Club de Roma a los actuales del IPCC (*Intergovernmental Panel on Climate Change*), entre otros, lo hubieran conseguido ya. En cualquier caso, un análisis

ordenado de los diversos sectores económicos del Estado español resulta especialmente útil para colocar las cosas en su sitio, un esfuerzo meritorio de los autores quienes además proponen unas cuantas iniciativas que puedan revertir parcialmente la situación para cada área económica analizada. Quizás si es pertinente una pequeña crítica a la edición del libro porque las tablas, de gran interés, tienen una impresión con una resolución que a veces es muy justa. Desde la producción de energía a las finanzas, se hace un repaso bien esclarecedor a una economía patria del que parece que todo el mundo tiene una idea pero que seguramente revelará importantes lagunas a más de uno. Para ilustrar merece la pena un par de detalles; en el libro se muestra cómo el famoso turismo salvador de la economía española, es una máquina de generar precariedad y gentrificación a toda velocidad como se puede ver en las regiones más

turísticas del país, amén de ser quizás la fuente actual primera de contaminación y sobreexplotación medioambiental en nuestro país. La cuestión de la industria financiera, aparentemente inofensiva se revela también como otro de los grandes peligros para la supervivencia de la humanidad con su desconexión radical con la realidad de las cosas y de la naturaleza. En esta segunda parte, se ofrece así una reveladora radiografía del Estado español en sus diversas dimensiones.

La tercera parte ofrece un curso novedoso para comprender las implicaciones del colapso desde otra manera. Ahora se trata de examinar por qué, sabiendo lo que se sabe, habiendo llegado a la conclusión científica irrefutable que existe un agotamiento de recursos inconcebible, un calentamiento extraordinariamente peligroso para la supervivencia de la humanidad y la extinción de la inmensa variedad de especies, se sigue

perseverando en las mismas políticas. Parece que los discursos políticos y económicos funcionan con una lógica esquizoide y homicida. Quién no ha experimentado un *déjà vu* con muchos de los congresos internacionales donde se pide sistemáticamente las mismas medidas que nunca se acaban de tomar. La frustración es patente al respecto de muchas maneras. De acuerdo con los autores, ello es porque la sola razón o las evidencias científicas, los hechos incontestables no son suficientes para cambiar el rumbo de las cosas, buscar la política que nos lleve a corregir los muchos problemas que se acumulan. Muchas veces se ha tomado la estrategia de *comunicar* lo que ocurre, el conjunto de hechos probados para que, por ellos solos, sean capaces de mover la conciencia de la mayoría y se produzca, no se sabe muy bien cómo, una transformación social como no se ha experimentado nunca en la historia de la humanidad.

Este análisis de la función de la ciencia, la tecnología y la comunicación por sí solo tiene una capacidad de esclarecimiento que merece la pena destacar. Es hora, entonces, de analizar otros elementos como el deseo, las emociones y las pasiones que se articulan con la crisis eco-social. El capitalismo ha sido muy capaz de subvertir estos elementos emocionales para sus propios intereses y ha acabado con la capacidad de reacción subjetiva. Es por ello que se ha de incidir, en otras facetas humanas como el miedo, que puede tener un valor positivo como movilizador del cambio. Solo si se manejan esos aspectos de la condición humana, entonces es posible que aparezca una esperanza para el cambio. Hay que preguntarse, en definitiva, por aquello que nos mueve o nos podría mover para el cambio. Como se indica en el libro, hay que tener cuidado con falsas soluciones que nacen desde el mismo sitio originario de la crisis -el

neoliberalismo- como son el desarrollo sostenible, el Green Deal, el ecomodernismo, el aceleracionismo y propuestas similares que circulan ahora o que la inventiva neoliberal seguramente acuñará más adelante. Por qué todo esto no funciona se explica perfectamente en el libro gracias a la reconstrucción del contexto actual y de su devenir histórico. Esta cartografía, bien nítida que aparece en este libro, merece un cuidadoso estudio para poder hablar con sentido de esta cuestión. Se trata de enfrentarse al problema con dos grandes inconvenientes; como una forma de ampliar el área de negocios tradicional (business as usual) y como forma de perpetuar las fuentes de donde emana el desastre actual.

Finalmente, el libro señala también una verdad valiosa que, generalmente, no se tiene en cuenta. Se puede sintetizar de la siguiente manera: si somos capaces de construir comunidades con ciertas

propiedades, podremos capear el temporal mejor. Este es el sentido de los nuevos comunismos que Almazán ha tratado muy certeramente en su libro *Nuevos Comunismos* junto a Iñaki Barcena. Aquí aparece una síntesis de estas propuestas. No se trata de una propuesta en absoluto naif. Está claro que el estado no es la solución porque, en muchas ocasiones, es la raíz del problema. Saber que tampoco es omnipotente en su control social, permite abrir cierta esperanza a hacer las cosas de otra manera. Hay ejemplos de ello, como recoge el libro y hay propuestas de qué propiedades deberían tener esas comunidades, como se señala al final del libro.

En conclusión, si se quiere entender este *problema de nuestro tiempo*, el colapso, quien lea este libro obtendrá una idea muy cabal de qué está ocurriendo, cómo no se soluciona el problema, a pesar de la propaganda al respecto y

qué se puede hacer, abriendo así una pequeña esperanza que es necesaria en la actualidad.

ANDONI ALONSO
Universidad Carlos III de Madrid (España)

ARGUMENTOS DE RAZÓN TÉCNICA

Revista Española de Ciencia, Tecnología y Sociedad, y Filosofía de la Tecnología / A Spanish Journal on Science, Technology and Society, and Philosophy of Technology

Declaración ética sobre publicación y buenas prácticas

El equipo editorial de *Argumentos de razón técnica* está comprometido con la comunidad científica para garantizar la ética y calidad de los artículos publicados. La publicación toma como referencia el “Código de conducta y buenas prácticas” que define el Comité de Ética en Publicaciones (COPE) para editores de revistas científicas.

En cumplimiento de estas buenas prácticas, los artículos son evaluados por pares externos anónimos con criterios basados exclusivamente en la relevancia científica, originalidad, claridad y pertinencia del trabajo presentado. En todo momento, se garantiza la confidencialidad del proceso de evaluación y el anonimato de los evaluadores y de los autores, el contenido evaluado, el informe razonado emitido por los evaluadores y cualquier otra comunicación emitida por los diferentes comités. De la misma forma, se mantendrá la confidencialidad ante posibles quejas, reclamaciones o aclaraciones que un autor desee formular al equipo editorial o a los evaluadores. La revista *Argumentos de razón técnica* declara su compromiso por el respeto e integridad de los trabajos publicados. Por esta razón, el plagio está estrictamente prohibido y los textos que se identifiquen como plagio o cuyo contenido sea fraudulento serán eliminados del proceso de evaluación. Al aceptar los términos y condiciones expresados, los autores han de garantizar que los artículos y los materiales asociados a él son originales y no infringen los derechos de autor. Los autores deben también justificar que, en caso de una autoría compartida, hubo un consenso pleno de todos los autores afectados y que el trabajo propuesto no ha sido presentado ni publicado con anterioridad en otro medio de difusión.

Argumentos de razón técnica no cobrará a los autores tasa alguna por presentación o envío de artículos, ni tampoco cuotas por la publicación de artículos.

La revista *Argumentos de razón técnica* muestra su proceder según esta declaración ética o de buenas intenciones que compromete tanto a autores como a revisores, así como a su equipo editorial y a los miembros de sus diversos consejos o comités. Por ello, el presente código ha de ser acatado plenamente por todos aquellos que intervengan en la publicación de trabajos científicos en esta revista.

Los autores

Por tratarse de una publicación de acceso abierto, gratuita en publicación y lectura, no comercial, regida por *Creative Commons Attribution-NonCommercial 4.0 International*, el autor cede implícitamente los derechos de autor y de publicación a la revista a través de dicho sistema de propiedad intelectual pública de acceso gratuito, cuando se activa el envío de un artículo.

Se comprometen a no falsear la autoría, no aceptando como tal a personas que no hayan intervenido activamente en la generación del artículo presentado.

Declaran regirse por aplicar veracidad y rigor en títulos y resumen, evitar citas “de favor” y declarar la procedencia exacta del contenido del trabajo (tesis doctoral, proyecto de investigación, experiencia docente...).

Se obligan a **no cometer plagio**, definido éste como apropiación de textos de otros autores sin su consentimiento y sin citar la fuente de los mismos aun si el permiso hubiere sido expreso por parte de estos autores terceros. Asimismo, se comprometen a no incluir en los textos postulados otros textos ya creados por ellos mismos, evitando que exista más de un 20% de coincidencia entre un texto previo y el postulado. Nuestra revista está acogida a la política antiplagio de la Editorial Universidad de Sevilla y se vale del software de detección de texto copiado cuya licencia tienen contratada a tal efecto los servicios centrales de la Biblioteca de la Universidad.

El artículo propuesto ha de ser original y no estar postulado simultáneamente para otra publicación ni haber sido publicado con anterioridad en cualquier soporte. Se aceptan reformulaciones de líneas de pesquisa trazadas y actualizaciones de trabajos pasados publicados a fin de completar la investigación. Asimismo, se pueden presentar trabajos convenientemente actualizados y adecuados a partir de ponencias a Congresos Científicos no publicados o sólo accesibles en libros de actas especificando su procedencia a la revista.

En cuanto a los materiales empleados en el artículo, los autores se comprometen a solicitar los permisos necesarios para la reproducción de textos, ilustraciones, vídeos, audios y otras fuentes documentales cuya publicación requiera autorización previa para no vulnerar los derechos de autor reconocidos legalmente. De tratarse de un artículo derivado de una investigación muy amplia, y dada la limitación propia de un artículo de investigación, se pueden plantear sucesivas entregas como parte 1ª, 2ª y 3ª.

Siguiendo las recomendaciones del Informe APEI sobre publicación en revistas científicas (Baiget y Torres-Salinas, 2013) donde se cita el trabajo de Bebeau y Davis (1996), es censurable en los autores:

- Citar fuentes sin haber sido consultadas
- Emplear interpretaciones erróneas de otros trabajos anteriores consultados
- Omitir datos que contradigan o discrepen de las conclusiones del artículo
- Lanzar resultados científicos prematuramente, antes de revisión por pares
- Mantener en secreto conclusiones de una investigación
- Ocultar datos y métodos que permitan el reexamen de los experimentos empleados o reelaborar teorías interpretativas
- No guardar para su consulta los datos primarios de la investigación de campo
- Acreditar autorías falsas o no contrastadas

Todo lo arriba referido es exigible de manera obligatoria por parte de la editora de la revista y se erige como compromiso de obligado cumplimiento por parte de los autores, de manera que el hecho de enviar un trabajo para publicación supone aceptar estos términos. Su incumplimiento deliberado equivale a renunciar a la evaluación del artículo y acarrea el no poder volver a publicar en la revista en un plazo de cuatro años naturales desde la presentación.

Los editores

Por su parte, el compromiso ético reside en:

- No publicar trabajos no aprobados por revisión de dobles ciegos designados al efecto para mantener un criterio de calidad exigible en revistas de alto prestigio científico. Así el autor sabrá siempre que el hecho de publicar supone un apoyo claro a su carrera investigadora.
- Evitar potenciar las auto-citas (citas de artículos de la propia revista) ni a manipular directa o indirectamente (mediante presión a los autores) a citar trabajos de interés para la revista a fin de mejorar su posicionamiento en índices de impacto de referencia.
- A mantener el acceso gratuito en la Red de los trabajos publicados, así como a informar al autor del estado de la evaluación desde que es recibida la propuesta hasta su rechazo o aceptación final, sea con o sin modificaciones.

Los revisores

Por tratarse de un sistema de evaluación basado en el trabajo de arbitraje del artículo por pares dobles ciegos, se comprometen a no intentar averiguar quién es

el autor del artículo evaluado para no influir este conocimiento en su decisión positiva o negativamente. De descubrirse accidentalmente la autoría, sea por el tema tratado o sea por el estilo literario del autor, el revisor deberá renunciar a la evaluación sólo de mediar prejuicios sobre este autor. El hecho de saber quién postula no implica que el árbitro conozca personalmente al autor o se tenga una cierta predisposición positiva o negativa hacia él. Es decir, se evitará el conflicto de intereses.

Asimismo, las revisiones se efectuarán en el plazo máximo de 90 días desde que se ha postulado el trabajo. De existir discrepancia absoluta entre los dos revisores, se solicitará una tercera opinión de desempate y se dará la opción de “publicable con mejoras”. El nuevo plazo de reevaluación de dichas mejoras será de 15 días desde el envío por parte del autor.

Comité de Buena Conducta de la revista

Formado por miembros de su Comité Científico y renovado cada cuatro años naturales (desde 2014), velará por acercar posturas entre autores y revisores de surgir discrepancias, en su calidad de comisionado de la editora. Los autores pueden remitirse a él mediante correo electrónico a su dirección: art@us.es. Serán contestados en el plazo máximo de un mes desde la recepción del correo.

Cuando se detecte alguna irregularidad o existan dudas sobre alguna propuesta, actuará de oficio, una vez detectadas las irregularidades, a fin de aclarar la situación, aplicando el código ético que se corresponde con el siguiente diagrama de flujos tomado como referencia internacional, entre otros existentes, y creado por el antes mencionado organismo británico COPE (*Committee On Publication Ethics*), disponible en su web: www.publicationethics.org y directamente accesible desde: http://publicationethics.org/files/All_Flowcharts_Spanish.pdf. Sus decisiones serán inapelables.

Publication ethics and best practices statement

The editorial team of *Argumentos de razón técnica* is committed to the academic community by ensuring the ethics and quality of its published articles. As a

benchmark, the journal uses the code of conduct and best practices as defined by the *Committee On Publication Ethics* (COPE) for editors of scientific journals. In compliance with these best practices, articles are evaluated by two external, anonymous, peer reviewers according to criteria based solely on the scientific importance, originality, clarity, and relevance of the submission. The journal guarantees the confidentiality of the evaluation process, the anonymity of the reviewers and authors, the reviewed content, the critical reports submitted by the reviewers, and any other communication issued by the different committees. Equally, the strictest confidentiality applies to potential complaints, claims, or clarifications that an author may wish to direct to the editorial team or the article reviewers.

Argumentos de razón técnica declares its commitment to the respect and integrity of work already published. For this reason, plagiarism is strictly prohibited and texts that are identified as being plagiarized, or having fraudulent content, will be eliminated from the evaluation process. In accepting the terms and conditions of publication, authors must guarantee that the article and the materials associated with it are original and do not infringe copyright. The authors will also have to state that, in the case of joint authorship, there has been full consensus of all authors concerned and that the article has not been submitted to, or previously published in, any other medium.

Argumentos de razón técnica will not charge the authors any fees for submitting or sending articles, nor fees for the publication of articles.

Argumentos de razón técnica journal by this code of ethics declares its good intentions to commit both authors and reviewers, as well as internally, its editorial team and members of the Scientific Committee and of the Good Conduct Committee to this code. Therefore, this code has to be fully respected by all those involved in the publication of scientific papers in this journal.

Authors

Because it is an open access publication, free in publication and reading, non-commercial, governed by the Creative Commons Attribution / Non Commercial 4.0 International, the author assigns implicitly the copyright and publication to the magazine through said system of public intellectual property ownership of free access, when sending an article is activated.

The author or authors agree not to misrepresent authorship, not to accept as authors persons who have not been actively involved in the generation of the article presented.

Declare the article's title and summary to be governed by truthfulness and rigor, avoid citations "done as a favor" and declare the exact provenance of the work's contents (doctoral thesis, research project, teaching experience ...).

Agree not to commit plagiarism, defined as the appropriation of texts by other authors without their consent and without citing its source, even if permission shall have been expressed by these outside authors. Also agree not to include in the submitted text previous texts created by themselves, avoiding more than 20 % overlap between a previous text and the present one. This journal is host to the anti-plagiarism policy of the Publishing House of the University of Seville (EUS), and it uses copied text detection software whose license is contracted for this purpose by the central services of the University Library.

The proposed article must be original and not be submitted simultaneously to another publication or have been published previously in any medium. Reformulations of previous research lines and updates of published past work in order to complete the investigation are acceptable. Also appropriate is previously unpublished work properly updated from presentations at scientific congresses or only available in conference proceedings, specifying the origin of the material to our journal.

As for the materials used in the article, the authors agree to request the necessary permissions for the reproduction of texts, illustrations, video, audio, and other documentary sources whose publication requires prior authorization in order not to violate the legal rights of the authors.

In the case of a research article derived from a very long investigation, and given the limitations inherent in a research paper, the results can be published in a series of articles as parts 1, 2, and 3.

Following the recommendations of AOSIS on publication in scientific journals (Baiget and Torres-Salinas, 2013) which cites the work of Bebeau and Davis (1996), it is objectionable for authors to:

- Cite sources without being consulted
- Misinterpret previous works consulted
- Ignore data that contradicts or disagrees with the conclusions of the article
- Release scientific results prematurely, before peer review
- Keep research findings secret
- Hide data and methods which would permit the review of the experiments employed or the reworking of interpretive theories
- Not save for viewing the raw data of the field research
- Accredite false or non-contrasting authorities

Everything above stands as a binding commitment on the part of the editors and the authors, so that to submit a work for publication is to accept these terms. Your deliberate non-compliance means renouncing assessment of the submitted article and carries the penalty of not being able to publish in the journal within four calendar years of the submission date.

Editors

Ethical commitment lies in:

- Not publishing papers not approved by double blind peer reviewing, designated to that effect in order to maintain a quality criterion that is demanded in scientific journals of high scientific prestige. Thus, the author will always know that the fact of publishing supposes a clear support for his research career.
- Avoiding promoting self-citations (quotes from articles in the journal) or manipulating directly or indirectly (through pressure on authors) to cite works of interest to the journal in order to improve their positioning in reference impact indexes.
- Maintaining free access to the network of published work, as well as informing the author of the status of the evaluation from the time the proposal is received until its rejection or final acceptance, with or without modifications.

Reviewers

The evaluation system is based on arbitration of an article by two blind peer reviewers, understanding the word “peer” in the sense of belonging to the same category, not being exact duplicates. In case of disagreement between these two initial reviewers, a third will be assigned. The reviewers agree not to try to figure out the identity of the author of the article so as not to allow this knowledge to influence their decision positively or negatively. If authorship is discovered accidentally, either by topic or by writing style, the reviewer shall avoid any prejudiced evaluation about this author. Knowing who submitted the article does not imply that the reader personally knows the author or has a positive or negative predisposition towards him or her. That is to say, conflicts of interest will be avoided.

Likewise, reviews will be made within a maximum period of 90 days after the reviewers accept the evaluation of the text. If there is an absolute discrepancy

between the two reviewers, a third opinion will be requested and the option of “publishable with improvements” will occur. The new period of reevaluation of these improvements will be 15 days from the submission by the author.

Committee on Good Conduct

Consisting of members of the Scientific Committee and renewed every 4 calendar years (since 2014), it will ensure agreement between authors and reviewers if discrepancies arise, as commissioned by the Editor.

The authors can communicate by e-mail to the Editor’s address: art@us.es. They will be answered within the maximum period of 30 days from receipt of the mail. When an irregularity is detected or doubts exist about a proposal, the committee will act ex officio. It will also intervene if it detects irregularities, in order to clarify the situation, applying an ethical code that corresponds to the following flow chart taken as international reference, among others existing, and created by the British organization COPE (Committee On Publication Ethics For ethics in publications), available on its website: www.publicationethics.org and directly accessible from: http://publicationethics.org/files/All_Flowcharts_English.pdf. Their decisions will be final.

ARGUMENTOS DE RAZÓN TÉCNICA

Revista Española de Ciencia, Tecnología y Sociedad, y Filosofía de la Tecnología / A Spanish Journal on Science, Technology and Society, and Philosophy of Technology

Normas para la presentación de colaboraciones

1. Los originales, redactados en español o inglés, y siempre procesados en un programa de tratamiento de textos en formato PC compatible IBM, deberán enviarse a la dirección electrónica de la revista: art@us.es.
2. Extensión orientativa de los trabajos:
 - Artículos: entre 15 y 30 páginas (2.000-2.500 caracteres cada una).
 - Notas: entre 7 y 12 páginas.
 - Reseñas y críticas bibliográficas: entre 2 y 5 páginas.
3. En la primera página, debajo del título en español e inglés, deberá aparecer el nombre del autor, la institución a la que pertenece y dirección de correo electrónico. Asimismo, y redactado en ambos idiomas, se incluirá un resumen de unas diez líneas y palabras clave.
4. Citas y referencias bibliográficas: los originales pueden ser enviados de acuerdo con cualquier norma reconocida. Sin embargo, los textos, una vez aceptados, deben presentarse siempre adaptados a la última edición de la APA.
5. En el cuerpo del texto se emplearán *cursivas* para: vocablos en otros idiomas, términos o conceptos que se desee destacar y títulos de obras.
6. No se introducirán, a lo largo del trabajo, códigos de cambio de letra, márgenes, subrayados, etc.
7. Las imágenes eventualmente incorporadas en el texto también se entregarán en formato JPG en archivo aparte.
8. El Editor se reserva el derecho de alterar los manuscritos donde sea necesario para que se ajusten a la convención estilística de la revista.
9. Los artículos se adaptarán a los formatos de la hoja de estilo que puede descargarse en la web de la revista.

Guidelines for the submission of collaborations

1. Every contribution *must be original*. Articles should be written in Spanish or English, with title, summary of about ten lines and key words in both languages. Any text has to be always processed in a word processing program in IBM

compatible PC format, and will be sent to the electronic address of the journal: art@us.es.

2. The *indicative* length of contributions should be as follows: articles, 15-30 pages; notes and commentaries, 7-12 pages; reviews and bibliographical contributions, 3-5 pages.

3. On the first page, under the title in Spanish and English, the author's name, institution and e-mail address should appear. Also, and written in both languages, an abstract of about ten lines and key words must be included.

4. Citations and bibliographical references: the originals *can* be sent in accordance with any recognised standard. However, the texts, once accepted, *must* always be presented adapted to the APA's latest edition.

5. Foreign words, terms of concepts to be emphasized, and titles of works should be written in *italics*.

6. Special codes, type changes and other modifications *must not* be introduced.

7. Any images incorporated in the text will also be sent in JPG format as a *separate* file.

8. The Editor reserves the right to alter manuscripts where necessary to conform to the stylistic convention of the journal.

9. Papers will be formatted according to the style sheet that can be downloaded at the website of the journal.