THE JANUSIAN FACE OF FACIAL RECOGNITION, PART 1: ITS SUBFACE, INTERFACE, AND SURFACE[1]

Devon Schiller
University of Vienna
devonschiller@gmail.com

Abstract

Facial recognition has a *Janusian* face. In the instant of its interaction, facial recognition brings many faces together into relation. It is not only visually representational but also computationally *re*-presentational. To make facial recognition *knowable*, therefore, one needs first to make its many faces *visible*. One must expose, so to say, the functional relationalities between rhizomatic facialities within these commercial products and their proprietary computing. Only then can the ways facial recognition technology systems work, and how they are either adversely misused or beneficially used, be substantively challenged or tactically critiqued. Toward this end, from the critical standpoint of a computational semiotics both Peircean and pragmatist, I apply a method of *semiotic deblackboxing*.

[1] This research was supported by a DOC Fellowship from the Austrian Academy of Sciences.

Digital Age in Semiotics & Communication, Vol. VIII, 2025, 103–125

https://doi.org/10.33919/dasc.25.8.6

I divide my inquiry between two distinct investigations: here in part one, I explicate this theoretical approach; then in part two, I explore its practical application. Across these two parts, I probe how, and in what ways, the artificial intelligence of facial recognition not only comprises a computational, mechanical, and technological system, but also constitutes a semiotic system. The question is this: Does the artificial intelligence system of a facial recognition technology have *semiotic agency*? That is, is it therefore able to process the action of a sign that has *genuine* triadic sign relations rather than *non-genuine* dyadic quasi-sign relations? In other words, is artificial intelligence actually either *artificial* or *intelligent*? As I argue here, this *Janusian or multiple relation* is the intelligent result, not of the semiotic machine, its efficient causation, and allopoiesis, but of the semiotic animal, our final causation, and autopoiesis.

Keywords: biometric identification, computational semiotics, facial recognition technology (FRT), probabilistic artificial intelligence (ProbAI), semiotic deblackboxing.

> "Yet what god shall I say you are, double-formed Janus? […] Reveal, at the same time, the reason why you alone of the heavenly ones see both what is behind and what is in front. […] Then holy Janus, amazing in his double-appearance, suddenly presented his pair of faces to my eyes. I was terrified…" – Ovid ([8 CE] 2011: I.89–96)

1. Introduction

Facial recognition has a *Janusian* face. At least analogically speaking, the probabilistic artificial intelligence (ProbAI) of facial recognition technology (FRT) has not one but two faces. One face looks backward into the past and its *input*. And one gazes forward into the future and its *output*. Both so-called "faces" principally serve as a spatiotemporal passage to yet another face. The Janusian face of facial recognition is, therefore, not only visually representational but also computationally *re*-presentational. In an instant of its interaction, a facial recognition technology system brings far more than one or two faces together into relationalities. There are the human faces of its designers, users, and targets. And there are the machine faces of its subface, interface, and surface. There are the images of faces used to train its artificial intelligence, computer vision, and machine learning. And there are the faces of individuals that it detects, extracts, and classifies. There are the frontal views of dynamic faces detected from the face, its muscles, and movements. And there are the thin slices of static faces extracted from the continuous flow of this behavioral stream. There are the faces of the individuals who are being recognized. And there are the faces of the community who are doing the recognizing. For this reason, to make facial recognition *knowable*, one needs first to make its many faces *visible*. One must expose, so to say, the functional relationalities between rhizomatic facialities within these commercial products and their proprietary computing. Only then can the ways facial recognition technology systems work, and how they are either adversely misused or beneficially used, be subversively challenged or tactically critiqued.

From the critical standpoint of a computational semiotics, I apply a method of *deblackboxing*. Here I explicate how, and to what extent, the artificial intelligence of facial recognition is made up from not only a computational, mechanical, or technological system but also a semiotic system. In a black box, as manifestly illustrated by Martin Irvine, "semiotic principles" have been implemented within the design yet remain "'blackboxed,'" or "hidden from view," and "closed off" from the understanding of its users (2022: 205–206). Of course, across computational semiotics alone, several traditions of semiotic thought potentially afford for critical perspectives into facial recognition and how this black box then works (see, for review, Meunier 2022). These include, for instance, more linguistic or logical methods as well as more dyadic or triadic models. Toward a critical deblackboxing, however, I advance a Peircean and pragmatist approach (see, for example, Andersen, Hasle, and Brandt 1997; Nöth 2002; Nake 2008; Sørensen, Thellefsen, and Thellefsen 2020; Irvine 2022), in which I apply the triadic sign model of Charles Sanders Peirce (1839–1914), one of, if not *the* most, foundational thinkers behind pragmatism and semiotics. The Peircean model of triadic signs is constituted by a "*representamen*," "*object*," and "*interpretant*" (CP 2.228). Consequently, in Peirce's classification, "the science of semiotic[s] [also] has three branches" (CP 2.229). This "trivium," which Peirce refers to variously across his writings after he introduced it with "On a New List of Categories" in 1867, is constituted by "logic," "grammar," and "rhetoric" (CP 1.559). Based upon these branches, like Irvine's "'deblackboxing' method" (2022: 205), Bent Sørensen, Torkild Thellefsen, and Martin Thellefsen propose a program for studying the semiotics of what they call "technological artefacts" by way of their "technological grammar, technological logic, and technological rhetoric" (2020: 254). In these papers, I focus on the grammar. Peirce characterizes "Speculative Grammar" as "the general theory of the nature and meanings of signs, whether they be icons, indices, or symbols" (CP 1.191). This grammar concerns the ground conditions whereby the representamen of a sign both carries information and conveys meaning, such as by its efficient and final causation, immediate and dynamic object, as well as quasi and genuine semiosis, and likewise by its signal and sign interrelation, spacetime instantiation, as well as tokenization and retokenization. From the critical standpoint of a computational semiotics both Peircean and pragmatist, I do not seek to reveal the computational software or mechanical hardware obfuscated inside the opaque box. Rather, I deblackbox relationalities between the many signs from which the box takes its meaning.

The question is this: Does the artificial intelligence system of a facial recognition technology have *semiotic agency*? That is, can facial recognition create *in and of itself* "[t]he action of a sign," which, in Peirce's own terms, is not only "dyadic" but also "triadic" and not only "dynamical" but also "intelligent" (CP 5.472)? Can facial recognition create "a *genuine* triadic relation" rather than a non-genuine "collocation of dyadic relations," or, as Peirce himself phrases it, "a 'degenerate' form of triadic relationship" (CP 6.322)? Inside the black box, is artificial intelligence actually either *artificial* or *intelligent*? With facial recognition technology during human-computer interaction, the human act of sign interpretation is brought into relation with the computer operation of signal determination. But the human interactor and the facial recognition are not equals in this relation. As I argue here, *this Janusian or multiple relation* is the intelligent result, not of the semiotic machine, its efficient causation, and allopoiesis, but of the semiotic animal, our final causation, and autopoiesis.

2. The Ethics of Visibility: From Closed/Opaque, to Open/Transparent, and Everything In-between

Facial recognition is a black box in most, but not all, cases at this moment in our history, when the Fourth Industrial Revolution and the Information Age have accelerated forward such technological advances. Developed since the 1960s, facial recognition became commercially available for applied configuration by non-expert users in large-scale systems in the mid to late 2000s. As historicized by Kelly Gates, a theorist in science and technology studies, "military needs," as for military perception technologies from a military-industrial complex, have "influenced the development" of facial recognition technology ever "since its inception as a research program" in the mid twentieth century (2011: 101). Amidst the intelligence community practices of the post-9/11 cultural imaginary, the industrialist advocates for artificial intelligence began their work behind the scenes of government to accelerate facial recognition technology from theoretical in-the-lab experiments to practical in-the-wild applications. Researched and developed for the government and its security, Gates argues, facial recognition "more closely resembl[es] […] a 'black box,'" that is, a functional technology positioned in society so as to be "virtually indispensable" for the people and their safety (2011: 101). The sales pitch went something like this: where humanity failed, technology would succeed; by identifying a face, threats would be individualized. At least, that was the claim. Though alleging "technical neutrality," Gates accounts, "the technology promised to […] protec[t] civilization" from a "generalized and racialized 'face of terror,'" all while this "'facialization' of terrorism," and "nonstate forms of political violence," actually fell back upon "an implicit classifying logic [of] deviant facial types" (2011: 101). Therefore, the principles behind the design, as Gates herself attests, require not only "legal frameworks of privacy rights" but also ethical guidelines for "structural inequality" (2011: 199). And generally, Gates advocates, "the design of these technologies," and the black boxes within them, "needs to be more widely understood at every level of their development" (2011: 199), as I attempt here.

Given the facial ethos of the twenty-first century, with its Orwellian phantasms, and its panoptical overwatch, omnipresent eye-in-the-sky facial recognition is often even imputed with a mythical identity or a religious quality. For instance, Massimo Leone, a cultural and religious semiotician, was principal investigator on Face Aesthetics in Contemporary E-Technological Societies (FACETS), a Consolidator Grant Project supported by the European Research Council, from 2019 to 2024. With popularized cases and religionized terms, Leone goes so far as to claim that the "'black box'" is "designed to mystify […] digital entities," such as "algorithmic images" with their "augur signs," by imbuing the "genesis" of the box with an "auratic character" and "halo of mystery" (2024: 426–429). Although Leone himself addresses neither the expert technical literature nor the extensive theoretical literature about such black boxes, he correctly recognizes their "dual functionality" (2024: 427), both in technical practices, and in rhetorical strategies. The meaning behind this metaphor, whether *the black box of artificial intelligence* generally or *the black box of facial recognition* specifically, is perhaps understood best using the cognitive metaphor theory or conceptual metaphor theory, as it has been variously called, developed in cognitive science and linguistic studies beginning in the 1980s (see, for example, Lakoff and Johnson 1980). In terms of an ecological, embedded, embodied, enactive, and extended cognition, or what is known as 5E cognition, the source domain of the *black box* is mapped in the mind onto the target domain of the *facial recognition* (Fig. 1). That is, *the black box* *cognitive metaphor*, both cognized in our technical imagery and conceptualized in our technological imaginary, is constructed at least in part out of what cognitive linguists George Lakoff and Mark Johnson term “a container schema” and “source-path-goal schema” (1999: 32). A container schema is, like other image schemas, not only “conceptual” but also “instantia[ble],” in Lakoff and Johnson’s view, with “an inside, a boundary, and an outside,” where this “physical boundary […] impose[s] forceful and visual constraints,” “protect[ing] the container’s contents” while “render[ing] them inaccessible to vision” (1999: 32). The source-path-goal schema, like the container schema, Lakoff and Johnson suggest, is useful for conceptualizing “spatial relations,” both “topological” and “trajector[ial],” in which something moves from a first, imagined point to a final, intended position “at a given time” (1999: 33). From the perspective of this metaphorization, facial recognition, like “faculty psychology” in Lakoff and Johnson’s survey, can be compared with a “flow chart,” where a box or boxes “represen[t] various separate faculties carrying out their functions,” including “algorithms” or “metaphorical machines” and their operations upon some information (1999: 413). This cognitive metaphor about facial recognition and its non-visible processuality, like other such metaphors analyzed by Lakoff and Johnson, is motivated in the mind through pre-linguistic image schemas, where “seeing what’s in the box correlates with knowing what’s in the box” (1999: 48), grounded by interactions between the human organism and our semiotic artifacts.

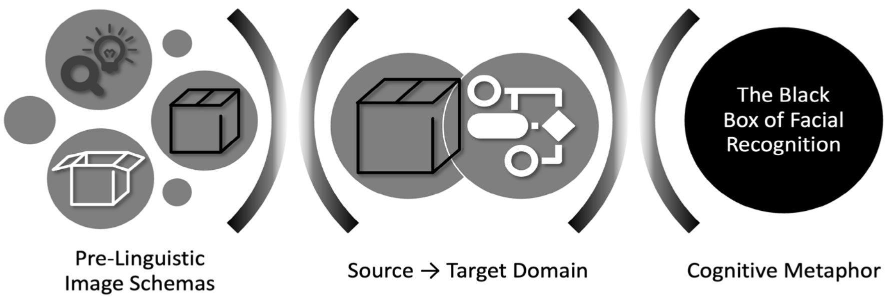

Figure 1: A process diagram for *the black box* cognitive metaphor, from the source domain of the black box to the target domain of the facial recognition, based upon pre-linguistic image schemas, including the container and source-path-goal schemas. © Schiller 2025.

Almost immediately upon the public release of this private research about facial recognition in the mid to late 2000s, scholars from across disciplines began to call for the *unveiling*, *unpacking*, or *uncovering* – that is, the *unboxing* – of proprietary black boxes in facial recognition technologies. More "closed" and "*opaque*" than "open" and "*transparent*" facial recognition in the "social-technical infrastructure[s]" of our everyday lives, Lucas Introna, a scholar in the social study of information systems, points out, make it "particularly difficult" to scrutinize, which, reciprocally, "creates unprecedented opportunities [for] "invisible" micro-politics" (2005: 77). Consequently, Introna calls for a "*dis*closive ethics" that would *make visible*: first, "a place for ethics" to be "immediately present" in the "actual operation[s] of power" in technology; second, the "values and interests," whether "intentional or unintentional," in facial recognition from its "technical detail[s] to social practices" (2005: 75, 79). However, in the case of artificial intelligence, "there is no singular black box to open, no secret to expose," as appraised by Kate Crawford, the co-founder and former director of the AI Now Institute in New York (2021: 12). Rather, Crawford finds, artificial intelligence is constituted by "a multitude of interlaced systems of [technocratic] power" (2021: 12). "Complete transparency," Crawford traces, is an "impossible goal" (2021: 12). To better understand the way artificial intelligence works and its "role in the world," Crawford resolves, one should rather focus on the "material architectures" and "contextual environments" behind artificial intelligence systems "by tracing how they are connected" (2021: 12), that is, their *interrelations*. Joy Buolamwini, who founded the Algorithmic Justice League while at the MIT Media Lab, calls attention to the fact that computer scientists themselves do not always know exactly how "some weights are strengthened [,] and others are weakened," when training an intelligence (2023: 53). Consequently, Buolamwini herself concludes, the current methods in computer science "do not allow us to explain in detail" the ways in which an artificial intelligence technology identifies "a pattern like a face" or "outputs a response" (2023: 53). Although "the term 'black box' [is] used to describe" such artificial intelligence systems, Buolamwini acknowledges, it is important and necessary to "closely examine the […] systems being developed" (2023: 53). But significantly, Buolamwini suggests, the "training data" is most critical to understand the risks involved in applying an artificial intelligence like facial recognition, and we must open the box and shed light upon where "the data comes from, who collected it, and how it is organized" (2023: 53). Most recently, for example, the European Parliament adopted the Artificial Intelligence Act on March 13, 2024. For this landmark piece of a supranational legislation, the member states of the European Union prepared amendments to proposed regulation on the "machine learning capacities" of artificial intelligence systems, among other operational aspects, where the Commission in Amendment 19 refers to outputs "difficult for humans to understand, monitor and trace back" to inputs, with "complex and opaque characteristics," that is, a "black box element," which "impact[s] accountability and explainability" (2024: 11–12). Perhaps semiotics affords possible solutions.

3. Critical Deblackboxing: The Black Box, Its Input, and Output

The term *black box* originally entered the scientific literature around the mid twentieth century in the form of a cognitive metaphor, design principle, and epistemic model, that is, a thought experiment. Indeed, the black box is first given its full treatment by the pioneers in cybernetic theory who then preceded the field of artificial intelligence. Although the earliest known usage of the term "black box" specifically derives from slang dictionaries, between the seventeenth century and the nineteenth century, where it refers to "a lawyer." And the pioneering conceptions of a black box, or at least of black box-like systems, can be found in works on thermodynamic transfer or about electrical engineering, including James Clerk Maxwell's scientific textbook *Theory of Heat* in 1871, Franz Breisig's patent application *Method and Arrangement for Determining Crosstalk in Multicircuit Systems* in 1921, and Wilhelm Cauer's magnum opus *Theory of Linear AC Circuits* in 1941. It is Norbert Wiener who then famously compares what he foundationally calls a *black box* and *white box*. A computer scientist, mathematician, and originator of cybernetics, Wiener studied at Harvard under Josiah Royce (1855–1916) and can himself "be considered a philosophical pragmatist," as John Durham Peters and Benjamin Peters contend, who views messages not only "as informational but [also] as operational" (2016: 157–158). In *Cybernetics: Or Control and Communication in the Animal and the Machine* in 1948, Wiener extends Maxwell's "anthropomorphic demon" thought experiment, about some sorting agency that structurally controls a closed system, to the acting on and the changing of "information," that "must be carried" by one process or another, in "physical, chemical, [or] biological system[s]" ([1948] 1961: 57–58). First, Wiener characterizes a "'black box'" as "a piece of apparatus" with an input and an output, which performs an operation "on the present and [the] past of the input," where information is inaccessible about "the structure by which this operation is performed" ([1948] 1961: xi n1). Then, Wiener characterizes a "'white box'" as a "similar network," where "the relation between input and output" is designed in accordance with "a definitive structural plan for securin[g] previously determined input-output relation[s]" ([1948] 1961: xi n1). Although Wiener himself acknowledges that "'black box'" and "'white box'" are simply "convenient and figurative" terms, which, at least in his time, were "of not very well determined usage" ([1948] 1961: xi n1). That is, the 1) nonvisible, opaque black box, 2) the visible, transparent white box, and 3) the semi-visible, translucent gray box a shade in between white and black, are epistemic models of information processing. These models thereby represent *the flow of information* over periods of time, for instance, from the *outside* of an environment to the *inside* of a machine and back *inside out* again. Between the *input* stimuli and the *output* response, the information is transformed in some way or other by the box, its structure, and functions. However, what is observable about the operations *internal* the box itself may only either be 1) *nothing*, as is the case with the black box; 2) *everything*, as is the case with the white box; or 3) *something*, as is the case with the gray box; and, therefore, may have to be inferred from the relation of the change, that is to say, from the box's *external* behavior.

In the manifold sense of its metaphorical significance, a black box can almost be anything and can almost be anywhere. Although a black box should not be confused with quantum physicist Erwin Schrödinger's *cat-in-a-box*, behavioral psychologist B.F. Skinner's *operant chamber*, or any such boxes. To know a box, therefore, one must experience it. Black boxes can be found, in the terms of an Uexküllean biosemiotics, not only in our *Umwelten* (outer environments), like in our machines and prostheses, but also in our *Innenwelten* (inner environments), like in our minds and bodies. Ross Ashby, a psychiatrist, and progenitor of cybernetics, mentions how "[i]n our daily lives we are confronted at every turn with systems whose internal mechanisms are not fully open to inspection" (1956: 86). Furthermore, Ashby maintains, "Black Box theory" can lead to philosophy about "the mechanism that, for whatever reason, is not wholly accessible," which perhaps is paramount "for those who study the brain" ([1958] 1991: 256). One of the many black boxes that each and every day we humans do experience is the very face itself, that is, the relation between facial behavior and mental phenomena. Indeed, as we could well consider it, the artificial black box of facial recognition technology is useful for studying the natural black box of human facial behavior. For a scientific epistemology of the black box, Ashby goes on to explicate, an Experimenter must "act on the Box," by "allowing the Box to affect" them and their apparatuses, thereby "coupling" themselves "to the Box, so that the two together form a system with feedback" (1956: 87). Ashby diagrams this interaction: "Box ⇆ Experimenter"

(1956: 87). The Box and Experimenter, Ashby expounds, "can be viewed as a compound system," in which the Experimenter "acts on the Box" when they stimulate it, the Box acts on the Experimenter when it reacts, and "each ac[t] on the other" ([1958] 1991: 303). "The interactions that occur between box and experimenter," Ashby defines, "are therefore subject to the laws of communication" ([1958] 1991: 303), or, more broadly, the logics of semiotics. For this reason, Ashby then recognizes, the study of a box, and whether, or to what extent, "the Box is behaving in a machine-like way[,] does not require study of its internal details" ([1958] 1991: 304). But rather, Ashby finds, the "*functional connections within a Black Box can be deduced from observations made from without*," such that "to find something of the connections [,] does not demand the opening of the Box" ([1958] 1991: 304). The box thus becomes not only a cognitive metaphor (for how something odd some*when* occurs between input and output within the operations of a system) but also an epistemic model (for the ways in which humans and technologies interact with one another).

Many cybernetic theorists in the mid twentieth century view the epistemic significance of this systems experience from a pragmatist perspective. In a phenomenological approach to his systems theory, philosopher and physicist Mario Bunge presents the black box as "a fiction," that he portrays as a "*genera[l]*" *model* (1963: 346, 357), which is applicable to diverse phenomena, and is aimed at system dynamics. As popularized by Bunge in the early 1960s, "black box theory" represents "a set of concrete systems into which stimuli (*S*) impinge and out of which reactions (*R*) emerge," where "the various kinds of stimuli and responses" are pictured as "signals," which travel along the "many channels (*C*)" that connect "the box with its environment" (1963: 346, parentheses added). The principal variable that Bunge problematizes is "time" (*t*), and the "time delay" or the "reaction time" (*T*), "of the box for [a] channel" (1963: 347). As Bunge points out, here "the principle of antecedence" holds that "[t]he effect cannot occur before the cause" (1963: 357). That is (Fig. 2), the reaction (R^n) at time (t^n) is determined by stimuli (S^n) at all times (t^1, t^2, *etc.)* prior to and including this time (t^n). To Bunge, therefore, "other meaning could be attributed to the variables" for time, like interpretation, processing, representation, or transformation, to an extent that "the theory is almost *abstract*" (1963: 357). Taking after Ashby's treatment, Bunge concludes the box "constitution" to be "altogether irrelevant" because one accounts only for "the behavior of the system" (1963: 346). Writing black box theory across some thirty years, cybernetician Ranulph Glanville characterizes the black box as a "*gedenkendexperiment* (thought experiment)" (2009: 153), much like

Maxwell's demon. To Glanville, a black box 1) is "believed to be distinct," 2) has "observable (and relatable) inputs and outputs," and 3) is "black (that is, opaque to the observer)" (1982: 1). Therefore, Glanville himself traces, "the blackness of a black box," analogically speaking, consists not in the non-visibility of its operations, but in the visible "change in signals" between input and output that one then can interpret to be "caused by the black box" (1982: 1). To put this pragmatically, what matters here is less the inner character of how a black box works and more the outer behavior of its semiotic actions. Wiener, Ashby, Bunge, and Glanville, among others, more or less share this position, ultimately, taking a perspective that is pragmatist. In terms of Peirce's own *pragmatic maxim*, we must "[c]onsider what effects, that might conceivably have practical bearings, we conceive the object of our conception to have," such that "our conception of these effects is the whole of our conception of the object" (CP 5.402). It is not what a black box is on the inside, but rather *what it does* that defines it.

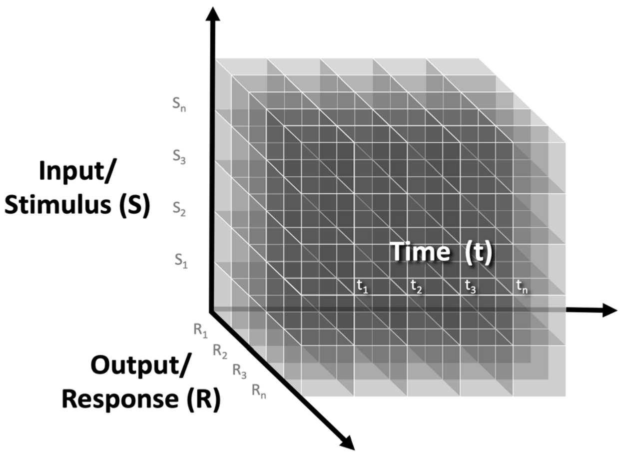

Figure 2: An epistemic model of the black box with input/stimuli (*S*), output/response (*R*), and time (*t*). © Schiller 2025.

But what a black box does and how it does it, idiomatically speaking, might not be visible *on its face* and should not be taken *at face value*. At least in part, the critical deblackboxing in computational semiotics is realized in response to *social constructivist theory* in science and technology studies. And the two approaches generally start from similar grounds. Already in the late 1970s, historian of science and technology Edwin Layton, among others, admonishes *black-boxist* approaches to the study of a technology that represent *input-output* relations with little to no concern for what happens in between. Layton calls for “an understanding of technology from [the] inside,” which includes, for instance, “as a body of knowledge and as a social system” (1977: 198). And Layton calls out any treatment of technology “as a ‘black box’” the “contents and behavior” of which “may be assumed to be common knowledge” (1977: 198). Taking further this foundation in the mid 1980s, Wiebe Bijker, Thomas Hughes, and Trevor Pinch, in their classic treatise on a *technological constructivism*, advance the approach of “‘thick description’” ([1987] 2012: xliii), as first coined by philosopher Gilbert Ryle ([1971] 2009: 497), and further championed by anthropologist Clifford Geertz (1973: 6). Bijker, Hughes, and Pinch apply this approach toward “looking into” not only “the black box of technology” but also “the black box of society” ([1987] 2012: xliii). However, in the early 1990s, political scientist Langdon Winner takes a critical stance toward any social constructivism in science and technology studies, or what is sometimes called *the social construction of technology*. In Winner’s own words, it is not “enough to provide clearer, well-nuanced explanations of technological development” ([1991] 1993: 375). For Winner, social constructivism fails to “call into question the basic commitments [of] technological society” and lacks “a generation position” on the very “patterns under study,” like in a “Marxis[t],” or in a “Heideggerian[ist],” perhaps in a feminist or queer “theory of technology” ([1991] 1993: 375). Simply put, social constructivism addresses technological creation not technological consequence. By the turn of the century, as Irvine himself “notes” (2022: 225n3), philosopher Bruno Latour problematizes the way in which “scientific and technological work is made invisible by its own success” (1999: 304). That is to say, in Latour’s post-humanist view and sociology of science, “[w]hen a machine runs efficiently, […] one need focus only on its inputs and outputs and not on its internal complexity” (1999: 304). “[P]aradoxically,” therefore, Latour postulates, “the more science and technology succeed, the more opaque […] they become” (1999: 304). In a “technical mediation,” such as facial recognition, and its “folding of time and space,” Latour recommends a “reversible blackboxing” (1999: 183–184), that is, a deblackboxing. Latour calls upon inquirers to “[o]pen the black boxes” that are around them, and “examine the assemblies” hidden there deep inside, each of which is, respectively, “itself a black box” (1999: 185). But Latour then asks “[h]ow far *back* in time, [and] *away* in space, should we retrace our steps to follow all those silent entities” (1999: 185)? Which is to say, how far should one trace relationalities from the black boxes through the black boxes to the black boxes, and so on, ad infinitum, that contribute to interactions between a human and a technology? A problem always persists, as Latour has lamented: "the construction of artifacts" in and of itself cannot "account for facts" because these "[n]onhuman[s]" also have agency (1999: 185). Semiotics, however, can reveal this (inter)relationality between human and machine.

4. The Semiotic Machine: Its Subface, Interface, and Surface

The black box in facial recognition, like any other computer system, and unlike many, if not most, other technologies or tools, is structurally, functionally, and constitutively semiotic. Facial recognition is what Mihai Nadin terms a "*semiotic machine*" (2007: 74),[2] not only retrospectively from a top-down critical perspective, but also relationally from its bottom-up design principles. A semiotic machine is "*Janus*-headed," Peter Bøgh Andersen, Per Hasle, and Per Aage Brandt claim, since the "bivalent function" of this "technical artifact" links "two *interface[ing]*" systems (1997: 566): that of the human and that of the computer. Of course, the signs that relate to the human and the signs that relate to the computer are not the same thing and should not be transposed. Both still are *signs*, however. They are *something that stands for something else*, as the Hellenistic Stoics first defined sign function in the third century BCE. But the "machine semiosis" of a semiotic machine, Andersen, Hasle, and Brandt maintain, consists of "processes that take place inside machines, between machines, and [outside machines] between them and their human users" (1997: 548). Tools are not machines. They are too simple. Chemicals are not machines. They are too similar. And importantly, Andersen, Hasle, and Brandt also infer, "living beings are not machines" (1997: 548). Neither are we designed by a human person. Nor do we exist for a human purpose. Indeed, the very handiness or utility itself of a semiotic machine clearly emerges from its cognitive "autonomy," Andersen, Hasle, and Brandt conclude (1997: 566). Therefore, at first sight it would seem, linguist Winfried Nöth laments, a "*semiotic machine*" is "a contradiction in terms" (2002: 84). Even machines, especially computers, however, "*are* involved in sign processes," as Nöth himself concedes (2002: 84). For this reason, in recent years an increasing number of computational semioticians have taken issue with the term semiotic machine. Irvine, for instance, finds that "'computers' are not usefully defined as 'machines'" and introduces the term "cognitive-semiotic artefacts" (2022: 208). Similarly,

[2] Nadin first introduced the critical term of "semiotic machine" in a number of lectures at the Rochester Institute of Technology in New York as early as 1983.

Sørensen, Thellefsen, and Thellefsen adopt the term "technological artefacts," not only for "semiotic entities" but also as "intrinsically semiotic," because they are "*some-thing*," both "designed" and "used," in "processes of signification" (2020: 253–254). Simply put, the pragmatic utility of a hammer tool is to drive in a nail. Similarly presented, the pragmatic utility of a facial recognition technology system is to recognize that which is facial. When it is working, that is all that facial recognition is meant to do. That is all that it does. But the context is more complex. The Janusian face of facial recognition does operate through signs. What is more, facial recognition technology, as an embedded system, or a distributed system, involves many other technologies that operate through signs. Of course, not each of these technologies is semiotic in and of itself. Together, however, they all become entangled with sign processes and semiotic productions.

More and more, not only computational semioticians but also computer scientists widely discuss and highly debate whether, and to what extent, probabilistic artificial intelligence systems, which include, for instance, facial recognition technology systems, constitute a deterministic machine, an indeterministic machine, or a semiotic machine. The "constitutive properties" that characterize a machine consist of its "*organization*" and its "*structure*," with "[t]he *use* to which a machine can be put" part of how to describe, explain, and observe "the machine in a context wider than the machine itself," according to Humberto Maturana and Francisco Varela (1980: 77–78, emphasis added), two biologists who are intellectually famous because they first introduced the concept of *autopoiesis*, that is, system complexity via self-creation, to biology and cybernetics in the early 1970s. The "*organization* of the machine," Maturana and Varela contend, is constituted by "[t]he relations that define a machine as a unity" and that "determine the dynamics of interactions and transformations which it may undergo as such a unity" (1980: 77). Whereas, the "*structure*" of the machine, Maturana and Varela contrast, is constituted by "[t]he actual relations which hold among the components which integrate a concrete machine in a given space" (1980: 77). Human-made machines, however, Maturana and Verela hold, "are all made with some purpose, practical or not, but with some aim (even if it is only to amuse)" specified by humans, although "purpose, aim, or function" are not properties of the machine in and of itself (1980: 78). In his foundational text on this artificial intelligence, James Fetzer, a philosopher of science turned Holocaust, 9/11, and Sandy Hook denier and conspiracy theorist, makes a distinction between: first, "*deterministic* causal systems," for which, "given the same input, the same output invariably occurs (without exception);" and second, "*inde-*

terministic (or "probabilistic") causal systems," for which, "given the same input, one or another output within the same class of outputs invariably occurs (without exception)" (1990: 37). For a causal system to also be a semiotic system, Fetzer decisively further delineates, "it must be a system for which something can stand for something (else) in some respect or other," where that something, that is, that sign "can affect the (actual or potential) behavior of that system" (1990: 37). Indeed, as Fetzer identifies, from the critical standpoint of a computational semiotics it becomes "apparent that for something to affect the behavior of a causal system does not mean that it has to be a sign for that system" (1990: 38). Consequently, one of the principal problems for semiotic studies, Fetzer concludes, is to "distinguis[h] semiotic causal systems from other kinds of causal systems" (1990: 38), as I do here with facial recognition technology.

Indeed, many leading thinkers in computer science today define the computer not through mechanic principles but through "implicit," if not explicit, "semiotic principles," Irvine highlights (2022: 218), like *interpretation*, *processing*, *representation,* and *transformation.* For instance, computer scientist Peter Denning systematically considers in his transformation model of these information processes: first, how "[a] representation is a pattern of symbols that stands for something" else; second, that "[t]he association between a representation and what is stands for can be recorded," as in a database memory or in a dataset table; and third, "a science of information" can therefore be grounded in "the observable affects (signs and referents) without a precise definition of 'meaning'" or signification (2012: 808). This model of information, in Denning's view, may perhaps be most "useful when the computations appear to be strings or streams" (2012: 808), like the continuous feedback loop of a facial recognition technology. Indeed, Denning himself infers, "definitions of information," more often than not, have "an objective component (signs and the things represented by signs) and a subjective component (the meanings)" (2012: 808). Given this "notion of a representation," Denning reasons, "representation-transformation" becomes a "reference model," where "an information process is a sequence of representations," and computation is the transformation from one representation "of the sequence to the next," so that, in the "continuous world," each "time and space," however "infinitesimal," is thus "controlled by a representation" (2012: 808). As Peter Denning and Craig Martell clearly explicate, "computing emphasizes the *transformation* of information," through systematic structures that "are not just descriptive [but actually] *generative*," like how "[a]n algorithm is not just a description of a [problem-solving] method" but actually "causes a machine to solve the problem," where "information caus[es] action" (2015: 16–17). Today, many characterizations of computational methods, as Irvine also identifies, apply the "view of sign-actions […] that Peirce first developed" (2022: 218). At the center of such computation is less representation *from* signs and more operationalization *of* signs, which cause actions with controlled operations, or, in Peircean terms, through law and rule-governed interpretants of sign representamen.

Of course, in a linguistic-verbocentric *semiological* approach, like by Saussure or by Hjemslev, one applies a dyadic or bi-relative sign model to describe the relations between signifier and signified. From this perspective, a communication is channeled, and message is mediated, from the user, through the very facial recognition itself, to the targets. Such *human-computer interaction* (HCI) constitutes a "*one-shot message*," informatician Clarisse Sieckenius de Souza infers, because the computer system more or less "tells" its human users how they should interact with it to achieve this or that effect (2005: 84). "The computer is the channel," in de Souza's interpretation, whereby "higher-level designer-to-user" and "lower-level" user-to-system messages are conveyed as part of a "computer-mediated *metacommunication*" (2005: 88). In the structuralist view of this *semiotic engineering*, a sign, its relations, and functions are principally considered a secondary communication. This "metacommunication," to use psychiatrist Jurgen Ruesch and anthropologist Gregory Bateson's term, is "a communication about communication" that includes all "cues" and "codes" for how information should be interpreted (1951: 209). That is to say, the computer-mediated metacommunication in facial recognition technology is between user and technology, the technology and target, and target and user. Either communication is conveyed at a higher-level between a human sender and a human receiver, that is, user-to-target. Or communication is carried at a lower-level between a human sender and a computer receiver, that is, user-to-technology and target-to-technology.

However, in a logical-gnoseological *semiotic* approach, like by Peirce or by Morris, one applies a triadic or tri-relative sign model to describe the relations between representamen, object, and interpretant. From this perspective, facial recognition technology is composed not only from a system of semiotic *artifacts* (such as computer systems, digital cameras, and display screens), but also from a system of semiotic *agents* (such as the facial recognition software itself, facial detection, extraction, and classification algorithms, their designers, users, and targets). Those who interact with the semiotic machine of this facial recognition do not reductively play just the passive role of a viewer, spectator, or audience – that is, a receiver. Rather, they are an *interactor*. They are an *interpreter*. They are themselves a "semiotic agent," Irvine maintains, who is definitively "presuppose[d]" by the design principles of the black box (2022: 208). In this view, facial recognition technology is greater than the sum of its parts. It is emergent. And it is irreducible.

The semiotic machine of facial recognition is characterized by triadicity, not only in regard to its relationality, but also in terms of its thinginess. Together, however, this human, computer, and human-computer interaction, as theorized by computer scientist, pioneering computer-based media artist, and computational semiotician Frieder Nake,[3] constitute an "*algorithmic sign*" (Hinterwaldner et al. 2014: 293). To analytically specify the algorithmic sign, Nake draws upon Peirce, who explains, for example, that "[a] sign is something, *A*, which brings something, *B*, its *interpretant* sign, determined or created by it, into the same sort of correspondence (or a lower implied sort) with something, *C*, its *object,* as that in which itself stands to *C*" ([1902] 1976: 54). As Peirce also portrays, "a *representamen*" is "that character of a thing by virtue of which, for the production of a certain mental effect," or "*interpretant*," it "may stand in place of another thing," its "*object*" (CP 1.564). In the terms of this triad, a sign therefore has three parts: its *representamen*, *object*, and *interpretant*. Each of these parts can also be a sign in and of itself, when interpreted as such. And each can in turn lead to another sign, potentially interpreted *ad infinitum*. From this theoretical framework, Nake schematizes the algorithmic sign as "a sign [with] two interpretants," one "generated by a human, [and] the other by [a computer]" (Hinterwaldner et al. 2014: 293). When something, like a digital image of a human face, is extracted from a video and enrolled into a computer, like in a facial recognition technology system, Nake proposes, "it exists in two forms," and "gains a two-fold ontology," where it must be "visible," or at least somehow "perceivable," by the human, and it must be "computable by the machine" (Hinterwaldner et al. 2014: 293). Semiotically speaking, Nake suggests, "both acts" by the human and the computer become "acts of interpretation" (Hinterwaldner et al. 2014: 294), which *in the very act itself* generate what Peirce terms interpretants. When engaging with facial recognition technology, the face under *semiosis*, that is, the face of the interactor, is being interpreted by two interpreters: the human and computer, independently and interrelatedly as well as concurrently and

[3] For transparency, I studied under Frieder Nake during my master's degree in the histories of media art from Danube University Krems, when his courses were held at Benedictine monastery Göttweig Abbey, in Lower Austria in May 2015.

continuously. As noted by Nake, the "algorithmic sign" has three faces (Fig. 3): its *inner* face or "*sub*face," *between* face or "*inter*face," and *outer* face or "*surf*ace" (2008: 104, emphasis mine). This cognitive metaphor of the Janusian face or the multi-relational face is "justified" to Nake principally because of the "one-to-one correspondence" between the inner action and the outer appearance of the computer, its hardware, and software (2008: 105). Beyond the interdependent sympathetic causality between human psychology and facial behavior, there is an imperatively stable correspondence in the facial triadicity of the semiotic machine, that is, between its subface, interface, and surface.

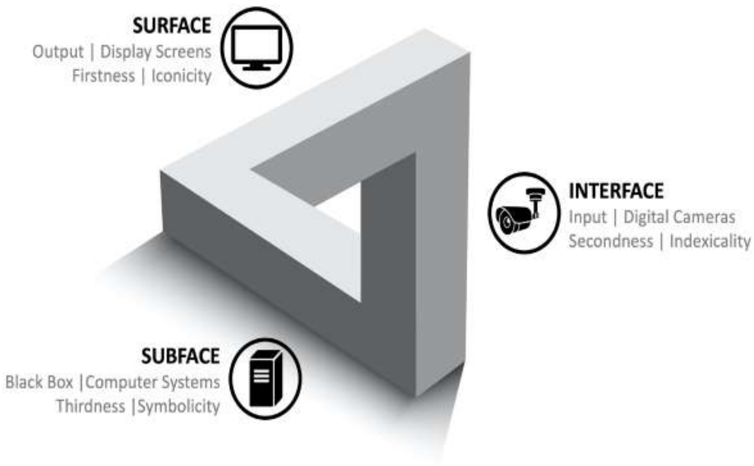

Figure 3: A triadic model, visually diagrammed as a Penrose triangle, that represents the homologies, which are generally extensive but not specifically exclusive, in facial recognition between: semiotic face (subface, interface, and surface); computational software (black box, input, and output); mechanical hardware (computer systems, digital cameras, and display screens); phenomenological categories (thirdness, secondness, and firstness); and sign categories (symbolicity, indexicality, and iconicity). © Schiller 2025.

5. Conclusion

From the critical standpoint of a computational semiotics both pragmatist and Peircean, "the pragmatic dimension of sign processing," Nöth confirms (2002: 95), is a fundamental criterion of a full semiosis. Peirce himself discriminates between the human, the machine, and their minds. In his theory of a quasi-semiosis, Peirce maintains that machines work like minds at least in some, if not in all, respects. He accounts not only for the differences between human semiosis and machine quasi-semiosis but also for the similarities between the human mind and machine minds. Perhaps moreover, Peirce maintains that human minds work like machine minds in some degree and to some extent. In Peirce's own words, a human can be considered "as a machine which turns out," for example, a "sentence expressing a conclusion" after "the man-machine" has "been fed with a […] statement of fact, as premiss" or proposition (CP 2.59), analogously speaking, like in facial recognition, its templates, and matches. As Nöth also notes, the human mind is like a machine mind only when it solves "a task that a logical or calculating machine can *also* solve," that is, when it follows the rules "of a predetermined algorithm in a quasi-mechanical way" (2002: 88-89). Although the most "[f]ormal" and the "least important part[s] of reasoning […] may be performed by a machine," Peirce "hold[s] that reasoning is the observation of relations, mainly by means of diagrams and the like" (qtd. in Ketner and Stewart 1984: 209). Reasoning, to Peirce, therefore, "is a living process" (qtd. in Ketner and Stewart 1984: 209).

However, Peirce primarily hesitates, reasoning cannot be performed "by *the unaided brain*," and, in fact, it "needs the cooperation of the eyes and hands," in the "art of this experimentation," to "use all [different] kinds of diagrams and devices for aiding the imagination" (qtd. in Ketner and Stewart 1984: 209, emphasis added). As posited here by Peirce, reasoning may be aided by semiotic artifacts for off-loaded cognition, as in facial recognition and its artificial intelligence. Generally, Nake and Grabowski find, the "*inter*face [is] the *face*" of such a software (2006: 67). To separate the interface between human, computer, and their interaction from the surface of the display screens and the subface of the computer system, Nake and Grabowski speculate, would "render software *faceless*" (2006: 67). A computer system, Nake and Grabowski specify, "cannot exist without [a] face;" indeed, a computer system exists *only when* it "show[s] its face;" and *without a face*, "it does not exist at all" (2006: 67). With facial recognition technology during human-computer interaction, the human act of sign interpretation is brought into relation with the computer operation of signal determination. Human-computer interaction, however, is neither bilateral nor reciprocal in "a true sense," Nake infers (2008: 107). In the final analysis, this relation is the result, not of artificial intelligence or the artificial mind, but of human intelligence and the human mind. With the artificial intelligence that is facial recognition, like the oracles of olden times, the one true face among its Janusian relationalities is the face of *those who use it*.

References

Andersen, P. B., P. Hasle and P. A. Brandt. 1997. Machine Semiosis (*Maschinensemiose*). In Posner, R., K. Robering, and T.A. Sebeok (eds.). *Semiotik: Ein Handbuch zu den zeichentheoretischen Grundlagen von Nature und Kultur/Semiotics: A Handbook on the Sign-Theoretic Foundations of Nature and Culture, Theilband 1/Volume* 1. Berlin: Walter de Gruyter, 548–570.

Ashby, W. R. 1956. *An Introduction to Cybernetics*. London: Chapman & Hall.

Ashby, W. R. (1958) 1991. General Systems Theory as a New Discipline. In Klir, G. J. (ed.). *Facets of Systems Science*, Second Edition, International Federation for Systems Research International Series on Systems Science and Engineering 7. Boston: Springer, 299–308. Available at: https://doi.org/10.1007/978-1-4899-0718-9_16. Originally published in Bertalanfly, L. von, A. Rapoport (eds.). *General Systems: Yearbook of the Society for the Advancement of General Systems Theory*, Vol. 3. Ann Arbor, MI: The Society for General Systems Research, 1–6.

Bijker, W. E., T. P. Hughes and T. Pinch. 2012 [1987]. General Introduction from the Original Edition. In Bijker, W. E., T. P. Hughes and T. Pinch (eds.). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. 25th Anniversary Edition. Cambridge, MA: The MIT Press, xxxix-xliv.

Bunge, M. 1963. "A General Black Box Theory". *Philosophy of Science*, Vol. 30, No. 4, 346–358. Available at: https://doi.org/10.1086/287954.

Buolamwini, J. 2023. *Unmasking AI: My Mission to Protect What is Human in a World of Machines.* New York: Random House.

Crawford, K. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven, CT: Yale University Press.

Denning, P. J. 2012. "Opening Statement: What is Computation?". *The Computer Journal*, Vol. 55, No. 7, 805–810.

Denning, P. J., C. H. Martell. 2015. *Great Principles of Computing*. Cambridge, MA: MIT Press.

European Parliament. 2024. Artificial Intelligence Act: MEP'S Adopt Landmark Law. *European Parliament News*, 13 March 2024. Available at: https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law.

Fetzer, J. H. 1990. *Artificial Intelligence: Its Scope and Limits*. Studies in Cognitive System (COFS) 4. Dordrecht: Kluwer.

Gates, K. A. 2011. *Our Biometric Future: Facial Recognition Technology and the Culture of Surveillance*. New York: New York University Press.

Geertz, C. 1973. Thick Description: Toward an Interpretive Theory of Culture. In Geertz, C. *The Interpretation of Cultures: Selected Essays*. New York: Basic Books, 3–30.

Glanville, R. 1982. "Inside Every White Box There Are Two Black Boxes Trying to Get Out." *Behavioural Science,* Vol. 12, No. 1, 1–11. Available at: https://doi.org/10.1002/bs.3830270102.

Glanville, R. 2009. "Black Boxes". *Cybernetics and Human Knowing: A Journal of Second Order Cybernetics, Autopoiesis and Cybersemiotics*, Vol. 16, No. 1–2, 153–167.

Hinterwaldner, I., L. Manovich, M. Warnke and F. Nake. 2014. "Roundtable: From Algorithmic Signs to Cultural Analytics". *Zeitschrift für Ästhetik und Allgemeine Kunstwissenschaft*, Vol. 59, No. 2, 291–302. Available at: https://doi.org/10.28937/1000106251 (accessed 14 March 2024).

Introna, L. D. 2005. "Disclosive Ethics and Information Technology: Disclosing Facial Recognition Systems". *Ethics and Information Technology*, Vol. 7, No. 2, 75–86. Available at: https://doi.org/10.1007/s10676-005-4583-2.

Irvine, M. 2022. Semiotics in Computation and Information Systems. In Pelkey, J., S. W. Matthews (eds.). *Bloomsbury Semiotics, Volume 2: Semiotics in the Natural and Technical Sciences*. London: Bloomsbury Academic, 203–237.

Ketner, K. L., A. F. Stewart. 1984. "The Early History of Computer Design: Charles Sanders Peirce and Marquand's Logical Machines". *The Princeton University Library Chronicle*, Vol. 45, No. 3, 187–224. Available at: https://www.jstor.org/stable/26402393.

Lakoff, G., M. Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G., M. Johnson. 1999. *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. New York: Basic Books.

Latour, B. 1999. *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Layton, E. 1977. Conditions of Technological Development. In Spiegel-Rösing, I., D. de Solla Price (eds.). *Science, Technology and Society: A Cross-Disciplinary Perspective*. London: Sage, 197–222.

Leone, M. 2024. "Semiotics of the Black Box: On the Rhetorics of Algorithmic Images". *Visual Communication*, Vol. 23, No. 3, 426–451 Available at: https://doi.org/10.1177/14703572241247120.

Maturana, H. R., F. Varela. 1980. *Autopoiesis and Cognition: The Realization of the Living*. Boston Studies in the Philosophy and History of Science (BSPS) 42. Dordrecht: D. Reidel Publishing Co. Originally published in Spanish: 1972. *De Maquinas y Seres Vivos*. Santiago: Editorial Universitaria S.A.

Meunier, J.–G. 2022. *Computational Semiotics*. Bloomsbury Academic.

Nadin, M. 2007. "Semiotic Machine". *The Public Journal of Semiotics*, Vol. 1, No. 1, 57–75. Available at: https://doi.org/10.37693/pjos/2007.1.8815.

Nake, F. 2008. Surface, Interface, Subface: Three Cases of Interaction and One Concept. In Seifert, U., J. H. Kim and A. Moore (eds.). *Paradoxes of Interactivity: Perspectives for Media Theory, Human-Computer Interaction, and Artistic Investigations*. Transcript Verlag, 92–109.

Nake, F., S. Grabowski. 2006. The Interface as Sign and as Aesthetic Event. In Fishwick, P.A. (ed.). *Aesthetic Computing*. The MIT Press, 53–70

Nöth, W. 2002. "Semiotic Machines". *S.E.E.D. Journal (Semiotics, Evolution, Energy, and Development)*, Vol. 3, No. 3, 81–99.

Ovid. (8 CE) 2011. *Times and Reasons: A New Translation of Fasti*. (transl. Anne Wiseman and Peter Wiseman. Oxford: Oxford University Press.

Peirce, C.S. (1902) 1976. *The New Elements of Mathematics, Volume 4: Mathematical Philosophy*. Carolyn Eisele (ed.). The Hague: Mouton Publishers. (Cited as NEM.)

Peirce, C. S. (1931–1958). *The Collected Papers of Charles Sanders Peirce*. 1931–1935, Vols. 1–6, Charles Hartshorne and Paul Weiss (eds.). 1958, Vols. 7–8, ed. Arthur W. Burks (ed.). Cambridge, MA: Harvard University Press. (Cited as CP.)

Peters, J. D., B. Peters. 2016. "Norbert Wiener as Pragmatist." *Empedocles: European Journal for the Philosophy of Communication*, Vol. 7, No. 2, 157–172. Available at: https://doi.org/10.1386/ejpc.7.2.157_1.

Ruesch, J., G. Bateson 1951. *Communication: The Social Matrix of Psychiatry*. New York: W. W. Norton and Co.

Ryle, G. (1971) 2009. The Thinking of Thoughts: What is 'le Penseur' Doing? In Ryle, G. *Collected Papers, Vol. 2, Collected Essays: 1929–1968*. London: Routledge, 494–510.

Sørensen, B., T. Thellefsen and M. Thellefsen. 2020. A Peircean Semiotics of Technological Artefacts. In Jappy, T. (ed.). *The Bloomsbury Companion to Contemporary Peircean Semiotics*. London: Bloomsbury, 253–275.

Souza, C. S. de. 2005. *The Semiotic Engineering of Human-Computer Interaction*. Acting with Technology 3. Cambridge, MA: The MIT Press.

Wiener, N. (1948) 1961. *Cybernetics: or the Control and Communication in the Animal and the Machine*. Second Edition. Cambridge, MA: The MIT Press.

Winner, L. (1991) 1993. "Upon Opening the Black Box and Finding it Empty: Social Constructivism and the Philosophy of Technology". *Science, Technology, & Human Values*, Vol. 18, No. 3, 362–378. Available at: https://doi.org/10.1177/016224399301800306. Originally published in Pitt, C., E. Lugo (eds.). *The Technology of Discovery and the Discovery of Technology: Proceedings of the 6th International Conference of The Society for Philosophy and Technology*. Blacksburg, VA: Society for Philosophy and Technology, 503–519.