Perspectiva Filosófica, vol. 49, n. 5, 2022

ANOTHER BRICK IN THE WALL: THREATS TO OUR AUTONOMY AS SENSE-MAKERS WHEN DEALING WITH MACHINE LEARNING SYSTEMS[1] [2]

Outro tijolo na parede: Ameaças à nossa autonomia como fazedores de sentido ao lidar com sistemas de aprendizagem de máquinas

---

Camila Leporace[3]

ABSTRACT

Autonomy, as proposed by the enactive approach to cognition, is the capacity that living organisms have to follow norms constituted by their own activity. This concept is linked to the concepts of *sense-making* and *participatory sense-making*, the former encapsulating the cognizer's ability to bring forth a world of meaning through its coupling with the environment – and being affected by its surroundings on an ongoing basis – and the latter being an extension of this idea, which applies to interactive processes in which at least two agents find themselves involved in. In this essay I advocate that, when dealing with machine learning systems, which cannot be considered autonomous, the agent or cognizer cannot sustain his or her autonomy in the same way as would be possible in an encounter with another agent. The reasoning is developed in three threads: the unbalanced encounter in which the cognizer's autonomy is threatened; the reduction of the range of experiences an autonomous agent could have; the lack of awareness of the cognizer concerning rules and potential risks of the systems he is dealing with. Even though these risks are focused on in the essay, the opportunities offered by machine learning systems are also recognized. To take advantage of them, it

---

[1] DOI: https://doi.org/10.51359/2357-9986.2022.252618

[2] This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. My thanks are also directed to PUC-Rio and the University of Coimbra. I am especially grateful to Ezequiel Di Paolo (Ikerbasque - Basque Foundation for Science); Mark Coeckelbergh (University of Vienna); Ralph Ings Bannell (PUC-Rio); Giovanni Rolla (UFBA) and Eros Carvalho (UFRGS) for the careful reading and comments which were essential to the development of this essay.
O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Direciono os meus agradecimentos também à PUC-Rio e à Universidade de Coimbra. E agradeço especialmente a Ezequiel Di Paolo (Ikerbasque - Basque Foundation for Science); Mark Coeckelbergh (Universidade de Viena); Ralph Ings Bannell (PUC-Rio); Giovanni Rolla (UFBA) e Eros Carvalho (UFRGS) pela leitura atenta e pelos comentários que foram essenciais para o desenvolvimento deste ensaio.

[3] Pontifícia Universidade Católica do Rio de Janeiro e Universidade de Coimbra (Portugal). E-mail: camiladoutoranda@gmail.com. ORCID: https://orcid.org/0000-0001-8293-949X.
Camila De Paoli Leporace é doutoranda em Educação pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio, Brasil) e pela Universidade de Coimbra (Portugal). Mestre em Educação pela PUC-Rio e graduada em Jornalismo pela mesma instituição, tem como principais interesses de pesquisa a mente e a cognição humana, a aprendizagem e as articulações entre inteligência artificial e educação.

is necessary to seek a balance that encompasses the inherent human capacity for intersubjectivity permeated by affectivity.

Key-words: Autonomy; Enactivism; Machine learning; Sense-Making; Participatory Sense-Making

RESUMO

A autonomia, tal como proposta pela abordagem enativa para a cognição, é a capacidade que os organismos vivos têm de seguir as normas constituídas pela sua própria atividade. Esse conceito está ligado a outros dois conceitos do enativismo, de *sense-making* e de *participatory sense-making* – o primeiro referindo-se à capacidade do cognoscente de produzir um mundo de sentido através da sua ligação com o ambiente, e de ser afetado pelo meio de forma contínua, e o último constituindo uma extensão desta ideia, que se aplica a processos interativos em que pelo menos dois agentes se encontram envolvidos. Neste ensaio, defendo que, ao lidar com sistemas de *machine learning*, que não podem ser considerados autônomos, o agente ou cognoscente não pode sustentar a sua autonomia da mesma forma que seria possível sustentá-la num encontro com outro agente. A argumentação desenrola-se em três pilares: o encontro desequilibrado em que a autonomia do cognoscente é ameaçada; a redução do leque de experiências que um agente autônomo poderia ter e o desconhecimento do cognoscente quanto às normase às ameaças potenciais dos sistemas com os quais está a lidar. Ainda que o ensaio seja centrado nesses riscos, as oportunidades oferecidas por sistemas algorítmicos como os de aprendizagem de máquina são também reconhecidas. Para aproveitá-las, aponta-se que é necessário buscar um equilíbrio capaz deenglobar a intersubjetividade, inerentemente humana, permeada pela afetividade.

Palavras-chave: Autonomia; Enativismo; Machine learning; Sense-Making; Participatory Sense-Making.

INTRODUCTION

We have been increasingly inhabiting environments permeated by machine learning algorithms. Embedded in this kind of *algorithmsphere*, we are at the same time the providers and the consumers of these machines' operations, which essentially rely on our data to perform. Since *big data* also means big responsibility (DE FILIPPI, 2014), AI powered machines have proven to be a double-edged sword[4]. Consequently, if, on the one hand, machine learning and other algorithmic systems may be powerful technological

[4] See Bannell (2017) for an overview of some important issues concerning machine learning and education.

resources, on the other it is necessary to analyze the possible threats which the use of these systems may bring to light. Not by chance, there is a global effort around legislation, evaluation and control mechanisms for algorithm-based systems so that they operate transparently, in compliance with the data they gather from their users and their purposes.

"Learning" in machine learning basically means to predict future values based on the past; it is, fundamentally, statistics. Still, much has been discussed about the meaning of 'learning' in this context, being the possible level of autonomy that can be attributed to machine learning systems one of the hot related topics. This deliberation has led to controversial issues in the ethics of artificial intelligence. Autonomy can be connected to queries concerning responsibility and accountability regarding the machines' actions and decision-making processes, for example. The term autonomy, which derived from the Greek *autos* (*self*) and *nomos* (*law*), captures the idea of self-government, the agent's ability to decide how to act (VÉLIZ, 2021, p. 490)[5]. Although there are different interpretations for that concept, autonomy is also a recurrent topic in education. A largely accepted definition would be something like the students being able to perform, decide and move forward by themselves, making improvements and achieving results as an outcome of their own actions. *Autonomy* is also one of the basal concepts proposed under the enactivist approach for the mind and cognition, together with *emergence*, *experience*, *embodiment* and *sense-making*.

In this essay, my intention is to provide a critical analysis of machine learning systems for *human* learning – taking education here in a broader sense, that is, as the ongoing development of human beings. More specifically, I focus on the possible consequences that may emerge when cognizers, i.e., human learners, make use of machine learning systems with the purpose of perceiving and experiencing the world. Enactivism, as developed by Di Paolo, Cuffari and De Jaegher (2018); Di Paolo, Rohde and De

[5] In the cited article, Carissa Véliz develops the idea that algorithms are like moral zombies. She mentions the eight ways how Nomy Arpaly defines autonomy as the term has been conceptualised in the literature – as personal efficacy; psychological independence; having a moral right to self-determination; authenticity; having a coherent self-image; being heroic; self-governance and being responsive to reasons. The author argues that among these concepts the ones that seem relevant to morality are self-governance and reasons-responsiveness, although they do not apply to "moral zombies or algorithms", according to the author (See Véliz, 2021, pp. 487-497).

Jaegher (2010); De Jaegher and Di Paolo (2007) depicts humans as *sense-makers*, able to engage in intersubjective activities with other human beings. In this framework, affection is an intrinsic dimension of the cognitive processes humans entangle themselves in. To be a *sense-maker* is to be affected by the environment and by others. In their coupling with the world, *sense-makers* are autonomous beings, able to maintain an ongoing energy and matter exchange with their surroundings. Autonomy, in the enactivist sense, does not imply an independence from the environment; on the contrary, it relies on the very means to maintain it: a codependent relationship with the cognizer's milieu.

That being said, I take the following questions as a starting point: how can we approach humans' autonomy when we take machine learning systems as their surroundings? Are we still autonomous beings, in the enactivist sense[6], when embedded in these algorithmic niches? How do these environments affect our perception and learning experience, also in the enactivist sense, as cognizers? We are embedded in a world that is not just a source of inputs for us to process, but an arena for us to act out and bring forth the results of our perception-action cycles. We produce something that is used as data and the data we produce are applied by neural networks to make predictions[7]. Taking enactivism as a departure, it is necessary to comprehend what kind of world of meaning we are able to constitute around us through the mutual process of generating data and being shaped by the very data we generate.

What I would like to do, then, is to articulate some of the underlying concepts from enactivism with the discussion about the possible threat machine learning could engender to our autonomy as cognitive beings. My decision to focus on the autonomy of cognizers (and not so much on the machines' putative autonomy) is due to the fact that it is a meaningful cause

6 The concept will be introduced further in this essay.

7 "Machine learning algorithms analyze the data to identify patterns and to build a model which is then used to predict future values (for example […] by identifying patterns in photographs of named people, it predicts who is shown in other photographs; and by identifying patterns in medical symptoms, it predicts a specific diagnosis). In other words, machine learning may be considered a three-step processs (analyze data, build a model, undertake an action) that is continuously iterated (the outcomes of the action generate new data, which in turn amends the model, which in turn causes a new action). It is in this sense that the machine is learning" (Holmes, Bialik and Fadel, 2019, p. 89).

of concern for education, maybe more than the autonomy of the machines themselves, although undoubtedly the latter intertwines with the former. Especially in times when algorithms and *big data* raise a feeling of insecurity in society in many ways – questions particularly relevant to education[8] – I believe we should discuss the digital environment in which we live and that is, at the same time, shaped by us. Enactivism represents a rich framework for this analysis.

After introducing the concepts of *autonomy, experience* and *sense-making* in enactivism, I argue that our autonomy as sense-makers tends to acquire new contours when we deal with machine learning systems. One reason why this happens, which is developed here, is because these systems cannot be considered *autonomous sense-makers*; hence, there is no interaction when we deal with them. Interaction is, after all, something that takes place between two or more agents[9], according to the branch of enactivism adopted here (DI PAOLO, ROHDE & DE JAEGHER, 2010). Given that the *algorithmsphere* is the environment being built around us, what emerges from our coupling with this world is reconfigured. As a consequence, the norms and constraints usually applicable to the environment as we understand it may not apply to this scenario. This prompts us to rethink the ways in which that balance could be preserved. To develop my argument, I also resort to the concept of *participatory sense-making*, seeking to reflect on our relationship with algorithmic systems on the basis of this concept.

My main concern is the hypothesis that algorithmic systems, although being helpful in some circumstances as technological extensions of our minds (CLARK, 2003, 2011, 2014), end up restricting the cognizer to a more or less fixed network of interactions and possibilities. Consequently, this could configure a harmful circularity that would possibly reduce the cognizer's autonomy and its possibility for expanding its knowledge and experience beyond that constrained universe. I believe that, while investigating human autonomy in a world that has been valuing (maybe overvaluing) the abilities of machines, or at least has been privileging discussions on them,

---

[8] See Bannell (2017) for an overview of some important issues concerning machine learning and education.

[9] I assume that machine learning systems are not cognizers; further on the essay I will explain why.

we can recover some of the important features which are essential for education – at least if we consider education's potential from an enactivist point of view. Namely, supporting a shift from the cognitivist to the enactivist paradigm (STEWART, GAPENNE E DI PAOLO, 2010) could enable us to depict the cognizer – in a broad learning perspective – as an agent who must be able to make good choices that ultimately assure its integrity, individuality and ongoing learning experience. The cognizer must build his own perspective for thinking, being regarded as a subject capable of changing and developing continuously, consequently being able to lay down his own path[10] actively collaborating and sharing with others.

Although I am aware of relevant ethical, political, social and cultural implications of the algorithmic society that we constitute, my focus here are the impacts on human learning and perceiving, taking a new approach to cognition as a departure – being enactivism an approach not yet so popular in the field of education. This, for sure, is not a dimension detached from the ethical, political and aesthetic ones; however, I would like to look carefully at what happens to our ability to learn and perceive the world when we deal with artificial systems that are supposed to "learn" – about us. That is, I seek to investigate the maintenance of human autonomy, in the enactivist sense, in a world where the possibility of machines' self-determination mobilizes attention.

On the following section I will introduce the concept of *autonomy* as proposed by enactivists as well as the other major concepts that constitute this approach. After doing so, I will address the issue of how the agent makes sense of the world, according to enactivism. I will then go on introducing the concepts of *sense-making* and *participatory sense-making*, key ideas for my reasoning, and, afterwards, focus on the discussion about whether the agent can sustain its autonomy when dealing with machine learning systems with the goal of learning and developing.

[10] As Thompson (2007, p. 13) reminds us, "Borrowing the words of the poet Antonio Machado, Varela described enaction as the laying down of a path in walking: 'Wanderer the road is your footsteps, nothing else; you lay down a path in walking' (Varela 1987, p. 63).

2. Autonomy

The enactivist conception of the mind, as developed by Di Paolo, Cuffari and De Jaegher (2018), De Jaegher and Di Paolo (2007); Colombetti (2014) and others, implies the inseparability between cognizers and their surroundings. It has been emerging as an alternative approach to the traditional cognitivist paradigm that depicts the human cognitive system as an input processor and an output generator, being the mind (essentially equated to the brain) responsible for processing information that goes into this "sandwich" of inputs and outputs (HURLEY, 2001). In the enactivist approach, action meets perception: cognizers must act in the world if they are to grasp it, their bodies being intertwined with the environment in an ongoing manner. It is also an important feature of enactivism to connect cognition and affection, bridging this gap; the mind is intrinsically or constitutively affective, according to Colombetti (2014).

Following the enactivist proposal, a cognizer's very being depends on the coupling between the operations and energy exchanges that go on inside of it and those that take place outside of it. Consequently, the environment is vital for the organism, as without it cognizers would not even be able to constitute themselves. However, it is not essential only in the sense that it offers something to be absorbed by the being, as if organisms were a kind of crunching machines (CLARK, 2014), but as part of a network in which they are embedded and that becomes meaningful to them as they live and act in it. As already mentioned, together with the concepts of *emergence*, *sense-making*, *experience* and *embodiment*, *autonomy* is one of the main concepts grounding the enactivist approach to mind and cognition. Before actually explaining the concept of autonomy as proposed by enactivists, it may be useful to introduce their ideas regarding individuation. The relations between the cognizer and the world are "inherently meaningful" (DI PAOLO, 2018), and

> [T]hese relations cannot be meaningful unless individuation is an ongoing, open, precarious process; i.e., a non-stationary one. The possibility of unpredictable, frame-transforming changes is inherent to being a cognitive sys-

> tem, even in the particular circumstances where these changes are not actually occurring (DI PAOLO, 2018, p. 76).

According to Di Paolo, Rohde & De Jaegher (2010), the living organisms' autonomy comes exactly from their self-generated identity as distinct entities. Being autonomous means following "laws setup by their own activity", which denotes that

> A system whose identity is fully specified by a designer and cannot, by means of its own actions, regenerate its own constitution, can only follow the laws contained in its design, no matter how plastic, adaptive, or lifelike its performance. In order for a system to generate its own laws, it must be able to build itself at some level of identity. If a system "has no say" in defining its own organization, then it is condemned to follow an externally given design like a railroad track (DI PAOLO, ROHDE AND DE JAEGHER, 2010, p. 37).

Self-constitution, in enactivism, is linked to the idea of *operational closure:*

> A system is operationally closed if, for any given process *P* that forms part of the system (1) we can find among its enabling conditions other processes that make up the system, and (2) we can find other processes in the system that depend on *P*. This means that at some level of description, the conditions that sustain any given process in such a network always include those conditions provided by the operation of the other processes in the network, and that the result of their global activity is an identifiable unity in the same domain or level of description (DI PAOLO, ROHDE AND DE JAEGHER, 2010, p. 38).

Although the autonomous agent defines its individuality by constituting itself in an ongoing basis, its cognitive activity relies on the continuous exchanges with the environment it finds itself immersed in. As Hans Jonas (1966) would put it, and enactivists embrace, the life of the agent bears a relation of *needful freedom* with the environment. When acting in the environment, the agent is not merely affected by it in a passive way, i.e., responding to "external perturbations" but "in fact actively and asymmetrically *regulate* the conditions of their exchange with the environment, and in doing so, enact a world or cognitive domain" (DI PAOLO, ROHDE AND DE JAEGHER, 2010, p. 38).

The enactivist concept of autonomy has its roots in the concept of *autopoiesis*, proposed by Maturana and Varela in the 1980s (see MATURANA & VARELA, 2019). At the time they developed it, they were devoted to searching for an explanation of what makes a system a living one, beginning with the minimum viable life form: the very conception of life. Having highlighted the cell as this basic unit, Maturana & Varela (2019, p. 40) demonstrated that "the biological foundations of cognition cannot be understood by examining solely the nervous system". In the search for comprehending the mind, they proposed, we must examine "how the cognitive processes are founded in the totality of the living being"; hence the need to understand features connected to the organization of life.

The idea that living beings are characterized by actively and continuously producing themselves is the autopoietic organization Maturana and Varela speak of. Based on the constant exchanges it establishes with its environment, a cell remains alive. Its identity is delimited by its membrane, which is, at the same time, generated by the cellular activity itself and responsible for delimiting the cell as a unit (MATURANA & VARELA, 2019, pp. 52-53). Yet, although the membrane circumscribes the singleness of the cell in relation to the environment in which it is embedded, it does not turn the cell into an entity completely detached from its environment, because the cell's maintenance requires a constant fluidity in both directions (inner-outer and outer-inner) so that this unity can be sustained. If there is an imbalance between the cell and the environment, however, and if that imbalance becomes such that this unit falls apart, the cell loses its identity: it disappears into its milieu (MATURANA & VARELA, 2019, pp. 54-55).

It must be elucidated that, for a living being to be considered an autonomous cognizer in the enactivist sense of the concept, it must possess another capacity, which is the one of *adaptivity* (DI PAOLO, 2005). This is exactly the capacity of cognizers to seek this equilibrium, constantly monitoring and improving their state, then being able to identify "tendencies that bring them closer to the boundary of viability" (Di Paolo, Rohde and De Jaegher, 2010, p. 50). In doing so, if the cognizer is able as well to counteract these tendencies, it can be considered a sense-maker. In the next section I will introduce he idea of sense-making as proposed by enactivists.

3. A world of meaning emerges

Meaning making needs to be analyzed from a specific perspective in enactivism. Following the logic of this paradigm, there is no such thing as ready-made information to be absorbed by the cognizer, as the cognitivist approaches would imply. The world exists even though the cognizer cannot perceive it. However, the elements of the world will only be available for the cognizer, then constituting its network of meaningful objects, when they affect the cognizer somehow. That is why enactivism is an approach that characterizes cognition as "affective", in the sense that "the sense-making living system is inherently purposeful (…) its autonomous organization makes it into a system that continuously aims or endeavors to be itself and as such is 'concerned' about its continuation" (COLOMBETTI, 2014, p. 18).

To comprehend this, we need to focus on the concept of *sense-making*. According to enactivism, even a cell can make sense of the world, of its environment, since its living activity is guided by the elements that, being around it, actually mean something for its survival. All sense-makers are *affected* by the elements of the world that mean something to them and to their organization, as Colombetti (2014) puts it. This is not the same as saying that a cell, for example, thinks and acts through reasoning and decision-making processes; rather, the cell is oriented towards the elements that surround it in a way that can assure its life. In other words, the cell is guided by the elements which are relevant for its operation and maintenance. In a way that parallels this, we act, cognitively; namely, while acting to perceive the world in which we inhabit, we attribute meaning to the elements that guide our activity. This suggests that, at the same time that we act to perceive the world, we generate a world of meaning around us. It is a world that makes sense to our activity and that is, therefore, able to sustain our autonomy. That is why sense-making is closely related to the notion of *Unwelt*, as coined by Uexküll ([1934] (2010). "For a living system to be a sense-making system is to live in a world that is always an Unwelt, namely, an environment that has a specific significance or value for it" (COLOMBETTI, 2014, p. 17).

We are sense-makers essentially because we have bodies; we are embodied agents – a condition that leads to one more of the foundational concepts from enactivism, namely the one of *embodiment*. “The body is material, dynamic and self-organizing” (DI PAOLO, BUHRMANN & BARANDIARAN, 2017, p. 5). The activity of sense-makers in the world depends “nontrivially” on the body, according to Di Paolo, Rohde and De Jaegher (2010), who postulate that

> [f]or the enactivist the body is the ultimate source of significance; embodiment means that mind is inherent in the precarious, active, normative, and worldful process of animation, that the body is not a puppet controlled by the brain but a whole animate system with many autonomous layers of self-constitution, self-coordination, and self-organization and varying degrees of openness to the world that create its sense-making activity (DI PAOLO, ROHDE & DE JAEGHER, 2010, p. 42).

While acting in the world and bringing forth the environment, cognizers or sense-makers live in the world and experience it; this fourth basic concept of enactivism, *experience*, is far from the idea of “an epiphenomenon or a puzzle as it is for cognitivism”, being otherwise “intertwined with being alive and immersed in a world of significance”. Experience “goes beyond data to be explained” (DI PAOLO, ROHDE & DE JAEGHER, 2010, p. 43). The following example, which resembles Dreyfus’s (2014) concept of *skillful coping,* can be useful for a further comprehension of experience in an enactivist sense:

> Becoming a wine connoisseur is certainly an achievable goal but expertise in this field (as in any other) is not obtained through gaining the right kind of *information* but through the right kind of *transformation* — one that can only be brought about by appropriate time-extended training (experimenting, making mistakes, and so on). Experience is altered in a lawful manner through the process. It is itself a skillful aspect of embodied activity (DI PAOLO, ROHDE & DE JAEGHER, 2010, p. 44).

After being introduced to the concepts of *autonomy, sense-making, embodiment and experience* in enactivism, it is necessary to understand the concept of *emergence* in the context of this approach. According to Di Paolo, Rohde and De Jaegher (2010), the idea of emergence appears in different domains regarding debates in metaphysics and epistemology and has been revived over the last three decades due to the advent of the sciences of complexity. As these authors explain, however, enactivism adopts a pragmatic application of the concept that originates from self-organization. "Emergence is used to describe the formation of a novel property or process out of the interaction of different existing processes or events" (Thompson, 2007; Thompson and Varela, 2001; *apud* Di Paolo, Rohde & De Jaegher, 2010, p. 40). The cell as the basic life unit is again useful to illustrate the concept:

> We find the clearest example of emergence in life itself. The property of continuous self-production, renewal, and regeneration of a physically bounded network of molecular transformations (autopoiesis) is not to be found at any level below that of the living cell itself (DI PAOLO, ROHDE & DE JAEGHER, 2010, p. 41).
> It seems ill-conceived to call any of the component parts (a protein, the DNA strands, *etc.*) alive: these are just physical structures that can be isolated, the material substrate of the living cell that is constantly changed and renewed. It is undeniable, however, that the phenomenon of life is as real as it could be (ROHDE, 2010, p. 22).

In the next section, I will introduce the concept of *participatory sense-making*, which, as an extension of the idea of *sense-making*, is relevant for the reasoning I intend to develop here.

4. Is human-machine interaction (with machine learning) actually an interaction?

In addition to reframing the way we conceive the individual's relationship with the world in order to comprehend cognition, enactivism as proposed by De Jaegher and Di Paolo (2007) postulates a new approach to social cognition[11]. *Participatory sense-making* is the point of departure for this approach. The idea is that, from the interaction between two individuals (at least), something new emerges, which simply would not exist if this interaction did not happen; from then on, the outcomes from this interaction themselves start to define its course, that is, if it continues or is interrupted,

[11] See De Jaegher and Di Paolo (2007) and De Jaegher and Di Paolo (2008).

as well as its fluidity and balance. Participatory sense-making is, then, the coordination of intentional activity in interaction, whereby individual sense-making processes are affected, and new domains of social sense-making can be generated that were not available individually (De Jaegher and Di Paolo, 2007, p. 497). The concept of *participatory sense-making* extends the concept of *sense-making*; the former is the active engagement with another social agent, whereas the latter is the general active coupling with the world.

At this point, it should be elucidated how the authors define interaction:

> An enactive approach to social understanding starts from the study of interaction and coordination. *Interaction* is here understood as the coupling between an agent and a specific aspect of its world: another agent. Interaction is the mutual interdependence (or bidirectional, co-regulated coupling) of the behaviors of two social agents (DI PAOLO, ROHDE & DE JAEGHER, 2010, p. 60).

And social interaction:

> Social interaction is the regulated coupling between at least two autonomous agents, where the regulation is aimed at aspects of the coupling itself so that it constitutes an emergent autonomous organization in the domain of relational dynamics, without destroying in the process the autonomy of the agents involved (though the latter's scope can be augmented or reduced) (DE JAEGHER & DI PAOLO, 2007, p. 493).

To mention a daily situation, for example, a conversation with a friend, if she at some point demonstrates through her gestures or other features that she is upset about something we said, the direction of the conversation may change, or the interaction may even be bluntly interrupted. If, on the other hand, our friend offers us a smile, it may represent a "green light" to go on speaking more confidently about the difficult subject we selected or one that makes us nervous. Thus, the very elements that emerge from the interaction are those that constitute it and are as well the ones which are capable of sustaining it. If this interaction stops at any time, that is, it "shuts down", it means that these elements were unable to remain coordinated in a way that could sustain it. In a sense, the same goes for the individual's relationship with the environment in general: it needs to continue to support itself, through both its internal and external operations, both of which must remain in constant exchange and looping – this leads to the definition of *autonomy* and also *emergence* as proposed by the enactivist approach. Just as it is not possible for an interaction to be sustained by only one of the individuals involved, exactly because there should be at least two, the analysis of the interaction cannot be reduced to individual behavior.

Once again, it can be seen that this is valid in relation to the cognizer and the environment: the capacity to sustain the living being is not in one *or* the other; it is in the ongoing exchanges between them in a way that ensures the permanence of this living being's life. The ability to interact, therefore, arises from the interaction itself. It is not possible to previously conceive of individuals as "interactors"; "individuals co-emerge as interactors with the interaction" (DE JAEGHER & DI PAOLO, 2007, p. 492). On the other hand, the autonomy of each of those involved in the interaction process cannot be disrupted, since, if it is broken, it also interrupts the interaction, according to the enactivist perspective. As De Jaegher and Di Paolo put it, resorting to dance as a form of interaction,

> [C]ouple dancing involves moving each other, making each other move, and being moved by each other. This goes for both leader and follower. Following is part of an agreement and does not equate with being shifted into position by the other. If the follower were to give up her autonomy, the couple dancing would end there, and it would look more like a doll being carried around the dance floor. The same goes for conversations: each partner must engage from an autonomous standpoint (DE JAEGHER & DI PAOLO, 2007, p. 494).

To look at the explanation of what happens in an interaction between two individuals, from the perspective of enactivism, can be a good key for us to seek an understanding of the human relationship with machine learning systems. The dance example suggests the following reasoning: at first, dancers are equal partners when engaged in such an interaction, as their individual autonomy is sustained during the process. It means that they have a balanced individual "weight" in the process that emerges from their coupling, even if one is an experienced dancer and the other is learning its first movements, for example. Following this reasoning, one could ask: what is the distribution of this "weight" when the "dance partner" for a human being is a machine learning system? Are we like puppets carried by machine learning systems? Or would these systems be like puppets carried by us? Maybe not exactly (one way or the other), but could we then make different "moves" depending on how this coupling takes place or on how much we know about this coupling or about the partner, so as to manage it, to cope with it? Would it somehow threaten our autonomy as sense-makers?

This question comes in the wake of others, one of which I have already mentioned: are machine learning systems autonomous themselves? Although this question raises others and is beyond the scope of this essay, as it would itself demand an extended discussion, I propose that we take a brief look at it through enactivist lenses. First of all, living beings are not the only autonomous systems we can account for. According to Thompson (2007, p. 44), autonomy can be observed in a system of a very different nature, that is, without a biochemical organization like that of the cell, and without a membrane made of material stuff; this is the case of a colony of insects, which form a network whose frontier or "membrane" is social and territorial, not material like the cell's. Autonomy could also be observed in an artificial system (THOMPSON, 2007, p. 50). However, there is a primordial aspect here: autonomous beings or systems are distinguished from systems determined by their external side, which are called heteronomous systems. The following quotation by Rohde (2010) may help clarify this argument:

> The constraints imposed on self-maintaining processes of identity generation are of a *mechanical* nature. Living organisms are bound by the laws of physics but the possibilities to reorganise themselves and, with them, the world of meaningful interactions they bring forth, are open-ended. This open-endedness contrasts with the explicit design of adaptive circuits in computationalist approaches, e.g., in the discipline of machine learning (ROHDE, 2010, p. 20).

Machine learning systems are developed based on certain parameters, embedded in the algorithms. It is said that, as they operate, they "learn", meaning that data already internal to them become raw material for new data to be generated (without these data, these kinds of systems cannot exist). Although they may be fed by data that initially came from outside (human activity external to the system), data also emerges from its inside operation insofar as the systems are used, in a circular movement. That said, should these systems be classified as autonomous or heteronomous? This depends on the perspective adopted. As Thompson (2007) postulates, the paradigm for the exchanges between a heteronomous system and the environment surrounding it is the paradigm of inputs that come from the outside to be processed and then generate outputs. When the expected output does not occur, it is understood that there was an error in the system. On the other hand, "[t]he paradigm for interaction with an autonomous system is a conversation, in which unsatisfactory outcomes are seen as breaches of understanding" (VARELA, 1979, p. xii, *apud* THOMPSON, 2007, pp. 37-38).

At first, a connectionist system would probably be considered heteronomous; but the artificial neural networks that constitute machine learning systems are not only fed by information from the outside. So, it seems that machine learning systems oscillate between these axes. However, as Rohde (2010, p. 20) explains, "Even if machine learning is a blossoming field as part of software engineering, such algorithms are *functionally* constrained by in-built rules". That is, the rules already exist in the systems before they start operating and were previously designed; so, it is unlikely that new rules will be generated during the operation of these circuits. On the other side, in the course of human interactions new rules are created all the time.

Also, according to Thompson (2007), the human mind emerges from processes of self-organization that intimately interconnect the brain, the body and the environment at various levels. Following this reasoning, machine learning systems would only be able to have a mind like the human mind if they had bodies intertwined with the environment and a specific organization like the ones found in living beings[12]; and it seems that, so far, only embodied, organic, sensorimotor and intersubjective creatures can be

[12] One of the most important features for enactivism is the organization of the being and its strategies to sustain it while embedded in the environment; so, if it were possible to instantiate the organization of a living being/a cognizer/a sense-maker in an artificial system, it would be possible to have an artificial mind. But the being would still lack a lived history, an idea which is linked to the concept of enactive becoming; and it would make it harder for an artificial agent to become a human *person* (see Di Paolo, 2020).

taken as cognitive agents[13]. Lacking an organic body, which would be the first mandatory feature for them to have cognition, systems lack the intrinsic purposiveness inherently found in autonomous systems (COLOMBETTI, 2014). Therefore, maybe machine learning systems can still present a certain level of autonomy, as, even if they don't have organic bodies, they are still able to sustain a balance between their interior and exterior that keeps them operating. However, I believe it is not possible to say that it would be the same to relate, or "dance", with an artificial machine learning system as it would be to do so with another human being.

I will now move on to articulate some more ideas that hopefully will clarify this reasoning a little more.

5. Does dealing with machine learning systems allow us to be autonomous sense-makers?

For all the reasons stated before, it may seem feasible to say that there can be some level of autonomy in a machine learning system; however, it also seems coherent to say that there isn't any participatory sense-making when we, humans, always ready to become interactors, engage with these systems. This is especially so because the activity of making sense of the world recruits the body as a whole, as well as emotions, and cannot be dissociated from all of these elements – which are still present when humans interact with machine learning systems – not in the system, of course, but in humans. Participatory sense-making, however, requires the presence of at least two bodies, therefore two minds; two agents; two interactors; to sum up, two autonomous beings or sense-makers.

As De Jaegher and Di Paolo (2007, p. 498) say: "[W]e encounter situations where, through coordination of sense-making, one of the interactors

---

[13] Di Paolo and Thompson (2014) say that "To the extent that robots fail to exhibit the dynamics of living systems, they are not self- regulating, autonomous agents, and their *sense-making* is not physically grounded in autopoietic and metabolic processes. Thus, there is good reason to think that they cannot have emotions like ours and are incapable of making sense of their surroundings via affective framing. This inevitably makes their ways of interpreting their surroundings very different from the *sense-making* that creatures like us carry out on a regular basis". Zebrowski and McGraw /2021), however, discuss an approach to construed artifacts like robots that, founded on habits, could enable them to become genuine meaning-making systems.

is oriented towards a novel domain of significance that was part of the sense-making activity of the other". We can see from this quotation that participatory sense-making leads to the unexpected, enables surprise to take place, in a way that can only happen in the presence of two individual sense-makers – although the very "possibility of unpredictable, frame-transforming changes is inherent to being a cognitive system" (DI PAOLO, 2018, pp. 76-77). The author also postulates that unpredictability is connected the fact that individuation should be an ongoing, precarious process.

Can the same amount or degree of unpredictability be expected when dealing with a machine learning system? Probably not, but how does this impact our autonomy? If the possibility of the unexpected itself cannot account for our autonomy alone, maybe it can at least account for it partly. But how? Predictability – or unpredictability – seems to be one of the keys for guiding interaction. If it orients participatory sense-making encounters, it is, in the first place, a basic feature of sense-making. Positing the differences between the concept of *behavior* as a motor program and as a constantly unfolding kinetic, Sheets-Johnstone mentions a hunt as a good example for analyzing the way we move in the world, which I mention here since it is linked to this issue:

> A hunt is not and cannot be a specific and repeatable sequence of actions, both because the world is not the same from one day to the next or possibly even from moment to moment with respect to terrain or weather, for example, and because the movement of living creatures is not the same from one day to the next or even from moment to moment. In essential ways, it is unpredictable (SHEETS-JOHNSTONE, 2010, p. 176).

We can't predict what will happen when dealing with the world because it changes all the time, and so do we. Similarly, it does not seem possible to predict everything that will emerge from the encounter of two humans, as the meeting itself generates something new which comes into being exactly from the interaction. This extends to language, too:

> Living language is directed, signified and resignified, vitalized, framed, and sculpted by the carnality of interactive engagements between real people in ways that cannot be fully determined nor fully predicted by the capabilities, intentions, and experiences they bring to the encoun-

> ter, nor by the broader constraints of social norms, environmental situation or grammar rules. The locus where language is enacted is the often neglected middle term between the individual mind and the broader sociocultural patterns (DI PAOLO, CUFFARI & DE JAEGHER, 2018, p. 132).

I will now develop my reasoning by showing how, when dealing with machine learning systems, the cognizers find themselves in situations that cannot be depicted as participatory sense-making encounters and, consequently, 1) lacking interaction/another interactor, do not allow the sense-maker to adjust itself to the other, generating an unbalanced encounter in which the cognizer's autonomy is threatened; 2) tend to reduce the kinds of experience lived by the cognizer, consequently provoking a decrease of the chances for the cognizer to remain an autonomous being in the maximum possible sense of its experience in the world; 3) being situations in which norms are hidden from the cognizer, generate an unfair scenario for the cognizer's autonomy, as acting in the world in a way that ensures autonomy would demand being aware of its rules and potential risks.

5.1. The unbalanced encounter in which the cognizer's autonomy is threatened

Social interaction involves the goals of (at least) two individuals. Both of them are oriented toward some kind of objective, and the encounter itself is a result of both individuals' plans and expectancies, even if they may change during the encounter. As a consequence, the level of autonomy achieved depends on these goals. Let me extend my argument with another example from De Jaegher and Di Paolo (2007, pp. 500-501). In the situation they mention, which I call situation (1), a pantomime game is happening in which a person from a group of participants tries to make gestures representing something that their group needs to guess (e.g. the title of a movie) to score points in the game. The goal is to make it easier for team players to guess. The person doing the gestures then tries her hardest to "learn" from what the group is guessing wrong until she improves her gestures and the participants get the answer right. What happens in this context? This person adjusts her movements based on what the other members of her group demonstrated, so that she could help them guess it, which led the whole group to score points together. Participants also adjusted their guesses to the gestures of the person who acts out the mime.

Now let me explain what I call situation (2), not mentioned by De Jaegher and Di Paolo, but which I have made up as an extension of their example: what if, for some reason, the participant responsible for miming wished to confuse the members of his group instead of leading them to the right answer about what that series of gestures meant? The player would probably gesture differently, working to increase the distance between the participants' responses from correct guesses so that the group would end up failing. In both examples, there is a goal to be achieved and intentions involved. A machine learning system has its targets, too. But, to what extent do these goals operate as situation (1) or situation (2)? That is, are the goals of these systems meant to act as the ones of a team player like in situation (1) or more like the ones of someone who intends to confuse us, as in situation (2)?

There are (at least) two important features connected to these questions: algorithms don't have real intentions, although they are created to sustain a certain logic that should lead to the reach of specific goals. In the first place, machine learning systems don't feel anything and are not affected by the environment; they are not sense-makers; the way they "behave" depends on how the system was programmed, the purposes and goals that back the programming. Built-in rules are embedded in the algorithms. Even so, in a sense there would be no difference between dealing with a machine or with a human being; after all, the intentions and goals (or *apparent* intentions and goals) of one or the other can always vary and sense-makers should adjust their moves according to what they perceive about the other and the situation – even if the other is a system.

However, if we look closer and consider that the system may seem like an interactor but in fact is not, we must realize that a system cannot make adjustments based on our contingencies, feelings, gestures, guesses, facial expressions; at least it would not be able to make adjustments as if it was in an interaction, something that necessarily requests at least two cognitive beings according to the enactivist proposal of participatory sense-making. These systems keep following their goals blindly, indifferent to sense-making and to the sense-maker, whatever happens to the cognizer. Being blind to the other (the sense-maker), machine learning systems may confuse us sometimes or lead to outcomes that would fit their "needs", but not ours. Because there is no way an interactor can regulate its moves if the other part doesn't do the same, it becomes an unbalanced situation, where the cognizer can't modulate himself to search for what would be better for him, his experience, his autonomy and so on. Consequently, it could be said his autonomy is threatened.

5.2. A narrower range of kinds of experiences

There is another slightly different way the imbalanced situation of dealing with machine learning systems may threaten the cognizer's autonomy, one that takes us back to the prediction-or-surprise problem I raised at the beginning of this section. Besides hardly ever being "neutral" – as they are created and fed by humans, and humans are not "neutral" – and being inflexible to the sense-maker, those systems always operate based on predictions, i.e., on what has happened in the past. These predictions neglect, for example – because algorithms neglect – bodily contingencies, social history, interactions with others and the possible changes in prospective adjustments, being blind to the context, the feelings involved and so on; there is, after all, no real interaction going on. If prediction-based machines can't cope with their users and sustain a real interaction with them, actually transforming, growing and evolving with them, when they block the possibility of mutual adjustment they prevent the new from emerging.

Eventually, something new in this sense can only appear as a result of real interactions, a genuine participatory sense-making context. Dealing with machine learning systems, consequently, leaves little room for surprise and flexibility, for mutual accommodation and transformations that could emerge. One could ask: does less surprise lead to less autonomy or simply to a reduction in the quality of that user's experience? Following the enactivist logic, I do not believe that it is possible to dissociate these two, as the aesthetics of experience seems to be a significant dimension intrinsic to the autonomy of the individual[14], which is linked to the concepts of emergence and experience. The aesthetics of experience is closely connected to affection as well, being emotions an essential dimension of the human cognitive processes. Why (and how) would an impoverishment of the quality of lived experience impact the autonomy of the cognizer? To fully understand this part of my reasoning, I find it important to bring back the aforementioned concept of *adaptivity* as developed by enactivism.

> An autopoietic entity can be robust to perturbations without the logical necessity to actively monitor its own state and act to improve the conditions for continued autopoiesis. Only *adaptive* autopoietic entities that improve the conditions for continued autopoiesis, by actively monitoring their own state, identifying at least some tendencies that bring them closer to the boundary of viability and counteracting these tendencies can be actual 'sense-makers' (DI PAOLO, ROHDE & DE JAGHER, 2010, p. 50).

As Di Paolo et al say: "[a]daptivity (Di Paolo, 2005) is what enables living bodies to distinguish a situation as a risk or an opportunity, to tell the difference between good or better, bad and worse" (Di Paolo, Cuffari & De Jaegher, 2018, pp. 32-33). When dealing with the world the agent gets to know by its own experience – in its organic, sensorimotor and intersubjective dimensions – it can evaluate the viability of its actions and so move away from situations that threaten its autonomy.

> [s]ense-making is the capacity of an autonomous system to adaptively regulate its operation and its relation to the environment depending on the virtual consequences for its own viability as a form of life. Being a sense-maker implies an ongoing (often imperfect and variable) tuning to the world and a readiness for action. Through the combination of material and precarious self-individuation and adaptive regulation of the relations to the environment, sense-making naturalizes the concept of vital norms and lies at the core of every form of action, perception, emotion, and cognition, since in no instance of these is the basic structure of concern and caring ever absent. This is

[14] The idea of aesthetics as an inescapable dimension of our cognitive activity is a mark of the work of the pragmatist John Dewey – see Dewey (2008), and has been explored by Mark Johnson (2017, 2018), well known for his ideas regarding embodied cognition. Johnson (2018, p. 2) defines us as "*homo aestheticus* – creatures of the flesh, who live, think, and act by virtue of the aesthetic dimensions of experience and understanding".

> constitutively what distinguishes mental life form from other material and relational processes (DI PAOLO, CUFFARI AND DE JAEGHER, 2018, p. 33).

The more we know about the world, the more we experience it in its totality, being immersed in disturbance and perturbations that demand acquiring skills to cope with them. Different kinds of experience seem to enlarge our possibilities to develop such abilities and become *skillful copers*, in the sense proposed by Dreyfus (2016). It seems that, on the other hand, diminishing the possibilities of experiencing different demands may result in a decrease in the development of these skills.

5.3) The lack of awareness of rules and potential risks

Although it is necessary for a sense-maker to safeguard his or her autonomy by continuously acting in the world and improving his abilities while dealing with it, enactivism has also something important to say about the normativity that guides the interaction. In order to keep safe and oriented towards what seems to be the best for him, the cognizer needs to be aware of these norms, as well as the risks he or she may face when coping with the world. The agent can also collaborate to reframe theses norms. But how would the sense-maker be able to decide what is best for its life when dealing with a system guided by norms he or she does not know and didn't contribute to establishing? Would the agent be able to make the best possible choices for its adaptive activity in this scenario? In the first place, given that machine learning systems are based on predictions, which are rooted in the past, they tend to limit these possibilities for the sense-maker (this takes us back to subsection 5.2). Second, this is aggravated by the fact that the systems are generally oriented towards some pre-defined goals which are not transparent (this takes us back to subsection 5.1).

The "normativity" created by the machine is determined by the goals it is expected to achieve. It tends to repeat the past, which is something that constrains future possibilities, thus creating a more restricted environment for the human cognizer. This problem is increased by the fact that the system's user is not always aware that the system has been programmed with certain pre-defined objectives, and that it obeys this strict pre-formatted regulation. Maybe we should highlight that what is sold by digital marketing agencies and big data advocates, in general, is that prediction-based machines will always know what you want or need based on your previous choices and actions (the footprints you leave behind that are used to make predictions about yourself and deliver content to you). However, even if there is awareness regarding how these systems work, it is often unusually hard to guess what these goals would be. They are like a black box, as the systems don't actually interact with us, having no minds, no bodies, no expressions and no emotions to do so. So, how could this situation be aligned with the way autonomy in enactivism is shaped? It seems it could not. When we deal with other people, unless we find ourselves in a situation of unequal relations of force or domination, we can try to reinstate equilibrium and act together in order to sustain our autonomy: keeping the balance between our inside operation and the outside; including the other; keeping the equilibrium of the interaction while keeping our autonomy as individual cognizer.

As I have mentioned, I will not go deep into the ethical implications of AI in this essay (for a detailed analysis on these matters, in a way I consider complementary to the ideas I expose here, see Coeckelberg, 2020); however, at this point, I should endorse that maybe it is not possible to detach the cognitive activity from the ethical implications of machine learning systems. After all, when the agent couples with its environment choices should be made, and the consequences of this coupling matter to the agent.

6. Discussion

As we have seen, cognitive beings or sense-makers seek not only a fluid, ongoing coupling with the environment to safeguard their autonomy, but the best possible coupling. In this "dance" with the world, cognizers learn about the norms that guide the universe in which they are embedded and can act and think accordingly, even contributing to modify them at some level. On the other hand, the normativity present in a machine learning system, or in several interconnected systems, is fixed and externally given, besides being, in general, unknown to the cognizers. This may hinder our possibilities for action, since the engagement to which we are used to, one that is regulated by our mutual bodily actions of intersubjectivity, allows our experiences to be distinguished by surprise, unpredictability and dealing with others whilst constantly adapting ourselves.

We can say that the individual *learns* from its coupling with the world, but this *learning* is far beyond learning in the sense featured in a machine learning system – which cannot be considered an agent, as a system disconnected from the experience of others and mainly oriented by the choices of the ones who were responsible for creating it, who fix certain parameters into the system through algorithms and models. We must remember, once again, that algorithms are not neutral in themselves (O'Neil, 2016); by definition, they are like snippets of the world, snapshots. A cutout is, after all, always partial, leaving out a large part of the picture (note that sometimes the blind spot is exactly the most important part of the picture that should never have been left out in the first place!). In this sense, the world is *still* its own best model, as Brooks (1991, p. 15) says.

In order to sustain our autonomy, not only seeking mere survival but also trying to find the best matches for our expectations, we should have the world within our reach. Yet, not a reduced world, which leaves out essential parts of our minds – our bodies and their organic, sensorimotor and intersubjective dimensions and all the possibilities for engaging with the world and others that emerge from them; the emergent features of all these corporeal-material networks. Once restricted to a certain concept of mind that is limited to the skull, to which enactivism is a theoretical-practical alternative, one essentially reduces learning to the brain's activity. To match the activity of making sense of the world to mere information decoding and processing corresponds to a classification of that which is present in the world that would fit into an atomized Cartesian ontology where the parts are perceived before the whole. We, however, cope with the world as it is, as an assembly, a continuum, with all its instabilities and inaccuracies (see Dewey, 2008) and its precariousness (DI PAOLO, CUFFARI & DE JAEGHER, 2010, for example). And it is precisely these experiences that allow us to perceive the environment in its rich aspects and multifaceted dimensions, allowing us to proceed then to an analysis of the parts, if we wish to do so. It is precisely throughout this process that we can develop our capacities as autonomous sense-makers. The cell that follows sugar particles for its survival does not need to know the sugar formula to make this movement; it simply does it. Similarly, we don't need to know the chemical composition of something to deal with it in our daily coping with the world (although we could, if we were chemists at a lab, for example). For doing so, each being counts on its physicality, bodily features, perceiving abilities and so on. We should be able to experience the whole world and not a pre-determined perspective, to sense-make, to make as we sense. As Varela, Thompson and Rosch (2016) say, we "lay down a path in walking".

In the case of machine learning, since we do not know each system individually, at least we must be aware of the way they work, generally; eventually, knowing the principles that underpin them, it is possible to contrast their operation with our own cognitive activity. This would help us not lose sight of our sense-making and autonomous abilities. If it can be argued, as I have done in this essay, that our autonomy may be restricted by machine systems, I believe that this can at least prepare us for this eventuality. The space for bodily, intersubjective experiences, the space for surprise and the emergence of something new is, after all, essential, especially for education and the formation of a human being. To experience the world is to possess an autonomous character, in the sense of being able to seek the best possibilities for our subject-world coupling. One should not lose sight of the fact that the new environments created for us and by us, sense-makers, after the advent of machine learning, do not maintain the same characteristics of environments created from the participatory sense-making of subjects with interconnected physical bodies. Perhaps this is the most important aspect, or at least one of them, for an analysis mainly focused on possible consequences for education. In a sense, it seems that the best extension for a human being, more than technology, is still "the other"[15].

I would like to highlight that autonomy, in the enactivist sense, exists and is maintained in the coupling with environments that, although unstable and regularly subjected to changes, are not hostile to experience (in

[15] It could be mentioned here the concept of empathy as developed by phenomenologists (Zahavi, 2021).

a wider sense both of the concept of experience and the concept of hostility). These couplings, after all, provide the means for a constant becoming (DI PAOLO, 2020), as living beings, especially linguistic beings, find themselves in an ongoing dynamic flow. This state of becoming requires the constant search for the best possible coupling between subject and world. So, if machine learning, today, offers restricted possibilities for the subject to couple with the world, this must be kept in mind before making decisions that involve the inclusion of such systems in educational experiences, the objective of which is to form an autonomous subject. This aspect cannot be detached from ethical issues and possible threats to human integrity and privacy. Machine learning systems have been introduced into the educational scenario as learning technologies, and they are also used for evaluating and selecting students for certain opportunities – applications that may influence their future. Coeckelbergh (2020) alerts to the vulnerability to which we as machine learning users have been exposed, considering that our data is collected in a way that not always cares for our privacy and the right to know what happens to all this data – and make choices accordingly.

We could react negatively to this if we knew what is going on, and then not consenting would be a way of avoiding the loss of our autonomy, for example. When Coeckelbergh (2020, p. 111) says that "Machines can be agents but not moral agents since they lack consciousness, free will, emotions, the capability to form intentions, and the like", and comments on how data science recruits humans to prepare, analyze and interpret the results brought up by machine learning networks, he is implying that machines should not be left alone in decision making processes that can impact humans lives. Although machine learning systems may end up "making decisions", in a merely statistical sense, they lack features which are essential for these decision-making processes. One could argue that humans may lack the needed sense of justice and fairness for these processes as well, but at least we can ask and try to answer the question of what a fair society is, while machine learning systems remain mostly black boxes (or at least boxes with content and possible outcomes which cannot be coordinated by us). Moreover, and in a primordial way, from the enactivist standpoint – encapsulated in the concept of participatory sense-making – when dealing with humans we can at least adjust expectancies and have a human-to-human, mind-to-mind, body-to-body dialogue. This only happens when dealing with others; it does not occur when dealing with blind, mindless, disembodied, emotionless artificial systems. These systems may raise concrete walls around us, instead of opening up fluid roads for us to keep building our own paths; in this sense, our data becomes another brick in the wall. Maybe we should start taking a look between the flaws in the bricks.

CONCLUSION

In this essay, I have developed the idea that, when cognitive agents deal with machine learning systems, their autonomy – in the enactivist sense of the concept – may be threatened. In order for it to be fully maintained, it would be necessary for the agent to be able to interact with the system in a similar way as it interacts with another cognitive agent, constantly readjusting expectations and collaborating in a readjustment of the norms involved in this interaction. However, this is not possible when we take into account the machine learning systems we have been dealing with. This is a concern particularly pertinent when it comes to human learning and development in a broad sense. We are, after all, entangled with machine learning systems in a continuous manner.

As I mentioned in the introduction, these powerful technological resources are double-edged, that is, they may offer opportunities as well as risks. So, by making an effort to provide a critical perspective, although I shed light on one of these dimensions, I don't deny the other. I believe it is by evaluating what happens when we deal with algorithm-based resources that we can comprehend how to get the best out of them. So, besides pointing out some of the possible threats of such technologies, I have also attempted to highlight the importance of human interaction in an algorithmic world. Balance, once more, is the key idea. Balance is essential for a sense-maker to deal with the environment the same way it is important for us, as a society, in our search for ways of coupling and coping with such resources. We are connected to digital and algorithmic technologies but, first and foremost, we are and have always been connected to each other. Thus, making sense of the world together, through participatory sense-making, involves making sense of human-technology relations and figuring out the challenges linked to them.

Recebido em 31/03/2022

Aprovado em 13/05/2022

REFERÊNCIAS

BANNELL, R. Uma faca de dois gumes. In: FERREIRA, Giselle dos Santos. *Educação e Tecnologia: abordagens críticas*. Rio de Janeiro: SESES, 2017.

BROOKS, R. A. Intelligence without representation. Artificial Intelligence, 47, 139–159, 1991.

CLARK, A. *Natural-Born Cyborgs. Minds, Technologies and the Future of Human Intelligence*, New York: Oxford University Press, 2003.

CLARK, A. Supersizing the mind: embodiment, action, and cognitive extension. Oxford: Oxford University Press, 2011.

CLARK, A. *Mindware*. Cambridge: MIT Press, 2014.

COECKELBERGH, M. *AI Ethics*. Cambridge and London: MIT Press, 2020.

DE FILIPPI, P. Big data, big responsibilities. *Internet Policy Review,* 3(1), pp. 1-12, 2014. https://doi.org/10.14763/2014.1.227

DE JAEGHER, H. e DI PAOLO, E. A. Participatory Sense-making: An enactive approach to social cognition. *Phenomenology and Cognitive the Sciences*, 6, pp. 485-507, 2007.

DE JAEGHER, H. e DI PAOLO, E. A. Making Sense in Participation: An Enactive Approach to Social Cognition. In: *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions*. F. Morganti, A. Carassa, G. Riva (Eds.) Amsterdam, IOS Press, pp. 33-47, 2008.

DEWEY, J. *The Later Works of John Dewey, Volume 1, 1925 - 1953: 1925, Experience and Nature*. 1ª ed. Illinois: Southern Illinois University Press, 2008.

DI PAOLO, E. Autopoiesis, adaptivity, teleology, agency. Phenomenol. Cogn. Sci. 4, 429–452, 2005. doi: 10.1007/s11097-005-9002-y

DI PAOLO, E. Extended Life. *Topoi* **28,** 9, 2009. https://doi.org/10.1007/s11245-008-9042-3

DI PAOLO, E. A., BURHMANN, T. E BARANDIARAN, X, E. *Sensorimotor Life. An Enactive Proposal*. Oxford: Oxford University Press, 2017.

DI PAOLO, E. The Enactive Conception of Life. In: NEWEN, A., DE BRUIN, L. & GALLAGHER, S. *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press, 2018.

DI PAOLO, E. A.; CUFFARI, E. C. & DE JAEGHER, H. *Linguistic Bodies. The Continuity between Life and Language*. Cambridge: MIT Press, 2018.

DI PAOLO, E. A. Enactive becoming. *Phenomenology and the Cognitive Sciences*, 1–27, 2020. https://doi.org/10.1007/s11097-019-09654 -1

DI PAOLO, E., ROHDE, M. & DE JAEGHER. Horizons for the Enactive Mind: Values, Social Interaction, and Play. In: Stewart, J., Gapenne, O. e Di Paolo, E. *Enaction – Toward a New Paradigm for Cognitive Science*. Cambridge: MIT Press, 2010.

DI PAOLO, E. e THOMPSON, E. The enactive approach. In: SHAPIRO, L. *The Routledge Handbook of Embodied Cognition*. London: Routledge Taylor & Francis Group, pp. 86-96, 2014.

DREYFUS, H. What Computers Still Can't Do. MIT Press: New York, NY, USA: 1992.

DREYFUS, H. Skillful Coping – *Essays on the phenomenology of everyday perception and action.* Oxford: Oxford University Press, 2016.

GIBSON, J. *The Ecological Approach to Visual Perception*. New York: Psychology Press, 1986.

JOHNSON, M. *The meaning of the body*: *aesthetics of human understanding*. Chicago & London: The University of Chicago Press, 2007.

JOHNSON, M. *Embodied mind, meaning and reason - how our bodies give rise to understanding*. Chicago: University of Chicago Press, 2017.

JOHNSON, M. *The aesthetics of meaning and thought – the bodily roots of Philosophy, Science, Morality and Art.* Chicago: University of Chicago Press, 2018.

HOLMES, W., BIALIK, M. e FADEL, C. *Artificial Intelligence in Education – Promises and Implications for Teaching and Learning*. Boston: The Center for Curriculum Redesign, 2019.

HURLEY, S. Perception And Action: Alternative Views. *Synthese* 129, 3–40, 2001. https://doi.org/10.1023/A:1012643006930

MATURANA, H. e VARELA, F. *Autopoiesis and Cognition*. Boston Studies in the Philosophy and History of Science, vol. 42. Springer, Dordercht, 1980.

MATURANA, H. e VARELA, F. *A Árvore do Conhecimento – As bases biológicas da compreensão humana*. São Paulo: Palas Athena, 2019.

O'NEIL, C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York: Crown, 2016.

ROHDE, M. *Enaction, Embodiment, Evolutionary Robotics*. Paris: Atlantis Press, 2010.

SHEETS-JOHNSTONE, M. Thinking in Movement: Further Analyses and Validations. In: *Enaction – Toward a New Paradigm for Cognitive Science*. Cambridge: MIT Press, 2010.

STEWART, J., GAPENNE, O. e DI PAOLO, E. *Enaction – Toward a New Paradigm for Cognitive Science*. Cambridge: MIT Press, 2010.

THOMPSON, E. *Mind in Life. Biology, Phenomenology, and the Sciences of Mind*. Cambridge: Harvard University Press, 2007.

VARELA, F., THOMPSON, E. and ROSCH, E. *The Embodied Mind – Cognitive Science and Human Experience*.

VÉLIZ, C. Moral Zombies: Why Algorithms Are Not Moral Agents. *AI and Society* 36 (2): 487–97, 2021. https://doi.org/10.1007/s00146-021-01189-x.

ZAHAVI, D. A fenomenologia e o(s) problema(s) da intersubjetividade. In Bannell, R. I., Mizrahi, M., Martins dos Santos Ferreira, G. (Orgs.) *(Des)educando a educação: Mentes, Materialidades e Metáforas*. Rio de Janeiro: PUC-Rio, 2021.

ZEBROWSKI, R. L., e McGRAW, E. Autonomy and Openness in Human and Machine Systems: Participatory Sense-Making and Artificial Minds. Journal of Artificial Intelligence and Consciousness 08 (02): 303–23, 2021. https://doi.org/10.1142/s2705078521500181.