AI-informed acting: an Arendtian perspective

Daniil Koloskov[1]

Abstract

In this paper, I will investigate the possible impact of weak artificial intelligence (more specifically, I will concentrate on deep learning) on human capability of action. For this goal, I will first address Arendt's philosophy of action, which seeks to emphasize the distinguishing elements of action that set it apart from other forms of human activity. According to Arendt, action should be conceived as *praxis*, an activity that has its goal in its own very performance. The authentic meaning of action includes the "passion" for articulation of one's own individuality; I can only manifest myself as a distinct personality insofar as I introduce myself as a novel beginning to the web of human interactions demonstrating both my relevance and distinction from others. From this Arendtian standpoint, I will analyse the impact of deep learning in modern AI from two possible angles. First, I will argue that *the direct interaction* between AI and action is impossible. Since AI operates on the principle of efficiency, it can neither suggest certain goals for action for us nor overtake their implementation because action is not guided by the instrumental need to be efficient but by the existential desire to be someone. Second, I will also analyse the possibility of *the indirect impact* of AI on action. More specifically, I analyse neural network's ability to circulate actions among individuals based on mathematical calculation. As I will argue, the efficiency of this circulation that surpasses human cognitive capacities can potentially organize a broader network of interaction among individuals and serve as a catalyst for the ability to act.

Keywords Arendt · Action · AI · Neural networks · Publicity · Transcendental philosophy

✉ Daniil Koloskov
    daniilkoloskov94@gmail.com

1 Philosophical Faculty, University of Hradec Králové, Hradec Králové, Czechia

Springer

1 Introduction

The rapid development of neural networks over the past ten years has already brought about so many changes in our daily lives that it can be classified as the beginning of a revolution by any measure. Yet it is also clear that this revolution is still at its dawn; we are still to find out what its consequences would mean for us. It remains to be seen, for example, what this development of AI would mean for the organization of our labour and distribution of wealth. We still have a lot to figure out about the way it can participate in our cultural life, medical services or other areas of human existence. But even despite being unclear and, to a large extent, unpredictable, the neural network revolution already poses a great need to reflect upon its significance. One such area, of course, concerns ethical dimension. Can our reliance on neural networks be considered ethically acceptable? In what cases? Should we even continue exploring this area? Another area is epistemic: how exactly can we trust the conclusions generated by them? How should we reconsider our ideas about justification in light of the development of AI? Now, one area that I want to explore in this paper has received less scholarly attention; it remains, nonetheless every bit as pressing as others. Namely, I would like to explore the consequences that our increasing reliance on AI might have on our capacity to act.

For this goal, I will turn to Arendt's phenomenology of action. First, this is because the phenomenological approach – and I proceed from the assumption that Arendt's approach is phenomenological – provides a unique angle of investigation of the development of AI in the context of human interaction. Concentrating on the lived experience and its essential structures, phenomenology is capable of grasping and fixating on the living core of the meaning of the process, which to an increasing extent no longer relies on the very quest for meaning and is no longer accessible to human comprehension. Second, Arendt's phenomenological strategy centres around explicating what is irreducibly distinct about action distinguishing it from other elements of human activity. This would make it possible to provide a 'cleaner' analysis of influences that AI usage poses directly to the ability of action.

In the first section, I will briefly introduce the well-discussed notion of action in Arendt's philosophy and present a widespread objection to her account, which will be relevant in the context of discussing the role of AI. In the second section, I will describe how ontological motivation of action, which is done for the sake of being someone, is still incompatible with the functioning of neural networks. Furthermore, in the final chapter, I will argue in favour of indirect interaction between neural networks and the capacity of action. Namely, I will demonstrate that despite the fact that AI cannot motivate or overtake human action, it can be seen as an efficient mediator and facilitator of action that makes it possible to establish a wider network of interactions among individuals.

2 Action in Arendt's philosophy

When it comes to Arendt's conception of action, it is important to hold in mind her methodological background. Investigating our 'active engagement' in the world she introduces the *structural distinction* among different types of our *vita activa* – labour, work and action. The point of this distinction does not consist of outlining three isolated activities but to introduce partitioning *into* our active life in order to emphasize its different elements, the active principles that are at work in our everyday dealings with the world. We should be careful enough not to confuse this essentially transcendental analysis (see also Braun, 1994 and Loidolt, 2018 for a more detailed discussion of this strategy), which investigates the organization and essential principles of our practical life with an empirical analysis, which would classify different activities into work, labour and action. Her distinction is not an empirical distinction among several entities; it is more like a *border*, the value of which is measured according to its capacity to emphasize rather than isolate. All modalities of *vita activa* presuppose one another and make sense only in juxtaposition to one another, a juxtaposition that presupposes the holistic context of the human condition. As Loidolt puts it, those distinctions are not meant to designate particular practices "that we can refrain from – we can only refrain from particular activities" (Loidolt, 2018, 114). So, although Arendt gives a wide range of empirical examples when she explains her distinctions, we should remember that those are examples that simply *better* manifest some principle of active life.

The distinction itself is relatively straightforward. Labour concerns the activity of life maintenance. Every organic element has to continuously reproduce itself, search for sources of energy, consume and digest it, maintain a certain state of physical health etc. Labour is connected to the *necessities of life* that are imposed on us by the very need of staying alive and preserving its normal flow. As Arendt puts it "the labor of our body which is necessitated by its needs is slavish" (Arendt, 1958, 83). The most crucial aspect of the activity of labour concerns its cyclicity: the cycle of labour leaves nothing behind other than the need to start over. The results of labour are the product of consumption widely conceived – they disappear in the course of their use leaving us with the need of repeating the cycle. The self-destruction character of labour means that it is "worldless" (Arendt, 1958, 115) as it does not establish any kind of 'transcendence' broadly conceived, i.e., anything beyond itself; while the labouring "human body … is also thrown back on itself, concentrates on nothing but its own being alive and remains imprisoned in its metabolism with nature without ever transcending or freeing itself from the recurring cycle of its own functioning" (ibid.).

Then there is work. Unlike labour, work is characterized by a certain degree of stability. If labour consumes everything it gives rise to without a trace, work culminates in the production of *things* that transcend the cycles of consumption. The products of work "guarantee the permanence and durability without which a world would not be possible at all" (ibid., 94): their stable presence structures the behaviour of individuals and creates the stable patterns of their interactions, which makes it possible for individuals to reckon with others on a permanent basis. Work, in other words, brings individuals together by opening up the shared space to which newborns arrive. It is "a

basic world-access" (Loidolt, 2018, 115) as it opens up what Heidegger has described as 'wordliness' of the world, i.e., the stable net of references that assigns places and functions to entities. Things are, thus, not only "parts of the world" (Arendt, 1958, 182) but also its hinges, which indicate the constant being-with-others. As Voice sums up nicely, "the presence of the other and the world both requires and helps establish an objective reality that is distinct from the subjective mind of the individual for whom the other and the world would otherwise be indistinguishable from mere subjective representations" (Hayden 2014, 40). So, while labour is concerned with cycles of life where an end of a cycle also flows into the beginning of a new one, the work opens up for itself the domain of the future. It posits something like a planned end and subordinates particular activities with regard to this end; work relies on the *instrumental organization* of activity. While remaining tightly bound to the situation of labour and its needs, the work inscribes a certain stability in it by learning how to use the materiality and contingency of the world for its goals. In this sense, speaking empirically, life and work interpenetrate rather than represent two isolated activities. (Braun, 1994, 29f; 36).

The third modality of human activity should be thought of in the same way. Action does not refer us, in the first place, to a specific practice. Rather it targets an activity, which is constitutive of human reality as we know it and which can be outlined in juxtaposition to other constitutive activities of work and labour. Action accounts for an essential element of human reality that they leave unaccounted for. Work and labour only characterize us as "earth-bound creatures" (Arendt, 1958, 11); they "have the task to provide and preserve the world for, to foresee and reckon with, the constant influx of newcomers who are born into the world as strangers" (Ibid., 9). Both of them either follow directly from our condition (as a set of biological imperatives that we happen to have) or from our adaption to this condition (expressed in products of work that we have inscribed into the cyclicality of life). Labour and work can only explicate *what* we are as following from this condition; they address what is *generalizable* about us, our "qualities, gifts, talents, and shortcomings" (Arendt, 1958, 179). The specifics of Arendtian approach, however, consists of emphasizing that the question of *what* we are cannot exhaust our way of being because our conditions "never condition us absolutely". The question of what we are must be complemented with the question of *who* we are, which would address the irreducible individuality of our perspective that is always something *more* than just its conditions. There is more to us than just having needs and mastering them (the "necessary and useful"); there also is an "authentic human" way of experiencing them and reacting *to* them from the standpoint of irreducible individuality that we are. The 'who' of our existence *shines through* our condition.

Action addresses exactly this individual aspect of human existence. Fusing the transcendental and existential explanatory strategies, Arendt stresses an ontological motivation of action. By its very definition, human existence is a bearer of radical novelty and irreducible individuality (which is why Arendt endows great philosophical significance to the fact of our natality); Arendt translates this ontological *fact* into a driving element of human experience claiming that one of the fundamental elements of *vita activa* consists of the objectively unexplainable 'desire' (which Honig describes more precisely as "the agonal passion for distinction" (Honig 1995, 159) to

express and "articulate" this individuality, i.e., to articulate one's own character of a new beginning. In its authentic sense, action always concerns the articulation of one's own position in the world; it seeks to *actualize* an individual's objective difference from everything else into an articulated distinction that would *manifest* this originality in the world, something he would become actively aware of. Arendt describes this articulation as a "second birth, in which we confirm and take upon ourselves the naked fact of our original physical appearance" (ibid., 176). Articulating the irreducible individuality, action manifests the *who* of an individual rather than his *what*. It demonstrates that despite all possible determinations that we might have, our own originality can only be located by pointing to the 'bearer' of all possible what-like determinations without being exhausted by them. Without taking into account the authentic desire for self-expression, we would fail to address human reality as distinctly human (Arendt, 1958, § 24).

Thus, action loosens its emphasis on what is done and becomes also preoccupied with "*how* it appears in the world; *how* it unfolds in special circumstances; *how* it actively responds to its conditions" (Loidolt, 2018, 121) because it is this *how* that manifests the surplus of objectivity that we are. In this sense, action is akin to a statement: instead of being concentrated exclusively on this or that state of affairs, it uses the latter as an *occasion for self-articulation*. When acting, I am going in two directions at the same time: I simultaneously act in the world making something concrete and expect that this concrete action will help me to articulate myself as a new beginning. The objective world becomes a stage where I announce myself demonstrating the *originality* of my perspective. For this reason, action is not a pragmatic category (if we conceive pragmatic in a narrow way): it is not pursued for the sake of some external goal but for the sake of this self-articulation. This "self-sufficient" or "self-contained" character (Villa, 1996, 21) of action makes it possible for Arendt to characterize it as *praxis* in an Aristotelian sense of the word, an activity that contains its goal within itself. The meaning of action consists of performing it and appearing as a particular who in the course of this performance; "a lived experience of action is a lived experience of performance and appearance" (Loidolt, 2018, 200).

As such a performance, action always addresses others through speech. In order to appear as this unique who, I require others to *whom* I would appear. My character of an irreducibly new beginning can only be expressed in *irreducible terms*, i.e., in terms of other irreducibilities who have already articulated themselves; it must gain a foothold in their perspectives demonstrating how my beginning is relevant for them. This can only happen thanks to language. Language, as already Saussure has pointed out, has precisely this nature of making an *internal difference* where the meaning of a particular sign is determined by its difference from other signs. In the same way, by individuating myself through speech I am able to articulate and present my distinctiveness from others, which does not isolate me from them but rather connects us internally showing how their perspectives are complemented by mine and mine by them through our difference. Manifesting myself as a new beginning, I enrich their perspectives. This is a kind of distinction that creates a connection between the distinguished elements. To present myself as a unique perspective, my action must resonate with others, must tackle their stories and individuate me in *their* terms. My action, so to say, spreads in waves among other individuals influencing their perspectives

and progressively individuating them to the same extent it individuates me. Action, in this sense, is always action-provoking: since to act means always to "act[] upon beings who are capable of their own actions, reaction, apart from being a response, is always a new action that strikes out on its own and affects others" (Arendt, 1958, 190). Thus, Arendt concludes that action can "never move in a closed circle" always going beyond what has been intended by it and introducing new individuals into the zone of its impact.

But just as action does not really belong to an actor, neither does it constitute ownership of those who have been impacted by it. Action is not addressed to someone in particular; it is essentially different from dialogue, which aims at mutual understanding. Rather, action belongs to what Arendt describes as the "worldly in-between" (ibid., 182), the context where it unfolds. It addresses the space that is built up among individuals where my interests intersect other interests. The "web of human relations" is irreducible both to the first-person perspective of an actor and to the second-person perspective of the witness of action because it methodologically privileges the very appearance of action over both its issuer and recipient (Loidolt, 2018, 200). This space distinguishes individuals exactly to the extent it divides them; being as different from norms of "strategic rationality" as it is from "ideal speech conditions" (see Canovan, 1983 for a comparison between Habermas's and Arendt's approaches), it is more akin to layered tiles that connects individuals but preserves this chaotic, spontaneous flow, which is impossible to predict and which never reaches complete harmony. By acting I insert myself into this intersection of interests of particular individuals and, so to say, carve a new path into this constellation of interests forcing others to react and recalculate their own trajectories. Thus, Arendt stresses that the web of human interaction has to be conceived before particular actions as it appears as a condition of the possibility of any action (Arendt, 1958, 184).

The important aspect of Arendt's approach consists of her emphasis on the fact that action is "the only activity that goes on directly between men without the intermediary of things or matter" (ibid., 7). We should remember what this lack of intermediary means. Action is not isolated from work and labour as a separate activity. Were it so, it would eventually lack any content and become unintelligible. Arendt's point is merely that while acting I always, *at the same time*, do *more* than just pay attention to everyday business as I insert myself into the constellation of agencies; [i]t is thus not the intentions of the agent that generate the action's meaning, but rather how the appearing deed and the doer connect to the whole of the web (Loidolt, 2018, 201). Her distinction among three types of activity makes it possible for her to thematize the existential and phenomenological aspect of acting, the simple and almost self-obvious fact that resources and economic interests *will not bring themselves into motion by themselves*. In order to obtain their meaning and their motivating power they need to be taken up by individuals; only in the context of *who* we are do things become meaningful enough to solicit us to commit an action. The public world is this world of *phenomenalizaiton* where objective influences are appropriated and become motives and obstacles *for us*. Action, in this sense, cannot be further from production because intention rather initiates than completes action; action is "boundless" (Arendt, 1958, 190) having an inexhaustible number of consequences, which are *not meant* to reach any state of completion. An actor is always a "doer and a sufferer" because by exposing himself into the open, he lets others define the meaning of his undertakings.

This also means that "action is the political activity par excellence" (ibid., 9). As an ability to establish a new beginning, which goes "against the overwhelming odds of statistical laws"(ibid., 178), action shouldn't be measured according to the ability to oversee the implementation of its goals but rather by the ability to motivate others to *take up* this beginning and "act in concert", i.e. to *act back* and *articulate back* their difference in the context of this new beginning. By distinguishing *and* connecting ourselves through action we come to realize that our differences are that of our respective perspectives on the same world. This insight stays at the foundation of Arendt's approach to power. According to her, not only does power in its primordial sense not rely on violence; it is even opposed to it. Power's fundamental constituent is that of speech: as she puts it, "[p]ower is actualized only where word and deed have not parted company" …. "where words are not used to veil intentions but to disclose realities, and deeds are not used to violate and destroy but to establish relations and create new realities" (ibid., 200). Initiating beginnings and articulating them makes it possible to disclose these new realities by endowing new meaning on objectivity in the context of *who we are*; doing so goes hand in hand with transforming individuals into a group, which is bound together by the interactivity of their actions.

Before going into the next section, I want to recall one objection against this approach, which will help us to clarify the potential role of AI. A number of authors take the desire for distinction as something that eventually "instrumentalizes action and judgment" and "comes at the expense of the initiatory or performative dimension" (Villa, 1996, 70). From this perspective, agents use everyday matters only *in order to* distinguish themselves; the matters themselves remain irrelevant. As Pitkin wittily (although misguidedly) notes, it would model political actors as "posturing little boys clamoring for attention" (Pitkin, 1981, 338). Benhabib sees this problem as following from the "essentialism" of Arendt's account (Benhabib, 2003, 125). She interprets the desire for distinction as this inner essence of human beings, something interior that operates as a preceding cause and further explains why we take up certain challenges as calling for action. Contrary to this objection, we need to stress that the desire to individualize ourselves is not seen by Arendt as an inner essence. Following Heideggerian guidelines, she consistently denies any possible fixed essence that we can ascribe to human beings. The desire for individualization is rather seen as an *ecstatic movement towards the world* in the context of which human beings can acquire any possible determinations. Otherwise, this desire would itself remain a universalistic, non-individual characteristic of humankind, which wouldn't match the radicality of Arendt's enterprise to think of novelty as a constitutive feature of human existence. The desire to individualize ourselves is not an object in us but our way of relating ourselves to the world. Thus, individualization can only take place *in* the world, by taking up some real possibilities of individualization (i.e., novel beginnings) that the world offers.

3 Irreducible action: AI and existential matters

Before proceeding to the analysis of the role of AI in coordinating action, let me first specify that when talking about AI, I mean specifically the so-called deep neural networks (henceforth DNNs), which have managed recently to bring the AI capacities to a new level Those algorithms, which are still qualified as 'weak AI', i.e., which are designed to deal with specific, narrow set of tasks (the notion 'strong AI' is thus left outside of the scope of the current paper), have managed to revolutionize the field because of a new way of machine learning that is called 'deep learning'. First, the operator 'feeds' the DNN data or input and specifies the conditions of success ("cost efficiency") or the output, i.e., what counts as labelled data (e.g. a chair, a car etc.). The principle of breaking down the information, however, is set to be calculated by the network itself. The operator specifies two main parameters for these calculations: the number of layers and the number of "nodes" in each layer. Here, a node is a place where computation is performed: it fires when it finds sufficient stimuli. A layer is a number of organized modes that process the information from the previous layer and pass it to the next one (other than the input and output layers, deep neural networks also have several hidden layers that mediate the procession of the information). Programmed to maximize its efficiency in meeting the success conditions, DNN itself qualifies certain patterns of information as a node. The specific feature of DNNs is that some of those patterns of information are no longer comprehensible to us: we cannot possibly understand what a certain pattern of information has to do with increasing the success rate of the DNN but it remains a fact that they manage to organize data even more efficiently than humans. Thus, the work of DNNs appears to be a 'black box' for its operators who can only orient at the efficiency of the DNN without understanding how it arrives at its 'conclusions.'

The question now is how the efficiency of DNNs that goes beyond human comprehension can be related to human action. One idea would be that AI could have an impact on our ability to pick certain goals suggesting that some are not very reasonable and will not help us to articulate ourselves as a new beginning, while other goals have a better chance of doing so. This could have been a possible scenario if authors like Benhabib were right in criticizing Arendt for essentialism. That is to say, if realizing particular possibilities of individuation is an external goal for our desire for individuation, then AI would be efficient in offering us tools for reaching it. But we have seen that individuality of human existence is not explainable in terms of *what-like* determinations (as a rule, human beings do not search for their objective differentiating feature); it is explainable in terms of *who* we are, i.e., in terms of a certain activity (a certain 'how') of our appearance to others for which we ultimately remain responsible. Picking up and articulating certain intentions to act is not something that we do for the sake of individualization; *it is the very process of individualization*. Benhabib and like-minded authors miss this very *simultaneous character* of 'how' and 'what', which eventually forces them to treat the 'how' of our existence as a more fundamental, preceding 'what'. Recognizing it, along with the ecstatic and processual character of the desire for distinction, we would also have to admit that the task of articulating one's own distinction goes hand in hand with picking up certain challenges and tasks. There is no delay between the desire to be individual and for-

mulating certain goals as a means of realizing this desire; we individualize ourselves by formulating intentions and we formulate intentions by individualizing ourselves. From this perspective, it is nonsensical to assume that DNN can somehow advise us in picking up our goals because it would mean we want DNN to individualize itself instead of individualizing ourselves. This would mean, in other words, that we have delegated AI the ontological task of being who we are.

But neither can it formulate *the ways of reaching this goal* for us. The operation of neural networks is behaviouristic in the classic sense of the word: we feed the machine data, formulate the conditions of success and the network suggests the most *efficient* way of reaching this goal, the precise mechanism of which frequently goes beyond the level of our comprehension. With action, this kind of solution does not work because action is not guided by the principle of efficiency. Efficiency is a category that belongs to the domain of work where the goal is external to the activity itself. Action, on the contrary, is characterized as *praxis* having its own goal in itself, in its own realization. *What* is done in action is not separable from *how* it's done exactly because it is this *how* that emphasizes the ontological constituent of action. While acting I present myself as a *particular doer* with a certain style that manifests my non-substantial individuality or my *who*. Thus, the how of action cannot be overtaken by AI without losing its character of action. Indeed, significant actions are frequently inefficient bringing no immediate gains but are often linked to quite immediate risks; judged by the category of efficiency, most actions would be qualified as unreasonable. But they are not instrumental activities. They are *living gestures* that present me to others, which cannot be judged by their efficiency without losing their character of action. Again, AI is fundamentally incapable of grasping this ecstatic, for-the-sake-of-itself character of action. AI's ability to mimic and supplement the content of action is of no use to the actors themselves for whom their own personality is at stake.

Even if we assume that AI can indeed intervene at any of those stages, we would have to acknowledge that its efficiency endangers the very possibility of action. Arendt systematically claims that action rarely reaches its original intention. Being introduced into the immensely complex web of interests of equal standing, an action is often faced with an impossible challenge. The narrowly conceived instrumental success of action, however, cannot measure adequately its true significance – it is the ability to affect other perspectives and its sheer impact on the public world that makes actions memorable (in this sense, Arendt also stresses that an 'evil deed' can also be qualified as an authentic action (Arendt, 1958, 206). The action accumulates its weight by motivating others to respond. Having obtained those miraculous tools of increasing efficiency, however, we might see that this efficiency can even endanger the possibility of action in the long run. Earlier some goals were unachievable because of the complexity of the public world. Rendering them achievable because of AI intervention necessarily leads to the simplification of the public world as it would reduce others' capacity to act back and offer resistance to action, which is what eventually individualizes it and makes it memorable. The increased efficiency of action, in other words, undermines our status as the *sufferers* of our actions. The very same thing that makes action inefficient renders it polyphonic and entrenches it

in the constellation of other interests by situating it against unsurpassable limitations posed by others.

Actions receive their complexity, nuanced and witty character, not as a result of sheer cognitive effort. It is not because actors are concerned about the efficiency of their actions that they have to think their actions through. Rather, it is because action requires others as those "whose company we may wish to join" (Arendt, 1958, 177): in order to announce myself as a new beginning, I need to introduce myself into the web of interactions among individuals, which is *already* complex and differentiated. Finding a place here and expressing my irreducibility in its complex terms requires an action of *comparable* complexity. The complexity of action, in such a way, is a result of a progressive act of situating ourselves in the field of human interactionIts efficiency is always constrained by a certain level and can not surpass the level when it would endanger the balance of the web of interaction, which places individual actors in comparably equal positions. God (ibid., 202) or anyone whose capabilities are expressively lower than that of others cannot be inserted in the field of human interaction. Here, efficiency is only a *procedural means*: an acting individual cannot be much less efficient than others but neither can it be much more efficient.

So, if we take action alone and its specific motivation, there is little point of contact between it and the development of AI. Acting individuals have no reason to rely on AI for advice because action is an existential matter. Thus, we could postulate the fundamental antagonism between AI and action: faced with the increased efficiency of AI, action might find itself faced with the explicit need to be less efficient than possible in order to be affective enough, in order to leave this 'in-between' alive and leave other the very possibility to respond.

4 Dependent action: alleviating the cognitive challenge

Demonstrating this fundamental antagonism between action and AI, we should nonetheless hold in memory the particular methodological setting where this demonstration takes place. As we have seen above, Arendt's method consists of introducing the active structuration *within* our active life. She seeks to emphasize certain fundamental principles of the active life by distinguishing them against one another. Thus, we have only demonstrated so far that there cannot be *direct interaction* between AI and action, which is conceived as *one* of the principles of active life. This, however, also leaves us with the question of whether it is possible to identify some *indirect impact* that AI might have on our ability to act, which is conceived in the *holistic context* of active life.

To clarify this point, we should also recall that even though one of Arendt's primary goals is to stress the irreducibility of action to other components of active life, she acknowledges (without elaborating on it in detail) that the web of human interaction remains "bound to the objective world" (Arendt, 1958, 183). The spontaneity of action does not occur *ex nihilo*: "the matters of the world of things in which men move" might not be able to explain sufficiently the occurrence of the public world of action but they do create a *necessary* condition for the deployment of the common world by offering certain *material predispositions* or, as she puts it, "material prereq-

uisites" (ibid., 201) that can be further taken up in free and spontaneous action. The goals and contradictions of work and labour, "the objective, worldly interests" like our needs in food or security are *appropriated* by acting individuals who use them as a foothold in order to articulate their unique perspective and announce themselves as a new beginning. The web of human interaction, thus, is not a formal interaction among people that is deprived of any 'material' content. Rather, it consists of constant surpassing of such content by appropriating it through the web of intersubjective relations. But this also means that the appropriation through action, which "discloses new realities" and redefines the meaning of objectivity, remains reliant on its condition in one way or another. As we are about to see, the impact of AI on coordinating actions can be located exactly on this level of interconnectedness between the material world and the public world of action.

For this goal, let me resort to a hypothetical situation. Imagine a coal mine with a particular division of labour where most people are working in a condition close to slavery. The exploited work from the dust to dawn, receive no reward other than minimal proviant facing severe health hazards. Then there are exploiters who obviously attain more advantageous positions, secure all the profits and threaten the exploited with violence. The cooperation among individuals here mostly follows from the objective needs of the process itself: the production of coal requires slave labour and an extreme level of coercion. This enterprise does not include the element of action for the most part; the actors here participate in a complex activity, which is nonetheless explained in terms of the 'objective' material needs. At this stage, exploiters are "private individuals" that merely have "interests in common" (ibid., 35). Furthermore, imagine also several other neighbouring enterprises with slightly different working conditions (e.g., a farm or even an office), which nonetheless preserve the coercive and generally non-political character of interaction among individuals.

Potentially, the exploiters can articulate their own position through action. Let's say, one of the exploiters might brutally suppress the resistance of exploited and articulate his action by claiming for himself the status of an apex predator holding that the exploited should be equated to prey. Through this action, the exploiter would appeal to those whom he considers to be of equal standing: predators, such as they are, are justified to commit acts of violence not because they are forced to but because that's who they are. By doing so, he would call other exploiters to express themselves in similar terms, thus, soliciting them to articulate the *sharedness* of their situation. If he succeeds in appealing to others, the "objective world" of work and necessities of life would be transformed into the public world of action, where individuals act out of their own agency as persons. Their particular identities or *whos* would no longer correspond to anything objective but articulate this sheer surplus of being, which every human existence brings to the world. It is the objective world that attains a new meaning based on the performative disclosure of their *whos*. Creating this space of interaction, they would become able to exercise their domination with much greater efficiency and flexibility, which would likely predetermine the occasional but disparate resistance of the exploited to failure.

Of course, the same possibility is open to the exploited. Their resistance does not necessarily have to be reduced to *mute* resistance against suffering. Someone, for example, might refuse to be exploited by denying his status as a prey and his condi-

tion of suffering as a natural order. He might attempt to protest against his situation of being exploited because the exploitation is *unjust* and goes below any possible human dignity. Thus, he would step up saying that he would be better dead than to live the life of a slave. By doing so, the protesting individual would posit himself as a person, as a particular *who* that is worthy of being recognized; at the same time, this positing himself would appeal to other exploited soliciting them to articulate their own objective position of suffering as well. In case this appeal is successful, the exploited would be able to articulate the sharedness of their situation where their own status of exploited *becomes unjust* and *unbearable*, which would unleash a certain potential of acting together and weave them into the web of human relations.

Both options represent the spontaneous occurrence of action *that does not have to* arrive objectively speaking but that *can* arrive nonetheless and reformate completely the landscape of human interaction. As we know, however, it took far more time for the second possible action to occur. For the most part of history, the actions undertaken by the exploited failed to accumulate the critical mass – they were unable to mobilize the sufficiently broad group of exploited and end up gaining only local improvements or are just sadly and brutally suppressed. Even after the development of labour power, which has made alternative forms of organization of labour conceivable, the action of a coal miner addressed mostly coal miners. To be more historically concrete, we could think of Engels's analysis of German's Peasant's War where peasants despite having all the necessary objective resources for success have been nonetheless suppressed by the nobility because peasants were unable to organize themselves in a single web of interaction and find a way of expressing their shared goals and demands. Despite the fact that retrospectively we can easily see a certain similarity in the *objective situation* of exploited, their common suffering and common interests, we also have to acknowledge that exploited have failed to *express* their interests and the originality of their perspectives sufficiently radically. The occasional actions of the exploited relied too much on the language of their exploiters; as a result, the exploited were unable to see other local protests and acts of insubordination as potentially relevant for them, which constrained the interactive potential of their actions.

This is, of course, a typically Arendtian point to make. We must acknowledge that action happens where we did not expect it and does not happen where we expect it exactly because action is a manifestation of human spontaneity and creativity that is further unexplainable. As a new beginning, every human being brings with himself the unpredictable possibility, which cannot be deduced from the materiality of the world. Arendt is adamant in stressing that we should not be seduced into treating this originality as a "façade" (ibid., 183) of more fundamental processes that can be sooner or later exhaustively explained in terms of the objective dynamics of the development of labour power. The irreducibility of action is, thus, *best demonstratable* based on examples of when it goes against the obviousness of the objective world.

There is, however, also a subtler point involved. Namely, we could still acknowledge that despite being irreducible to their objective conditions, some actions are more likely to occur and some less. It was *harder* for the exploited to articulate their solidarity and transform their objective condition of work and labour in the shared

public world of action. This is because they were faced with a *wider challenge* of *expressing a more encompassing objective situation* that can potentially include many more individuals in the public world of action. The level of generality and abstractness of this objective condition (as well as the "physical" gap among the respective individual positions in the objective world) was so high that it required an action of the most outstanding magnitude, so even the most marginalized and suppressed individuals would recognize that it potentially affects them as well. This outstanding action, which could be compared in its magnitude with Jesus's revelation (whom Arendt frequently treats as a political actor (see, for example, Arendt, 1958, 248; Dolan, 2020), launches the long history of inter-actions that eventually brings us to recognition of the universal human dignity along with the corresponding attempt to build something like a community of universal justice. So, while the solidarity among exploiters relies on a much more immediate objective predisposition, we have required generations of technological, scientific and moral development in order to formulate the possibility of acting in the name of the shared solidarity of sufferers. In this sense, we should also acknowledge that action, while being an existential act in the first place, also includes a *cognitive aspect* since it relies on the objective world that supplies predispositions for action. More specifically, action presupposes an ability to *recognize* the shared objective situation of individuals and an ability to generalize this objective predisposition understanding to *whom* this action might be addressed.

Acknowledging that articulating the objective predispositions in action has different coefficients of complexity, we open up an interesting way of accounting for the impact of AI. We have recently seen the most striking example of it. A number of news channels and researchers have turned their attention to the role the *TikTok* social network and its algorithms have played in the mobilization of action even in depoliticized countries where the public realm is largely disarrayed (see, for example, Brodovskaya et al., 2022; Boffone, 2021; Hautea et al., 2021; Sofya Glazunova and Malmi Amadoru 2023). To be more specific, we have seen, for example, that the Russian protest movement, which has been largely demolished after the protests in 2012 with most significant actors being killed, imprisoned, ostracized or discredited (for a more detailed overview of Russian democratic protest movement see Gabowitsch, 2017) has suddenly come to life in 2018[1]. A number of anticorruption investigations have brought dozens of thousands of people onto the streets demanding the democratization of the country and the resignation of top officials; they have done so in spite of the brutal measures undertaken by police forces and the real risk of persecution. The very fact of this outburst was surprising enough. For decades Russian society has been suffering from a high level of corruption among the political elite, a problem that has been widely acknowledged by the population and resulted in a systematically

---

[1] In what follows, I will largely concentrate on one particular case of how DNNs mobilize human action taking for example Russian protest movement in 2018. However, I am convinced that this account has wider applicability. For example, the protest movement in Iran in 2022 demonstrated similar patterns when the feministic movement "Hair for Freedom" virally spread in *TikTok* and *Instagram* involving not only teenage females but also various groups of males into the protests (see Navarro & Peres-Neto, 2023; Afary & Anderson, 2023). See also Bossetta (2018) for analysing how DNNs algorithms in Snapchat, Instagram etc. have affected the 2016 US selection.

low level of public confidence. In this sense, a series of anti-corruption investigations did not bring anything new to the table and could be hardly seen as a sufficient cause for the protests. Furthermore, the social makeup of protesters was also unexpected as the actions of protest leaders originally focused on students and young adults as their main target group. The investigations themselves as well as the claims and demands of the leaders were often reasonably complex and presupposed a certain level of education of their recipients. Yet, the most numerically significant and loud part of the protesters turned out to be teenagers many of whom haven't finished high school. The protests have gained their scope because opposition leaders were able to involve those social groups that they were not meaning to involve in the first place. Out of nowhere, Russian high schools have been suddenly moved to the forefront of political life.

Again, by itself, it shouldn't be *too* surprising given that action is surprising according to its very essence. As Arendt stresses many times, the action spirals out of the agent's control the moment it has been thrown in the open. It becomes capable of addressing people it was never meant to address and having consequences that it was never meant to have because its existential meaning is to *let oneself to be weaved* into the web of human interactions and to be defined in *their* terms, from *their* perspectives. The interesting part comes from this implicit cognitive challenge that we have identified above. One of the natural constrainers towards the mobilization of action concerns *its potential scope*: using the objective situation as a foothold, the impact of action also remains limited by the agent's ability to recognize how many individuals might share this foothold with him and, thus, how many individuals might be potentially affected by the action. In a similar way, since action relies on idiosyncratic language and endows new meaning on objectivity, other individuals might easily overlook that this action springs from the shared objective predisposition. So, coal miners might not see the protest of office workers as related to them because office workers protest for paid vacation, while coal mine workers want to alleviate their health hazards even though both of those groups are placed in the position of the exploited. In the same way, suffering from the pettiness, boorishness and tyranny of high schools, Russian teenagers were not able to see how those phenomena represent the general style of political organization of the country and, thus, failed to see how protests against corruption address their situation as well. Actions might fail to reach their critical mass exactly because the distance between an acting individual and those who witness his action is too great and requires some intermediary steps in order to be weaved into a single web of action.

It is here where AI has proved to be impactful. *TikTok's* DNNs aim at learning when a given user prefers certain content and this by breaking down the information about the content and its users according to a set of characteristics, i.e., 'nodes,' those quasi-principles of organization of data. The quasi-principles, which, as we remember, are no longer accessible to human comprehension, have suddenly started recommending political content to teenagers sharing information about protests and actions of protesters. The processing power of DNNs have concentrated on the mathematical analysis of what is generalizable and analysable about their audience: they have made their racial and educational profiles, considered their age group and hobbies, their type of humour, etc. Naturally, DNNs have no means of addressing what

is unique about their perspectives – they can only analyse the calculatable, *objective predispositions* of their irreducible perspectives. Proceeding this way, they were able to detect a certain similarity between the material predispositions of protesters who have articulated their positions into public actions and the material predispositions of a yet politically uninvolved part of Russian society. And having detected it, DNNs have started offering particular content 'assuming' that it might be relevant for such 'yet inactive' individuals.

This situation is different from most common uses of 'weak AI' when algorithms (for example, in medicine or banking) are charged with the task of *making* the most efficient decision. In those areas individuals *delegate* their decision-making and end up privileging the efficiency of the decision over its meaningfulness; in the Arendtian sense, they thus eliminate the element of acting from their activity. Here the 'opacity' of AI's decision-making becomes an issue for individuals who might reach their goals but who cannot understand properly *how* they did it and whether this result is sustainable, ethically acceptable etc. While we shouldn't overlook the dangers for human plurality that are coming from this opacity (see, for example, Rouvroy and Berns 2013, Tufekci, 2018, Ziarek 2022 for investigating such dangers), we should also hold in mind that DNNs relations with 'action as such' should be assessed differently. DNNs are indeed *efficient* in circulating those actions, which they 'assume' to be relevant for others; but Sharing information by itself would not suffice for soliciting action. To use Arendt's words, the mathematical language of DNNs "no longer lend themselves to normal expression in speech and thought" (Arendt, 1958, 3); the results of statistical analysis are just meaningless data that is not compensated by its efficiency and represent no value for action in itself. That I am statistically likely to find certain content relevant does not say anything about whether or not I will actually find it so. Action becomes dependent on the objective world only insofar as it finds a way to use this objectivity as a foothold for the deployment of something that no longer corresponds to any objective state. Therefore, DNN's algorithms can only succeed in motivating my action insofar as I appropriate those statistical regularities and interpret them in the context of this unique perspective that I am.

In this sense, DNNs recommendations only concern this necessary condition indicating the possible *occasion* for action with greater insistency and efficiency than other media; this is because they statistically 'assume' that some other actors have managed to present themselves by relying on the same objective predisposition that I have. By doing so, they *alleviate* the cognitive challenge connected with the need to recognize the similarity in the objective predispositions among potential actors. However, they do not resolve the sufficient condition of action leaving the existential task intact: I still have to make sense of this similarity. Recognizing *what* exactly is shared among us refers me to the task of understanding *how* it becomes possible for me to articulate myself based on this shared predisposition. Correspondingly, there is always a chance that those AI-recommended actions won't appeal to me; even when confronted with information about some shared objective presuppositions, I always have the possibility of safely ignoring it as I would miss how they can be relevant to the uniqueness of my perspective. The troubling opacity of AI, in such a way, does not affect action. Opacity only concerns the infrastructure of action operating on the level of its necessary condition and suggesting individuals to react upon one another;

but if action takes place, it takes "between men without the intermediary of things" translating this opaque, meaningless statistical suggestion into the meaningful language. In the end, it doesn't matter for what opaque reason DNNs have offered me certain content; what matters is that I managed to appropriate this opacity and saw it as relevant for me, as soliciting my action.Not only that. Other than providing this indirect occasion for action, DNNs also help to facilitate *a certain way* of presenting actions, so they would *better appeal* to others. Since DNNs establish the circulation of information based on the efficiency rate of its consumption, they are also sensitive to any increase in content's ability to appeal to others. They momentarily pick up and circulate those 'actions', which are able to better express the convergence in the objective predispositions among actors. Thus, they further reduce the complexity of the cognitive challenge. DNNs reinforce a powerful loop effect, which is typical of action in general: profound actions solicit others to act back and articulate their own perspectives, which further affects an even greater number of individuals. The mathematization of this exchange of actions leads to weaving individuals into the web of interactions at an unseen rate: they become capable of communicating the relevance of actions much faster and to a broader number of people mathematically emphasizing those actions, which would be most efficient in expressing the similarities among potential actors. As in the case of the sudden outburst of protests in Russia, the claims and actions of oppositionists have first been communicated to teenagers via *TikTok* and then motivated the latter to further express these claims and actions in a language that is much closer to their background. Inside this loop where most communicatively efficient expressions gave birth to even more efficient ones, hours-long anticorruption investigations, debates and interviews of opposition leaders, the references to classic authors and the prominent histories of the past have been sliced into short video clips, mixed up with internet culture, memes, local jargon etc. and have been virally spread across the country. As a result, the social basis of the protests has been substantially extended and the web of interactions among individuals became much more extensive than before.[2]

Of course, the protests were successfully suppressed by the government, the consequences of which are too well known nowadays. What they have demonstrated, nonetheless, is that AI's explosive impact on our single most explosive capacity will make the latter even harder to control in future.

---

[2] One concern that might arise in this context is that by connecting people to one another this way, DNNs might not only facilitate human plurality but also endanger it; by bringing individuals together based on statistical likeness, they might as well give voice to fascist or racist people. To avoid this worry, we should recall the inherent riskiness of Arendt's notion of action that doesn't want any safeguards against 'wrong' people entering the public domain but rather relies on its power to counter-act and gather others in this counter-action. Action does not seek to secure any particular regime but throws itself into the open increasing the turbulency of our being-with-one-another being prepared to encounter hostile others as well.The discussion of the possible impact of DNNs on action should be distinguished from parallel usages of Arendt that see.

5 Conclusion: AI as a facilitator of action

We have seen, on the one side, that AI is clearly an antagonistic force when it comes to the phenomenon of action. While AI's main point of reference is the principle of efficiency, the for-the-sake-of-itself character of action points to a different direction. It is, thus, nonsensical to assume that AI can threaten our ability to act. Neither can it overtake acting from us or oversee its implementation. But this doesn't mean that the development of AI is irrelevant to action. AI can also *escalate* our ability to act by appearing as a *mediational element of action*. It eases up the communication of action to others and increases its potential appeal to others, helping it to reach them in a more extensive fashion by *alleviating the cognitive challenge* connected with the desire to act; AI facilitates the very recognition of the relevance of actions and helps to express it in more familiar terms.

In this sense, not only does AI not threaten our ability to act – *it makes action more of itself*. As we have seen, action is constituted exactly by its ability to go beyond what was intended by it, an ability that introduces actors into the web of human relationships where their actions obtain their full significance. By its definition, action is unpredictable: one way or another, I am not able to predict how others would react to my action and who would react to it since action takes place in the public domain, which is the domain of freedom and spontaneous self-presentation. Facilitating the interaction among individuals, AI only makes it more unpredictable, more ready to slip one's own control and go in unintended directions. Without intervening in my action directly, without telling me what to do and how to do it, the AI can effectively restructure the very landscape where the action takes place as it ensures greater communicability of action and enlarges its potential scope. So, instead of taking away an agent's spontaneity, AI only makes it possible to override certain cognitive limitations; reducing the complexity of communication of action and decreasing the number of intermediary steps that are needed in order to circulate it across the community, AI, at least potentially, offers an infrastructure for the shared public world, which will enlarge the possible extent and density of the web of human interaction.

Author contributions The article was written by a single author.

Data availability The data and materials used in the article can be accessed based on the references mentioned in the article.

Declaration

Ethical approval Non-applicable.

Informed consent Non-applicable.

Conflicts of interest There are no competing interests.

References

Afary, J., & Anderson, K. B. (2023). Woman, life, freedom: The origins of the uprising in Iran. *Dissent*, *70*(1), 82–98.

Arendt, H. (1958). *The Human Condition*. University of Chicago Press.

Benhabib, S. (2003). *The reluctant modernism of Hannah Arendt*. Rowman & Littlefield.

Boffone, T. (2021). *Renegades: Digital Dance Cultures from Dubsmash to TikTok*. Oxford University Press

Bossetta, M. (2018). The digital architectures of social media: Comparing political campaigning on Facebook, Twitter, Instagram, and Snapchat in the 2016 U.S. election. *Journalism & Mass Communication Quarterly*, *95*(2), 471–496. https://doi.org/10.1177/1077699018763307.

Braun, M. (1994). *Hannah Arendts Transzendentaler Tätigkeitsbegriff: Systematische Rekonstruktion Ihrer Politischen Philosophie Im Blick auf Jaspers Und Heidegger*. P. Lang.

Brodovskaya, E. V., Lukushin, V. A., & Davydova, M. A. (2022). Mobilization of protest activism on TikTok: Scale, features and threats. In A. V. Chugunov, M. Janssen, I. Khodachek, Y. Misnikov, & D. Trutnev (Eds.), *Electronic governance and open society: Challenges in Eurasia. EGOSE 2021* (Vol. 1529). Springer. Communications in Computer and Information Science.

Canovan, M. (1983). A case of distorted communication: A note on Habermas and Arendt. *Political Theory*, *11*(1), 105–116.

Dolan, F. (2020). Hannah Arendt and the political philosophy of Jesus of Nazareth. August 2020 Conference: Annual Meeting of the Hannah Arendt Circle At: University of California.

Gabowitsch, M. (2017). *Protest in Putin's Russia*. Polity Press.

Glazunova, S., & Amadoru, M. (2023). Anti-regime influentials across platforms: A case study of the free Navalny protests in Russia. *Media and Communication*, *11*(3), 187–202.

Hautea, S., Parks, P., Takahashi, B., & Zeng, J. (2021). Showing they care (or don't): Affective publics and ambivalent climate activism on TikTok. *Social Media + Society, 7*(2). https://doi.org/10.1177/20563051211012344.

Hayden, P. (Ed.). (2014). *Hannah Arendt key concepts* (1st ed.). Routledge.

Honig, B. (1995). Toward an agonistic feminism: Hannah Arendt and the politics of identity. In B. Honig (Ed.), pp. 135–166.

Loidolt, S. (2018). *Phenomenology of plurality*. Routledge.

Navarro, C., & Peres-Neto, L. (2023). Hair for freedom movement in Iran: Interreligious dialogue in social media activism? *Luiz Religions*, *14*(5), 602.

Pitkin, H. F. (1981). Justice: On relating private and public. *Political Theory*, *9*(3).

Rouvroy, A. (2013). B. Thomas (Ed.), Algorithmic governmentality and prospects of Emancipation. *Réseaux* 177 1 163–196 Translated by Libbrecht.

Tufekci, Z. (2018). *Twitter and tear gas: The power and fragility of networked protest*, Connecticut: Yale University Press

Villa, D. R. (1996). *Arendt and Heidegger: The fate of the political*. Princeton University Press.

Ziarek, E. (2022). Against digital worldlessness: Arendt, narrative, and the onto-politics of big data/AI technologies, *Postmodern Culture*, *32*(2).