

The Review of Politics 83 (2021), 555–577.

© The Author(s), 2021. Published by Cambridge University Press on behalf of University of Notre Dame. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use or in order to create a derivative work. doi:10.1017/S0034670521000474

## Old Facts, New Beginnings: Thinking with Arendt about Algorithmic Decision-Making

*Lisa Herzog*

**Abstract:** More and more decisions in our societies are made by algorithms. What are such decisions like, and how do they compare to human decision-making? I contrast central features of algorithmic decision-making with three key elements—plurality, natality, and judgment—of Hannah Arendt’s political thought. In “Arendtian practices,” human beings come together as equals, exchange arguments, and make joint decisions, sometimes bringing something new into the world. With algorithmic decision-making taking over more and more areas of life, opportunities for “Arendtian practices” are under threat. Moreover, there is the danger that algorithms are tasked with decisions for which they are ill-suited. Analyzing the contrast with Arendt’s thinking can be a starting point for delineating realms in which algorithmic decision-making should or should not be welcomed.

Algorithmic decision-making plays an ever-increasing role in our lives. Algorithms recommend books, music, and restaurants to us or allocate our shifts and tasks at work. When we apply for a loan, our application will be run through an algorithmic scoring system.<sup>1</sup> Algorithms support hiring and promotion decisions.<sup>2</sup> They decide about the prioritization of homeless

Lisa Herzog is professor of political philosophy, Faculty of Philosophy, University of Groningen, Oude Boteringestraat 51, 9712 GL Groningen, The Netherlands ([l.m.herzog@rug.nl](mailto:l.m.herzog@rug.nl)).

I thank Dana Mills, Keith Breen, Katya Assaf, and Garrath Williams for written comments, and audiences at the Wissenschaftskolleg, the Universities of Berlin (FU), Navarra, Duisburg, Jerusalem, Groningen, Stirling, and Lucerne, the ECPR Political Theory colloquium, and reviewers and the editor of *Review of Politics* for their valuable questions and comments.

<sup>1</sup>Cathy O’Neill, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (New York: Penguin Books, 2016).

<sup>2</sup>*Ibid.*; Don Peck, “They’re Watching You at Work,” *The Atlantic*, December 2013, 72–84.

people on waiting lists for housing.<sup>3</sup> They are used in predictive policing and in supporting parole decisions.<sup>4</sup> Thus, algorithms have long entered social and political spheres in which basic rights are at stake.

Many commentators have welcomed the advent of algorithmic decision-making: decision-making would be more objective rather than being marred by human biases; it would be more efficient; access to expert decision-making would be democratized. In some cases, algorithms may indeed be a force for good.<sup>5</sup> However, certain real-life applications have brought sobering, and even alarming, results.<sup>6</sup> Critics fear not only the end of privacy, but also that of due process.<sup>7</sup> They warn that algorithms might not live up to the promise to reduce biases and that new forms of discrimination might arise.<sup>8</sup> They also point out that the disadvantages of algorithmic systems may disproportionately hit marginalized groups.<sup>9</sup>

Many such criticisms come from the perspectives of ethics, justice, and the rule of law.<sup>10</sup> In contrast, this article develops a critique from the perspective of intersubjective practices. It shows how looking at algorithmic decision-making through the lens of Hannah Arendt's political thought helps us see some of its central strengths and weaknesses. The concerns thus identified would raise questions about the use of algorithms even if, counterfactually, the challenges of fairness, discrimination, bias, and so forth were all addressed. More specifically, I contrast algorithmic decision-making with egalitarian intersubjective practices of shared decision-making and draw on

<sup>3</sup>Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (New York: St. Martin's, 2017).

<sup>4</sup>Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias: There's Software Used across the Country to Predict Future Criminals," *Pro Publica*, May 23, 2016.

<sup>5</sup>See Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, and Cass R. Stunstein, "Discrimination in the Age of Algorithms" (working paper, February 5, 2019), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3329669](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3329669), accessed March 8, 2021; Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (London: Penguin Books, 2017).

<sup>6</sup>See Eubanks, *Automating Inequality*.

<sup>7</sup>Alex Campolo, Madelyn Sanfilippo, Meredith Whittaker, Kate Crawford, *AI Now Report 2017* (New York: AI Institute New York University, 2017), [https://ainowinstitute.org/AI\\_Now\\_2017\\_Report.pdf](https://ainowinstitute.org/AI_Now_2017_Report.pdf), accessed March 8, 2021; Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: Public Affairs Books, 2019).

<sup>8</sup>See Solon Barocas and Andrew D. Selbst, "Big Data's Disparate Impact," *California Law Review* 104 (2017): 671–732; Pauline T. Kim, "Data-Driven Discrimination at Work," *William & Mary Law Review* 58, no. 3 (2017): 857–936; Lisa Herzog, "Algorithmic Bias and Access to Opportunity," in *Oxford Handbook of Digital Ethics*, ed. Carissa Véliz (Oxford: Oxford University Press, 2021).

<sup>9</sup>O'Neill, *Weapons of Math Destruction*; Eubanks, *Automating Inequality*.

<sup>10</sup>For an overview see Brent Daniel Mittelstadt, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi, "The Ethics of Algorithms: Mapping the Debate," *Big Data & Society* 3, no. 2 (July–Dec. 2016), <https://doi.org/10.1177/2053951716679679>.

Arendt to carve out a number of central features of such practices: the ontological conditions of plurality and natality, and the possibility of judgment that arises from them. To be sure, all three features continue to be controversially discussed by commentators. My aim is not to resolve the tensions between different interpretations; rather, I focus on the aspects that are most relevant for understanding the contrast with algorithmic decision-making.<sup>11</sup>

Arendt developed her ideas against the background of the experience of totalitarianism, but also of the dangers of behaviorism, economism and the rule of bureaucracy, and of the thoughtless use of technologies. Algorithms, it might be said, only reinforce tendencies that are already present in behavioristic and economic forms of thinking, which have prevailed for a long time and which Arendt criticized as the rise of “the social” against “the political.”<sup>12</sup> But the advent of algorithms raises the stakes, and hence brings new urgency to these problems. Margaret Canovan writes that for Arendt the “special danger of modernity. . . was that those who felt the impulse to act tended to look for some kind of irresistible trend to side with, some natural or historical force with which they could throw in their lot.”<sup>13</sup> Today, this force might well be that of algorithms taking over more and more areas of life, making Arendt’s thought all the more relevant.

Three Arendtian concepts are central to my argument: plurality, natality, and judgment. Plurality, the emphasis on individuals encountering each other as *different* individuals rather than as members of a homogeneous species, is at the core of what for her defines politics. Natality, the ability to begin anew, is for her a property of humans that is just as important as the property that philosophers had traditionally emphasized, mortality. Judgment, finally, is a specific way of coming to an evaluation which, while often practiced by individuals, has an irreducible social dimension. These three features come together in what I call “Arendtian practices.” I argue that such practices, while having their original home in the political sphere, can also be found in other spheres of life.<sup>14</sup> In such intersubjective encounters,

<sup>11</sup>Another aspect of algorithms that would deserve attention from an Arendtian perspective is the way in which social media and their algorithmic governance shape the online public sphere. For reasons of scope, I cannot here address this issue.

<sup>12</sup>See Hanna F. Pitkin, *The Attack of the Blob: Hannah Arendt’s Concept of the Social* (Chicago: University of Chicago Press, 1998).

<sup>13</sup>Margaret Canovan, *Hannah Arendt: A Reinterpretation of Her Political Thought* (Cambridge: Cambridge University Press, 1992), 11.

<sup>14</sup>The relevance of her thought for reflecting on current developments justifies such an approach, which has recently been adopted by a number of writers (e.g., Patchen Markell, “Arendt’s Work: On the Architecture of *The Human Condition*,” *College Literature* 38, no. 1 [2011]: 15–44; Steven Klein, “‘Fit to Enter the World’: Hannah Arendt on Politics, Economics, and the Welfare State,” *American Political Science Review* 108, no. 4 [2014]: 856–69) and which reopens the discussion about her account of work (see also Andrea Veltman, “Simone de Beauvoir and Hannah Arendt on Labor,” *Hypatia* 25, no. 1 [2011]: 55–78; and Christopher Holman,

citizens relate to one another as equals and share the experience of acting together, a core experience of a democratic form of life.

Exploring algorithmic decision-making by contrasting it with Arendt's reflections about plurality, natality, and judgment is particularly illuminating because it points us to something distinctively human that might be lost when algorithmic decision-making shrinks the spaces in which "Arendtian practices" might arise, especially for less privileged members of society. This is not only a matter of justice, but also of the conditions of the possibility of democracy.<sup>15</sup> And while other authors, especially in the debate about deliberative democracy, have also emphasized the value and importance of intersubjective exchanges,<sup>16</sup> Arendt is unique in bringing out the dimensions of human plurality and natality, that is, the possibility of new beginnings.

In the next section I provide an overview of what I mean by "algorithmic decision-making." I then summarize Arendt's concepts of plurality, natality, and judgment and explain my notion of "Arendtian practices." The core of my argument is to show that algorithmic decision-making is diametrically opposed to such Arendtian practices. Hence, the question is which areas of life algorithms should be allowed to take over, and how spaces for Arendtian practices can be preserved. The Arendtian concepts of plurality, natality, and judgment provide a starting point for answering these questions. I conclude by describing the implications of this Arendtian perspective for the division of labor between humans and algorithms.

### Algorithmic Decision-Making

An algorithm is "a procedure for solving a mathematical problem . . . in a finite number of steps that frequently involves repetition of an operation; broadly: a step-by-step procedure for solving a problem or accomplishing some end especially by a computer."<sup>17</sup> As such, algorithms are nothing new: they are omnipresent in mathematics, but also in guidebooks for tasks such as cooking a meal. Past experiences about how to do these things are distilled into step-by-step instructions that others can follow. My focus, however, is on algorithms that are implemented in software.

---

"Dialectics and Distinction: Reconsidering Hannah Arendt's Critique of Marx," *Contemporary Political Theory* 10, no. 3 [2011]: 332–53).

<sup>15</sup>I read Arendt as a democratic theorist; the reasons will become clear in section 4.

<sup>16</sup>E.g., Nancy Fraser, "Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy," *Social Text* 25/26 (1990): 56–80; Iris Marion Young, *Inclusion and Democracy* (Oxford: Oxford University Press, 2000); Jane J. Mansbridge et al., "A Systemic Approach to Deliberative Democracy," in *Deliberative Systems: Deliberative Democracy at the Large Scale*, ed. James Parkinson and Jane Mansbridge (Cambridge: Cambridge University Press, 2013), 1–26.

<sup>17</sup>This definition is from Merriam Webster; similar ones are used in the literature.

What differentiates such algorithms is that they need to be formulated in a language that computers can understand. Natural language needs to be translated into formal language.<sup>18</sup> Another crucial difference is that with increasing computing power, computers can go through a large number of steps very quickly, and process amounts of data that human beings could never process (“big data”). Algorithms can be very simple or extremely complex; up to a certain point, however, it is possible to understand, step by step, what they do and how different inputs influence the outputs.<sup>19</sup>

This changes when elements of machine learning are involved. What they have in common is that the steps of the algorithms are not all defined beforehand; rather, the algorithms can find their own solutions for the given problems. They proceed by trial and error and are “taught” which strategies are promising, by receiving feedback on whether they have provided correct solutions. For example, when an image recognition program is “taught” to recognize horses in pictures, the programmers do not know what exactly the program does; they can only observe whether the success rate increases. This is what is usually meant when an algorithmic system is called an “artificial intelligence.”<sup>20</sup> In such cases it can become very difficult, even for the programmers themselves, to comprehend how the algorithms arrive at their results. The algorithms manage, for example, to categorize an email as “spam” or “not-spam,” but they do so in ways that are not following the same logic as human intuition. For example, they classify frequencies of words in texts (the most frequent ones being “our,” “click,” “remov,” “guarantee,” and “visit”), without any attention to the meaning of messages, and yet arriving at a relatively high percentage of correct decisions.<sup>21</sup>

A major challenge for the application of algorithms, and in particular machine learning, is the quality of data. There can be various problems, for example, differential error rates in different categories (such as between men and women who change their names when getting married), which can lead to distortions.<sup>22</sup> And depending on how algorithms are used, the feedback they receive might be distorted. For example, if a program is used for sorting applicants into those who get a job interview and those whose

<sup>18</sup>See Joseph Weizenbaum, *Computer Power and Human Reason* (San Francisco: Freeman, 1976).

<sup>19</sup>For an introduction to algorithms see Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein, *Introduction to Algorithms*, 3rd ed. (Cambridge, MA: MIT Press, 2009). From a philosophical perspective, Brian Christian and Tom Griffiths, *Algorithms to Live By: The Computer Science of Human Decisions* (London: HarperCollins, 2017), provides illuminating discussions.

<sup>20</sup>See Larry Hauser, “Artificial Intelligence,” *The Internet Encyclopedia of Philosophy*, <https://iep.utm.edu/art-inte/>, accessed March 8, 2021.

<sup>21</sup>See Jenna Burrell, “How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms,” *Big Data & Society* 3, no. 1 (Jan.–Jun. 2016): 1–12, <https://doi.org/10.1177/2053951715622512>. The examples are from p. 8.

<sup>22</sup>Kim, “Data-Driven Discrimination at Work,” 886.

applications are immediately rejected, there is no possibility of getting feedback about the rejected candidates.<sup>23</sup> This is why, in order to arrive at meaningful evaluations, one has to consider algorithmic decision-making in the social contexts in which it is used: it is the combination of various social factors *and* the programs that leads to certain outcomes.<sup>24</sup>

### Plurality, Natality, Judgment

To ground the contrast with algorithmic decision-making, in this section I briefly review three core elements of Arendt's political thought: the notions of plurality, natality, and judgment. Plurality, for Arendt, is an ontological condition of human life and a key feature of the political realm. In her distinction between labour, work, and action in *The Human Condition*, Arendt holds that "Human plurality, the basic condition of both action and speech, has the twofold character of equality and distinction."<sup>25</sup> In the public realm of politics, individuals become visible in their individuality, which is always more than the "what" of their different features. They answer the question "Who are you?" by action and speech.<sup>26</sup> This openness and mutual visibility of the public realm stand in stark contrast to the darkness of the private realm in which labor and work take place. Hence, Arendt's rejection of a model of politics that exclusively follows the model of "work" or "production"<sup>27</sup> also implies that politics requires that individuals see and hear each other.

The existence of such a realm is important for Arendt not least because it is the space in which natality, in its political form, has its place. Philosophers have emphasized throughout the centuries that humans are mortal. Arendt emphasizes that they are also "natal": with each human being who is born, something new comes into the world, and human beings have the power to make new beginnings. "The miracle that saves the world, the realm of human affairs, from its normal, 'natural' ruin is ultimately the fact of natality. . . . It is. . . the birth of new men and the new beginning, the action they are capable of by virtue of being born."<sup>28</sup>

Natality implies an openness towards the future that stands in sharp contrast to the ever-sameness of biological processes. Human beings can build new worlds; together, they can take their fate into their own hands. Beginning, "the supreme capacity of man," is crucial for understanding what it means to act in the shared human world—and it is "identical with

<sup>23</sup>O'Neill, *Weapons of Math Destruction*, chap. 6.

<sup>24</sup>E.g., Campolo et al., *AI Now Report 2017*, 36.

<sup>25</sup>Hannah Arendt, *The Human Condition* (Chicago: University of Chicago Press, 1958), 157.

<sup>26</sup>*Ibid.*, 178–79.

<sup>27</sup>*Ibid.*, chap. 31.

<sup>28</sup>*Ibid.*, 247.

man's freedom."<sup>29</sup> She uses the metaphor of a second birth: "With word and deed we insert ourselves into the human world, and this insertion is like a second birth, in which we confirm and take upon ourselves the naked fact of our original physical appearance."<sup>30</sup>

Lastly, there is Arendt's "elusive"<sup>31</sup> notion of judgment, the capacity made possible by the ontological conditions of plurality and natality, which has led to some controversy among commentators. Of interest in the present context is Arendt's emphasis on the social nature of judging, even when done by a single individual: in judging, one evaluates new phenomena against a background of comparable phenomena. It takes into account how other human beings, from their specific perspectives, see these phenomena, in an appeal to "common sense" (*sensus communis*). In her discussion of "Socratic dialogue" in *Thinking and Moral Considerations*, Arendt describes the internal dialogue, the "two in one," that human beings hold when they reflect about their actions. This inner dialogue does not provide us with positive prescriptions what to do.<sup>32</sup> It tells us what *not* to do and undermines "our fixed habits of thought and the accepted rules of conduct."<sup>33</sup>

This creates the space for judgment in the proper sense: a kind of judging that does justice to particulars and yet arrives at intersubjectively valid results. Individuals need to adopt what Arendt, drawing on Kant's *Third Critique*, calls an "enlarged mentality": one that stands in a dialogue with how other individuals would describe a specific phenomenon. For her, "without 'the test of free and open examination,' no thinking and no opinion-formation are possible."<sup>34</sup> When we judge a particular thing (the beauty of a flower, the virtue of an individual, the greatness of a historical moment, etc.), we imagine the perspectives of others and the *sensus communis* about it, which allows us to transcend our own limited perspective.

This is what connects judgment—which might, from what has been said so far, be understood purely as a *moral* practice—to the *political* realm. Judgment is, for Arendt, "one of the fundamental abilities of man as a political being insofar as it enables him to orient himself in the public realm, in the

<sup>29</sup>Hannah Arendt, *The Origins of Totalitarianism* (New York: Harcourt, Brace, 1973), 479.

<sup>30</sup>Arendt, *Human Condition*, 176–77.

<sup>31</sup>Bryan Garsten, "The Elusiveness of Arendtian Judgment," in *Politics in Dark Times: Encounters with Hannah Arendt*, ed. Seyla Benhabib (Cambridge: Cambridge University Press, 2019), 318.

<sup>32</sup>Hannah Arendt, "Thinking and Moral Consideration," in *Responsibility and Judgment*, ed. J. Kohn (New York: Schocken Books, 2003), 187–88.

<sup>33</sup>Maurizio P. d'Entreves, "Hannah Arendt," *The Stanford Encyclopedia of Philosophy*, Fall 2019 ed., ed. Edward N. Zalta, <https://plato.stanford.edu/archives/fall2019/entries/arendt/>, accessed March 8, 2021.

<sup>34</sup>Hannah Arendt, *Lectures on Kant's Political Philosophy*, ed. Ronald Beiner (Chicago: University of Chicago Press, 1992), 40, quoting from Kant's "Reflection on Anthropology."

common world.”<sup>35</sup> As such, judgment is a “democratic world-building practice”:<sup>36</sup> it creates the shared world in which human beings can live together. Some commentators have argued that Arendt seems to shift from a notion of judgment of *actors* to one of *spectators* in *The Life of the Mind*.<sup>37</sup> Spectators are better able to take on a plurality of perspectives, while actors, in the heat of the moment, might lose the detachment that is required to imagine the perspectives of others.<sup>38</sup> Moreover, actors only rarely find themselves in the egalitarian but contestatory scenario that Arendt’s understanding of judgment presupposes.

Nonetheless, the difference between actors and spectators should not be understood as a categorical one.<sup>39</sup> After all, human beings often take on both roles: even those who are busy actors at day can think about their actions at night. Those who seem to be mere spectators can decide to become active when they feel that they can add a specific perspective. Moreover, there is always the possibility of a dialogue between actors and spectators. In fact, Arendt herself mentions the possibility that “actor and spectator become united” in moral judgment.<sup>40</sup> Similarly, spectatorship and action can go hand in hand in what I call “Arendtian practices”: human beings can *both* observe and act, emphasizing the one or the other depending on the occasion. This notion of Arendtian practices provides my point of reference for contrasting human and algorithmic decision-making.

### Arendtian Practices beyond Politics

Plurality, natality, and judgment are related and complement each other in fleshing out a specific vision of how human beings “do politics.” For Arendt, they have their primary place in the political sphere, which she famously contrasts with the private sphere of labor and work. But clearly, her notion of politics cannot be equated with the institution and processes

<sup>35</sup>Hannah Arendt, “The Crisis in Culture,” in *Between Past and Future* (New York: Viking, 1961), 221.

<sup>36</sup>Linda M. G. Zerilli, *A Democratic Theory of Judgment* (Chicago: Chicago University Press, 2016), 30–31.

<sup>37</sup>See Ronald Beiner, “Hannah Arendt on Judging,” in *Lectures on Kant’s Political Philosophy*, 89–156.

<sup>38</sup>Arendt, *Lectures on Kant’s Political Philosophy*, 55; Dana R. Villa, *Politics, Philosophy, Terror: Essays on the Thought of Hannah Arendt* (Princeton: Princeton University Press, 1999), 103.

<sup>39</sup>Keith Breen, *Under Weber’s Shadow: Modernity, Subjectivity and Politics in Habermas, Arendt and MacIntyre* (Farnham: Ashgate, 2012), 119–20; Villa, *Politics, Philosophy, Terror*, 90–99; David L. Marshall, “The Origin and Character of Hannah Arendt’s Theory of Judgment,” *Political Theory* 38, no. 3 (2010): 367–93; but see also Shmuel Lederman, “The Actor Does Not Judge: Hannah Arendt’s Theory of Judgment,” *Philosophy & Social Criticism* 42, no. 7 (2016): 727–41.

<sup>40</sup>Arendt, *Lectures on Kant’s Political Philosophy*, 75.



of representative democracies. Arendt's frequent references to the Greek polis might be thought to suggest she was hopelessly nostalgic, while her discussion of the French and American Revolutions<sup>41</sup> might create the impression that politics in her emphatic sense is only, if ever, possible in extraordinary circumstances.

One might therefore be tempted to dismiss her vision of politics as an ideal type that is hardly ever instantiated in real life. Given Arendt's many pessimistic remarks about her own period, one could indeed wonder whether plurality, natality, and judgment can *ever* be realized in modernity. The focus on labor and the rise of "the social," or what Arendt calls the "victory of the *animal laborans*,"<sup>42</sup> threaten to submerge the realm in which action and speech would be possible.<sup>43</sup> With the rise of consumerism, and the organization of work along Fordist lines, society is understood—and organized!—as a realm of predictable processes that stifle individuality and the creative search for new beginnings.<sup>44</sup> Politics degenerates into administration, and where it leaves space for something akin to action, it does so only for a handful of privileged individuals.

There is another possible way of reading Arendt, however, which clings to her vision of politics in a way that might be more optimistic than she herself was. Such a reading can follow Seyla Benhabib in seeing judgment—and with it, ultimately, also the conditions of plurality and natality—as not belonging exclusively to the political realm, but rather as concerning the possibility of moral judgment in general.<sup>45</sup> It can also include the dimension of "world building" inherent in judgment that Linda Zerilli emphasizes,<sup>46</sup> which has a more political dimension, but is, arguably, also not limited to "political" arenas in the traditional sense.

This reading holds that opportunities for judgment and action can arise in many social spheres (and questions the way in which Arendt is sometimes read in a "territorial" of separate social spheres, as Patchen Markell puts it).<sup>47</sup> Human beings are spectators and actors, not only in politics, but also in many other areas of life in which they encounter others: as colleagues at work, as members of neighborhood associations or NGOs, as parents of children who go to the same schools, and so forth. While such areas are often not organized in a formally democratic way, there is at least the potential of

<sup>41</sup>Hannah Arendt, *On Revolution* (New York: Viking, 1963).

<sup>42</sup>Arendt, *Human Condition*, chap. 45.

<sup>43</sup>Patricia Bowen-Moore, *Hannah Arendt's Philosophy of Natality* (New York: St. Martin's, 1989), 125–29.

<sup>44</sup>See also Ari Elmeri Hyvönen, "Invisible Streams: Process-Thinking in Arendt," *European Journal of Social Theory* 19, no. 4 (2016): 538–55.

<sup>45</sup>Seyla Benhabib, "Judgment and the Moral Foundations of Politics in Arendt's Thought," *Political Theory* 16, no. 1 (1988): 29–51.

<sup>46</sup>Zerilli, *Democratic Theory of Judgment*.

<sup>47</sup>Patchen Markell, "The Rule of the People: Arendt, *Archê*, and Democracy," *American Political Science Review* 100, no. 1 (2006): 1–14.

encountering others as equals, without a presumption of hierarchy. And often, in order to arrive at judgments in these spheres, we do imagine the perspectives of others, in ways that resemble the “enlarged mentality” that Arendt describes. This also means that the possibility of beginning something new is not limited to one social sphere; what matters for this to happen is “an agent’s attunement to its character as an irrevocable event, and therefore as a new point of departure,” as Markell puts it.<sup>48</sup>

Of course, the fact that individuals reflect and form judgments about questions in various spheres of life does not, by itself, mean that there are also social spaces for real exchanges of opinions, or “common worlds.” Many decisions are not at all taken in a way that resembles Arendt’s concepts of judgment or action. Often, unequal power relations make this hard or impossible. Moreover, if encounters follow the logic of what Arendt calls “the social,” standardized behavior is expected, with no room for deviation from the preestablished patterns. This holds in particular for the realm of capitalist wage labor.

And yet, time and again, forums for open discussions on an equal footing do emerge, at least in societies that preserve the right political conditions, such as freedom of speech and freedom of association. It can then be possible to encounter others along the lines described by Arendt: presenting our opinions to one another, trying to understand each other’s position, and coming to shared judgments. As Keith Breen notes, “Arendtian freedom concerns the vocal ability to engage and participate with others in matters of common significance whose status is uncertain and thus amenable to debate.”<sup>49</sup> Encountering each other as equals and exchanging perspectives, so that something new can emerge, is what I see as the core of “Arendtian practices.”

In a discussion about Arendt and Habermas, Craig Calhoun emphasizes that Arendt does not “tie her idea of public space to the state in the way Habermas does his notion of public sphere.” This allows for the possibility that “the occasions of public action may be multiple, each involving different mixes of people.”<sup>50</sup> Such freedom can appear in various spaces, not only in the political realm but also, for example, when patient groups push for change in the health system.

“Arendtian practices” are valuable both in themselves, as a core element of human sociability, and in their role in shaping the habits and practices of democratic societies. Arguably, no democratic society can flourish, and remain democratic over extended periods of time, without spaces in which citizens can have this experience. The joint “management of collective affairs” can shape the habits and develop the skills that citizens need for democracies

<sup>48</sup>Markell, “Arendt’s Work,” 7.

<sup>49</sup>Breen, *Under Weber’s Shadow*, 113.

<sup>50</sup>Craig Calhoun, “Plurality, Promises, and Public Spaces,” in *Hannah Arendt and the Meaning of Politics*, ed. Craig Calhoun and John McGowan (Minneapolis: University of Minnesota Press, 1997), 251.

to flourish.<sup>51</sup> The rulers of nondemocratic societies, in contrast, often try to suppress such practices, out of (probably well-founded) fears that individuals could get a taste of what it means to act together.

One contested question is whether, for Arendt, such practices could also arise in the world of work (in a non-Arendtian sense)—which is, after all, where many algorithms are currently being used. Arendt famously distinguishes between “labor,” or the cyclical, repetitive activities tied to biological necessities, and “work,” or the creation of the material world around us. In some places, she seems to reject the realms of “labor” and “work” as potential public spaces.<sup>52</sup> Some have criticized her “bifurcation of freedom and necessity and equation of necessity with labour and nature, an equation that results in labour and, by association, the economic being erroneously denied a uniquely human status.”<sup>53</sup> Arendt does indeed describe *animal laborans* as incapable of “action and speech”;<sup>54</sup> the same seems to hold for *homo faber*, where speech serves merely for the giving of orders for instrumental purposes.

But human beings are always more than *animal laborans*, even when they enter their workplaces. What Arendt has in mind when talking about the modern workplace seem to be Fordist and Tayloristic factories. But many workplaces are different, and even Fordist factories have canteens and changing rooms and permit union meetings and common festivities.<sup>55</sup> And while individuals may have work together as homogeneous equals, as Arendt notes,<sup>56</sup> such spaces also offer opportunities for meeting one another as distinct individuals and deliberating in a noninstrumental way.

Some readers have indeed rejected a reading of Arendt that sees the different spheres she describes as completely separate.<sup>57</sup> Klein argues that what Arendt feared was not so much the replacement of politics by economics as the disappearance of certain institutions that make it possible for economic issues to be debated and negotiated in the ways that she describes by drawing on the notions of judgment and action.<sup>58</sup> Arendt can thus indeed be read as holding that “all spheres of life, including labor and work, can be informed by an appreciation of the human need to be recognized and known through engaging in meaningful, nondeterministic, personally creative, and ultimately ‘public’ endeavors.”<sup>59</sup>

<sup>51</sup>Carole Pateman, *Participation and Democratic Theory* (Cambridge: Cambridge University Press, 1970), 43.

<sup>52</sup>Arendt, *Human Condition*, 149–51.

<sup>53</sup>Breen, *Under Weber’s Shadow*, 199.

<sup>54</sup>Arendt, *Human Condition*, 215.

<sup>55</sup>E.g., Huw Beynon, *Working for Ford* (London: Allen Lane, 1973).

<sup>56</sup>Arendt, *Human Condition*, 214–15.

<sup>57</sup>E.g., Markell, “Rule of the People.”

<sup>58</sup>Klein, “Fit to Enter the World.”

<sup>59</sup>Zerilli, *Democratic Theory of Judgment*, 226.

Numerous studies in economic sociology confirm that the account of work as a soulless “system” driven by instrumental reason is too simplistic. Isabelle Ferreras, in an ethnographic study of supermarket cashiers, has underlined the degree to which work has an expressive, noninstrumental dimension.<sup>60</sup> Lyn Spillman studied the role of business associations as social spaces in which complex intersubjective processes of meaning making take place, and in which sociability and collegiality can develop.<sup>61</sup> When Ronald Beiner holds that “the real danger in contemporary societies is that the bureaucratic, technocratic, and depoliticized structures of modern life encourage indifference and increasingly render men less discriminating, less capable of critical thinking, and less inclined to assume responsibility,”<sup>62</sup> it seems that one has to take into account *all* spheres in which judgment and joint action are, at least potentially, possible. These spaces matter both individually—because each one is valuable for the human individuals who take part in it; and in sum—because the fewer such spaces exist, the more the democratic ethos of a society is threatened.

However, one important reason why such spaces could continue to exist, even within the hierarchies of workplaces, was what organization research has called the “problem of control”: controlling workers who execute complex tasks is difficult and costly, hence there is hardly ever full control.<sup>63</sup> Often, one needs to give individuals and groups a certain degree of autonomy in order for them to organize themselves and to react to the specificities of new cases or new localities; sometimes, this can even lead to a rejection of the given instrumental logic and to the formulation of new goals. Arguably, the bureaucratic and technocratic tendencies that Arendt deplored in her own time still left open quite some space in which individuals could encounter each other on an equal footing, as distinct individuals, exchange their opinions, and form judgments. For example, when colleagues discuss their work, they can arrive at shared judgments about how to reorient that work. In other words, they can not only reflect, instrumentally, about how to achieve certain goals, but also reflect about new goals.<sup>64</sup> At least this holds in workplaces that are not completely algorithmically governed, such as certain warehouses in which workers’ activities are entirely determined by software.

<sup>60</sup>Isabelle Ferreras, *Critique politique du travail: Travailler à l’heure de la société des services* (Paris: Presses de Sciences Po, 2007).

<sup>61</sup>Lyn Spillman, *Solidarity in Strategy: Making Business Meaningful in American Trade Associations* (Chicago: University of Chicago Press, 2012).

<sup>62</sup>Beiner, “Hannah Arendt on Judging,” 113.

<sup>63</sup>Anthony Downs, *Inside Bureaucracy* (Boston: Little, Brown, 1967).

<sup>64</sup>On the possibility of deliberation in firms see also Andrea Felicetti, “A Deliberative Case for Democracy in Firms,” *Journal of Business Ethics* 150 (2018): 803–4; and Felix Gerslsbeck and Lisa Herzog, “The Epistemic Potentials of Workplace Democracy,” *Review of Social Economy* 78, no. 3 (2020): 307–30.

With the advent of algorithms, many such spaces that offer the potential for creating Arendtian practices might disappear, and individuals might be conditioned, even more than is already the case, towards instrumentalist instead of deliberative, or hierarchical instead of egalitarian, modes of encounter.<sup>65</sup> In the next section, I show that algorithmic decision-making is fundamentally different from such practices, with regard to the three dimensions of plurality, natality, and judgment.

### An Arendtian Critique of Algorithmic Decision-Making

When human beings encounter each other in Arendtian practices, their distinctness as individuals and their equal status shape their interaction—as captured in Arendt’s notion of *plurality*. Take, for example, a number of book lovers who run into each other in a book shop, engage in a discussion about a recent bestseller, start recommending other novels to each other, and finally decide to start a book club. They might convince individuals to start reading something that is completely different from their previous reading habits. This process is rather different from the way in which algorithms provide buying recommendations for books. Here, individuals appear as statistical constructs, as “whats” not “whos”: sums of data points that may include things like age, sex, location, and in particular previous purchases. The algorithms compare these data points to others, categorizing, comparing, and aligning them and then spitting out recommendations.<sup>66</sup>

It might be objected that the more data points are available, the more the algorithmic programs can tailor their recommendations to each specific individual. But this is, emphatically, not the same as the kind of individuality that matters for Arendtian practices. It is, rather, a more fine-grained cross-sectional analysis of data in which individuals continue to be treated as members of specific categories—and only of those categories that are captured by the algorithms. It can hardly capture the true quality of social relations and the meaning that a certain book can have for an individual who recommends it to someone else. Or take the example of job applications—here, algorithms will often look for similarities with existing employees, but they can hardly capture the more nuanced questions about how candidates would fit into the social relations of existing teams.

Another element that is missing in the algorithmic decision is the openness, the space of mutual visibility, that is constitutive of human plurality. Users

<sup>65</sup>To be sure, algorithms are not the only factor here—competitive pressures and managerial ideology certainly also play a role. But algorithms can often interact with, and reinforce, these other factors.

<sup>66</sup>There is a family resemblance, which I cannot explore here, to Arendt’s criticism of rational choice theory as a tool of political analysis in “Lying in Politics” (Hannah Arendt, “Lying in Politics: Reflections on the Pentagon Papers,” in *Crises of the Republic* [San Diego: Harcourt Brace, 1972], 37).

may or may not know what inputs go into a system, and they often also do not know what exactly the algorithmic system is supposed to produce. For example, applicants may have no reliable information about what a company considers a “good employee” and has, hence, encoded in its hiring programs. Indeed, the company may carefully guard the list of criteria, since if it were to become public knowledge, individuals might start to game the system by tweaking the information they share. To be sure, lack of transparency and strategic behavior often mar human hiring processes as well—but in these, there is at least the *possibility* of challenging decisions and demanding an explanation, in a frank encounter, which cannot be done if an algorithm made a decision and even the company itself maybe does not quite understand what it is doing.

Algorithmic decision-making is also diametrically opposed to what Arendt describes as *natality*. Instead, it is tied to the past. The potential here, as described earlier, is the possibility of encoding past knowledge so that current users do not have to repeat the same errors over and over again. But the danger is to thereby tie the present to the past in a way that makes newness impossible. Catherine O’Neill puts this point strongly: “Big Data processes”—including the algorithms used to analyse the data sets—“codify the past. They do not invent the future.”<sup>67</sup> Algorithms work with past data (which in some cases reach up to the present but might also be much older) and they are based on the assumption that the same correlations that held in the past will also hold in the future. For example, if applicants with a certain university degree were a good fit for the job in the past, it is assumed that this will also be true for the next round of applicants. An organization that uses algorithmic systems “must assume that its future applicant pool will have the same degree of variance as its current employee base.”<sup>68</sup> This is also why structural biases, e.g., discrimination against minorities, are reproduced by algorithms.<sup>69</sup>

The way in which Arendt describes *judgment* is also very different from the way algorithms make decisions. One difference has to do with the lack of transparency mentioned above. When humans are confronted with an algorithm, they often do not know how the outcomes are produced. In contrast, in processes of judgment—especially in ones that are truly intersubjective, not just imagined—we can see how the process of mutual adjustment of perspectives comes about. In “The Crisis in Culture,” Arendt mentions the notion of “wooing” others, which appears in similar ways in Kant and in Greek political philosophy.<sup>70</sup> One of its features is that it is an open, inviting mode of trying to convince others, which excludes, for example, manipulative rhetorical tricks. In the *Lectures on Kant*, Arendt emphasizes the importance of

<sup>67</sup>O’Neill, *Weapons of Math Destruction*, 203.

<sup>68</sup>Barocas and Selbst, “Big Data’s Disparate Impact,” 687.

<sup>69</sup>Kim, “Data-Driven Discrimination at Work,” 872.

<sup>70</sup>Arendt, “Crisis in Culture,” 222.

“holding oneself and everyone else responsible and answerable for what he thought and taught.”<sup>71</sup> This is, of course, an idealization. But the way in which algorithmic decision-making works is at the very opposite end of the spectrum. This can have to do with business secrecy, lack of knowledge about coding, and the intrinsic intransparency of many algorithmic systems,<sup>72</sup> but also with the ways in which algorithms are embedded in social hierarchies in which those who are subject to the algorithms’ decisions are often not able to hold those who use them accountable.<sup>73</sup>

One might object, here, that there can be forms of “newness” brought about by algorithms. There are examples of AI generating “art,” for instance, composing music. But these follow patterns that were already established and recombine them, usually producing mediocre versions of more of the same.<sup>74</sup> And natality cannot be reduced to creativity; it also has an intersubjective dimension and includes the *will* to do something new. Take, for example, IBM’s Watson’s recipe creation function, which recombines flavors, sometimes creating “new” recipes.<sup>75</sup> But what it cannot do is to completely switch categories—for example, to realize that the friends you have invited over are not in the mood for trying out new recipes, and that it would be better to order pizza and talk through their relationship problems—which requires judgment.

Thus, there are important limits on the “newness” that algorithms can create. They also have to do with one important difference between human and artificial intelligence, which is that the latter is domain specific.<sup>76</sup> Hence, it cannot recognize the need for new decisions that cut across domains. It needs to be taught beforehand what matters for its task (and it may well lead to a self-fulfilling prophecy in which these features come to

<sup>71</sup>Arendt, *Lectures on Kant*, 41. On responsibility in Arendt see also Garrath Williams, “Disclosure and Responsibility in Arendt’s *The Human Condition*,” *European Journal of Political Theory* 14, no. 1 (2015): 37–54. There are also important questions about responsibility for algorithmic systems, but for reasons of scope I cannot discuss these here.

<sup>72</sup>Burrell, “How the Machine ‘Thinks.’”

<sup>73</sup>Danielle K. Citron, “Technological Due Process,” *Washington University Law Review* 85, no. 6 (2007): 1249–1313.

<sup>74</sup>It might in fact be glitches that lead to what seem to be truly creative suggestions, as when a buying algorithm cannot distinguish which products I order for myself and which ones as presents, and comes up with a surprising “mixture” between. It is through such mistakes that software sometimes produces genuine surprises, but ironically it is a feature that programmers try to iron out.

<sup>75</sup>Alexandra Kleeman, “Cooking with Chef Watson, I.B.M.’s Artificial-Intelligence App,” *New Yorker*, Nov. 28, 2016, <https://www.newyorker.com/magazine/2016/11/28/cooking-with-chef-watson-ibms-artificial-intelligence-app>, accessed March 8, 2021.

<sup>76</sup>See Nick Bostrom and Eliezer Yudkowsky, “The Ethics of Artificial Intelligence,” in *The Cambridge Handbook of Artificial Intelligence* (Cambridge: Cambridge University Press, 2014), 316–34.

matter most). For many decision-making situations, however, “what matters” cannot be determined beforehand, but needs to be found out in the process. Algorithms might be able to find new *means* but cannot find new *ends*.

It seems no exaggeration to say that where algorithms make decisions, there is no “common world” in the Arendtian sense, in which individuals jointly interpret the reality they encounter, exchange opinions, and come to judgments.<sup>77</sup> When one encounters an algorithmic decision-making system, it does not make sense to try to convince it—even less so than in the case of a rule-bound bureaucrat, who exemplifies the “rule by nobody” that Arendt warned against.<sup>78</sup> If one wants to bring about a different outcome, the only thing to do is try to game the system, by second-guessing which parameters one has to insert in order to get the results one wants. The mindset is a purely strategic one, and while such maneuvers also happen between human beings, algorithms do not leave *any other choice* but the manipulative road. What is lost when we are faced with algorithms is the constitutively shared nature of our human reality that we are all responsible for interpreting together.<sup>79</sup> Beiner captures well what is at stake here: “judgment has the function of anchoring man in a world that would otherwise be without meaning and existential reality: a world unjudged would have no human import for us.”<sup>80</sup>

Another feature of human agency that gets lost in algorithmic decision-making are the various checks on the unpredictability of action. Of particular relevance here are Arendt’s reflections on the role of promising and forgiveness, and the dangers of acting upon the realm of nature—e.g., through nuclear technology—where the possibility of “undoing” actions does not exist in the same way.<sup>81</sup> Arendt discusses the role of modern science at quite some length in *The Human Condition*.<sup>82</sup> Some aspects of her concerns also apply to algorithms, notably the rise of instrumental thinking and the fact that scientists take decisions that affect societies as a whole.<sup>83</sup> As she writes about “the action of the scientists”: “since it acts into nature from the

<sup>77</sup>There can certainly be complex arrangements in which algorithms and human beings interact, for example in financial markets (see Donald MacKenzie, “Material Signals: A Historical Sociology of High-Frequency Trading,” *American Journal of Sociology* 123, no. 6 [2018]: 1635–83). On a purely phenomenological level, this might look like “world-making,” but it is not the Arendtian sense of “world.”

<sup>78</sup>Hannah Arendt, “Reflections on Violence,” *New York Review of Books*, Feb. 27, 1969.

<sup>79</sup>Bowen-Moore, *Hannah Arendt’s Philosophy of Natality*.

<sup>80</sup>Beiner, “Hannah Arendt on Judging,” 152.

<sup>81</sup>Arendt, *Human Condition*, 238–39.

<sup>82</sup>On the historical and intellectual context see Waseem Yaqoob, “The Archimedean Point: Science and Technology in the Thought of Hannah Arendt, 1951–1963,” *Journal of European Studies* 44, no. 3 (2014): 199–224.

<sup>83</sup>Arendt, *Human Condition*, e.g., 272, 324; see Roni Hirsch, “Bounded Action: Hannah Arendt on the History of Science and the Limits of Freedom,” *Philosophy and Social Criticism* 46, no. 4 (2020): 441–45, for a discussion.



standpoint of the universe and not into the web of human relationships, [it] lacks the revelatory character of action as well as the ability to produce stories and become historical, which together form the very source from which meaningfulness springs into and illuminates human existence.”<sup>84</sup> Similarly, algorithms “going wild” can have the same unpredictability as human action, but without the corrective mechanism provided by promising and forgiveness.<sup>85</sup>

Thus, *some* spaces for judgment need to be maintained, from an Arendtian perspective—but are there reasons to fear that algorithms would make this impossible or even just less likely? After all, not *all* social spheres need to be, or indeed could be, governed by what I have called “Arendtian practice.” And if algorithms could contribute to liberating individuals from labor or work, they might indeed help create *more* space, or, more concretely, more time, for Arendtian practices.<sup>86</sup> So, do we not run the risk of overlooking the emancipatory potentials of algorithms?

This is a valid point, but it misses the thrust of my argument. My point is not to reject algorithmic decision-making; there are many places where algorithms can be very useful tools. Instead, I want to warn against them crowding out, often for the sake of alleged cost reductions,<sup>87</sup> spaces for Arendtian practices, which have specific, irreplaceable value. This concern, which is in the spirit of Arendt’s critique of bureaucracy, economism, scientism, and other deterministic approaches, takes two forms.

The first is the *overall* space for Arendtian practices in society; this point is structurally similar to her admonition of the rise of the “social” that crowds out spaces for political action. We might see a phenomenon of “algorithm creep”: of algorithms slowly taking over far more space than they should.<sup>88</sup> Such a scenario could come about because each single choice to introduce an algorithmic decision-making system, taken by itself, may seem harmless—and there may be a coordination problem among those who decide about the introduction of algorithmic decision-making systems. The summative effect might be considerable, and it might be far harder to return to nonalgorithmic decisions once the programs have been introduced. One factor to take into account in this context is what psychologists call “automation bias”: humans

<sup>84</sup> Arendt, *Human Condition*, 324.

<sup>85</sup> On forgiveness in Arendt see Bowen-Moore, *Hannah Arendt’s Philosophy of Natality*, 58–60 and 147–49.

<sup>86</sup> E.g., Tegmark, *Life 3.0*.

<sup>87</sup> The issue of cost reductions is documented, for example, in the case studies discussed in Eubank, *Automating Inequality*.

<sup>88</sup> John Danaher, “The Threat of Algocracy: Reality, Resistance and Accommodation,” *Philosophy & Technology* 29, no. 3 (2016): 245–68, analyzes a similar scenario from a republican perspective, under the heading of “algocracy.” His focus is on the threats of opacity and the impossibility of human understanding and participation in decision making.

tend to trust automated systems more than may be appropriate.<sup>89</sup> When Joseph Weizenbaum developed an early form of artificial intelligence, a program called ELIZA that could mimic a therapist, he was shocked to see the extent to which humans shared intimate details with it.<sup>90</sup> Even programs that are meant to be merely auxiliary tools might be given far too much power, with humans giving up their own responsibility and simply going with “what the computer said.”

The second reason for concern is that algorithms might be applied to questions for which they are simply not suitable. Here, Arendt’s notions of plurality, natality, and judgment can in fact help develop criteria for thinking through various cases. These may not be the only relevant criteria—for example, in a situation of emergency there may be reasons to prioritize efficiency and speed in order to satisfy urgent needs—but they provide *pro tanto* reasons.

With regard to plurality, one key question is whether it is appropriate to treat individuals according to predefined categories, or whether their distinctness needs to be taken into account. In a discussion of “technological due process,” Danielle Citron draws a basic distinction between rules, standards, and combinations of both: “A rule prescribes *ex ante* an outcome for a particular fact scenario. . . . On the other hand, a standard requires decision makers to exercise discretion, applying *ex post* policies to events.”<sup>91</sup> While rules have the advantage of predictability, standards permit decision makers to tailor an outcome to the facts, increasing the likelihood of an “ideal” ruling.<sup>92</sup> Algorithms, however, are rule based—therefore, “decisions best addressed with standards should not be automated.”<sup>93</sup> While this is certainly true, it may be questioned whether it is sufficient as a guideline. One important question is that of alternatives: Will those who make discretionary decisions really make them according to the requirements of the situation, or is it likely that considerations that should be irrelevant (e.g., some form of private gain) will intrude into the decision-making process? Importantly, the notion of plurality, in conjunction with the notion of judgment, implies that the most adequate alternative may often not be decisions by single individuals with their own idiosyncracies and biases. Rather, an appropriate way of judging particulars may be to bring together a group of diverse people who can discuss face to face, on an equal footing, and try to achieve an “enlarged mentality” together.

<sup>89</sup>Linda J. Skitka, Kathleen L. Mosier, Marie Burdick, and Bill Rosenblatt, “Automation Bias and Errors: Are Crews Better Than Individuals?,” *International Journal of Aviation Psychology* 10, no. 1 (2000): 85–97.

<sup>90</sup>Weizenbaum, *Computer Power and Human Reason*, 141.

<sup>91</sup>Citron, “Technological Due Process,” 1301.

<sup>92</sup>*Ibid.*, 1302.

<sup>93</sup>*Ibid.*, 1304.

With regard to natality, an important question is whether it is appropriate to “go on” with existing patterns, or whether there might be a need to take a new tack. Hence, at a minimum, the desirability of the existing patterns needs to be scrutinized: Are they such that we can, with a clear conscience, continue with them? Many data sets reflect past injustices, such that algorithms come to be biased against women or ethnic minorities;<sup>94</sup> such phenomena could, however, be corrected if the data were prepared, and the algorithms trained accordingly. But in other cases, completely new approaches might be needed—think, for example, about the shift in the treatment of drug addicts that turned away from seeing it as a mere medical problem, and started to take the social contexts into account. Such a shift could not have been initiated by an algorithmic program; it required human agency and political struggles.<sup>95</sup>

Relatedly, there is the question of judgment. Judgment is not needed, or if it is, only minimally, when one can proceed with given ends in mind (or rather, in the code), in one single domain, and when all that matters is finding the best strategy for getting there. But one may well ask how many decisions in the realm of human affairs are like this. Earlier, I quoted Breen’s formula about “matters of common significance whose status is uncertain and thus amenable to debate.”<sup>96</sup> In such situations, the ends are not predefined, and there are no clear-cut, uncontroversial standards that one could apply for describing outcomes as successes or failures.

In contrast, algorithmic decision-making is ideal for situations in which the status is certain, there is no need for debating the goals, and the criteria of success are clear-cut. This is why computers and artificial intelligence are extremely good at games: there, the “set of rules” is “complete and consistent.”<sup>97</sup> But many social situations do not have the character of games. Think, for example, about many challenges of social work, where the aim is not—or should not be—to simply tick boxes and achieve a certain outcome, but rather to pay attention to the specific situation of each individual and to reflect, together with them, what the goal of the intervention is actually supposed to be.

In stark contrast to games, many decisions that humans have to take are multidimensional and we cannot expect all parameters to fit into predefined categories—we may not even know yet which categories will be relevant for understanding them. In other words, we do not yet have general categories under which to subsume the particulars, but rather need to find the right—

<sup>94</sup>See Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).

<sup>95</sup>As described, for example, in Maxwell A. Cameron, *Political Institutions and Practical Wisdom: Between Rules and Practices* (New York: Oxford University Press, 2018), 95–99.

<sup>96</sup>Breen, *Under Weber’s Shadow*, 113.

<sup>97</sup>Weizenbaum, *Computer Power and Human Reason*, 44.

new!—categories for grasping what matters about the particulars; and here, judgment is needed. Again, social work is a case in point: there might be biographical issues that do not fall under the predefined categories of a questionnaire. The danger, obviously, is to resort too quickly to algorithms, thereby closing down the avenue for judgment, and for a deeper understanding of the problems. In such cases, we try to fit particulars into general categories that are not right for them, and maybe even impose on them inadequate value judgments that are built into the algorithms.<sup>98</sup>

Seeing algorithms as ready solutions for various economic, organizational, and civic tasks requires understanding these tasks as mere problems of administrative weighing and prioritization—but they might, in fact, arise out of deep, and maybe even tragic, value conflicts. Weizenbaum, one of the early pioneers of AI research, warned against such a reductionism of rationality to “logicality” early on, mentioning Arendt’s skepticism towards formalistic analyses of politics.<sup>99</sup> The technocratic, model-based approach to conflict that appeared in the “rational choice” view of the Cold War, and in technocratic approaches to conflicts in general, amounts to a denial of “the very possibility of the collision of genuinely incommensurable human interests and of disparate human values, hence the existence of human values themselves.”<sup>100</sup>

Here, an interesting connection can be drawn to Arendt’s essay “Truth and Politics.” Arendt contrasts the “element of coercion” in truth with the “representative thinking” of citizens in which the opinions of others are taken up. It would be a misunderstanding to read this essay as holding that “truths”—or, to use a less highly charged, but by now also controversial, term, “well-established facts”—are irrelevant for politics; Germany *did* invade Belgium and not the other way round.<sup>101</sup> Knowledge about facts that are so clearly established that they require no discussion is indeed the kind of knowledge that could be encoded into algorithms: for example, political agents can and should use pocket calculators whose software encodes certain mathematical laws. The problem, however, is that few facts are as clear, and as clearly recognizable, as mathematical laws; even “scientific” facts are discovered in ongoing discussions and interpretations of the evidence by scientists.<sup>102</sup> Their significance and relevance in evaluating situations and possibilities must also be a matter of judgment. This means that the boundary between the realms of facts and opinions is notoriously

<sup>98</sup>Nicholas Diakopoulos, *Algorithmic Accountability Reporting: On the Investigation of Black Boxes* (New York: Tow Center for Digital Journalism Columbia University, 2014) [http://www.nickdiakopoulos.com/wp-content/uploads/2011/07/Algorithmic-Accountability-Reporting\\_final.pdf](http://www.nickdiakopoulos.com/wp-content/uploads/2011/07/Algorithmic-Accountability-Reporting_final.pdf), accessed March 8, 2021.

<sup>99</sup>Weizenbaum, *Computer Power and Human Reason*, 13–16.

<sup>100</sup>*Ibid.*, 14.

<sup>101</sup>Hannah Arendt, “Truth and Politics,” in *Between Past and Future*, 239.

<sup>102</sup>Jon Simons, “Politics and Truth, Immanence, Practice and Constellations,” *Social Epistemology* 15, no. 1 (2001): 43–44.

contested—and hence the question of which facts could be embedded in algorithmic programs and which ones need to remain debatable in the realm of opinions is equally contested.

In summary, what is needed is a conscious, responsible delineation of the kinds of decisions that we can hand over to algorithms with a clear conscience, the ones where algorithms and humans can collaborate in productive ways, and the ones where it would be irresponsible to let algorithms take over, because the potential for “Arendtian practices,” and the need for attention to the dimensions of plurality, natality, and judgments, would be lost. When decisions about the introduction of algorithms are made in practice, however, this often happens against a background in which there is *already* too little space for Arendtian practices, for example because the logic of “the social” had already been given far too much space. In that sense, the problem is not new, but with new technological possibilities of algorithmic decision-making it gains new urgency.

These questions lead back to the issues of justice I alluded to in the introduction. In addition to the problem of reinforcing old, or creating new, injustices through algorithmic decision-making, one crucial question is for *whom* algorithms are used. Given the way in which algorithmic systems blend into existing hierarchies of power and status, we should hardly expect these systems to take away decision space in the executive suites and other realms in which the rich and the powerful meet. It is much more likely that such spaces will shrink for the middle and lower ranks of workplace hierarchies, and for those who receive services, e.g., medical services, from overburdened and understaffed welfare-state institutions. This is another sense in which many algorithms are unlikely to bring about anything new, but might rather reinforce existing injustices.<sup>103</sup>

Think, for example, about a team of street-level bureaucrats whose daily meeting in which they distribute tasks and discuss cases is replaced by an algorithmic allocation system; this takes away an opportunity in which they could not only discuss instrumental questions, but also deliberate together about goals that need to be rethought, or new starts that need to be made. Not even the very few moments in which workers, patients, or the recipients of public services could encounter one another as individuals, exchanging opinions and forming judgments together, might be preserved if their work is constantly algorithmically allocated, tracked, and controlled. An experience that is, arguably, crucial for maintaining a democratic ethos would then become a privilege for the few.

## Conclusion

I have contrasted decision-making by algorithmic systems with “Arendtian practices” shaped by the elements of plurality, natality, and judgment.

<sup>103</sup>Herzog, “Algorithmic Bias and Access to Opportunity.”

I have argued that the advent of algorithmic decision-making might shrink the spaces for Arendtian practices, in the realm of work and elsewhere. There are good reasons to be concerned about these changes, not only because the overall sum of Arendtian practices might be reduced, especially for less privileged members of society, but also because there are good reasons not to allow algorithms to take on certain kinds of decisions: decisions that require attention to human difference and awareness of changing circumstances, for which there are not predefined rules, and which require the integration of different perspectives rather than the mere application of established categories. The Arendtian notions of plurality, natality, and judgment thus provide helpful criteria for thinking about the distinction between algorithmic and human decision-making and seeing the strengths and weaknesses of each.

The arguments I have developed are *pro tanto* considerations. They apply in addition to other considerations—of justice, of the rule of law, of stability, etc.—that also matter for deciding where algorithmic decision-making should or should not have a place, and which the literature has so far mostly focussed on. To be sure, in order to judge concrete cases, one has to turn from the idealized versions of algorithmic and human decision-making and judgment towards concrete evaluations as to how each would work in specific scenarios. This, however, in turn shows the irreducibility of human judgment: in order to *evaluate* the usefulness of new technologies, including algorithmic systems, we cannot rely on algorithmic decision-making itself. Otherwise, we run the risk of ending up in the state that Arendt warned against at the beginning of *The Human Condition*: as “thoughtless creatures at the mercy of every gadget which is technically possible.”<sup>104</sup>

Might algorithms one day become so human-like that they would indeed be capable of matching our abilities of judgment, and of respecting plurality and starting something new? It is worth noting just how far we are from such scenarios—the impressive achievements of artificial intelligence are all highly domain-specific. While one can speculate about the future of artificial intelligence and ask whether it might, one day, resemble human cognition,<sup>105</sup> the Arendtian perspective points in a different direction: Why should one try to mimic what is specific about humans, rather than delegate *other* tasks to robots, which are far less central to what it means to live a human life, and liberate humans for genuinely human activities? In 1976, Weizenbaum asked, “why are there still poets?” referring to the existence of forms of intelligence above and beyond that “logicality” that can be encoded in programs.<sup>106</sup> But even if algorithms could one day write poems, what would be the point? Often, it is the *writing* of the poem, and the human context in which it took place, that matters most.

<sup>104</sup> Arendt, *Human Condition*, 10.

<sup>105</sup> Bostrom and Yudkowsky, “Ethics of Artificial Intelligence.”

<sup>106</sup> Weizenbaum, *Computer Power and Human Reason*, 247.

---

While the point of this article is to illuminate the nature of Arendtian practices and algorithmic decision-making—and the focus is thus on the *contrast*—the possibility of fruitful collaboration and even of coconstitutive processes can and should not be excluded (at least not on the basis of the arguments provided here). One interesting question is to what extent computational processes can help make explicit the *sensus communis* about an issue, for instance, by aggregating individual judgments and making them visible. However, they cannot replace the processes in which individuals, as a (real or imagined) group, *make sense* of the results.

Algorithms can certainly help us to analyze patterns of behavior, and even help us to diagnose our own biases.<sup>107</sup> There is no point in trying to beat them when it comes to executing predefined, repetitive tasks—but as Arendt noted, this just proves that rational calculation is not “the highest and most human of man’s capacities.”<sup>108</sup> Algorithms can take on repetitive tasks that build on firmly established, uncontroversial knowledge. This could, ideally, free up spaces for truly human activities characterized by plurality, natality, and judgment. Leah Bradshaw’s warning seems more relevant than ever: “The fact of natality places the burden upon us to ensure the continuity of the species, to guarantee that the new human being will be able to act into the future, and to not obstruct the possibilities for his creative potential.”<sup>109</sup>

<sup>107</sup>Kim, “Data-Driven Discrimination at Work,” 872.

<sup>108</sup>Arendt, *Human Condition*, 172.

<sup>109</sup>Leah Bradshaw, *Acting and Thinking: The Political Thought of Hannah Arendt* (Toronto: University of Toronto Press, 1989), 111.