

## Identità personale e algoritmi. Una questione di filosofia morale [*Personal identity and algorithms. An issue of moral philosophy*]

by Simona Tiribelli (2023) Milano, Carocci Editore, 131 pp.

Lorenzo Olivieri 

University of Bologna

The ubiquity and potentialities of algorithms, as well as their supposed opacity, have transformed them into a modern myth, an “algorithmic drama” running across various domains in the struggle to understand their increasing influence on societies and individual lives (Ziewitz 2016). *Personal identity and algorithms. An issue of moral philosophy* (original title in Italian: *Identità personale e algoritmi. Una questione di filosofia morale*) by Simona Tiribelli represents a precious attempt to disentangle one of the “acts” of this drama, namely algorithms’ impact on personal identity. Assistant Professor in Ethics at the University of Macerata and Director for AI and Technology Ethics at the Institute for Technology & Global Health in Boston, Tiribelli has written extensively on moral freedom, algorithmic decision making and the ethical principles for Artificial Intelligence. Her book relies on ethical theories and concepts to enquire about how algorithms interfere with moral autonomy, eventually shaping the process of self-formation and the possibility of choosing, pursuing and endorsing values, aims, personal projects and beliefs. In doing this, the major quality of the book is to move away from recent discussions about personal identity, shifting the focus from the *protection* of digital identity to the *construction* of personal identity.

As Tiribelli points out in the first pages of the book, the contemporary debate concerning digital identity and the ethics of Artificial Intelligence is mostly informed by legal notions and underpinned by an informational conception of personal identity (see Mittelstadt et al. 2016 for a review). By conceiving of the self as a data subject or as an informational agent, who produces data and who is fed by data, this body of scholarship tends to analyze the relation between digital technologies and personal identity mostly in terms of informational privacy. Yet, what remains problematic from an ethical perspective is, according to Tiribelli, the process of construction of personal identity, how algorithms interfere with and re-model people’s freedom of choice and action. What matters is not just to protect personal data, but to understand how those data *become* personal and with which consequences. This point constitutes the theoretical pillar and starting point of Tiribelli’s proposal: the pursuit of ethical values, and the choices made accordingly, are what define people as unique and specific individuals. Personal identity must hence be conceived as an open and genuine per-

sonal project which, however, is increasingly shaped and threatened by the datafication of life and by the pervasiveness of digital environments.

To articulate these issues, the book is divided into three chapters. Chapter One delineates the notion of moral freedom which is later deployed to address the impact of algorithms on personal identity. Unlike metaphysical or socio-political freedom, moral freedom is defined by its normative dimension: it depends on internal, self-reflexive approval and it concerns what we ought to do and what has value for us. The ethical-normative dimension of moral freedom can be further articulated along a positive dimension (the freedom to determine values, ideas, and beliefs genuinely and uniquely) and a negative dimension (the independence from any relevant constraints on action). Drawing on this, Tiribelli identifies autonomy and the availability of alternative and morally heterogeneous options as the two *conditions sine qua non* for the exercise of moral freedom. To properly guarantee freedom of choice, individuals must have the possibility to act rationally and authentically without any forms of coercion and manipulation. Moreover, within the social contexts in which people must choose and act, it is only the presence of a plurality of options embodying different values, beliefs, reasons, and ideas – namely, the availability of morally heterogeneous options – that ensures and guarantees agent's reflection on, endorsement and pursuit of specific values. This element is particularly relevant since it allows us to avoid conceiving moral freedom as complete independence of the subject, but rather to stress both its individual and socio-relational dimension. Overall, Chapter One provides a well-thought theory of moral freedom, which brings together several threads of scholarship and it is clear also to readers not familiar with moral philosophy.

Building on the concepts described in the previous chapter, Chapter Two analyzes how algorithmically generated knowledge interferes with the two *conditions sine qua non* of moral freedom. By relying on models, patterns and correlation, algorithms (such as machine learning, deep learning systems, and recommender systems) profile users along groups and categories to predict individuals' behaviors and steer their actions. The first consequence is that algorithms reduce the options available to the users. To elaborate on this argument, Tiribelli proposes to understand algorithms as new choice architectures (Thaler and Sunstein 2008). What is crucial, from an ethical point of view, is how these systems define and shape the content available to users, including the possibility to access informational, socio-economic and socio-cultural opportunities. By governing and pre-selecting what is shown as well as what is *not* shown, the algorithms deployed by search engines or by social networking systems hence structure the options available to the users. Yet, unlike traditional and institutional choice architectures, algorithms are based on rules and criteria established to achieve goals set by commercial parties and hence aim to maximize, for economic purposes, users' clicks and engagement. Moreover, the deployment of echo-chambers, filter bubbles, epistemic bubbles and confirmation bias tend to reduce, both qualitatively and quantitatively, the range of options available to the users. In this way, algorithms promote contexts of choice which are "characterized by a reduced level of socio-relational heterogeneity and, hence, defined by a lower possibility to meet the unexpected, the different, the alternative, the differing, even the new" (p. 74, *my translation*). The second interference concerns algorithms' negative impact on moral autonomy and emotions. User profiling and targeting tend to recommend emotionally charged content to trigger specific reactions and behaviors, such as online shopping,

click-through rates, and subscription websites. Therefore, algorithms reverse moral reasoning: rather than epistemologically contributing to users' choices, they become the cause of their actions. The delegation and deskilling of moral and critical reasoning progressively affect and inhibit individuals' reflective endorsement, eventually jeopardizing moral autonomy. The second chapter does a precious job of connecting widely known issues of algorithmic governance with ethical concepts. However, the chapter might look, at least to STS scholars, relatively poor in terms of empirical research and well-analyzed study cases. Throughout the chapter, Tiribelli proposes some examples like social networks, Google search, and Cambridge Analytica, but the discussion remains quite abstract. While a more empirical analysis was probably beyond Tiribelli's goals, the chapter might have benefited from engaging with other works which have addressed similar problems. For instance, Tanya Kant's *Making It Personal. Algorithmic Personalization, Identity, and Everyday Life* (2020) is a book animated by the same concerns and which, thanks to its more ethnographic but less ethically oriented focus, would have represented an interesting reference point.

Chapter Three concludes the book with *pars construens* delineating an ethical framework of personal identity in the algorithmic era. On the one hand, in contemporary technologically mediated societies, it seems impossible, and unreasonable, to achieve complete control over one's personal data. On the other hand, a major problem of informational accounts of personal identity is the struggle to identify, *a priori*, what counts as personal. To solve this conundrum, Tiribelli proposes to focus on the ethically relevant gap between what *concerns* us and what *characterizes* us, where the latter has to do with what "we endorse as cause and reason of our choices and actions; hence, what we choose as relevant and which become connotative of our identities" (p. 101, *my translation*). Drawing on this, Tiribelli singles out three ethical dimensions which are fundamental to guarantee the construction of personal identity according to the values, beliefs, and projects we decide to pursue openly and authentically. First, an epistemic dimension, namely the possibility of creating, experimenting, and testing new thoughts and ideas. Second, a socio-relational dimension, which concerns the exposure to heterogeneous interpersonal relations and social practices to genuinely create significant social bonds. Lastly, a normative dimension, which refers to the freedom to follow, endorse and realize what we consider relevant, in terms of values, aims, and plans. To ensure the respect of these dimensions, Tiribelli closes the chapter by proposing two ethical criteria for the design of algorithms: intelligibility and heterogeneity. The criterion of intelligibility prescribes the design of algorithmic systems where users are informed about the profiling process to which they are subjected and they can modify the parameters underlying that process. This condition would allow users to alleviate the asymmetry between how they are algorithmically seen and treated and who they really are, eventually providing them with the possibility to control the options and information shaping their contexts of choice. The criterion of heterogeneity aims at guaranteeing users' exposure to alternative and unknown ways of thinking in order to keep them inclined to critical reasoning and open to mutual recognition and understanding.

Overall, the book has the important merit of advancing the debate about personal identity and algorithms by focusing on the process of self-formation. However, from an STS perspective, the book suffers from two major problems. First, as already noted, Tiribelli does not discuss ethnographic research which has explored the influence of algorithms on people's

choices (for example, Graham 2018; Cohn 2019). While this choice allows Tiribelli to develop clear-cut arguments, these tend to remain relatively abstract and general. The engagement with STS scholarship and media studies would have contributed to better substantiate her arguments as well as to make them more nuanced. The second problem concerns the essentialist account of identity which seems to underpin the book's theoretical framework. Tiribelli clarifies that metaphysical issues about the self are beyond the scope of the book, yet throughout her analysis the self emerges as a rather fixed and static entity, whose defining properties are impenetrable and independent of the agency of other entities. This might sound particularly problematic in the light of post-human and performative accounts (Barad 2003) which, on the contrary, stress the mutual shaping of subjects and objects. Similarly, the notion of authenticity, which plays an important role in the book as one of the sources of moral freedom, is underdeveloped, especially regarding what it means to be authentic in technologically mediated contexts (Beerends and Aydin 2023). Despite these points, the book offers a very clear and rigorous ethical framework for understanding the impact of algorithms on personal identity. It also provides ethical criteria that can be helpful beyond the realm of moral philosophy, especially to scholars interested in the assessment and design of algorithm governance.

## References

- Barad, Karen (2003) *Posthumanist Performativity: Toward an Understanding of How Matter Comes to Matter*, in "Journal of Women in Culture and Society", 28(3), pp. 801-831.
- Beerends, Siri and Aydin, Ciano (2021) *Negotiating Authenticity in Technological Environments*, in "Philosophy & Technology", 34(4), pp. 1665-1685.
- Cohn, Jonathan (2019) *The Burden of Choice: Recommendations, Subversion, and Algorithmic Culture*, New Brunswick (NJ), Rutgers University Press.
- Graham, Timothy (2018) *Platforms and Hyper-Choice on the World Wide Web*, in "Big Data & Society", 5(1), pp. 1-12.
- Kant, Tanya (2020) *Making it Personal: Algorithmic Personalization, Identity, and Everyday Life*, Oxford (UK), Oxford University Press.
- Mittelstadt, Brent D., Allo, Patrick, Taddeo, Mariarosaria, Wachter, Sandra and Floridi, Luciano (2016) *The Ethics of Algorithms: Mapping the Debate*, in "Big Data & Society", 3(2), pp. 1-21.
- Thaler, Richard and Sunstein, Cass R. (2008) *Nudge: Improving Decisions About Health, Wealth, and Happiness*, New Haven, Yale University Press.
- Ziewitz, Malte (2016) *Governing Algorithms: Myth, Mess, and Methods*, in "Science, Technology, & Human Values", 41(1), pp. 3-16.