



Better than they know themselves? Algorithms and subjectivity

Liran Razinsky¹

Accepted: 23 November 2023 / Published online: 15 December 2023
© The Author(s), under exclusive licence to Springer Nature Limited 2023, corrected publication 2023

Abstract

The paper explores the widely circulated idea that algorithms will soon be able to know people “better than they know themselves.” I address this idea from two perspectives. First I argue for the particular subjective qualities of experience and self-understanding issuing from our engagement with the world and the constitutive role of our reflexive relation to ourselves. These are not “known” by the algorithms. I then address our fundamental opacity to ourselves and the biased, partial, and limited nature of human self-understanding. Our failure to know ourselves is however essential to our subjectivity and therefore, to know a subject in a perfect way that bypasses these limitations is actually not to know them. Taken together, both directions show that while algorithmic knowledge of humans can be vast, and can outperform their own knowledge, it remains foreign to their subjectivity and cannot be said to be better than self-understanding.

Keywords Self-knowledge · Subjectivity · Algorithms · AI · Big data

Introduction

Everyday life is exceedingly mediated by algorithms. It isn’t only activities in the virtual world; real life, too, comes dressed in algorithms. Many choices we make, the culture we consume, our knowledge of the world as we receive it through news sites and social networks, our social ties and our possibilities in terms of work, health and so on, are curated for us, pushed on us, and decided for us by algorithms. In an attempt to tailor their recommendations to us, to sell us the right product at the right time, to make the most accurate evaluation of us, these algorithms are purported to know us in the best way possible. What is the measure of knowing people best? This paper analyzes a popular answer to that question whereby high-tech

✉ Liran Razinsky
liranr@biu.ac.il

¹ The Program for Hermeneutics and Cultural Studies, Bar Ilan University, Ramat Gan, Max and Anna Web St, Israel



entrepreneurs, corporations and journalists argue that algorithms know people so well that their understanding of people exceeds their own self-understanding. By means of this discussion I hope to say something more far reaching about the algorithmic understanding of people and its limits.¹

“Google will know the answer to your question before you have asked it,” says Ray Kurzweil, famous prophet of AI. “It will have read every email you’ve ever written, every document, every idle thought you’ve ever tapped into a search-engine box. It will know you better than your intimate partner does. Better, perhaps, than even yourself” (Cadwalladr 2014).

Similar statements abound: “What happens when technology knows more about us than we do?” wonders Poppy Crum in her popular TED talk (2018), where she celebrates empathic technology’s ability to know much more about people than what they think they are sharing. The same idea is everywhere to be found in tech and general journals: “Recommender Systems —Know your users better than they know themselves” (Gama Rose 2021); “AI [...] will know us better than we know ourselves” (McCarthy-Jones 2020); “Big data knows you better than you know yourself” (Shaw 2016); “Google knows you better than you know yourself” (Carmichael 2014); “When Facebook knows you better than you know yourself” (Evans 2015). The jealous appropriation of self-knowledge is posited as a challenge, a summit to be conquered: “Perhaps the biggest claim smart apps and AI-bots can make is that sooner or later, they’ll know users better than they know themselves” (Koksall 2018).

Everyone then—Facebook, recommender systems, AI, Google and big data industry in general—seem to know us better than we know ourselves. These examples celebrate, or warn against, the power of algorithms to analyze huge amounts of data concerning us, apply to them machine learning procedures and come up with an understanding of humans that exceeds what they know about themselves because, unlike humans, the algorithms tell the truth. Thus, for example, explains one of these articles, “we all have stories we tell ourselves, stories about who we really are. All Google Now does is reveal whether these stories check out. They often don’t” (Carmichael 2014). The algorithms have no use for self-deception, are supposedly not biased or selective in the picture they construct of human individuals, and work on huge amounts of data while monitoring people below the level of awareness. All this goes into the claim that they know us better than we know ourselves.

This notion of algorithms knowing us intimately to the furthest possible extent is not, I argue, marginal or just a manner of speaking. It is, rather, a crucial cornerstone of the algorithmic justification regime: Results of algorithmic processing are to be trusted and adopted because they exceed human knowledge, being better than what we can figure out for ourselves. This holds in general and also with regard to knowledge about oneself. The algorithms are to be heard not only for their objectivity but also for their specificity. Their recommendations, advice and judgments

¹ I wish to thank Yoav Ronel, Omri Blum, Shira Shemi and Asaf Hazan for brainstorming and for painful, helpful feedback, and Mark Joseph, Yaron Wolf, Zohar Kaufman, Robert Rosenberger, Arnon Keren, Dror Yinon and most of all Hili Razinsky for helping push ideas further. This paper is part of a research project funded by the Israel Science Foundation (no. 1426/21).



are not generic but specific to us. We are to accept algorithmic interventions in our lives precisely because they are tailored to us and based on thorough information about us. Large amounts of data allow for that high a degree of personalization. The deal, as it were, is that the algorithms can supposedly tailor the world for us and we should help them do so by allowing as much information about us to be available for them. We need them to know us personally (an idea propagated by Google itself. For example, Eric Schmidt’s “you give us more information about you, [...] and we can improve the quality of our searches” [Thompson 2010]).

The same comparison of algorithmic and self-knowledge can be heard from its promoters as well as its critics. “Who knows you best in the world?,” world renowned historian and critic Yuval Noah Harari asks, in a discussion (with Harris and Thompson) in *Wired*: “[W]hen you are two years old it’s your mother,” but soon enough it’s you who knows you best. But today, corporations and governments “know you better than you know yourself. [...] You know it’s the oldest advice in the book: Know yourself. But in the past you did not have competition.” (Harari et al. 2018).

“What will happen to society, politics, and daily life,” goes the concluding line in Harari’s book *Homo Deus* (2017), “when non-conscious but highly intelligent algorithms know us better than we know ourselves?” What troubles him is the precision of the dataistic monitoring of our thoughts and behaviour and the possibility of influencing them, subconsciously, in shaping what we believe and the way we act.

Many, then, believe the ultimate algorithmic understanding of human individuals would be one that surpasses their own self-understanding.² Why, however, this particular criterion? Why consider knowledge produced by algorithms not just as good, excellent, or even complete and definitive, but better than people’s knowledge of themselves, as the apex of the scale?

Numerous questions and preliminary observations immediately offer themselves. Do we know ourselves? To the extent that we do, can one know us, in a *similar way*, better? Can one compare the knowledge produced by the algorithm to self-knowledge? Would it not necessarily be of a different kind? What does this, namely the idea that the algorithm will know us better than we know ourselves, imply for humanity’s longstanding cultural project of knowing itself, and for personal efforts to know oneself from the Greek *gnothi seauton*, through the history of Western philosophy, to the advent of psychoanalysis, among many other way-stations?

In the statements I reviewed above, self-knowledge is taken to be the ultimate form of knowledge, the fortress to be conquered if algorithmic knowledge is to prove victorious. It is as if the effort or the feat is to outdo human subjects at the very heart of what characterizes them as human subjects.

² Two more nuanced critics of this idea are Andrejevic (p. 3), discussed later, and Hong. Hong explores a technological fantasy, particularly pronounced in followers of the quantified self movement, of “data’s intimacy,” the “promise” that “machines will know us better than we know ourselves and that through these friendly devices, we will also achieve better self-knowledge” (Hong 2020, p. 79). Hong argues there is a pressure on subjects to transform in a way that will make them more suitable for algorithmic understanding (pp. 26, 161).



Kurzweil, in the above quotation, first compares the algorithmic knowledge about you to how your partner knows you and then makes a more ambitious claim—nah, the algorithmic knowledge will even exceed the knowledge you have of yourself. But we are not dealing here with degrees: Your understanding of yourself is not slightly better than that of your spouse about you. It is structurally of a different kind. The very uniqueness that makes it an achievement target for algorithms also makes this algorithmic effort necessarily fail. This will be the first point I will develop below.

The second point will bring to the fore the fact that self-knowledge, rather than being the best possible kind of knowledge about oneself, is actually very partial, lacking and biased, and to a large extent not something that demands much effort to compete with. The fortress is weak and its ruler misinformed. To quote Freud:

The ego feels uneasy; it comes up against limits to its power in its own house, the mind. [...] Psycho-analysis [...] can speak thus to the ego:— ‘[...] a part of the activity of your own mind has been withdrawn from your knowledge [...] You feel sure that you are informed of all that goes on in your mind [...] in spite of the most obvious evidence that a great deal more must constantly be going on in your mind than can be known to your consciousness. [...] your intelligence service breaks down [...] the news that reaches your consciousness is incomplete and often not to be relied on. (Freud 1917, pp. 141–143)

Freud’s prosopopoeic address to an over-confident subject is just one way to refer to the fact that the fortress has more windows than walls, and one does not have to adopt psychoanalysis to see the frailty of self-knowledge.

The two points, then, address two almost contradictory aspects of self-knowledge: it is a constitutive kind of knowing that differs from any other and cannot be imitated, and yet it is a deficient knowledge full of holes and biases. The contradiction, I will show, is to some extent only apparent, as the holes are part and parcel of the fragile fortress, of which the constant failure of self-understanding is an essential part.

Trying to dig further into what is meant by “better than” we can point to a few possibilities: Harari and Harris are mostly concerned by the possibility of influencing people without their knowing. Algorithms will be able to influence what you buy, how you act, or your political opinions and behaviour, and yet you’ll be altogether unaware of that influence. It is as if the algorithm can sneak behind the back of the rational subject of the Enlightenment, who can and should think and decide for itself (Fisher [2022] is a good and thorough recent exposition of this line of thought).

Then there is the aspect of time. Algorithms can know what you want before you do. Kurzweil’s words above hint at this, as does a famous quote from Erik Schmidt, Google’s former CEO: “most people don’t want Google to answer their questions. They want Google to tell them what they should be doing next” (Jenkins 2010).

Another aspect of “better than” is “more fully than,” more extensively, as Kurzweil’s quote, again, seems to hint. The algorithmic understanding would be all-encompassing, and would bring data into its calculations derived from all events, all moments, all aspects, allowing it to surpass the more cognitively limited human knowledge. The latter can only deal with so many concerns, themes, issues; and



being biased and in other ways restricted (by individual abilities, handicap, perspective, feelings) the human subject will skip over much of the information it could have had. Human individuals can also not go into the minute resolution of the algorithmic eye and lack the processing power of the machine learning algorithms that are used to make sense of data.³ A related possibility of understanding “better than” is that algorithmic knowledge of people will be without the inevitable biases and myopia characteristic of human self-understanding.

While much of this is probably correct, in the current paper I seek to point to the limitations of the idea of “better than” that stem from the very nature of human self-understanding and experience.

Preliminary clarifications

Before we begin we need to clarify a few notions. First, what is an algorithm? While a narrow definition describes a series of lines of code that receives an input and issues an output, clearly this is not what is intended here. Critical algorithm studies have long made clear that the algorithm is not merely the source code of a specific computational process (see Bucher [2018, pp. 19–40] for elaboration, or Finn’s perspective [2017 pp. 15–56] of algorithms as “culture machines” [p. 16]). “Algorithm” or “algorithmic” in what follows is shorthand for a much broader definition, a technological-cultural-social-economic agent, a vast knowledge machine based on the collection of large amounts of data, abstracted from their initial source and analyzed with machine learning algorithms (here in the narrow sense) to produce certain classifications, predictions, profiles and clusterings which are then put to use for various commercial and other needs by corporations and institutions.

The panoply of possibilities, merely in the random series of references mentioned above, regarding who exactly will be in possession of the superior knowledge, is already sufficient evidence that the problem is not with a series of code instructions. Is it “technology,” “smart apps and AI-bots,” “non-conscious but highly intelligent algorithms” (vague technological entities using AI)? Is it “Google”/“Facebook”/“corporations and governments” (major social/political/economic players that develop and use algorithms), is it “big data” (referring to the large-scale collection and analysis of data) or “recommender systems” (a specific kind of software/service with commercial goals)? We are clearly dealing with a whole that includes technological, human, cultural, ideological and commercial elements. In saying that algorithms “know” I am condensing a complicated, multi-actor process, involving the corporations and other institutions that collect the data, deploy the algorithms, and make use of the results, the programmers and data scientists who develop and maintain the algorithms and analyze the results, as well as other human actors who translate algorithmic results into human terms (Kotliar 2020), the data-mining process, and the wider system within which massive amounts of data are extracted from

³ The two forms of “thinking” are of course fundamentally different in their very way of functioning. The comparison here concerns only comprehensiveness and bias.



our many daily activities online and offline, to be stored, sold and analyzed. This vast matrix of activity includes the legal infrastructure allowing for the use of this data, the interaction of various algorithms (Facebook or Netflix are not just one algorithm), the data feedback the algorithms receive, the potential knowledge that can be culled from their results and the commercial products that use these results. Also included in the definition is the broader ecology of which all of this is part (legal, sociocultural and economic) and the ideology of computation behind the scenes.

As a result of this wider perspective, the idea of “algorithmic knowledge,” knowledge formed at the interface of so many factors, is itself wide-ranging. It is knowledge that has no subject. It can be classificatory (a person is deemed to be of a certain kind) and is mostly aimed towards its use in decision-making and more concretely for prediction. Limiting ourselves to the technical side, and beyond the data about past activities itself, that is in a sense more solid, there is the problem that the algorithmic knowledge is unknown: It is impossible to make sense of what happens “inside,” what a certain inner layer of a deep neural network algorithm may “represent” (see Fazi 2021, p. 59). In addition, the clusters formed and patterns discerned have no definable human meaning. The data of the initial training set and the data collected later become coefficients of the various connections between different layers and nodes of the algorithm. And again, the result for the most part is the production of a certain classification, clustering, and an associated prediction. Often, in practical uses, there is a retranslation of algorithmic knowledge (a certain cluster or pattern) to human knowledge (“this is a trustworthy worker”) (see Kotliar 2020).

The idea of “knowing” in the statements discussed above is therefore stretched to the extreme and condenses a lot. And yet it is a commonplace in both the tech industry and critical research to speak of knowledge, and not only of what has already happened, but of future actions and behaviour. More important is what this knowledge is made of and aimed at. This is a knowledge based on data and information-processing and designed for prediction. It is knowledge that does not pass through the human mind. The data collected is not mediated by human consciousness, and “bypasses” human meaning-making processes (Rouvroy 2013, pp. 143). Rouvroy describes this as “data behaviorism”: a “new way of producing knowledge about future preferences attitudes, behaviours or events without considering the subject’s psychological motivations, speeches or narratives, but rather relying on data” (p. 143).⁴

⁴ The results and uses of algorithms occur in a human context; they are embedded in culture and are the objects of cultural imagination. Research has addressed the “algorithmic imaginary” of users (Bucher 2018), or folk theories (Ytre-Arne and Hallvard 2021) about algorithms, fantasies relating to algorithmic objectivity and purity (Hong, pp 18-19), “faith in computation” (Finn 2017 p. 9) and the “data imaginary” surrounding data and its widescale use (Beer 2019, pp. 14–33), that is, “the type of power we imagine data to hold” (14), an imaginary that is “cultivated and fostered as part of a set of commercial agendas [...] founding the legitimacy of data-led practices” (19). The current paper adds to that discussion by examining a certain way of imagining what algorithms can do, what is within their power, and what is their place with regard to humans; it looks at the idea of algorithms as all-knowing and as able, as it were, to x-ray the interior lives of human beings. But it cannot be subsumed in the perspective. The point is not just to discuss what journalists, developers, corporations or critics think algorithms can do, but what they really can do. I point to a mismatch between a certain image of their powers (to know people) and the actual impossibility of knowing humans to the desired and fantasized extent.



Subjective “knowledge”

Knowing humans

The main reason that algorithmic knowledge cannot be measured against our self-understanding is that it is of a necessarily different kind. To explain, let me try and briefly describe, in a very coarse manner, this uniqueness of self-understanding. I will take it as a challenge to do it without jargon and without leaning on any particular philosophical system; in this way I hope to maintain a level of discussion general enough to demonstrate the essential elements that the algorithmic perspective cannot capture, rather than to delve into a meticulous and painstaking discussion of this or that philosophical view.⁵

To begin with, the term [self-] “knowledge” here is inexact and one should talk instead of an ongoing process of self-understanding, the capacity to direct attention to oneself and the attempt to make sense of one’s experience, process and interpret it. It is an ongoing directedness towards ourselves rather than some definite end-product. “Knowledge,” in what follows should be read in that sense.

To simplify, and for the sake of clarity, let me present two aspects of the uniqueness of self-understanding, although they are actually one and the same. One would be concern or engagement and the other reflexivity.

For us the “data” of perception are never data. We experience what happens in the first person as happening to us. We feel it as part of concern, of a kind of engagement with the world. We care about things (if not actually in the sense of having empathy for them)—we are engaged with the world and therefore nothing appears or happens to us without this subjective involvement.

Perceptions and experiences are integrated into our psychic life. They diffuse, as it were, into our minds as a whole, connecting associatively and affectively to prior experiences. They are interpreted both immediately and retrospectively, perceived in

⁵ For elements of this description, see the following (more on endnotes 6 and 8). On the nature of experience, see Merleau-Ponty (e.g. 1945, pp. 240–281) (on engagement, p. 247). On irreducible reflexivity, see Sartre (1943 pp. 190–192) and in general, on the element of self-relatedness, his notion of prereflective self-consciousness (e.g. pp. 19–20) (referring to a more immediate reflexivity, rather than the second-order, higher reflexivity I also mention). On the difference of self-understanding from any other form of knowledge (the first- as opposed to third-person attitude) see Moran: “Whatever knowledge of oneself may be, it is a very different thing from the knowledge of others” (2001 p. 31). For the immediacy of some sense of self-knowledge as “a fundamental form of self-apprehension” and also for the authority of self-knowledge as opposed to others’ knowledge of you, see Moran (pp. 29–31; 9–11). Moran also refers to how such immediacy is constitutive of being a person (p. 30). About the irreducible subjective character of experience, see Nagel. Nagel reminds us that “every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view” (p. 437). “It appears unlikely”, he argues, “that we will get closer to the real nature of human experience by leaving behind the particularity of our human point of view” (p. 444) (but see endnote 7).

For a more analytic review of philosophical approaches to the problem of self-knowledge see Jopling (2000, pp. 1–19) and Renz (2017, pp. 1–18). For a perspective from cognitive psychology, see Neisser (1993, pp. 3–19).



certain ways based on prior experiences and contents of our psychic world, and in general, become psychically metabolized.

Because of this internal processing—not necessarily conscious—of our perceptions, sensations, thoughts and feelings, they are never what they “objectively” are, if such a thing even exists. “Experience” results from an ongoing transmutation of life as it happens, a kind of constant metabolization inside of everything lived through.

A second aspect of our self-understanding is reflexivity (again, separated only for the sake of clarity). Unlike with any other knowledge of objects in the world we know from the outside, we are here at once the knower and the known: the object and the subject of knowing in one and the same cognitive act. We look at ourselves from within, we are involved. We have some degree of immediacy to ourselves. The very characteristic that could seem a serious limitation to self-knowledge (it is not objective and remote) is what makes it unique and inimitable.

We are engaged, then, with whatever happens to us and with whatever we do, and have a constant relation to ourselves. This relation can be at the level of immediate experience (to be conscious of oneself) or at higher level of reflective thought, of self-understanding as a project. We have a relationality to our experience, our thoughts and our feelings.

Reflexive self-understanding is not like any other form of knowledge in the sense that it has a defining and creative aspect. This internal gaze directed at ourselves is a structural, constitutive element in our subjectivity. To an extent it could even be said to be what forms subjectivity: “The ‘I,’” writes Butler (1997), “is defined by this capacity for reflective self-relation or reflexivity” (p. 22).

The uniqueness of self-understanding is not in its totality or range, or in being all-encompassing and very detailed, but in its reflexivity: it is the directing of the gaze to oneself that makes it unique. It is a particular kind of relation, a relationality.

Lastly, we could mention a certain unending dynamic aspect, where things keep being examined and are actively at work inside us. Psychic contents are hardly ever at rest, and the examination and reinterpretation of them continues whenever they are evoked. There is therefore no bottom-line knowledge of ourselves but more like an ongoing dynamic interpretative process, conscious and unconscious.⁶

Where does algorithmic knowledge stand with regard to this? While algorithms can know certain things earlier than we do, this knowledge is of a different kind than our self-knowledge. Algorithmic knowledge of humans is based on data but the data does not address the subjective aspect of experience.⁷ Data gathering can be linked

⁶ One need not adopt an individual/self-centered approach to see the irreducibility of the relation to oneself and the first person basis of experience. The same holds even if we adopt a more situated and contextual approach to the self. Self-understanding is indeed not isolated. It is woven with an understanding of the environment and of others. Neisser presents “the ecological self” and “the interpersonal self” (1993, pp. 3–19): in both, one co-perceives oneself as part of perceiving the environment (the ecological self) or other people (the interpersonal self) in an immediate way. In addition, understanding oneself involves an understanding of the other and vice versa.

⁷ I do not just mean here what some philosophers have called qualia, the “what it’s like” of experience but the whole fact of being engaged with the world, which gives a certain quality to experience. I refer also to how things are experienced in relation to our psychic life, and to our reflexive approach to them.



to the reality we perceive, to our actions, small or big, to bodily changes and even, as emotional AI develops, to a proxy sampling of inner states. Algorithmic knowledge is built on a mass of such data and can be quite encompassing. It can monitor us almost totally and be based on a log of everything that goes on; but it fails to know the engagement and concern we have with ourselves and our experience and cannot access the gaze we turn on ourselves. It cannot sense that inner perspective on ourselves.

More generally, the data gaze necessarily remains external (“what algorithms cannot gauge,” writes Fisher, “is precisely the critical, reflexive events that take place in” “the private space of subjectivity,” 2022, p. 21). It can fathom what goes on psychically only indirectly and when something is actualized in the world (as an action that leaves a trace, or as a bodily change of the kind empathetic AI is so eager to home in on), only when it becomes data. (“What one types, clicks, likes, links [...] is socially configured as both an analog for and the totality of what one thinks, who one is, and what one will do. [...] any element that falls outside of this regime of positively measurable action [...] is denied existence.” Franklin 2015, p. 81. See also Rouvroy 2016, p. 47).

The algorithms know us well but we have access (rather, we *are* that access) to a dimension that cannot be known from outside. Perhaps, one could argue, this dimension is unimportant, but, be that as it may, it cannot be taken in by the algorithmic gaze. This means that algorithmic knowledge is always, inevitably, one step behind, even if in other aspects it can be far ahead of us, in knowing, for example, things that we hide from ourselves and that become clear only when a lot of data is collected and a pattern emerges. What it cannot do is imitate or exceed self-knowledge that is by nature inaccessible to it.⁸

Algorithmic knowledge is based on data about you, lots and lots of it. What it does not pick up is the meaning of experience for you, meaning that comes from engagement and reflexivity. This matters as, in us, the very meaning of sense “data” is transformed, from its first reception as experience, and then in our processing and examination of it. We are shaped and reshaped by meaning all the time. The algorithm may well know what you did and what entered the perceptual system

⁸ The promise of algorithms to know us better is never explained analytically in any of the examples I looked at above, given that it is, as I emphasize here, more of an ideological belief. But if we do adopt an analytic perspective on what “knowing ourselves” can mean, we come to a similar conclusion about the difficulty of that promise. Renz (2017) distinguishes four types of self-knowledge: “Self knowledge of one’s actual states” (sensations, perceptions, passing thoughts); “self knowledge of one’s standing attitudes” (beliefs, values, aims, preferences, intentions); “self knowledge of one’s dispositional properties” (character traits, capacities); and “self knowledge of one’s being subject to the human condition” (9–11). With the first two, our acquaintance is “irreducibly first-personal” (9) and we enjoy an “epistemic privilege bound up with the first-person perspective” (10) (Moran 2001, pp. 9–11, would agree). The third type is more likely gleaned from external information about us such as the views of others. Jopling (2000), too, lays stress on the common trait in salient traditions of self-knowledge, namely that “the self [...] must confront itself by itself (pp. 15–16). Algorithmic knowledge, I have argued, cannot access that level of self knowledge. Similarly, in Neisser’s analysis of self-knowledge and aspects of the self (1993, pp. 3–5), the ecological self, the interpersonal self, the remembered (temporally extended) self, and the private self all include this unavoidable first-person perspective/experience. Only what he terms the conceptual self is more available to algorithms.



(not quite what you perceived though) but not how it was experienced (except very coarsely: anxiety, joy, and so on), or the nature of one's secondary reflection on what happened. It does not know how what was perceived or done is connected to some or other associative—conscious or unconscious—network, or in which context of our inner life it is then reworked, what kind of object relations, in psychoanalytic terms, are affected by it and affect it, how it is seen by you, and in short, what it is for you, what the meaning for you is of what occurred or was done. The knowledge of oneself, by contrast, includes that meaning.

One can know, using data, something about subjects, a lot perhaps (and often knowledge that is not available to them, for example that a certain health condition is evolving), but never know their own knowledge, exactly because being a subject is being more than oneself, more than what simply happens to oneself or what one does. And this more, this little extra mile, is what remains outside the scope of algorithmic curiosity. As such it serves as the envied object of algorithmic method, a goal to conquer and replace.

The odd thing is that the promoters of algorithmic data-led knowledge, in believing an algorithm can know us better than we know ourselves, seem to completely ignore these levels of subjectivity. It seems that there is very little understanding of what subjectivity is in the prolific algorithmic efforts to map humans, along with an extreme reduction of it. It is not surprising then that when the human subject is reduced to the overall data he or she produces, to external sampling and measurement (the attempt, that is, to know a person through objective traces left in the world, traces that do not pass through one's consciousness of meaning-making. See again Rouvroy 2013. p. 143), one can boast about how a technology that samples often and across many parameters has the capacity to understand the subject better than they understand themselves.⁹

⁹ A word of caution is due here. It is important to stress that my concern is not with an idea of subjectivity whereby the rational, independent and critical subject acts rationally and with maximal knowledge and judgment. In his critique of algorithms and subjectivity, Fisher (2022), basing himself in part on Habermas (pp. 14–18), refers to this idea of subjectivity, formed during the Enlightenment, as being “excluded” by algorithmic knowledge and brings convincing examples of how judgment and political deliberation are “bypassed” by the algorithmic environment (pp. 1, 12, 64). Regarding that subject, characterized by its rational, independent judgment or critical abilities, I would definitely concede that algorithms do know better; but this is simply because that image has never been true. Even if this myth has had an important role in Western society for at least 200 years, in shaping, for example, a view of the political sphere, it was always only a construction, while actual humans are much less deliberative and more manipulable and influenceable, and all the more so in the economy of desires and actions which are themselves shaped by the practices of consumer capitalism and which are pushed further by its offshoot “surveillance capitalism” (Zuboff).

I find it much more fruitful to examine algorithmic knowledge not with regard to the rational subject of the Enlightenment but with regard to the psychoanalytic subject, whose subjectivity consists not in knowing but in lack of knowing. Andrejevic recalls the psychoanalytic insight whereby “subjects are non-transparent to themselves, since they are characterized by an internal gap or split—that of the unconscious—which renders them non-self identical” (Andrejevic 2020 pp. 1–8). The subject that cannot be preempted and automated, unless it is bypassed, is for him psychoanalytic (see also pp. 133–164). Like Andrejevic, I think that as long as we limit ourselves to a mythical model of subjectivity we will find ourselves repeatedly caught short by just how well the algorithms can know the subject, work in its stead and “hack” it. What the algorithms can much less easily approach is the subject as formulated by psychoanalysis.



To conclude, the algorithmic understanding of humans necessarily remains incommensurable with self-knowledge. The algorithmic gaze can never know our knowledge of ourselves and can never, therefore, exceed it: it is of a structurally different kind. One's entire array of reflexive approaches to oneself enter one's functioning, one's being. The algorithm does not sense that. To try to sense below or behind the subject (by tracking eye movements, skin conductance, or mouse movements, or any other data) is to miss the subject. As Rouvroy explains, "algorithmic governmentality is without subject: it operates with infra-individual data and supra-individual patterns without, at any moment, calling the subject to account for himself" (2013, pp. 144–145; Reigeluth makes a similar point, 2017, pp. 3–4, 9–10). Algorithms can know, based on data and before we do, things *about* us, but this is different from knowing something which is part of self-knowledge, which really would be to *know us*. It is that attempt to compete with our own knowledge, where our own understanding is constitutive of who we are, that makes for a much more difficult challenge for algorithms. Let's look at an example.

Reflexivity, I claimed, is constitutive of what can be termed subjectivity. To pursue this point and illustrate further the difficulty with the attempt to compare algorithmic knowledge of humans to their self-understanding (and to regard the former as better), let's look at the example Harari gives in warning against the dangerous potential of algorithmic knowledge of humans (Harari et al. 2018). Harari mentions his realization at age 21 that he was homosexual, and asks what if an algorithm, tracking his eye movements, say, could have inferred long before, perhaps at 14, that he was gay from how his eyes focused on men rather than women on the beach. Such an algorithm could have known he was gay before he himself did. He often asks himself, he tells us, what he thought he was doing in those adolescent years, pretending to be straight and not knowing he was gay. How could he not know *that* about himself?

So the algorithm is dangerous because it might know us better than we know ourselves, it might know one is gay before one does, or, to put it differently, it might know that gay is what one "really" is, contrary to one's own belief that one is straight. This information can, Harari warns, be used to "hack" the subject as it were, knowing things about them that they themselves are unaware of. His concern, however, seems to ignore a major point. Harari's homosexuality at 14, given by a sophisticated device tracking his eye movements which could detect his frequent focus on men—the gayness that the algorithm could "know"—is not the gayness of a 21-year-old Harari who comes to think of himself as gay. The algorithm knows something better, in the sense of earlier, but this something has nothing to do with knowing Harari as gay or knowing Harari at all. The knowledge Harari himself forms of himself, later, belatedly, through possible resistances, crises and difficulties or joys and liberation—of course I do not know, I invent here—is of a different kind. It is (self-)knowledge that is essential to what constitutes him as gay. Harari's understanding of himself as gay, whenever that understanding was formed and in



whatever way it was qualified, has a completely different role in his homosexuality than machine-inference from his eye-movements.¹⁰

When a subject comes to know themselves as gay, it has a constitutive role in their subjectivity. This coming to know, as well as any array of concerns around it, is essential to what makes—constitutes—one as gay. The point is not that one is what one takes oneself to be, but that whatever psychic work is done around one's homosexuality is essential to what that homosexuality is. The process of self-understanding cannot be bypassed. Algorithmic knowledge is not at any way here better than the subject's knowledge of themselves. The algorithm may well “predict” that one is gay beforehand, can make that classification, but it will never be a knowledge of gayness in the same way one understands oneself as gay.

Human realities are also flexible and in a flux, and to know a subject is not to know some immutable essence waiting to be discovered. Harari did not have “gay” inscribed on his forehead to be discovered by the algorithm. The degree of complexity lost in this reduction is absurd. Self-understanding of sexual orientation involves forces that operate in various directions, multiple layers of our personality, social influences and psychic forces. What we are, what we want and how we conceive of ourselves is shaped in a gentle or abrupt, gradual or quantum-leaping, ongoing process that involves the active participation of wounds, contingencies, perspectives and experience.

Why compare to those clueless humans anyway?

Let us approach from another angle. A strange dynamic is involved in the claim that algorithms know us better than we know ourselves, in that the very claim for the excellence of the algorithm presupposes the quality of the self-knowledge that is depreciated. Algorithmic knowledge, we are told, works on data, and it is therefore objective and reliable; it is not biased by omissions or wilful perceptions, it is not handicapped by short sight due to insufficient information, and is not conditioned by pre-existing theory (for a list of how data analytics companies promote data products see Beer 2019, pp. 21–31). And yet at its highest achievement and most excellent level, it manages to overcome... what? Human self-knowledge, which is considered full of holes, biased, error-prone and inferior. In comparing though, it is as if the promoters of algorithmic knowledge actually admit that this other knowledge, human self-knowledge, is not as despised as they claim but rather constitutes a criterion and a major challenge to outdo.

The basic perspective of both the data-led tech world and the neoliberal frame within which it works is actually that people fail to know themselves. Zuboff shows how “surveillance capitalism” (the ideological, social and economic worldview in

¹⁰ Harari the person or his own homosexuality are of course not the issue here; I merely take up the personal example he gives. Harari's intention is to warn us against the algorithmic manipulation of humans, and as someone who practices meditation and promotes self-understanding as an antidote to algorithmic practices of knowing us he obviously holds a complex view of the human (Harari et al. 2018).



which algorithmic knowledge is pursued), constantly seeking to extract people's data, is built on the premise (she terms it "instrumentarian power") that humans are manipulable, and their behaviour can be engineered (Zuboff 2019 pp. 376–9).¹¹

The behaviouristic psychology on which the entire algorithmic project of influencing people is built sees people as fundamentally not knowing themselves.¹² Skinner's notorious position on the belief in personal freedom is summarized by Zuboff, as she shows how his thinking grounds the algorithmic worldview:

Under the behaviorist's gaze, these lacunae of ignorance that we mistake for free will are queuing up for explanation [...] [C]onceptual confusion [...] cloaks our deepest ignorance in the sacred robes of freedom and dignity. Skinner argued that our allegiance to these lofty notions is simply the way we protect ourselves from the hard truths of "unsuspected controlling relations between behavior and environment." (Zuboff 2019 pp. 367–368)

Humans indeed do not know themselves very well. They are not aware of what makes them act, and indeed, they can very easily be manipulated. Place the chocolate bar stand by the register and people will buy them even when they did not intend to. The marketing world, behavioural economics, empirical psychology and the entire capitalistic infrastructure of the algorithmic world embody this very well. Nudging, curating, recommending, promoting, influencing: all the capitalist algorithmic techniques hone in on exactly this human weakness to influence us behind our backs, so to speak, to mislead us, channel us, and shape our consumerist behaviour.

Why compare then, if human knowledge is supposedly so inferior, so frail and humans are so easy to manipulate? If the claim is actually that people fail to understand themselves and the algorithms should be given the task instead, what's the big deal in showing that algorithmic knowledge outperforms human self-knowledge?

Humans are a unique object of knowledge, an "object" that knows itself: a subject. They have a reflexive relation to themselves, and it is essential to knowing what they are. Only humans need to be known so far as to include that element too. It is

¹¹ One should remember that so-called predictive algorithms not only predict what people want, but actively contribute to shaping the world they predict. By deciding what a certain person wants or is, they help fashion and curate a field of possibilities for that person, and this further promotes behaviour that will bring the actual individual closer to the prediction. They also have a hand in shaping the more general economy in which desires are formed.

¹² Surveillance capitalism, Zuboff writes, "reimagines us" through the lens of what she calls "instrumentarianism": "the instrumentation [...] of behavior for the purposes of modification, prediction, monetization, and control" (Zuboff 2019, p. 352). This form of power has its origins, she argues, in B. F. Skinner's "radical behaviourism" [...] and his dream of a 'technology of behavior'" (p. 353). The idea is that human behavior "could be effectively predicted and shaped" (p. 369) by being manipulated completely from the outside by controlling external conditions. There is no real free will and there is no need to look inside the person (pp. 366–369). "The human being is recast as an 'it,' [...] an 'organism among organisms'" (p. 363). Following that tradition, instrumentarianism "is profoundly and infinitely indifferent to our meanings and motives. Trained on measurable action, its only care is that whatever we do is accessible to its ever-evolving operations of rendition, calculation, modification, monetization, and control" (p. 360). (See also Franklin 2015, pp. 89–94 for a similar perspective on cybernetics as the origin of contemporary capitalism).



to the bottom of the well of the human that one should try, as it were, to descend to try and understand it, even if it is too deep to fully do so. The attempt to know people so well as to outdo precisely their own self-knowledge assumes their difference from other objects, which do not know themselves. Thus, a certain uniqueness of the human subject is posited in the very attempt to negate it.

On human knowing as failing to know

The uniqueness of self-knowledge is that one is necessarily implicated in knowledge about oneself. The inherent confines of self-knowledge arise from that very implication. One is involved and therefore necessarily fails to see right. This is the second characteristic of our self-knowledge I now wish to consider: not its subjective quality but rather its incompleteness and limitations. Self-knowledge is indeed not good knowledge and so aiming to know “better” than it, is really no big deal. Self-knowledge has a lot to do with not-knowing, with opacity to ourselves and with failing to know. We are not good at knowing ourselves and it is commonplace in some ways that we know ourselves less than do many around us.

In her brilliant *Giving an Account of Ourselves*, Judith Butler (2005) explains how the account we can give of ourselves inevitably remains partial. For Butler, “the subject is opaque to itself, not fully translucent and knowable to itself” (p. 19) and she lists several aspects of selfhood that are necessarily left out. The body is a major aspect of our being that almost by nature resists our understanding: “To be a body is [...] to be deprived of having a full recollection of one’s life” (p. 38). Butler focuses attention on the very discourse that to begin with frames what we can think and tell about: “Any discourse,” she writes following Foucault, “any regime of intelligibility, constitutes us at a cost. Our capacity to reflect upon ourselves, to tell the truth about ourselves, is correspondingly limited by what the discourse, the regime, cannot allow into speakability” (p. 121). Also outside the frame of self-knowledge are “norms that facilitate my telling about myself but that I do not author” (p. 39), the very structure of address (when you give an account to others), which is social and trans-individual by nature. Furthermore there are “primary relations, irrecoverable, that form lasting and recurrent impressions on the history of my life” and “a history that establishes my partial opacity to myself” (p. 39). Even eliminating the outward-facing nature of the account of oneself, that account remains not fully narratable, full of holes, and it cannot encompass who we are because we are not independent and isolated.

In fact, it is more correct to define our relation to ourselves as one of mis-knowledge than of knowledge. The knowledge we have of ourselves is thoroughly distorted (as Freud argues in the quote cited above). It is self-serving and egocentric. We magnify certain things and minimize others. We overlook things uncomfortable to us, we suffer biases, many of which are systematic and mislead us always in the same direction. We rarely understand the influence of social forces and contexts on us; how we are shaped within a certain contingent, historical, specific context and how so many of our choices, preferences and habits belong to our class, nation, milieu more than they do to us as individuals.



We are also ill-informed about our inner reality. This holds both at the level of a superficial unawareness of the reasons for our acts, an unawareness repeatedly indicated, as discussed earlier, by empirical psychology, as when we are driven to make a certain purchase without seeing that we were led there, as well as, more interestingly, at the level of a deeper failure to know ourselves of the kind psychoanalysis insists on: Much of what motivates us, influences us and is at work in us, goes the psychoanalytic perspective, happens outside our awareness. The unconscious is not some passive reservoir of repressed emotions but a thinking, feeling, psychic reality that processes external and inner reality and negotiates with it without conscious involvement.¹³ So when we try to understand ourselves we often do so blind to what truly motivates us, causes us to act, and governs our perspective on reality. We do so without knowing some of the sources of our emotions, the inner workings of our psyche, or of the struggles within us. We do not know how our behaviour is being guided and our perception distorted by fantasies we do not know we have, and past scenarios and relationships that we reproduce (Lear 2005, p. 149). Then we interpret ourselves, and tell a story about ourselves, but too much of the real picture is shut out of the narrative.¹⁴

Moreover, it is important to note that not only are our unconscious desires, thoughts and perspectives part of us, but so are the mechanisms which impede us from acknowledging them, and which leave them to act on us and in us without any decision or agreement on our part. Psychoanalysis stresses that the ongoing failure of self-understanding is not passive but active: We invest a lot of effort in not knowing ourselves.¹⁵ Substantial psychic energy is put into hiding our desires and our thoughts from ourselves, disguising them, channelling them to other, neutral directions, and preventing them from erupting into consciousness. These mechanisms and active manners of failing to see ourselves are, in the psychoanalytic perspective, a substantial part of who we are.

So all in all, it seems that it is perhaps not that difficult to know us better than we know ourselves. But the picture gets complicated because of another factor: not only do we fail to know ourselves and the reality we are part of, but this very failure is also essential to who we, personally, are. It is precisely the absence of “proper” or “full” self-awareness, the specific biases of perception, the failure to perceive reality as it is, that most characterizes every human individual. If our knowledge was flawless we would all look at the world the same and see the same thing and yet we all see different things. We notice certain elements and miss others, focusing on one thing and misperceiving another, in a unique melange of selection and error.

Psychoanalysis teaches us that while we fail to know ourselves, this failure is essential to who we are. To put it in a catchy form, what characterizes us most, what

¹³ As is apparent already in Freud’s formulation of the concept (1915, p. 189). By “thinking,” to be clear, I am not referring to rational thought and to the way we believe we are normally thinking.

¹⁴ Psychoanalysis is of course not the first philosophy to observe human irrationality but it can be said to be the one that has most systematically mapped and explained it. For more on non-psychoanalytic traditions that emphasize failures of self-knowledge, see Jopling (2000 pp. 5–9).

¹⁵ In part through “defence mechanisms,” theoretically developed by ego-psychology.



is most definitive of what we are, is not what we know about ourselves but what we do not know and our ways of not knowing. Each of us is blind to other aspects of reality, each tends to see things in one's own particular way. Our uniqueness as subjects is rooted in this blindness and distortion, these biases and selections, the false enhancements and minimizations peculiar to us, the specific scenarios under which we perceive, and therefore misperceive the world and ourselves.

To conclude, our knowledge of ourselves is highly selective, distorted and biased but the biases and selections are essential to who we are. To try to understand humans without understanding this is to misunderstand them.

Consider for example the following: Your friends might know your tendency to have outbursts of rage, and they might even know quite well what circumstances tend to trigger them. Meanwhile you hardly acknowledge you have an anger issue. The point is that your denial of your angry feelings and behaviour is part of who you are. The anger outbursts are part of you but so is not knowing about them, along with the specific mechanisms developed not to know them. These mechanisms, the will to change that behaviour to the extent you are somewhat aware of it, the flickering awareness of the problem, are all part of you, as much as the gradual transformation of all these over time. Similarly, the manner with which you handle angry feelings, the causes, sources and consequences of the outbursts, the dynamics they lead you into, all these are woven into every aspect of your life. In her work on ambivalence, Hili Razinsky (2016) demonstrates how contradictory feelings and thoughts, knowing and not knowing, feeling a certain way and resisting it, loving and hating the same object, are fundamentally intertwined and cannot be known separately. One's judgment that X is good and that it is bad, like acknowledgement and lack of acknowledgement of a part of oneself, are bound together and create each other.

Or perhaps your co-workers know you are a narcissist—it is so blatant in fact—and yet you are blissfully unaware of it. More often than not the narcissist does not know he is one. Perhaps, if he really knew it, he would acknowledge a great need for self-love and its origin in feeling fundamentally unloved, or he would know how base, worthless and despicable he feels. But the narcissism is exactly his way not to feel that, not to connect to this very difficult experience, and to gain redress instead in excessive self-worth. The narcissism has a purpose and for this very reason the narcissist cannot know what the narcissism is supposed to redress: he cannot allow himself to know he is narcissist. To understand the narcissist without understanding his fundamental refusal to feel and connect to himself is not to know him.

Subjectivity then is centred around self-opacity. Those who develop and deploy algorithms do not of course want to imitate it. They strive to know best about us, to know us completely. They want to gather as much data as possible and analyze it in order to know the most about us. But given that what we are includes this extreme selectiveness of our gaze, the result is that the more the algorithms approach the ideal of knowing us completely, the farther they get from actually knowing us as we are, and it makes the project of knowing us paradoxical and impossible. We are opaque to ourselves, and attempting to understand us fully, based on the data about us that will supposedly reveal our inner secrets, is to fundamentally misrepresent what we are and in this sense is a profound failure to know.



To illustrate this point let me revisit a point from earlier research of mine. In a best-seller entitled *Total Recall* (2009), high-tech guru Gordon Bell and Jim Gemmell put forth a vision of personal data collection that is so thorough and complete that it can actually serve as a vicarious personal archive or memory. Everything one does, sees or hears, down to the tiniest elements, is documented and this exhaustive logging of life results in a kind of digital archive of one's memories. The authors enumerate the superiority of "e-memory" over biological memory with its biases, selections and distortion-laden "creativity." The automatic archive will be much more exact and objective. It will also be all-encompassing: rather than choosing some elements over others, selecting and preferring as humans do, the archive will record everything. Full and comprehensive, exact, and unbiased, e-memory will represent the subject far better than is done in one's own (weak, bio-) memory.

However, as I have argued, the very attempt to create an exact representation of the subject colossally fails as it leaves behind exactly the subject it was supposed to capture (L. Razinsky 2020, pp. 194–195). For if one's selections and biases in perception and memory, the creative refashioning of memorized events, are eliminated in favour of a kind of absolute, complete and flat memory (not over-noticing certain things nor under-noticing others), then this is no longer a representation of the subject since the subjective is exactly in those overs and unders, in bias and selection:

One's own perspective, one's own subjectivity as possibly expressed in one's preferences and choices, is left out of the digital representation of oneself. What one ends up with is an archive of oneself where one is left out. [...] The attempt to create an I – through exhaustively documenting it – for which nothing is negligible or insignificant, ends in a rejection of that very I. (Razinsky 2017, p. 161)¹⁶

The failures of self-perception belong to us. They are an essential part of how we are, constituting our unique spectacles through which we see and fail to see the world and ourselves. Dirty and opaque, distorting and deceiving, full of blind spots, these are *our* glasses.

Conclusion

To sum up, in the third part of the article I made a claim for an ongoing, engaging self-relation whereby everything that happens, happens in the first person, to us, is processed in the human interior, echoing in body and mind, and interacts with and is adapted to memories, prior experiences and affects. Experiences, thoughts and feelings are also constantly examined and reinterpreted, and this reflexive attention to ourselves is what defines subjectivity. Though constant relating to oneself is inimitable, it does not necessarily lead to good "knowledge": in the fourth part, I added that we perform very badly in what is supposedly our forté, namely to know ourselves.

¹⁶ In the case of *Total Recall* the issue is mostly with exhaustive data collection rather than with algorithmic processing but the point is the same.



It is both these elements—the gaze directed inward and its inherent limitation—that make up, together, the human subject.

Human beings, then, have a special relation to ourselves, a reflexive quality probably unique to our species. We know ourselves directly in experience and interpret ourselves, but at the same time we also fail to know ourselves. Neither of these aspects, the unique kind of knowledge we have of ourselves or the ways in which this knowledge is full of holes and biases in a way that is part of who we are, can be imitated in external, algorithmic knowledge about us.¹⁷ One cannot therefore know us better than we know ourselves, even if we barely know ourselves at all.

The question of where the idea that algorithms do, or will soon, know us better than we know ourselves comes from is too large to discuss here.¹⁸ Suffice to say that the entire presence of algorithms in our everyday, whether seen from the side of the corporations or from the point of view of the users who interact with them, use them and are used by them, is always more than pragmatic. We expect the algorithms to be able to tell us what to do, how to act; we trust them assuming they may indeed know, better than we do, what is good for us. We are less interested in generic recommendations of what to do, preferring personalized advice tailored to *us*, based on an understanding of who *we* are. The algorithm has watched us in so many situations, one thinks, it cannot be easily fooled, isn't biased, uses objective measures and is "smart." The idea the algorithm knows us better than we know ourselves is in a sense needed for our everyday reliance on it, even if this background assumption is not always spelled out.

It might be said that whenever we expect some input from AI, say a recommendation system to present us with something we are as yet unfamiliar with (a song, perhaps), we assume it knows something about us slightly beyond what we ourselves know. When we want more of what we like we assume the algorithm knows us well enough to supply us with it.

On an even more fantasmatic level, we feel the algorithm *knows*. It knows in ways we cannot, it knows profoundly and completely. We can lean on that knowledge, probably more than on knowledge concocted by humans, we think. This implies that knowledge can be achieved, that there is a truth to know and a foundation to build on (Just one component of the idea of algorithms knowing us better than we do is the promise that there is, in us, more to know than we do.) Applied to humans, the idea that this technology can know not only general things but know us, intimately, has great appeal. The belief it knows me better than I know myself is the edge of this fantasy, an idea, as I have tried to show, that has to remain on the fanatasmatic level.

Algorithmic knowledge, we saw, is knowledge of a different kind and excludes the essential element in that knowledge we have of ourselves. More broadly speaking, it can be said that the reduced and flattened view of the subject expressed in the idea that algorithms can know people better than they know themselves is a result of

¹⁷ So not only is there a general misrepresentation of the the essence of subjectivity, there is a particular misrepresentation of each individual subject when their own distortions are not taken into account.

¹⁸ This question is of course part of a larger one concerning the promises of algorithmic knowledge especially with regard to self-knowledge. For a discussion see Hong (pp. 13–29, 76–113).



a theoretical ignorance, and the troubling thing is that this ignorance is ideological. The data-led industry and the algorithmic gaze seem to look at a subject without any understanding of what it is to be a human being, drawing their view, at most, from behavioural psychology, cognitive and social experimental psychology, marketing and behavioural economics. But philosophical understanding, as well as an understanding coming from critical theory, is completely lacking and as a result the image formed is completely flat.

The tech industry could evoke the pragmatic argument and say that for any practical matter, knowing about people through their data is enough without the need to go into the complexities of the human mind or the social formation of the subject. And yet this would be only a partial answer, because the real project is not just economic-ideological, a further step in capitalism (Zuboff). There is a project of knowledge involved, namely the idea that the data-revolution can give us a deeper insight into the world and the possibility of knowing it better, to a degree impossible before. It is in the context of this deeper project of knowledge that I suggest understanding the comparison of algorithmic knowledge of people to their own self knowledge.

The project of knowledge is first of all a matter of curiosity and the will to know. Ed Finn, pointing to the encyclopaedic aspiration to exhaustive knowledge of the world, which he finds in the efforts of the large tech companies (2017, pp. 57–85), shows how such efforts to know the world in its entirety are bound up with intimate knowledge of human subjects. If we want the personal assistant to give us the right answer, it needs not only to know everything about the world but also everything about us. “Anticipation,” Finn argues, “requires building a machine that understands not just public culture but private intention” (2017, p. 73). “The twinned desires embedded in the heart of effective computability” he argues, are “the quest for universal knowledge and perfect self-knowledge” (p. 13).

But on another level, the project of knowledge is one of order and control, of rendering the world completely decipherable and readable, and as a result, controllable. In presenting this point, I’d like to also explain the larger stakes of this paper:

The bigger social-political-economic project behind the expansion of the information economy is one of full computation of the world, “the desire to make the world effectively calculable” (Finn 2017, p. 26). “Digitality,” Franklin writes, “promises to render the world legible, recordable, and knowable” (Franklin 2015, p. 19). Algorithms can automate, control and regulate that world smoothly, and they do so objectively and in the best way. To be run by algorithms, the world has to be understood by algorithms—to this end it is to be conceived as consisting of information alone, of data. This of course includes the human and social worlds “reconceptualized as information-processing systems” (Franklin 2015, p. xv).

The origins of the full computation project are in the early day of cybernetics, where the idea was that the human is like a computer and that it can essentially be black boxed: Franklin (pp. 89–105) shows how the leading thinkers of cybernetics methodically “black boxed” the inner lives of humans. “All interest in internal function is discarded in favour of measurable behaviors understood as input and output” (p. 91). Statistical analysis is enough and there is no need to understand internal elements (p. 92). Consciousness, particularly, is excluded (p. 93). This is part of what allows the equivalence of the human mind and the computer to be established,



where the mind is seen as an information-processing unit. Franklin sees this tendency as crucial to the contemporary information economy (p. xiv) or what he calls the “control-era” (p. 91), with its “conceptual as well as technical digitization of the world” (p. xxii). Tiqqun, the bad boy(s) of this line of critique, opts for the stronger formula: “Cybernetics is warfare directed against all that lives” (p. 32). This form of governmentality is “a negation of everything that escapes regulation [...] of all the behavioural fluctuations that ultimately would not follow from natural law” (p. 27).

But of course vision and reality don’t coincide. Humans are much less predictable than early cybernetic thinking assumed. The coarse prototype of a pilot’s movements explored by Norbert Wiener in an attempt to model and predict where the aircraft would be so that it could be shot down, doesn’t come close to the complexities of human behaviour which render predictability much more difficult. As a result, one needs to make the black box smaller. One needs to get much more granular and precise, and at the same time much more total in the collection of data. It means still looking from outside, at external data, but on a much smaller scale, looking at every micro-element (and collecting masses of data points reflecting such tiny elements). As if what is considered to be the outside is shrunk in order to look externally at smaller and smaller elements.

In parallel, there is the need to go inward: In today’s algorithmic governmentality, it is not enough to just collect data about every little action. The prediction this yields is never enough; humans are simply too complicated. Today, therefore, we are witnessing a kind of technological race to know our interiority, which Zuboff calls “machine invasion of human depth” (p. 255). From knowing one’s personality to knowing what one attends to, what one feels while watching a video, to the attempt with emotional AI to know feeling (Krum), and now to the use of electrodes or EEG to know what is, quite literally, on your mind. This is to go against the logic of the black box somewhat, as if there is an understanding that ultimately people cannot be black-boxed and need to be thoroughly known to be predicted; but this is still knowing according to their data, that is from the outside, rather than by their own human perspective.

Returning to the general endeavour of complete datafication and computation, we can say that what stands in its way is the human. As Rouvroy puts it, the aim of the algorithmic project is “the neutralization of all in life that is irreducible to calculation” (2016, p. 49, my translation). Algorithms, she argues, try to get around human subjectivity (2014, pp. 115–16) and “minimiz[e] the uncertainty associated with human agency” (2013, p. 155).

Now we can better see the stakes of claiming to know people better than they know themselves. Humans are the stubborn obstacle to a smooth, fully automated and controllable world. Andrejevic shows how the complexity of the human has to be flattened out, automated, for it to become object of algorithmic knowledge (pp. 1–8, 133–167). Remarking precisely on the idea of “better than,” he sees it as an attempt to automate the subject (p. 3), as part of the effort to automate society altogether and render it smooth and frictionless. To do what subjects do in their stead, algorithms have to overcome exactly “the moment of uncertainty, unpredictability inconsistency, or resistance posed by the figure of the subject” (pp. 2, 140), that is, to make humans more predictable. “The figure of the subject,” he adds, “poses an



obstacle for automation and thus a potential point of resistance to its logics and discourses” (p. 133). If the algorithms don’t know humans better than they know themselves, if humans indeed have an edge, I argue—if they aren’t fully decipherable but keep some of the opacity of their self-understanding—the larger computational view of the world fails.

If one reason to try to outdo human self-knowledge is to understand the human completely, thereby neutralizing the human freedom to act and interpret the world on some basis other than algorithmic, thus eliminating any possible source of uncertainty for the algorithmic perspective, any hidden unknown areas, a second reason is a kind of competition: If the algorithms can outdo humans in understanding humans, their victory is complete, they can become the sole means for understanding, interpreting and influencing the world. One can make do with the computational worldview alone. Humans no longer need to understand the world in their human way since even the thing they most understand, themselves, is better understood by algorithms. Full computation is possible since there is no other agent that can make sense of the world or be the source of valid knowledge of it. In this second option the human is not an undecipherable opacity but a rival, as it were, offering a human understanding of the world. If the rival can itself be thoroughly known, its perspective on all the rest loses its strength and claim to superiority. In both these directions then, understanding humans better than they understand themselves means the world as a whole can be fully understood, predicted, and controlled through an algorithmic computational perspective alone.

Human self-understanding is the opponent to be overcome. It is that knowledge that algorithmic knowledge seeks to replace, and against which it competes. And it is this knowledge which, fundamentally, it does not understand.

Acknowledgements I wish to thank the wonderful researchers of my research group, *Subjectivity and Data*, Yoav Ronel, Omri Blum, Shira Shemi and Asaf Hazan for brainstorming and for painful, helpful feedback, and to Mark Joseph, Yaron Wolf, Robert Rosenberger and most of all Hili Razinsky for helping push ideas further.

References

- Andrejevic, M. 2020. *Automated Media*. New York: Routledge.
- Beer, D. 2019. *The Data Gaze: Capitalism, Power and Perception*. Los Angeles: Sage.
- Bell, G., and J. Gemmill. 2009. *Total Recall: How the E-memory Revolution Will Change Everything*. New York: Dutton.
- Bucher, T. 2018. *If Then: Algorithmic Power and Politics*. Oxford: Oxford UP.
- Butler, J. 1997. *The Psychic Life of Power. Theories in Subjection*. Stanford: Stanford UP.
- Butler, J. 2005. *Giving an Account of Oneself*. New York: Fordham University Press.
- Cadwalladr, C. 2014. Are the robots about to rise? Google’s new director of engineering thinks so... Interview with Ray Kurzweil. *The Guardian*. 22 Feb 2014, <https://www.theguardian.com/technology/2014/feb/22/robots-google-ray-kurzweil-terminator-singularity-artificial-intelligence>.
- Carmichael, J. 2014. Google Knows You Better Than You Know Yourself. *The Atlantic*. August 19, <https://www.theatlantic.com/technology/archive/2014/08/google-knows-you-better-than-you-know-yourself/378608/>.
- Crum, P. 2018. Technology that knows what you’re feeling. *TED*, https://www.ted.com/talks/poppy_crum_technology_that_knows_what_you_re_feeling/transcript.



- Evans, J. 2015 When Facebook Knows You Better Than You Know Yourself. *TechCrunch+*, October 24, 2015, <https://techcrunch.com/2015/10/24/when-facebook-knows-you-better-than-you-know-yourself/>.
- Fazi, M.B. 2021. Beyond Human: Deep Learning, Explainability and Representation. *Theory, Culture & Society* 38 (7–8): 55–77.
- Finn, E. 2017. *What Algorithms Want: Imagination in the Age of Computing*. Cambridge MA: MIT Press.
- Fisher, E. 2022. *Algorithms and Subjectivity. The Subversion of Critical Knowledge*. London: Routledge.
- Franklin, S. 2015. *Control. Digitality as Cultural Logic*. Cambridge: The MIT Press.
- Freud, S. 1915. The Unconscious. In *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, vol. XIV, ed. J. Strachey, 159–215. London, UK: Hogarth Press.
- Freud, S. 1917. A Difficulty in the Path of Psycho-Analysis. In *The Standard Edition of...*, Volume XVII, London: Hogarth Press, pp. 135–44.
- Gama Rose, Ch. D. 2021. Recommender Systems—Know your users better than they know themselves Medium, June 27, <https://prianjali98.medium.com/recommender-systems-know-your-users-better-than-they-know-themselves-4568eef3d4ad>.
- Harari, Y.N. 2017. *Homo Deus: A brief History of Tomorrow*. New York: Harper.
- Harari, Y. N., Harris, T. and Thompson N. 2018. When Tech Knows You Better Than You Know Yourself. *Wired*. OCT 4 <https://www.wired.com/story/artificial-intelligence-yuval-noah-harari-tristan-harris/>.
- Hong, S.-H. 2020. *Technologies of Speculation. The Limits of Knowledge in a Data-Driven Society*. New York: New York UP.
- Jenkins Jr., H. W. 2010. Google and the Search for the Future, Wall Street Journal, August 14, <http://www.wsj.com/articles/SB10001424052748704901104575423294099527212>
- Jopling, D.A. 2000. *Self-Knowledge and the Self*. New York: Routledge.
- Koksal, I. 2018. Artificial Intelligence May Know You Better Than You Know Yourself. *Forbes*, Feb 27, <https://www.forbes.com/sites/ilkerkoksal/2018/02/27/artificial-intelligence-may-know-you-better-than-you-know-yourself/?sh=a04cb574058a>.
- Kotliar, D.M. 2020. The Return of the Social: Algorithmic Identity in an Age of Symbolic Demise. *New Media & Society* 22 (7): 1152–1167.
- Lear, J. 2005. *Freud*. New York: Routledge.
- McCarthy-Jones, S. 2020. Artificial intelligence is a totalitarian’s dream – here’s how to take power back. *The conversation*, August 12, <https://theconversation.com/artificial-intelligence-is-a-totalitarians-dream-heres-how-to-take-power-back-143722>.
- Merleau-Ponty, M. 1945. *Phénoménologie de la perception*. Gallimard.
- Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton UP: Princeton.
- Neisser, U. 1993. *The Perceived Self: Ecological and Interpersonal Sources of Self-Knowledge*. Cambridge: Cambridge University Press.
- Razinsky, H. 2016. *Ambivalence: A Philosophical Exploration*. London NY: Rowman & Littlefield.
- Razinsky, L. 2017. Everything: Totality and Self-Representation, from Past to Present”. *SubStance* 46 (3): 150–172.
- Razinsky, L. 2020. The Dream of Absolute Memory: On Digital Self-Representation. *LIT* 31 (2): 182–201.
- Reigeluth, T. 2017. Recommender Systems as Techniques of the Self? *Le Foucauldien* 3 (1): 1–25.
- Renz, U. 2017. *Self-Knowledge. A History*. Oxford: Oxford UP.
- Rouvroy, A. 2013. The End(s) of Critique. Data Behaviourism Versus Due Process. In *Privacy, Due Process and the Computational Turn*, ed. M. Hildebrandt and K. de Vries, 143–167. London NY: Routledge.
- Rouvroy, A. 2015. Intervention. in La Régime de vérité numérique. *Débat Entre Antoinette Rouvroy Et Bernard Stiegler*, Socio 4: 105–132.
- Rouvroy, A. 2016. L’art de ne pas changer le monde (Interview). *La Revue Nouvelle* 8: 44–50.
- Sartre, J.-P. 1943. *L’Être et le Néant*. Gallimard.
- Shaw, C. M. 2016. Big data knows you better than you know yourself. *The New American*, August 10 <https://thenewamerican.com/big-data-knows-you-better-than-you-know-yourself/>.
- Thompson, D. 2010. “Google’s CEO: ‘The Laws Are Written by Lobbyists,’” *The Atlantic*, October 1 <https://www.theatlantic.com/technology/archive/2010/10/googles-ceo-the-laws-are-written-by-lobbyists/63908/>.
- Ytre-Arne, B., and M. Hallvard. 2021. Folk Theories of Algorithms: Understanding Digital Irritation. *Media, Culture & Society* 43 (5): 807–824.



Zuboff, Sh. 2019. *The Age of Surveillance Capitalism*. New York: PublicAffairs.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Liran Razinsky is an associate professor in the Program for Hermeneutics and Cultural Studies at Bar Ilan University, Israel, working on psychoanalysis, literature, critical theory and the critique of digital technology. He has published *Freud, Psychoanalysis and Death* (Cambridge UP, 2013), co-edited *Writing the Holocaust Today* (Rodopi, 2012), and edited a special issue of *YFS* on Jorge Semprun, as well as published papers on digital memory. His current project, *The Human Being in the Age of Data*, investigates how human subjectivity changes with the rise of data-led algorithmic understanding of individuals, and how algorithmic knowledge interacts with human-centered worlds of meaning.

