



## UvA-DARE (Digital Academic Repository)

### Mark of the Human: On the Concept of the Digital Human Being

Roessler, B.

**DOI**

[10.21552/edpl/2021/2/5](https://doi.org/10.21552/edpl/2021/2/5)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

European Data Protection Law Review

**License**

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care>)

[Link to publication](#)

**Citation for published version (APA):**

Roessler, B. (2021). Mark of the Human: On the Concept of the Digital Human Being. *European Data Protection Law Review*, 7(2), 157-160. <https://doi.org/10.21552/edpl/2021/2/5>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# Mark of the Human: On the Concept of the Digital Human Being

Beate Roessler\*

Ian McEwan's novel *Machines Like Me*<sup>1</sup> is set in an alternative 1982 United Kingdom: The war against the Falklands has been lost, the miners' strike is still on, unemployment is rising by the day, John Lennon as well as John F Kennedy are still alive – and above all, so is Alan Turing ('Sir Alan Turing, war hero and presiding genius of the digital age'). Turing has been working successfully on AI and the construction of a robot, and the first set of these robots is on sale: 12 Adams and 13 Eves as they have been subtly called. The protagonist, Charles Friend, spends the little inheritance he received after the death of his mother on buying one of them, and since he is too late for an Eve, he gets an Adam. The plot of the novel has different threads: there is the relationship with Miranda, Charlie's upstairs neighbor who he fell in love with long ago and who he starts dating. Miranda, after some time, has an affair with Adam; furthermore, she herself not only has a difficult personal history which she lies about and which is only revealed little by little, leading to the terrible unfolding, in which Charlie kills Adam. This thread in the complicated plot is important because it forces Miranda and Charles to lie – and after Adam has found out about this piece in Miranda's past, he intends to inform the police, since, as a robot he can't lie. He wants to be, he has to be, relentlessly upright. Also, rather uncannily, in the last third of the novel an increasing number of suicides by some of the Adams and Eves is being reported. But the main plot is simple: Charlie buys Adam, programs him together with Miranda, develops a rather friendly relationships with him and in the end kills him.

There are four points of interest here:

- The idea and the process of *programming Adam*. With the robots, there comes a 470-page online handbook about how to program them, but Charlie writes: 'I couldn't think of myself as Adam's 'user': I'd assumed there was nothing to learn about him that he could not teach me himself. But the manual in my hands had fallen open at chapter 14. Here, the English was plain: preferences; personality parameters. Then a set of headings – Agreeableness. Extraversion. Openness to experience. Consciousness. Emotional stability. [...] Glancing at the next page I saw that I was supposed to select various settings on a scale of one to ten.' Charles feels uncomfortable to choose the settings since he is very well aware of their reductive character.

---

DOI: 10.21552/edpl/2021/2/5

\* University of Amsterdam. For Correspondence: <B.Roessler@uva.nl>

1 Ian McEwan, *Machines Like Me* (Jonathan Cape, 2019)

- The problem of self-knowledge and decision-making, and Turing declares at the end of the novel: ‘I think the A-and-Es [the Adams and Eves] were ill equipped to understand human decision-making, the way our principles are warped in the force field of our emotions, our peculiar biases, our self-delusion and all the other well-charted defects of our cognition. Soon these Adams and Eves were in despair. They couldn’t understand us because we couldn’t understand ourselves. Their learning programs couldn’t accommodate us. If we didn’t know our own minds, how could we design theirs and expect them to be happy alongside us?’
- The problem of lying: as Turing explains to Charles, ‘Machine learning can only take you so far. You’ll need to give this mind some rules to live by. How about a prohibition against lying? [...] But social life teems with harmless or even helpful untruths. How do we separate them out? Who’s going to write the algorithm for the little white lies that spare the blushes of a friend? [...] We don’t yet know how to teach machines to lie.’
- The robotic corporeality of Adam and the relationship between Adam and Miranda. After having slept with Adam, Miranda insists that he is not more than a vibrator in human-like form, that he is ‘a fucking machine’ Whereas Charles’ take on the situation is rather different: ‘Listen’, I said, ‘if he looks and sounds and behaves like a person, then as far as I’m concerned, that’s what he is.’

All of these themes seem to be vital for understanding the characteristic differences between Charles and Adam, between human beings and robots, and therefore for understanding the essential characteristics of human beings. Corporeality, finiteness, vulnerability and the (subtle, competent, possibly deviant) use of symbols are among the classic characteristics of the human being. Very clearly, what is at issue in the novel is the disorder, the messiness of being human, being thrown into the world without a ‘program’, and the inability to cope with that world. The ability to distinguish between ‘humans’ and ‘machines’ has been approached and problematized primarily in one direction, such as under the Turing test: can machines become human like? But the other direction, the question of whether humans can become human like is increasingly brought to the fore by academics and experts worldwide. For example, in a recent article, Acquisti and others write, summarizing their thesis: ‘Technologies, interfaces, and market forces can all influence human behavior. But probably, and hopefully, they *cannot alter human nature*.’<sup>2</sup>

What I’m interested in is the following: what does it mean that we hope technologies don’t change our human nature and what would this human nature be? Or put differently, what would it mean to change human nature through technologies and why would it be bad to do so? There has been quite some discussion of this or similar problems in the literature and the most helpful and intriguing is to my mind Frischmann

<sup>2</sup> Alessandro Acquisti et al, ‘Secrets and Likes: The Drive for Privacy and the Difficulty of Achieving It in the Digital Age’ (2020) 30 The Journal of Consumer Psychology 4.

and Selinger's *Re-Inventing Humanity*. Frischmann writes in an article in *The Guardian* (2015): 'Alan Turing wondered if machines could be human-like, and recently that topic's been getting a lot of attention. But perhaps a more important question is a reverse Turing test: can *humans become machine-like* and pervasively programmable?' This latter question is the topic of their book. And in the Introduction to their book, they write: "As we collectively race down the path toward smart techno-social systems that efficiently govern more and more of our lives, we run the risk of *losing ourselves* along the way. We risk becoming increasingly predictable, and, worse, programmable, like mere cogs in a machine.'<sup>3</sup> To quote one last passage, this time by Pasquale: 'The future that [the robot] Adam imagines ... reduces the question of human perfectibility to one of transparency and predictability. But disputes and reflection on how to lead life well are part of *the essence of being human*.'<sup>4</sup>

In this picture, we have Turing on the one hand, trying to build a computer which could be mistaken to be human. And on the other hand, we have Frischmann, Selinger and Pasquale, who show us that people – humans – are becoming more and more similar to machines. Acquisti e.a. who I quoted earlier, argue in the same direction: we should be critical of developments which endanger the humanity of human beings. Both sides assume – intuitively quite plausibly - that we know what a 'human being' is and where the limits lie that we should not cross: technologies shouldn't change our human nature. In most of the literature of the last 20 or so years, the criticism of technological developments that come dangerously close to a 'change in nature' is being spelled out in terms of (normative) practical concepts, like freedom, (relational) autonomy, agency, privacy, democracy; also in terms of 'capabilities', the 'freedom to play' etc.

My question is 'what is this human nature that we do not want to change'? And how can we answer this question in such a way that the answer gives us even better ways of criticizing technologies that we believe cross that line? I think there is nothing like *the human nature*, but that human nature is not purely contingent either. The concept or idea should be able to give us critical guidance in talking about the digital society without running the risk of calling 'human' whatever humans learn to do under digitally changing conditions. Also, we shouldn't commit the naturalistic fallacy: the fallacy which is about illegitimately deriving normative statements from descriptive findings. In a first step, 'human nature' could be seen as one of the 'thick concepts' which Clifford Geertz, Bernard Williams or Martha Nussbaum describe, concepts which are not purely normative or purely descriptive, but express elements of both dimensions. Thick concepts are both action-guiding and world-guided and avoid the naturalistic fallacy, or at least that is the idea. 'If a concept of this kind applies,' Williams writes, 'this often provides someone with a *reason for action* ... At the same time, their application is *guided by the world*.' Therefore, it is more complicated than simply claiming that, say, anthropology is a descriptive science, while ethics is supposed to give us an-

---

3 Brett Frischmann and Evan Selinger, *Re-Engineering Humanity* (Cambridge University Press, 2018).

4 Frank Pasquale, *New Laws of Robotics: Defending Human Expertise in the Age of AI* (Belknap Press, 2020) 209.

swers to questions about how we should act. Human beings as human beings always already live in their biological nature, but at the same time in a texture, fabric of norms and concepts that determine, or govern, or shape, the relationship to the person herself, to others, and to the world.

The question of what human or humanity means and how one can develop a critical understanding of human nature can - as an anthropological approach - complement the practical, ethical-political criticism. The anthropological approach adds a historical dimension - from the past and into the future - to the practically normative discourse: this is how we were, how we want to be or how we do not want to be.

Then how should we develop a concept of human nature that is neither too abstract, nor biologicistic, nor (completely) culturally relative? 'Human nature' should not be seen as justifying (which would again be threatened by the naturalistic fallacy) but as reinforcing a network of normative and 'thick' concepts that can give us criteria or standards for the good, the right human digital life. In trying to spell out the irreducible characteristics of human life, I'm taking my lead from novels (and some films). They help to describe situations and phenomena that we would call human, and they help to indicate where we should set the limits of technology.

What should stand at the end of such a research program: a critical anthropological concept linked with normative basic concepts that could be used as criteria in the evaluation of digital developments. So far, this is not much more than a research program. But the idea is that this approach can contribute to clarify the question of the individual and collective good human life and the limits of technological interventions in this life.