

Machine anthropology or will robots talk about us behind our back?

Alec Balasescu *

Royal Roads University Victoria, Canada

Abstract. This article proposes an anthropological approach on technology's own discourse on humanity – or whatever humanity will become – and proposes the term “Machine Anthropology” to designate this discourse. While it may resonate with Agamben's “anthropological machine” – through which humans constantly recreate themselves in reference to, and against, the animal – Machine Anthropology refers to something slightly different: it is pointing to a potentiality of the future, when humanity itself will be the object of analysis for the informed gaze of Artificial Intelligence. Thus, Machine Anthropology designates possible images of humanity through the sensorial and analytic apparatus of the future global technological network. This article in the inaugural issue of the Journal of Future Robot Life will be twofold. It will explore the possible implications of the emergence of Machine Anthropology, while also speculating on the triad AI, humanity, and nature. While Agamben's anthropological machine constantly re-creates nature as opposing referent for culture/humanity, how will the introduction of a third dimension impact this binome? What will be the shape/understanding of nature, if it would exist at all, in the AI algorithms, and further more in the future Machine Anthropology? These type of questions are important today when the rapid advancement of technology is concomitant with the setting of rapid climate change disruptions. If we are approaching an AI future, will climate – and ultimately nature – be relevant for AI?

1. MACHINE ANTHROPOLOGY OR WILL ROBOTS TALK ABOUT US BEHIND OUR BACK?

Storytelling is the fundamental technology of culture building and transmission. We usually think of technology as something material, but the materiality of technology itself is always accompanied by the stories we are telling about it, about what it does, how to use it, and for what purposes. Together, storytelling and material creation – from needle to space rocket, passing through the plough – constitute the software and the hardware of culture in anthropological sense. The stories we created, and are part of, narrate us as much as we create them. The cultural narratives are re-telling the tale of humanity, in a continuous process of co-creation. Thus, if storytelling is a technology of culture, “the story” can be considered the first form of artificial intelligence, in the sense of an extraneous form of human-like intelligence incorporated in something other than the human body. Up to now the source of, and the public for, the stories were humans, both individually and collectively. Things may change.

At some point in mid-2017 few tech publications headlines signaled a story of two chatbots shut down by Facebook apparently because they started to talk to each other in a previously non-existent language. The news spread in the larger media and, depending on what one reads, it may have sounded like the new era of machines taking over in a Terminator-like scenario was at its dawn. What really happened was slightly different: the two bots were a Deep Learning type of software training themselves to negotiate. The programs performed – and even outperformed – human negotiators, and they were interacting with each other in conformity with the algorithm. However, at some point they started to take shortcuts from human language, continuing to perform as programmed – that is negotiating with each other – but using the abbreviated, more direct language. They did not do anything

*Royal Roads University Victoria, British Columbia, Canada. E-mail: alec.balasescu@royalroads.ca.

outside the script, however they innovated the manner in which they were doing it, since the programmers omitted to incentivise the bots to use the human language in the process. (Gizmodo). The programmers shut the bots down, and corrected the error.

The general public reaction and understanding of the event varied, and in itself expresses the anxiety around what AI is, or more accurate, what AI might become. The distinction between AI – specialized Artificial Intelligence as we have and use today – and AGI – Artificial General Intelligence, or a human like type of intelligence, that may or may not be attainable in a foreseeable future, depending on whom you ask – is yet underused in general public discourses. Many times AI and AGI are conflated when artificial intelligence is discussed. The anxiety around the event of the two bots chatting to each other had two sources:

1. the autonomy they displayed in their actions
2. the apparent act of creation of a new language, inaccessible to humans.

For the sake of argument, let's assume that in an indefinite future AI will be able to autonomously communicate, create meaning outside of the task they were designed for, and generate a language or languages of their own. Let's call this "AGI". While this is not enough to qualify the bots as "alive" in biological terms, it is enough to generate "culture", as a set of artefacts created by the robots, meanings attached to them and transmissible through a specific language that was co-created in this process. If the robots are able to generate a culture of their own, the question "are these bots alive?" returns in another guise: "in what way are these bots alive?" pushing us towards a redefinition of life itself. Revisiting the "*zoe vs bios*" distinction, as used by Rabinow (1999) or Agamben (1998) in their writings, where *zoe* is life itself stripped of its particularities, while *bios* is life plus its details (political in Agamben's term, or cultural in Rabinow's), the culture-creating robots have a future life that may be defined as the subtraction of *zoe* from *bios*. The simple mathematical representation of this would be:

1. Bios = Zoe <plus> Cultural/Political
2. FRL (Future Robot Life) = Bios <minus> Zoe
3. FRL = Cultural/Political

Within this self-referential logic, the condition of life for robots is not the presence of *zoe* – of life as we understand it – but the presence of culture building capacities, which are, at least speculatively, foreseeable. It is important to underline from the beginning that the materiality of AGI is, and will continue to be, a disputed territory. In the above speculative scheme, AGI's materiality is part of the shared cultural realm between humans and robots, as it is a result of a creative act, at the moment performed by humans in collaboration with machines (technology). Will that materiality move towards a realm comparable with *zoe*, and to what effects? A new type of self generating materiality, a *robo-zoe*?

Let's keep these questions in mind while moving forward into exploring a part of the possible cultural production of FRL: its anthropology.

What could Machine Anthropology be? The short answer: the aggregate of AI conversations about us, humans. One of the possible (and necessary) creations of FRL will be its own discourse on humanity – or whatever humanity will become. I call this "Machine Anthropology", using a term that emerged during the collaboration between Roy Bendor, currently at Technical University Delft, and myself, on a fictional research abstract as a response to a call for papers for Critical Alternatives 2015.

While the term is chosen to resonate with Agamben's (2002) "anthropological machine" – through which humans constantly recreate humanity as a category in reference to, and against, the animal realm – Machine Anthropology refers to something slightly different: it is pointing to a potentiality of the future, when humanity itself will be the object of analysis, and perhaps polar opposition for the informed gaze of Artificial Intelligence. Thus, an exercise in Machine Anthropology can only be speculative, pondering on how we may look like through the sensorial and analytic apparatus of a global technological network endowed with AGI.

The questions arising from this possible scenario are endless. The central idea here is that the stories about ourselves will no longer originate with(in) us, we will not be narrated only by our own stories, but by the stories told by something else, within a different cultural space, and possibly in a non-human language. Assuming that we will continue to collaborate and co-exist with AGI, that very knowledge will have a direct impact on humans' daily practices. As our own stories reshape ourselves, AGI stories about us will also reshape us – and them in their turn.

In his essay "The Storyteller" Walter Benjamin (1973) observed, at the time of writing in the early to mid 20th century, a decline of storytelling caused by the transformation of techniques of sharing information – from oral to written (the novel) and from theatre to cinematography. Benjamin has a particular view on cinematography, observing its non-engaging character and relating it with the rise of totalitarianism. However, our storytelling habit did not die away, despite the advances in technology and the ever shifting changes in the form of storytelling.

In our current, let's call it pre-FRL time, has the storytelling already been replaced by dry information sharing? Hard to say, the arguments can go both ways. Probably most important for story telling is the presence of the public and the direct interaction with it, regardless of the scale. This generates and/or activates empathy based responses, both from the public and from the narrator. Mediated contact – such the one facilitated by communication devices – significantly reduces the possibility of expressing reciprocity and the generation of empathy.

In this light, we can wonder about the way in which AGI will tell stories about humanity. How will Machine Anthropologists navigate the data aggregates in order to tell stories about us? Or, will it do it at all? Already one of the fears expressed in many critical approaches on AI is the lack of empathy in algorithms – and even lack of objectivity for that matter (see Galison, 2019). Is it possible for the future AGI to replicate the human storyteller, and to what extent? I imagine – and I may be absolutely wrong – that our own stories will not be told anymore, but that specialized narrative machines will generate them, along with their own stories. It somehow already happens, the film industry working more often than not with screenplay generators based on narrative algorithms, while some publications are experimenting with AI written articles. AI generated art proves to be more than a fad and sells for thousands of dollars.

Now, in order to stay true to my theoretical stance, I do not think that there is a separation between humans and machines, yet. We are, as we have always been, in full process of co-creation; we have always been cyborgs (Harraway 2016 [1984]; Warnier 1999; Weber 1984 [1909]) but, at the same time we, humans, need to affirm that separation, in the same way we need to affirm the separation from the "animal" or "nature" while constantly producing through discourse that very "nature". The question is: will the machines ever feel and respond to that same need of separation from humans? And what happens when they start to do so?

It may be argued that the process of creation of Machine Anthropology already started: the easiest example is the digital world, in which personal digital prints create profiles as algorithmic interpretations of an individual's behaviour. A future leap towards a fully fledged Machine Anthropology

happens when aggregated data will be subject of AI analysis in order to translate and interpret human collective behaviour, producing “cultural readings” of that behaviour. Based on that, decisions may be, and have already been taken, that influence our daily lives. Targeted marketing, including political marketing, and all the phenomena in their trail from fake news to skewed elections have continued to make the headlines in the past 2-3 years. We can thus already see the socio-political effects of a series of co-created stories about ourselves.

The question is what will happen when we will face a fully fledged Future Robot Life. The “Machinity” may start to produce itself an autonomous discourse about humanity, while re-producing itself in constant oppositional reference to Humanity. The Machine Anthropology can thus be considered the necessary ideological prosthetic of the creation of a self-conscious, self-reliant AGI. The networked technological being (let us call it this for lack of a better descriptive) will create its own interpretations of what humanity is, in its various manifestations, thus creating a more or less unifying identity for itself, and entering into the competition of “owning” humanity through digitally created discourse about humanity.

Nature itself as a category will be further removed from Future Robot Life. Let us revisit the equation $FRL = Bios <minus> Zoe$. Whatever remains in between, the machine culture, is doubly removed from nature if there is any perception of nature at all. The genealogy of AI creation and development shows how pre-existent biases get to be built into algorithms. Up to now, AI seems to be vastly influenced by a projection of what intelligence means for a narrow population segment, with narrowly defined characteristics and cultural biases that include sexism, and racial and political biases (Broussard, 2018). This reflects the way in which algorithms tend to reproduce and augment biases. The social effects are truly unforeseeable. It is remarkable how artists offer critiques to this machine behaviour, while sometimes needing to work around the very biases that their art critiques. (Amaro, 2018).

Even if we make abstraction of these biases, we seem to be far from agreeing on what intelligence means, but we are projecting mostly a disembodied definition of it on the artificial one (Richardson, 2017). Robotics, or embodied AI, despite its existence as an independent field, captures much more Hollywood’s and military imagination, and the attention of manufacturers, but much less that of corporations that look for fast applicability. Robots appeared as, and continue to be, the working class of AI, despite the fact there are many arguments for embodiment as the *sine qua non* condition for approaching Artificial General Intelligence.

The equation “Bios <minus> Zoe” thus results in a narrowly defined space, an idealized intelligence that would or does operate within the confines of that definition, explicitly containing an idealized and culturally marked measure of humanity and of what human intelligence is. The body seems not to be part of it, but only a necessary tool that may or may not have an input in the process of “acquiring” intelligence for a machine. However, the body is humanity’s unmediated link to nature, hence the efforts in domesticating and/or negating it in the narrow definitions of intelligence that for the moment seem to be the blueprint for AI.

Furthermore, if we think more generally about the category “nature”, we may find ourselves facing one of the most fundamental changes that AGI – and even the current use of AI – may bring about. If “nature” is an idealised category that opposes, and simultaneously creates humanity through that opposition. “Idealized humanity” (void of nature?) will be what opposes and (literally) creates Future Robot Life. This is precisely where the materiality of Future Robot Life comes into question: will the *robo-zoe* allow for a conceptualization of nature? And if yes, in what terms? I think this question is fundamental for the way in which future life in general, robotic or not, will continue on our planet, and in the outer reachable realms.

To conclude, the category “artificial” is highly disputed, and anthropologically explainable as a by-product of the process of culture creation: in order to define ourselves, we humans invented “nature” and we’ve relegated it to the outside of humanity in order to create humanity. We behaved up until recently (and ruthlessly so) as if we are outside nature. At the same time, and even more now, we use the category artificial to name everything that is created by us. Since we decided that we are not part of nature, this comes as a logical (should I say “natural”?) consequence. So we placed ourselves between the “natural” rock and the hard place of “artificiality”. They both exist in our minds, and their existence as cultural categories have immense impact on how we live our lives today.

Within the context of Climate Change as hardstop for human activity, the question of the place of Nature within the AI realm needs to be addressed from an extra-algorithmical, non-optimization-oriented perspective.

REFERENCES

- Agamben, G. (1998). *Homo Sacer. Sovereign Power and Bare Life*. Stanford: Stanford University Press.
- Agamben, G. (2002). *The Open. Man and Animal*. Stanford, California: University of Stanford Press.
- Amaro, R. (2018). “As If” in *e-flux* <https://www.e-flux.com/architecture/becoming-digital/248073/as-if/> accessed August 2019.
- Benjamin, W. (1973). The storyteller. In W. Benjamin (Eds.), *Illuminations* (pp. 83–109). London: Fontana. Edited with an Introduction by H. Arendt.
- Broussard, M. (2018). *Artificial Unintelligence. How Computers Misunderstand the World*. Cambridge, Massachusetts and London, England: The MIT Press.
- Galison, P. (2019). Algorithmic dream of objectivity. In J. Brockman (Ed.), *Possible Minds. 25 Ways of Looking at AI* (pp. 231–240). New York: Penguin Press.
- Haraway, D.J. (2016). *A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century*. University of Minnesota Press [1984].
- McKay, T. (2017). *No, Facebook Did Not Panic and Shut Down an AI Program That Was Getting Dangerously Smart* in *Gizmodo*, 07/31/2017. Available at: <https://gizmodo.com/no-facebook-did-not-panic-and-shut-down-an-ai-program-1797414922>.
- Rabinow, P. (1999). *French DNA. Trouble in Purgatory*. Chicago and London: University of Chicago Press.
- Richardson, K. (2017). *An Anthropology of Robots and AI. Annihilation Anxiety and Machines*. New York and London: Routledge.
- Warnier, J.-P. (1999). *Construire la Culture Matérielle: L’homme Qui Pensait Avec Ses Doigts*. Paris: Presses Universitaires de France.
- Weber, M. (1984). Energetic’ theories of culture. *Mid-American Review of Sociology*, 9(2), 33–58. [1909]. Translated by Jon Mark Mikkelsen and Charles Schwartz.