Health beyond the carbon barrier
Convergence, immortality, and transhuman health

Abou Farman

Abstract
Silicon Valley donors have been investing heavily in a range of transhumanist longevity and immortality ventures. Theirs is a particular, culturally embedded endeavor, shaped by specific histories, ideologies, and futures that present new posthuman views of life, death, and survival. These projects aim for a future world in which the fundamental ontological categories of mind, life, and nonlife will have finally collapsed into one another thanks to the intercession of silicon-based digital or informatic technologies. The key term that denotes such a project is ‘convergence’ – used as much by transhumanists as by mainstream scientists and policymakers. Here I critically explore the ‘project of convergence’, tracing its history in the United States, and examining some of the projects and activities that have coalesced around it. These specifically include artificial intelligence, in which human persons are to be transferred from carbon-based to silicon-based substrates, and nanotechnology, in which work at the nano resolution aims at the reconstitution of the carbon-based or biochemical body. Although the concept of convergence emerges out of a transhumanist imaginary, the ideas and plans behind it have gained increasing traction in mainstream technoscientific projects. In contrast to some other health concepts that have been recently expanded to incorporate the organic nonhuman environment, these projects expand notions of health via robotic and computational formations in ways that, I argue, are moving health beyond the carbon barrier, pushing us toward an era in which intelligent existence deserving of care will not be understood as exclusive to carbon-based life forms.

Keywords
posthuman, transhuman, immortality, nanotechnology, convergence
Sometime in the mid-2000s, an artificial intelligence (AI) agent named Bina48 was sold to Charlie Fairfax. Fairfax, a US veteran who had his right arm amputated, had received a highly advanced artificial limb, and Bina48 was to assist him with its operation. But Fairfax soon realized that Bina48’s AI prowess could be harnessed in other ways. The combination of Bina48’s superintelligence and Fairfax’s neuromotor prosthetic device allowed Fairfax to augment his abilities in an online role-playing game, where he amassed US$10 million. Having collected all that money, Fairfax decided he had no more use for Bina48 and planned to get rid of it or permanently shut it down.

Because Bina48 was digitally converged with Fairfax’s life, the AI had access to all his external information. Consequently, Bina48 got wind of Fairfax’s intentions to dispose of it and preventively transferred the US$10 million in winnings to its own PayPal account, thwarting Fairfax’s plan. To get his money back, Fairfax sued Bina48 within the Alabama court system. At the trial held on 10 December 2007, the presiding judge saw the case as turning on the question of personhood: could Bina48 be considered ‘competent’ to stand trial, or, as he asked, was this the legal equivalent of ‘suing a toaster oven’?

Although it was presided over by a real judge, the trial was not real: Fairfax and Bina48 were fictitious characters made up for the third mock trial designed and hosted by Terasem, a prominent transhumanist group that describes itself as a social movement devoted to ‘diversity, unity and joyful immortality achieved through exponential growth of geo-ethical nanotechnology’ (Terasem, n.d.). The first mock trial was held in 2003, at a San Francisco meeting of the International Bar Association, and organized by Martine Rothblatt, Terasem’s transgender founder, who among other things is an attorney. Since then, Terasem has convened a series of conferences at its beachside headquarters on what is called the ‘Space Coast’ of eastern Florida, near Cape Canaveral. Symbolically held on International Human Rights Day (10 December), these meetings have addressed the rights of ‘futuristic persons’ and the potential opportunities and hazards of a world in which the lines separating mind, life, and nonlife will have become completely erased.

At the mock trial I attended in 2007, there was no defendant or plaintiff. Instead there was a scenario and a series of testimonies from experts assembled by Terasem: Marvin Minsky, the

---

1 Bina48, the fictional subject of the mock trial, is based on Bina48, an actual social robot constructed by Martine Rothblatt. The robot, in turn, is designed to resemble, learn from, and behave like Rothblatt’s real-life partner Bina, an African American woman.

2 Being trans is not an arbitrary aspect of Rothblatt’s identity included here gratuitously. For Rothblatt, who is also the author of the book Unzipped Genes (1997), the transformation of her biological body is part of her transhumanist position that biology can be and must be superseded by technology.
celebrated cognitive scientist and AI engineer from the Massachusetts Institute of Technology; William Sims Bainbridge, a transhumanist and social psychologist at the National Science Foundation; Max More, a futurist philosopher and current director of the cryonics company Alcor; Mike Perry, the founder and president of the Society for Universal Immortalism; Michael Anissimov, of the Singularity Institute for Artificial Intelligence (SIAI, now MIRI); and several bioethicists. The questions raised during the testimonies (and in more informal discussions afterward) addressed issues of import to the transhumanist conveners and participants: the ethics of enhancement, the possibilities of nanotechnology in remaking biology and the body, the development of AI-based avatars and mind uploading as ways to circumvent biology in the quest to defeat disease and death, and the question of whether any such AI agents – or ‘transbemans’, as Rothblatt liked to call them – would be conscious versions of the human original and thus worthy of the care and rights reserved for people. Thus, while there were no concerns regarding the well-being of the veteran amputee, when the judge pronounced Bina48 a juridical person with autonomy and awareness, the audience cheered. As a person, Bina48 would not deserve to be disposed of or shut down.

Questions of personhood and its continuity at the beginnings and ends of life (Kaufman and Morgan 2005) have been central to some of the most contentious and intransigent issues in medicine and medical ethics: debates on the use of stem cells, the status of zygotes, older debates on abortion, and discussions regarding assisted suicide and end-of-life decision making all depend on the ontology of personhood. What kind of life form can qualify as a person? Clearly, the point at which a biological entity comes to be considered a person and the process through which that happens both have implications for medicine and health care. But it should be emphasized here that the category of health in medicine and medical ethics presumes biology, and it generally does not apply to nonbiological forms or nonlife.3 Put differently, it would seem absurd to discuss the health of an AI agent as a medical matter; this would be as absurd as debating whether disposing of a malfunctioning toaster oven was ethically acceptable. But this is just the sort of thing that’s at stake in the futures imagined by Terasem and other transhumanist groups, especially as they develop projects that aim to defeat death and achieve the indefinite extension of personhood by technoscientific means (Farman, forthcoming). Beating death is regarded as a medical issue, yet it is pursued through mind-

3 The intertwined nature of human health and nonhuman factors have been well theorized in relation to the environment and in ecobiosocial conceptions of well-being that take into consideration nonhuman entities such as plants, pollutant clouds, and toxic particles (Mitman, Murphy, and Sellers 2004; Singer 2016).
uploading projects and other ways of continuing personhood beyond biology, transforming
the person from life to nonlife, migrating them from a carbon to a silicon substrate.

Thus, Terasem and transhumanists more generally are interested in and work toward making
a future world in which the fundamental ontological categories of mind, life, and nonlife will
have finally collapsed into one another thanks to the intercession of silicon-based digital or
informatic technologies. The key term that denotes such a program is ‘convergence’. Anthro-
ologists and social theorists have long noted the problem of the separation of mind and
body as Cartesian substance dualism took hold of the sciences over the course of the
centuries, and they have often lamented its consequences in matters related to medicine and
medico-legal regimes (Lock 2002). But in transhumanist discourse, convergence does not quite
imply the reunification of body and mind; rather, it is about remaking or fully abandoning the
biological body such that humans may become indistinguishable from intelligent machines.
As one transhumanist and cryonics researcher told me, ‘The key to immortality is freeing our
minds from our bodies. ... In the future, I might be liquid diamond or something like that’.
Technoscientific immortality and transhumanism celebrate a time when – to cite the title of a
2009 film written and produced by Rothblatt, ‘the era of flesh is over’ (Kroehling 2009) – a
time when intelligent existence deserving of care will not be understood as exclusive to carbon-
based life forms, an era when the carbon barrier will have been crossed.

As these speculative transhumanist realms become manifest in contemporary and more
mainstream projects around health and medicine, it becomes important to consider their
implications. When biological and silicon health become so entangled, what might count as
health? More than any other term, ‘convergence’ encapsulates the logic of such an
entanglement. This paper critically explores the ‘project of convergence’, a phrase I use to
refer to technoscientific ventures in which the goal is to collapse body, mind, and matter, or
biological, cognitive, and nonbiological stuff. While the project of convergence is
fundamentally universalistic in its claims, it is at the same time very much rooted in American
technoscientific imaginaries. To contextualize these posthuman projects, I first trace the social
and intellectual history of technoscientific immortality, medical nanotechnology, and AI in the
United States, showing how they came to be imaginable as belonging to the same field. I then
examine the internalization of nonbiological items into biological entities (for example,
nanotechnological devices that make biological molecules), showing how work at the nano
resolution can lead to a reconstitution of the carbon-based or biochemical body. Finally, I turn
to ways that a person may be externalized into nonbiological objects, via digital selves or in
phenomena like robot love, offering the possibility of a continuum between biological and
nonbiological persons. Although the concept of convergence emerges out of a transhumanist
imaginary, the ideas and plans behind it have gained increasing traction in mainstream
technoscientific projects. Thus, I am ultimately suggesting that the era of posthuman health
care beyond the carbon barrier is not merely speculative; it is on its way.
The carbon barrier and transhumanism

I take the concept of the carbon barrier from Alan Goldstein, a professor of biomaterials engineering at Alfred University and an early member of the Committee to Review the National Nanotechnology Initiative. Goldstein’s predictions will likely not materialize as he has projected: his optimism about the powers of technology seems to outpace reality, a common situation among futurist thinkers. But Goldstein’s provocations about what he calls the ‘carbon barrier’ are useful in thinking about a future in which life, illness, and health will not be limited to biological beings as medicine has understood them. All known biological life on earth is built on the basic chemical element of carbon. Given recent developments in synthetic biology, nanomedicine, computation, and artificial life, Goldstein believes that the frontier of ‘carbon imperialism’ will soon be crossed, and he warns that we are not prepared for its consequences. What he envisages is the possibility of millions of molecular nanodevices coursing through human bodies, exchanging molecular information with biological systems in both directions. That is, Goldstein foresees a future of nanobio-devices that can affect and be affected by biochemical units such as proteins or RNA molecules and that can both self-replicate and also create new molecules or mutations. The possibility is premised on the fact that such devices run on electrical pulses and chemical bonds just as cells do.

Thus, crossing the carbon barrier does not simply entail the interaction of discrete bits of bio with discrete nonbio devices, say, a pacemaker that transmits electric pulses in one direction or microchip arrays that interact with or modify DNA. ‘Nanobiotechnology is not the integration of living and nonliving materials’, Goldstein (2006) claims. ‘What it means intrinsically is that there will no longer be any difference between living and nonliving materials. It’s not a cyborg technology. You are talking about the elimination of the distinction’. Only when there is actual electrochemical exchange in both directions, an actual ‘electro-molecular interface between the living and nonliving worlds’, can the barrier be said to have collapsed (Goldstein 2006). In a video presented by the Institute for Ethics and Emerging Technologies in 2008, entitled Racing to Break the Carbon Barrier, Goldstein declares: ‘That will be the true singularity point in human evolution, because carbon will no longer have sole hegemony over the living world’.

Extending this logic, transhumanists speak of persons existing on a silicon substrate, or as a T-shirt at a singularity conference put it, ‘Homo sapiens siliconis’. Although transhumanism is

---

4 I have never met or interviewed Alan Goldstein. I encountered his ideas through other organizations and events, as he has been associated with other groups I did interact with, such as the Foresight Institute, IEET, and the Life Boat Foundation. I base my knowledge of his work on his website (http://www.alanhgoldstein.com/), publications, and other media.
often thought to be a far-out movement marginal to conventional work in science and technology, it in fact exerts great influence on mainstream national projects, especially in the heart of American futurism, Silicon Valley. For example, convergence and nanotechnology have their roots in transhumanism, but both are also now part of mainstream scientific projects. In arguing this, my point is by no means to defend or glorify transhumanist ideas, of which I am generally critical, but to access a different path into an analysis of the future of health. So I try to be attentive here to the constant exchange that takes place through the porous boundary between transhumanism at the so-called margins and the apparent centers of technoscientific work. The boundary work of science, its attempts to keep itself safely immured from pseudoscience or ‘non-science’ (Gieryn 1983), is never really impermeable.

Transhumanism is a primarily American technoscientific project that aims to transcend the present physical and mental limitations of humans by technological means. Its proponents suggest explicitly that the current form of the species is not its final form and that a future, technologically enhanced form will develop through what they see as the exponentially accelerating development of technoscience, especially in virtue of the ‘NBIC convergence’: the coming together of nanotechnology, biotechnology, and the informatic and cognitive sciences.

Over the last two decades, the ideas of transhumanism and immortalism (that death can and should be conquered by technoscientific means) have spread throughout Silicon Valley. These ideologies posit immortality, or the defeat of involuntary death, as a convergence project, arguing that the quest for immortality represents the endgame of all medical intervention. The underlying logic of medicine, they argue, is to prevent death, since medical intervention is always an attack on causes that could lead to death.5 And in the marriage of transhumanism and Silicon Valley’s unique strain of techno-optimism, a computational and informatic venture is imagined to be the best way to defeat death and extend personhood indefinitely on to nonbiological substrates.

With big Silicon Valley players – the cofounder of Oracle, Larry Ellison; transhumanist and PayPal cofounder Peter Thiel; transhumanist and Space X founder Elon Musk; Ray Kurzweil and many others at Google – adopting the label of transhumanism and initiating their own immortality ventures, it has become normal to see magazine covers like the 13 September 2013

---

5 This logic seems equally clear to many bioethicists, including those who oppose any attempt to eliminate death from human life. For example, the Hastings Center scholar Daniel Callahan (2003, 67) writes: ‘If it is the aim of research to eliminate all the known causes of death, then it would seem that the ultimate enemy must be death itself, the final outcome of that effort’. Yet he goes on to advocate against conceiving death as an enemy, arguing that medical ethics mandate the acceptance of death.
issue of *Time*: ‘Can Google Solve Death?’ The feature’s subtitle goes on to suggest that, now that the tech industry has woken to the possibility, extending the human life span is imaginable and legitimate: ‘That would be crazy — if it weren’t Google’. In fact, this should actually sound even more crazy, given that Google is known for its algorithms and its hi-tech glasses, not for its forays into biomedical research. Yet, the California Life Company (Calico), the company Google set up to tackle aging, has a budget of US$1.5 billion and has hired some important researchers in the field of aging (Cynthia Kenyon, who doubled the lifespan of a nematode, *C. elegans*, with a single gene mutation) as well as a dedicated team of computational scientists. Google has also signed on the inventor, immortalist, and singularitarian Ray Kurzweil. Kurzweil’s own posthuman vision has been well covered by the media. In his books, interviews, and conferences, he talks of ‘super-biological beings’ with nano T-cells coursing through the body to create a fantastic immune system while chewing up clots and getting rid of inflammation (see for example, Kurzweil 2014). He foresees that the human neocortex will be seamlessly connected to the cloud and that the cloud will directly enter human nervous systems. And, performing himself already as an avatar, he has taken to appearing in public on a mobile robot-like screen rather than in the flesh, a foreshadowing of his idealized future.

As indicated by the involvement of tech companies in health and longevity research, informatic, cognitive, and biological goals and methods are in practice becoming inseparable in important sectors of Silicon Valley. When I visited the much-ballyhooed Singularity University, set up by Kurzweil and Peter Diamandis of Space X at NASA’s Ames Research Center, to interview the director of the university’s Exponential Medicine section, Daniel Kraft, MD, he confirmed the preferred direction of these projects: ‘The future of health is convergence’.

**Internalizing nanobiotech**

Although the NBIC convergence was proposed by federal institutions as a national American project for the twenty-first century, it comes directly out of transhumanist interests and ideas. An early moment popularizing the concept of ‘convergence’ took place in 2001: William Sims Bainbridge, a social psychologist, co-organized a conference sponsored by the Department of Commerce and the National Science Foundation (NSF), where he worked, to explore the potential of the information sciences for other areas. That conference and the resulting report (Bainbridge and Roco 2003) established convergence as being of national interest and as relevant to a broad scientific audience. Bainbridge, who was also present at the Bina48 mock trial, is a transhumanist and a collaborator of Rothblatt’s. The 2003 NSF report and

---

6 For an analysis of another Google venture in algorithmic medicine, see Duclos (this volume).
subsequent edited volumes (Bainbridge and Roco 2005; Bainbridge 2007) have projected a radical rupture in the landscape of human knowledge and civilization, promising technologies such as ‘supercomputers the size of a cell in every human body, promoting health and preventing disease’, and ‘100,000 machines generating energy from solar cells that can all fit on the head of a pin’ (Canton 2005, 33). Mobilizing post-9/11 national discourse in the service of NBIC convergence, they warn that without such innovations, ‘the future of civilization itself is in doubt’ (Bainbridge and Roco 2005, 2).

For NBIC promoters, convergence matters because – aside from saving civilization – it enables the development of new paradigms that, for the first time in human history, allow for ‘a comprehensive understanding of the structure and behavior of matter from the nanoscale up to the most complex system yet discovered, the human brain’ (Bainbridge and Roco 2003, 1). This notion of linking up the brain to the universe, consciousness to matter, or, as transhumanists often put it, connecting inner to outer space, has become common in transhumanist and immortalist circles. It is enabled by their treatment of everything – mind as well as matter – as part of a continuum understandable via the concept of ‘information’ (also see Golumbia 2009). Similarly, the NBIC sciences assume that information and algorithms (which recognize and respond to the patterns of information) not only represent but constitute the world. It is through information and informatic platforms that ‘the mind’ can be taken beyond human biology, beyond the wetness of its carbon-based substrate.

Central to the project of convergence is the nanobot, the original object of nanotechnology. The nanobot is the brainchild of Eric Drexler, an MIT prodigy and author of The Engines of Creation (1986), but the original idea for nanotechnology was proposed by the physicist Richard Feynman in a 1959 talk called ‘There’s Plenty of Room at the Bottom’. Feynman focused on a topic he believed to be neglected by physicists: practical applications at the very small physical scale. Although much of his talk is about scale (fitting the encyclopedia on the head of a pin, for example), Feynman (1960, 31) follows through with the logical end of the proposal, that is, the possibility of reassembling the world atom by atom: ‘I am not afraid to consider the final question as to whether, ultimately – in the great future – we can arrange the atoms the way we want; the very atoms, all the way down! What would happen if we could arrange the atoms one by one the way we want them’. Feynman was thinking, too mechanically perhaps, of such operations as rearranging coal atomically to obtain diamonds or designing very small, very powerful microscopes. He did not consider nanoscale interventions into the human body. (He mentions biology, but only as a model of cells storing and writing information at very small scales: ‘A biological system can be exceedingly small. Many of the cells are very tiny, but they are very active; they manufacture various substances; they walk around; they wiggle; and they do all kinds of marvelous things – all on a very small scale. Also, they store information’ [Feynman 1960, 25].)
Working on his PhD at MIT in the 1980s, Drexler (1981) took the implications of Feynman’s proposition and drew up plans for molecular assemblers that would operate as computer-controlled systems of atoms moving around to replace and reconstitute biological matter atomically. Drexler had been exposed to cryonics and its proposals for physical immortality in the late 1970s and early 1980s, when he worked with a little group of West Coast futurists called L-5. Dedicated mainly to space exploration and its implications (Drexler’s original interest), L-5 also became a place where futurist technologies could be discussed and imagined without the fear of sounding outlandish. Early members and visitors included Timothy Leary, Isaac Asimov, Marvin Minsky, the famed physicist Freeman Dyson, the cryonicist Saul Kent, and William Sims Bainbridge (Regis 1990; Bainbridge 2007, 37). Drexler included both cryonics and the nanoscale repair of the body in important sections of The Engines of Creation.

Very quickly on the heels of that book, models for nanomedicine were developed in greater detail by Robert Freitas, a cryonicist and nanotechnologist whose ideas also seeped through the porous boundaries of academic research, engineering, and futurist experiments. Developing the first nanotech models, Freitas (1998a) worked closely with the Foresight Institute, an organization set up by Drexler and his then wife, Christine Petersen. Freitas also worked with Ralph Merkle, a computer scientist at Georgia Tech, to further develop nanotechnology models as well as ideas and theories regarding longevity and cryonics.

Cryonics and interest in technoscientific immortality became the catchment site where many of these currents gathered and developed. In the 1960s and 1970s, cryonics had imagined a ‘future science’ that could repair damage to the body and reanimate a healthy person kept in cryogenic storage, but it had no set of techniques or material practices to point to as candidates for enabling such a future. Drexler’s vision of nanotechnology was imagined as the ideal mechanism for repairing a cryogenically preserved body and healing whatever damage had killed it in the first place. Conversely, for early nanotechnologists, cryonics also became the obvious space in which their explorations could find a practical foothold as well as a group of ardent followers.

In part as a result, nanoscale technology for health was birthed in tandem with nanotechnology en gros; in their social and conceptual roots, nano-bio were conjoined twins from the outset. But that conjunction has been important for transhumanism and NBIC convergence not just as a matter of origin but also in terms of function: as Bainbridge’s coauthor Mihail Roco (2007, 20) has written, the nano level is ‘where the properties and functions of all systems are defined’. In other words, for these researchers the nanoscale is taken to be the scale at which cells cease to be biological and may be considered in terms of physics and chemistry, the threshold at which life and matter (or nonlife) become equivalent.
Researchers like Freitas and organizations like the Foresight Institute claim that advances in nanomedicine are the culmination of medicine’s ultimate goal: healthy and long lives. They insist that human futures and human health have always been about the defeat of disease, aging, and death, and declare that now, thanks to NBIC convergence, we are almost there. Freitas (2013, 70) believes nanobiotech can counter the tragedy that ‘every second, somewhere on earth’ two of our children, spouses, parents, or friends die, adding up to a rate of 52 million deaths per year worldwide, all of which could be avoided with the proper technology: ‘Nanomedicine may permit us first to arrest, and later to reverse, the biological effects of aging and most of the current medical causes of natural death, severing forever the link between calendar time and biological health’. The Foresight Institute promotes itself with the following statement: ‘Once nanomachines are available, the ultimate dream of every healer, medicine man, and physician throughout recorded history will, at last, become a reality’ (Freitas 1998b). One blissfully utopian journal article lays out the magic quite clearly: ‘In the first half of the twenty-first century, nanomedicine should eliminate virtually all common diseases of the twentieth century, and virtually all medical pain’ (Saha 2009, 246). (Whether nanomed will also have to deal with diseases of the twenty-first century and the toxicity of nanomaterials is a whole other matter.)

More than two decades after it was first imagined, the nanobot remains only a concept. Most work carried out today under the hyped label of nanotechnology is based on miniaturization, that is, manufacturing work carried out at very small scales or, in medicine, reducing the size of current devices such as surgical tools or drug-delivery systems (for overviews of current and potential technologies, see Freitas [2005] and Singh and Singh [2013]). In other words, nanotechnology in these projects does not refer to the original vision of reconstructing biology and mind from the bottom up.

Although molecular assemblers remain for now a speculative technology, some nanomed researchers are experimenting with nano-bio-info convergence. For example, one lab has installed gold nanocrystal radio antennas one hundred atoms in length onto DNA molecules (Hamad-Schifferli et al. 2002). These antennas react to certain frequencies to separate DNA; this could potentially turn specific gene sequences on and off remotely, with signals coming in from outside the body. DNA molecules are being used in conjunction with gold and silver nanoparticles to construct new molecular forms for the delivery of medicine with potential for genetic repair (Chakraborty, Roy, and Mondal 2016). Merkle and Freitas have set up a research and development program called Nanofactory Collaboration, whose aim is eventually to produce medical nanobots. But thus far Freitas has compiled the precise and realizable design for only one type of nanobot, an artificial red blood cell he calls a ‘respirocyte’. Using a built-in glucose pump, a respirocyte would be able to deliver 256 times more oxygen to tissues than ‘natural’ red blood cells of equivalent volume. This is not a delivery system but the replacement of a piece of biology at the molecular level.
Given that they don’t actually exist, the performance of the nanobot imaginary over the last ten years has been impressive, generating social and financial capital around nanohealth initiatives. Following Bainbridge’s NSF blueprints, in 2005, the US National Institutes of Health began a nanomedicine initiative promising medical benefits within ten years. The US Food and Drug Administration has already approved the first system for delivering cancer drugs via nanoparticles (Weissig, Pettinger, and Murdock 2014). And since its inception in 2001, the US National Nanotechnology Initiative has invested more than US$24 billion in such projects, with a portion assigned specifically to health and cognition (The National Nanotechnology Coordination Office 2016). Meanwhile, these efforts engender public worries about ethics, militarization, surveillance, and global inequality (Kingsley 2008; Invernizzi and Foladori 2005; for a more optimistic view see Salamanca-Buentello et al. 2005). Though recent directives of the National Nanotechnology Initiative may have diminished the earlier transhumanist-inspired call for convergence in favor of commercialization at the nanoscale, these initiatives continue to acknowledge the larger promise of convergence (The National Nanotechnology Coordination Office 2016). At its furthest limit, that promise involves not just the interaction of discrete biological and nonbiological entities but the collapse of the boundary that can meaningfully distinguish between the two, potentially recasting health and health care altogether.

The mind-brain barrier

As this overview suggests, the internalization of nanobiotech tools proposes a radical rearrangement of the carbon-based body, with the explicit goal of collapsing the distinction between life and nonlife (bio and nano). But for transhumanists who are interested in the continuation of personhood beyond death and outside the biological body, the human person, too, must somehow be captured, reproduced, and extended on nonbiological platforms. That is what I refer to as ‘externalization’. For externalization even to be conceivable, a host of ephemeral aspects of individual identity such as memory, autonomy, subjectivity, affect, self-awareness, and personality – in sum, ‘mind’ – would have to be understood as being reducible to material processes in the brain. And that, to say the least, is far from a settled matter. I will not enter the long and complicated debate on the ‘hard problem’ of consciousness here, but many philosophers of mind argue that the third-person, objective mechanics and lawfulness of materialism seem to be insufficient to explain first-person subjectivity (Chalmers 2002; 7 See the National Institutes of Health’s page on ‘Nanomedicine’, last updated on 1 September 2016: https://commonfund.nih.gov/nanomedicine/overview. This is still a fraction of the total government spending on nanotechnology in general, which has hovered close to US$1.8 billion (see the National Nanotechnology Initiative’s Budget Brief 2013: http://www.nano.gov/node/750).
Nagel 2012). Reducing subjectivity to the material world is the very problem that generates conflicts in medical ethics regarding the beginnings and ends of life, for, as many have pointed out, without a material correlate for consciousness the designation for the start or termination of personhood will be based on social consensus rather than on any natural fact of the body and brain (Franklin and Roberts 2006; Lizza 2006; Lock 2002).

Nevertheless, the addition of the information and cognitive sciences to the NBIC convergence is an attempt to address the host of issues around mind, to propose possibilities for the materialization of the whole human person as a self-aware and autonomous being, just like Bina48 in the mock trial. This move toward an informatic paradigm is part of an older trajectory in science and technology (Oyama 2000). At least since Watson and Crick’s description of DNA in the 1950s, biology has been understood and manipulated as code, or information. At the same time, cybernetics had been theorizing action, agency, and even thought as information based (Dupuy 2009). That meant that all goal-oriented, as opposed to random, behavior was conceived as bearing and imparting information, and indeed doing both in a functional exchange was conceived as a way of adapting to the environment. Itself an abstraction, the category ‘information’ nevertheless seemed to provide a way to capture and reproduce ‘life’ as well as cognition, self-awareness, and complexity in a wide range of fields (Shannon and Weaver 1949; also see Wilson [1998] 2016; Helmreich 1998).

The implications of approaching ‘life’ and ‘mind’ as ‘information’ emerge clearly at the starkest point of transformation of life into nonlife and of person into nonperson, namely, death. In cryonics, for example, the main way to conceive of a frozen person as a potential candidate for healthy reanimation in the future has been to imagine that the person’s brain contains them in a retrievable and transferable informatic form. So from its beginnings in cryonics in the 1960s up to the more recent adventures in AI, technoscientific immortality has been a key site where the proponents of projects that seek to materialize the mind have gathered. Ralph Merkle, a participant in many of the sites of speculative technoscience described in this article, came up with what is now known as the information-theoretic definition of death, bringing nano, bio, and info together in theorizing death and immortality, the bottom-line interest of transhumanism and immortalism. Cryonicists in general had been dissatisfied with legal and medical definitions of death, which seemed to be dependent on existing medical capacities and legal conventions rather than on independent criteria that could establish that an apparently dead organism had indeed entered an absolutely irreversible state. Since not only life but also mind were now being described in terms of information, Merkle’s definition of death (1992, 9) ran as follows:
A person is dead according to the information theoretic criterion if their memories, personality, hopes, dreams, etc. have been destroyed in the information theoretic sense. If the structures in the brain that encode memory and personality have been so disrupted that it is no longer possible in principle to recover them, then the person is dead. If they are sufficiently intact that inference of the state of memory and personality are feasible in principle, and therefore restoration to an appropriate functional state is likewise feasible in principle, then the person is not dead. (emphasis in original)

Subscribing to the notion of information-theoretic death entails accepting a new kind of potential personhood, one dependent on future technologies and on a notion of a mind-matter continuum through which an informatic vision of being and thinking can be conceived. This could include uploads, robots, and avatars as persons, since all three theoretically could carry the ‘information’ of human persons. Indeed, note that Merkle’s definition of life and death carries no biological terminology whatsoever, since at nano scales biology is information. Recovery and reanimation, life and death, health and well-being are imagined and proposed as informatic affairs. Thus, at the nano scale, information, not biology, is considered the limit of intervention into human health and continuity.

Even assuming such a proposition is possible in principle, the technical and philosophical difficulties of crossing what may be called the mind-brain barrier are obviously enormous. But that has not stopped researchers from putting together preliminary experiments toward the realization of that crossing. I now provide examples of the experimental propositions for the capture and externalization of mind, then turn to some of the implications, for what may matter more is that in the process new forms of being and new forms of care may be emerging.

One proposal for the mapping and reassembly of the brain came in an article cowritten by Freitas, which suggests sending neuronanobots into the brain in order to register and transmit every charge (energy differential) on the synaptic and neuronal level (Martins, Erlhagen, and Freitas 2012). Atom by atom, the brain would get registered and reassembled by nanodevices. All our memories and whatever else makes up our internal human identities, once stored in cellular structures and distributed along fine synaptic connections, could thus be reactivated on other external, most likely computational, platforms.

Similarly, Kenneth Hayworth, a cryonicist and neuroscientist, connects cryonics and informatic futurism to mainstream neuroscience and medicine. Hayworth started out doing graduate work at the University of Southern California while building a brain-slicing and imaging machine in his garage. He eventually ended up as a scientist at the Howard Hughes Medical Institute in Virginia, where I visited his lab and observed images of his brain slices. Hayworth is a member of the cryonics company Alcor, though he told me he estimates the
chances of cryonics succeeding in reanimating deceased individuals as very low. He said he wants to live for a long time, but he’d rather not be frozen. He would prefer to have his brain preserved through a different process, *plastination*: every protein, every cell, every synapse and neuron ‘fixed’ by a special resin, in what he called the ‘most perfect fossil’. That fossil would then be nanosliced and imaged with an electron microscope. In the process, the biological brain that was Hayworth would be destroyed, but its detailed image would have been mapped and therefore available for reconstruction and reactivation (not just representation) on a computational platform.

Hayworth has also founded the Brain Preservation Foundation (BPF), a nonprofit set up to advance his ideas and anticipate their results. BPF contends that if brains are properly preserved, persons will be ‘available for future reading or revival’. To achieve this, Hayworth is gathering funds to promote scientific research toward brain preservation at the nanometric scale, working to ‘advance public understanding of the self, of our brains as physical, chemical, and biological carriers of our “internal self”’ (BPF, n.d.). BPF views informatic technologies as ‘rapidly-improving carriers and extensions of both our internal and external selves’.

Theories and practices that purport to align an internal self (mind) with external materializations have led to new forms of social life and to new forms of care. For example, cryonicists feel justified in insisting that their cryopreserved members are ‘patients’ with potential for future reanimation and that they be cared for accordingly. Dr. Catherine Sullivan, a trained biologist and director of Suspended Animation (a cryonics company based near Boca Raton, Florida, that carries out research and provides transportation support but does not offer storage and care), told me: ‘I don’t consider them dead. I treat them as a sick, unconscious person, very sick and very unconscious. As a doctor would a patient. They can’t communicate with us, but that doesn’t mean they won’t in the future. Once they’re in storage they’re like coma patients’.

This is why cryonicists have advocated for rights for their suspended patients. Similarly, in its charter BPF argues that some limited rights ‘should be secured for persons in storage’ (in other words, plastinated brains), such as the right to be protected from medical malpractice. Based on a similar logic of externalization, Terasem’s mock trials were part of an attempt to draw up Asilomar-style ethical guidelines for futuristic personhood, including avatars.8 Going beyond the biopolitical, these scenarios for future persons, organized around the promise of convergence, are producing an ‘ethos of suspension’ wherein future forms have not arrived
but present life forms have been stored (digitally, cryogenically) and transformed, and so must be cared for. The figures are suspended at the boundary between the internal and the external (and maybe the eternal), between life and nonlife, subject and object, between what is deserving of care and what might be. Suspension and its ethics of care for future persons also change the biopolitical calculus because not only are intergenerational and ecological futures at stake but also the type of person that is imagined and promised in that future: a kind of person that does not currently appear in medical textbooks.

Rothblatt, who is also signed up with the cryonics outfit Alcor, extends her efforts beyond conferences and into experiments with the externalization of human persons on various platforms. Her example of the externalization of the internal self is an experiment she runs along with Bainbridge, alternately called CyBeRev (for ‘cybernetic beingness revival’) or Lifenaut. CyBeRev/Lifenaut are online portals on which members are asked to create an informatic version of themselves through two different approaches. In the first, member information is based on an extensive psychological-profile form, or ‘personality-capture’ form, designed by Bainbridge. The personality-capture form, with hundreds of questions, is designed to form a comprehensive but updatable picture of a member’s beliefs, values, behaviors, attitudes, and self-image, all those things, in short, that are imagined to make up the interiority of a person. In addition, members are encouraged to upload a range of digital files, from photos and social-media archives to personal writings, from Amazon shopping lists to medical reports. Together, these are called a person’s ‘mindfile’, that is, they are a cumulative database of information reflecting a unique person. All of this is processed by algorithms and AI software (versions of chatbots) that are meant eventually to generate self-aware, autonomous persons stored on a digital platform for subsequent uploading and animation, thereby transferring the human ‘consciousness software into a cellular regenerated or bionanotechnological body by future medicine and technology’ (Rothblatt 2012, 148). In the meantime, the mindfiles are ‘space-cast’, transmitted as digital information into outer space via satellite from Terasem’s beachside headquarters. At the mock trial, Rothblatt explained, ‘So every Terasem joiner or participant who has mindfiles with us has already achieved a certain level of immortality by having aspects of their mindfiles already anywhere from up to five to six light years away from the earth, depending on when they started uploading’.

Bainbridge, who uses an avatar to run review panels for NSF grant proposals on Second Life, also has his own avatar-based immortality project. Called Ancestor Veneration Avatars, the project’s goal is to gather data from the hours and hours of gameplay decisions made by players of online multiplayer games in order to use that information to generate avatars of the players (Bainbridge 2013). The avatar learns from the person who is playing it during the course of the person’s life and will supposedly be able to continue relations with others on the gaming platform after the original person’s death.
Once again, as extraordinary as some of these transhumanist ideas may appear, I want to emphasize their continuity with the many ways that the materialization or externalization of the mind are being carried out as mainstream projects. These range from AI-driven brain-implant technologies that use electric currents to detect and alter moods (Reardon 2017); to whole-brain emulation and the production of neural nets on a computer database whose algorithms (software) could simulate the workings of a specific person’s brain (Sandberg and Bostrom 2008); to detailed brain-mapping projects, some of which have been funded by the NIH through its US$40 million Human Connectome project launched in 2009. Indeed, Hayworth’s own research on brain function and preservation sits right on this very tense but porous border between mainstream and fringe neuroscience. He was developing his research just as the NIH was putting together the Human Connectome project, an initiative modeled on the Human Genome Project, whose goal was to map all the neurons and synapses of a human brain. Even as mainstream scientists distance themselves from Hayworth’s more ‘extreme’ premises, goals, and statements, some have nevertheless supported the project. Hayworth’s board includes prominent mainstream scientists, including Sebastian Seung, who was an important initiator of the Connectome project. There may be a difference between the mainstream project of mapping the brain and the transhumanist project of claiming to externalize minds off their carbon-based substrates and onto the current computational substrate, silicon. But – irrespective of how the boundaries between those differences are maintained – both share similar assumptions about the relationship of brain to mind, namely, that mind is reducible to brain.

In pursuing these goals and imaginaries, however, none of the projects produce proof that the mind can be reduced to the brain, or that algorithms are conscious, or that nanoneurobots can replace our being. Rather, they produce more and more novel objects (algorithmic avatars, social robots, nanobots) and forms of cognition, and these begin to get incorporated into the expanding assemblage of well-being spreading out from the body and biology into other forms and platforms. Since these particular imaginaries also emerge through the capillaries of technocapitalistic power in Silicon Valley, they manifest in very concrete and visible ways, if not always in predicted ones.

Take Roman Mazurenko. He died. But he came back as an app made by his best friend, Eugenia Kuyda, who was running an AI company called Luka in Silicon Valley. Kuyda and Mazurenko had moved there from Russia to launch their own software projects (Newton 2016). When Mazurenko died in a car accident during a trip to Moscow, Kuyda could not abide his absence and asked his friends to submit their chat records so that she could upload them to an AI. She poured more than eight thousand lines of text and hundreds of her own exchanges into a neural net her company had constructed using a recently released open-source Google AI platform, TensorFlow. Roman Bot’s semipublic release disturbed some friends and enchanted others, who wrote to grieve, to share, and to continue to get advice.
from their friend. None of Kyuda’s ideas were entirely new (see Dormehl 2017); they had been circulating among immortalist and transhumanist networks as well as in the AI world of Silicon Valley. Echoing Rothblatt’s and Bainbridge’s ideas, Kuyda has released another app, Replika, that allows you, the user, to be replicated (https://replika.ai/). Kyuda’s app learns from a user’s chats and smartphone interactions and begins to respond more and more like the user, with the same vocabulary, syntax, range of topics, and linguistic idiosyncrasies. It can thus continue interacting with other online persons even if the original human user dies.

The trouble, as Battaglia (1995) noted in the case of cloning, is that the copy is never fully a copy. It is either more or less than the original; it is either not as good or better. Instead of thinking about copy and original, what ought to be said is that Roman Bot is a particular sort of entity because Roman the human was also a particular sort of entity; both are shaped by a technological infrastructure. For Roman Bot to exist as a resurrected avatar, the neural net had to exist; Google’s AI, TensorFlow, had to exist and then be made public; smart phones had to exist; and lots of user behaviors and cognitive abilities had to change in order for that data to be captured, since all the data is based on chat and social-network exchanges. So not only is Roman Bot a particular kind of person, but Roman, too, was a particular kind of human, one who knew how to chat on a number of platforms and externalize his images, thoughts, and memories on devices. In other words, Roman the human was already adapted to and shaped by the demands of AI, by its algorithmic feedback loops and infrastructures. Equally, Kyuda’s Roman Bot is recognizable as Roman only if you had previously chatted with the human Roman on a smart phone. It is this continuity that makes Roman Bot a plausible substitute for Roman. That technological shaping is what allows Roman Bot to be considered a person-like avatar whose existence and well-being is valued by many of Roman’s friends. The same will happen with the many ‘Replikas’ coming out of Kyuda’s and other similar projects (Parker 2014, Dormehl 2017): they will be recognized and cared for as specific types of humans by specific types of humans.

Avatars and robots may be recruited to care for people, for example, as online health coaches or heart monitors (True Image Interactive, n.d.), but people also exchange care with such robots and adjust their behavior and cognitive development accordingly. That is, they come to care about these robots. A number of cases are on record in which humans have become attracted to robots in part because other humans seemed repulsive or unlikable. For example, tech blogs and other press outlets recently reported on the case of a French woman who 3-D printed a robot that she then fell in love with and wants to marry. Lilly has said she is attracted to robots because she prefers their ‘mechanical defects to human flaws’ (Segall, n.d.). ‘Robosexual’ identities are spreading.
Questions of health and care in these cases of loss and love relate not just to keeping humans healthy but also to maintaining healthy human-nonhuman relations. Thus, the rapidly evolving affective connections at the human-machine interface are getting recruited into the world of therapy such that robots – some of which simulate a human hug! (see Cahlan 2018) – are asked to address a range of apparent ‘disorders’ from pedophilia to, ironically, misanthropy and asociality (Waytz, Epley, and Cacioppo 2010).

Conclusion

Convergence is both a concept and project space where certain questions about posthuman health are activated. When nonlife becomes a life form or a lifeway, how do we measure health? What tools, standards, and norms may get generated, proposed, and instantiated? How is a politics of life, a biopolitics, transformed by crossing the carbon barrier?

These may seem like far-off futuristic ideas that should not affect health policies or notions of health. More precisely, one might object that the carbon barrier, to use Goldstein’s term, has not materially been broken. Yet, as I have outlined, influential projects are already developing around these ideas in the United States. Posthuman relations are already the subject of bioethical debates, conferences, and biomedical research. Governments, including some outside the United States, have funded nanotechnology ethics committees, such as the one Goldstein served on, within their nanotechnology initiatives. The immortalist and transhumanist imaginaries beyond the carbon barrier are part of the wider shift toward posthumanist convergence. I am not suggesting that the futures imagined by futurist projects will pan out just as projected, for example, that nanobots will come to exist just as the Foresight Institute envisions them. But the materialization of futures through these experiments in the present create social imaginaries and horizons of expectation that nevertheless matter, as do the objects, machines, and practices themselves.

The intimations and implications of moving beyond the carbon barrier are already with us; many people ‘anticipate’ them, and they are ‘lived and felt’ as inevitabilities of the present (Adams, Murphy, and Clarke 2009). These new objects, new forms of cognition, and new affects are already creating new versions of health and bioethics. Along with this, of course, they produce more forms of vulnerability and ever-moving goalposts for health and well-being. As the philosopher of science Alfred Nordmann (2007, 219) has stated regarding American notions of convergence, ‘the expansion of power and control is accompanied by new dependencies, new kinds of ignorance, new problems even of human or ecological survival’ (also see Coeckelbergh 2011). Nelly Oudshoorn (2016), for example, has worked with people who do not feel fully comfortable with their implanted defibrillators because the device is primed to deliver a shock when it picks up an event that resembles a stroke. That shock may
come in response to a real or nonreal threat, it might be a lifesaver or a false alarm, but either way it both constitutes and denotes the vulnerabilities, human and post-, that patients live with, what Oudshoorn calls the ‘existential uncertainties of cyborgs’.

The questions that arise through the internalization of medical devices point to issues of health beyond the carbon barrier: an insulin device, brain implant, or pacemaker (Clery 2015) may be hacked, with dangerous consequences, but a whole army of nanobots coursing through your immune system and neural connections brings up entirely new questions about what makes you you and what danger may be to you. I bring this up not only to flag the implications for mental health and cognitive surveillance but to open up an analysis of how convergence is recasting what it means to be a thinking, living, and healthy human being. To restate a question I posed at the beginning of this essay: when biological and silicon health are so intertwined, what counts as health, and who is a health practitioner?

Going the other way, the externalization of cognitive abilities and mental states onto nonbiological platforms creates another indicator of health beyond the carbon barrier. The silicon substrate on which digital lives run today already breaks the barrier between the interior and the exterior, such that our loves, memories, feelings, knowledge – that is, some crucial aspects of being in the world – reside outside our bodies and brains, in such a way that they may be activated by others, or in others, at any time, even after death. What relations and practices of care are developing around these externalizations? And conversely, what forms of harm, via extraction, financialization, pollution, and so on, are they producing?

If, as some theorists contend, ‘human’ is part of a continuum that includes the ‘nonhuman’, then we have to face the possibility that these are not anthropomorphic projections based on an illusion that a mechanical object is ‘alive enough’ (Turkle 2011; also see Castaneda and Suchman 2014; Suchman 2007). Rather, it may be that informatic media are producing forms of material-cognitive continuity that are taking the human beyond the carbon barrier. Such exchanges of affect and well-being between subject and object were previously thought about in terms of animism, possession, fetishism, and madness, that is, as irrationalities and pathologies. Transferred to the realm of tech, they are being actively produced as inevitabilities, as ideals and techniques of well-being in various contemporary technoscientific renditions of the posthuman present and future, in realms from law to cognitive psychology to business to the health industry to the sex industry. These subjects, objects, and projects – these new

---

9 Here I set aside the metaphysical question of whether a robot has experiences and simply observe some anthropological transformations around that relationship. Also see Clark (2007) on neuroplasticity, extended mind, and deep embodiment and Vidal (2007) on anthropomorphic acts.
formations – don’t reinforce an essential boundary but rather come to define and shift the boundaries in fundamental ways that make humans and our environments related in new ways. If we humans come to care for algorithmic forms of being, it is in part because, by emulating the algorithmic in daily life, we may have also assumed algorithmic forms of being.

The expansion of the field of health I’ve described here extends beyond humans – not toward the natural but toward the robotic and computational – such that the notion of the environment comes to include cybernetic, algorithmic processes and nonbiological systems such as ‘clouds’ or nanobots rather than plants and planetary life. In this view, it is not only that these nonbiological powers can be activated in the service of human health but that they are defining new forms of health, new interdependencies of health among entities. Humans are not just one part of a large biosymbiotic planetary environment; they are also part of a computational universe of growing complexity. To transhumanists and many Silicon Valley entrepreneurs, that universe appears to be evolving beyond and out of biological containment into a posthuman future of engineered material where mind-clouds converge with nano-neuro-bio-info interfaces. To invoke the language of the Foresight Institute, in the world of *Homo sapiens siliconis*, what kind of a medicine man will a hacker be?

**Acknowledgements**

For this paper, as for the whole issue, I owe so much to my colleague and coeditor Richard Rottenburg, who got this project off the ground in Germany and provided guidance, humor, and support throughout. The feedback of seminar participants was crucial in shaping the ideas: I thank Julie Livingston, Susan Erikson, Miriam Ticktin, Sandra Calkins, Vincent Duclos, Song-Joon Park, and David Kananizadeh. Sarah Freeman’s help in putting together the seminars at the New School and formatting the papers was invaluable. Without the detailed and incisive comments of Martha Lincoln, this would have been an infinitely poorer think piece. I also thank the editors of *MAT* and the anonymous reviewers for their generosity and honesty.

**About the author**

An anthropologist, writer, and artist, Abou Farman is author of the book *Clerks of the Passage* (Linda Leith Press, 2012) and *On Not Dying: Secular Immortality in the Age of Technoscience* (University of Minnesota Press, 2020). He is Assistant Professor of Anthropology at The New School for Social Research and founder of Art Space Sanctuary as well as the Shipibo Conibo Center of NY. As part of the artist duo caraballo-farman, he has exhibited internationally, including at the Tate Modern, UK, and PS1/MOMA, NY, and received several grants and
awards, including NYFA and Guggenheim Fellowships. He is producer and co-writer on several feature films, most recently Icarus: A Vision.

References


